Bifurcation Analysis of Large Networks of Neurons

by

Wilten Nicola

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Applied Mathematics

Waterloo, Ontario, Canada, 2015

Author's Declaration

I hereby declare that I am the sole author of this thesis.

I understand that my thesis may be made electronically available to the public.

Abstract

The human brain contains on the order of a hundred billion neurons, each with several thousand synaptic connections. Computational neuroscience has successfully modeled both the individual neurons as various types of oscillators, in addition to the synaptic coupling between the neurons. However, employing the individual neuronal models as a large coupled network on the scale of the human brain would require massive computational and financial resources, and yet is the current undertaking of several research groups. Even if one were to successfully model such a complicated system of coupled differential equations, aside from brute force numerical simulations, little insight may be gained into how the human brain solves problems or performs tasks.

Here, we introduce a tool that reduces large networks of coupled neurons to a much smaller set of differential equations that governs key statistics for the network as a whole, as opposed to tracking the individual dynamics of neurons and their connections. This approach is typically referred to as a mean-field system. As the mean-field system is derived from the original network of neurons, it is predictive for the behavior of the network as a whole and the parameters or distributions of parameters that appear in the mean-field system are identical to those of the original network. As such, bifurcation analysis is predictive for the behavior of the original network and predicts where in the parameter space the network transitions from one behavior to another.

Additionally, here we show how networks of neurons can be constructed with a mean-field or macroscopic behavior that is prescribed. This occurs through an analytic extension of the Neural Engineering Framework (NEF). This can be thought of as an inverse mean-field approach, where the networks are constructed to obey prescribed dynamics as opposed to deriving the macroscopic dynamics from an underlying network. Thus, the work done here analyzes neuronal networks through both top-down and bottom-up approaches.

Acknowledgements

I would like to thank first and foremost my supervisor, Prof. Sue Ann Campbell. Without her guidance, wisdom, and patience, I wouldn't have made it this far and I am eternally grateful.

I would also like to thank my committee members, Prof. Matthew Scott, Prof. Francis Poulin, Prof. Chris Eliasmith and Prof. Bard Ermentrout in addition to my collaborators Prof. Frances Skinner, Prof. Cheng Ly, Prof. Bryan Tripp, Dr. Felix Njap, Dr. Katie Ferguson, Dr. Muhammad Dur-e-Ahmad for all their helpful suggestions and the many meetings and conversations we've had over the years

I would also like to thank my friends and co-workers that I've met over the last 5 years in Waterloo: Darryl Jones, Wendy Guan, Alex Shum, Boglarka Soos, Jared Penney, Killian Miller, and Clinton Loo.

Dedication

This is dedicated to my parents, Evelyn and Beuoder, my brother Martin, and my grandfather, Ekshiresh Gorgis.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

The human brain contains on the order of one hundred billion neurons [83], the primary computational unit of animal nervous systems. These cells come together to perform specific functions in large networks, often containing thousands of individual neurons each with thousands of connections [160]. For example, in a region of the brain referred to as the hippocampus, the neuronal networks are functional units that are involved in the encoding and recollection of memories, in addition to spatial localization and navigation [6].

While the function of various regions of the brain and their associated networks has been known since the time of Wilder Penfield [144], very little is known about how collections of coupled neurons can perform these functions. Fortunately however, the individual neurons, and their electro-chemical coupling to each other at connections referred to as synapses are well understood from both a biological and a mathematical perspective for decades now [87]. Given our mathematical insights into the functioning of neurons and their connections to one another, one can approach the problem of understanding how networks function from either a bottom-up or a top-down perspective.

For example, one can model an entire neuron in exquisite detail down to the molecular level and with modern computing resources, attempt to further model an entire network of these neurons. This is the approach taken by the Blue Brain Project [114], a European collaboration which seeks to simulate human scale neuronal networks down to the molecular level by 2017 using a Blue Gene supercomputer. Using this approach, one can attain a great deal of knowledge using these very large scale realistic simulations. For example, one can run in-silico experiments that are either not feasible or not ethical with actual neuronal networks. However, with realistic simulations come incredibly large parameter spaces and intractable analyses. Unfortunately this approach is entirely confined to simulations and very little can be said about how networks function (and even less can be said without a Blue Gene super computer). This limits what one can infer about the macroscopic functioning of a network. This is one of only a few large network approaches [5, 50]

For a more top-down perspective, one can also approach the understanding of neuronal networks if one makes very large sacrifices in realism, by using very simple neuron models and connections. With simpler computational units, one can easily train the network to fix the connection weights between neurons to perform any arbitrary function. In essence, the minute details of a relatively simple network are fixed so that a macroscopic behavior can emerge.

This is the approach taken in the field of artificial neural networks (ANNs), where the "neurons" are sigmoid functions or other simple threshold functions representing the information processing that a neuron performs on its inputs, and the networks are compositions of linear combinations of these neurons [82]. This field started with the seminal paper by Warren Mcculloch and Walter Pitts [118] and quickly grew thereafter particularly in the direction of how to train an ANN to approximate any function or task [155] and the construction of more specialized kinds of networks [88]. One can use ANNs to perform a great deal of tasks such as function estimation, object recognition, motor control, to name a few [82]. However, given the abstractness of the computational elements, one can not immediately conclude anything about biological neuronal networks. While the parameter space is significantly smaller then efforts like the Blue Brain Project, the parameters are abstractions that are not easily comparable to measurable parameters of actual neurons. Furthermore, the sigmoids themselves should actually be thought of as the population response for a network of neurons, as opposed to an individual neuron, [190].

One method of applying the theory behind artificial neuronal networks with spiking neurons is to replace the sigmoids used in ANN theory with the firing rates of spiking neuron models. This is the core idea behind the Neural Engineering Framework (NEF) [51, 52] and results in spiking neural networks that can generate spiking models of path integration [37], working memory [167], visual attention [19], motor control [46], various cognitive functions [15, 148], and many others. This has culminated in the most sophisticated spiking neuron model of the human brain to date [53]. While the synaptic weights coupling these spiking neurons in the NEF approach are immediately determined through a quadratic optimization problem, little can be said analytically about the resulting networks as the optimization still introduces a numerical step that must be solved.

While both these approaches have their purposes and applications, it is clear that due to the tradeoffs, a different perspective on biological neuronal networks is needed. In particular, starting with realistic models of neurons and their synaptic connections, what is needed is an equation that can predict the function and behavior of the resulting network at a macroscopic level, while being much simpler to analyze and understand. This is typically referred to as a mean-field approach. In the context of neural systems, the original idea of a mean-field approach originates with Jack Cowan and Hugh Wilson [190]. In the broader context, mean-field approaches appear in many different fields with many different names. In the context of computational neuroscience, there are many differing applications, definitions, and formulations of what a mean-field system should be which will be covered in greater detail in the relevant chapters.

In mean-field theory, one attempts to reduce the behavior of large numbers of interacting units to a small set of representative equations that govern the macroscopic behavior of the entire system. The system we will consider is a spiking neuronal network. We will analytically derive a small system of mean-field equations for the statistical properties that describe the networks behavior. The neuron models we will use are substantially more realistic than the sigmoid functions used in artificial neural networks and can be fit to actual neurons. While less sophisticated than the models used by other approaches, these models we will use are substantially more tractable analytically, and thus represent an optimal trade off between realism and tractability. As the mean-field system is analytically derived from the original network, any parameters or distributions of parameters that are present from the mean-field system are identical to those in the original network. Thus, bifurcation analysis of the mean-field system is predictive of the behavior in the original neuronal network.

Additionally, in the final chapter we will demonstrate the relationship between mean-field theory and NEF generated networks, in addition to determining analytical solutions to the synaptic weights for NEF networks in the limit that these networks are large. For large networks, one can show that the NEF approach is effectively a mean-field approach with heterogeneous coupled neurons. By using the mean-field perspective, one can now compute various important statistical quantities analytically with a NEF synthesized network.

This thesis is organized into 8 chapters, which are listed and summarized below. Much of the research presented in this thesis is original, and was conducted with researchers in the examining committee and outside of it.

Chapter 1. Introduction

A general introduction for this thesis.

Chapter 2. Mathematical Prerequisites

In Chapter 2 we will introduce the mathematical prerequisites required to read this thesis in addition to introducing dynamical systems theory, topological equivalence and bifurcation theory, in addition to some important results in the asymptotic expansions of integrals that will be required at various points in this thesis

Chapter 3. Biological Prerequisites

In Chapter 3, we will introduce the various ways one can model neurons, synapses, and large networks. We will introduce the primary networks we consider for mean-field derivations in this chapter, in addition to briefly introducing other types of related neuronal modeling approaches. We will also introduce population density equations in this chapter. These are a necessary step in deriving a mean-field system.

Chapter 4. Derivation of Mean-Field Systems for a Homogeneous Network of All-to-All Coupled Spiking Neurons

In Chapter 4, we will derive the mean-field system for a network of two-dimensional integrate-and-fire neurons with all-to-all coupling through a series of analytical reductions

and approximations. We will additionally discuss when the mean-field system is an appropriate descriptor of the behavior of the original network of neurons through spectral analysis of the population density equations. This research was performed with Prof. Sue Ann Campbell and was published in [133, 134]. My contribution to [134] was derivation of the mean-field system in addition to numerical bifurcation analysis. The spectral analysis that appears in [133] while unique to the particular system in that paper, is similar to other analyses in the literature.

Chapter 5. Mean-Field Systems for More Realistic Networks

Chapter 5 extends the derivation and spectral analysis conducted in Chapter 4 to networks of neurons with heterogeneity, noise, or sparse coupling. Additionally, we look at a specific application of mean-field systems in predicting where in the parameter space networks of fitted hippocampal CA1 pyramidal neurons display prominent theta oscillations. The work on mean-field systems and heterogeneous networks was conducted with Prof. Sue Ann Campbell and was published in [135]. The work on mean-field systems and noisy networks was conducted with Prof. Sue Ann Campbell and Prof. Cheng Ly (Virginia Commonwealth University) and will appear in [136]. The work on mean-field systems with sparsity, and the case study was conducted with Prof. Sue Ann Campbell, Prof. Frances Skinner (Toronto Western Research Hospital), Dr. Felix Njap, and Dr. Katie Ferguson and will appear in [63]. My contribution in all cases was derivation of the mean-field system for the underlying network, and numerical applications of the mean-field system.

Chapter 6. Bifurcation Analysis of the Mean-Field System for a Homogeneous Network of Neurons

In Chapter 6, we analyze the bifurcation structure of the mean-field system derived in Chapter 4. This work was performed with Prof. Sue Ann Campbell and will appear in [133]. Much of the work on local bifurcations of equilibria in this chapter is analytical, while the work on global bifurcations is numerical. Additionally, we present some numerical work on the mean-field system for a network of neurons with noise, which was derived in Chapter 5.

Chapter 7. The Neural Engineering Framework, Inverse Mean-Field Theory, and Analytical Solutions to Weight Matrices for Networks of Neurons with Prescribed Dynamics

Chapter 7 Introduces the Neural Engineering Framework and recasts the optimization problem as an inversion problem for a linear operator in the limit of large networks. This leads to analytical solutions for the synaptic weights that couple neurons with specific types of firing rates. Various results about the convergence and scaling of the analytical solution are presented, in addition to its non-uniqueness. The relationship between mean-field theory and the NEF approach is also elucidated. This work has been submitted for publication. The research was conducted with Prof. Matthew Scott and Prof. Bryan Tripp (University of Waterloo). The analysis present here was conducted by myself in conjunction with Prof. Scott while much of the numerics were performed by myself with the aid of Prof. Bryan Tripp.

Chapter 8. Conclusions and Future Directions

Chapter 8 summarizes the results present in this thesis and discusses related mean-field approaches and their comparison to our approach in addition to comparing the NEF approach with another constructive network approach that has recently been published. Furthermore, we outline some extensions for both mean-field theory and the Neural Engineering Framework.

Chapter 2

Mathematical Prerequisties

In this chapter we will introduce some important mathematical prequisites needed to understand the thesis in 2.1 before moving on to introduce dynamical systems theory 2.2. We will introduce the idea of topological equivalence in section 2.3. We will also introduce the closely related idea of a bifurcation, which can be thought of as a qualitative change in a dynamical system, in section 2.3. Finally, we will introduce stable, unstable, and center-manifolds of the equilibria of a dynamical system in section 2.4. The center-manifold theorem is a particularly important topic in the context of bifurcation analysis. Finally, important tools used throughout the rest of this thesis are asymptotic expansions of specific classes of integrals. These will be covered in section 2.5.

The mathematical introduction in this chapter will be necessarily abrupt. One can find more detail in the various textbooks on dynamical systems and bifurcation theory that aided in the preparation of this chapter [17, 99, 100, 145, 189]. The general introduction to dynamical systems theory and bifurcation analysis (sections 2.2,2.3) follows closely from [100], with some of the newer non-smooth bifurcation analyses from [48, 164, 166]. The section on asymptotic expansions of integrals follows closely from [17].

2.1 Preliminary Mathematical Definitions

Here, we will quickly define much of the preliminary mathematical infrastructure required for the rest of the thesis and the introduction to dynamical systems theory. We will assume that the reader has a basic understanding of linear algebra and sequences and series.

2.1.1 Metric Spaces, Banach Spaces, and Contraction Mappings

Definition 2.1.1. Metric Space A metric space is a set X and a distance function $\rho(x, y) : X \times X \to \mathbb{R}^+$ defined for all $x, y \in X$ such that the following hold

1. $\rho(x, y) = 0$ if and only if $x = y$
2. $\rho(x, y) = \rho(y, x)$ (symmetry)
3. $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ (triangle inequality)

The function $\rho(x, y)$ is referred to as a metric of the space X if it satisfies these three conditions. While very basic in their definition, these spaces still have a very rich theory.

For example, the set $\mathbb{R}^n$ with

$$\rho(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

is a metric space, as is the set of continuous functions on the interval $[a, b]$, denoted by $C_{[a,b]}$, with metric

$$\rho(f, g) = \max_{a \leq t \leq b} |f(t) - g(t)|.$$

One of the most important results in the theory of metric spaces is the contraction mapping theorem (also known as the Banach fixed point theorem [13]). Consider equations of the form

$$Ax = x \tag{2.1}$$

where $A : X \to X$, and a solution to this system, x^* is referred to as a fixed point. The contraction mapping theorem establishes the existence and uniqueness of a fixed point of (2.1).

Theorem 2.1.1. ***Contraction Mapping Principle*** *Suppose there exists a number* $\alpha < 1$ *such that*

$$\rho(Ax, Ay) \leq \alpha\rho(x, y), \tag{2.2}$$

then A *is a contraction map and the equation* (2.1) *has a unique fixed point.*

The proof of this statement can be found in any standard functional analysis textbook (see [99]). This is one of the most useful theorems in all of functional analysis and can be used to prove both the inverse, and implicit function theorems in basic analysis in addition to proving the classical existence and uniqueness theorem for ordinary differential equations [99].

A more restrictive type of space is a Banach space

Definition 2.1.2. Banach Space A Banach space is a complete vector space X with a function $\|x\| : X \to \mathbb{R}^+$ that satisfies the following conditions:

1. $\|x\| \geq 0$

2. $\|\alpha x\| = |\alpha|\|x\|$

3. $\|x + y\| \leq \|x\| + \|y\|$.

The function $\|x\|$ is referred to as the norm of the Banach space. All Banach spaces are metric spaces with the metric

$$\rho(x, y) = \|x - y\|$$

however not all metric spaces are Banach spaces.

Definition 2.1.3. Linear Operator A linear operator is a map $L : X \to Y$ from the Banach space X to the Banach space Y that satisfies the following conditions:

1. $L(x_1 + x_2) = L(x_1) + L(x_2), \quad \forall x_1, x_2 \in X$

2. $L(\alpha x) = \alpha L(x), \quad \forall x \in X, \alpha \in \mathbb{R}$.

The linear operator is referred to as bounded if it additionally satisfies the following condition

3. $\|L(x)\|_Y \leq M\|x\|_X$, for all $x \in X$, for some $M \in \mathbb{R}^+$.

Linear operators allow us to extend the concept of differentiability to more general spaces. This will be useful later on when we need to linearize abstract operators.

Definition 2.1.4. Fréchet Derivative Let W and V be two Banach Spaces, and $U \subset W$. A function $f : U \to V$ is called Fréchet differentiable at $x \in U$ if there exists a bounded linear operator $Df : V \to W$ such that

$$\lim_{h \to 0} \frac{\|f(x + h) - f(x) - Dfh\|_W}{\|h\|_V} = 0.$$

The linear operator Df is the Fréchet derivative (or derivative) of $f(x)$

2.2 Introduction to Dynamical Systems

In order to define a dynamical system, we need to first define some preliminary mathematical objects.

Definition 2.2.1. State Space A state space is the set, X, of all possible states a system can display

The term state space is often used interchangably with the term phase space. For example, consider a system that is initially measured to be at some initial state, x_0 when $t = 0$, where $x_0 \in \mathbb{R}^n$. After some time $t \in \mathbb{R}$, the system is measured again and the resulting state is x_t. Then $x_0, x_t \in X$ and the process that transformed the initial state x_0 to x_t is referred to as the evolution operator.

Definition 2.2.2. Evolution Operator: The map $\phi^t : X \to X$, defined by

$$\phi^t x_0 = x^t$$

is known as the evolution operator or flow of the dynamical system.

Not any map is admissable as an evolution operator of a dynamical system. In particular, we will only consider the maps that satisfy the following condition as admissable evolution operators:

- Continuity: $\phi^0 = \boldsymbol{e}$
- Autonomy: $\phi^{t+s} = \phi^t(\phi^s)$

where e is the identity operator in X, such that $ex = xe = x$. The autonomy condition is also more frequently referred to as the semi-group property. We will refer to evolution operators that satisfy the continuity and autonomy conditions as admissable. For some dynamical systems, the current state x_0 uniquely determines previous states, for $t < 0$. For these systems, there is an additional property which is sometimes imposed:

- Invertibility: $\phi^{-t}\phi^t = \phi^0 = e$

However, we will follow [100] and not specially demand invertibility of our dynamical system. The state space and flow operator uniquely define a dynamical system

Definition 2.2.3. Dynamical System: A dynamical system is a pair (X, ϕ^t) where X is a state space and ϕ^t is an admissable evolution operator

In the context of this thesis, we will primarily consider the case where $t \in \mathbb{R}$ and $X \subset \mathbb{R}^n$. While infinite-dimensional dynamical systems do arise in the derivation process of a mean-field system, we primarily use these as an intermediate to arrive at the finite-dimensional mean-field system. Other types of dynamical systems can occur. For example when the flow operator ϕ^t operates for discrete values of t, such as $t \in \mathbb{Z}$.

Starting from an initial point, x_0, one can define an ordered set generated by the evolution operator, ϕ^t

Definition 2.2.4. Trajectory: A trajectory is the ordered set defined by

$$T(x_0) = \{x \in X : x = \phi^t x_0, t \in T\} \subset X$$

The continuity condition on the evolution operator forces the trajectories to be continuous, while the autonomy condition allows us to relate a continuous time dynamical system to a system of differential equations, as we will see in theorem 2.2.1. There are two types of trajectories that are of particular importance.

Definition 2.2.5. Equilibria An equilibrium is a point x^* such that $\phi^t x^* = x^*$ for all $t > 0$

Definition 2.2.6. Cycle A cycle is a trajectory $T(x_0)$ that satisfies $\phi^{t+\bar{T}}x = \phi^t x$ for all $x \in T(x_0)$, and some $\bar{T} > 0$. The quantity $\bar{T}$ is referred to as the period of the limit cycle.

Equilibria and cycles are said to be isolated if there are no other equilibria or limit cycles in a small neighbourhood around them. Isolated cycles are called limit cycles. One can use trajectories to partition the entire state space, X. The resulting quantity is called a phase portrait.

Definition 2.2.7. Phase Portrait The phase portrait of a dynamical system is a partitioning of the state space into orbits.

We will primarily consider dynamical systems defined by autonomous differential equations:

$$\dot{x} = f(x, \alpha), \quad x \in \mathbb{R}^n, \alpha \in \mathbb{R}^p \tag{2.3}$$

The quantity α is a parameter of the dynamical system, and is static in time. In general, we do not have an explicit solution to (2.3), and thus one typically does not have an evolution operator on hand to analyze. However, one can easily show that continuous time dynamical systems that satisfy the semigroup property lead to a system of differential equations.

Theorem 2.2.1. ***Dynamical Systems on*** $\mathbb{R}^n$ ***I*** *Let X be an open subset of $\mathbb{R}^n$, then every dynamical system on X gives rise to a system of ordinary differential equations.*

Proof. Define $\phi^t x = \Phi(t, x)$ and

$$f(x) = \frac{d}{dt}\phi_t(x)\Big|_{t=0} \tag{2.4}$$

and set $x(t) = \phi^t(x)$. Then we have

$$\begin{aligned} x'(t) &= \lim_{s\to 0} \frac{\phi^{t+s}(x) - \phi^t(x)}{s} && (2.5)\\ &= \lim_{s\to 0} \frac{\phi^s(\phi^t x) - \phi^0(\phi^t x)}{s} && (2.6)\\ &= \frac{d}{ds}\left(\phi_s(\phi^t x)\right)\Big|_{s=0} && (2.7)\\ &= f(\phi^t x) = f(x(t)) && (2.8) \end{aligned}$$

with a similar derivation for parameter dependent systems. □

The proof of this theorem follows closely from [31]. Moreover, every system of differential equations defines a dynamical system, as we shall see in section 2.3.

Given the correspondence between dynamical systems and autonomous ordinary differential equations, we can define the flow as

$$\phi^t(x_0, \alpha) = x(t, x_0, \alpha)$$

where $x(t, x_0, \alpha)$ is the solution to the initial value problem

$$x' = f(x, \alpha), \quad x(0) = x_0 \tag{2.9}$$

Thus, if we can determine when solutions to 2.9 exist, we can establish the existence of the evolution operator for a dynamical system. This is taken care of by the Picard Existence Theorem.

Theorem 2.2.2. ***Picard Existence Theorem*** *Consider $f(x, \alpha)$, which is defined and continuous as a function of x on an n dimensional domain containing the point x_0 at $t = t_0$. Suppose that each component of f, $f_i(x, \alpha)$ satisfies a Lipschitz condition of the form*

$$|f_i(x, \alpha) - f_i(y, \alpha)| \leq M \max_{1 \leq i \leq n} |x_i - y_i| \tag{2.10}$$

on the domain. Then there is an interval $|t - t_0| \leq \delta$ in which the system of differential equations (2.3) has a unique solution satisfying the initial value problem $x(t_0) = x_0$.

The proof of this theorem can be found in any standard textbook on dynamical systems theory, or in textbooks on functional analysis [99]. Unfortunately, the mean-field systems we see in this thesis do not necessarily satisfy a Lipschitz property globally in the phase space, as they have unbounded derivatives in specific regions. For this reason, while we cannot guarantee the uniqueness of solutions, we can guarantee their existence with the following theorem

Theorem 2.2.3. *Suppose that $f(x, \alpha)$ is a continuous function of x on an n-dimensional domain containing the point x_0 at $t = t_0$. Then there is an interval $|t - t_0| \leq \delta$ in which the system of differential equations has a solution satisfying the initial value problem $x(t_0) = x_0$*

Thus any system that is continuous (but not necessarily differentiable) is guaranteed to have a solution (that is not necessarily unique). While the Picard Existence Theorem guarantees the existence of unique solutions over some interval $|t - t_0| \leq \delta$, the interval of existence for solutions can be extended

Theorem 2.2.4. ***Extended Existence to Picard Existence Theorem*** *Suppose that f is continuous and Lipschitz on the domain D. If $x(t)$ is a solution to the initial value problem (2.9) on some interval, then it can be extended to a maximum interval of existence, (α, β). Moreover, if (α, β) is a maximal interval of existence, then $(t, x(t))$ tends to the boundary of D as $t \to \beta^-$ and $t \to \alpha^+$.*

2.2.1 Non-Smooth Dynamical Systems

The class of systems we will consider in this thesis are primarily non-smooth, yet still continuous. These systems are referred to as piecewise-smooth continuous in the literature (PWSC). While we will exclusively see PWSC or smooth mean-field systems, we will introduce some of the general theory of non-smooth systems as a point of reference for the PWSC system. Following [48], we will define the degree of smoothness of a non-smooth system. First, we will rewrite (2.3) as

$$\dot{x} = \begin{cases} F_1(x,\alpha) & H(x,\alpha) > 0 \\ F_2(x,\alpha) & H(x,\alpha) < 0 \end{cases}, \quad x \in \mathbb{R}^n, \alpha \in \mathbb{R}^m \tag{2.11}$$

where the equation $H(x,\alpha) = 0$ defines the switching manifold for the system.

Definition 2.2.8. Degree of Smoothness Consider the difference of derivatives

$$\Delta_{ik}(x) = \frac{\partial^k F_1}{\partial x_i^k} - \frac{\partial^k F_2}{\partial x_i^k}$$

let $d \geq 0$ be the smallest integer such that $\Delta_{ik}(x)$ is continuous for all $i = 1, 2, \ldots n$, and $0 \leq k < d$. Then d is the degree of smoothness of (2.11)

For example, the system

$$\dot{x} = \begin{cases} x & x < 0 \\ -5x & x > 0 \end{cases}$$

has a degree of smoothness of 1 while the system

$$\dot{x} = \begin{cases} x+1 & x < 0 \\ -5x & x > 0 \end{cases}$$

has a degree of smoothness of 0. Systems with a degree of smoothness of 0 are referred to as Filippov systems, due to the pioneering work of Aleksei Fedorovich Filippov [64]. They display phenomena that are not possible with systems of higher degrees of smoothness, such as sliding flow, where trajectories are forced onto the switching manifold, [48].

2.2.2 Stability Theory

Stability is a property of sets contained in the state space of a dynamical system. The simplest set to consider is an equilibrium point however stability can be defined for limit-cycles (or any sets for that matter). We will focus here on equilibria as the majority of the stability analysis in this thesis is the local analysis of various kinds of equilibria of mean-field systems.

Definition 2.2.9. Stability An equilibrium point x^* is said to be stable if for any $\epsilon > 0$, and any $t_0 \in \mathbb{R}^+$, there exists a $\delta(t_0, \epsilon)$ such that

$$\|x_0 - x^*\| \leq \delta \rightarrow \|x(t, x_0, \alpha) - x^*\| < \epsilon, \quad t \geq t_0.$$

Definition 2.2.10. Asymptotically Stable An equilibrium is said to be asymptotically stable if it is stable, and $\forall t_0 \in \mathbb{R}^+$, there exists a $\eta(t_0) > 0$ such that if $\|x_0 - x^*\| \leq \eta$, then

$$\lim_{t \to \infty} \|x(t, x_0, \alpha) - x^*\| = 0.$$

The equilibrium is globally asymptotically stable if η is arbitrary.

Definition 2.2.11. Unstable An equilibrium is said to be unstable if it is not stable.

Note that there are stronger definitions that are also possible, such as uniform stability, uniform asymptotic stability, or exponential asymptotic stability. Furthermore, it is possible for equilibria to satisfy

$$\forall t_0 \in \mathbb{R}^+, \quad \exists \eta(t_0) > 0, \quad \text{s.t.} \quad \|x(t, x_0, \alpha) - x^*\| \leq \eta, \quad \Rightarrow \lim_{t \to \infty} \|x(t, x_0, \alpha) - x^*\| = 0$$

yet not be stable. These are referred to as attractors. See figure 2.1 for an example of various kinds of equilibria and how their stability properties are related to the phase portrait near the equilibrium. Figure 2.1(a) contains what is called a stable node, which occurs when the nearby trajectories converge to the equilibrium in a direct path, while figure 2.1(b) contains a stable focus, where the trajectories spiral inwards towards the equilibrium.

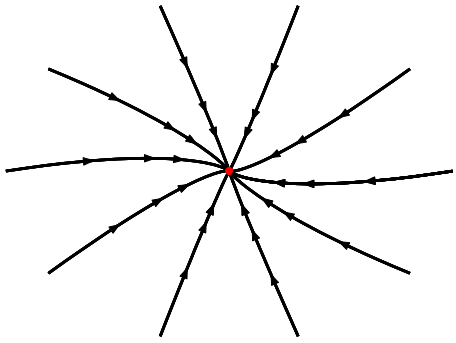
The local stability of an equilibrium can be ascertained in certain cases quite easily. Consider (2.16), then by definition 2.2.5, an equilibria, x^*, of (2.16) satisfies the condition

$$\dot{x}\Big|_{x=x^*} = f(x^*(\alpha), \alpha) = 0 \tag{2.12}$$

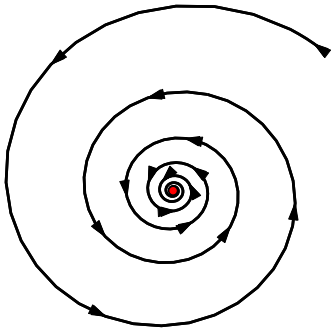
where we have made the parameter dependence of x^* explicit. Suppose that we perturb the system from its equilibrium, $x^*(\alpha)$, with the perturbation $\Delta_x = x - x^*(\alpha)$, In which case we can write

$$\begin{aligned} \dot{\Delta}_x &= f(x, \alpha) = f(\Delta_x + x^*(\alpha), \alpha) && (2.13) \\ &= Df(x^*(\alpha), \alpha)\Delta_x + O(\Delta_x^2) && (2.14) \end{aligned}$$

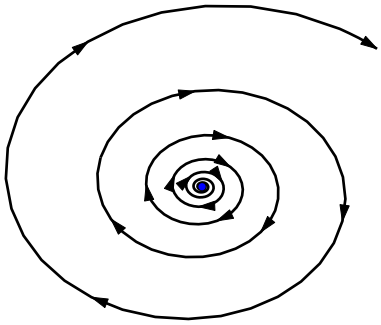
where we have used Taylor's theorem in equation (2.14) after noting that $f(x^*(\alpha), \alpha) = 0$. It should be clear that for small perturbations Δ_x, the matrix $Df(x^*(\alpha), \alpha)$ determines how the perturbations Δ_x decay.



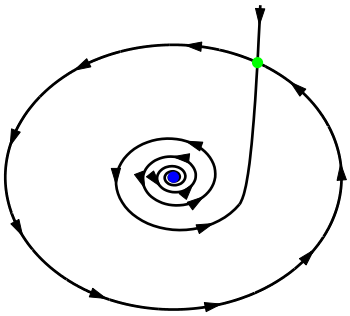
(a) Stable Node



(b) Stable Focus



(c) Unstable Focus



(d) Attracting, but not Stable

Figure 2.1: Various kinds of equilibria and their stability properties in $\mathbb{R}^2$. In figure 2.1(a) is a stable node, which occurs when the eigenvalues of the Jacobian are real and negative. Figure 2.1(b) shows a stable focus, which occurs when the eigenvalues have negative real parts and non-zero imaginary parts. Figure 2.1(c) is an unstable focus which occurs when the eigenvalues have positive real parts and non-zero imaginary parts. Figure 2.1(d) is an unstable equilibrium (green), but still attracting. The equilibrium has what is referred to as a homoclinic orbit.

Definition 2.2.12. Linearization The matrix $Df(x^*(\alpha), \alpha)$ is called the Jacobian matrix of (2.16). The system

$$\dot{\Delta}_x = Df(x^*(\alpha), \alpha)\Delta_x \tag{2.15}$$

is called the linearization of (2.16).

The linearization is a crucial tool in determining the stability of equilibria. Additionally, it can also determine how the trajectories behave locally independent of their stability. For example, nodes have Jacobians with real eigenvalues while focuses have Jacobians with imaginary eigenvalues.

Theorem 2.2.5. ***Local Asymptotic Stability*** *Suppose that all the eigenvalues of $Df(x^*)$ have a negative real part. Then the equilibrium x^* is locally asymptotically stable.*

Theorem 2.2.6. ***Unstable Equilibrium*** *Suppose that an eigenvalue of $Df(x^*)$ has a positive real part. Then the equilibrium x^* is unstable*

The proofs of both theorems can be found in any standard dynamical systems textbook (see for example [145]). Unfortunately, one cannot determine the stability properties of an equilibrium with eigenvalues that have zero real parts. However, the center manifold theorem allows us to resolve even these cases with a great deal of work.

Definition 2.2.13. Hyperbolic Equilibria An equilibrium is called hyperbolic if none of the eigevalues of $Df(x^*)$ have zero real parts

As we shall see in the next section, hyperbolic equilibria of dynamical systems are easy to understand using the linearization and topological equivalence.

2.3 Topological Equivalence and Bifurcation Theory

2.3.1 Topological Equivalence

The topological equivalence of two dynamical systems implies that the systems are qualitatively identical. To formalize this notion, we need some initial mathematical machinery.

Definition 2.3.1. Homeomorphism A homeomorphism $h : X \to Y$ is an invertible map such that both h and $h^{-1} : Y \to X$ are continuous.

With the concept of a homeomorphism in hand, we can define topological equivalence

Definition 2.3.2. Topological Equivalence A dynamical system $\{\mathbb{R}^n, \phi^t\}$ is called topologically equivalent in a region $U \in \mathbb{R}^n$ to a dynamical system $\{\mathbb{R}^n, \psi^t\}$ in a region $V \in \mathbb{R}^n$ if there is a homeomorphism $h : \mathbb{R}^n \to \mathbb{R}^n$ mapping orbits of the first system in U to orbits of the second system in V, preserving the direction of time

Note that topologically equivalence is a mapping of sets in U to sets in V, two subsets of the state spaces of the dynamical systems on $\mathbb{R}^n$ with evolution operators ϕ^t and ψ^t respectively. While there are much stronger conditions one can use that force a closer relationship between ϕ^t and ψ^t (such as topological conjugacy or orbital equivalence), topological equivalence is a weak enough condition to yield applicable theory. A more useful form of the definition is the following:

Definition 2.3.3. Topological Equivalence II A dynamical system $\{\mathbb{R}^n, \phi^t\}$ is called topologically equivalent in a region $U \in \mathbb{R}^n$ to a dynamical system $\{\mathbb{R}^n, \psi^t\}$ in a region $V \in \mathbb{R}^n$ if and only if there is a homeomorphism $h : \mathbb{R}^n \to \mathbb{R}^n$ and for each $x \in U$ there is a differentiable function $t(x, \tau)$ defined for all $\tau \in \mathbb{R}$ such that $\frac{\partial t}{\partial \tau} > 0$ and

$$h(\phi^{t(x,\tau)}(x)) = \psi^{\tau}(h(x))$$

One can also define topological equivalence for parameter dependent systems as done in [100]

Definition 2.3.4. Topological Equivalence for Parameter Dependent Systems: Consider the two dynamical systems defined by

$$\dot{x} = f(x, \alpha), \quad x \in \mathbb{R}^n, \alpha \in \mathbb{R}^p \tag{2.16}$$
$$\dot{y} = g(y, \beta), \quad y \in \mathbb{R}^n, \beta \in \mathbb{R}^p \tag{2.17}$$

Let U_α and V_β be two parameter dependent regions in the phase space of the dynamical system defined by (2.16) and (2.17), respectively. The system (2.16) is called topologically equivalent to the system (2.17) if there is

i. a homeomorphism of the parameter space $p : \mathbb{R}^m \to \mathbb{R}^m$, $p(0) = 0$

ii. a parameter dependent homeomorphism of the state space $h_\alpha : \mathbb{R}^n \to \mathbb{R}^n$, $h_\alpha(U_\alpha) = V_{p(\alpha)}$, $h_0(0) = 0$

such that for all α, h_α maps orbits of the first system in U_α onto orbits of the second system with $\beta = p(\alpha)$ in $V_{p(\alpha)}$ preserving the direction of time.

We can use topological equivalence to establish the following theorem

Theorem 2.3.1. ***Dynamical Systems on*** $\mathbb{R}^n$ ***II*** *Suppose that* $f(x, \alpha) \in C^1(X)$*, then there is a function* $F(x, \alpha) \in C^1(X)$ *such that*

$$x' = F(x, \alpha) \tag{2.18}$$

defined a dynamical system on X *such that* (2.18) *is topologically equivalent to* (2.16)

A natural question that emerges is when are two systems topologically equivalent? One result that helps resolves this question for systems with hyperbolic equilibria is the Hartman-Grobman theorem [73, 81].

Theorem 2.3.2. ***Hartman-Grobman Theorem*** *Let* $f \in C^1(E)$*,* $x^* \in E$*, and* ϕ^t *be the flow of* (2.16)*. Suppose that* x^* *is a hyperbolic equilibrium. Then* (2.16) *is locally topologically equivalent to its linearization in a neighbourhood around* x^**.*

Given our definition of topological equivalence, it should be clear that the stability properties of equilibria are topological properties that are common to equivalent systems, and the Hartman-Grobman establishes this formally. More generally, we have the following theorem from [100]

Theorem 2.3.3. ***Topological Equivalence of Hyperbolic Equilibria*** *Consider* x^* *and* y^**, two hyperbolic equilibria of* (2.16) *and* (2.17)*. If both equilibria have* n_- *and* n_+ *eigenvalues with negative real parts and positive real parts respectively, then* (2.16) *and* (2.17) *are locally topologically equivalent in neighbourhoods around* x^* *and* y^**.*

These latter two theorems also immediately imply that the stability of a hyperbolic equilibria is a topologically conserved quantity. This can be established without these theorems however. Now that we know when two systems are topologically equivalent, we can easily determine when they are not.

2.3.2 Bifurcation Theory

We now have the necessary machinery to define what a bifurcation is. In particular, we can use the definition of topological equivalence on the two systems

$$\dot{x} = f(x, \alpha), \quad x \in \mathbb{R}^n, \alpha \in \mathbb{R}^p \tag{2.19}$$

$$\dot{x} = f(x, \alpha^*), \quad x \in \mathbb{R}^n, \alpha^* \in \mathbb{R}^p \tag{2.20}$$

Normally, one would expect that if α and α^* are in a small neighourhood of one another, these two systems should be qualtitatively similar and thus topologically equivalent. The emergence of a non-equivalent system defines a bifurcation.

Definition 2.3.5. Bifurcation: Consider the system (2.19) and (2.20). Suppose that at α^*, the system 2.20 is not topologically equivalent to the system 2.19 in a neighbourhood around α^*, with $\alpha \neq \alpha^*$. Then the system (2.20) has undergone a bifurcation and $\alpha = \alpha^*$ is a bifurcation point.

While this seems like a very vague definition of a bifurcation, the key insight is that topological equivalence is mathematically an equivalence class. Thus, if we can generate a set of simple dynamical systems that act as representatives for specific bifurcations and are analytically tractable, then we can understand the phase portrait of a system that has an identical bifurcation. These simple systems are referred to as topological normal forms. These systems typically have the form

$$\dot{\eta} = g(\eta, \beta) \tag{2.21}$$

where $g(\eta, \beta)$ is a low order polynomial.

Definition 2.3.6. Topological Normal Form System (2.21) is called a topological normal form for the bifurcation if any generic system with the equilibrium satisfying the same bifurcation conditions at $\alpha = \alpha^*$ is locally topologically equivalent to (2.21) near the origin.

We need to take note of a few things before we proceed. First, the systems belonging to the equivalence class defined by the normal form have to be "generic"in the sense that they satisfy a specific set of conditions. These conditions ensure the existence of a homeomorphism that transforms any member of the equivalence class into the normal form. This will yield a set of nondegeneracy conditions that have to be satisfied for a system to be considered generic. The second point is that this definition is for local bifurcations of equilibria. These definitions can similarly be extended for global bifurcations, however the work in this thesis is predominantly concerned with local bifurcations of equilibria.

Bifurcations can be categorized by their co-dimension

Definition 2.3.7. co-Dimension The co-dimension of a bifurcation is the number of independent conditions determining the bifurcation

Loosely speaking, the higher the co-dimension (or codim for short) of a bifurcation, the more complex the dynamical system. In this thesis, we will see smooth and nonsmooth bifurcations of co-dimension 1 and 2.

2.3.3 Saddle-Node Bifurcation

The saddle-node bifurcation is one codimension-1 bifurcation of the system (2.16). It occurs when the system has two equilbria that collide with one another at a point. The equilibrium point at the collision has a zero eigenvalue and is often called a saddle-node equilibrium.

Theorem 2.3.4. ***Saddle-Node Bifurcation Theorem*** *Consider the one dimensional system*

$$\dot{x} = f(x, \alpha) \quad x \in \mathbb{R}, \alpha \in \mathbb{R} \tag{2.22}$$

where f is smooth, $f : \mathbb{R}^1 \times \mathbb{R}^1 \to \mathbb{R}^1$ and has $\alpha = 0$, $x = 0$ as an equilibrium point with $\lambda = f_x(0, 0) = 0$. Assume that the following non-degeneracy conditions are satisfied

i $f_{xx}(0, 0) \neq 0$

ii $f_\alpha(0, 0) \neq 0$

Then the system (2.22) is locally topologically equivalent to one of the following systems

$$\dot{y} = \beta \pm y^2$$

Either of the normal forms are very simple to analyze. For example, the system $\dot{y} = \beta + y^2$ has two equilibria, $y_\pm = \pm\sqrt{-\beta}$ and the stability of these two equilibria are determined by $\lambda(y_\pm) = 2y_\pm = \pm 2\sqrt{-\beta}$ and thus $y_+ = \sqrt{-\beta}$ is a stable equilibrium while $y_- = -\sqrt{-\beta}$ is an unstable equilibrium. The two collide at $\beta = 0$, with $y_0 = 0$ forming a semi-stable equilibrium (the saddle-node), and disappear for $\beta > 0$. The bifurcation point occurs at this collision.

2.3.4 Andronov-Hopf Bifurcation

The Andronov-Hopf bifurcation is another co-dimension 1 bifurcation. The Andronov-Hopf or Hopf bifurcation occurs when a limit cycle emerges from an equilibrium point that undergoes a change in stability [7]. All of the qualitative features of the Andronov-Hopf Bifurcation can be determined through analyzing the system.

$$\dot{x}_1 = \alpha x_1 - x_2 - x_1(x_1^2 + x_2^2) \tag{2.23}$$
$$\dot{x}_2 = x_1 + \alpha x_2 - x_2(x_1^2 + x_2^2) \tag{2.24}$$

which should not be surprising, given that this is in fact the topological normal form for this bifurcation. Transforming this system into polar coordinates yields

$$r' = \frac{x_1 x_1' + x_2 x_2'}{r} = \frac{x_1(\alpha x_1 - x_2 - x_1 r^2) + x_2(x_1 + \alpha x_2 - x_2 r^2)}{r} = \alpha r - r^3 \tag{2.25}$$
$$\theta' = \frac{x_1 x_2' - x_2 x_1'}{r^2} = \frac{x_1(x_1 + \alpha x_2 - x_2 r^2) - x_2(\alpha x_1 - x_2 - x_1 r^2)}{r^2} = \frac{r^2}{r^2} = 1 \tag{2.26}$$

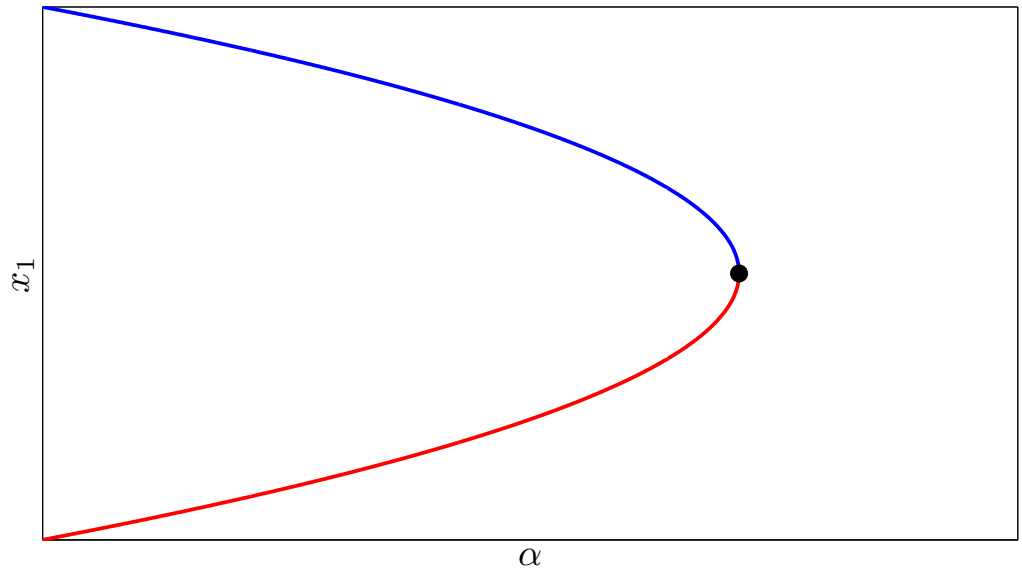


Figure 2.2: Shown above is the bifurcation diagram for the system $\dot{x} = \alpha + x^2$. For $\alpha < 0$, there are two equilibria, $x = \sqrt{-\alpha}$ which is unstable (blue) and $x = -\sqrt{-\alpha}$ which is stable (red). These two equilibria collide at the bifurcation point (black dot).

For $\alpha > 0$, it's clear that this system has a limit cycle with $r = \sqrt{\alpha}$, and for $r > \sqrt{\alpha}$, $r' < 0$ and for $r < \sqrt{\alpha}$, $r' > 0$, thus the limit cycle is stable. The equilibrium point at the origin, $r = 0$ is stable for $\alpha < 0$, and unstable for $\alpha > 0$. At $\alpha = 0$, the system reduces to $r' = -r^3$, and as $r' < 0$ for $r > 0$, the equilibrium point $r = 0$ is stable. To conclude, the only invariant sets for this system are a stable equilibrium point for $\alpha \leq 0$, an unstable equilibrium point for $\alpha > 0$, and a stable limit cycle for $\alpha > 0$. The linearization of the original system shows that the equilibrium points are foci. As a stable limit cycle and an unstable equilibrium point emerge from a stable equilibrium point, this is the supercritical Andronov-Hopf bifurcation point. The subcritical case reverses the stability properties of the equilibria and limit cycles.

Theorem 2.3.5. ***Andronov-Hopf Bifurcation Theorem*** *Consider the two dimensional system*

$$\dot{x} = f(x, \alpha) \qquad x \in \mathbb{R}^2, \alpha \in \mathbb{R} \tag{2.27}$$

where f is smooth, and for all sufficiently small $|\alpha|$, has at the equilibrium $x = 0$ with eigenvalues

$$\lambda_{1,2}(\alpha) = \mu(\alpha) \pm i\omega(\alpha)$$

where $\mu(0) = 0$ and $\omega(0) = \omega > 0$. Let the following nondegeneracy conditions be satisfied

i $\mu'(0) \neq 0$

ii $l_1(0) \neq 0$

Then the system is locally topologically equivalent to one of the following

$$\frac{d}{dt}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \beta & -1 \\ 1 & \beta \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \pm (y_1^2 + y_2^2)\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

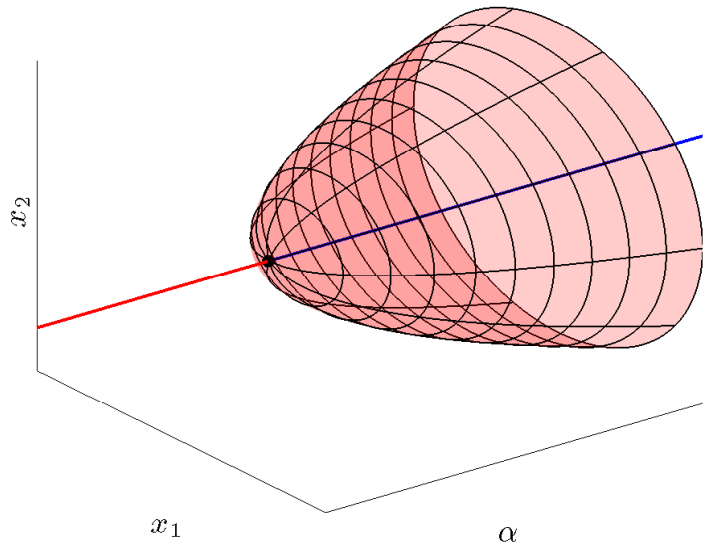
The quantity $l_1(0)$ is referred to as the first Lyapunov Coefficient, and emerges in the reduction process to the normal form. It is quite laborious to compute by hand, however if one can write the dynamical system (2.27) at $\alpha = 0$ in the form

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} P(x,y) \\ Q(x,y) \end{pmatrix} \tag{2.28}$$

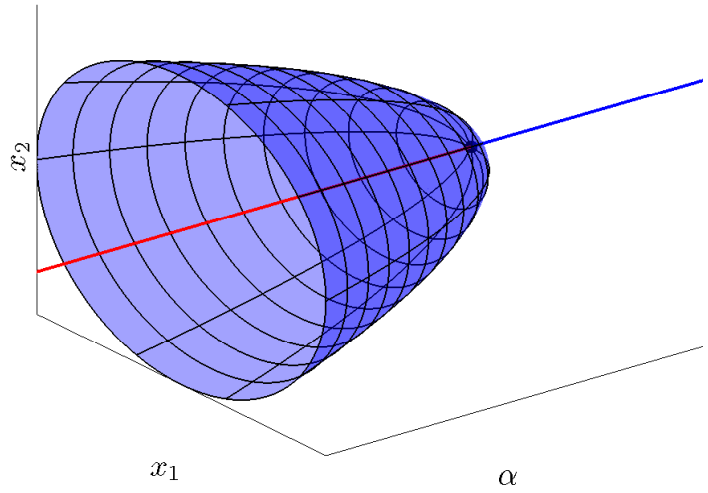
then the first Lypaunov coefficient can be computed by using the formula

$$\begin{aligned} l_1(0) &= \frac{1}{8\omega}\left(P_{xxx} + P_{xyy} + Q_{yyx} + Q_{yyy}\right) \\ &+ \frac{1}{8\omega^2}\left(P_{xy}(P_{yy} + P_{xx}) - Q_{xy}(Q_{xx} + Q_{yy}) - P_{yy}Q_{yy} + P_{xx}Q_{xx}\right) \end{aligned} \tag{2.29}$$

the derivation of which can be found in [74]. The bifurcation diagram is shown in 2.3.



(a) Supercritical Hopf Bifurcation



(b) Subcritical Hopf Bifurcation

Figure 2.3: Shown above are the two possible Andronov-Hopf bifurcations. In figure 2.3(a), the system undergoes a supercritical Hopf bifurcation resulting in the emergence of a stable limit cycle when the parameter α is varied. The equilibrium point changes from stable to unstable. In 2.3(b), the equilibrium point changes from stable to unstable again, but this occurs simultaneously with the disappearance of an unstable limit cycle.

2.3.5 Co-Dimension 1 Non-smooth Bifurcations in Piecewise Smooth Continuous Systems

In general, the concept of topological equivalence has not been extended to nonsmooth dynamical systems [48]. Thus, one cannot consider the unfoldings for the bifurcations here as normal forms in the classical sense. However, they do provide important information about the local behavior of a system that undergoes a discontinuity induced bifurcation. Returning to our non-smooth dynamical systems definition:

$$\dot{x} = \begin{cases} f_1(x,\alpha) & H(x,\alpha) > 0 \\ f_2(x,\alpha) & H(x,\alpha) < 0 \end{cases} \tag{2.30}$$

As the system is piecewise smooth continuous, then one can write

$$f_2(x,\alpha) = f_1(x,\alpha) + H(x,\alpha)G(x,\alpha) \tag{2.31}$$

where $G(x,\alpha) \neq 0$ if $H(x,\alpha) = 0$ [48]. Additionally, we can define the following points.

Definition 2.3.8. Pseudo-Equilibrium A pseudo-equilibrium is an equilibrium $x^*(\alpha)$ of $f_1(x,\alpha)$ or $f_2(x,\alpha)$ such that $H(x^*,\alpha) = 0$.

More informally, a pseudo-equilibrium is an equilibrium that lies on the switching manifold.

Definition 2.3.9. Virtual Equilibrium A virtual equilibrium is an equilibrium of $f_1(x,\alpha)$ ($f_2(x,\alpha)$) where $H(x,\alpha) < 0$ (> 0)

Using the definitions in [48], we can consider what are called boundary equilibrium bifurcations (BEB). These are co-dimension 1 bifurcations of non-smooth systems.

Definition 2.3.10. Boundary Equilibrium Bifurcations A boundary equilibrium bifurcation occurs at $\alpha = \alpha^*$ if $f_1(x^*,\alpha^*) = 0$, $H(x^*,\alpha^*) = 0$, and

$$\det\left(\left.\frac{\partial f_i(x,\alpha)}{\partial x}\right|_{(x^*,\alpha^*)}\right) \neq 0, \quad i = 1,2 \tag{2.32}$$

The final condition ensures that the pseudo-equilibrium of (2.30) is a hyperbolic equilibrium of $\dot{x} = f_1(x,\alpha)$ and $f_2(x,\alpha)$, and thus no smooth bifurcations are simultaneously occurring. If condition (2.32) is not satisfied, there is a higher co-dimension bifurcation.

There are two co-dimension 1 cases that occur and are commonly referred to as persistence and nonsmooth fold bifurcations.

Definition 2.3.11. Persistence Bifurcation The system (2.30) exhibits a persistence bifurcation at $\alpha = 0$ if when α is varied, one branch of regular and one branch of virtual equilibria cross at the boundary and exchange properties. Letting $x_1(\alpha)$ ($x_2(\alpha)$) be the equilibria of $f_1(x,\alpha)$ ($f_2(x,\alpha)$) then this yields

$[\alpha < 0\ ]$ $f_1(x_1(\alpha), \alpha) = 0$, $H(x_1(\alpha), \alpha) > 0$ and $f_2(x_2(\alpha), \alpha) = 0$, $H(x_2(\alpha), \alpha) > 0$

$[\alpha > 0\ ]$ $f_1(x_1(\alpha), \alpha) = 0$, $H(x_1(\alpha), \alpha) < 0$ and $f_2(x_2(\alpha), \alpha) = 0$, $H(x_2(\alpha), \alpha) < 0$

Definition 2.3.12. Nonsmooth Fold Bifurcation The system (2.30) exhibits a nonsmooth fold bifurcation at $\alpha = 0$ if when α is varied, two branches of regular equilibria collide and become virtual equilibria after the collision. Letting $x_1(\alpha)$ ($x_2(\alpha)$) be the equilibria of $f_1(x, \alpha)$ ($f_2(x, \alpha)$) then this yields

$[\alpha < 0\ ]$ $f_1(x_1(\alpha), \alpha) = 0$, $H(x_1(\alpha), \alpha) > 0$ and $f_2(x_2(\alpha), \alpha) = 0$, $H(x_2(\alpha), \alpha) < 0$

$[\alpha > 0\ ]$ $f_1(x_1(\alpha), \alpha) = 0$, $H(x_1(\alpha), \alpha) < 0$ and $f_2(x_2(\alpha), \alpha) = 0$, $H(x_2(\alpha), \alpha) > 0$

These bifurcations are local, co-dimension 1 bifurcations of equilibria for piecewise-smooth continuous systems. If one linearizes the functions $f_1(x, \alpha), f_2(x, \alpha)$ and $H(x, \alpha)$, one can easily prove the following theorem from [48], for piecewise smooth continuous systems.

Theorem 2.3.6. ***Persistence/Nonsmooth Fold Bifurcation Theorem*** *Consider system (2.30) and suppose that at $\alpha = \alpha^*$ the system has a boundary equilibrium point, $x_1(\alpha^*) = x_2(\alpha^*)$ and define the following matrices*

$$A = \left.\frac{\partial f_1(x,\alpha)}{\partial x}\right|_{(x_1(\alpha^*),\alpha^*)}, \quad B = \left.\frac{\partial f_1(x,\alpha)}{\partial \alpha}\right|_{(x_1(\alpha^*),\alpha^*)} \tag{2.33}$$

$$C = \left.\frac{\partial H(x,\alpha)}{\partial x}\right|_{(x_1(\alpha^*),\alpha^*)}, \quad D = \left.\frac{\partial H(x,\alpha)}{\partial \alpha}\right|_{(x_1(\alpha^*),\alpha^*)} \tag{2.34}$$

$$E = G(x_1(\alpha^*), \alpha^*). \tag{2.35}$$

Assuming that the following nondegeneracy conditions hold

$$D - CA^{-1}B \neq 0 \tag{2.36}$$

$$1 + CA^{-1}E \neq 0 \tag{2.37}$$

Then a persistence bifurcation occurs if $1 + CA^{-1}E > 0$ and a nonsmooth fold occurs if $1 + CA^{-1}E < 0$

While the systems we consider are piecewise smooth continuous, the majority of the literature is not directly applicable due to the specific form of discontinuity in the higher order derivatives that arise in the derived mean-field systems. In particular, as we shall see in Chapter 4, the square-root functions that emerge yield problems in applying the existing literature on piecewise smooth continuous systems. We include the co-dimension 1 bifurcations here as a point of comparison. Higher co-dimension bifurcations in piecewise smooth continuous systems have been analyzed however in the literature [48, 67, 107, 164, 165, 166]

2.3.6 Bogdanov-Takens Bifurcation

The Bogdanov-Takens bifurcation is a co-dimension 2 bifurcation that occurs in a smooth system when an equilibrium has two zero eigenvalues at the bifurcation point. As with other high co-dimension bifurcations, it often results in lower co-dimension bifurcations emanating out from the bifurcation point.

Theorem 2.3.7. ***Bodganov-Takens Bifurcation*** *Suppose that a planar system,*

$$\dot{x} = f(x, \alpha), x \in \mathbb{R}^2, \alpha \in \mathbb{R}^2$$

with smooth f, has at $\alpha = 0$ the equilibrium $x = 0$ with a double zero eigenvalue, $\lambda_{1,2}(0) = 0$. Assume that the following nondegeneracy conditions hold:

i. The Jacobian is non-zero

ii. $a_{20} + b_{11}(0) \neq 0$ *(see [100] for formula)*

iii $b_{20} \neq 0$ *(see [100] for formula)*

iii the map

$$(x, \alpha) \to (f(x, \alpha), tr(Df(x, \alpha)), det(Df(x, \alpha)))$$

is regular at $(0, 0)$

then the system is locally topologically equivalent near the equilibrium to one of the following normal forms

$$\eta_1' = \eta_2 \tag{2.38}$$
$$\eta_2' = \beta_1 + \beta_2\eta_1 + \eta_1^2 \pm \eta_1\eta_2 \tag{2.39}$$

The nondegeneracy quantities can be found in [100], in addition to the full proof of this theorem. The analysis of the topological normal form reveals four branches of co-dimension 1 bifurcations that emerge

[SN1] There is a branch of saddle-node bifurcations, where a stable node and a saddle coalesce

[SN2] There is a second branch of saddle-node bifurcations where an unstable node and a saddle coalesce

[H1] There is a branch of Andronov-Hopf Bifurcations. The branch is either super-critical or sub-critical, resulting in an unstable or stable limit cycle, depending on which specific normal form the original system is topologically equivalent to.

[SH1] The limit-cycle generated from the Andronov-Hopf bifurcation is destroyed through the interaction with a saddle point generated from the saddle-node bifurcation. This is called a homoclinic bifurcation (or saddle homoclinic) as a homoclinic limit cycle exists at the bifurcation point.

2.3.7 Cusp-Bifurcation

The cusp bifurcation occurs when one of the nondegeneracy conditions of the saddle-node bifurcation are violated. In particular, when $f_{xx}(0,0) = 0$, a higher co-dimension bifurcation occurs as we now have a second condition that the system must satisfy.

Theorem 2.3.8. ***Cusp Bifurcation*** *Suppose that a one-dimensional system*

$$\dot{x} = f(x, \alpha), x \in \mathbb{R}^1, \alpha \in \mathbb{R}^2$$

with smooth f has at $\alpha = 0$ the equilibrium $x = 0$, with $\lambda = 0$, $f_{xx}(0,0) = 0$. Assume that the following nondegeneracy conditions hold

i. $f_{xxx}(0,0) \neq 0$

ii. $f_{\alpha_1} f_{x\alpha_2}(0,0) - f_{\alpha_2} f_{x\alpha_1}(0,0) \neq 0$

Then the system is locally topologically equivalent to one of the following

$$\eta' = \beta_1 + \beta_2 \eta \pm \eta^3$$

This system is characterized by the presence of two branches of saddle-node bifurcations that collide with one another tangentially at the cusp bifurcation point.

2.3.8 Bautin/Generalized Hopf Bifurcation

The Bautin bifurcation occurs when the first Lyapunov coefficient of the Hopf bifurcation is 0, thus it occurs when a Hopf bifurcation is degenerate.

Theorem 2.3.9. ***Bautin Bifurcation*** *Suppose that a planar system*

$$\dot{x} = f(x, \alpha), x \in \mathbb{R}^2, \alpha \in \mathbb{R}^2$$

with smooth f has the equilibrium $x = 0$ with eigenvalues $\lambda_{1,2} = \mu(\alpha) \pm i\omega(\alpha)$ where $\omega(0) = \omega > 0$. For $\alpha = 0$, let the Bautin bifurcation conditions hold:

$$\mu(0) = 0, \quad l_1(0) = 0$$

where $l_1(\alpha)$ is the first Lyapunov coefficient. Assume that the following nondegeneracy conditions hold

i. $l_2(0) \neq 0$, *where $l_2(0)$ is the second Lyapunov coefficient (see [100])*

ii. The map $\alpha \to (\mu(\alpha), l_1(\alpha))$ is regular at $\alpha = 0$

then the system is locally topologically equivalent to one of the following

$$\dot{z}' = (\beta_1 + i)z + \beta_2 z|z|^2 \pm z|z|^4$$

which can be converted to their planar form via $\eta_1 = Re(z)$, $\eta_2 = Imag(z)$

Analysis of the normal form reveals three co-dimension 1 bifurcation branches

[H-Sup] A super-critical branch of Andronov-Hopf bifurcations, creating a stable limit cycle

[H-Sub] A sub-critical branch of Andronov-Hopf bifurcations, creating an unstable limit cycle

[SNP] A branch of saddle-node of limit cycles bifurcation. These occur when a stable and unstable limit cycle collide and destroy each other. This is similar to the saddle-node bifurcation of equilibria.

2.4 Stable, Unstable, and Center Manifolds

The stable, unstable, and center manifolds are sets of trajectories in the phase space that have the same common behavior. To simplify some of the theorems, we will primarily consider the system without parameters:

$$\dot{x} = f(x), \quad x \in X. \tag{2.40}$$

We can also assume that if this system has an equilibrium, x^*, then it is at the origin $x^* = 0$, without loss of generality.

Definition 2.4.1. Local Stable and Unstable Manifolds Let x^* be an equilibrium of (2.40) and N a neighbourhood around x^*. The local stable and unstable manifolds of x^* are defined by

$$\begin{aligned} S &= \left\{x \in N | \phi^t(x) \to x^* \text{ as } t \to \infty \text{ and } \phi^t(x) \in N, \forall t \geq 0\right\} \\ U &= \left\{x \in N | \phi^t(x) \to x^* \text{ as } t \to -\infty \text{ and } \phi^t(x) \in N, \forall t \leq 0\right\} \end{aligned}$$

These local manifolds contain all the local trajectories that converge to the equilibrium point either in forward time or reverse time. They can be defined globally as well

Definition 2.4.2. Global Stable and Unstable Manifolds The global stable and unstable manifolds of x^* are defined by

$$W^s(x^*) = \cup_{t \leq 0} \phi^t(S) \tag{2.41}$$
$$W^u(x^*) = \cup_{t \geq 0} \phi^t(U) \tag{2.42}$$

The reason why these particular sets are important is because they are invariant sets in the phase space.

Definition 2.4.3. Invariant Sets A set is called invariant with respect to ϕ^t if $\phi^t(S) \subset S$ for all $t \in \mathbb{R}$. The set is positively invariant (negatively invariant) with respect to ϕ^t if $\phi^t(S) \subset S$ for all $t \geq 0$ ($t \leq 0$)

Invariant sets allow us to simplify the analysis of dynamical systems as the trajectories that begin in these sets are confined to these sets. If the invariant set considered is of lower dimension than the full dynamics, the analysis of the dynamics on the invariant set becomes significantly easier than the original higher dimensional dynamical system. Simple kinds of invariant sets include equilibria and limit cycles. The existence of the local stable and unstable manifolds is established by

Definition 2.4.4. Existence of the Stable Manifold Let $f \in C^1(X)$ and $x^* \in X$ be a hyperbolic equilibrium point of the system (2.40). Then there exists local stable and unstable manifolds S and U that are of the same dimension as n_- and n_+, the number of eigenvalues of the Jacobian matrix that have negative and positive real parts.

While the stable and unstable manifolds are important for analyzing certain global behaviors, the local behavior of hyperbolic equilibria is adequately explained by the linearization alone. However, the real importance of these constructions is when we have nonhyperbolic equilibria.

Theorem 2.4.1. ***Center Manifold Theorem I*** *Let $f \in C^r(X)$, $r \geq 1$, and let $x^* \in X$ be an equilibrium point of the differential equation (2.40). There exists manifolds $W^s(x^*), W^u(x^*), W^c(x^*)$ which are n_-, n_+ and n_0 dimensional where these quantities are the number of eigenvalues of $Df(x^*)$ that have negative, positive, and zero real parts, respectively. These manifolds are invariant with respect to the flow ϕ^t and the stable and unstable manifolds are unique, and $C^r(X)$. The center manifold is $C^{r-1}(X)$ and need not be unique.*

The proof can be found in [32]. The real importance of the center manifold theorem is that it allows us to analyze the behavior of a system with a non-hyperbolic equilibrium by looking at a much simpler, topologically equivalent system. This will be demonstrated by the following two theorems

Theorem 2.4.2. ***Center Manifold Theorem II*** *Let $f \in C^r(X)$, where $0 \in X$, and $r \geq 1$. Suppose that $f(0) = 0$ and $Df(0)$ has n_- eigenvalues with negative real part, n_+ eigenvalues with positive real part, and $n_0 = n - n_- - n_+$ eigenvalues with zero real part. Then the differential equation (2.40) can be written in the form*

$$x' = Cx + F(x, y, z) \tag{2.43}$$
$$y' = Py + G(x, y, z) \tag{2.44}$$
$$z' = Qz + H(x, y, z) \tag{2.45}$$

where $(x, y, z) \in \mathbb{R}^{n_0} \times \mathbb{R}^{n-} \times \mathbb{R}^{n+}$, F, G *and* H *only contain quadratic and higher order terms. The matrices* P *and* Q *are negative definite and positive definite, respectively while the matrix* C *has eigenvalues with zero real parts exclusively. Furthermore, there exist a* $\delta > 0$ *and functions* $h_1 \in C^r(B(\delta), \mathbb{R}^{n-})$ *and* $h_2 \in C^r(B(\delta), \mathbb{R}^{n-})$ *that define the local manifold*

$$W^{n_0}_{loc}(0) = \{(x, y, z) \in \mathbb{R}^{n_0} \times \mathbb{R}^{n-} \times \mathbb{R}^{n+} | y = h_1(x), \quad z = h_2(x), \quad \|x\| < \delta \in\}$$

and satisfies

$$\begin{aligned} Dh_1(x)\left[Cx + F(x, h_1(x), h_2(x))\right] - h_1(x) - G(x, h_1(x), h_2(x))) &= 0 \qquad (2.46) \\ Dh_2(x)\left[Cx + F(x, h_1(x), h_2(x))\right] - h_2(x) - H(x, h_1(x), h_2(x)) &= 0 \qquad (2.47) \end{aligned}$$

for $\|x\| < \delta$. *The flow on* $W^c_{loc}(0)$ *is defined by the solutions of*

$$x' = Cx + F(x, h_1(x), h_2(x)) \qquad (2.48)$$

for all $x \in \mathbb{R}^c$ *with* $\|x\| < \delta$

The proof of this theorem can once again be found in [32]. There are two important things to note about this theorem

1. The local center manifold can be determined by the partial differential equations (2.46)-(2.47)
2. The original n-dimensional system 2.3 has an embedded n_0-dimensional dynamical system on the center manifold

Thus we can solve for the center manifold, and the invariant dynamics on it. These two points should immediately demonstrate the real power of the following theorem

Theorem 2.4.3. ***Center Manifold Theorem III*** *Let* $f \in C^1(X)$, *and* $0 \in X$. *Suppose that* $Df(0) = diag[C, P, Q]$ *where* $[C, P, Q]$ *are as in Theorem 2.4.2. Then there exists* C^1 *functions* $h_1(x)$ *and* $h_2(x)$, *satisfying* (2.46)-(2.47) *in a neighbourhood of the origin such that the system* (2.40) *is topologically equivalent to the* C^1 *system*

$$\begin{aligned} x' &= Cx + F(x, h_1(x), h_2(x)) \qquad (2.49) \\ y' &= Py \qquad (2.50) \\ z' &= Qz \qquad (2.51) \end{aligned}$$

The proof can be found in [32].

Theorem 2.4.3 is one of the most powerful results in dynamical systems for various reasons. We can immediately note that the dynamics of (2.49)-(2.51) are much simpler then the original system, yet still locally topologically equivalent. The dynamics of x, y, z

are independent of one another, and the dynamics of x and y are explicitly integrable, and either decay exponentially (y) or diverge exponentially (z). Thus, the only dynamics that actually have to be analyzed for a non-hyperbolic system are the dynamics of the center manifold equation, (2.49). If the system (2.40) is high dimensional with a non-hyperbolic equilibria where $n_0 \ll n$, then we can immediately reduce the local behavior to a much lower dimensional dynamical system.

Furthermore, we have seen that non-hyperbolic equilibria show up at bifurcation points for the various local bifurcations. Thus, it is natural to think of applying the center manifold. All these theorems and definitions can easily be extended to parameter dependent systems. More importantly however, we have seen that the topological normal forms and the bifurcation theorems were only defined in section 2.3.2 for specific dimensions. For example, the saddle-node bifurcation and its normal form reduction were performed on a one-dimensional dynamical system. However, the saddle-node bifurcation (and any other bifurcation) can occur for higher dimensional dynamical systems. One can use the center manifold theorem to establish when these bifurcations are non-degenerate as we can still establish topological equivalence with a normal form, only we are establishing the equivalence with the dynamical system on the center manifold ((2.49)) as opposed to the full n-dimensional dynamics.

For a concrete example, one can use the center manifold analysis to prove Sotomayor's theorem

Theorem 2.4.4. ***Sotomayor's Theorem*** *Suppose that the system (2.3) has x^* as an equilibrium at $\alpha = \alpha^*$ which has the Jacobian $A = Df(x^*, \alpha^*)$ with a single zero eigenvalue. Let v be the eigenvector of A and w be the eigenvector of A^T both corresponding to the zero eigenvalue of their respective matrices. Furthermore, assume that A has k eigenvalues with negative real part and the following nondegeneracy conditions are satisfied*

$$\left\langle w, \left.\frac{\partial f}{\partial \alpha}\right|_{(x^*,\alpha^*)} \right\rangle \neq 0, \quad \left\langle w, \left. D^2 f(x,\alpha) \right|_{(x^*,\alpha^*)} \neq \right\rangle 0 \tag{2.52}$$

Then the system has a generic saddle-node bifurcation.

Note that the $\langle\,\rangle$ brackets in this theorem denote the standard inner product in $\mathbb{R}^n$. This is the only time they will be used to denote this. In later chapters they will be used to denote a population average. The proof of Sotomayor's theorem can be found in [145].

2.5 Asymptotic Expansions of Integrals

Throughout this thesis, we will see integrals of the form

$$I(x) = \int_a^b f(t) \exp(x\phi(t))\, dt \tag{2.53}$$

or integrals of the form

$$I(x) = \int_a^b f(t)e^{ix\psi(t)}\,dt. \tag{2.54}$$

This section will review results that determine their asymptotic behavior when $x \to \infty$. The bulk of this section is from [17].

2.5.1 Watson's Lemma

Watson's Lemma resolves the asymptotics of integrals of the form (2.53) under the special case that $\phi(t) = -t$ and $a = 0$. The only requirement imposed is that $f(t)$ is continuous on $[0, b]$ and that

$$f(t) \sim t^\alpha \sum_{n=0}^{\infty} a_n t^{\beta n}, \quad t \to 0^+$$

where $\alpha > -1$ and $\beta > 0$. If these two conditions hold, then one can show the following asymptotic behavior

$$\int_0^b f(t)\exp(-xt)\,dt \sim \sum_{n=0}^{\infty} \frac{a_n \Gamma(\alpha + \beta n + 1)}{x^{\alpha+\beta n+1}}, \quad x \to \infty \tag{2.55}$$

where $\Gamma(x)$ is the gamma function. The bulk of the proof of this can be found in [17].

2.5.2 Laplace's Method

Laplace's method extends Watson's lemma to the general integral form (2.53). Unlike Watson's lemma however, Laplace's method has several sub-cases that have to be considered. In particular, these cases are delineated by where the maximum of $\phi(t)$ is on the interval $[a, b]$ and how many of the derivatives of $\phi(t)$ vanish at this point. Suppose that the max of $\phi(t)$ is at $t = a$, then $\phi'(a) < 0$ and one can show that

$$I(x) = \int_a^b f(t)\exp(x\phi(t))\,dt \sim -\frac{f(a)\exp(x\phi(a))}{x\phi'(a)}, \quad x \to \infty. \tag{2.56}$$

If, however, the max is at $\phi(t) = b$, then $\phi'(b) > 0$ and one can show that

$$I(x) = \int_a^b f(t)\exp(x\phi(t))\,dt \sim \frac{f(b)\exp(x\phi(b))}{x\phi'(b)}, \quad x \to \infty \tag{2.57}$$

If, however, there is some interior max of $\phi(t)$, $\phi(c)$ for $a < c < b$, then $\phi'(c) = 0$ and $\phi^{(p)}(c) < 0$, with p being an even integer and one can again show

$$I(x) = \int_a^b f(t)\exp(x\phi(t))\,dt \sim \frac{2\Gamma(p^{-1})(p!)^{p^{-1}}}{p(-x\phi^{(p)}(c))^{p^{-1}}} f(c)\exp(x\phi(c)), \quad x \to \infty \tag{2.58}$$

2.5.3 The Method of Stationary Phase

The method of stationary phase deals with the case where $\phi(t)$ is complex valued, and $\phi(t) = i\psi(t)$:

$$I(x) = \int_a^b f(t) \exp(xi\psi(t))\, dt \tag{2.59}$$

This yields an oscillatory integrand with faster oscillations as $x \to \infty$. The leading order asymptotic behavior of the integral is given by

$$\int_a^b f(t) \exp(xi\psi(t))\, dt = \frac{f(t)}{ix\psi'(t)} \exp(ix\psi(t))\Big|_{t=a}^{t=b}, \quad x \to \infty \tag{2.60}$$

which is valid provided that $\phi'(t)$ is non-zero for all $t \in [a, b]$. Suppose however there was some c where $\phi'(c) = 0$. Then we can merely rewrite (2.59) as the

$$I(x) = \int_a^c f(t) \exp(xi\psi(t))\, dt + \int_c^b f(t) \exp(xi\psi(t))\, dt = I_1(x) + I_2(x) \tag{2.61}$$

and one can merely consider integrals where $\phi'(t) = 0$ at a boundary point. This simultaneously deals with all cases that emerge. Suppose that this occurs at the boundary point $t = a$ for the first $p - 1$ orders $\phi'(a) = \phi^{(2)}(a) = \ldots \phi^{(p-1)}(a) = 0$. Then one can show that

$$I(x) \sim f(a) \exp\left(ix\psi(a) + \frac{ip\pi}{2}\right) \left(\frac{p!}{x|\psi^{(p)}(a)|}\right)^{1/p} \frac{\Gamma(1/p)}{p} \quad x \to \infty. \tag{2.62}$$

See [17] for the full details.

Chapter 3

Biological and Mathematical Prerequisites

In this chapter, we will introduce neurons and their general functions in section 3.1, in addition to the various approaches one can use to model neurons in section 3.2. From there, we will see how realistic models for neurons can be coupled together with realistic models for synapses to form networks in section 3.3. As the networks present in the brain are large, we can asymptotically look at the large network limit, and derive a partial differential equation (PDE) that governs the evolution of the entire network in section 3.4. This PDE will serve as an intermediate equation for the purposes of arriving at a mean-field system in the subsequent chapters.

3.1 Neurons, Action Potentials and Synapses

In order to process and relay information, the mammalian nervous system uses a network of cells called neurons that are coupled together at junctions called synapses, as shown in Figure 3.1. The primary way neurons perform any function is through the maintenance of a transmembrane voltage potential. This potential difference across the plasma membrane of any neuron is maintained through an electrochemical gradient across the membrane by the different concentrations of ions on the inside and outside of the plasma membrane. This electrochemical gradient can be controlled and altered by ion channels on the plasma membrane in addition to ion pumps [170].

While the potential difference across the membrane can remain static, neurons actually use oscillations in the membrane potential to convey information. These oscillations are typically referred to as spikes (see Figure 3.1). This is due to the fact that the oscillations typically contain a slow component and a fast upward swing. It is thought that networks of neurons convey information in the rate of spiking, in the synchrony of spiking across neurons in the network, and in the precise timing between spikes. The spikes are

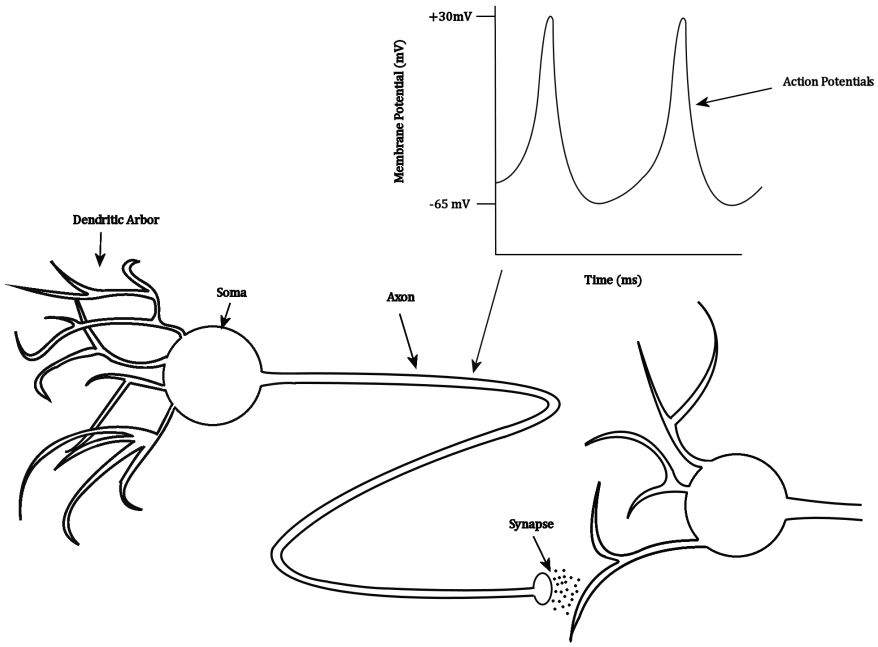


Figure 3.1: Neurons generate action potentials (also referred to as spikes) by integrating all the signals from their dendritic arbor at the junction between the neurons cell body, or soma, and the axon terminal. The action potentials propagate down the axon and synapse onto multiple targets typically (one is shown for simplicity). The action potential reaches a terminal bouton, the end of the axon terminal. Here, it synapses onto the dendritic arbor of another neuron. At the synapse, the presynaptic neuron secretes neurotransmitter into the synaptic cleft as a result of the action potential. The neurotransmitter causes changes in the permeability of the postsynaptic membrane to ions, thereby producing postsynaptic currents. The currents from multiple synapses subsequently sum up again at the action potential initiation site in the post synaptic neuron, which may or may not cause action potentials, depending on the relative magnitude and direction of the currents.

transmitted through biological cables attached to the neurons called axons. These axons arrive at other neurons and connect through synapses. The spikes cause an influx of ions at the presynaptic neurons which triggers release of chemicals called neurotransmitters. These neurotransmitters diffuse across the synaptic cleft and arrive at the postsynaptic neuron and thereby trigger the opening (or closing) of ion channels. This has the effect of generating (or inhibiting) the creation of currents in the postsynaptic neuron. All the cumulative effects of the synapses on the post-synaptic neuron sum together in the soma to either initiate (or stop) spike generation in the post-synaptic neurons axon.

While this entire process is fairly complicated, both neurons and synapses can be modeled to a great deal of accuracy. Thus, with sufficient computing resources, one can model an entire network. In section 3.2, we will introduce how neurons and synapses, and thus networks, can be modeled.

3.2 Modeling Neurons

One can model neurons in various ways, depending on the level of detail required. For a more bottom-up approach, conductance based models are used which are based on the underlying biophysics of action potential generation, as described in section 3.2.1. However, due to the complexity involved in a conductance based model, various simplifications have been proposed. Integrate-and-fire models reproduce the phenomenological behavior of the conductance based models and can be fit to actual neurons, as described in section 3.2.2. An even simpler approach is to merely use a sigmoid curve to model the firing rate of spikes, as described in section 3.2.3.

3.2.1 Conductance Based Models

The individual ion channels on a plasma membrane stochastically transition between different states. The different states can make the plasma membrane either more permeable (such as open states) or less permeable (such as closed or inactivated states) to the flow of ions. These channels can be formally described as Markov processes. For example, the sodium channel studied in the Nobel prize winning work of Hodgkin and Huxley [87] can be described by Markov equations in figure 3.2. With enough ion channels, present in the plasma membrane, one can formally transition to a system of differential equations known as a conductance based model.

One can model a neuron with great detail by using a conductance based model. In a conductance based model, the variation of the potential difference of a neuron, V is modeled as a differential equation along with various gating variables m, n, h, etc. that are related to the kinetics that can be thought of as the proportion of ion channels that are in any particular state. For example, the original Hodgkin-Huxley model is given by the four

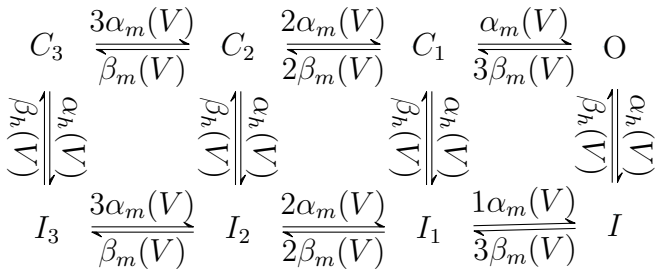


Figure 3.2: The Markov chain describing the sodium channel opening and closing. The state C_i corresponds to i gates being closed, while the state I_i corresponds to the state that i gate are inactivated while state I is an inactivated state that the open state can transition to. The open state O is the only state that admits current flow. The transition rates of the state changes are voltage dependent.

differential equations:

$$\begin{aligned} C\dot{V} &= -I_{Na} - I_K - I_L + I \\ &= \bar{g}_{Na} m^3 h(E_{Na} - V) + \bar{g}_K n^4 h(E_K - V) + \bar{g}_L(V - E_L) + I \qquad (3.1) \\ \dot{m} &= \alpha_m(V)(1-m) - \beta_m(V)m \qquad (3.2) \\ \dot{n} &= \alpha_n(V)(1-n) - \beta_n(V)n \qquad (3.3) \\ \dot{h} &= \alpha_m(h)(1-h) - \beta_h(V)h. \qquad (3.4) \end{aligned}$$

Thus, the membrane potential for the voltage is modeled as an RC circuit. The quantity $E_i - V$ is typically referred to as the driving force, while the term in front of it is referred to as the conductance, $g(t)$, and is the reciprocal of the resistance. For example, the conductance of the sodium current is $g_{Na}(t) = \overline{g_{Na}} m^3(t)h(t)$ where $\bar{g}_{Na}$ is the unitary synaptic conductance of an indivdual sodium ion channel.

A sample voltage trace ($V(t)$) generated from these equations is shown in Figure 7.2(a). An important thing to immediately note is that it appears the spiking behavior present in these conductance models (and real neurons) is an oscillation that has two time scales. The actual action potential or spike occurs on a very fast time scale, while the rest of the oscillation occurs on a much slower time scale. This is stereotypical of most neurons and has been exploited by many individuals to generate much simpler models of neurons which we will see in section 3.2.2.

Conductance based models have had a long and storied history in computational neuroscience. Starting with the original Hodgkin-Huxley model, various currents have been modeled with conductance based formalism. Other conductance based models include the Morris-Lecar model of the barnacle muscle fiber [127], the Connor-Stevens model [38], and the Wang-Buzsaki model [187], to name a few. Smooth, phenomenonological models that reproduce the underlying behaviors of these conductance based models have also been suggested. Some examples include the Hindmarsh-Rose model [86], and the Fitzhugh-Nagumo model [65, 128].

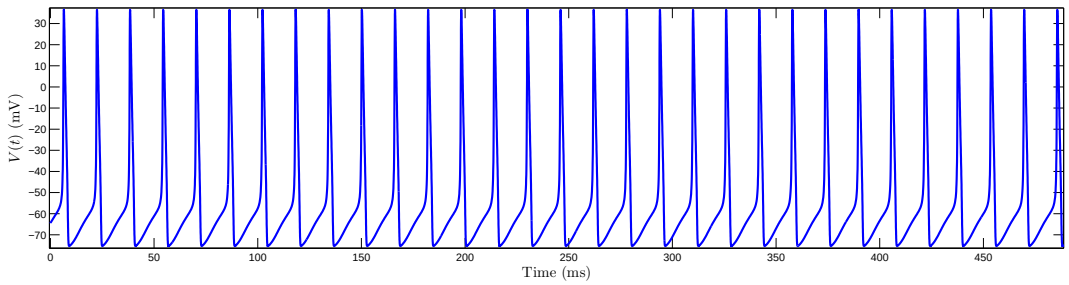
An important thing to note is that while these equations are based on the underlying physics of ion channels and neurons, they are still a simplification. A typical neuron is a cell with many long and tapered processes. In order to fully model a neuron, one decomposes these processes into compartments. Each compartment is modeled by a conductance based model that varies depending on the ion channel(s) present in each compartment. These compartments are then coupled together through currents that travel between adjacent compartments. The equations (3.1)-(3.4) on their own can still be used to model neurons however, and this approach is typically referred to as a point model. In a point model, the differential equations model the neuron only at a single point, typically assumed to be the axon-hillock, the starting point of spike generation in a neuron. Unfortunately, point models that use conductance based equations cannot reproduce the rapid upstroke of an action potential and other modelling approaches are required [131].

A multi-compartment conductance based model has many advantages over other forms of neuron modeling. In particular, with a suitably fine compartmentalization, one can model a neuron with a great deal of accuracy. Additionally, all the parameters of a conductance based model can experimentally be fit to an actual neuron, although with a great deal of difficulty, as the underlying model is based on the underlying biophysics of the cells. Furthermore, a multi-compartment model can reproduce the rapid upstroke of an action potential where as a single-compartment or double compartment conductance based model usually cannot [117, 131]

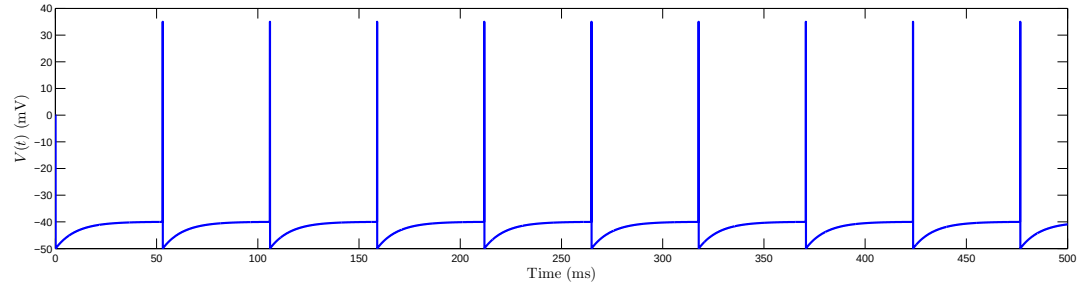
While these models are exceptionally accurate for the behavior of an actual neuron, they have several draw backs. First, generating a conductance based model from a biological neuron is an arduous process requiring intracellular recordings and the fitting of many parameters. For an idea of how difficult it is to fit even a single-compartment conductance based model, the authors in [147] have generated a database of 1.7 million simulations by varying 8 parameters for a single-compartment conductance based model with the goal of researchers using this database to reproduce the behaviors they find in actual neurons, as opposed to ad-hoc methods of picking parameters or fitting them to data. Additionally, they are very complicated to analyze once put into large networks. Furthermore, given that each neuron is governed by multiple differential equations, simulating a network to any biological scale requires extreme computing resources, hence the need for supercomputers [91, 114]. Finally, due to the analytical intractability, it is very difficult to determine the possibly low dimensional dynamics of an entire network via any kind of reduction.

3.2.2 Integrate-and-Fire Models

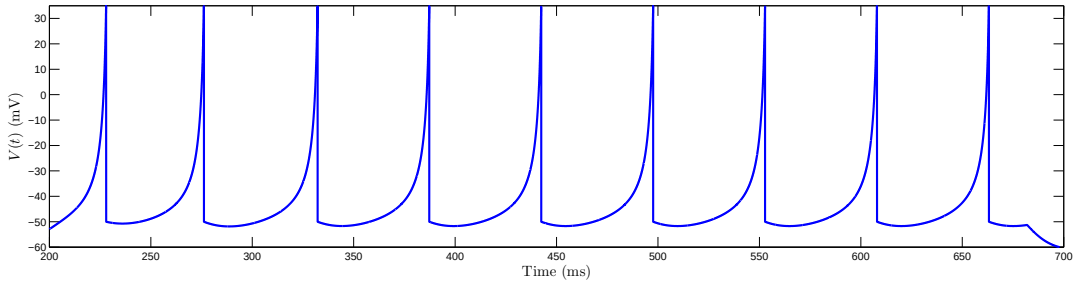
Due to some of the disadvantages of multi-compartment modeling, various simplifications have been proposed. One of the simplest approaches for modeling is the integrate-and-fire approach. In particular, this approach models the spiking as a discontinuous oscillator. For example, if we consider a single neuron as simply a passive RC circuit with no spike



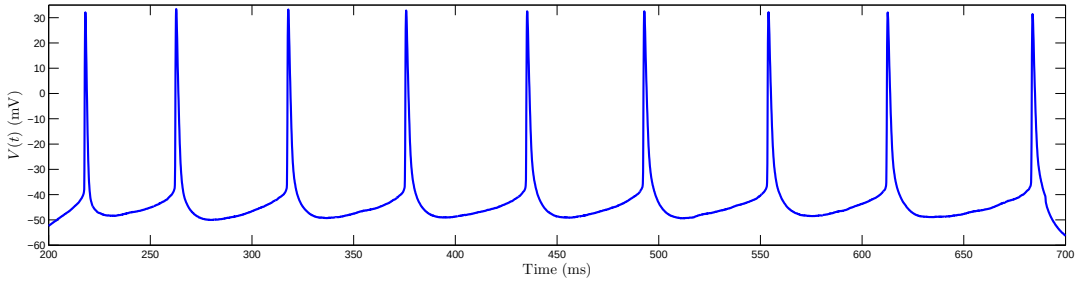
(a) Hodkin Huxley Model



(b) Leaky Integrate and Fire Model



(c) Izhikevich Model



(d) Layer 5 Pyramidal Neuron Recording.

Figure 3.3: A comparison of spikes generated from various sources. Figure 7.2(a): The standard Hodgkin-Huxley model, a conductance based model consisting of the four ODEs (3.1)-(3.4). Figure 7.2(b): A leaky integrate-and-fire model. Note that the model does not spike and the spikes are drawn in. Figure 7.2(d): An Izhikevich model fit to a layer 5 pyramidal neuron. Figure 6.2(d): The layer 5 pyramidal neuron, at the same current. The model parameters were taken from [92].

generation mechanism, then one obtains the equation:

$$C\dot{V} = -\frac{V}{R} + I(t) \tag{3.5}$$

where $I(t)$ is an input current, and C and R are the capacitance and resistance of the circuit. Then $V(t)$ hits a specific threshold voltage value, V_T, the neuron is said to fire a spike. It is subsequently immediately reset to a reset potential, V_R:

$$V(t^-) = V_T \Rightarrow V(t^+) = V_R \tag{3.6}$$

and thus the result is a discontinuous oscillator, shown in Figure 7.2(b). This is referred to as a leaky integrate-and-fire (LIF) neuron [1, 25, 105]. From our previous discussion, the quantity that is thought of as a spike is the rapid upswing and downswing in the membrane potential that occurs on a much faster time scale then the rest of the oscillation. Thus, the LIF model is incapable of actual spike generation. This is due to the fact that without the designated voltage resets and threshold, the model simply tends towards the equilibrium $V = IR$ for a constant I. When $I \geq V_T/R$, then the neuron is said to fire, however the firing and spikes are essentially artificial. Additionally, the leaky integrate and fire neurons have a constant firing rate for fixed I. This is in general not true of actual neurons, which display a phenomenon called spike frequency adaption, where the firing rate of spikes often decreases with time for a fixed current I. Given these constraints, this model is too simple to model the majority of neurons.

However, a class of two-dimensional spiking neuron models has recently been introduced that can faithfully model the action potential upswing in real neurons. This general class of models differs from the typical linear/leaky integrate and fire neuron in the sense that the dynamics are nonlinear and thus capable of the upswing generation. This is due to the fact that the nonlinearity present in these models leads to a phenomenon referred to as "finite escape time", where the differential equations can diverge or blow up in finite time. To stop this divergence, a peak voltage is set at v_{peak} and the neurons are immediately reset at v_{reset}. The advantage that this has is that the peak voltage and reset voltage can be more easily measured from actual neuron then the threshold voltage, as in the LIF neuron. Additionally, the finite blow up allows one to approximate the rapid upstroke of biological neurons better then a single compartment conductance based model [152]. The models also display spike frequency adaptation through a recovery variable.

Models of this class include the Izhikevich neuron [89], the Adaptive Exponential (AdEx) neuron [23, 130], and the Quartic integrate and fire neuron [177]. The general class of models can be conveniently written in a dimensionless form:

$$\dot{v} = F(v) - w + I \tag{3.7}$$

$$\dot{w} = a(bv - w), \tag{3.8}$$

where v represents the nondimensionalized membrane potential, and w serves as an adaptation/recovery variable [177]. Time has also been non-dimensionalized. For the purpose

of the subsequent sections, we will define $\mathbf{x} = (v, w)$, and denote the dynamical system above in vector form as

$$\dot{\mathbf{x}} = \mathbf{G}(\mathbf{x}) = \begin{pmatrix} F(v) - w + I \\ a(bv - w) \end{pmatrix}. \tag{3.9}$$

The dynamical equations (96)-(97) are supplemented by the following discontinuities

$$v(t^-_{spike}) = v_{peak}, \rightarrow v(t^+_{spike}) = v_{reset} \tag{3.10}$$

$$w(t^+_{spike}) = w(t^-_{spike}) + w_{jump}. \tag{3.11}$$

This particular notation was formally introduced by Touboul (2008), along with a full bifurcation analysis of this general family of adapting integrate and fire neurons in [177]. The nonlinear $F(v)$ for the particular neuron models we will consider is given by

$$F(v) = v(v - \alpha) \quad \text{Izhikevich Model [89]} \tag{3.12}$$

$$F(v) = \exp(v) - v \quad \text{Adaptive Exponential Model [23, 130]} \tag{3.13}$$

$$F(v) = v^4 - 2av \quad \text{Quartic Model [177]} \tag{3.14}$$

When fit to actual neurons, these models can faithfully reproduce spike times and voltage traces from intracellular recordings. For example, in figure 7.2(d) an Izhikevich model was fit to the voltage trace of a layer 5 pyramidal neuron, shown in figure 6.2(d). While the class of 2-dimensional integrate and fire neurons might not be as accurate as conductance based models, they are significantly easier to fit as they have far less parameters. Additionally, they can phenomenologically reproduce many of the behaviors of actual neurons with only two differential equations per neuron, making them ideally suited for large network simulations. In section 3.3, we shall see how these models can be coupled together using models of synapses to form networks.

Before discussing integrate-and-fire models further, we need to note several modifications of the standard models. For both the leaky-integrate-and-fire model, and the AdEx, adaptive threshold models have also been suggested that fit neuronal data better then the models without adaptive thresholds [12, 98]. Furthermore, specialized forms of the Izhikevich model have also been suggested. Their are numerous modifications Izhikevich himself has suggested in [92]. Examples include modifying the dynamics so that they are piecewise smooth continuous so that the model better fits the action potential half-width [49, 92]. Additionally, when the adaptation variable is removed, and $\alpha = 0$, the resulting model is referred to as the quadratic integrate-and-fire model [24, 70, 106] and can be transformed to what is referred to as the theta-model or the Ermentrout-Koppell canonical model [55, 60, 75] by the change of variables $v = \tan(\theta/2)$, assuming that $v_{peak} = \infty$ and $v_{reset} = -\infty$. The theta-model is given by the equations

$$\dot{\theta} = 1 - \cos(\theta) + (1 + \cos(\theta))I \tag{3.15}$$

$$\theta(t^-) = \pi, \rightarrow \theta(t^+) = -\pi \tag{3.16}$$

where I is the applied current to the model. Finally, there is an additional model in the literature that is referred to as the perfect integrate-and-fire model [129] or alternatively

as the simple integrate-and-fire model [66]. For this model, $F(v) = 0$ and this results in some fairly unique behaviors (see [66]). As such, we will not consider it further.

An important advantage that integrate-and-fire models have is that they can be simulated much faster then conductance based models as they require fewer operations to numerically integrate [91]. Furthermore, switching to an event-based simulation strategy yields even greater speed increases and is feasible for certain nonlinear integrate and fire models as well [22, 154, 176]. Thus, these models are far more amenable for large network simulations, even at the scale of millions of neurons [50].

In addition to faster simulations, integrate-and-fire models can be fit substantially quicker then conductance based models. There are multiple techniques now for fitting integrate-and-fire neurons rapidly using either intracellular voltage recordings [12, 79, 80] or just the raster plots [98, 153]. In particular, The dynamic IV curve technique introduced in [12] also verifies that the dynamics of pyramidal neurons are well described by the AdEx model. The dynamic IV technique also only requires around 50 recorded spikes from an intracellular recording. This minimal duration of recording has allowed the authors in [80] to record and fit from large numbers of pyramidal neurons in multiple layers of the cortex to such a degree that the distributions of heterogeneity in all the parameters can be estimated. Simpler integrate-and-fire models with adaptation also appear to fit neuronal data more accurately then more complicated integrate-and-fire models [98].

However there are a few issues with integrate-and-fire neurons that one has to be aware of. As the systems are non-smooth due to the voltage reset, one has to be careful of what numerical integration schemes are used. For example, regardless of the scheme of integration used in between successive spikes, the order of error in the integration scheme will always be $O(\Delta t)$ sets $v(t + \Delta_t) = v_{reset}$ if $v(t + \Delta_t) > v_{peak}$ [162]. One can obtain higher order integration schemes if one uses a spike-time interpolant however [162].

Integrate and fire models are quickly becoming one of the dominant tools for modelling neurons and simulating neuronal networks due to their many advantages over other types of models. They have been fit to many neurons in many different brain areas such as hippocampal CA1 inhibitory and excitatory neurons [62, 89, 92], various kinds of cortical pyramidal neurons [12, 79, 80]. From a more phenomenological perspective, networks of integrate-and-fire neurons have been used to analyze various phenomena both general and specific. Some general examples include the use of integrate-and-fire models for emergent network properties like bursting [9, 49, 77, 104, 132, 134], problems in learning and memory [34, 35, 95, 156], the analysis of cortical gain control [3, 26, 33, 121], the affects of neuronal heterogeneity [76, 120], the affects of sparsity on coding and synchrony [21, 179]. Examples of more specific applications include the emergence of grid cells via oscillatory interference theory [139, 191] or studies on schizophrenia [10, 96].

Integrate and fire models have also branched out into tasks in data mining and machine-learning that would typically be performed using artificial neural networks. Indeed, the entire Neural Engineering Framework approach is based on the relationship between spiking neurons and specific artificial neural networks.

3.2.3 Artificial Neural Networks

The idea behind artificial neural networks (ANNs) is to ignore the actual spiking in neurons and instead encode the average rate of spiking for either the individual neurons, or a population of neurons. These average rates are modeled as sigmoid functions. Examples include the following:

$$f(x) = \tanh(x) \quad \text{(Hyperbolic Tangent Function)} \tag{3.17}$$

$$f(x) = \frac{1}{1+\exp(-x)} \quad \text{(Logistic Function)} \tag{3.18}$$

$$f(x) = \int_{-\infty}^{x} \exp(-x'^2/2)\, dx', \quad \text{(Error Function)} \tag{3.19}$$

These functions, and other typically used functions are differentiable to all orders and bounded. The boundedness can be thought of as a physical criterion as neurons themselves have bounded firing rates with regards to their inputs. The sigmoids are more formally justified by the work of Wilson and Cowan, [190] and can be thought of as population response variables. This literature on artificial neural networks is vast, and we make no attempt at summarizing it here. The majority of this thesis will not consider ANNs aside from their relationship to the Neural Engineering Framework where one replaces the sigmoids above with firing rate functions for spiking neuron models. More modern artificial neural network approaches also use rectified linear neurons, with the firing rate given by

$$f(x) = xH(x) \quad \text{(Rectified Linear Unit)} \tag{3.20}$$

which have unbounded firing rates [71]. The firing rate for theta neurons is also unbounded and has been considered in the context of artificial neural networks [119].

3.3 Modeling Synapses

Networks of neurons are coupled together through changes in the synaptic current. We will consider two methods for modeling synapses: kinetic synapses and pulse coupling. Both of these methods however share one thing. They both attempt to model the proportion of open ion channels that connect the presynaptic neuron i to the postsynaptic neuron j, which we will refer to as s_{ij}, for $i, j = 1, 2, \ldots N$. Assuming that the ion channels all have a uniform synaptic conductance. The variable s_{ij} is assumed to be in the interval $[0, 1]$, being the proportion of ion channels open. There are two primary modeling approaches to synapses: pulse coupling and kinetic coupling. These will be considered separately, and we will finally state the full network equations.

3.3.1 Kinetic Synapses

Like conductance based models, kinetic models of the synapse are derived from the underlying biophysics of ion channel opening. In particular, the channels are represented with

a number of open and closed states, and they stochastically transition from these states, similar to chemical reaction kinetics. For example, a simple synapse with just open and closed schemes can take the form

$$\mathrm{O} \quad \underset{k_{-1}(V)}{\overset{k_1(V)}{\rightleftharpoons}} \quad \mathrm{C} \tag{3.21}$$

where the transition rates between open and closed states are voltage dependent. These can also be substantially more complicated and include the affects of any secondary messengers such as calcium or magnesium blocks for NMDA channels. To be fully accurate, one has model each primary component from the arrival of the action potential at the presynaptic terminal to the opening (or closing) of the ion channels in the post-synaptic membrane. Kinetic models are set up for each protein involved in the transmission process and these are transformed into differential equations, on the assumption that the number of proteins is suitably large. This would involve many differential equations finally feeding into a differential equation for $s_{ij}(t)$. To avoid the complication associated with such a large number of differential equations, approximations to these equations were derived in [43, 44, 45]. These approximations were based on a separation of time scales. The primary result was that the system of equations that linked the presynaptic voltage to the postsynaptic gating variable was drastically reduced. For example, $s_{ij}(t)$ is now given by a kinetic gating equation:

$$s'_{ij}(t) = (1 - s_{ij}(t))\alpha(V)[T] - \beta(V)s_{ij}(t) \tag{3.22}$$

where $\alpha(V)$ and $\beta(V)$ are voltage dependent transition probabilities (although they need not actually depend on the voltage), and $[T]$ is the quantity of neurotransmitter released into the synaptic cleft as a result of presynaptic firing. This particular differential equation corresponds to the kinetic scheme (3.21) with a single open and closed state, with $k_1(V) = \alpha(V)[T]$ and $k_{-1}(V) = \beta(V)$ and is again due to [45]. It is valid for AMPA receptors/channels, an excitatory synaptic receptor. A table of synaptic channels, and their simplifications due to [45] is included in Table 3.1

There are two forms for $[T]$ suggested in [45]. The first, which we will refer to as kinetic synapse I directly links the presynaptic voltage V_j to the amount of neurotransmitter released:

$$[T](V_j) \quad = \quad \frac{T_{max}}{1 + \exp(-(V_j - V_p)/Kp)} \quad \text{Kinetic Type I} \tag{3.23}$$

where T_{max} is the maximum amount of transmitter released. The second, which we will refer to as kinetic synapse II, is given by;

$$[T](V_j) = \begin{cases} T_{max} & t_{j,k} < t < t_{j,k} + \bar{t} \\ 0 & \text{otherwise} \end{cases} \quad \text{Kinetic Type II} \tag{3.24}$$

In this model, a rectangular pulse of transmitter is released when presynaptic neuron j fires a spike. We will primarily use pulse coupling, which we will cover in the next section. It is possible to asymptotically approximate kinetic type-I channels as a kind of pulse coupling, as we will later show.

Receptor	Neurotransmitter	Kinetic Equations
AMPA Receptor	Glutamate	$I_{AMPA} = \bar{g}_{AMPA} s_{ij}(V_i - E_{AMPA})$ $s'_{ij} = (1 - s_{ij})\alpha(V)[T] - \beta(V)s_{ij}$
NMDA Receptor	Glutamate	$I_{NMDA} = \bar{g}_{NMDA} B(V_i) s_{ij}(V_i - E_{NMDA})$ $s'_{ij} = (1 - s_{ij})\alpha(V)[T] - \beta(V)s_{ij}$ $B(V) = \frac{1}{1+\exp(-0.062V)[Mg^{2+}]/3.57}$
GABA$_A$ Receptor	GABA	$\bar{g}_{GABA_A} s_{ij}(V_i - E_{GABA_A})$ $s'_{ij} = (1 - s_{ij})\alpha(V)[T] - \beta(V)s_{ij}$
GABA$_B$ Receptor	GABA	$I_{GABA_B} = \bar{g}_{GABA_B} \frac{s_{ij}^n}{s^n + K_d}(V - E_K)$ $r'_{ij} = K_1[T](1 - r_{ij}) - K_2 r_{ij}$ $s_{ij} = K_3 r_{ij} - K_4 s_{ij}$

Table 3.1: Table of simplified synaptic ion channel equations from [45]. Nominal parameter values can be found in [43, 44, 45]

3.3.2 Pulse Coupling

In the case of pulse coupling, the presynaptic firing of neuron j is assumed to cause a transient change in the value of $s_j(t)$, in particular after neuron j fires a spike, a pulse function $E(t)$ is added to $s_j(t)$:

$$s_j(t) = \sum_{t_{j,k}<t} E(t - t_{j,k}). \tag{3.25}$$

where $t_{j,k}$ is the kth spike fired by neuron j. There are different functions proposed for $E(t)$ in the literature. Examples include the single exponential synapse:

$$E(t) = s_{jump} \exp\left(\frac{-t}{\tau_s}\right), \tag{3.26}$$

the double exponential synapse:

$$E(t) = \frac{1}{\tau_R \tau_D}\left(\exp\left(\frac{-t}{\tau_R}\right) - \exp\left(\frac{-t}{\tau_D}\right)\right), \tag{3.27}$$

where τ_R is the rise time, and τ_D is the decay time, and the alpha synapse

$$E(t) = \alpha^2 t \exp(-\alpha t). \tag{3.28}$$

where $1/\alpha$ is both the rise and decay times [61]. These motivation for pulse-coupling comes from an examination of post-synaptic conductance changes at single synapses which often have these exponential like pulses after presynaptic firing. Each synaptic function $s_j(t)$ under pulse coupling can be described by a simple differential equation. For example, If we consider the simple exponential function, then

$$s'_j = -\frac{s_j}{\tau_s} + s_{jump} \sum_{t_{j,k}<t} \delta(t - t_{j,k})$$

where $t_{j,k}$ is the k as integration immediately yields

$$\begin{aligned} s_j(t) &= s_{jump} \sum_{t_{j,k}<t} \exp\left(-\frac{t-t_{j,k}}{\tau_s}\right) + C\exp(-t/\tau_s) \\ &= \sum_{t_{j,k}<t} E(t-t_{j,k}) + C\exp(-t/\tau_s) \end{aligned}$$

where the final term can be ignored if we set $s(0) = C = 0$, or in general as it decays exponentially and is unaffected by spiking activity.

For pulse coupled networks, the synaptic coupling function $s(t)$, in the case of all-to-all connectivity (and indeed other cases), can be formally described by a linear system of ordinary differential equations with Dirac delta functions on the right hand side every time a neuron in the network fires a spike [61]. This is typically done after rescaling $s(t)$ to absorb the $\frac{1}{N}$ term. For example, the simple exponential synapse is governed by the ordinary differential equation

$$\frac{ds(t)}{dt} = -\frac{s}{\tau_s} + \frac{s_{jump}}{N}\sum_{j=1}^{N}\sum_{t_{j,k}<t}\delta(t-t_{j,k}). \tag{3.29}$$

The alpha function synapse and the double exponential synapse have a system of two first order ODEs that describe their dynamics and can be derived in the same manner as the single exponential synapse. For reference purposes, these systems are:

$$\frac{ds}{dt} = -\frac{s(t)}{\tau_R} + h \tag{3.30}$$

$$\frac{dh}{dt} = -\frac{h}{\tau_D} + \frac{1}{N\tau_R\tau_D}\sum_{j=1}^{N}\sum_{t_{j,k}<t}\delta(t-t_{j,k}) \tag{3.31}$$

for the double exponential synapse and

$$\frac{ds}{dt} = -\alpha s(t) + h \tag{3.32}$$

$$\frac{dh}{dt} = -\alpha h + \frac{\alpha^2}{N}\sum_{j=1}^{N}\sum_{t_{j,k}<t}\delta(t-t_{j,k}) \tag{3.33}$$

for the alpha synapse. One should note that all these ordinary differential equations are linear, and they all contain the term

$$j(t) = \frac{1}{N}\sum_{j=1}^{N}\sum_{t_{j,k}<t}\delta(t-t_{j,k}) \tag{3.34}$$

which is very important in the large network dynamics and represents a kind of network averaged firing rate as we will show later.

One advantage in pulse-coupling is the simplicity in fitting to actual neurons. As the kinetics of the pulse do not depend on the voltage, one can immediately measure the postsynaptic current or conductance and fit a simple pulse to it without the need of determining the voltage dependent terms. Unfortunately however pulse coupling has other disadvantages. If we interpret $s_j(t)$ as a proportion, then it is clear that $s(t)$ is also a proportion and should be inside the interval $[0, 1]$, however for sufficiently fast presynaptic firing, $s(t)$ exceeds one for pulse coupling. Additionally, it is not based on the underlying biophysics of synapses and thus is less biologically plausible.

The time course of $s_j(t)$ for a single spike is shown for all the different synaptic models considered in Figure 3.4.

3.3.3 Network Equations

The total synaptic conductance of a post-synaptic neuron i due to the presynaptic neurons $j = 1, 2 \dots N$ is then given by

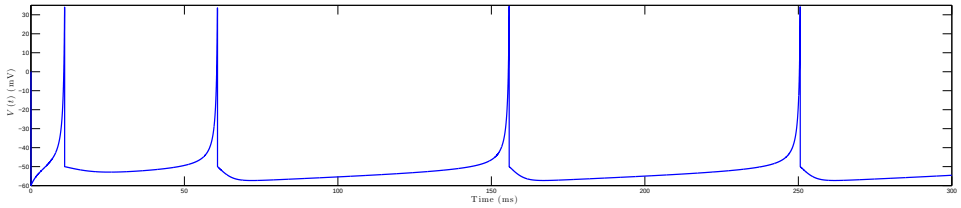
$$g_i(t) = \sum_{j=1}^{N} g_{ij} s_{ij}(t) \tag{3.35}$$

where g_{ij} denotes the maximal synaptic conductance at the synapse connection neuron i to neuron j. We will assume that g_{ij} is a Bernoulli random variable given by

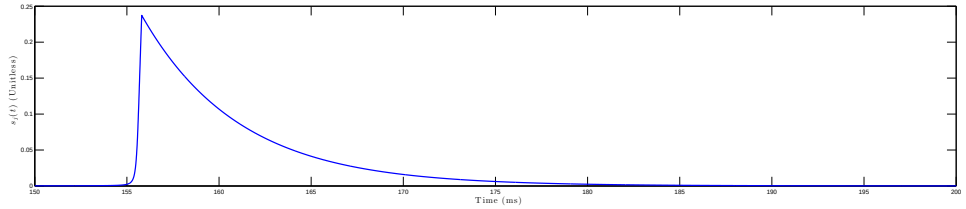
$$g_{ij} = \begin{cases} \bar{g}_i & \text{with probability } p \\ 0 & \text{with probability } 1 - p \end{cases} \tag{3.36}$$

when we have $p = 1$, the network is referred to as an all-to-all coupled network. When $p \ll 1$, then the network is sparsely coupled. We will look at sparsely coupled networks in greater detail in Chapter 4. For now, we will restrict our selves to $\bar{g}_i = \frac{g}{N}$, and $p = 1$. In the sections on heterogeneity and sparsity, we will demonstrate how to proceed without making these assumptions. Additionally, one can make the much less strict assumption that $s_{ij}(t) = s_j(t)$. In this case, neuron j activates the same proportion of ion channels at each synapse it makes. The reason we can make this assumption is that it relies on the transient effects of synapses not mattering after an extended period of time, and the fact that the same number of spikes fired by neuron j arrive at each connection it makes. These assumptions vastly simplify the structure of the network, allowing us to proceed more easily in the mean-field system derivation. However, it is important to state that they are not necessary to make, and they are also not biologically implausible. While actual neurons also display synaptic failure, where a presynaptic spike fails to elicit a postsynaptic response, we leave the incorporation of synaptic failure for future work. The resulting equation for the synaptic conductance to neuron i becomes:

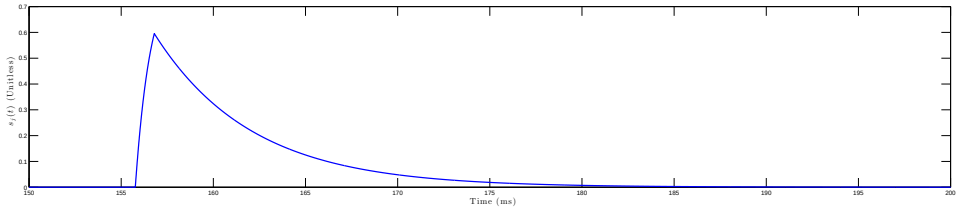
$$\bar{g}_i(t) = \sum_{i=1}^{N} \bar{g}_i s_{ij}(t) = \frac{g}{N} \sum_{j=1}^{N} s_j(t) = g s(t) = g(t) \tag{3.37}$$



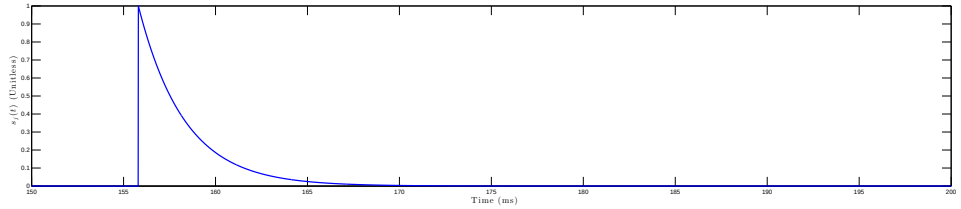
(a) Izhikevich Model Neuron



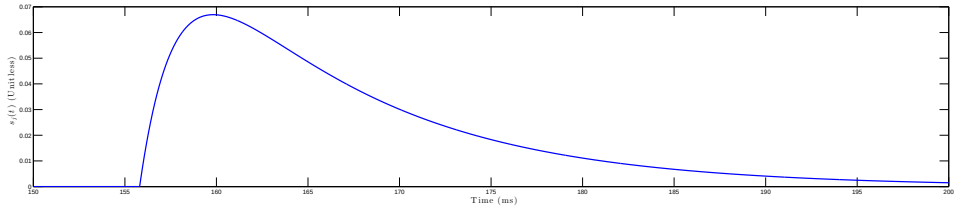
(b) Kinetic Type I



(c) Kinetic Type II



(d) Single Exponential



(e) Double Exponential

Figure 3.4: A comparison of the pulse waveforms for $s_j(t)$ generated via Kinetic Type I (3.4(b)), Kinetic Type 2 (3.4(c)), single exponential (3.4(d)) and double exponential coupling (3.4(e)) from a single Izhikevich neuron (3.4(a)). The waveforms are shown in the 150-200 ms range.

where $s(t)$ is the network average average proportion of synaptic channels open.

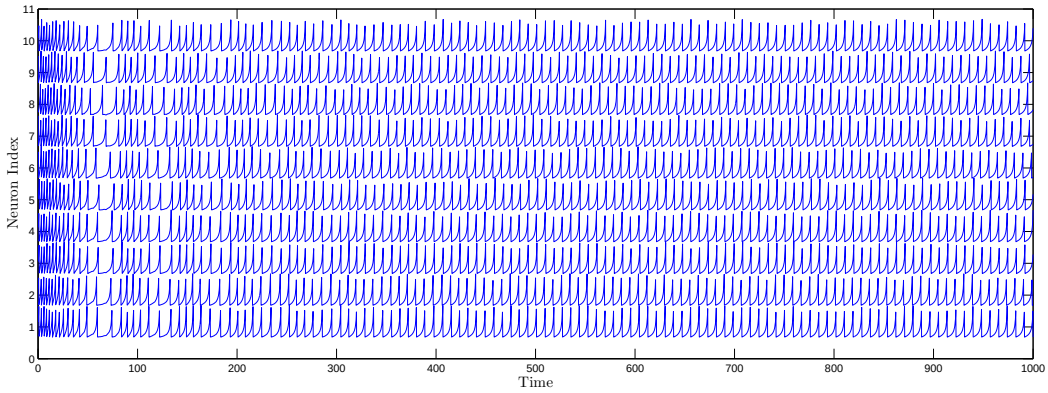
In summary, the approximation of a network of all-to-all coupled neurons that we will consider is given by system of discontinuous ODE's

$$\begin{aligned}
\dot{v}_i &= F(v_i) - w + I + gs(t)(E_r - v_i) && (3.38)\\
\dot{w}_i &= a(bv_i - w_i) && (3.39)\\
v_i(t^-_{spike}) &= v_{peak}, \to v_i(t^+_{spike}) = v_{reset} && (3.40)\\
w_i(t^+_{spike}) &= w_i(t^-_{spike}) + w_{jump} && (3.41)\\
s(t) &= \frac{1}{N}\sum_{j=1}^{N} s_j(t) && (3.42)
\end{aligned}$$

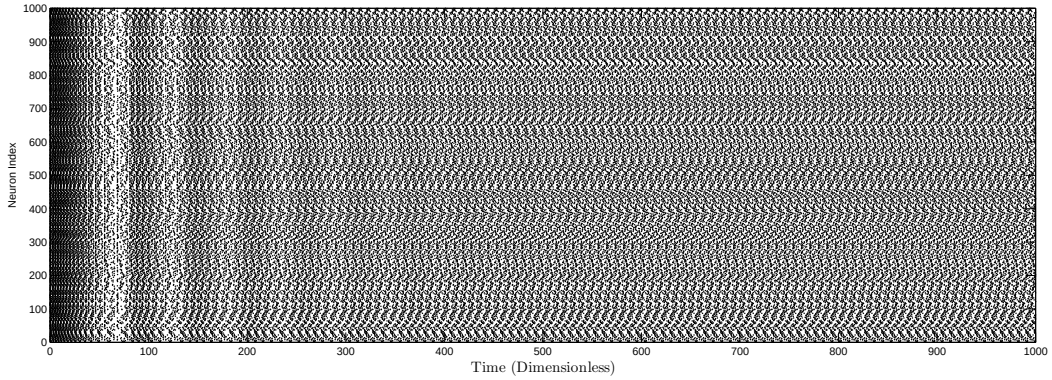
for $i = 1, 2, \ldots N$, and where the dynamics of $s(t)$ depend on the specific synapse implemented.

Figure 3.5 shows a simulation of a network of 1000 neurons with the voltage trace of 10 randomly chosen neurons shown in Figure 7.4(a), and the entire raster plot of spike times in Figure 7.4(b). The network settles on a tonic firing steady state, where the neurons all fire spikes with a constant firing rate asynchronously across the network. If we plot the $s(t)$ and the mean of $w_i(t)$, $\langle w \rangle$, these two variables appear to settle on a steady state equilibrium, as shown in Figure 7.4(c). If we adjust some of the parameters we and simulate the network again we can observe different behaviors. In particular, in 3.6, the network appears to transition to bursting, a oscillatory steady state where the neurons alternate between firing a volley of spikes and a quiescent stage periodically. Again, if we plot $s(t)$ and $\langle w \rangle$, it appears that the network bursting results in a limit cycle for these variables.

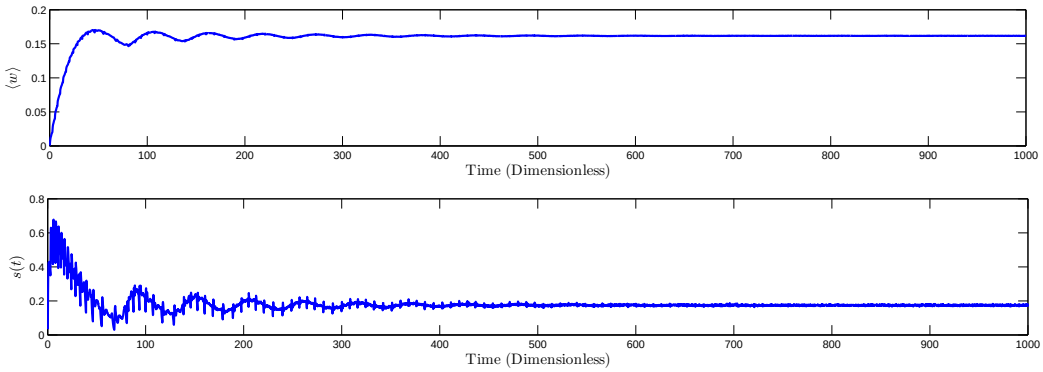
More importantly however is the fact that $\langle w \rangle$ and $s(t)$ seem be to enough to predict the behavior of the network. That is, there appears to be a low dimensional (in this case two-dimensional) dynamical system that can predict the behavior of the entire network. We will later see that as $N \to \infty$, the equations for the dynamics of $s(t)$ and $\langle w \rangle$ can be determined. The resulting system is a closed system of ordinary differential equations.



(a) Voltage trace for 10 Neurons

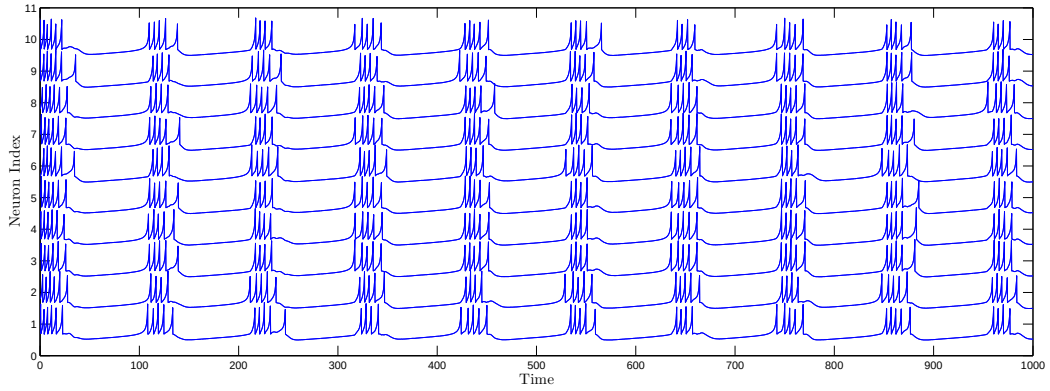


(b) Raster plot for the entire network

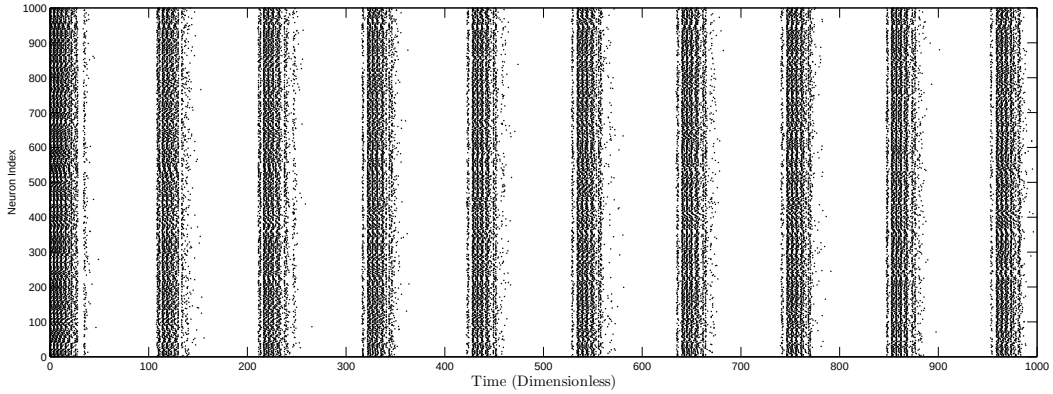


(c) $s(t)$ and $\langle w \rangle$, the mean adaptation variable.

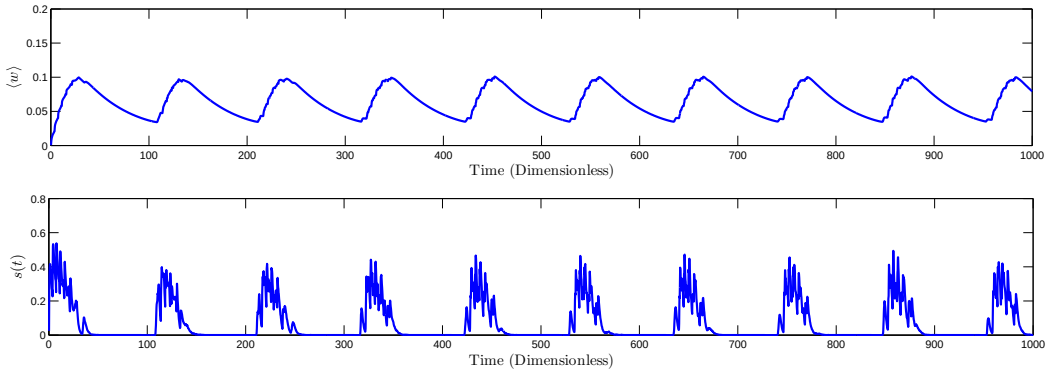
Figure 3.5: A network of 1000 Izhikevich neurons is simulated with single-exponential synaptic coupling. The neurons settle on a tonic firing asynchronous steady state. In this state, the neurons all reach an identical steady state firing rate, and fire asynchronously across the network. This manifests itself as a steady state equilibrium for $s(t)$ and $\langle w \rangle$, the mean-adaptation across the network, as shown in figure 7.4(c).



(a) Voltage trace for 10 Neurons



(b) Raster plot for the entire network



(c) $s(t)$ and $\langle w \rangle$, the mean adaptation variable.

Figure 3.6: A network of 1000 Izhikevich neurons is simulated with single-exponential synaptic coupling. The neurons settle on a synchronous bursting steady state. In this state, the neurons fire synchronous bursts of action potentials across the network. This manifests itself as a steady state limit cycle for $s(t)$ and $\langle w \rangle$, the mean-adaptation across the network, as shown in figure 7.4(c).

3.4 Network Model Reduction

3.4.1 Derivation of the Population Density Equation

One way to analyze large systems with quantities that are conserved is the continuity equation. This equation applies to neural networks since the total number of neurons in a network is a conserved quantity. In order to apply it, we need to define a density function, ρ. Consider a region Ω in phase space at time t. Using the notation introduced in the previous section, the vector in phase space is $\mathbf{x} = (v, w)$. Thus, we are not considering the synaptic components as part of phase space. This is due to the fact that only the average $s_j(t)$ appears in the equation for each neuron. Population density methods involve analyzing the proportion of neurons that lie in Ω at time t:

$$P_N(\Omega, t) = \frac{1}{N} \sum_{i=1}^{N} \chi_\Omega(\mathbf{x}_i), \tag{3.43}$$

where χ_Ω is the conventional indicator function.

Assuming that the boundaries of Ω are piecewise smooth, then the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \chi_\Omega(\mathbf{x}_i) = P(\Omega, t) \tag{3.44}$$

always exists for a fixed time and arbitrary region of phase space Ω. The reason for this is that the above is merely the expectation of the indicator function

$$P(\Omega, t) = E(\chi_\Omega(\mathbf{x})). \tag{3.45}$$

Thus the neurons have a well defined probability distribution, with a density function $\rho(\mathbf{x}, t)$ defined by

$$P(\Omega, t) = \int_{\mathbf{X}} \chi_\Omega(\mathbf{x}) \rho(\mathbf{x}, t)\, d\mathbf{x} = \int_\Omega \rho(\mathbf{x}, t)\, d\mathbf{x}. \tag{3.46}$$

The time evolution of this probability distribution is simply a conservation equation

$$\frac{d}{dt} P(\Omega, t) = \int_\Omega \frac{d}{dt} \rho(\mathbf{x}, t)\, d\mathbf{x} = \text{ flux into } \Omega \text{ - flux out of } \Omega. \tag{3.47}$$

Note that at each point in the phase space, there is an associated vector field determined by the underlying differential equations of the network. With this in mind, the flux is caused by the bulk flow of neurons along the vector field. If $\mathbf{n}$ is the outward normal to the boundary of Ω, $\partial\Omega$, then the proportion of neurons leaving Ω at a particular point at time t, at a point $\mathbf{x}$ on $\partial\Omega$ is

$$\begin{pmatrix} F(v) - w + I \\ a(bv - w) \end{pmatrix} \cdot \mathbf{n} \rho(\mathbf{x}, t) = \mathbf{G}(\mathbf{x}) \cdot \mathbf{n} \rho(\mathbf{x}, t) \tag{3.48}$$

Note that this is without coupling, which we will add in later on. Thus we have the flux into/out of Ω in its integral form:

$$\mathbb{J}_{drift} = -\int_{\partial\Omega} \mathbf{G}(\mathbf{x}) \cdot \mathbf{n}\rho(\mathbf{x}, t)\, dS. \tag{3.49}$$

Using the divergence theorem, assuming of course that Ω has a piecewise smooth boundary, then we have

$$\mathbb{J}_{drift} = -\int_{\partial\Omega} \mathbf{G}(\mathbf{x}) \cdot \mathbf{n}\rho(\mathbf{x}, t)\, dS = -\int_{\Omega} \nabla \cdot (\mathbf{G}(\mathbf{x})\rho(\mathbf{x}, t))\, d\mathbf{x}. \tag{3.50}$$

Defining $J(\mathbf{x}, t) = \mathbf{G}(\mathbf{x})\rho(\mathbf{x}, t)$, then by the Dubois-Reymond Lemma [93], we have the following

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot J(\mathbf{x}, t). \tag{3.51}$$

Note that (3.51) is the "vanilla" form of the equation, without any specifics related to the connectivity of the neurons or their discontinuities. In this particular case, the flux is a vector of two components, $\mathbf{J}(\mathbf{x}, t) = (J_V(v, w, t), J_W(v, w, t))$. Additionally, if our neurons have resets, then we have boundary conditions defined in terms of the flux. The boundary condition has to incorporate both the discontinuity in the adaptation current and the reset in the membrane potential yielding:

$$J_V(v_{peak}, w, t) = \lim_{v \to v_{reset}^+} J_V(v_{reset}, w + w_{jump}, t) - \lim_{v \to v_{reset}^-} J_V(v_{reset}, w + w_{jump}, t). \tag{3.52}$$

This boundary condition only applies when we consider $v \in [-\infty, v_{peak}]$. Note that the discontinuity in the flux at $v = v_{reset}$ implies that the density function is non-smooth. However, if we restrict ourselves to the interval $[v_{reset}, v_{peak}]$, then the boundary condition becomes

$$J_V(v_{peak}, w, t) = J_V(v_{reset}, w + w_{jump}, t). \tag{3.53}$$

With this boundary condition, we have removed a source of non-smoothness in the population density equation (which may still have a non-smooth solution, depending on the initial conditions).

When we couple the network together, the v component of the flux becomes s dependent, and is given by

$$J_V(v, w, s, t) = (F(v) - w + I + gs(e_r - v))\rho(v, w, t). \tag{3.54}$$

The flux is particularly interesting as it is intuitively the mass (proportional) flow rate along a specific direction in phase space. In the large network limit, the flux ends up being closely related to $j(t)$ in equation (3.34).

Aside from being called the continuity equation, (3.51) is also referred to as the Liouville equation. In the context of Neuroscience, there are similar derivations to that shown above contained in various sources [137, 138]. The continuity equation has been used

to analyze the stability of the asynchronous state of a network of one dimensional non-adapting coupled neurons (predominantly integrate and fire) by various authors [2, 169, 174, 181, 183]. Another use for the continuity equation is to eliminate the need for direct simulation of large networks, instead numerically solving the continuity equation [8, 113, 137, 138]. More analytical treatments of the partial differential equation (3.51), including resolving its spectrum of eigenvalues, solution to the steady state, and general features as a linear operator can be found in [2, 97, 168, 169, 174, 181]. Other applications involve using population density equations to analyze phase models [111, 112] or the vast literature on Kuramoto oscillators (see [173] for a review). More specific applications include an analysis of a pair of coupled population density equations representing the population of neurons in the subthalamic-Nucleus and the external segment of the globus pallidus, two structures that are part of the basal ganglia [124, 125, 126].

3.4.2 Relationship Between the Flux and the Firing Rate

In a finite network with pulse coupling, the ODE governing the synaptic coupling (see section 1.2.2) often contains the term

$$j(t) = \frac{1}{N}\sum_{j=1}^{N}\sum_{t_{j,k}<t}\delta(t - t_{j,k}). \tag{3.55}$$

We will show later that approximations can be made for the kinetic synapse models that also involve $j(t)$. We will show that in the limit that $N \to \infty$,this quantity converges to the network averaged instantaneous firing rate, which we will denote by $\langle R_i(t)\rangle$, where $\langle\,\rangle$ corresponds to averaging across the network. We will think of $j(t)$ and its limit as a distribution.

One should be very careful in defining the appropriate "mean firing rate" as there are at least three definitions of a kind of averaged firing rate present in the literature. For an excellent source on rate codes, see [70].

First, define the function $n_i(t)$ to be the number of spikes fired by the ith neuron in the time interval $[0, t]$. One can relate $j(t)$ to the average of $n_i(t)$ across the network. This is given by

$$\int_0^t j(x)\,dx = \frac{1}{N}\int_0^t \sum_{j=1}^{N}\sum_{t_{j,k}<t}\delta(x - t_{j,k})\,dx = \langle n_i(t)\rangle. \tag{3.56}$$

Then we define $\langle R_i(t)\rangle$ as

$$\langle R_i(t)\rangle = \lim_{\Delta t\to 0}\frac{1}{\Delta t}\lim_{N\to\infty}\sum_{i=1}^{N}\frac{n_i(t+\Delta t) - n_i(t)}{N}. \tag{3.57}$$

Thus, the network averaged firing rate in this sense is the limit of the population activity as $\Delta t \to 0$ [70]. However, one should note that we also obtain the equation

$$\langle R_i(t)\rangle = \lim_{\Delta t\to\infty}\frac{\langle n_i(t+\Delta t)\rangle - \langle n_i(t)\rangle}{\Delta t} = \frac{d}{dt}\langle n_i(t)\rangle, \tag{3.58}$$

after rearranging the limits. Thus, we can replace the differential equation for $s(t)$ with a simpler differential equation where we replace $j(t)$ with $\langle R_i(t)\rangle$.

Now, all that remains is to relate $\langle R_i(t)\rangle$ to the flux. As the flux is a kind of directional flow rate of the proportion of neurons at a particular point in phase space, the firing rate can be computed by integrating the flux vectors across the peak-voltage surface, $v = v_{peak}$. For example, the flux for a system of one dimensional neurons, $J(v,t)$ is merely the proportion of neurons that flow across the point v in phase space per unit time at time t. Thus, for a network of linear integrate and fire neurons for example, the firing rate is $J(v_t, s, t)$, the flux through the threshold. However, for the class of neurons we are dealing with, there is a specific peak voltage, and the resulting firing rate is actually $\int_W J_W(v_{peak}, w, s, t)\, dw$ as we have to integrate the v component of the flux on the line $v = v_{peak}$ over the other state variable, w, in phase space. Note that only the first component of the flux contributes to the firing rate, as the second component, $J_W(v, w, t)$ is in a direction parallel to the firing rate boundary, $v = v_{peak}$. Thus, for the two-dimensional models in consideration, we have

$$\langle R_i(t)\rangle = \int_W J_V(v_{peak}, w, s, t)dw. \tag{3.59}$$

Using equations (3.51), (3.58), and (3.59),and the results of section 3.4, our model is now a PDE coupled to an ODE or system of ODE's. In the case of the simple exponential synapse, we have

$$\begin{aligned} \frac{\partial \rho(v,w,t)}{\partial t} &= -\frac{\partial}{\partial v}\left((F(v) - w + I + gs(E_r - v))\rho(v,w,t)\right) \\ &- \frac{\partial}{\partial w}\left(a(bv - w))\rho(v,w,t)\right) \qquad (3.60) \\ \dot{s} &= -\frac{s}{\tau_s} + s_{jump}\int_W J_V(v_{peak}, w, s, t)dw. \qquad (3.61) \end{aligned}$$

The ordinary differential equations for the double exponential and alpha synapse can also be derived in a similar matter. As a shortcut, merely replace $j(t)$ with $\langle R_i(t)\rangle$. Equations (3.60)-(3.61) are not analytically solvable in the vast majority of cases. To our knowledge, an analytical solution only exists for the case where there is no coupling $g = 0$ and no adaptation variable w_i. There are, however, various perturbation approximations that one can use to study the solutions.

Numerical solution of equations (3.60)-(3.61) are also possible. However, to the best of our knowledge only a first order method has been successfully implemented in the literature for the full system of equations [125].

There is a simpler alternative to numerical methods that can be implemented to determine the bifurcation types this system displays, and yield the approximate bifurcation curves. This is the mean-field approach which involves reducing the system consisting of (3.60)-(3.61) to a small system of differential equations for $\langle w\rangle$ and $s(t)$. The mean-field approach is a sequence of analytical steps and approximations that one performs to yield this reduction and can easily be extended to far more elaborate cases than the network

we are currently working with. In particular, we will later apply the mean-field approach to networks with heterogeneous neurons, white noise, sparse coupling, and multiple subpopulations. In each case, the final set of equations will govern them mean adaptation variable(s) and the mean synaptic gating variable(s).

Chapter 4

Derivation of the Mean-Field System for a Homogeneous Network of All-to-All Coupled Spiking Neurons

In this chapter, we will proceed in deriving the mean-field system for a network of homogeneous, all-to-all coupled spiking neurons using a population density equation as an intermediate, as introduced in the previous chapter. To arrive at the final mean-field system will require the application of a set of approximations. At each step in the reduction process, we will ascertain the effect of the approximations on the overall error for the mean-field system, as compared to simulated networks of neurons. The sequence of approximations will be a first order moment closure approximation (section 4.1), a small w_{jump} expansion, and a separation of time scales (section 4.2.2). The approximation that dominates the error is a separation of time scales, as will be shown by numerical simulations. The final mean-field system is actually the $O(1)$ solution to the slow system. The $O(1)$ solution to the fast system is also resolvable. Finally, we will show that the stability properties of the asynchronous solution(s) to the moment-closure reduced population density equation are determined to order $O(\tau_w^{-1})$ by the stability of the corresponding steady state(s) of the mean-field system. This will be done by resolving the spectral equation for the moment-closure reduced PDE in section 4.4.

4.1 Moment Closure Reductions

4.1.1 First Order Moment Closure

In this section we will show how the moment closure assumption can be used to replace the PDE (3.60) by one for the marginal density, $\rho_V(v,t)$, and an ODE governing the mean

of w. To begin, we write the density in its conditional form

$$\rho(v, w, t) = \rho_W(w|v, t)\rho_V(v, t), \tag{4.1}$$

and substitute into (3.60) to obtain

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\rho_V(v,t)\rho_W(w|v,t)\right) &= -\frac{\partial}{\partial v}\left((F(v) + I - w + g(e_r - v)s)\rho_V(v,t)\rho_W(w|v,t)\right) \\
&- \frac{\partial}{\partial w}\left(a(bv - w)\rho_V(v,t)\rho_W(w|v,t)\right) \\
&= -\frac{\partial}{\partial v}\left(G_1(v,w,s)\rho_V(v,t)\rho_W(w|v,t)\right) \\
&- \frac{\partial}{\partial w}\left(G_2(v,w)\rho_V(v,t)\rho_W(w|v,t)\right). \qquad (4.2)
\end{aligned}$$

Next, we integrate with respect to w, yielding the first order conditional moments $\langle w|v\rangle$ which can be subsequently reduced with the first order moment closure assumption $\langle w|v\rangle = \langle w\rangle$. Integrating (4.2) with respect to w over the entire phase space W and using the normalization condition on the conditional density of w and the moment closure assumption yields

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\rho_V(v,t)\right) &= -\frac{\partial}{\partial v}\left((F(v) - \langle w|v\rangle + I + g(e_r - v)s)\rho_V(v,t)\right) \\
&- \left(a(bv - \langle w|v\rangle)\rho_V(v,t)\rho_W(w|v,t)\right)\Big|_{\partial W} \qquad (4.3) \\
&= -\frac{\partial}{\partial v}\left((F(v) - \langle w|v\rangle + I + g(e_r - v)s)\rho_V(v,t)\right) \qquad (4.4) \\
&\approx -\frac{\partial}{\partial v}\left((F(v) - \langle w\rangle + I + g(e_r - v)s)\rho_V(v,t)\right), \qquad (4.5)
\end{aligned}$$

where the second term in (4.3) vanishes because $\rho_W(w|v, t)$ is zero on the boundary, ∂W. We will redefine the flux as

$$J(v, \langle w\rangle, s, t) = (F(v) - \langle w\rangle + I + gs(e_r - v))\rho_V(v, t) = G_1(v, \langle w\rangle, s)\rho_V(v, t). \tag{4.6}$$

Integrating both sides of the boundary condition (4.9) yields:

$$\begin{aligned}
\int_W J_V(v_{peak}, w, s, t)\, dw &= \int_W \rho_W(w|v_{peak}, t)\rho_V(v_{peak}, t)G_1(v, w, s)\, dw \\
&= \rho_V(v_{peak}, t)G_1(v, \langle w|v_{peak}\rangle, s) \\
&= J(v_{peak}, \langle w|v_{peak}\rangle, s, t) \qquad (4.7) \\
\int_W J_V(v_{reset}, w + w_{jump}, s, t)\, dw &= \int_W \rho_W(w + w_{jump}|v_{reset}, t)\rho_V(v_{reset}, t)G_1(v, w, s)\, dw \\
&= J(v_{reset}, \langle w|v_{reset}\rangle, s, t). \qquad (4.8)
\end{aligned}$$

Subsequently applying first order moment closure yields the new boundary condition

$$J(v_{peak}, \langle w\rangle, s, t) = J(v_{reset}, \langle w\rangle, s, t). \tag{4.9}$$

We now derive an ordinary differential equation for $\langle w \rangle$. Recall the differential equations for v_i and w_i are given by

$$\begin{aligned} \dot{v}_i &= G_1(v_i, w_i, s) \\ \dot{w}_i &= G_2(v_i, w_i). \end{aligned}$$

The corresponding continuity equation is

$$\frac{\partial \rho(v,w,t)}{\partial t} = -\frac{\partial}{\partial v}\left(\rho(v,w,t)G_1(v,w,s)\right) - \frac{\partial}{\partial w}\left(\rho(v,w,t)G_2(v,w)\right),$$

and the mean adaptation variable is given by

$$\langle w \rangle = \int_V \int_W w\rho(v,w,t)\, dwdv.$$

Differentiation with respect to time and application of the continuity equation yields

$$\langle w \rangle' = -\int_V \int_W w\frac{\partial}{\partial v}\left(\rho(v,w,t)G_1(v,w,s)\right)\, dwdv - \int_V \int_W w\frac{\partial}{\partial w}\left(\rho(v,w,t)G_2(v,w)\right)\, dwdv.$$

Applying integration by parts and interchanging the order of integration as needed then gives:

$$\begin{aligned} \langle w \rangle' &= -\int_W w\Big[\rho(v,w,t)G_1(v,w,s)\Big]\Big|_{\partial V} dw \\ &- \int_V \left[\Big[wG_2(v,w)\rho(v,w)\Big]\Big|_{\partial W} - \int_W G_2(v,w)\rho(v,w)\, dw\right] dv \\ &= \langle G_2(v,w)\rangle - \int_W w\Big[\rho(v,w,t)G_1(v,w,s)\Big]\Big|_{\partial V} dw - \int_V \Big[wG_2(v,w)\rho(v,w,t)\Big]\Big|_{\partial W} dv. \end{aligned}$$

Note that $G_2(v,w)$ is linear, hence $\langle G_2(v,w)\rangle = G_2(\langle w\rangle, \langle v\rangle)$. Thus, returning to the flux notation, the equation can be rewritten

$$\langle w \rangle' = G_2(\langle w\rangle, \langle v\rangle) - \int_W w\Big[J_V(v,w,s,t)\Big]\Big|_{\partial V} dw - \int_V \Big[wJ_W(v,w,t)\Big]\Big|_{\partial W} dv \quad (4.10)$$

$$= G_2(\langle w\rangle, \langle v\rangle) - \int_W w\left(J_V(v_{peak},w,s,t) - J_V(v_{reset},w,s,t)\right)\, dw, \quad (4.11)$$

$$= G_2(\langle w\rangle, \langle v\rangle) - \int_W w\left(J_V(v_{peak},w,s,t) - J_V(v_{peak},w-w_{jump},s,t)\right)\, dw, \quad (4.12)$$

where we have used the assumption that $\rho(v,w,t) = 0$ on ∂W (which implies $J_W(v,w,s,t) = 0$ on ∂W) and the boundary condition (4.9).

To proceed further, we assume that $w_{jump} \ll 1$ and apply a Taylor expansion followed by integration by parts on line (4.13) to yield

$$\begin{aligned}
\langle w\rangle' &= G_2(\langle w\rangle, \langle v\rangle) \\
&- \int_W w \left(J_V(v_{peak}, w, s, t) - J_V(v_{peak}, w, s, t) + w_{jump} \frac{\partial J_V(v_{peak}, w, s, t)}{\partial w} + O(w_{jump}^2) \right) dw \\
&= G_2(\langle w\rangle, \langle v\rangle) - w_{jump} w J_V(v_{peak}, w, s, t)\Big|_{\partial W} + w_{jump} \int_W J_V(v_{peak}, w, s, t)\, dw + O(w_{jump}^2) \qquad (4.13) \\
&= G_2(\langle w\rangle, \langle v\rangle) + w_{jump} \int_W J_V(v_{peak}, w, s, t)\, dw + O(w_{jump}^2) \qquad (4.14)
\end{aligned}$$

Thus the partial differential equation (3.60) can be replaced by the partial differential equation (4.5) subject to the boundary condition (4.9) in addition to the ordinary differential equation (4.14).

We now have a first order partial differential equation and a pair of coupled ordinary differential equations describing our system:

$$\begin{aligned}
\frac{\partial}{\partial t}\rho(v,t) &= -\frac{\partial}{\partial v}\left((F(v) - \langle w\rangle + I + g(e_r - v)s)\, (\rho(v,t)) \right) \qquad (4.15) \\
\langle w\rangle' &= a(b\langle v\rangle - \langle w\rangle) + w_{jump} \int_W J_V(v_{peak}, w, s)\, dw + O(w_{jump}^2) \qquad (4.16) \\
s' &= -\frac{s}{\tau_s} + s_{jump} \int_W J_V(v_{peak}, w, s)\, dw, \qquad (4.17)
\end{aligned}$$

where we have dropped the V subscript on $\rho_V(v,t)$ for convenience. Thus far, everything we have done is exact, aside from the first order moment closure assumption. One level of approximation comes from dropping all the higher order terms in the expansion in (4.16), yielding

$$\langle w\rangle' \approx a(b\langle v\rangle - \langle w\rangle) + w_{jump} \int_W J_V(v_{peak}, w, s)\, dw.$$

Evaluating the integrals in equations (4.16)-(4.17) by applying the moment closure assumption and (4.6), we get the following dynamical system.

$$\begin{aligned}
\frac{\partial}{\partial t}\rho(v,t) &= -\frac{\partial}{\partial v}\left((F(v) - \langle w\rangle + I + g(e_r - v)s)\, (\rho(v,t)) \right) \qquad (4.18) \\
\langle w\rangle' &= a(b\langle v\rangle - \langle w\rangle) + w_{jump} J(v_{peak}, \langle w\rangle, s, t) \qquad (4.19) \\
s' &= -\frac{s}{\tau_s} + s_{jump} J(v_{peak}, \langle w\rangle, s, t). \qquad (4.20)
\end{aligned}$$

At this point we need another level of approximation to yield a system that is simple enough to analyze using bifurcation theory. This will take the form of a quasi-steady state approximation, based on a separation of time scales. However, prior to proceeding

further, we will numerically simulate the system of equations consisting of (4.18)-(4.20) and compare it to the actual network for various values of w_{jump} and fixed values of all the other parameters. The population density equation is solved using a Runge-Kutta 4,5 scheme implemented using ODE45 in MATLAB, with a first order in space differencing scheme. The ODEs for s and $\langle w \rangle$ are simultaneously integrated in the finite differencing scheme. The actual network consists of 3000 simulated neurons. The results are shown in figure 4.1. Note that as we are only using a first order finite differencing scheme, the leading order error is diffusive and thus the higher frequency components of the solution are not adequately picked up by the numerical solution. However, from the numerical simulations, it appears that the first order moment closure reduction employed does not contribute significantly to the error.

4.1.2 Second Order Moment Closure

For a homogeneous network, one can apply a higher order moment closure assumption with a straightforward physical meaning. In particular, we will consider the assumption

$$\langle w^2|v\rangle - \langle w|v\rangle^2 = \sigma^2_{w|v} = 0. \tag{4.21}$$

Multiplying equation (3.60) by w, integrating with respect to w followed by differentiation with respect to time yields:

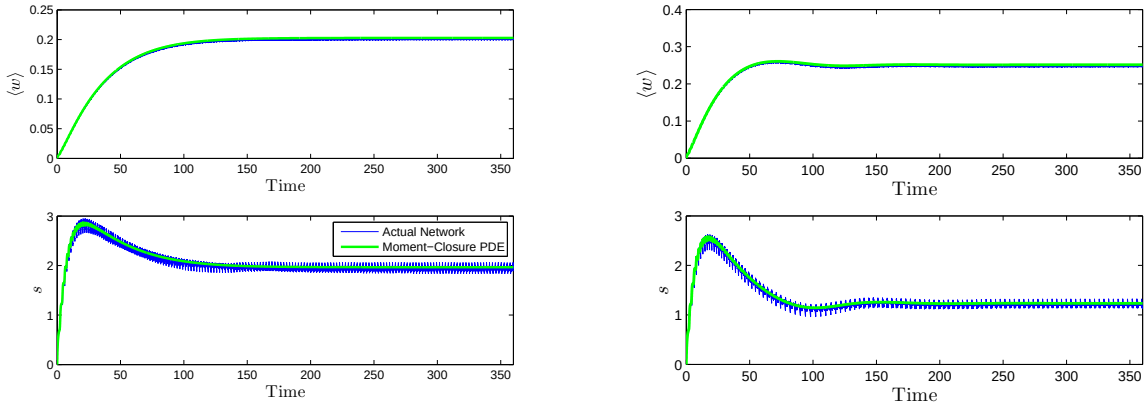
$$\begin{aligned}
\frac{\partial}{\partial t}\left(\rho_V(v,t)\langle w|v\rangle\right) &= -\frac{\partial}{\partial v}\left[\langle w|v\rangle G_1(v,s,\langle w|v\rangle)\rho_V(v,t)\right] + G_2(v,\langle w|v\rangle)\rho_V(v,t) &(4.22)\\
&= \frac{\partial \rho_V(v,t)}{\partial t}\langle w|v\rangle - G_1(v,s,\langle w|v\rangle)\rho_V(v,t)\frac{\partial \langle w|v\rangle}{\partial v} + G_2(v,\langle w|v\rangle)\rho_V(v,t)\\
\rho_V(v,t)\frac{\partial \langle w|v\rangle}{\partial t} &= -\rho_V(v,t)G_1(v,s,\langle w|v\rangle)\frac{\partial \langle w|v\rangle}{\partial v} + G_2(v,\langle w|v\rangle)\rho_V(v,t) &(4.23)
\end{aligned}$$

As every term in eq. (4.23) contains $\rho_V(v,t)$, we can factor it out (assuming it is non zero on $[v_{reset}, v_{peak}]$ for all t) which results in the following closed form equation for $\langle w|v\rangle$:

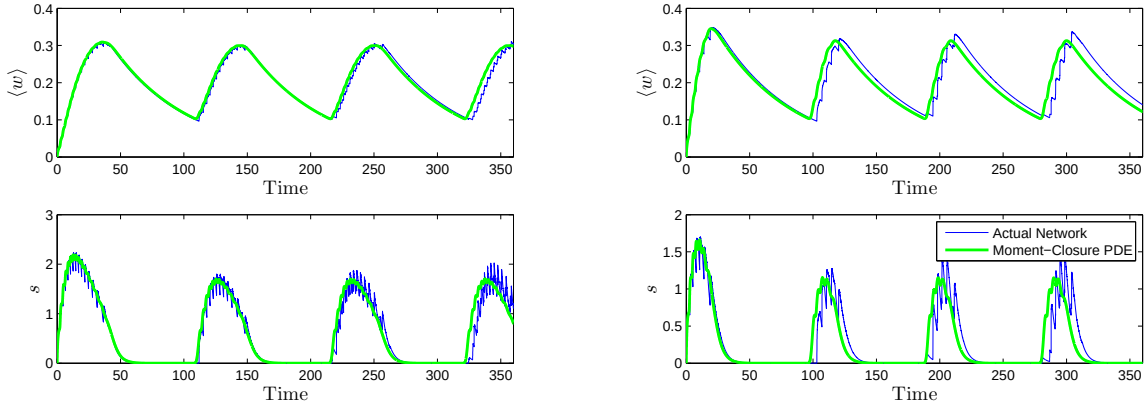
$$\frac{\partial \langle w|v\rangle}{\partial t} = -G_1(v,s,\langle w|v\rangle)\frac{\partial \langle w|v\rangle}{\partial v} + G_2(v,\langle w|v\rangle) \tag{4.24}$$

One can derive a boundary condition for equation (4.24) by deriving the differential equation for $\langle w\rangle$ in two ways; by integration with respect to v first and then w, and also interchanging the order of integration. Equating the differential equations for $\langle w\rangle$ that result, yields the following necessary equations for equivalence:

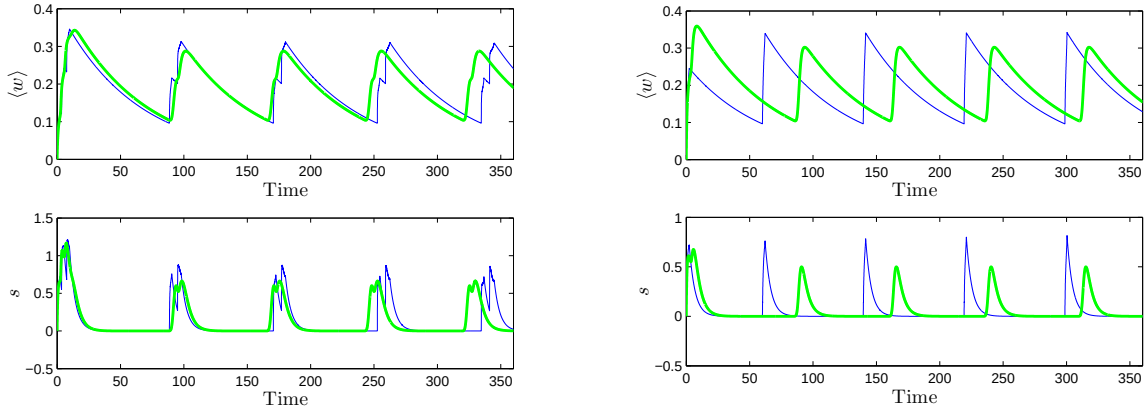
$$\begin{aligned}
\langle w|v_{reset}\rangle - \langle w|v_{peak}\rangle &= w_{jump} &(4.25)\\
\sigma^2_{w|v}\rho_V(v,t)\Big|_{\partial V} &= 0 &(4.26)
\end{aligned}$$



(a) $w_{jump} = 0.007$ (b) $w_{jump} = 0.015$



(c) $w_{jump} = 0.03$ (d) $w_{jump} = 0.06$



(e) $w_{jump} = 0.12$ (f) $w_{jump} = 0.24$

Figure 4.1: A comparison of a simulated network of 3000 neurons (blue) versus the numerical solution to an approximation to the population density equation (green). The population density equation was solved numerically with a first order in space, and fourth order in time finite differencing scheme. The parameter set is for a chattering neuron, and is from [92]

The second equation is immediately taken care of by the higher order moment closure assumption, (4.21). The first equation forms a boundary condition on $\langle w|v\rangle$. Coupling this partial differential equation to the PDE (4.4) for $\rho_V(v,t)$ and the ODE for s given by equation (3.61) gives the following system:

$$\begin{aligned}
\frac{\partial \rho_V}{\partial t} &= -\frac{\partial}{\partial v}\left(G_1(v,s,\langle w|v\rangle)\rho_V\right) && (4.27)\\
\frac{\partial \langle w|v\rangle}{\partial t} &= -G_1(v,s,\langle w|v\rangle)\frac{\partial \langle w|v\rangle}{\partial v} + G_2(v,\langle w|v\rangle) && (4.28)\\
\dot{s} &= -\frac{s}{\tau_s} + s_{jump}J(v_{peak},\langle w|v_{peak}\rangle,s,t) \\
&= -\frac{s}{\tau_s} + s_{jump}G_1(v_{peak},s,\langle w|v_{peak}\rangle)\rho_V(v_{peak},t) && (4.29)
\end{aligned}$$

where $v \in [v_{reset}, v_{peak}]$.

One can interpret the assumption $\sigma^2_{w|v} = 0$ statistically as the random variable w being a function of the random variable v, $w = g(v) = \langle w|v\rangle$ in which case the density in w will be determined by the standard change of variables formula:

$$\rho_W(w) = \rho_V(g^{-1}(w))\left|\frac{d}{dw}(g^{-1}(w))\right|.$$

We have simulated this system and it improves on the first order moment closure approach by providing more details and accuracy of the distribution of w by accurately approximating $\langle w|v\rangle$ (not shown). However, as the system is still too complicated for the purposes of analysis, we will primarily consider first order moment closure reduced system given by equations (4.18)-(4.20).

4.2 The Derivation of the Mean-Field System: A Separation of Time Scales

4.2.1 A Separation of Time Scales

In the type of neural models we are considering, the shortest time scale is typically the membrane potential time constant and the longest time scale is that of the adaptation variable. However, for inhibitory synapses or certain excitatory synapses, the longest timescale may be that of the s variable. As our starting system (4.18)-(4.20) is dimensionless, the membrane potential time constant is 1. Assuming the adaptation time scale is much longer, $\tau_w \gg 1$, we can regard $\epsilon = \tau_w^{-1}$ as a small parameter. Then, introducing the "slow time" $\tilde{t} = \tau_w^{-1} t$, the system becomes

$$\epsilon \frac{\partial}{\partial \tilde{t}} \rho(v, \tilde{t}) = -\frac{\partial}{\partial v} \left((F(v) - \langle w \rangle + I + g(e_r - v)s) \left(\rho(v, \tilde{t}) \right) \right) \tag{4.30}$$

$$\frac{d}{d\tilde{t}} \langle w \rangle = b\langle v \rangle - \langle w \rangle + \tau_w w_{jump} J(v_{peak}, \langle w \rangle, s, \tilde{t}) \tag{4.31}$$

$$\gamma \frac{ds}{d\tilde{t}} = -s + \tau_s s_{jump} J(v_{peak}, \langle w \rangle, s, \tilde{t}) \tag{4.32}$$

where the flux is defined equation (4.6). If $\tau_s = O(\tau_w) = \tau_w \gamma$, we can apply a quasi-steady state to equilibriate the PDE $\rho_V(v, t)$ and yield a system of coupled ODE's for s and $\langle w \rangle$. We will refer to (4.30)-(4.32) as the slow system. The resulting $O(1)$ solution to the slow system determines the mean-field system for the network. If however we consider our original time scale, then using $\epsilon = \tau_w^{-1}$ as before, we have the following.

$$\frac{\partial}{\partial t} \rho(v, t) = -\frac{\partial}{\partial v} \left((F(v) - \langle w \rangle + I + g(e_r - v)s) \left(\rho(v, t) \right) \right) \tag{4.33}$$

$$\frac{d\langle w \rangle}{dt} = \epsilon \left[b\langle v \rangle - \langle w \rangle + \lambda_w J(v_{peak}, \langle w \rangle, s, t) \right] \tag{4.34}$$

$$\frac{ds}{dt} = \epsilon \gamma^{-1} \left[-s + \lambda_s J(v_{peak}, \langle w \rangle, s, t) \right], \tag{4.35}$$

where $\lambda_s = \tau_s s_{jump}$, and $\lambda_w = \tau_w w_{jump}$. We will refer to equations (4.33)-(4.35) as the fast system. The $O(1)$ solution to the fast system is also analytically tractable, as we shall see. Much of our work on the fast system and perturbation analysis follows the approach of [2, 76]. Indeed, resolving the full spectral equation to a population density equation very similar to (4.30) was first done in [2]. We will first tackle the solution to the slow system, as the mean-field system of differential equations immediately follows without much trouble. In Section 4.4, we will analyze the full spectral equation and fast system.

4.2.2 The Slow System: The Mean-Field Derivation

To solve for the $O(1)$ system, we set $\epsilon = 0$ in equations (4.30)-(4.32). The resulting system implies that the density $\rho(v, t)$ reaches its steady state density $\rho(v)$, rapidly relative to s

and $\langle w \rangle$:

$$\begin{aligned} 0 &= \frac{\partial}{\partial v}\left[\left(F(v) - \langle w \rangle + I + g(e_r - v)s\right)\left(\rho(v)\right)\right] \\ &\Rightarrow \left[F(v) - \langle w \rangle + I + g(e_r - v)s\right]\rho(v) = J(\langle w \rangle, s), \end{aligned} \tag{4.36}$$

i.e., the flux is independent of v. Note that we have not specified which solution space the steady state in $\rho(v)$ belongs to. For now, we will assume a continuous solution, instead of a merely integrable (or weak solution) to equation ((4.36)). If $G_1(v, s, \langle w \rangle) > 0$, then we can write down the following steady state $\rho(v)$:

$$\rho(v) = \frac{J(\langle w \rangle, s)}{F(v) - \langle w \rangle + I + g(e_r - v)s)} = \frac{J(\langle w \rangle, s)}{G_1(v, s, \langle w \rangle)} \tag{4.37}$$

and, using the normalization condition on the density, we can solve for the flux:

$$J(\langle w \rangle, s) = \left[\int_V \frac{dv}{F(v) - \langle w \rangle + I + g(e_r - v)s}\right]^{-1}. \tag{4.38}$$

Using equation (4.38) and the ordinary differential equations for $\langle w \rangle$ and s, we now have a closed system of non-linear autonomous ODE's. Hence it would appear that we can use classical bifurcation theory to analyze the resulting bifurcations. However, one has to be careful when the denominator inside the integrand of (4.38) is not strictly positive on the phase space, that is $G_1(v, \langle w \rangle, s) < 0$ for some $v \in [v_{reset}, v_{peak}]$. If this occurs, then the flux defined (4.38) changes signs in the phase space. In this case one can no longer assume the density is continuous on $[v_{reset}, v_{peak}]$ or flux is constant ($\frac{\partial J}{\partial v} = 0$). For a more intuitive explanation, one can consider the pseudo-neuron with "average" dynamics

$$v' = F(v) - \langle w \rangle + I + g(e_r - v)s, \tag{4.39}$$

For the majority of neuronal models we consider (all models except the leaky integrate and fire model), a sign change in the flux corresponds to a saddle-node bifurcation for (4.39). Let v_{sn} denote the value of v when (4.39) undergoes a saddle-node bifurcation, additionally, regarding s, $\langle w \rangle$ as fixed parameters, then consider

$$I - I^*(s, \langle w \rangle) = \min_{v \in [v_{reset}, v_{peak}]} G_1(v, s, \langle w \rangle) = F(v^*) + gs(e_r - v^*) + I - \langle w \rangle \tag{4.40}$$

where v^* is the value of v in $[v_{reset}, v_{peak}]$ where $G_1(v, s, \langle w \rangle)$ attains its minimum. We must consider two cases:

Case I: $v_{sn} \in [v_{reset}, v_{peak}]$

If $v_{sn} \in [v_{reset}, v_{peak}]$, then $v* = v_{sn}$ and for $I - I^*(\langle w \rangle, s) < 0$, (4.39) has two fixed points $v = v_{\pm}(s, \langle w \rangle)$, with $v_{-}(s, \langle w \rangle)$ being the stable fixed point. For $I - I^*(\langle w \rangle, s) > 0$, (4.39)

has no fixed points, and for $I - I^*(\langle w\rangle, s) = 0$, the mean-neuron has one fixed point. The line $I - I^*(\langle w\rangle, s) = 0$ thus forms a switching manifold in the phase space of s and $\langle w\rangle$. For $I > I^*(\langle w\rangle, s)$, there are no fixed points and the system is in a tonic firing state and thus we have:

$$\langle R_i(t)\rangle = J(\langle w\rangle, s) = \left[\int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v) - \langle w\rangle + I + g(e_r - v)s}\right]^{-1} \tag{4.41}$$

$$\rho(v) = \frac{J(\langle w\rangle, s)}{F(v) - \langle w\rangle + I + g(e_r - v)s} \tag{4.42}$$

When the system (4.30)-(4.32) undergoes the saddle-node bifurcation, $I - I^*(s, \langle w\rangle) \leq 0$, and a pole emerges in the region of integration for (4.41). The solutions approach a stable fixed point (for $I - I^*(s, \langle w\rangle) < 0$), thus it seems reasonable to take the firing rate to be zero:

$$\langle R_i(t)\rangle = 0, \quad \text{if} \quad I \leq I^*(\langle w\rangle, s). \tag{4.43}$$

$$\rho(v) = \delta(v - v_-(s, \langle w\rangle)) \tag{4.44}$$

Also note that as we have $\rho(v)$, we can immediately compute $\langle v\rangle$, and in fact, any moment of any function of v through

$$\langle g(v)\rangle = \int_V \rho(v)g(v)\,dv = \begin{cases} \int_V \frac{g(v)J(s,\langle w\rangle)}{G_1(v,s,\langle w\rangle)}\,dv & I - I^*(s, \langle w\rangle) > 0 \\ g(v_-(s, \langle w\rangle)) & I - I^*(s, \langle w\rangle) \leq 0 \end{cases} \tag{4.45}$$

Indeed, in section 4.3, we will prove that $\langle R_i(t)\rangle \to 0$ as $I - I^*(s, \langle w\rangle) \to 0^+$. Additionally, we will resolve the leading order asymptotic behavior of $\langle R_i(t)\rangle$. However, for now we note that in this case, we obtain

$$\langle w\rangle' = a(b\langle v\rangle - \langle w\rangle) + w_{jump}\langle R_i(t)\rangle \tag{4.46}$$

$$s' = -\frac{s}{\tau_s} + s_{jump}\langle R_i(t)\rangle \tag{4.47}$$

$$\langle R_i(t)\rangle = \begin{cases} \left[\int_V \frac{dv}{G_1(v,s,\langle w\rangle)}\right]^{-1} & I > I^*(s, \langle w\rangle) \\ 0 & I < I^*(s, \langle w\rangle) \end{cases} \tag{4.48}$$

$$\rho(v) = \begin{cases} \frac{\langle R_i(t)\rangle}{G_1(v,s,\langle w\rangle)} & I > I^*(s, \langle w\rangle) \\ \delta(v - v_-(s, \langle w\rangle)) & I < I^*(s, \langle w\rangle) \end{cases} \tag{4.49}$$

$$\langle v\rangle = \int_V v\rho(v)\,dv. \tag{4.50}$$

This is the general mean-field model for the class of neural models we are looking at with simple exponential coupling. The following subsections will cover the specifics associated with the different neural models, and other types of synaptic coupling, in addition to multiple coupled populations.

Using the same parameters as Figure 4.1, we have simulated the mean-field system, as shown in Figure 4.2. For w_{jump} not too large, the mean-field system does quite well in reproducing the qualitative and quantitative behaviors of the full network. However, for larger values of w_{jump}, the mean-field system is accelerating with regards to the network level oscillations. The bulk of this error is due to the separation of time scales, as we can see numerically in figure 4.2 which compares the mean-field system, the moment closure reduced population density equation, and the original network.

Case II: $v_{sn} \notin [v_{reset}, v_{peak}]$

In this case, the saddle-node point does not occur in $[v_{reset}, v_{peak}]$. Thus, $I - I^*(s, \langle w \rangle) > 0$ even if the mean-neuron is already past the saddle-node bifurcation point, and two equilibria have emerged. So long as the two equilibria, $v_\pm \notin [v_{reset}, v_{peak}]$, then everything is defined as before. Only when either v_+ or v_- enter into $[v_{reset}, v_{peak}]$ is a pole generated at one of the boundary points, and the integral is no longer defined. Indeed, when this occurs we have $v^* = v_{reset}$ or $v^* = v_{peak}$ and $I - I^*(s, \langle w \rangle) = 0$ and thus equations (4.46)-(4.49) are still valid. One may wonder as to why we have separated out these two cases when the equations are precisely the same. The reason is that $\langle R_i(t) \rangle$ has a different asymptotic behavior as $I - I^*(s, \langle w \rangle) \to 0^+$ when $v_{sn} \in [v_{reset}, v_{peak}]$ versus $v_{sn} \notin [v_{reset}, v_{peak}]$. As we shall see, the asymptotics for the mean-field system here more closely resemble the leaky integrate and fire neuron then the type-1 normal form firing rate. However, $\langle R_i(t) \rangle \to 0$ even in this case, which we will prove in section 4.3.

Different Neural Models

For the majority of the neural models we have considered, $F(v)$ is a convex function, as $F''(v) > 0$, and the only transition from tonic firing to non-firing is through the switching manifold as described in Case I above. Effectively, the neurons in the network undergo a saddle-node bifurcation in a mean sense when $I - I^*(s, \langle w \rangle) = 0$. For the models we have thus far considered, the values of $I^*(s, \langle w \rangle)$ and $v_-(s, \langle w \rangle)$, the stable node past saddle-node bifurcation, are given by:

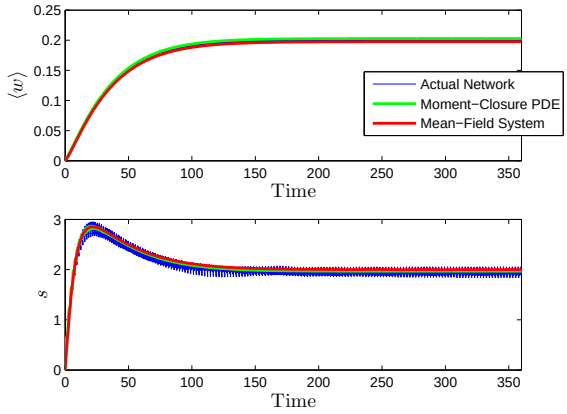
$$I^*(s, \langle w \rangle) = w - gse_r + \frac{(\alpha + gs)^2}{4} \tag{4.51}$$

$$v_-(s, \langle w \rangle) = \frac{\alpha + gs}{2} - \sqrt{I^*(s, \langle w \rangle) - I} \quad \text{(Izhikevich)} \tag{4.52}$$

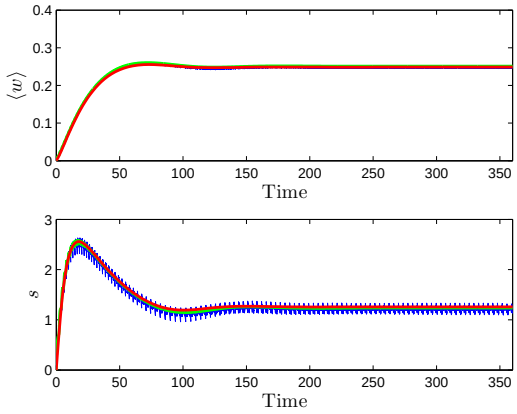
$$I^*(s, \langle w \rangle) = w - gse_r + (1 + gs)(\log(1 + gs) - 1) \tag{4.53}$$

$$v_-(s, \langle w \rangle) = \frac{I - \langle w \rangle + gse_r}{1 + gs} - W\left(-\frac{1}{1 + gs} \exp\left(\frac{I - \langle w \rangle + gse_r}{1 + gs}\right)\right) \quad \text{(AdEx)} \tag{4.54}$$

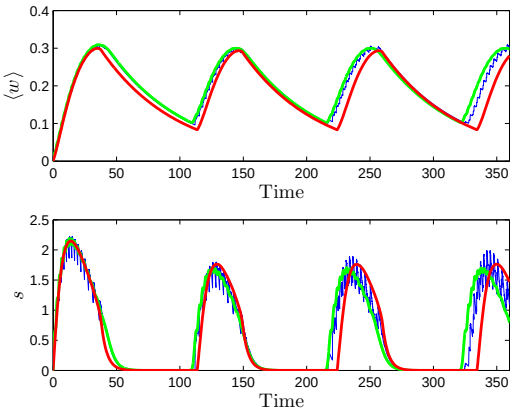
$$I^*(s, \langle w \rangle) = w - gse_r + 3\left(\frac{gs + 2a}{4}\right)^{4/3} \quad \text{(Quartic)} \tag{4.55}$$



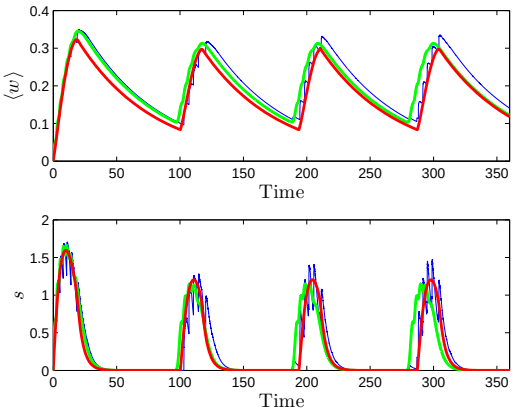
(a) $w_{jump} = 0.007$



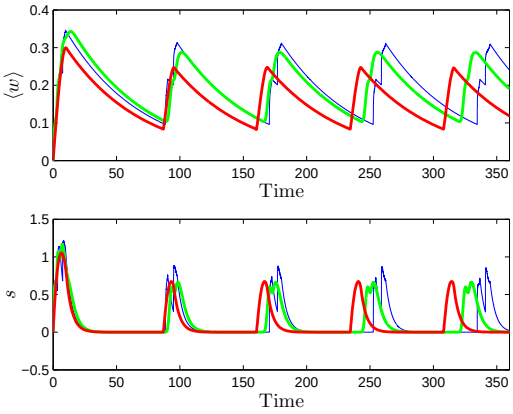
(b) $w_{jump} = 0.015$



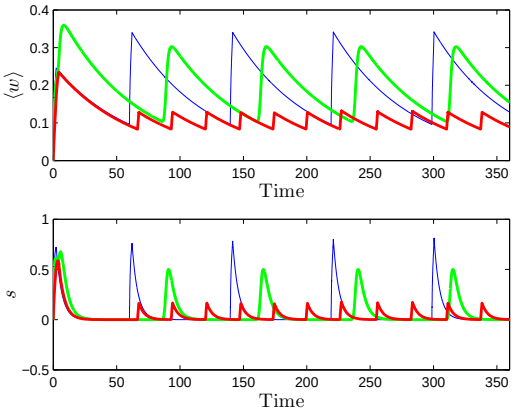
(c) $w_{jump} = 0.03$



(d) $w_{jump} = 0.06$



(e) $w_{jump} = 0.12$



(f) $w_{jump} = 0.24$

Figure 4.2: A comparison of a simulated network of 3000 neurons (blue) versus the numerical solution to an approximation to the population density equation (green equations (4.30)-(4.32)) and the mean-field system (red, equations (4.46)-(4.49)). The population density equation was solved numerically with a first order in space, and fourth order in time finite differencing scheme. The parameter set is for a chattering neuron, and is from [92].

where the $v_-(s, \langle w \rangle)$ for the quartic model is not included due to its complexity, and W is the Lambert function [39].

For the leaky integrate and fire neuron, however, the corresponding mean-field system that displays the transition from firing to quiescence outlined in Case II above. The resulting density and firing rates are:

$$\rho(v) = \begin{cases} \frac{\langle R_i(t) \rangle}{-\frac{v}{\tau_m} + I + gs(e_r - v) - \langle w \rangle} & I > I^*(s, \langle w \rangle) \\ \delta(v - v_-(s, \langle w \rangle)) & I < I^*(s, \langle w \rangle) \end{cases} \tag{4.56}$$

$$\langle R_i(t) \rangle = \begin{cases} \left[\int_V \frac{dv}{-\frac{v}{\tau_m} + I + gs(e_r - v) - \langle w \rangle} \right]^{-1} & I > I^*(s, \langle w \rangle) \\ 0 & I < I^*(s, \langle w \rangle) \end{cases} \tag{4.57}$$

Here, the minimum for the vector field in the average neuron dynamics actually occurs at $v = v_{peak}$ as $F'(v) < 0$, and thus $F(v)$ is strictly decreasing. Thus, the $I^*(s, \langle w \rangle)$ and $v_-(s, \langle w \rangle)$ are given by

$$I^*(s, \langle w \rangle) = -\frac{v_{peak}}{(1 + \frac{1}{\tau_m})} + I + gse_r - \langle w \rangle \tag{4.58}$$

$$v_-(s, \langle w \rangle) = \frac{I - \langle w \rangle + gse_r}{1/\tau_m + gs} \quad \text{(LIF)} \tag{4.59}$$

For both the leaky integrate and fire model and the Izhikevich model, the integrals for the firing rates are resolvable analytically when $I > I^*(s, \langle w \rangle)$. For the Izhikevich model we have:

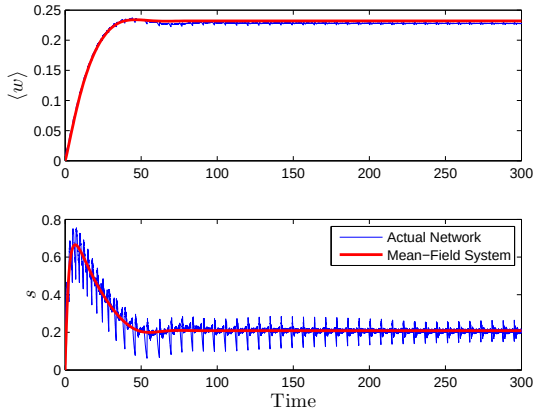
$$\langle R_i(t) \rangle = \frac{\sqrt{I - I^*(s, \langle w \rangle)}}{\arctan\left(\frac{v_{peak} - \frac{1}{2}(\alpha + gs)}{\sqrt{I - I^*(s, \langle w \rangle)}} \right) - \arctan\left(\frac{v_{reset} - \frac{1}{2}(\alpha + gs)}{\sqrt{I - I^*(s, \langle w \rangle)}} \right)} \quad I > I^*(s, \langle w \rangle) \tag{4.60}$$

and for the LIF model we have:

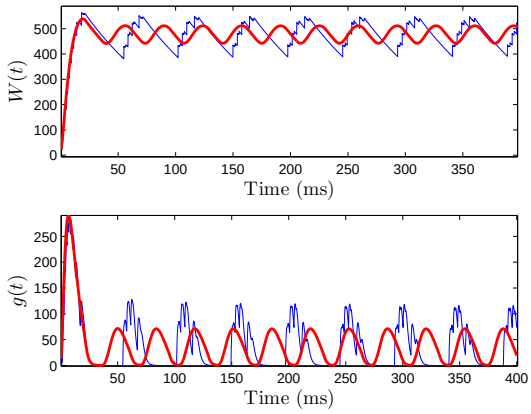
$$\langle R_i(t) \rangle = \frac{1/\tau_m + gs}{\log\left(\frac{-v_{reset}/\tau_m + gs(e_r - v_{reset}) + I - \langle w \rangle}{-v_t/\tau_m + gs(e_r - v_t) + I - \langle w \rangle} \right)} \quad I > I^*(s, \langle w \rangle). \tag{4.61}$$

For the other models, $\langle R_i(t) \rangle$ has to be numerically computed as a function of the parameters s and $\langle w \rangle$ using standard numerical integration methods to evaluate the integral (4.41) at each time step when simulating the mean-field system (i.e. trapezoidal method or quadrature). In either case however, the mean-field system can adequately predict quantitative and qualitative characteristics of the simulated network, as shown in Figure 4.3 for all the neuron models in consideration.

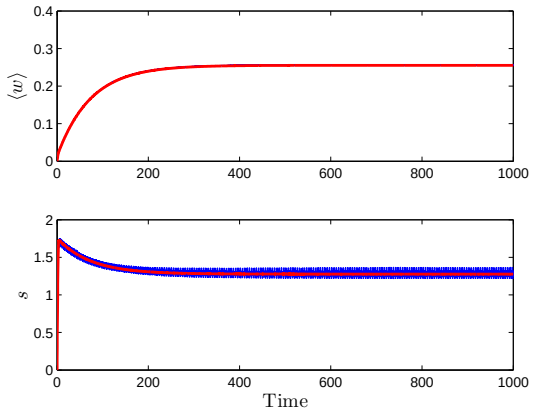
Although the firing rate functions $\langle R_i(t) \rangle$ are fairly complicated, asymptotically they do simplify quite substantially in two cases. We will see in section 4.3 that provided that certain assumptions are met, the firing rates asymptotically approach expressions that are similar to either the Izhikevich firing rate (equation (4.60), or the leaky integrate and fire firing rate (equation (4.61)) as $I - I^*(s, \langle w \rangle) \to 0$.



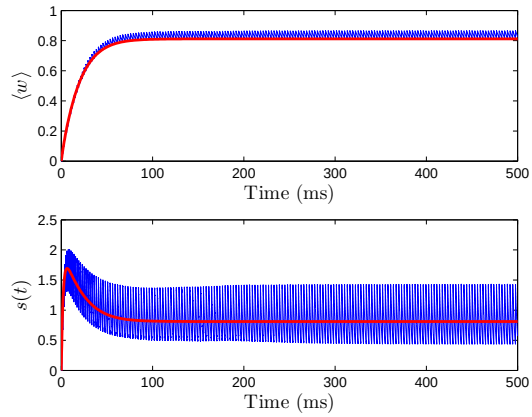
(a) Izhikevich Model



(b) AdEx Model



(c) LIF Model



(d) Quartic IF model

Figure 4.3: A comparison of the mean-field system (red) versus the actual network moments (blue) for the different models in consideration. All the models are coupled with simple exponential synapses in an all-to-all coupled network.

4.2.3 Synaptic Coupling Models

Of the synaptic coupling models we have looked at, the mean-field system can immediately be written down in the case of pulse coupling models. For example, for the double exponential synapse, we have the following mean-field system for an arbitrary neural model:

$$s' = -\frac{s}{\tau_R} + h \tag{4.62}$$

$$h' = -\frac{h}{\tau_D} + \frac{1}{\tau_R \tau_D}\langle R_i(t)\rangle \tag{4.63}$$

$$\langle w\rangle' = a(b\langle v\rangle - \langle w\rangle) + w_{jump}\langle R_i(t)\rangle \tag{4.64}$$

Setting $1/\alpha = \tau_R = \tau_D$ yields the α synapse. The single exponential synapse is effectively the default model we have dealt with already. For kinetic synapses however, the situation is more complicated.

Recall that for kinetic synapses, the equation for s_j was given by:

$$s_j' = (1 - s_j)\alpha[T]_j - \beta s_j \tag{4.65}$$

where $[T]_j$ varies depending on whether or not we have kinetic type I or kinetic type II synapses. For type I synapses, we have

$$[T] = \frac{T_{max}}{1 + \exp(-(v_j - v_p)/k_p)} = g(v_j)$$

while for type-II synapses,

$$[T] = \begin{cases} T_{max} & t_{j,k} < t < t_{j,k} + \bar{t} \\ 0 & \text{otherwise} \end{cases} = g(v_j)$$

where $t_{j,k}$ is the kth spike fired by the jth neuron. In order to derive the mean-field system for the kinetic synapses, one actually has to start with a 3-dimensional population density equation with the spatial variables (v, s, w) and apply the same moment closure approximations as in the preceding section. Applying the first order moment closure assumption to the Type I synapse in addition to the assumption that $\langle sg(v)\rangle \approx \langle s\rangle\langle g(v)\rangle$ immediately yields the differential equation

$$\langle s\rangle' = \alpha(1 - \langle s\rangle)\langle g(v)\rangle - \beta\langle s\rangle \tag{4.66}$$

where $\langle g(v)\rangle$ is computed using equation (4.45).

For the type II synapse, we can actually make a simpler reduction than straight forward averaging. In particular, we will approximate the dynamics of $s_j(t)$ with a linear system, that of the double exponential synapse:

$$\dot{s}_i = -\frac{s_i}{\tau_R} + h_i \tag{4.67}$$

$$\dot{h}_i = -\frac{s_i}{\tau_D} + \frac{A}{\tau_D \tau_R}\sum_{t<t_{i,k}} \delta(t - t_{i,k}) \tag{4.68}$$

where A is a weighting factor. We will use this two-dimensional linear synapse to approximate the more complicated Type-II synapse. We will do so by imposing the following constraints:

1. The type-II synapse and the double exponential synapse have the same synaptic rise and decay times.

2. The type-II synapse and the double exponential synapse have the same area underneath a pulse on $[0, \infty)$.

Note that both kinetic type II and double exponential synaptic models have analytical solutions. If a spike occurs at time $t = 0$, then the pulses are given by:

$$E_{Destexhe}(t) = \begin{cases} s_\infty \left(1 - \exp\left(-\frac{t}{\tau_s}\right)\right) & 0 < t < \bar{t} \\ s_\infty \left(1 - \exp\left(-\frac{\bar{t}}{\tau_S}\right)\right) \exp(-\beta(t - \bar{t})) & t \geq \bar{t} \end{cases} \tag{4.69}$$

$$E_{Exponential}(t) = \frac{A}{\tau_R - \tau_D}\left(\exp\left(-\frac{t}{\tau_R}\right) - \exp\left(-\frac{t}{\tau_D}\right)\right) \tag{4.70}$$

where

$$\tau_S = \frac{1}{\alpha T_{max} + \beta} \tag{4.71}$$

$$s_\infty = \frac{\alpha T_{max}}{\alpha T_{max} + \beta} \tag{4.72}$$

Now, it should be clear that to force the first constraint, we need to set

$$\tau_R = \tau_S, \quad \tau_D = \frac{1}{\beta}. \tag{4.73}$$

To force the second constraint, as the area underneath $E_{Exponential}(t)$ is A, then we can merely set A to be the area underneath the type II pulse, which is given by:

$$A = \int_0^\infty E_{Destexhe}(t')\, dt' = s_\infty \left[\bar{t} + \left(\frac{1}{\beta} - \tau_S\right)\left(1 - \exp\left(-\frac{\bar{t}}{\tau_S}\right)\right)\right] \tag{4.74}$$

Given the fact that we now have valued of the parameters A, τ_R and τ_D that force those two constraints, we can compare the type II and double-exponential pulses. This is shown in figure 4.4. The single pulse in figure 4.4(a) is fairly accurate, however, to be useful, the synapses must perform similarly for more then a single spike. When 5, 50, and 100 Hz background spiking is provided to each synapse, the steady state oscillation is similar between the two synaptic models, as shown in figure 4.4. However the similarity breaks down for spiking faster than around 150-200 Hz (not shown). Fortunately, spike rates of

less then 100Hz are suitable for most biological neurons. The resulting mean-field system is thus:

$$\langle s\rangle' = -\frac{\langle s\rangle}{\tau_R} + \langle h\rangle \tag{4.75}$$

$$\langle h\rangle' = -\frac{\langle h\rangle}{\tau_D} + \frac{A}{\tau_R\tau_D}\langle R_i(t)\rangle \tag{4.76}$$

$$\langle w\rangle' = a(b\langle v\rangle - \langle w\rangle) + w_{jump}\langle R_i(t)\rangle \tag{4.77}$$

where A and τ_R, τ_D are specifically determined by (4.74) and (4.73). A comparison of networks of Izhikevich models coupled together with the different synaptic models and the corresponding mean-field systems is shown in Figure 4.5.

4.2.4 Multiple Sub-Populations

For multiple sub-populations, the derivation of the mean-field system via a population density equation is as straight forward as before. In particular, consider a network consisting of K sub-populations, with N_j neurons in each network, for $j = 1, 2 \dots K$. Then the network equations are given by

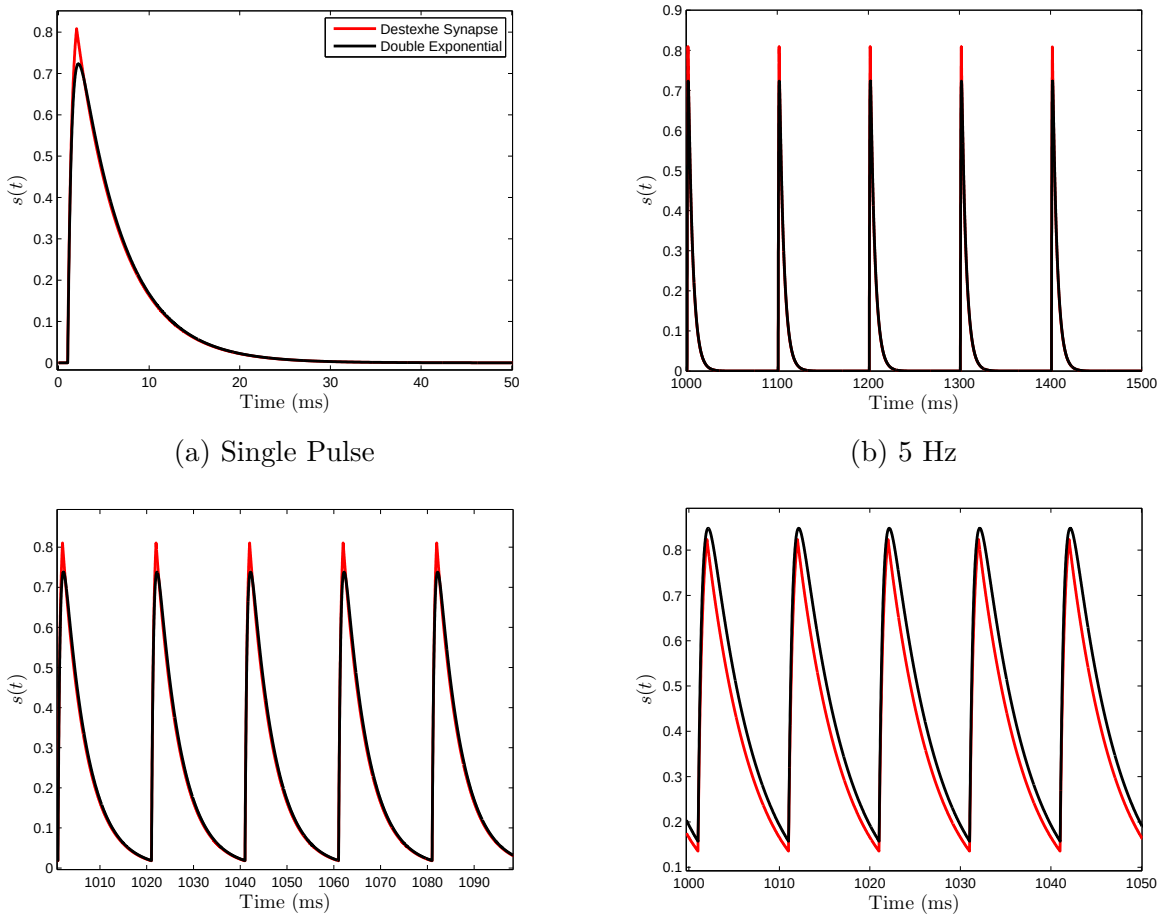
$$v'_{ij} = F_j(v_{ij}) - w_{ij} + I_j + \sum_{m=1}^{K} g_{mj}(e_{r,m} - v_{ij})s_m = G_j^v(v_{ij}, w_{ij}, s_1, s_2, \dots s_K) \tag{4.78}$$

$$w'_{ij} = a_j(b_j v_{ij} - w_{ij}) = G_j^w(v_{ij}, w_{ij}). \tag{4.79}$$

These are supplemented by the subpopulation specific resets in the voltage and jumps in the adaptation variables. The dynamics of s_j depends on which specific synaptic model is chosen, and $F_j(v)$ depends on which specific neuronal model is chosen for the jth subpopulation. One can immediately proceed to a population density equation for each subpopulation. The resulting equations are:

$$\frac{\partial\rho(v_j, w_j, t)}{\partial t} = -\frac{\partial}{\partial v_j}\left[G_j^v(v_j, w_j, s_1, \dots s_K)\rho(v_j, w_j, t)\right] - \frac{\partial}{\partial w_j}\left[G_j^w(v_j, w_j, \rho(v_j, w_j, t)\right] \tag{4.80}$$

for $j = 1, 2, \dots K$. One can then apply a first order moment closure approximation to each population density equation and resolve the slow system (equilibriate the population density equations) to obtain the mean-field model. For example, consider a network of two sub-populations of Izhikevich neurons with different parameter sets, coupled together with



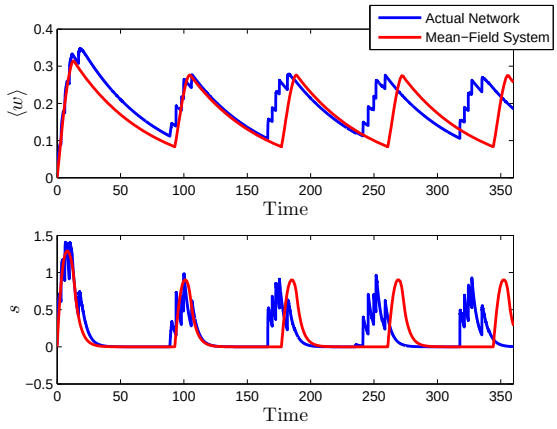
(a) Single Pulse

(b) 5 Hz

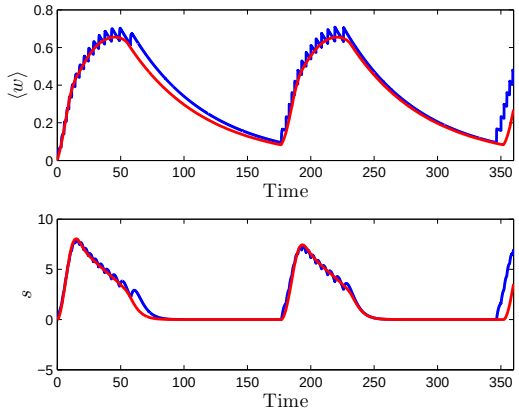
(c) 50 Hz

(d) 100 Hz

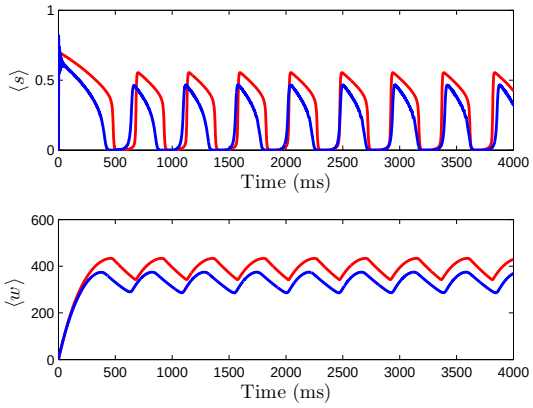
Figure 4.4: A comparison of the kinetic type II synapse model versus its double exponential approximation. A single pulse comparison is shown in figure 4.4(a). When background firing rates are at 5Hz, 50Hz, and 100 Hz, the double exponential approximation replicates the behavior of the type II kinetic synapses with adequate accuracy.



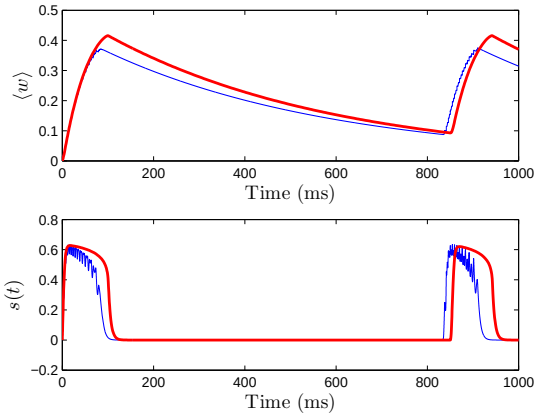
(a) Simple Exponential Synapse



(b) Alpha Synapse



(c) Type I Kinetic Synapse



(d) Type II Kinetic Synapse

Figure 4.5: A comparison of the mean-field system (red) versus the actual network moments (blue) for the different synaptic models in consideration. The network of neurons was an Izhikevich network, coupled together in an all-to-all fashion using the various types of synaptic models under consideration, with their corresponding mean-field approximations

simple exponential synapses. The network equations are given by

$$
\begin{aligned}
v'_{ij} &= v_{ij}(v_{ij}-\alpha_j) - w_{ij} + g_{j1}s_1(e_{r,1}-v_{ij}) + g_{j2}s_2(e_{r,2}-v_{ij}) + I_j && (4.81)\\
w'_{ij} &= a_j(b_j v_{ij} - w_{ij}) && (4.82)\\
v_{ij}(t_-) &= v_{peak,j} \Rightarrow \begin{cases} v_{ij}(t_+) = v_{reset,j} \\ w_{ij}(t_+) = w_{ij}(t_-) + w_{jump,j} \end{cases} && (4.83)\\
s'_j &= -\frac{s_j}{\tau_{s,j}} + s_{jump}\frac{1}{N_j}\sum_{i=1}^{N_j}\sum_{t<t_{j,i,k}} \delta(t-t_{j,i,k}) && (4.84)
\end{aligned}
$$

for $j = 1, 2$ and $i = 1, 2, \ldots N_j$. One can immediately write the moment closure reduced population density system:

$$
\begin{aligned}
\frac{\partial}{\partial t}\rho_1(v_1,t) &= -\frac{\partial}{\partial v_1}\left[(v_1(v_1-\alpha_1) - \langle w_1\rangle + g_{1,1}s_1(e_{r,1}-v_1) + g_{1,2}s_2(e_{r,2}-v_1) + I_1)\rho_1(v_1,t)\right]\\
&= -\frac{\partial}{\partial v_1}J_1(v_1,\langle w_1\rangle, s_1, s_2, t)\\
\frac{\partial}{\partial t}\rho_2(v_2,t) &= -\frac{\partial}{\partial v_2}\left[(v_2(v_2-\alpha_2) - \langle w_2\rangle + g_{2,1}s_1(e_{r,1}-v_2) + g_{2,2}s_2(e_{r,2}-v_2) + I_2)\rho_2(v_2,t)\right]\\
&= -\frac{\partial}{\partial v_2}J_2(v_2,\langle w_2\rangle, s_1, s_2, t)\\
s'_1 &= -\frac{s_1}{\tau_{s,1}} + s_{jump,1}J_1(v_{peak,1},\langle w_1\rangle, s_1, s_2, t)\\
s'_2 &= -\frac{s_2}{\tau_{s,2}} + s_{jump,2}J_2(v_{peak,2},\langle w_2\rangle, s_1, s_2, t)\\
\langle w_1\rangle' &= a_1(b_1\langle v_1\rangle - \langle w_1\rangle) + w_{jump,1}J_1(v_{peak,1},\langle w_1\rangle, s_1, s_2, t)\\
\langle w_2\rangle' &= a_2(b_2\langle v_2\rangle - \langle w_2\rangle) + w_{jump,2}J_2(v_{peak,2},\langle w_2\rangle, s_1, s_2, t)
\end{aligned}
$$

which results in the following mean-field system:

$$
\begin{aligned}
s_1' &= -\frac{s_1}{\tau_{s,1}} + s_{jump,1}\langle R_i(t)\rangle_1 \\
s_2' &= -\frac{s_2}{\tau_{s,2}} + s_{jump,2}\langle R_i(t)\rangle_2 \\
\langle w_1\rangle' &= a_1(b_1\langle v_1\rangle - \langle w_1\rangle) + w_{jump,1}\langle R_i(t)\rangle_1 \\
\langle w_2\rangle' &= a_2(b_2\langle v_2\rangle - \langle w_2\rangle) + w_{jump,2}\langle R_i(t)\rangle_2 \\
\langle R_i(t)\rangle_1 &= \begin{cases} \frac{\sqrt{I_1 - I_1^*(s_1,s_2,\langle w_1\rangle)}}{\arctan\left(\frac{v_{peak,1} - \frac{1}{2}(\alpha + g_{11}s_1 + g_{12}s_2)}{\sqrt{I_1 - I_1^*(s_1,s_2,\langle w_1\rangle)}}\right) - \arctan\left(\frac{v_{reset,1} - \frac{1}{2}(\alpha + g_{11}s_1 + g_{12}s_2)}{\sqrt{I_1 - I_1^*(s_1,s_2,\langle w_1\rangle)}}\right)} & I_1 > I_1^*(s_1, s_2\langle w\rangle_1) \\ 0 & I_1 \leq I_1^*(s_1, s_2\langle w\rangle_1) \end{cases} \\
\langle R_i(t)\rangle_2 &= \begin{cases} \frac{\sqrt{I_2 - I_2^*(s_1,s_2,\langle w_2\rangle)}}{\arctan\left(\frac{v_{peak,2} - \frac{1}{2}(\alpha + g_{21}s_1 + g_{22}s_2)}{\sqrt{I_2 - I_2^*(s_1,s_2,\langle w_2\rangle)}}\right) - \arctan\left(\frac{v_{reset,2} - \frac{1}{2}(\alpha + g_{21}s_1 + g_{22}s_2)}{\sqrt{I_2 - I_2^*(s_1,s_2,\langle w_2\rangle)}}\right)} & I_2 > I_2^*(s_1, s_2\langle w\rangle_2) \\ 0 & I_2 \leq I_2^*(s_1, s_2\langle w\rangle_2) \end{cases} \\
I_1^*(s_1, s_2, \langle w_1\rangle) &= \frac{(\alpha_1 + g_{11}s_1 + g_{12}s_2)^2}{4} - g_{11}s_1e_{r,1} - g_{12}s_2e_{r,2} - \langle w_1\rangle \\
I_2^*(s_1, s_2, \langle w_2\rangle) &= \frac{(\alpha_2 + g_{21}s_1 + g_{22}s_2)^2}{4} - g_{21}s_1e_{r,1} - g_{22}s_2e_{r,2} - \langle w_2\rangle \\
\rho_1(v_1) &= \begin{cases} \frac{\langle R_i(t)\rangle_1}{v_1(v_1-\alpha_1) + g_{11}s_1(e_{r,1}-v_1) + g_{12}s_2(e_{r,2}-v_1) + I_1 - \langle w_1\rangle} & I_1 > I_1^*(s_1, s_2, \langle w_1\rangle) \\ \delta(v_1 - v_{-1}(s_1, s_2, \langle w_1\rangle)) & I_1 \leq I_1^*(s_1, s_2, \langle w_1\rangle) \end{cases} \\
\rho_2(v_2) &= \begin{cases} \frac{\langle R_i(t)\rangle_2}{v_2(v_2-\alpha_1) + g_{21}s_1(e_{r,1}-v_2) + g_{22}s_2(e_{r,2}-v_2) + I_2 - \langle w_2\rangle} & I_2 > I_2^*(s_1, s_2, \langle w_2\rangle) \\ \delta(v_2 - v_{-2}(s_1, s_2, \langle w_1\rangle)) & I_1 \leq I_2^*(s_1, s_2, \langle w_2\rangle) \end{cases} \\
v_{-1} &= \frac{\alpha_1 + g_{11}s_1 + g_{12}s_2}{2} - \sqrt{I_1^*(s_1, s_2, \langle w_1\rangle) - I_1} \\
v_{-2} &= \frac{\alpha_2 + g_{21}s_1 + g_{22}s_2}{2} - \sqrt{I_2^*(s_1, s_2, \langle w_2\rangle) - I_2}
\end{aligned}
$$

4.3 The Asymptotics of $\langle R_i(t)\rangle$ in the Slow System

In this section, we will show that the firing rate:

$$\langle R_i(t)\rangle = \left(\int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v) - \langle w\rangle + gs(e_r - v) + I}\right)^{-1}$$

simplifies asymptotically as $I - I^*(s, w) \to 0$, which we defined as

$$I - I^*(s, \langle w\rangle) = \min_{v\in[v_{reset}, v_{peak}]} \{F(v) - \langle w\rangle + gs(e_r - v) + I\}$$

In effect, what this section will demonstrate is that as the trajectories of (s, w) tend toward the switching manifold, the dynamics become "nice" in the sense that the bifurcation structure of the system becomes analytically tractable, which we shall see in Chapter 5 of this thesis.

Recall the assumptions made on $F(v)$: $F''(v) > 0$ for all v and $F'(v) < 0$ as $v \to -\infty$ and $F'(v) > 0$ as $v \to \infty$. Returning to our mean neuron with these assumptions in mind:

$$\begin{aligned}
\dot{v} &= F(v) - \langle w\rangle + gs(e_r - v) + I && (4.85)\\
&= I + F(v^*) - \langle w\rangle + gs(e_r - v^*) + \frac{F''(v^*)}{2}(v - v^*)^2 + O((v - v^*)^3) && (4.86)\\
&= I - I^*(s, \langle w\rangle) + \frac{F''(v^*)}{2}(v - v^*)^2 + O((v - v^*)^3) && (4.87)
\end{aligned}$$

where v^* satisfies the equation $F'(v^*) = gs$. Then as this system undergoes a saddle-node bifurcation when $I - I^*(s, \langle w\rangle) = 0$, topological normal form theory tells us that the mean-neuron is locally topologically equivalent to

$$\dot{v} = I - I^*(s, \langle w\rangle) + \frac{F''(v^*)}{2}(v - v^*)^2 \qquad (4.88)$$

in the vicinity of $I - I^*(s, \langle w\rangle) = 0$ but this is merely a quadratic integrate and fire neuron. The topological reduction to the related theta model for a conductance based model undergoing a saddle-node on an invariant circle (SNIC) bifurcation was done in [60]. Given that, then we should expect that $\langle R_i(t)\rangle$ is somehow asymptotically related to the firing rate of the quadratic integrate and fire neuron (4.88) as $I - I^*(s, \langle w\rangle) \to 0$. The term $v^*(gs)$ corresponds to the minimum of $G_1(v, s, \langle w\rangle)$ where the switching manifold of our system is defined by $G_1(v^*(gs), s, \langle w\rangle) = 0$.

Depending on the location of $v^*(gs)$ relative to the interval $[v_{reset}, v_{peak}]$, we actually obtain different asymptotic expressions for the firing rate. In particular, suppose that $v^*(gs) \in [v_{reset}, v_{peak}]$. Then, as stated earlier, two equilibria v_- and v_+ emerge with $v_- < v_+$ and when $G_1(v^*(gs), s, \langle w\rangle) = 0$, then $v^*(gs) = v_{SN}$, a saddle-node bifurcation point for the mean-neuron. In the vicinity of this bifurcation,

$$\langle R_i(t)\rangle \to \frac{1}{\pi}\sqrt{\frac{F''(v^*(gs))}{2}}\sqrt{I - I^*(s, \langle w\rangle)} \quad \text{as} \quad I - I^*(s, \langle w\rangle) \to 0 \qquad (4.89)$$

We also have to consider the situation when $v^*(gs) \notin [v_{reset}, v_{peak}]$. In this situation we have either $v^*(gs) > v_{peak}$ or $v^*(gs) < v_{reset}$. For the first case, in the vicinity of the saddle node bifurcation we have $v_- > v_{peak}$ and v_- attracts $v(t)$. However, as soon as $v(t) = v_{peak}$, it is forced back to v_{reset} by the threshold condition and this process continues cyclically. In this case, we have tonic firing and $\langle R_i(t) \rangle > 0$. If $v^*(gs) < v_{reset}$, then, two equilibria emerge and close to the bifurcation point, we will have $v_- < v_+ < v_{reset}$. As v_+ is repelling, it forces $v(t)$ towards v_{peak} and tonic firing continues. However, the analysis done here is only valid locally. In particular, for $v^*(gs) > v_{peak}$, once $v_- = v_{peak}$ firing ceases. The same is true for $v^*(gs) < v_{reset}$, once $v_+ = v_{reset}$. See figure 4.6 for a graphical depiction of what happens to the mean neuron in all these cases. In both cases, will see that the leading order asymptotics for the firing rate are those of the leaky integrate and fire neuron regardless of the neural model used.

Throughout the subsequent derivations, we will assume that $F(v)$ is an analytic function. First we will look at $v^*(gs) \in [v_{reset}, v_{peak}]$, then we will deal with $v^*(gs) \notin [v_{reset}, v_{peak}]$.

4.3.1 Case I: The Type-I Firing Rate When $v^*(gs) \in [v_{reset}, v_{peak}]$

Both here, and in case II, we will effectively be applying Laplace's method/the saddle-point approximation [17]. Unfortunately, however, as the reciprocal of the firing rate is not specifically a Laplace type integral with an exponential function, we will have to proceed in a more direct and lengthy fashion. Note the reciprocal of the firing rate is given by:

$$\begin{aligned}
\frac{1}{\langle R_i(t) \rangle} &= \int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v) - \langle w \rangle + gs(er - v) + I} \\
&= \int_{v_{reset}}^{v_{peak}} \frac{dv}{I - I^*(s, \langle w \rangle) + \frac{F''(v^*)}{2}(v - v^*)^2 + \sum_{n=3}^{\infty} \frac{F^{(n)}(v^*)}{n!}(v - v^*)^n} \\
&= \int_{v_{reset}}^{v_{peak}} \frac{dv}{I - I^*(s, \langle w \rangle) + \frac{F''(v^*)}{2}(v - v^*)^2 + \sum_{n=3}^{\infty} \frac{F^{(n)}(v^*)}{n!}(v - v^*)^n}
\end{aligned}$$

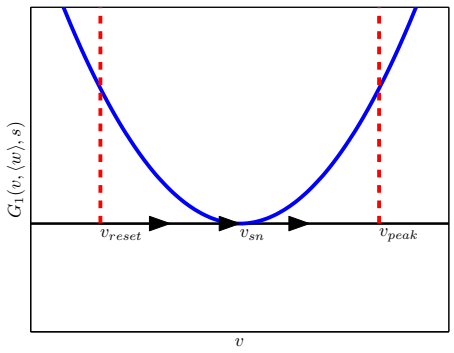
Now, to proceed further, we will have to make a very particular substitution. First, let $\kappa = I - I^*(s, \langle w \rangle) > 0$. Recalling the Izhikevich neuron, consider the substitution

$$z = \frac{v - v^*}{\sqrt{\kappa}}, \quad dz = dv/\sqrt{\kappa}$$

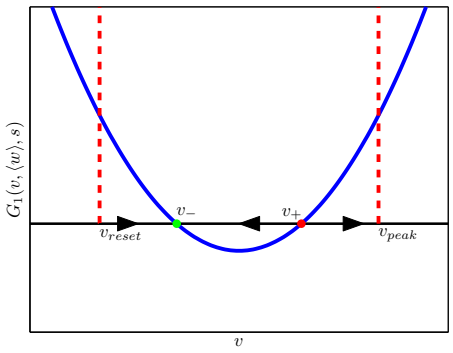
Applying this substitution yields the following:

$$\frac{1}{\langle R_i(t) \rangle} = \int_{(v_{reset} - v^*)/\sqrt{\kappa}}^{(v_{peak} - v^*)/\sqrt{\kappa}} \frac{\sqrt{\kappa}}{\kappa + \frac{F''(v^*)}{2} z^2 \kappa + \sum_{n=1}^{\infty} \frac{F^{n+2}(v^*)}{(n+2)!}(z\sqrt{\kappa})^{n+2}} dz \tag{4.90}$$

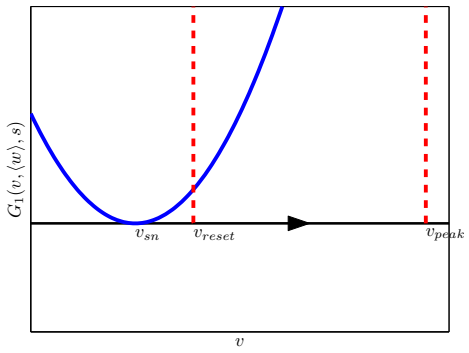
$$= \frac{1}{\sqrt{\kappa}} \int_{(v_{reset} - v^*)/\sqrt{\kappa}}^{(v_{peak} - v^*)/\sqrt{\kappa}} \frac{1}{1 + \frac{F''(v^*)}{2} z^2} \frac{1}{1 + \frac{\sum_{n=1}^{\infty} \frac{F^{n+2}(v^*)}{(n+2)!} z^{n+2} (\sqrt{\kappa})^n}{1 + \frac{F''(v^*)}{2} z^2}} dz \tag{4.91}$$



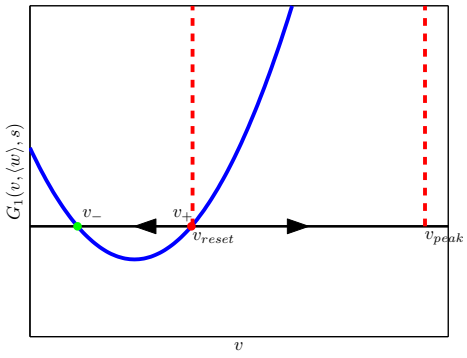
(a) $v_{sn} \in [v_{reset}, v_{peak}]$, $I - I^*(s, \langle w \rangle) = 0$)



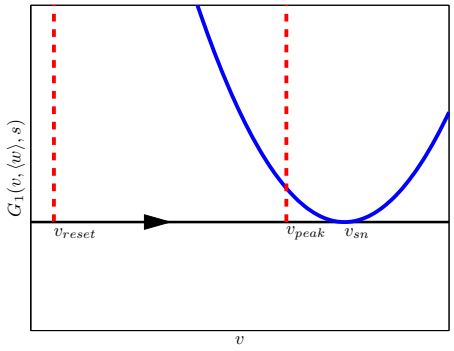
(b) $v_{sn} \in [v_{reset}, v_{peak}]$, $I - I^*(s, \langle w \rangle) < 0$)



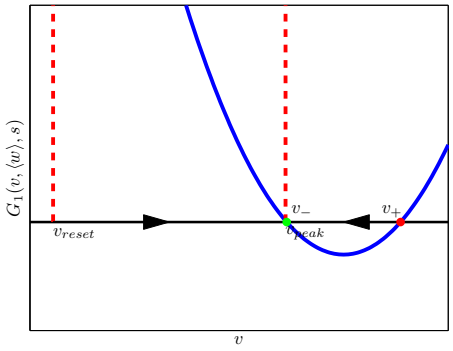
(c) $v_{sn} < v_{reset}]$, $I - I^*(s, \langle w \rangle) > 0$)



(d) $v_{sn} < v_{reset}$,$I - I^*(s, \langle w \rangle) = 0$)



(e) $v_{sn} > v_{peak}$, $I - I^*(s, \langle w \rangle) > 0$)



(f) $v_{sn} > v_{peak}$, $I - I^*(s, \langle w \rangle) = 0$)

Figure 4.6: The asymtotics of $\langle R_i(t) \rangle$ and the behavior of the mean neuron depend on the relationship between v_{sn}, v_{reset}, and v_{peak}. The blue curve denotes $G_1(v, s, \langle w \rangle)$, the black line denotes the $G = 0$ axis, and the arrows denote the vector field for the mean-neuron. When $v_{sn} \in [v_{reset}, v_{peak}]$, the immediate emergence of a pair of nodes when $I - I^*(s, \langle w \rangle) < 0$ stops firing. In this case, $\langle R_i(t) \rangle \propto \sqrt{I - I^*(s, \langle w \rangle)}$ as $I - I^*(s, \langle w \rangle) \to 0^+$. When $v_{sn} < v_{reset}$ or $v_{sn} > v_{peak}$, only when $v_+ = v_{reset}$ or $v_- = v_{peak}$ does the firing stop, respectively. In this case, v_{reset} or v_{peak} corresponds to the v^*, the minimum of G on $[v_{reset}, v_{peak}]$. In this case, $\langle R_i(t) \rangle$ tends to a leaky integrate and firing firing rate as $I - I^*(s, \langle w \rangle) \to 0^+$.

Now, here is where we have to be careful. In particular, we want to expand out the second term as a geometric series, but for that, we require

$$\left|\sum_{n=1}^{\infty}\frac{F^{n+2}(v^*)}{(n+2)!}z^{n+2}(\sqrt{\kappa})^n\right| \leq 1+\frac{F''(v^*)}{2}z^2 \tag{4.92}$$

In particular, we have to confine z to a neighbourhood. The reason for this is that the domain of integration is growing like $1/\sqrt{\kappa}$. Now it's clear that for $z \ll 1$, the condition (4.92) holds by default as the left hand sides leading order term is z^3, while the right hand side is $O(1)$. For large z, in particular for $z = O(1/\sqrt{\kappa})$, we have another requirement that must hold. Letting $M = \max\{|v_{peak}-v^*|, |v_{reset}-v^*|\}$, then for $\zeta M/\sqrt{\kappa} < z < M/\sqrt{\kappa}$, $0<\zeta<1$, then condition (4.92) is satisfied provided that:

$$\sum_{n=1}^{\infty}\frac{|F^{n+2}(v^*)|}{(n+2)!}\left(\frac{M}{\sqrt{\kappa}}\right)^{n+2}(\sqrt{\kappa})^n \leq 1+\zeta^2\frac{F''(v^*)}{2!}\frac{M^2}{\kappa} \tag{4.93}$$

$$\frac{1}{\kappa}\sum_{n=1}^{\infty}\frac{|F^{n+2}(v^*)|}{(n+2)!}M^{n+2} \leq 1+\zeta^2\frac{F''(v^*)}{2!}\frac{M^2}{\kappa} \tag{4.94}$$

Now, note that the dominant terms in the inequality are κ^{-1} as $\kappa \ll 1$ This immediately implies that we require the following to hold:

$$\sum_{n=1}^{\infty}\frac{|F^{n+2}(v^*)|}{(n+2)!}M^n \leq \zeta^2\frac{F''(v^*)}{2!} \tag{4.95}$$

Now, as the left hand side is a function of M, it can be made arbitrarily small to satisfy the inequality. Additionally, the series on the left is convergent assuming that $F(v)$ is analytic at v^*. So, in effect, if the second derivative dominates in a suitably small neighbourhood with a width of M around v^*, we are fine taking the geometric series. For the neural models we are considering, this translates to the condition that if v_{peak} and v_{reset} are suitably close to v^*, the threshold for firing, and if the second derivative is suitably dominant, then the asympotics we have derived here apply to the neurons firing rate.

Returning to our original integral:

$$\frac{1}{\langle R\rangle} = \frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{1}{1+\frac{F''(v^*)}{2}z^2}\frac{1}{\left(1+\frac{\sum_{n=1}^{\infty}\frac{F^{n+2}(v^*)}{(n+2)!}z^{n+2}(\sqrt{\kappa})^n}{1+\frac{F''(v^*)}{2}z^2}\right)}\,dz \tag{4.96}$$

$$= \frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{1}{1+\frac{F''(v^*)}{2}z^2}\left(1-\frac{\sum_{n=1}^{\infty}\frac{F^{n+2}(v^*)}{(n+2)!}z^{n+2}(\sqrt{\kappa})^n}{(1+\frac{F''(v^*)}{2}z^2)}+HOT\right)\,dz \tag{4.97}$$

$$= \frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{1}{1+\frac{F''(v^*)}{2}z^2} - \frac{1}{\sqrt{\kappa}}\sum_{n=1}^{\infty}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{F^{(n+2)}(v^*)z^{n+2}(\sqrt{\kappa})^n}{(n+2)!(1+\frac{F''(v^*)}{2}z^2)^2}\,dz+HOT$$

where HOT denotes higher order terms. Now, note that both the first and second term contain integrals with divergent boundaries. The first, we will denote as $1/\langle R_0\rangle = 1/\phi_0(\kappa)$,

where $\phi_0(\kappa)$ will be used as a gauge function later on in an asymptotic analysis and is merely the firing rate for the type I normal form. The second however, is the first term in the asymptotic sequence and we need to show that it is $O(1)$. If this is the case, then $1/\langle R_i(t)\rangle = \frac{1}{\Omega(\kappa)}$ diverges as $\kappa \to 0$ as the first term diverges, and the next highest order term in the expansion is $O(1)$. More formally, if we show that the next highest order term is $O(1)$ then we have

$$\lim_{\kappa\to 0}\Omega(\kappa) = \lim_{\kappa\to 0}\langle R_i(t)\rangle = \lim_{\kappa\to 0}\left(\frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}} \frac{dz}{1+\frac{F''(v^*)}{2}z^2+\sum_{n=1}^{\infty}\frac{F^{n+2}(v^*)}{(n+2)!}z^{n+2}\sqrt{\kappa}^n}\right)^{-1} = 0$$

The sum of integrals contains integrands that are functions of the form

$$\frac{z^m}{(1+az^2)^2}, \quad m = 3\ldots$$

with a divergent boundary. Now, we'll need to analyze two cases, $m = 3$ and $m > 3$. For $m = 3$, we have the following:

$$\int \frac{F'''(v^*)}{3!}\frac{z^3}{(1+\frac{F''(v^*)}{2}z^2)^2} = \frac{F'''(v^*)}{3!}\left(\frac{4}{2F''(v^*)^2(z^2F''(v^*)^2+4)}+\frac{2}{F''(v^*)}\log\left(1+\frac{F''(v^*)}{2}z^2\right)\right) \tag{4.98}$$

Which yields the following when evaluated with the region of integration

$$\begin{aligned}
&= \frac{F'''(v^*)}{3!}\left(\frac{4\kappa}{2F''(v^*)^2((v_{peak}-v^*)^2F''(v^*)^2+4\kappa)}-\frac{4\kappa}{2F''(v^*)^2((v_{reset}-v^*)^2F''(v^*)^2+4\kappa)}\right)- \\
&+ \frac{F'''(v^*)}{3!}\left(\frac{2}{F''(v^*)}\left[\log\left(1+\frac{F''(v^*)}{2\kappa}(v_{peak}-v^*)^2\right)-\log\left(1+\frac{F''(v^*)}{2\kappa}(v_{reset}-v^*)^2\right)\right]\right) \\
&= O(\kappa)+\frac{F'''(v^*)}{3!}\frac{2}{F''(v^*)}\log\left(\frac{2\kappa+F''(v^*)(v_{peak}-v^*)^2}{2\kappa+F''(v^*)(v_{reset}-v^*)^2}\right) = O(\kappa)+O(1)
\end{aligned}$$

thus, the dominant term is $O(1)$. For $m > 3$, we have

$$\frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{F^{(n+2)}(v^*)z^{n+2}(\sqrt{\kappa})^n}{(n+2)!(1+\frac{F''(v^*)}{2}z^2)^2}\,dz \quad n > 2 \tag{4.99}$$

The reason why we have a critical split is because for $m \geq 4$, the numerator over-powers the denominator on a region of integration that diverges. However, the integral is in fact

bounded:

$$\begin{aligned}
\left|\frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{F^{(n+2)}(v^*)z^{n+2}(\sqrt{\kappa})^n}{(n+2)!(1+\frac{F''(v^*)}{2}z^2)^2}\,dz\right| &\leq \frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{|F^{(n+2)}(v^*)||z|^{n+2}(\sqrt{\kappa})^n}{(n+2)!(1+\frac{F''(v^*)}{2}z^2)^2}\,dz \\
&\leq \frac{2}{\sqrt{\kappa}}\int_0^{M/\sqrt{\kappa}}\frac{|F^{(n+2)}(v^*)|z^{n+2}(\sqrt{\kappa})^n}{(n+2)!(1+\frac{F''(v^*)}{2}z^2)^2}\,dz \\
&\leq \frac{2}{\sqrt{\kappa}}\int_0^{M/\sqrt{\kappa}}4\frac{|F^{(n+2)}(v^*)|z^{n+2}(\sqrt{\kappa})^n}{(n+2)!(F''(v^*)z^2)^2}\,dz \quad (4.100) \\
&= \frac{8}{\sqrt{\kappa}}\int_0^{M/\sqrt{\kappa}}\frac{|F^{(n+2)}(v^*)|z^{n-2}\sqrt{\kappa}^n}{(n+2)!F''(v^*)^2}\,dz \quad (4.101) \\
&= \frac{8}{\sqrt{\kappa}}\frac{F^{(n+2)}(v^*)\sqrt{\kappa}^n}{(n+2)!(n-1)F''(v^*)^2}\frac{M^{n-1}}{\sqrt{\kappa}^{n-1}} \quad (4.102) \\
&= 2\frac{F^{(n+2)}(v^*)M^{n-1}}{(n+2)!(n-1)F''(v^*)^2} \quad (4.103)
\end{aligned}$$

Finally, putting it all together we have

$$\begin{aligned}
\left|\frac{1}{\sqrt{\kappa}}\sum_{n=1}^{\infty}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{F^{(n+2)}(v^*)z^{n+2}(\sqrt{\kappa})^n}{(n+2)!(1+\frac{F''(v^*)}{2}z^2)^2}\,dz\right| &\leq \left|\frac{2!F'''(v^*)}{3!F''(v^*)}\log\left(\frac{2\kappa+F''(v^*)(v_{peak}-v^*)^2}{2\kappa+F''(v^*)(v_{reset}-v^*)^2}\right)\right| \\
&+ O(\kappa)+\frac{2}{F''(v^*)^2}\sum_{n=2}^{\infty}\frac{F^{(n+2)}(v^*)M^{n-1}}{(n+2)!(n-1)} \quad (4.104)
\end{aligned}$$

Which, under generic assumptions about the growth of the magnitude of $F^{(n+2)}(v^*)$, we know the last term converges for sufficiently small M. Thus, we have

$$\lim_{\kappa\to 0}\Omega(\kappa)=\lim_{I-I^*(s,\langle w\rangle)\to 0^+}\langle R_i(t)\rangle=0$$

To simplify the asymptotic argument, let

$$K=\frac{2}{F''(v^*)^2}\sum_{n=2}^{\infty}\frac{F^{(n+2)}(v^*)M^{n-1}}{(n+2)!(n-1)}+\left|\frac{2!F'''(v^*)}{3!F''(v^*)}\log\left(\frac{F''(v^*)(v_{peak}-v^*)^2}{F''(v^*)(v_{reset}-v^*)^2}\right)\right|$$

that is K is all the $O(1)$ components in equation (4.104) which is the leading order term. Define our gauge function $\phi_0(\kappa)$ as

$$\phi_0(\kappa)=\left[\frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{dz}{1+\frac{F''(v^*)}{2}z^2}\right]^{-1}=\sqrt{\kappa}\left[\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{dz}{1+\frac{F''(v^*)}{2}z^2}\right]^{-1},$$

then we have

$$\begin{aligned}
|\Omega(\kappa)-\phi_0(\kappa)| &= |\Omega(\kappa)||\phi_0(\kappa)|\left|\frac{1}{\sqrt{\kappa}}\int_{(v_{reset}-v^*)/\sqrt{\kappa}}^{(v_{peak}-v^*)/\sqrt{\kappa}}\frac{\sum_{n=1}^{\infty}\frac{F^{n+2}(v^*)}{(n+2)!}(z\sqrt{\kappa})^n dz}{\left(1+\frac{F''(v^*)}{2}z^2\right)^2}+HOT\right| \\
&\leq |\Omega(\kappa)||\phi_0(\kappa)|\,|K+HOT| \quad (4.105)
\end{aligned}$$

From (4.105), we immediately have

$$\lim_{\kappa\to 0}\frac{|F(\kappa)-\phi_0(\kappa)|}{|\phi_0(\kappa)|}\leq \lim_{\kappa\to 0}|F(\kappa)|(K+HOT)=0 \tag{4.106}$$

Now, all that remains is to analyze the behavior of $\phi_0(\kappa)$ as a function of $I-I^*(s,w)$. Undoing our substitutions yields:

$$\begin{aligned}
\phi_0(\kappa) &= \sqrt{\frac{F''(v^*(gs))(I-I^*(s,w))}{2}}\frac{1}{\arctan\left(z_1(s)\sqrt{\frac{F''(v^*(gs))}{I-I^*(s,w)}}\right)-\arctan\left(z_2(s)\sqrt{\frac{F''(v^*(gs))}{I-I^*(s,w)}}\right)}\\
&= \sqrt{\frac{F''(v^*(gs))(I-I^*(s,w))}{2}}\left(\frac{1}{\pi-\sqrt{\frac{I-I^*(s,w)}{F''(v^*(gs))}}\left(\frac{1}{z_1(s)}-\frac{1}{z_2(s)}\right)+O(I-I^*(s,w))}\right)\\
&= \sqrt{\frac{F''(v^*(gs))(I-I^*(s,w))}{2}}\left(\frac{1}{\pi}-\frac{1}{\pi^2}\sqrt{\frac{I-I^*(s,w)}{F''(v^*(gs))}}\left(\frac{1}{z_1(s)}-\frac{1}{z_2(s)}\right)+O(I-I^*(s,w))\right)
\end{aligned}$$

where $z_1(s)=v_{peak}-v^*(s)$, $z_2(s)=v_{reset}-v^*(s)$. If we reconsider equation (4.90), the dominant term can be immediately seen as

$$\langle R_i(t)\rangle\sim\frac{1}{\pi}\sqrt{\frac{F''(v^*(gs))(I-I^*(s,w))}{2}}\quad\text{as}\quad I-I^*(s,w)\to 0$$

and this concludes the proof that

$$\langle R_i(t)\rangle\sim\frac{1}{\pi}\sqrt{\frac{F''(v^*(gs))}{2}}\sqrt{I-I^*(s,\langle w\rangle)}\quad\text{as}\quad I-I^*(s,\langle w\rangle)\to 0$$

4.3.2 Case II: The Leaky Integrate and Fire Rate When $v_{sn}\notin[v_{reset},v_{peak}]$

Consider again the reciprocal of the firing rate:

$$\frac{1}{\langle R_i(t)\rangle} = \int_{v_{reset}}^{v_{peak}}\frac{dv}{F(v)-\langle w\rangle+gs(er-v)+I}$$

In this case, while the switching manifold is still defined as

$$I-I^*(s,\langle w\rangle)=\min\{F(v)-\langle w\rangle+gs(e_r-v)+I\},$$

the minimum is at a boundary point. In particular, if $v_{sn}>v_{peak}$, Then

$$I-I^*(s,\langle w\rangle)=F(v_{peak})-\langle w\rangle+gs(e_r-v_{peak})+I$$

and if $v_{sn} < v_{reset}$

$$I - I^*(s, \langle w \rangle) = F(v_{reset}) - \langle w \rangle + gs(e_r - v_{reset}) + I$$

Now, we will deal with both cases simultaneously. Letting a correspond to either of the boundary points, then we can write the following:

$$\frac{1}{\langle R_i(t) \rangle} = \int_{v_{reset}}^{v_{peak}} \frac{dv}{I - I^*(s, \langle w \rangle) + F'(a)(v-a) + \sum_{n=2}^{\infty} \frac{F^{(n)}(a)}{n!}(v-a)^n} \quad (4.107)$$

To proceed further, we will again make a substitution. Let $\kappa = I - I^*(s, \langle w \rangle)$ and consider

$$z = \frac{v-a}{\kappa}, \quad \kappa dz = dv$$

In this case we have

$$\frac{1}{\langle R_i(t) \rangle} = \int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{\kappa dz}{\kappa + \kappa F'(a)z + \sum_{n=2}^{\infty} \frac{F^{(n)}(a)}{n!} z^n \kappa^n} \quad (4.108)$$

$$= \int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{dz}{1 + F'(a)z} \frac{1}{1 + \frac{\sum_{n=2}^{\infty} \frac{F^{(n)}(a)}{n!} z^n \kappa^{n-1}}{1 + F'(a)z}} \quad (4.109)$$

As in the previous case, we require the geometric series to proceed further. This implies that we need

$$\left| \sum_{n=2}^{\infty} \frac{F^{(n)}(a)}{n!} z^n \kappa^{n-1} \right| \leq |1 + F'(a)z| \quad (4.110)$$

Which is clearly valid for $z \ll 1$, and for $\zeta M/\kappa < z < M/\kappa$, where $M = |v_{peak} - v_{reset}|$, as $a = v_{reset}$ or $a = v_{peak}$ depending on the sign of $F'(a)$, we have:

$$\frac{1}{\kappa} \sum_{n=2}^{\infty} \frac{|F^{(n)}(a)|}{n!} M^n \leq \left| 1 + \zeta F'(a) \frac{M}{\kappa} \right|$$

which if we consider the dominant $O(\kappa^{-1})$ term in the inequalities, this is approximately

$$\sum_{n=2}^{\infty} \frac{|F^{(n)}(a)|}{n!} M^{n-1} \leq \zeta |F'(a)|$$

and for sufficiently small M, which in this case implies that if v_{peak} and v_{reset} are not too far apart, and the first order derivative dominates in this region, the asymptotics of the firing rate will match the ones derived in this section. Applying our geometric series yields the following:

$$\frac{1}{\langle R_i(t) \rangle} = \int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{dv}{1 + F'(a)z} - \int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{\sum_{n=2}^{\infty} \frac{F^{(n)}(a)}{n!} z^n \kappa^{n-1}}{(1 + F'(a)z)^2} + HOT$$

First note that the first term is divergent, as it is merely given by:

$$\int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{dv}{1+F'(a)z} = \frac{1}{F'(a)} \log(1+F'(a)z)\Big|_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \quad (4.111)$$

$$= \frac{1}{|F'(a)|} \log(1+|F'(a)z|)\Big|_{0}^{(v_{peak}-v_{reset})/\kappa} \quad (4.112)$$

which diverges as $\kappa \to 0$. Thus, we can define

$$1/\phi_0(\kappa) = \frac{1}{|F'(a)|} \log(1+|F'(a)|z)\Big|_{0}^{(v_{peak}-v_{reset})/\kappa}$$

and if the next leading order term is $O(1)$, then as before, we can use $\phi_0(\kappa)$ as a gauge function and argue that $F(\kappa) = \langle R_i(t) \rangle \to \phi_0(\kappa)$. Looking at the magnitude of the next order term:

$$\left| \int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{\sum_{n=2}^{\infty} \frac{F^{(n)}(a)}{n!} z^n \kappa^{n-1}}{(1+F'(a)z)^2} \right| \leq \int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{\sum_{n=2}^{\infty} \frac{|F^{(n)}(a)|}{n!} |z|^n \kappa^{n-1}}{(1+F'(a)z)^2} \quad (4.113)$$

$$= \sum_{n=2}^{\infty} \int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{\frac{|F^{(n)}(a)|}{n!} |z|^n \kappa^{n-1}}{(1+F'(a)z)^2} \quad (4.114)$$

As in the previous case, one can separately show the following:

$$\int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{|F''(a)|z^2\kappa}{2(1+F'(a)z)^2} dz = O(1) \quad (4.115)$$

$$\sum_{n=3}^{\infty} \int_{(v_{reset}-a)/\kappa}^{(v_{peak}-a)/\kappa} \frac{\frac{|F^{(n)}(a)|}{n!} |z|^n \kappa^{n-1}}{(1+F'(a)z)^2} = O(1) \quad (4.116)$$

Which, as before immediately yields

$$F(\kappa) = \left(\int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v) - \langle w \rangle + gs(e_r - v) + I} \right)^{-1} \sim |F'(a)| \left(\log \left(1 + \frac{|F'(a)|(v_{peak} - v_{reset})}{I - I^*(s, \langle w \rangle)} \right) \right)^{-1}$$

Note that this is merely the firing rate of a leaky integrate and fire neuron. While this is an important case to consider, we will primarily focus on Case I for the rest of this paper. The reason being is that this case is primarily valid for the leaky integrate and fire neuron by default, or when one places the uncoupled threshold for firing, $v^*(0)$ in biologically unrealistic places (either above or below the voltage reset).

Example

As a quick example, we will formally prove the first term in the asymptotic sequence of the function

$$\langle R \rangle = \left(\int_{-1}^{B} \frac{dv}{\kappa + v^2 + Av^3} \right)^{-1} \quad (4.117)$$

where A and B will be bounded to ensure the sequence is valid. Suppose that $B > 0$, then $v^* = 0 \in [-1, B]$

$$\begin{aligned}
\langle R\rangle = F(\kappa) &= \sqrt{\kappa}\left(\int_{-1/\sqrt{\kappa}}^{B/\sqrt{\kappa}} \frac{dv}{1+v^2+\sqrt{\kappa}Av^3}\right)^{-1} \\
\phi_0(\kappa) &= \sqrt{\kappa}\left(\int_{-1/\sqrt{\kappa}}^{B/\sqrt{\kappa}} \frac{dv}{1+v^2}\right)^{-1} \\
|F(\kappa)-\phi_0(\kappa)| &= \langle R\rangle\phi_0(\kappa)\left|\int_{-1/\sqrt{\kappa}}^{B/\sqrt{\kappa}} \frac{Av^3dv}{(1+v^2)^2} + \ldots + \int_{-1/\sqrt{\kappa}}^{B/\sqrt{\kappa}} \frac{(Av^3)^n(\sqrt{\kappa})^{n-1}}{(1+v^2)^{n+1}}(-1)^n + \ldots\right| \\
&\leq F(\kappa)\phi_0(\kappa)\left(\left|\int_{-1/\sqrt{\kappa}}^{B/\sqrt{\kappa}} \frac{Av^3dv}{(1+v^2)^2}\right| + \sum_{n=2}^{\infty}\int_{-1/\sqrt{\kappa}}^{B/\sqrt{\kappa}} \frac{|A|^n|v|^{3n}\sqrt{\kappa}^{n-1}}{(1+v^2)^{n+1}}\,dv\right) \\
&= F(\kappa)\phi_0(\kappa)\left(\frac{|A|}{2}\left|\log\left(\frac{\kappa+B^2}{\kappa+1}\right) + \frac{\kappa}{\kappa+B^2} - \frac{\kappa}{1+\kappa}\right| + \sum_{n=2}^{\infty}\int_0^{M/\sqrt{\kappa}} \frac{|A|^n v^{3n}\sqrt{\kappa}^{n-1}}{(1+v^2)^{n+1}}\,dv\right) \\
&\leq F(\kappa)\phi_0(\kappa)\left(\frac{|A|}{2}\left|\log\left(\frac{\kappa+B^2}{\kappa+1}\right) + \frac{\kappa}{\kappa+B^2} - \frac{\kappa}{1+\kappa}\right| + \sum_{n=2}^{\infty}\int_0^{M/\sqrt{\kappa}} |A|^n v^{n-2}\sqrt{\kappa}^{n-1}\,dv\right) \\
&= F(\kappa)\phi_0(\kappa)\left(\frac{|A|}{2}\left|\log\left(\frac{\kappa+B^2}{\kappa+1}\right) + \frac{\kappa}{\kappa+B^2} - \frac{\kappa}{1+\kappa}\right| + \sum_{n=2}^{\infty} \frac{|A|^n M^{n-1}}{n-1}\right) \\
&= F(\kappa)\phi_0(\kappa)\left(\frac{|A|}{2}\left|\log\left(\frac{\kappa+B^2}{\kappa+1}\right) + \frac{\kappa}{\kappa+B^2} - \frac{\kappa}{1+\kappa}\right| + |A|^2M\sum_{n=0}^{\infty} \frac{|A|^n M^n}{n+1}\right) \\
&= F(\kappa)\phi_0(\kappa)\left(\frac{|A|}{2}\left|\log\left(\frac{\kappa+B^2}{\kappa+1}\right) + \frac{\kappa}{\kappa+B^2} - \frac{\kappa}{1+\kappa}\right| + \frac{|A|}{1-|A|M}\right)
\end{aligned}$$

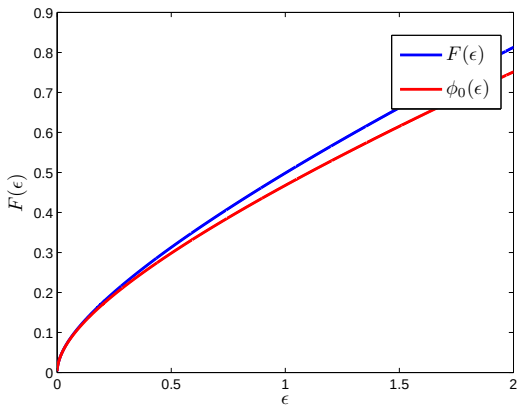
Thus, if $|A||M| < 1$, then the right hand side is bounded. Which immediately implies that $F(\kappa) \sim \phi_0(\kappa)$ as $\kappa \to 0$. Additionally, in order to have used the geometric series, we required that

$$|A|v^3\sqrt{\kappa} < 1 + v^2$$

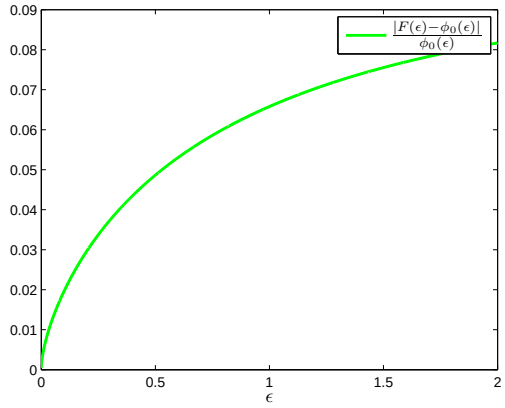
clearly this inequality is satisfied for $v \ll 1$, but it is also valid for $v = O(1/\sqrt{\kappa})$, in particular if

$$|A|\frac{M^3}{\kappa} < 1 + \frac{M^2}{\kappa}$$

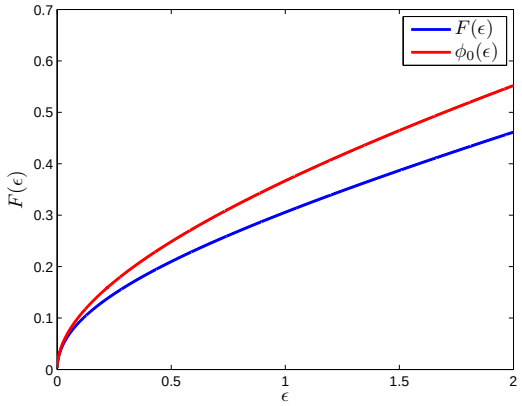
which holds for small κ and $|A|M < 1$



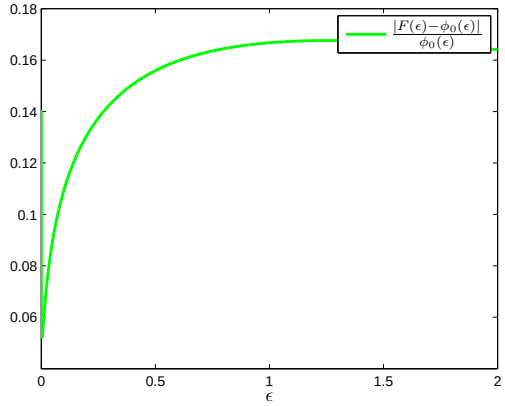
(a) $M = 4.5$, $A = 0.2$



(b) $\frac{|F(\kappa)-\phi_0(\kappa)|}{\phi_0(\kappa)} \to 0$ as $\kappa \to 0$



(c) $M = 10$, $A = 0.2$



(d) $\frac{|F(\kappa)-\phi_0(\kappa)|}{\phi_0(\kappa)} \neq 0$ as $\kappa \to 0$

Figure 4.7: The asymptotic expansion of $F(\kappa)$ has a first term of $\phi_0(\kappa)$ as $\kappa \to 0$ provided that the conditions on A and M are met.

4.4 The Full Spectrum of Eigenvalues

In Chapter 6, we will analyze the eigenvalues of the slow system via bifurcation analysis. Here however, we derive the spectral equation for equations (4.18)-(4.20). The derivation largely follows the one performed in [2], only instead of the weak-coupling limit analyzed in [2], we will analyze the mean-field limit as $\epsilon = \tau_w^{-1} \to 0$.

To simplify matters somewhat, we will assume that $b = 0$. The analysis can be reproduced however without this assumption. Recalling the population density equations for the marginal density and the mean coupling and adaptation variables:

$$\frac{\partial}{\partial t}\rho_V(v,t) = -\frac{\partial}{\partial v}\left[\left(F(v) - \langle w\rangle + gs(e_r - v) + I\right)\rho_V(v,t)\right] \tag{4.118}$$

$$s' = -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s}J(v_{peak}, \langle w\rangle, s, t) \tag{4.119}$$

$$\langle w\rangle' = -\frac{\langle w\rangle}{\tau_w} + \frac{\lambda_w}{\tau_w}J(v_{peak}, \langle w\rangle, s, t) \tag{4.120}$$

We will proceed by first demonstrating the existence of steady state solutions in different asymptotic regimes. The steady states we are considering are specifically asynchronous firing. Then, we will linearize the resulting operator around the steady states via the transformation from [2]. Then, we will determine the spectral equation. Note that we can interpret right hand side of the equations (4.118)-(4.120) as an operator working on the space $L_2([v_{reset}, v_{peak}]) \times \mathbb{R}^2$.

4.4.1 Existence of Steady State Solutions

The steady states are given by the following set of equations:

$$\rho_V(v) = \frac{\overline{\langle R\rangle}}{F(v) - \overline{\langle w\rangle} + g\bar{s}(e_r - v) + I} \tag{4.121}$$

$$\bar{s} = \lambda_s\overline{\langle R\rangle} \tag{4.122}$$

$$\overline{\langle w\rangle} = \lambda_w\overline{\langle R\rangle} \tag{4.123}$$

$$\frac{1}{\overline{\langle R\rangle}} = \int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v) - \overline{\langle w\rangle} + g\bar{s}(e_r - v) + I} \tag{4.124}$$

Now, clearly, the existence of a solution to (4.121)-(4.124) will depend on the existence of a solution to the algebraic equation

$$\frac{1}{\overline{\langle R\rangle}} = \int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v) - \lambda_w\overline{\langle R\rangle} + g\lambda_s\overline{\langle R\rangle}(e_r - v) + I} \tag{4.125}$$

or defining $x = \overline{\langle R\rangle}$, we get

$$x = \left(\int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v) + I - x\lambda_w + gx\lambda_s(e_r - v) + I}\right)^{-1} = \phi(x) \tag{4.126}$$

As we require a solution to $\phi(x) = x$, we should naturally proceed by the contraction mapping theorem/Banach fixed point theorem. We will consider two cases to ascertain the existence of an equilibria:

Case I: $I - I^*(0,0) > 0$

First, assume that $I - I^*(0,0) > 0$ and we restrict the analysis to $x \geq 0$. In this case we can immediately use the mean value theorem to yield:

$$\begin{aligned} |\phi(x) - \phi(y)| &= \left| \phi(c)^2 \int_{v_{reset}}^{v_{peak}} \frac{(g\lambda_s(e_r - v) - \lambda_w)dv}{(F(v) + I + g\lambda_s c(e_r - v) - \lambda_w c)^2} \right| |x - y| \\ &\leq |g\lambda_s N + \lambda_w| \phi(c)^2 \int_{v_{reset}}^{v_{peak}} \frac{dv}{(F(v) + I + g\lambda_s c(e_r - v) - \lambda_w c)^2} |x - y| \end{aligned}$$

where $N = \max\{|e_r - v_{peak}|, |e_r - v_{reset}|\}$, Now, using the Cauchy-Schwarz theorem and the integral definition of $\phi(c)$, we have

$$K(c) = \phi(c)^2 \int_{v_{reset}}^{v_{peak}} \frac{dv}{(F(v) + I + g\lambda_s c(e_r - v) - \lambda_w c)^2} > 1$$

for all $c \geq 0$, and, $K(0)$ is defined as we have assumed that $I - I^*(0,0) > 0$ (this implies that all the integrals involved in evaluating $K(c)$ are defined). This immediately implies that $K(c)$ is a continuous function for $c \geq 0$, and thus if we consider some interval $c \in [0, M]$ that is compact, then $K(c)$ is bounded on this interval by some value K^*. Thus, we have

$$|\phi(x) - \phi(y)| \leq K^*(g\lambda_s N + \lambda_w)|x - y|$$

thus, if

$$(g\lambda_s N + \lambda_w) \leq K^{*-1}$$

$\phi(x)$ is a contraction mapping and there exists a unique steady state solution to (5.126)-(4.124).

Case II: $I - I^*(0,0) \to 0$

Now, assume that $I - I^*(0,0) = 0$. In this case we can use the leading order asymptotics derived in section 4.3. There are two possible cases: the Type-I and LIF leading order asymptotics. However we will restrict ourselves to the Type-I asymptotics. For the Type-I case, the asymptotics yield:

$$x \sim \frac{1}{\pi} \sqrt{\frac{F''(v^*(g\lambda_s x))}{2}} \sqrt{I - I^*(\lambda_s x, \lambda_w x)} \quad \text{as} \quad I - I^*(\lambda_s x, \lambda_w x) \to 0$$

which implies that

$$x^2 \sim \frac{F''(v^*(g\lambda_s x))\pi^2}{2} (I + F(v^*(g\lambda_s x)) + g\lambda_s(e_r - v^*(g\lambda_s x)) - \lambda_w x) \quad \text{as} \quad I - I^*(\lambda_s x, \lambda_w x) \to 0$$

Now, we solve the $O(I - I^*)$ problem:

$$\frac{\pi^2}{2}x^2F''(v^*(g\lambda_s x)) = I + F(v^*(g\lambda_s x)) + g\lambda_s(er - v^*(g\lambda_s x)) - \lambda_w x$$

which in order to proceed, one has to Taylor expand in x which yields:

$$\begin{aligned}
\frac{\pi}{2}x^2F''(v^*(0) + v^{*\prime}(0)g\lambda_s x + O((g\lambda_s x)^2)) &= I + F(v^*(0) + v^{*\prime}(0)(g\lambda_s x) + O((g\lambda_s x)^2)) \\
\frac{\pi}{2}x^2F''(v^*(0)) + O(g\lambda_s x^3) &= I + F(v^*(0)) + F'(v^*(0))(v^{*\prime}(0)g\lambda_s x + O(g\lambda_s x)) \\
&\quad + \frac{F''(v^*(0))}{2}(v^{*\prime}(0)g\lambda_s x + O(g\lambda_s x))^2 \\
&\quad + g\lambda_s x(e_r - v^*(0) - v^{*\prime}(0)(g\lambda_s x) + O((g\lambda_s x))) \\
&= I + F(v^*(0)) + g\lambda_s x(er - v^*(0)) - \frac{1}{2}(g\lambda_s x)^2 v^{*\prime}(0) - \lambda_w x
\end{aligned}$$

where we have used the definition of $v^*(z)$, being the solution to the equation $F'(v^*(z)) = z$ to simplify the first few terms on the left hand side. The resulting lowest order polynomial is:

$$x^2\left(\frac{\pi}{2}F''(v^*(0)) + \frac{1}{2}g^2\lambda_s^2\right) + x\left(\lambda_w - g\lambda_s(e_r - v^*(0))\right) - (I + F(v^*(0))) = 0 \quad (4.127)$$

This quadratic can easily be shown to have up to two solutions in different regions of the (I, g) parameter plane. We will elaborate more on this in our chapter on the non-smooth bifurcation analysis of the resulting mean-field system.

4.4.2 Linearizing the Operator

To proceed further, we will need to linearize the operator. If we integrate the population density equation once in v we obtain a partial differential equation for the distribution function:

$$\frac{\partial P(v,t)}{\partial t} = \frac{\partial}{\partial t}\left(\int_{v_{reset}}^{v} \rho_V(v,t)\right) \quad (4.128)$$

$$\frac{\partial P(v,t)}{\partial t} = -G_1(v, s, \langle w\rangle)\frac{\partial P(v,t)}{\partial v} + G_1(v_{peak}, s, \langle w\rangle)\frac{\partial P(v,t)}{\partial v}\bigg|_{v=v_{peak}} \quad (4.129)$$

where we have used the boundary condition (4.9) to simplify (4.129). Note that this is different from the approach used in [2], however should yield the same eigenvalues at the end. By considering the distribution function, one can consider just perturbations of the steady state distribution function, as opposed to being forced to relate perturbations in the density function to perturbations in the flux, as in [2].

To simplify the steady state, we will use the Abbott/Vreeswijk transform from [2]:

$$y = \int_{v_{reset}}^{v} \frac{\overline{\langle R \rangle}}{G_1(v', \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \, dv' = \eta(v) \tag{4.130}$$

This transformation is invertible provided that $G_1(v, \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle}) > 0$ for all $v \in [v_{reset}, v_{peak}]$. This will be true if we are restricted to the tonic firing region of the parameter space. Applying the transformation to (4.129),

$$\begin{aligned}
\frac{\partial P(y,t)}{\partial t} &= -\frac{\overline{\langle R \rangle} G_1(\eta^{-1}(y), s, \langle w \rangle)}{G_1(\eta^{-1}(y), \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \frac{\partial P(y,t)}{\partial y} + \frac{\overline{\langle R \rangle} G_1(v_{peak}, s, \langle w \rangle)}{G_1(v_{peak}, \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \frac{\partial P(y,t)}{\partial y}\bigg|_{y=1} \\
s' &= -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s} \frac{\overline{\langle R \rangle} G_1(v_{peak}, s, \langle w \rangle)}{G_1(v_{peak}, \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \frac{\partial P(y,t)}{\partial y}\bigg|_{y=1} \\
w' &= -\frac{w}{\tau_w} + \frac{\lambda_w}{\tau_w} \frac{\overline{\langle R \rangle} G_1(v_{peak}, s, \langle w \rangle)}{G_1(v_{peak}, \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \frac{\partial P(y,t)}{\partial y}\bigg|_{y=1}
\end{aligned}$$

Now, note that the steady state distribution function after performing the Abbott/Vreeswijk transform is merely $P(y) = y$. Expanding about this steady state $P(y,t) = y + \epsilon_y(y,t)$, $s(t) = \epsilon_s(t) + \lambda_s \overline{\langle R \rangle}$, $\langle w \rangle = \epsilon_w(t) + \lambda_w \overline{\langle R \rangle}$. We immediately require $\epsilon(1,t) = \epsilon(0,t) = 0$ for all $t > 0$. Using this expansion in the system above gives

$$\begin{aligned}
\frac{\partial \epsilon_y(y,t)}{\partial t} &= -\overline{\langle R \rangle} \frac{G_1(\eta^{-1}(y), \lambda_s \overline{\langle R \rangle} + \epsilon_s(t), \lambda_w \overline{\langle R \rangle} + \epsilon_w(t))}{G_1(\eta^{-1}(y), \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \left(\frac{\partial \epsilon_y(y,t)}{\partial y} + 1 \right) \\
&+ \frac{\overline{\langle R \rangle} G_1(v_{peak}, \lambda_s \overline{\langle R \rangle} + \epsilon_s(t), \epsilon_w(t) + \lambda_w \overline{\langle R \rangle})}{G_1(v_{peak}, \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \left(\frac{\partial \epsilon_y(y,t)}{\partial y} + 1 \right)\bigg|_{y=1} \\
&= L_1(\epsilon_y(y,t), \epsilon_s(t), \epsilon_w(t)) \\
\epsilon_s'(t) &= -\frac{\epsilon_s(t)}{\tau_s} + \frac{\overline{\langle R \rangle} \lambda_s}{\tau_s} \left(\frac{g \epsilon_s(t)(e_r - v_{peak}) - \epsilon_w(t)}{G_1(v_{peak}, \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \right) \left(1 + \frac{\partial \epsilon_y(y,t)}{\partial y}\bigg|_{y=1} \right) \\
&+ \frac{\overline{\langle R \rangle} \lambda_s}{\tau_s} \frac{\partial \epsilon_y(y,t)}{\partial y}\bigg|_{y=1} = L_2(\epsilon_y(y,t), \epsilon_s(t), \epsilon_w(t)) \\
\epsilon_w'(t) &= -\frac{\epsilon_w(t)}{\tau_w} + \frac{\overline{\langle R \rangle} \lambda_w}{\tau_w} \left(\frac{g \epsilon_s(t)(e_r - v_{peak}) - \epsilon_w(t)}{G_1(v_{peak}, \lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})} \right) \left(1 + \frac{\partial \epsilon_y(y,t)}{\partial y}\bigg|_{y=1} \right) \\
&+ \frac{\overline{\langle R \rangle} \lambda_w}{\tau_w} \frac{\partial \epsilon_y(y,t)}{\partial y}\bigg|_{y=1} = L_3(\epsilon_y(y,t), \epsilon_s(t), \epsilon_w(t))
\end{aligned}$$

Now, we are going to analyze the operator

$$L(\epsilon_y(y,t), \epsilon_s(t), \epsilon_w(t)) = \begin{pmatrix} L_1(\epsilon_y(y,t), \epsilon_s(t), \epsilon_w(t)) \\ L_2(\epsilon_y(y,t), \epsilon_s(t), \epsilon_w(t)) \\ L_3(\epsilon_y(y,t), \epsilon_s(t), \epsilon_w(t)) \end{pmatrix}$$

In order to analyze this operator and its eigenvalue spectrum, we will first need to 1) define the Banach space on which this operator works on and 2) determine its Fréchet derivative. As we are dealing with the direct product of three Banach spaces, one can use the sum of their norms to define a norm. In particular, referring to this space as $\Omega = \mathbb{L}_2[0,1] \times \mathbb{R}^2$, we have the norm:

$$\begin{aligned} \|\boldsymbol{z}\|_\Omega &= \|\epsilon_y(y), \epsilon_s, \epsilon_w\| = \|\epsilon_y(y)\|_{L_2} + |\epsilon_s| + |\epsilon_w| && (4.131) \\ &= \left(\int_0^1 \epsilon(y',t)^2\, dy'\right)^{1/2} + |\epsilon_s(t)| + |\epsilon_w(t)| && (4.132) \end{aligned}$$

where we have used the L_1 norm on $\mathbb{R}$ and the L_2 norm on $\mathbb{L}_2[0,1]$ space. We will now proceed to find the Fréchet derivative of our linear operator. This implies resolving the limit:

$$\lim_{h\to 0} \frac{\|L(\boldsymbol{z}+\boldsymbol{h}) - L(\boldsymbol{z}) - DL(\boldsymbol{z})\boldsymbol{h}\|_\Omega}{\|h\|_\Omega} \tag{4.133}$$

where $\boldsymbol{h} = (h_1(y,t), h_2(t), h_3(t))$. Looking at the terms of the of the numerator independently:

$$\begin{aligned} L_1(\mathbf{0}+\boldsymbol{h}) &= -\overline{\langle R\rangle}\frac{G_1(\eta^{-1}(y), \lambda_s\overline{\langle R\rangle} + h_2, \lambda_w\overline{\langle R\rangle} + h_3)}{G_1(\eta^{-1}(y), \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\left(\frac{\partial h_1(y)}{\partial y} + 1\right) \\ &+ \overline{\langle R\rangle}\frac{G_1(v_{peak}, \lambda_s\overline{\langle R\rangle} + h_2, \lambda_w\overline{\langle R\rangle} + h_3)}{G_1(v_{peak}, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\left(\frac{\partial h_1(y)}{\partial y} + 1\right)\Bigg|_{y=1} \\ &= -\overline{\langle R\rangle}\left(\frac{\partial h_1(y)}{\partial y} - \frac{\partial h_1(y)}{\partial y}\Big|_{y=1}\right) \\ &- \overline{\langle R\rangle}\left(\frac{gh_2(e_r - \eta^{-1}(y)) - h_3}{G(\eta^{-1}(y), \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})} - \frac{gh_2(e_r - v_{peak}) - h_3}{G_1(v_{peak}, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\right) \\ &- \overline{\langle R\rangle}\frac{\partial h_1(y)}{\partial y}\left(\frac{gh_2(e_r - \eta^{-1}(y)) - h_3}{G(\eta^{-1}(y), \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\right) \\ &+ \overline{\langle R\rangle}\frac{\partial h_1(y)}{\partial y}\Big|_{y=1}\left(\frac{gh_2(e_r - v_{peak}) - h_3}{G_1(v_{peak}, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\right) \\ L_2(\mathbf{0}+\boldsymbol{h}) &= -\frac{h_2}{\tau_s} + \frac{\overline{\langle R\rangle}\lambda_s}{\tau_s}\left(\frac{gh_2(e_r - v_{peak}) - h_3}{G_1(v_{peak}, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\right)\left(1 + \frac{\partial h_1(y)}{\partial y}|_{y=1}\right) \\ &+ \frac{\overline{\langle R\rangle}\lambda_s}{\tau_s}\frac{\partial h_1(y)}{\partial y}\Big|_{y=1} \\ L_3(\mathbf{0}+\boldsymbol{h}) &= -\frac{h_3}{\tau_w} + \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\left(\frac{gh_2(e_r - v_{peak}) - h_3}{G_1(v_{peak}, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\right)\left(1 + \frac{\partial h_1(y)}{\partial y}\Big|_{y=1}\right) \\ &+ \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\frac{\partial h_1(y)}{\partial y}\Big|_{y=1} \end{aligned}$$

One can immediately ascertain a candidate for the linear operator $DL(0)$:

$$\begin{aligned}
DL_1(0)\boldsymbol{h} &= -\overline{\langle R\rangle}\left(\frac{\partial h_1(y)}{\partial y} - \frac{\partial h_1(y)}{\partial y}\bigg|_{y=1}\right) \\
&- \overline{\langle R\rangle}\left(\frac{gh_2(e_r-\eta^{-1}(y))-h_3}{G_1(\eta^{-1}(y),\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})} - \frac{gh_2(e_r-v_{peak})-h_3}{G_1(v_{peak},\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})}\right) \\
DL_2(0)\boldsymbol{h} &= -\frac{h_2}{\tau_s} + \frac{\overline{\langle R\rangle}\lambda_s}{\tau_s}\left(\frac{gh_2(e_r-v_{peak})-h_3}{G_1(v_{peak},\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})}\right) + \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\frac{\partial h_1(y)}{\partial y}\bigg|_{y=1} \\
DL_3(0)\boldsymbol{h} &= -\frac{h_3}{\tau_w} + \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\left(\frac{gh_2(e_r-v_{peak})-h_3}{G_1(v_{peak},\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})}\right) + \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\frac{\partial h_1(y)}{\partial y}\bigg|_{y=1}
\end{aligned}$$

as these are the only linear terms in the operator $L(\boldsymbol{h})$. We will forgo the formal proof that our candidate $DL(\mathbf{0})$ is in fact the linerization of L at $\mathbf{0}$. Now, we can proceed in looking for the spectrum of eigenvalues for this linear operator. In particular, we require that

$$D_L(\mathbf{0})\boldsymbol{h} = \mu\boldsymbol{h}$$

Which yields the following equations:

$$\begin{aligned}
\mu h_1(y) &= -\overline{\langle R\rangle}\left(\frac{\partial h_1(y)}{\partial y} - \frac{\partial h_1(y)}{\partial y}\bigg|_{y=1}\right) \\
&- \overline{\langle R\rangle}\left(\frac{gh_2(e_r-\eta^{-1}(y))-h_3}{G_1(\eta^{-1}(y),\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})} - \frac{gh_2(e_r-v_{peak})-h_3}{G_1(v_{peak},\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})}\right) \\
\mu h_2 &= -\frac{h_2}{\tau_s} + \frac{\overline{\langle R\rangle}\lambda_s}{\tau_s}\left(\frac{gh_2(e_r-v_{peak})-h_3}{G_1(v_{peak},\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})}\right) + \frac{\overline{\langle R\rangle}\lambda_s}{\tau_s}\frac{\partial h_1(y)}{\partial y}\bigg|_{y=1} \\
\mu h_3 &= -\frac{h_3}{\tau_w} + \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\left(\frac{gh_2(e_r-v_{peak})-h_3}{G_1(v_{peak},\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})}\right) + \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\frac{\partial h_1(y)}{\partial y}\bigg|_{y=1}
\end{aligned}$$

To simplify the notation somewhat, we will write down the following:

$$A(y) = \frac{g(e_r-\eta^{-1}(y))}{G_1(\eta^{-1}(y),\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})} \tag{4.134}$$

$$B(y) = \frac{-1}{G_1(\eta^{-1}(y),\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})} \tag{4.135}$$

which immediately lets us write the eigenvalue problem in the more compact notation:

$$\begin{aligned}
\mu h_1(y) &= -\overline{\langle R\rangle}\left(\frac{\partial h_1(y)}{\partial y} - \frac{\partial h_1(y)}{\partial y}\bigg|_{y=1}\right) - \overline{\langle R\rangle}\left[h_2(A(y)-A(1)) + h_3(B(y)-B(1))\right] \\
\mu h_2 &= -\frac{h_2}{\tau_s} + \frac{\overline{\langle R\rangle}\lambda_s}{\tau_s}\left(h_2A(1)+h_3B(1)\right) + \frac{\overline{\langle R\rangle}\lambda_s}{\tau_s}\frac{\partial h_1(y)}{\partial y}\bigg|_{y=1} \\
\mu h_3 &= -\frac{h_3}{\tau_w} + \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\left(h_2A(1)+h_3B(1)\right) + \frac{\overline{\langle R\rangle}\lambda_w}{\tau_w}\frac{\partial h_1(y)}{\partial y}\bigg|_{y=1}
\end{aligned}$$

Now, in order to solve this particular eigenvalue equation, we have to force the conditions $h_1(0) = h_1(1) = 0$, and we want non-trivial eigenspaces. That is, we want to specifically avoid the eigenspace $h_1(y) = 0$, $h_2 = h_3 = 0$. The former condition will allow us to solve for $h_1(y)$, and the latter condition will finally yield the spectral equation for the eigenvalues. Solving for $h_1(y)$ yields:

$$\begin{aligned} h_1(y)e^{y\mu/\overline{\langle R\rangle}} &= \frac{\overline{\langle R\rangle}}{\mu}(e^{y\mu/\overline{\langle R\rangle}} - 1)(h_1'(1) + h_2A(1) + h_3B(1)) && (4.136)\\ &- \int_0^y e^{y'\mu/\overline{\langle R\rangle}}(h_2A(y') + h_3B(y'))\,dy' + C && (4.137)\end{aligned}$$

where $C = h_1(0) = 0$. The condition that $h_1(0) = 0$ immediately gives us a solution for $h_1'(1)$:

$$h_1'(1) = \frac{\mu}{\overline{\langle R\rangle}}(e^{\mu/\overline{\langle R\rangle}} - 1)^{-1}\int_0^1 e^{y'\mu/\overline{\langle R\rangle}}(h_2A(y') + h_3B(y'))\,dy' - h_2A(1) - h_3B(1)$$

This allows us to solve for the eigenvectors $h_1(y)$:

$$\begin{aligned} h_1(y) &= \left(\frac{e^{y\mu/\overline{\langle R\rangle}} - 1}{e^{\mu/\overline{\langle R\rangle}} - 1}\right)\int_0^1 e^{(y'-y)\mu/\overline{\langle R\rangle}}(h_2A(y') + h_3B(y'))\,dy'\\ &- \int_0^y e^{(y'-y)\mu/\overline{\langle R\rangle}}(h_2A(y') + h_3B(y'))\,dy'\end{aligned}$$

in addition to simplifying the equations for h_2 and h_3:

$$\begin{aligned} \mu h_2 &= -\frac{h_2}{\tau_s} + \frac{\mu\lambda_s}{\tau_s(e^{\mu/\overline{\langle R\rangle}} - 1)}\left(\int_0^1 e^{y'\mu/\overline{\langle R\rangle}}(h_2A(y') + h_3B(y'))\,dy'\right) && (4.138)\\ \mu h_3 &= -\frac{h_3}{\tau_w} + \frac{\mu\lambda_w}{\tau_w(e^{\mu/\overline{\langle R\rangle}} - 1)}\left(\int_0^1 e^{y'\mu/\overline{\langle R\rangle}}(h_2A(y') + h_3B(y'))\,dy'\right). && (4.139)\end{aligned}$$

To simplify the notation, we will use the following:

$$\begin{aligned} \hat{A}(\mu) &= \int_0^1 e^{y'\mu/\overline{\langle R\rangle}}A(y')\,dy' && (4.140)\\ \hat{B}(\mu) &= \int_0^1 e^{y'\mu/\overline{\langle R\rangle}}B(y')\,dy' && (4.141)\end{aligned}$$

$$\begin{aligned} h_2\left(\mu + \frac{1}{\tau_s} - \frac{\mu\lambda_s}{\tau_s(e^{\mu/\overline{\langle R\rangle}} - 1)}\hat{A}(\mu)\right) + h_3\left(\frac{-\mu\lambda_s}{\tau_s(e^{\mu/\overline{\langle R\rangle}} - 1)}\hat{B}(\mu)\right) &= 0\\ h_2\left(-\frac{\mu\lambda_w}{\tau_w(e^{\mu/\overline{\langle R\rangle}} - 1)}\hat{A}(\mu)\right) + h_3\left(\mu + \frac{1}{\tau_w} - \frac{\mu\lambda_w}{\tau_w(e^{\mu/\overline{\langle R\rangle}} - 1)}\hat{B}(\mu)\right) &= 0\\ \boldsymbol{M}(\mu)\begin{pmatrix}h_2\\ h_3\end{pmatrix} &= 0\end{aligned}$$

Based on the fact we want non trivial eigenvectors for this system, we require $\det \boldsymbol{M}(\mu) = 0$, which yields:

$$\left(e^{\mu/\overline{\langle R\rangle}} - 1\right)\left(\mu + \frac{1}{\tau_s}\right)\left(\mu + \frac{1}{\tau_w}\right) - \left(\mu + \frac{1}{\tau_s}\right)\left(\frac{\lambda_w}{\tau_w}\mu\hat{B}(\mu)\right) - \left(\mu + \frac{1}{\tau_w}\right)\left(\frac{\lambda_s}{\tau_s}\mu\hat{A}(\mu)\right) = 0 \tag{4.142}$$

Equation (4.142) is the spectral equation which determines the eigenvalues of our linear operator, and thus the stability of the solution to equations (5.126)-(4.124). The transcendental equation (4.142) is notoriously difficult to solve. However, there are some perturbative approaches one can look at to try to glean some information from equation (4.142). This will be the approach taken in the next section.

4.4.3 The Mean-Field Limit

Previously, we derived the mean-field system of differential equations via a perturbation argument. We assumed that $\tau_w^{-1} = \epsilon$ was small, and $\tau_s = \tau_w\gamma$ where $\gamma = O(1)$. Applying the same assumption here yields spectral equation:

$$\left(e^{\mu/\overline{\langle R\rangle}} - 1\right)\left(\mu + \frac{\epsilon}{\gamma}\right)(\mu + \epsilon) - \left(\mu + \frac{\epsilon}{\gamma}\right)\epsilon\lambda_w\mu\hat{B}(\mu) - \frac{1}{\gamma}(\mu + \epsilon)\lambda_s\epsilon\mu\hat{A}(\mu) = 0$$

Now, we will look for perturbation solutions to this system, $\mu = \mu_0 + \epsilon\mu_1$. The $O(1)$ problem is immediately resolvable:

$$\mu_0^2(\exp(\mu_0/\overline{\langle R\rangle}) - 1) = 0$$

which implies that $\mu_0 = 0$ with multiplicity 2, or $\mu_0 = 2n\pi i\overline{\langle R\rangle}$. Ignoring the infinite set of eigenvalues for now, let us consider $\mu = \epsilon\mu_1$. The resulting system is

$$\epsilon^3\left(\frac{\mu_1}{\overline{\langle R\rangle}} + O(\epsilon)\right)\left(\mu_1 + \frac{1}{\gamma}\right)(\mu_1 + 1) - \epsilon^3\left(\mu_1 + \frac{1}{\gamma}\right)\mu_1\lambda_w\hat{B}(\mu_1\epsilon) - \epsilon^3\frac{1}{\gamma}(\mu_1 + 1)\lambda_s\mu_1\hat{A}(\mu_1\epsilon)$$

Now, we need to determine the leading order behavior of

$$\hat{A}(\mu_1\epsilon) = \int_0^1 e^{\epsilon y'\mu_1/\overline{\langle R\rangle}}A(y')\,dy = \int_0^1 A(y')\,dy' + O(\epsilon) \tag{4.143}$$

$$\hat{B}(\mu_1\epsilon) = \int_0^1 e^{\epsilon y'\mu_1/\overline{\langle R\rangle}}B(y')\,dy = \int_0^1 B(y')\,dy' + O(\epsilon) \tag{4.144}$$

Undoing all the substitutions we had before yields the following:

$$\int_0^1 A(y')\,dy' = \int_0^1 \frac{g(e_r - \eta^{-1}(y'))}{G_1(\eta^{-1}(y')\lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\,dy' \tag{4.145}$$

$$= \overline{\langle R\rangle}\int_{v_{reset}}^{v_{peak}} \frac{g(e_r - v)}{G_1(v, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})^2}\,dv \tag{4.146}$$

$$\int_0^1 B(y')\,dy' = \int_0^1 \frac{-1}{G_1(\eta^{-1}(y'), \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}\,dy' \tag{4.147}$$

$$= \overline{\langle R\rangle}\int_{v_{reset}}^{v_{peak}} \frac{-1}{G_1(v, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})^2}\,dv' \tag{4.148}$$

Noting that

$$\langle R_i(t)\rangle(s, \langle w\rangle) = \left(\int_{v_{reset}}^{v_{peak}} \frac{dv}{G_1(v, s, \langle w\rangle)}\right)^{-1} \tag{4.149}$$

it should be clear that we have the following:

$$\int_0^1 A(y')\,dy' = \frac{1}{\overline{\langle R\rangle}}\frac{\partial\langle R_i(t)\rangle(s, \langle w\rangle)}{\partial s}\bigg|_{(\lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})} \tag{4.150}$$

$$\int_0^1 B(y')\,dy' = \frac{1}{\overline{\langle R\rangle}}\frac{\partial\langle R_i(t)\rangle(s, \langle w\rangle)}{\partial w}\bigg|_{(\lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})} \tag{4.151}$$

Collecting all the $O(\epsilon^3)$ terms we have

$$0 = \frac{\mu_1}{\overline{\langle R\rangle}}(\mu_1 + \gamma^{-1})(\mu_1 + 1) - \frac{\mu_1}{\overline{\langle R\rangle}}\lambda_w(\mu_1 + \gamma^{-1})\frac{\partial\langle R_i(t)\rangle}{\partial\langle w\rangle}\bigg|_{(\lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}$$
$$- \gamma^{-1}\lambda_s\frac{\mu_1}{\overline{\langle R\rangle}}(\mu_1 + 1)\frac{\partial\langle R_i(t)\rangle}{\partial s}\bigg|_{(\lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})}.$$

After we factor out $\mu_1/\overline{\langle R\rangle}$

$$\left(\mu_1 + \frac{1}{\gamma}\right)(\mu_1 + 1) - \lambda_w\left(\mu_1 + \frac{1}{\gamma}\right)\frac{\partial\langle R_i(t)\rangle}{\partial\langle w\rangle}\bigg|_{(\lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})} - \frac{1}{\gamma}\lambda_s(\mu_1 + 1)\frac{\partial\langle R_i(t)\rangle}{\partial s}\bigg|_{(\lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})} = 0$$

The above is the characteristic polynomial for the Jacobian of the mean-field system of equations when we backsubstitute for $\mu_1 = \mu/\epsilon + O(\epsilon)$ (to leading order in ϵ). Thus, we know that the two solutions to μ_1 are equivalent to the mean-field eigenvalues for any steady state up to $O(\epsilon)$. Now, consider $\mu_0 = 2n\pi i\overline{\langle R\rangle}$. In this case, the $O(\epsilon)$ problem yields:

$$-4n^2\pi^2\mu_1\overline{\langle R\rangle} + 4n^2\pi^2\overline{\langle R\rangle}^2\lambda_w\hat{B}(2n\pi i\overline{\langle R\rangle}) + \frac{1}{\gamma}4n^2\pi^2\overline{\langle R\rangle}^2\lambda_s\hat{A}(2n\pi i\overline{\langle R\rangle}) = 0$$

Solving gives:

$$\begin{aligned}
\mu_1 &= \overline{\langle R\rangle}\lambda_w\hat{B}(2n\pi i\overline{\langle R\rangle}) + \overline{\langle R\rangle}\gamma\lambda_s\hat{A}(2n\pi i\overline{\langle R\rangle}) \\
&= -\lambda_w\overline{\langle R\rangle}\int_0^1 e^{2n\pi iy'}\frac{1}{G_1(\eta^{-1}(y'),\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})}\,dy' \\
&+ \frac{\lambda_s}{\gamma}\overline{\langle R\rangle}\int_0^1 e^{2n\pi iy'}\frac{g(e_r-\eta^{-1}(y))}{G_1(\eta^{-1}(y'),\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle})}\,dy' \\
&= \frac{\overline{\langle R\rangle}}{\gamma}\int_0^1 e^{2n\pi iy'}\frac{g\lambda_s(e_r-\eta^{-1}(y))-\gamma\lambda_w}{G_1(\eta^{-1}(y),\lambda_s\overline{\langle R\rangle},\lambda_w\overline{\langle R\rangle}}\,dy
\end{aligned}$$

Now, the integral terms are actually the Fourier coefficients of the functions $A(y)$ and $B(y)$. It should be clear that when the Fourier coefficients of $B(y)$ are negative, and the Fourier coefficients of $A(y)$ are positive, then the mean-field system forms a stable slow manifold that the solutions of the PDE system converge to provided that $\overline{\langle R\rangle} > 0$. To determine the sign of the integrands, there are two approaches one can take. First, we will look at a case where the integral equations are analytically resolvable.

The Fourier Coefficients for a Network of Quadratic Integrate and Fire Neurons

Consider the case when $G_1(v, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle})$ is for a network of Izhikevich/quadratic integrate and fire neurons:

$$\begin{aligned}
G_1(v, \lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle}) &= I - I^*(\lambda_s\overline{\langle R\rangle}, \lambda_w\overline{\langle R\rangle}) + F''(v^*)\frac{(v-v^*)^2}{2!} \\
&= \kappa + F''(v^*)\frac{(v-v^*)^2}{2!} \\
y &= \eta(v) = \int_{v_{reset}}^{v} \frac{\overline{\langle R\rangle}dv'}{\kappa + F''(v^*)\frac{(v-v^*)^2}{2!}} \\
&= \overline{\langle R\rangle}\sqrt{\frac{2}{F''(v^*)\kappa}}\left[\arctan\left((v-v^*)\sqrt{\frac{F''(v^*)}{2\kappa}}\right) - \arctan\left((v_{reset}-v^*)\sqrt{\frac{F''(v^*)}{2\kappa}}\right)\right] \\
v = \eta^{-1}(y) &= v^* + \sqrt{\frac{2\kappa}{F''(v^*)}}\tan\left[\frac{1}{\overline{\langle R\rangle}}\sqrt{\frac{\kappa F''(v^*)}{2}}y + \arctan\left((v_{reset}-v^*)\sqrt{\frac{F''(v^*)}{2\kappa}}\right)\right] \\
&= v^* + \sqrt{\frac{2\kappa}{F''(v^*)}}\tan(Cy+D) \\
G_1(\eta^{-1}(y), \lambda_s R, \lambda_w R) &= \kappa + \frac{F''(v^*)}{2}\frac{2\kappa}{F''(v^*)}\tan(Cy+D)^2 \\
&= \kappa(1+\tan^2(Cy+D)) \\
C &= \frac{1}{\overline{\langle R\rangle}}\sqrt{\frac{\kappa F''(v^*)}{2}} = \arctan\left((v_{peak}-v^*)\sqrt{\frac{F''(v^*)}{2\kappa}}\right) \\
&+ \arctan\left((v^*-v_{reset})\sqrt{\frac{F''(v^*)}{2\kappa}}\right) \\
&\sim \pi - \sqrt{\frac{2\kappa}{F''(v^*)}}\left(\frac{1}{v_{peak}-v^*} + \frac{1}{v^*-v_{reset}}\right), \quad \kappa \to 0 \\
D &= -\arctan\left((v^*-v_{reset})\sqrt{\frac{F''(v^*)}{2\kappa}}\right) \\
&\sim -\frac{\pi}{2} + \sqrt{\frac{2\kappa}{F''(v^*)}}\frac{1}{v^*-v_r}, \quad \kappa \to 0
\end{aligned}$$

As we only really want the real parts of μ_1 (although we can also exactly compute the imaginary parts), the sign of the real part of μ_1 can be determined by

$$\begin{aligned}
\mathrm{Re}\left(\int_0^1 e^{2n\pi i y'} B(y')dy'\right) &= \int_0^1 \frac{\cos(2n\pi y')(g\lambda_s\gamma(e_r - v^* - \sqrt{\frac{2\kappa}{F''(v^*)}}\tan(Cy+D)) - \lambda_w}{\kappa(1+\tan^2(Cy+D))}\,dy' \\
&= \frac{H_1(C,D)}{\kappa}\left(\frac{g\lambda_s\gamma(e_r-v^*)-\lambda_w}{-2n^2\pi^2+2C^2}\right) + \frac{H_2(C,D)}{\sqrt{\kappa}}\left(\frac{g\gamma\lambda_s}{(-2n\pi^2+2C^2)\sqrt{F''(v^*)}}\right) \\
H_1(C,D) &= 2C(\cos(C)^2\sin(D)\cos(D) + \left(\cos(D)^2 - \frac{1}{2}\right) - \sin(D)\cos(D)) \qquad (4.152) \\
H_2(C,D) &= 2\sqrt{2}C\left(\left(\cos(D)^2-\frac{1}{2}\right)\cos(C)^2 - \frac{1}{4}\sin(2C)\sin(2D) - \cos(D)^2 + \frac{1}{2}\right)
\end{aligned}$$

Provided that $H_1(C,D) > 0$, then $\mu_1 < 0$ if $\kappa \ll 1$ and the quantity $g(e_r - v^*(0))\lambda_s - \gamma\lambda_w > 0$. For the theta neuron, when the quantities H_1 and H_2 are evaluated at $D = -\pi/2$ and $C = \pi$ one obtains $H_1 = H_2 = 0$, and one has to compute a higher order perturbation to resolve this as the $O(\epsilon)$ contribution is 0. For all the other neuron models however, $H_1(C,D) > 0$ when $C < \pi$ and $D > -\frac{\pi}{2}$ locally near $C = \pi$ and $D = -\frac{\pi}{2}$ which is precisely the case when one expands out C and D asymptotically.

This implies that for the all the type-I neurons and any quadratic integrate and fire neuron with a finite peak/reset, as $I - I^*(s, \langle w\rangle) \to 0^+$, and $\tau_w, \tau_s \gg 1$, the mean-field system of equations forms a finite dimensional stable slow manifold. One should be careful however in terms of what this means. This is a stable slow manifold for the tonic firing solutions in the sense that the stability of the steady tonic firing solutions of the population density equation is exactly mirrored by the mean-field system in the vicinity of the switching manifold when τ_w and τ_s are large. This says nothing about other potential steady state solutions such as oscillatory solutions (bursting etc.)

4.4.4 The Fast System: The Eigenvalue Problem

Recall the fast system from section 4.4:

$$\begin{aligned}
\frac{\partial}{\partial t}\rho(v,t) &= -\frac{\partial}{\partial v}\left((F(v) - \langle w\rangle + I + g(e_r - v)s)\,(\rho(v,t))\right) \qquad (4.153) \\
\frac{d\langle w\rangle}{dt} &= \epsilon\left[b\langle v\rangle - \langle w\rangle + \lambda_w J(v_{peak}, \langle w\rangle, s, \tilde{t})\right] \qquad (4.154) \\
\frac{ds}{dt} &= \epsilon\gamma^{-1}\left[-s + \lambda_s J(v_{peak}, \langle w\rangle, s, t)\right], \qquad (4.155)
\end{aligned}$$

with fast time $t = \bar{t}/\epsilon$. The slow system that we previously analyzed comes from the standard fast-slow analysis where the fast variable(s) are evaluated at their equilibrium values. The other approach one can take in fast-slow analysis is to consider the slow variables as constants, and analyze the fast variables. The $O(1)$ solution for this system is

also resolvable, and as we shall see, will correspond to the eigenvalues $2n\pi i\overline{R}$ found earlier. We will again consider the differential equation for the distribution function:

$$P(v,t) = \int_{v_{reset}}^{v} \rho(v',t)\,dv'$$

Deriving the partial differential equation for the distribution function yields the following $O(1)$ problem:

$$\begin{aligned} \frac{\partial}{\partial t}P(v,t) &= -G_1(v,\langle w\rangle,s)\frac{\partial P}{\partial v} + G_1(v_{peak},\langle w\rangle,s)\frac{\partial P}{\partial v}\bigg|_{vpeak} && (4.156)\\ s' &= 0 && (4.157)\\ \langle w\rangle' &= 0 && (4.158) \end{aligned}$$

As $s' = 0$ and $\langle w\rangle' = 0$, s and $\langle w\rangle$ are constants and we can merely take them to be s^* and $\langle w\rangle^*$. Consider again the Abbott/Vresswijk transform:

$$y = R^* \int_{v_{reset}}^{v} \frac{dv'}{G_1(v', s^*, \langle w\rangle^*)} = \eta(v) \tag{4.159}$$

where s^* and w^* are the values for which this transform exists and is invertible. The quantity R^* corresponds to

$$R^* = \left(\int_{v_{reset}}^{v_{peak}} \frac{dv'}{G_1(v, s^*, \langle w\rangle^*)}\right)^{-1}$$

to confine $y \in [0,1]$. Also, if $I - I^*(s^*, w^*) > 0$, then $\eta(v)$ is invertible. Let

$$\begin{aligned} P(v,t) &= P(\eta^{-1}(y),t) = \hat{P}(y,t) && (4.160)\\ \Rightarrow \frac{\partial \hat{P}(y,t)}{\partial t} &= -R^*\frac{\partial \hat{P}}{\partial y} + R^*\frac{\partial \hat{P}}{\partial y}\bigg|_{y=1} && (4.161) \end{aligned}$$

The $O(1)$ problem for the fast system is completely analytically resolvable with classical approaches. In particular, let us look for a separable solution:

$$\begin{aligned} P(y,t) &= Y(y)T(t) && (4.162)\\ T'(t)Y(y) &= -R^*Y'(y)T(t) + R^*Y'(1)T(t) && (4.163) \end{aligned}$$

dividing both sides by $T(t)Y(y)$ yields the following eigenvalue problem:

$$\frac{T'(t)}{T(t)} = -R^*\frac{Y'(y)}{Y(y)} + R^*\frac{Y'(1)}{Y(y)} = \lambda \tag{4.164}$$

Now, we'll consider two cases, $\lambda = 0$ and $\lambda \neq 0$. The first case implies that

$$\begin{aligned} -R^*Y'(y) + R^*Y'(1) &= 0 && (4.165)\\ Y_0'(y) &= Y_0'(1) && (4.166)\\ Y_0(y) &= yY_0'(1) + C && (4.167) \end{aligned}$$

For the $\lambda = 0$ case, we will require that $Y(1) = 1$ and $Y(0) = 0$, this satisfies the conditions on the distribution function ($P(1,t) = 1, P(0,t) = 0$ for all t). This implies that

$$Y_0 = y \tag{4.168}$$

for $\lambda \neq 0$, we have the following equation:

$$-R^*Y'(y) + R^*Y'(1) = \lambda Y(y) \tag{4.169}$$

we subject this ODE to the conditions $Y(0) = 0$, and $Y(1) = 0$ which yields the following solution:

$$\lambda = 2n\pi R^* i, \quad n = 1, 2, \ldots \tag{4.170}$$

$$Y_n(y) = \frac{1}{2n\pi i}(1 - \exp(-2n\pi iy))Y_n'(1) \tag{4.171}$$

where $Y_n'(y)$ can be absorbed into the differential equation for $T_n(t)$ as a dummy constant. In particular, for $\lambda \neq 0$, we have the following for $T_n(T)$:

$$T_n(t) = A_n \exp(2n\pi R^* it) \tag{4.172}$$

In which case, we may write the entire solution to $P(y,t)$ as:

$$P(y,t) = y + \sum_{n\neq 0} \frac{A_n}{2n\pi i} e^{2n\pi i R^* t}(1 - e^{-2n\pi iy}) \tag{4.173}$$

where

$$P(y,0) = y + \sum_{n\neq 0} \frac{A_n}{2n\pi i}(1 - e^{-2n\pi iy}) \tag{4.174}$$

$$= y + \sum_{n=1}^{\infty} \frac{A_n}{2n\pi} \sin(2n\pi y) \tag{4.175}$$

Note that if we differentiate, then we have the following:

$$\rho_0(y) = 1 + \sum_{n=1}^{\infty} A_n \sin(2n\pi y) \tag{4.176}$$

and thus the coefficients A_n are the Fourier series coefficients of $\rho_0(y) - 1$. Note that this implies the firing rate, defined by the flux, oscillates around R^*:

$$J(v_{peak}, t) = R^* \frac{\partial P}{\partial y}\bigg|_{y=1} = R^* \left(1 + \sum_{n=0}^{\infty} A_n e^{2n\pi i R^* t}\right) \tag{4.177}$$

Thus, the firing rate, given by $J(v_{peak}, t)$ merely oscillates in time off the steady state firing rate R^*.

With a fast-slow analysis of the population density equation and the coupled equations for the mean adaptation and synaptic gating variables, one can analytically resolve both the fast and the slow system of the original coupled population density system.

Chapter 5

Mean-Field Systems for More Realistic Networks

In this chapter, we will consider the derivation of mean-field systems for more complicated networks of neurons. In particular, we will consider separately the effects of parameter heterogeneity (section 5.1), noise (section 5.2) and sparse coupling between neurons (section 5.3). For heterogeneity and noise we will proceed as in Chapter 4, by starting with a population density equation, applying a sequence of analytical reductions to arrive at a one-dimensional partial differential equation coupled to ordinary differential equations for the moments of interest, and subsequently decomposing the resulting equation to slow and fast subsystems. The slow system leads to variations of the mean-field system derived in chapter 4, while additional information can be gleaned from the fast system as well. Finally, we will show that the type of sparsity we will consider here (random connectivity) can effectively be treated as a source of heterogeneity.

5.1 Parameter Heterogeneity

We will start with our network of integrate and fire neurons:

$$\begin{aligned} \dot{v}_i &= F(v_i) - w_i + gs(e_r - v_i) + I && (5.1)\\ \dot{w}_i &= a(bv_i - w_i) && (5.2)\\ \dot{s} &= -\frac{s}{\tau_s} + \frac{s_{jump}}{N}\sum_{i=1}^{N}\sum_{t<t_{j,k}}\delta(t - t_{j,k}) && (5.3)\\ v_i(t^-) &= v_{peak}, \rightarrow \begin{cases} v_i(t^+) = v_{reset} \\ w_i(t^+) = w_i(t^-) + w_{jump} \end{cases} && (5.4) \end{aligned}$$

only now, we will let any of the parameters come from some random distribution that is thought to represent the natural heterogeneity present in real neurons. For example, we can

let the current to each neuron I_i be a normally distributed random variable, $I_i \sim \mathbb{N}(\mu_I, \sigma_I^2)$. To make the derivations more compact, we will denote the source(s) of heterogeneity by the vector of parameters $\boldsymbol{\beta}$, and the state variables v and w by $\boldsymbol{x} = (v, w)$ and rewrite the dynamics as:

$$\dot{\mathbf{x}} = \mathbf{G}(\mathbf{x}, \boldsymbol{\beta}, s) = \begin{pmatrix} G_1(\boldsymbol{x}, \boldsymbol{\beta}, s) \\ G_2(\boldsymbol{x}, \boldsymbol{\beta}) \end{pmatrix} \tag{5.5}$$

Note that for a specific heterogeneous parameter(s), G_1 and G_2 may not depend on $\boldsymbol{\beta}$, or all of the components of $\boldsymbol{\beta}$. However, for the sake of simplicity, we include the dependence in both equations.

Now, in order to proceed in deriving a mean-field system we will need to first look at the population density equation in section 5.1.1. It turns out that heterogeneity complicates the set of possible moment closure assumptions to the point where multiple "mean-field" systems can actually be generated depending on the specific moment closure assumption used. These mean-field systems all have different interpretations and potential applications. Additionally, they are all qualitatively different types of equations. These systems will be referred to as mean-field systems I, II, and III and will be derived in section 5.1.1. We will look at applications of these mean-field systems in section 5.1.2.

5.1.1 The Population Density Equation

The population density function, $\rho(\boldsymbol{x}, t)$ determines the density of neurons at a point in phase space, $\boldsymbol{x}$, at time t. Consider first the case of a homogeneous network, i.e., all the neurons have the same parameter values, denoted by $\boldsymbol{\beta}$. In the limit as $N \to \infty$, one can derive the following evolution equation for the population density function:

$$\frac{\partial \rho(\boldsymbol{x}, t)}{\partial t} = -\nabla \cdot \mathbf{J}(\boldsymbol{x}, \boldsymbol{\beta}, s, t) \tag{5.6}$$

where $\mathbf{J}$ is given by

$$\mathbf{J}(\boldsymbol{x}, \boldsymbol{\beta}, s, t) = \mathbf{G}(\boldsymbol{x}, \boldsymbol{\beta}, s)\rho(\boldsymbol{x}, t) = \left(J^V, J^W\right). \tag{5.7}$$

and must satisfy the boundary condition

$$J^V(v_{peak}, w, \boldsymbol{\beta}, s, t) = J^V(v_{reset}, w + w_{jump}, \boldsymbol{\beta}, s, t). \tag{5.8}$$

In the same limit, the differential equation for s becomes

$$s' = -\frac{s}{\tau_s} + s_{jump} \int_W J^V(v_{peak}, w, s, \boldsymbol{\beta}, t)\, dw \tag{5.9}$$

where the integral term is actually the network averaged firing rate, which we denote as $\langle R_i(t) \rangle$.

Now consider a heterogeneous network where the parameters vary from neuron to neuron to neuron, but are static in time. Then one can rewrite the equations for the individual oscillator as

$$\begin{aligned} \dot{v}_i &= G_1(\boldsymbol{x}_i, \boldsymbol{\beta}_i, s), && (5.10)\\ \dot{w}_i &= G_2(\boldsymbol{x}_i, \boldsymbol{\beta}_i), && (5.11)\\ \dot{\boldsymbol{\beta}}_i &= 0. && (5.12) \end{aligned}$$

In this case the flux contribution due to $\boldsymbol{\beta}$ is 0, and the evolution equation for the network is immediately given by

$$\frac{\partial \rho(\boldsymbol{x}, \boldsymbol{\beta}, t)}{\partial t} = -\nabla \cdot \mathbf{J}(\boldsymbol{x}, \boldsymbol{\beta}, s, t) \tag{5.13}$$

The density now has the vector of parameters, $\boldsymbol{\beta}$, as an independent variable. The flux consists of the vector $(J^V, J^W, 0)$, with $\boldsymbol{\beta}$ as an independent variable, as opposed to a fixed constant.

If the parameters are time varying, equation (5.13) will be unchanged. However, the final component of the flux will be non-zero. We will not consider this case further.

The equation for s is also different in the heterogeneous case:

$$s' = -\frac{s}{\tau_s} + s_{jump} \int_W \int_{\boldsymbol{\beta}} J^V(v_{peak}, w, s, \boldsymbol{\beta}', t)\, dw\, d\boldsymbol{\beta}'. \tag{5.14}$$

Now, in order to proceed analytically, as in section 4.1, we need a moment closure assumption. This is where the ambiguity lies in defining a mean-field system as multiple moment closure assumptions are actually possible. Note that for the most part, we will assume that the time scales of the adaptation variable and the synapses are homogeneous across the network of neurons. All other parameters can be either heterogeneous or homogeneous.

Derivation of Mean-Field I

We begin by writing out the density function in the conditional form

$$\rho(\mathbf{x}, \boldsymbol{\beta}, t) = \rho_{\boldsymbol{x}}(\mathbf{x}, t)\rho_{\beta}(\boldsymbol{\beta}|\boldsymbol{x}, t) \tag{5.15}$$

The continuity equation is then given by

$$\frac{\partial \left(\rho_{\boldsymbol{x}}(\mathbf{x}, t)\rho_{\beta}(\boldsymbol{\beta}|\boldsymbol{x}, t)\right)}{\partial t} = -\nabla \cdot \boldsymbol{J}(\boldsymbol{x}, s, \boldsymbol{\beta}, t). \tag{5.16}$$

Simple integration with respect to $\boldsymbol{\beta}$ yields the reduced continuity equation

$$\frac{\partial \rho_{\boldsymbol{x}}(\boldsymbol{x}, t)}{\partial t} = -\nabla \cdot \boldsymbol{J}(\mathbf{x}, \mathbf{s}, \langle \boldsymbol{\beta}|\mathbf{x} \rangle, \mathbf{t}). \tag{5.17}$$

This step is valid for all the non-dimensionalized models we consider as they are all linear in their dimensionless parameters (see [177]). The flux has also been redefined upon integration to

$$(J^V, J^W) = \rho_{\boldsymbol{x}}(\boldsymbol{x}, t)\,(G_1(\boldsymbol{x}, \langle\boldsymbol{\beta}|x\rangle, s), G_2(\boldsymbol{x}, \langle\boldsymbol{\beta}|x\rangle))\,.$$

We now apply the first order moment closure assumption $\langle\boldsymbol{\beta}|\boldsymbol{x}\rangle = \langle\boldsymbol{\beta}\rangle$ to yield the following PDE:

$$\frac{\partial \rho_{\boldsymbol{x}}(\boldsymbol{x}, t)}{\partial t} = -\nabla \cdot \boldsymbol{J}(\mathbf{x}, \mathbf{s}, \langle\boldsymbol{\beta}\rangle, \mathbf{t}). \tag{5.18}$$

It should be clear that this is equivalent to the continuity equation for a homogeneous network with parameter values fixed at $\langle\boldsymbol{\beta}\rangle$. Thus, the associated mean-field system is identical to the homogeneous case, only with the parameters fixed at $\langle\boldsymbol{\beta}\rangle$. This is the simplest assumption one can make in the heterogeneous case. For example, if we treat I as the source of heterogeneity for a network of Izhikevich neurons, with distribution $\rho_I(I)$, then the resulting mean-field system is

$$\langle w\rangle' = \frac{b\langle v\rangle - \langle w\rangle}{\tau_w} + w_{jump}\langle R_i(t)\rangle \tag{5.19}$$

$$s' = -\frac{s}{\tau_s} + s_{jump}\langle R_i(t)\rangle \tag{5.20}$$

$$\langle R_i(t)\rangle = \begin{cases} \left(\int_V \frac{dv}{v(v-\alpha)-\langle w\rangle+\langle I\rangle+g(e_r-v)s}\right)^{-1} & : H(\langle w\rangle, s, \langle I\rangle) \geq 0 \\ 0 & : H(\langle w\rangle, s, \langle I\rangle) < 0 \end{cases} \tag{5.21}$$

$$H(\langle w\rangle, s, \langle I\rangle) = \langle I\rangle - \langle w\rangle - \frac{(\alpha+gs)^2}{4} + ge_r s \tag{5.22}$$

$$\langle v\rangle = \begin{cases} \frac{\langle R_i(t)\rangle}{2}\log\left(\frac{(v_{peak}-\frac{\alpha+gs}{2})^2+H(\langle w\rangle,s,\langle I\rangle)}{(v_{reset}-\frac{\alpha+gs}{2})^2+H(\langle w\rangle,s,\langle I\rangle)}\right) + \frac{\alpha+gs}{2} & : H(\langle w\rangle, s, \langle I\rangle) \geq 0 \\ \frac{\alpha+gs}{2} - \sqrt{-H(\langle w\rangle, s, \langle I\rangle)} & : H(\langle w\rangle, s, \langle I\rangle) < 0 \end{cases} \tag{5.23}$$

Note that I in equations (5.21)-(5.23) has been replaced by $\langle I\rangle$. We treat this system as the baseline mean-field model for comparison purposes, in addition to direct numerical simulations of the network, and we denote this system of equations as mean-field one (MFI). We should expect this system to be an adequate approximation to the actual network for narrowly centered distributions of the parameter heterogeneity (small values of the variance, σ_β).

This set of differential equations is representative of a common approach taken when fitting actual neurons. In this approach, multiple estimates of parameters or measurements taken from multiple neurons are averaged to yield a single parameter value, which is really the mean parameter value, $\langle\boldsymbol{\beta}\rangle$. Simulations of homogeneous, large networks are then run with the parameters fixed at their mean values. As we shall see in subsequent sections, the behavior of a simulated heterogeneous network can differ substantially from that of MFI. We should also note that replacing a distribution of parameters with its mean is in general a dubious proposition as the measured parameter distributions from neurons can be quite broad [79, 80].

Derivation of Mean-Field II

Suppose that instead, we write the density function in the alternative conditional form

$$\rho(v,w,\boldsymbol{\beta},t) = \rho_W(w,t|\boldsymbol{\beta},v)\rho_V(v,t|\boldsymbol{\beta})\rho_\beta(\boldsymbol{\beta}). \tag{5.24}$$

Next we integrate the continuity equation with respect to w. This yields the following system

$$\begin{aligned}
\frac{\partial \rho_V(v,t|\boldsymbol{\beta})}{\partial t}\rho_\beta(\boldsymbol{\beta}) &= -\int_W \left(\frac{\partial J^V(v,w,s,\boldsymbol{\beta},t)}{\partial v} + \frac{\partial J^W(v,w,\boldsymbol{\beta},t)}{\partial w}\right) dw \\
&= -\frac{\partial}{\partial v} J(v,\langle w|v,\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta},t) - J^W(v,w,\boldsymbol{\beta},t)|_{\partial W} \\
&= -\frac{\partial}{\partial v} J(v,\langle w|v,\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta},t),
\end{aligned} \tag{5.25}$$

where the last term vanishes as J^W is assumed vanish on the boundary, and

$$J(v,\langle w|v,\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta},t) = \int_W J^V(v,w,s,\boldsymbol{\beta},t)\,dw. \tag{5.26}$$

since J^V is linear in w. We now make the first order moment closure assumption $\langle w|v,\boldsymbol{\beta}\rangle = \langle w\rangle$. Then to complete the system, we must derive a differential equation for $\langle w\rangle$:

$$\begin{aligned}
\langle w\rangle' &= \int_V\int_W\int_{\boldsymbol{\beta}} w\frac{\partial \rho(v,w,\boldsymbol{\beta},t)}{\partial t}\,d\boldsymbol{\beta}\,dw\,dv \\
&= -\int_V\int_W\int_{\boldsymbol{\beta}} w\left(\frac{\partial J^W}{\partial w} + \frac{\partial J^V}{\partial v}\right) d\boldsymbol{\beta}\,dw\,dv \\
&= \int_V\int_W\int_{\boldsymbol{\beta}} G_2(v,w,\boldsymbol{\beta})\rho(v,w,\boldsymbol{\beta},t)\,d\boldsymbol{\beta}\,dw\,dv \\
&- \int_W\int_{\boldsymbol{\beta}} w(J^V(v_{peak},w,s,\boldsymbol{\beta},t) - J^V(v_{reset},w,s,\boldsymbol{\beta},t))\,d\boldsymbol{\beta}\,dw \\
&= \langle G_2(v,w,\boldsymbol{\beta})\rangle \\
&- \int_W\int_{\boldsymbol{\beta}} w(J^V(v_{peak},w,s,\boldsymbol{\beta},t) - J^V(v_{peak},w-w_{jump},s,\boldsymbol{\beta},t))\,d\boldsymbol{\beta}\,dw \qquad (5.27) \\
&= \langle G_2(v,w,\boldsymbol{\beta})\rangle + \int_{\boldsymbol{\beta}}\int_W w_{jump}J^V(v_{peak},w,s,\boldsymbol{\beta},t)\,dw\,d\boldsymbol{\beta} + O(w_{jump}^2) \\
&\approx G_2(\langle v\rangle,\langle w\rangle,\langle\boldsymbol{\beta}\rangle) + \int_{\boldsymbol{\beta}} w_{jump}J(v_{peak},\langle w\rangle,s,\boldsymbol{\beta},t)\,d\boldsymbol{\beta}. \qquad (5.28)
\end{aligned}$$

Note that we have made the approximation $\langle G_2(v,w,\boldsymbol{\beta})\rangle = G_2(\langle v\rangle,\langle w\rangle,\langle\boldsymbol{\beta}\rangle)$ in addition to dropping the $O(w_{jump}^2)$ terms. Additionally, the substitution on line (5.27) comes from the boundary condition (5.8).

Applying our separation of time scales as in Chapter 4, yields the following equation for the steady state voltage independent flux, $J(\langle w\rangle,s,\boldsymbol{\beta})$:

$$J(\langle w\rangle,s,\boldsymbol{\beta}) = \begin{cases} \left[\int_V \frac{dv}{G_1(v,\langle w\rangle,s,\boldsymbol{\beta})}\right]^{-1}\rho_\beta(\boldsymbol{\beta}) & \text{if} \quad H(\langle w\rangle,s,\boldsymbol{\beta}) \geq 0 \\ 0 & \text{if} \quad H(\langle w\rangle,s,\boldsymbol{\beta}) < 0 \end{cases}. \tag{5.29}$$

We interpret the ratio $J(\langle w\rangle, s, \boldsymbol{\beta})/\rho_{\boldsymbol{\beta}}(\boldsymbol{\beta})$ as the parameter dependent (or conditional) network averaged firing rate, $\langle R_i(t)|\boldsymbol{\beta}\rangle$, based on the fact that

$$\int_{\boldsymbol{\beta}} J(\langle w\rangle, s, \boldsymbol{\beta})\, d\boldsymbol{\beta} = \langle R_i(t)\rangle + O(w_{jump})$$

In other words, the distribution of parameters induces a distribution of firing rates across the network, and the network averaged firing rate is the mean of the distribution.

In summary, the resulting mean-field equations are given by:

$$\langle w\rangle' = \frac{b\langle v\rangle - \langle w\rangle}{\tau_w} + \int_{\boldsymbol{\beta}} w_{jump}\langle R_i(t)|\boldsymbol{\beta}\rangle \rho_\beta(\boldsymbol{\beta})\, d\boldsymbol{\beta} \tag{5.30}$$

$$s' = -\frac{s}{\tau_s} + s_{jump}\int_{\boldsymbol{\beta}} \langle R_i(t)|\boldsymbol{\beta}\rangle \rho_\beta(\boldsymbol{\beta})\, d\boldsymbol{\beta} \tag{5.31}$$

$$\langle R_i(t)|\boldsymbol{\beta}\rangle = \begin{cases} \left[\int_V \frac{dv}{G_1(v,s,\langle w\rangle,\boldsymbol{\beta})}\right]^{-1} & : H(\langle w\rangle, s, \boldsymbol{\beta}) \geq 0 \\ 0 & : H(\langle w\rangle, s, \boldsymbol{\beta}) < 0 \end{cases} \tag{5.32}$$

$$H(\langle w\rangle, s, \boldsymbol{\beta}) = I - \langle w\rangle + \min_v (F(v) + g(e_r - v)s) \tag{5.33}$$

$$\langle v\rangle = \int_{\boldsymbol{\beta}} \langle v|\boldsymbol{\beta}\rangle \rho_\beta(\boldsymbol{\beta})\, d\boldsymbol{\beta} \tag{5.34}$$

where the forms of $G_1(v, s, \langle w\rangle, \boldsymbol{\beta})$ and $H(\langle w\rangle, s, \boldsymbol{\beta})$ depend on which specific neural model is used. We refer to equations (5.30)-(5.34) as mean-field two (MFII).

There are a few things to notice about this set of equations. In particular, It appears that MFII adds a degree of smoothness to the non-smooth MFI equations since the piecewise smooth continuous firing rate is integrated in (5.30)-(5.31). Additionally, MFI and MFII also differ in the order in which the integrations are carried out. In MFI, we integrate with respect to $\boldsymbol{\beta}$ first, and then apply the first order moment closure assumptions $\langle \boldsymbol{\beta}|\boldsymbol{x}\rangle = \langle \boldsymbol{\beta}\rangle$ and $\langle w|v\rangle = \langle w\rangle$. In MFII, we integrate with respect to w first, and then apply the moment closure assumption $\langle w|v, \boldsymbol{\beta}\rangle = \langle w\rangle$. Furthermore, if $\langle R_i(t)|\boldsymbol{\beta}\rangle$ does not actually depend on the heterogeneous parameter $\boldsymbol{\beta}$, such as when the heterogeneity is in w_{jump}, then MFI and MFII are identical.

The first order moment closure assumption used here can be weakened. This leads to the "mean-field" system in the next subsection, which is a different kind of system than MFI and MFII.

Derivation of Mean-Field III

Suppose that instead of assuming that $\langle w|v, \boldsymbol{\beta}\rangle = \langle w\rangle$, we make the weaker assumption that $\langle w|v, \boldsymbol{\beta}\rangle = \langle w|\boldsymbol{\beta}\rangle$. It turns out that this assumption yields a PDE, even when one makes the separation of time scales.

Applying this weaker moment closure assumption to (5.25) yields the following simplification of the continuity equation:

$$\frac{\partial \rho_V(v,t|\boldsymbol{\beta})}{\partial t}\rho_\beta(\boldsymbol{\beta}) = -\frac{\partial}{\partial v}J(v,\langle w|\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta},t). \tag{5.35}$$

Analysis of the slow system yields the following steady state flux:

$$\begin{aligned} J(\langle w|\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta}) &= \begin{cases} \left[\int_V \frac{dv}{G_1(v,s,\langle w|\boldsymbol{\beta}\rangle,\boldsymbol{\beta})}\right]^{-1}\rho_\beta(\boldsymbol{\beta}) & : H(\langle w|\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta}) \geq 0 \\ 0 & : H(\langle w|\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta}) < 0 \end{cases}, \\ H(\langle w|\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta}) &= I - \langle w|\boldsymbol{\beta}\rangle + \min_v(F(v)+g(e_r-v)s). \end{aligned}$$

An equation for the time variation of $\langle w|\boldsymbol{\beta}\rangle$ can be derived in a similar manner to the last section, yielding the following mean-field system:

$$\langle w|\boldsymbol{\beta}\rangle' = \frac{b\langle v|\boldsymbol{\beta}\rangle - \langle w|\boldsymbol{\beta}\rangle}{\tau_w} + w_{jump}\langle R_i(t)|\boldsymbol{\beta}\rangle \tag{5.36}$$

$$s' = -\frac{s}{\tau_s} + s_{jump}\int_\beta \langle R_i(t)|\boldsymbol{\beta}\rangle\rho_\beta(\boldsymbol{\beta})\,d\boldsymbol{\beta} \tag{5.37}$$

$$\langle R_i(t)|\boldsymbol{\beta}\rangle = \begin{cases} \left[\int_V \frac{dv}{G_1(v,\langle w|\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta})}\right]^{-1} & : H(\langle w|\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta}) \geq 0 \\ 0 & : H(\langle w|\boldsymbol{\beta}\rangle,s,\boldsymbol{\beta}) < 0 \end{cases} \tag{5.38}$$

Note that $\langle w\rangle$ can be computed via:

$$\langle w\rangle = \int_\beta \langle w|\boldsymbol{\beta}\rangle\rho_\beta(\boldsymbol{\beta})\,d\boldsymbol{\beta}. \tag{5.39}$$

We denote this system as mean-field three (MFIII). Note that the equation for $\langle w|\boldsymbol{\beta}\rangle$ is actually a PDE. This is due to the fact that the conditional moments, $\langle w|\boldsymbol{\beta}\rangle$, $\langle R|\boldsymbol{\beta}\rangle$, and $\langle v|\boldsymbol{\beta}\rangle$ are functions of both time, and the variable $\boldsymbol{\beta}$. This can be made more explicit if we write out $\langle w|\boldsymbol{\beta}\rangle = f(\boldsymbol{\beta},t)$, then the equations (5.36)-(5.38) can be written as

$$\begin{aligned} \frac{\partial f(\boldsymbol{\beta},t)}{\partial t} &= \frac{b\langle v|\boldsymbol{\beta}\rangle - f(\boldsymbol{\beta},t)}{\tau_w} + w_{jump}\langle R_i(t)|\boldsymbol{\beta}\rangle \\ s' &= -\frac{s}{\tau_s} + s_{jump}\int_\beta \langle R_i(t)|\boldsymbol{\beta}\rangle\rho_\beta(\boldsymbol{\beta})\,d\boldsymbol{\beta} \\ \langle R_i(t)|\boldsymbol{\beta}\rangle &= \begin{cases} \left[\int_V \frac{dv}{G_1(v,f(\boldsymbol{\beta},t),s,\boldsymbol{\beta})}\right]^{-1} & : H(f(\boldsymbol{\beta},t),s,\boldsymbol{\beta}) \geq 0 \\ 0 & : H(f(\boldsymbol{\beta},t),s,\boldsymbol{\beta}) < 0 \end{cases} \end{aligned}$$

This is a partial differential equation, however it is easier to deal with than most PDEs in the sense that it has no derivatives with respect to $\boldsymbol{\beta}$. The only complication is that the right hand side of the differential equation is non-smooth.

While this system should be more accurate than mean-field II, it has the drawback of being more difficult to analyze. The dependence on $\boldsymbol{\beta}$ forces one to discretize over a

mesh in $\boldsymbol{\beta}$ in order to work numerically with this system. This is the approach we used to numerically simulate this PDE, and is typically referred to as the method of lines in the literature. In order to compute the integrals while using the method of lines, the grid is non-uniform and generated with the density function $\rho_\beta(\boldsymbol{\beta})$. The integrals are subsequently replaced with averaging over the entire grid, which is precisely a Monte-Carlo method for estimating the integrals. This is explained in greater detail in the Appendix.

5.1.2 Applications and Numerical Simulations of the Slow Systems

Preliminary Numerical Simulations

We will first numerically simulate these three mean-field systems, and the corresponding networks of neurons with heterogeneity for comparison purposes.

For tonic firing (Figure 5.1(a)), even when the standard deviation is large, all mean-field systems approximate the network means $\langle g(t)\rangle$, and $\langle w(t)\rangle$ very well. The quantity $\langle g(t)\rangle$ is $gs(t)$ when g is homogeneous and $\langle g\rangle s(t)$ when g is heterogeneous. When the network is bursting, with $\langle I_{app}\rangle > I_{rh}$, we see a difference as to which mean-field system is superior. For small values of σ_I, we have numerically found that mean-field I is superior to mean-field II and III, however all the systems are quantitatively and qualitatively accurate (see Figure 5.1(b),5.1(c)). However, for larger values of σ_I, the amplitude error of MFIII is the smallest, and MFII is the worst approximation as it bifurcates to tonic firing prematurely.

When $\langle I_{app}\rangle$ is close to I_{rh}, we see even stronger differences between the three mean-field systems. For small to intermediate standard deviations, MFII and MFIII are clearly superior to MFI, having a smaller amplitude and frequency error (see Figure 5.2(a), 5.2(b)). However, for larger values of σ_I as shown in Figure 5.2(c), and 5.2(d), only MFIII is a qualitatively and quantitatively accurate representation of the behavior of the network. The amplitude and frequency error of MFI are very large, and MFII again bifurcates prematurely to tonic firing.

One should note that for $\langle I_{app}\rangle = O(I_{rh})$ and for large values of σ_I, the network can undergo a period doubling bifurcation. This is shown in Figure 7.4. The large standard deviation in the current forces different neurons into different regimes, such as tonic firing, bursting, alternate burst firing, and quiescence as seen in Figure 5.3(b). The fact that a small subpopulation of neurons are alternate bursters (i.e., burst with twice the period of the rest of the bursting neurons) appears as a period doubled limit cycle in the mean variables of the network, as seen in Figure 5.3(a). Only MFIII is able to approximate the period-doubled limit cycle with any degree of accuracy, as shown in Figure 5.4(c) and 5.4(b). Period doubling bifurcations are well known for their capability of inducing chaos. Given that MFIII accurately represents the period doubling bifurcation, it may be able to replicate any potential chaotic behavior. However, we leave further investigation of this interesting behaviour for future work. We do note however that this observation was

first made by [28] in their analysis of populations of bursting neurons that were models of pacemaker cells in the pre-Bötzinger complex [27]. They found that the period doubling also only occurred under what they refer to as extreme heterogeneity [28]. Additionally, they found a full cascade to chaotic behavior (see Figure 4 in [28]).

To summarize, all the mean-field systems are valid for firing parameter regimes, and MFI is valid for all parameter regimes with small σ_I, except for $\langle I_{app}\rangle = O(I_{rh})$. Mean-Field II and III are valid for bursting with $\langle I_{app}\rangle \gg I_{rh}$, and MFIII is the only valid approximation for $\langle I_{app}\rangle = O(I_{rh})$.

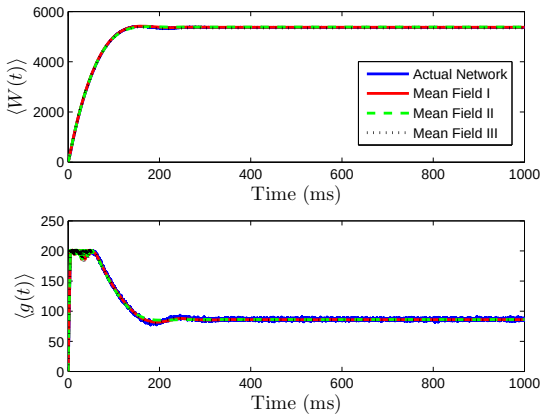
Dimensional Parameters		Dimensionless Parameters	
C	250 pF		
k	2.5 nS/mV		
V_R	-65 mV		
V_T	$V_R + 40 - \frac{b}{k} = 41.7$mV	$\alpha = 1 + \frac{V_T}{\lvert V_R\rvert}$	0.6215
V_{peak}	30 mV	$v_{peak} = 1 + \frac{V_{peak}}{\lvert V_R\rvert}$	1.461
V_{reset}	-55 mV	$v_{reset} = 1 + \frac{V_{reset}}{\lvert V_R\rvert}$	0.1538
W_{jump}	200 pA	$w_{jump} = \frac{W_{jump}}{k\lvert V_R\rvert^2}$	0.0189
τ_W	200 ms	$a = \left(\frac{\tau_W k\lvert V_R\rvert}{C}\right)^{-1}$	0.0077
$\hat{b}$	-1 nS	$b = \frac{\hat{b}}{k\lvert V_R\rvert}$	-0.0062
I_{app}	1000 - 5000 pA	$I = \frac{I_{app}}{k\lvert V_R\rvert^2}$	0.0776 - 0.3333
g_{syn}	0 - 600 nS	$g = \frac{g_{syn}}{k\lvert V_R\rvert}$	0 - 3.6923
τ_{syn}	4 ms	$\tau_s = \frac{\tau_{syn} k\lvert V_R\rvert}{C}$	2.6
s_{jump}	0.8		
N	1000		
σ_I	0 - 1000 pA		

Table 5.1: The parameters and distribution variances used in this paper. These parameters apply unless otherwise indicated. Rheobase for the dimensional parameter values is $I_{rh} =$ 1000 pA.

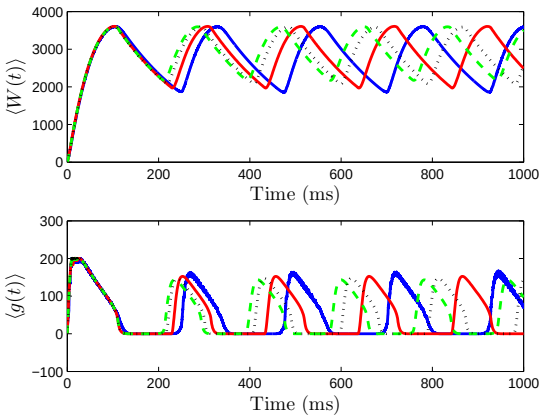
Steady State Distribution Estimation and Parameter Inversion

For a network of neurons with heterogeneity in the parameters, even if all the neurons are tonically firing, one cannot find a steady state average firing rate for the network, as in the case of a homogeneous network. The parameter heterogeneity creates a distribution of steady state firing rates across the network. While the mean-field equations by themselves can only determine the mean of this distribution, via the integral:

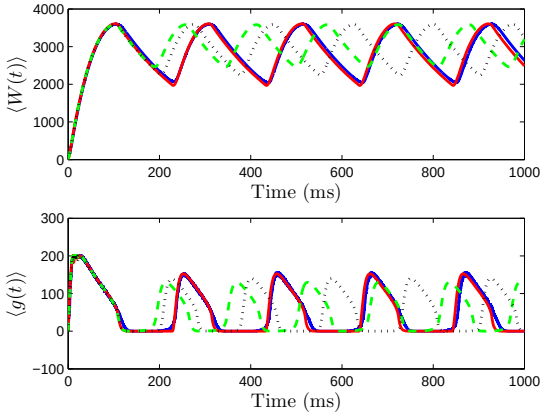
$$\langle R_i(t)\rangle = \int_{\boldsymbol{\beta}} \langle R_i(t)|\boldsymbol{\beta}\rangle \rho_{\boldsymbol{\beta}}(\boldsymbol{\beta})\, d\boldsymbol{\beta}$$



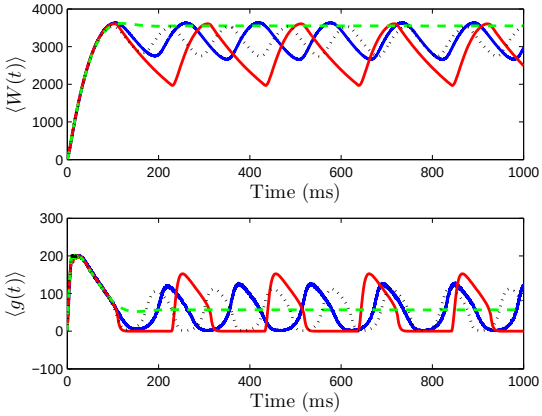
(a) $\langle I \rangle = 5000$ pA, $\sigma_I = 2000$ pA



(b) $\langle I \rangle = 3000$ pA, $\sigma_I = 200$ pA

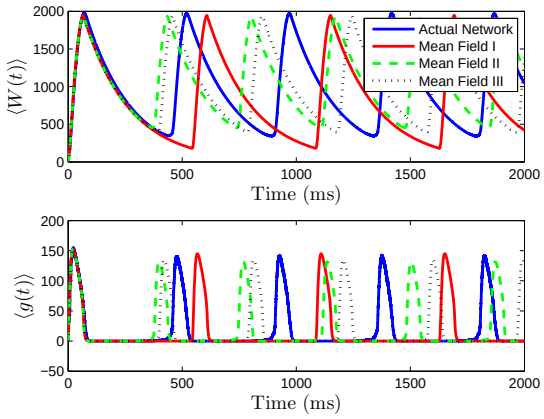


(c) $\langle I \rangle = 5000$ pA, $\sigma_I = 500$ pA

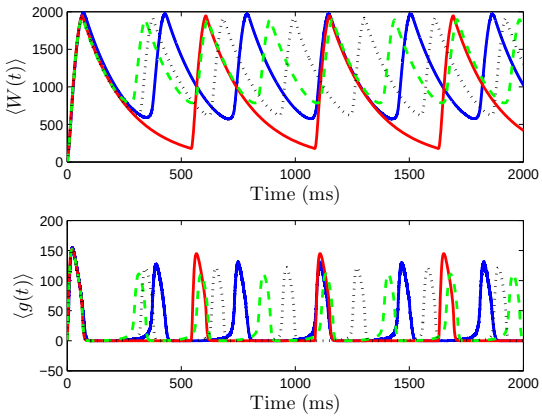


(d) $\langle I \rangle = 3000$ pA, $\sigma_I = 2000$ pA

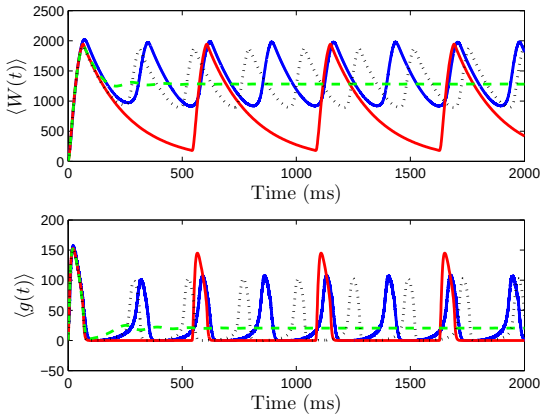
Figure 5.1: Numerical simulations of a network of 1000 Izhikevich neurons with parameters as in Table 5.1, except $g_{syn} = 200$ and the applied current which is normally distributed with mean and variance as shown. Blue is the network average of a given variable, red is MFI, green is MFII and black is MFIII. In this region, the mean-driving current is away from rheobase, $\langle I_{app} \rangle \gg I_{rh}$. All three approximations are quantitatively and qualitatively similar for small to intermediate sized variances in the distribution of currents. For small variances, MFI is the most accurate and for larger variances, MFIII is the most accurate. For large variance, MFII bifurcates back to tonic firing earlier than MFI and MFIII, as seen in (d)



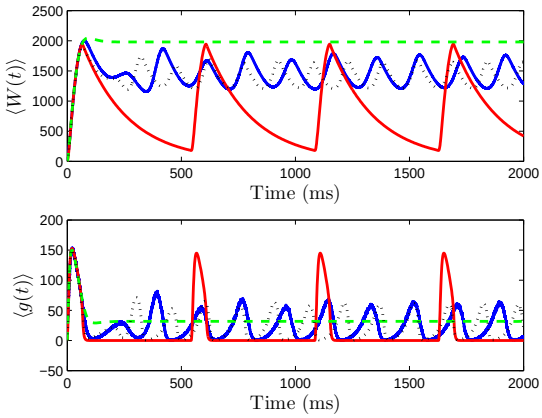
(a) $\langle I \rangle = 1200$ pA, $\sigma_I = 200$ pA



(b) $\langle I \rangle = 1200$ pA, $\sigma_I = 500$ pA

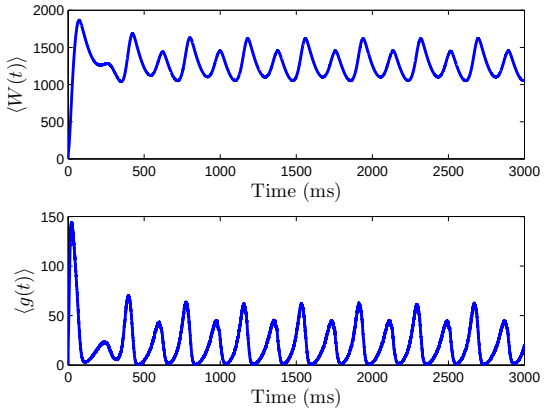


(c) $\langle I \rangle = 1200$ pA, $\sigma_I = 1000$ pA

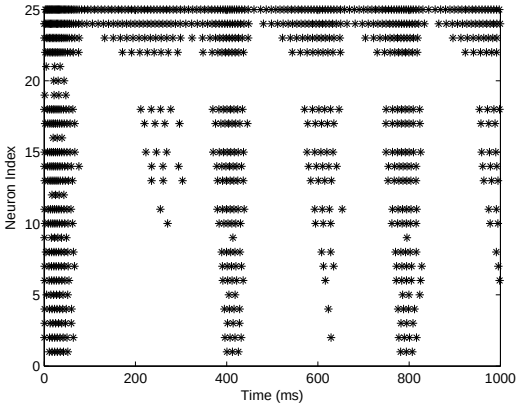


(d) $\langle I \rangle = 1200$ pA, $\sigma_I = 2000$ pA

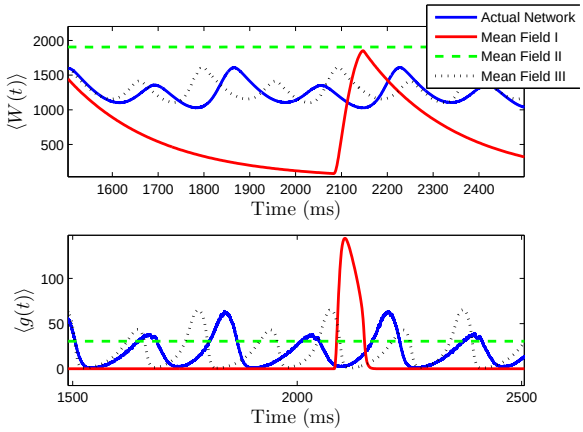
Figure 5.2: Numerical simulations of a network of 1000 neurons with parameters as in Table 5.1, except $g_{syn} = 200$ and the applied current which is normally distributed with mean and variance as shown. Blue is the network average of a given variable, red is MFI, green is MFII and black is MFIII. In these simulations, the mean-driving current is close to (and over) the rheobase. In all cases, MFI is the least accurate. This is because it depends only on $\langle I_{app} \rangle$. When $\langle I_{app} \rangle = O(I_{rh})$, even for small variance, many of the neurons have $I < I_{rh}$ and may not spike at all. (a),(b) For small values of σ_I, all three approximations are qualitatively and quantitatively accurate. (c),(d) For larger variance, $\sigma_I = O(I_{rh})$, only MFIII is qualitatively and quantitatively accurate. In this case, MFII bifurcates early to tonic firing.



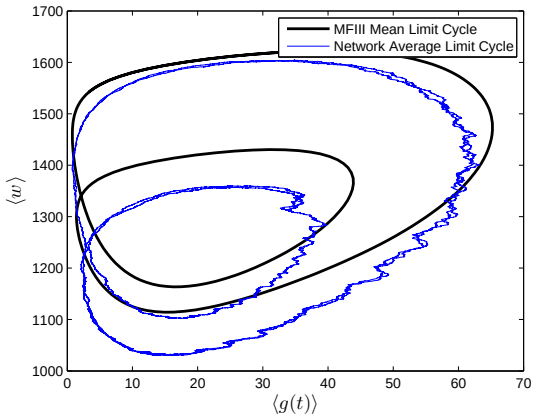
(a) $\langle I_{app} \rangle = 1100$ pA, $\sigma_I = 2000$ pA



(b) Raster plot of the simulation



(c) Mean-Field Equations at Same Parameter Values



(d) Steady state period-doubled limit cylce

Figure 5.3: Period doubled limit cycle in the heterogeneous network and in MFIII. The network consists of 5000 neurons, with parameters as in Table 5.1, except $g_{syn} = 200$ and the applied current which is normally distributed with mean and variance as shown in (a). (a) period-doubled limit cycle for the network shown in terms of the mean variables. (b) raster plot of 25 randomly selected neurons of the network arranged in order of increasing current. The behaviors include burst firing, alternate burst firing, tonic firing, and quiescence. (c) numerical simulations of the mean-field systems. Only MFIII is able to reproduce the period doubling behavior. (d) Comparison of the "phase portrait" of period doubled limit cycle for MFIII and the mean variables of network.

with an added assumption we can approximate the distribution of steady state firing rates for the network with a great degree of accuracy.

Consider a network with just one heterogeneous parameter, β. Assume that the steady state firing rate of each neuron in the network can be related to its value for the heterogeneous parameter: $R_i = g(\beta)$. Assume further that one can approximate this function by the steady state value of $\langle R_i(t)|\beta\rangle$:

$$g(\beta) \approx \langle R_i|\beta\rangle. \tag{5.40}$$

This is easily determined through direct simulation of MFIII, (5.36)-(5.38), until the system reaches steady state. Treating g as the transformation of a random variable, one can determine the steady state distribution of firing rates in the network, $\rho_R(r)$, through the transformation of random variables:

$$\rho_R(r) = \rho_\beta(g^{-1}(r)) \left| \frac{d}{dr} g^{-1}(r) \right|, \tag{5.41}$$

which can be found in any standard textbook on probability theory. Note that we must assume that $\langle R_i|\beta\rangle$ is monotonic and thus invertible for this procedure to be valid. Otherwise, the formula for the distribution function is more complicated.

We carried out this computation for a network of 1000 neurons with various distributions in either I, g, or w_{jump}. We numerically determined the distribution of steady state firing rates for the neurons in the full network through

$$R_i = \frac{1}{ISI_{i,last}}, \quad i = 1, 2, \ldots N \tag{5.42}$$

In these examples, bi-modal distributions in I, g, or w_{jump}, were generated by mixing normal unimodal distributions .This is one way of representing a network with two subpopulations of neurons with different parameters. The mean field approach gives an excellent approximation to the qualitative and quantitative properties of the steady state distribution of firing rates, as shown in the right column of Figure 6.8.

Many parameters for neuron models are difficult to measure directly using electrophysiology. However, a distribution of firing rates across a network of neurons is relatively easy to measure using intracellular recordings, or can be estimated using measurements from multi-electrode recordings and spike sorting algorithms, among other methods [29, 72]. We have seen above that, given a distribution of heterogeneities, MFIII can predict the steady state distribution of firing rates. Here we show that one can invert this process to yield a distribution of parameters given a steady state distribution of firing rates.

We assume that only the firing rate distribution is known, and denote it $\rho_R(r)$ as above. We then proceed as above, assuming that the steady state firing rate for a particular neuron is some function of the heterogeneous parameters $R_i = g(\beta)$ and that this function is well

approximated by $\langle R_i|\beta\rangle$. Under these assumptions, one can solve for the distribution of parameters β using

$$\rho_\beta(\beta) = \rho_R(g(\beta)) \left| \frac{d}{d\beta} g(\beta) \right| \tag{5.43}$$

which follows from standard statistical theorems on the transformations of random variables. Note that we need to assume that $\langle R_i|\beta\rangle$ is differentiable for this procedure to be valid. Note that this also allows us to perform an estimate of the parameter for any individual neuron given knowledge of its steady state firing rate, R_i via $\beta_i = g^{-1}(\beta_i)$

The primary problem we face in using this approach to approximate the distribution $\rho_\beta(\beta)$ is that we need to determine the steady state values of the function $\langle R_i|\beta\rangle$. However, a cursory look at the equations for MFIII shows that these in fact depend on $\rho_\beta(\beta)$, the function we are trying to find, through the equation for s:

$$\dot{s} = -\frac{s}{\tau_s} + s_{jump} \int_\beta \langle R_i(t)|\beta\rangle \rho_\beta(\beta)\, d\beta. \tag{5.44}$$

Fortunately, however, this problem disappears when we look at the steady state value for s:

$$\bar{s} = \tau_s s_{jump} \int_\beta \langle R_i|\beta\rangle \rho_\beta(\beta)\, d\beta = \tau_s s_{jump} \langle R_i\rangle. \tag{5.45}$$

Here $\langle R_i\rangle$ is the unconditioned steady state mean of the firing rate distribution. This information is readily available, as we have assumed we know the steady state distribution, $\rho_R(r)$, and determining the first moment is numerically trivial.

Putting the expression for $\bar{s}$ into the steady state equation for $\langle w|\beta\rangle$ yields a set of coupled equations:

$$\langle w|\beta\rangle = \tau_w w_{jump} \langle R_i|\beta\rangle, \tag{5.46}$$

$$\langle R_i|\beta\rangle = \begin{cases} \left[\int_V \frac{dv}{G_1(v,\langle w|\beta\rangle,\tau_s s_{jump}\langle R_i(t)\rangle,\beta)}\right]^{-1} & : H(\langle w|\beta\rangle, \bar{s}, \beta) \geq 0 \\ 0 & H(\langle w|\beta\rangle, \bar{s}, \beta) < 0 \end{cases}. \tag{5.47}$$

These may be solved for $\langle w|\beta\rangle$ and $\langle R_i|\beta\rangle$ by discretizing in β and numerically solving the resulting system at each grid point with any standard root finding algorithm.

Alternatively, one can set s to its equilibrium value in MFIII and numerically integrate the resulting equation:

$$\langle w|\beta\rangle' = -a\langle w|\beta\rangle + w_{jump}\langle R_i(t)|\beta\rangle, \tag{5.48}$$

$$\langle R_i(t)|\beta\rangle = \begin{cases} \left[\int_V \frac{dv}{G_1(v,\langle w|\beta\rangle,\tau_s s_{jump}\langle R_i(t)\rangle,\beta)}\right]^{-1} & : H(\langle w|\beta\rangle, s, \beta) \geq 0 \\ 0 & H(\langle w|\beta\rangle, s, \beta) < 0 \end{cases} \tag{5.49}$$

until it reaches steady state, which will determine $\langle w|\beta\rangle$ and $\langle R_i|\beta\rangle$. Note that this approach will only work if the tonic firing equilibrium of the original mean-field system MFIII is asymptotically stable.

We have implemented this approach as follows. A network of 1000 neurons is numerically integrated until it reaches its steady state firing rate. The distribution of firing rates over the network is found as described in above. The density function for this distribution, $\rho_R(r)$, is then estimated using the firing rate histogram. Equations (5.48)-(5.49) are numerically integrated until they reach steady state. We then substitute the estimate of $\rho_R(r)$ and the approximation $\langle R_i|\beta\rangle$ of $g(\beta)$ into (5.43) to determine the parameter distribution $\rho_\beta(\beta)$. The results are shown in 6.8. In the right column of each figure, the solid blue curve is the distribution of steady state firing rates from integration of the full network. In the left column of each figure the dashed red curve is the estimate of $\rho_\beta(\beta)$ found using the procedure above, while the blue curve is the actual parameter distribution used in the network simulation. We note that no information about the distribution of parameters is known in the estimation procedure, yet the numerical results are very accurate.

While network level inversion of a single heterogeneous parameter is an important step forward, this is performed under very strong assumptions. In particular, when performing this inversion, all of the heterogeneity in the firing rates is assumed to come from a single parameter. Additionally, all the other parameters are assumed to be known. These two assumptions are exceptionally strong and one has to take great care that they be reasonably satisfied before inverting actual recorded firing rates from neurons. In general, the former assumption is seldom appropriate as all the parameters are physical quantities that vary due to the measurement process alone, let alone the natural variability of these physical quantities. However, this is only a proof of principle attempt to show that it is possible to obtain information about the neuronal parameters with a mean-field system.

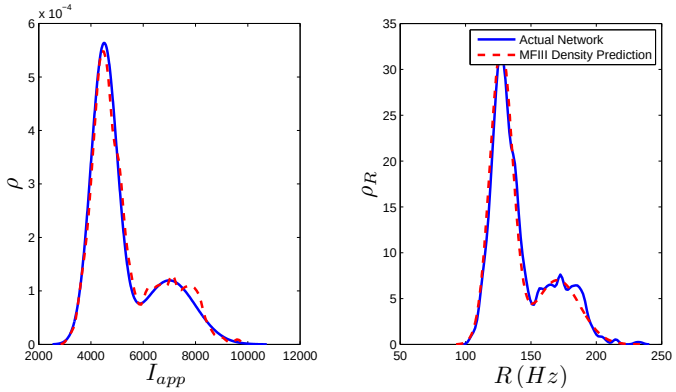
5.1.3 Inverse Mean-Field Theory

It is clear that heterogeneity alters the dynamics of a neuronal network. For example, we have shown that heterogeneity alone can induce the transition from tonic firing to bursting and vice versa, in addition to adding new dynamics, like period doubling bifurcations. Thus, one may ask the question, given a particular source of heterogeneity, say for example in the applied current I_{app}, what are the potential mean-field dynamic s of the network? More formally, using MFII as an example, given the mean-field system

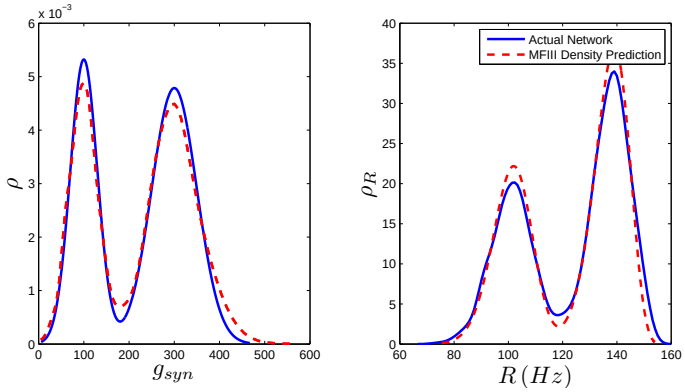
$$\langle w\rangle' = \frac{b\langle v\rangle - \langle w\rangle}{\tau_w} + \int_{\boldsymbol{\beta}} w_{jump}\langle R_i(t)|\boldsymbol{\beta}\rangle\rho_\beta(\boldsymbol{\beta})\,d\boldsymbol{\beta} = F(s,w) \tag{5.50}$$

$$s' = -\frac{s}{\tau_s} + s_{jump}\int_{\boldsymbol{\beta}}\langle R_i(t)|\boldsymbol{\beta}\rangle\rho_\beta(\boldsymbol{\beta})\,d\boldsymbol{\beta} = G(s,w), \tag{5.51}$$

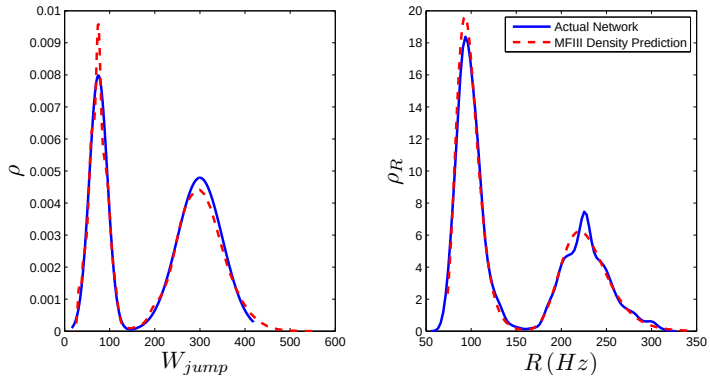
what functions $F(s,w)$ and $G(s,w)$, and thus potential dynamics, are possible using a particular heterogeneous distribution $\rho_\beta(\boldsymbol{\beta})$? The answer to this question is highly non-trivial as it concerns the representation space of very specific linear operators which we will discuss in Chapter 7. Additionally, as we will later see, to achieve any dynamics for our mean-field system, we require a certain degree of heterogeneity in the connectivity between



(a) Distribution of I_{app} with $\mu_1 = 4500$ pA, $\sigma_1 = 500$ pA, $\mu_2 = 7000$ pA, $\sigma_2 = 1000$ pA, $m = 0.7$



(b) Distribution of g_{syn} with $\mu_1 = 100$ nS, $\sigma_1 = 30$ nS, $\mu_2 = 300$ nS, $\sigma_2 = 50$ nS, $m = 0.4$



(c) Distribution of W_{jump} with $\mu_1 = 300$ pA, $\sigma_1 = 50$ pA, $\mu_2 = 75$ pA, $\sigma_2 = 20$ pA, $m = 0.6$

Figure 5.4: Bimodal distributions in I_{app}, g_{syn}, and W_{jump} lead to bimodal distributions in the firing rate. These bimodal parameter distributions are generated through distribution mixing of two normal subpopulations with standard deviations and means as indicated. The distribution of the firing rate or the distribution of the parameter can be computed using MFIII if one knows the complementary distribution. The calculations were carried out on a network of 1000 neurons.

neurons. For now however, we will merely state that given a network of integrate and fire neurons with a source of heterogeneity, $\rho_\beta(\boldsymbol{\beta})$, one can define a series of connection weights $\omega_{ij} = \omega(\boldsymbol{\beta}_i, \boldsymbol{\beta}_j)$ that couple the presynaptic neuron i to the post-synaptic neuron j such that arbitrary (smooth) dynamics are possible. Additionally, it turns out that the weights, ω_{ij} are non-unique. Using the exact same heterogeneous network, one can create infinitely many weight matrices ω_{ij} such that the macroscopic dynamics are identical. Furthermore, for specific types of firing rates, these weight matrices are analytically solvable. This will all be explored further in Chapter 7.

5.1.4 Spectral Analysis of the the Population Density Equation

Consider the moment-closure reduced population density equation that leads to the derivation of MFII:

$$\frac{\partial \rho_V(v,t|\beta)}{\partial t}\rho_\beta(\beta) = -\frac{\partial}{\partial v}\left(G_1(v,w,s,\beta)\rho_V(v,t|\beta)\rho_\beta(\beta)\right) \tag{5.52}$$

$$s' = -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s}\int_\beta G_1(v_{peak},w,s,\beta)\rho_V(v_{peak},t|\beta)\rho_\beta(\beta)\,d\beta \tag{5.53}$$

$$w' = -\frac{w}{\tau_w} + \frac{\lambda_w}{\tau_w}\int_\beta G_1(v_{peak},w,s,\beta)\rho_V(v_{peak},t|\beta)\rho_\beta(\beta)\,d\beta. \tag{5.54}$$

We will prove that the eigenvalues for the mean-field system MFII are a subset of the eigenvalues of the full spectral equation for this system. This will be similar to the work done in [76], which also follows closely from the original work in [2]. Unfortunately, solving for the infinite dimensional set of eigenvalues is more difficult than in Section 4.4, however the fast system is still analytically resolvable. The steady state of this system is given by the equations

$$G_1(v,w,s,\beta)\rho_V(v|\beta)\rho_\beta(\beta) = J(w,s,\beta) \tag{5.55}$$

$$J(\beta,s,w) = \rho_\beta(\beta)\left(\int_{v_{reset}}^{v_{peak}}\frac{dv}{G_1(v',w,s,\beta)}\,dv'\right)^{-1} \tag{5.56}$$

$$\rho(v,t|\beta) = \left(\int_{v_{reset}}^{v_{peak}}\frac{dv}{G_1(v',\bar{w},\bar{s},\beta)}\,dv'\right)^{-1}\frac{1}{G_1(v,\bar{w},\bar{s},\beta)} \tag{5.57}$$

$$\bar{s} = \lambda_s\overline{\langle \mathrm{R}\rangle}, \quad \bar{w} = \lambda_w\overline{\langle \mathrm{R}\rangle} \tag{5.58}$$

$$\overline{\langle \mathrm{R}\rangle} = \int_\beta J(\beta,\lambda_s\overline{\langle \mathrm{R}\rangle},\lambda_w\overline{\langle \mathrm{R}\rangle}) \tag{5.59}$$

To begin we collapse the conditioning so that we have $\rho(v,\beta,t) = \rho_V(v,t|\beta)\rho_\beta(\beta)$. Furthermore, consider the distribution function

$$P(v,\beta,t) = \int_{v_{reset}}^{v}\rho(v',\beta,t)\,dv' \tag{5.60}$$

$$P(v_{peak},\beta,t) = \rho_\beta(\beta), \quad P(v_{reset},\beta,t) = 0 \tag{5.61}$$

We will use the quantity $P(v,\beta,t)$ to extend the spectral analysis that we conducted in section 4.4 for the homogeneous system to the system with heterogeneity. To do this, we need to modify the Abbott/Vreeswijk transform. In particular, consider the following

$$y = \eta(v,\beta) = \frac{J(\lambda_w \overline{\langle \mathrm{R} \rangle}, \lambda_s \overline{\langle \mathrm{R} \rangle}, \beta)}{\rho_\beta(\beta)} \int_{v_{reset}}^{v} \frac{dv'}{G_1(v', \lambda_w \overline{\langle \mathrm{R} \rangle}, \lambda_s \overline{\langle \mathrm{R} \rangle}, \beta)}. \tag{5.62}$$

For ease of notation we will relabel the steady state $J(\lambda_w \overline{\langle \mathrm{R} \rangle}, \lambda_s \overline{\langle \mathrm{R} \rangle}, \beta)$ as $J(\beta)$. Then $\eta(v_{reset},\beta) = 0$, and $\eta(v_{peak},\beta) = 1$, which should be clear from equation (5.56). Working with $P(v,\beta,t)$, then we have the following partial differential equation

$$\frac{\partial P(v,\beta,t)}{\partial t} = -G_1(v,w,s,\beta)\frac{\partial P}{\partial v} + G_1(v_{peak},w,s,\beta)\frac{\partial P}{\partial v}\bigg|_{v=v_{peak}} \tag{5.63}$$

$$\tag{5.64}$$

Applying the transform 5.62 and defining $\hat{P}(y,\beta,t) = P(\eta^{-1}(y,\beta),\beta,t)$:

$$\begin{aligned}
\frac{\partial \hat{P}(y,\beta,t)}{\partial t} &= \frac{\partial P(\eta^{-1}(y,\beta),\beta,t)}{\partial t} = -\frac{G_1(\eta^{-1}(y,\beta),w,s,\beta)}{G_1(\eta^{-1}(y,\beta),\lambda_w \overline{\langle \mathrm{R} \rangle}, \lambda_s \overline{\langle \mathrm{R} \rangle},\beta)} \frac{J(\beta)}{\rho_\beta(\beta)} \frac{\partial \hat{P}}{\partial y} \\
&+ \frac{G_1(v_{peak},w,s,\beta)}{G_1(v_{peak},\lambda_w \overline{\langle \mathrm{R} \rangle}, \lambda_s \overline{\langle \mathrm{R} \rangle},\beta)} \frac{J(\beta)}{\rho_\beta(\beta)} \frac{\partial \hat{P}}{\partial y}\bigg|_{y=1} \\
s' &= -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s} \int_\beta \frac{G_1(v_{peak},w,s,\beta)}{G_1(v_{peak},\lambda_w \overline{\langle \mathrm{R} \rangle}, \lambda_s \overline{\langle \mathrm{R} \rangle},\beta)} \frac{J(\beta)}{\rho_\beta(\beta)} \frac{\partial \hat{P}}{\partial y}\bigg|_{y=1} d\beta \\
w' &= -\frac{w}{\tau_w} + \frac{\lambda_w}{\tau_w} \int_\beta \frac{G_1(v_{peak},w,s,\beta)}{G_1(v_{peak},\lambda_w \overline{\langle \mathrm{R} \rangle}, \lambda_s \overline{\langle \mathrm{R} \rangle},\beta)} \frac{J(\beta)}{\rho_\beta(\beta)} \frac{\partial \hat{P}}{\partial y}\bigg|_{y=1} d\beta
\end{aligned}$$

we will drop the hat from P from now on for convenience. Given the steady states (5.55)-(5.59), in addition to the boundary conditions on $P(v,\beta,t)$ given by equation (5.61), consider the perturbations

$$\epsilon_y(y,\beta,t) = P(y,\beta,t) - y\rho_\beta(\beta) \tag{5.65}$$

$$\epsilon_s(t) = s - \lambda_s \overline{\langle \mathrm{R} \rangle}, \quad \epsilon_w(t) = w - \lambda_w \overline{\langle \mathrm{R} \rangle} \tag{5.66}$$

One can easily double check that the boundary condition on $\epsilon_y(y,\beta,t)$ becomes $\epsilon_y(1,\beta,t) = \epsilon_y(0,\beta,t) = 0$ and that the steady states are indeed valid. After making the substitution

we have the following:

$$
\begin{aligned}
\frac{\partial \epsilon_y(y,\beta,t)}{\partial t} &= -\frac{G_1(\eta^{-1}(y,\beta),\epsilon_w+\lambda_w\overline{\langle \mathrm{R}\rangle},\epsilon_s+\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}{G_1(\eta^{-1}(y,\beta),\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\frac{J(\beta)}{\rho_\beta(\beta)}\left(\frac{\partial \epsilon_y}{\partial y}+\rho_\beta(\beta)\right)\\
&+ \frac{G_1(v_{peak},\epsilon_s+\lambda_s\overline{\langle \mathrm{R}\rangle},\epsilon_w+\lambda_w\overline{\langle \mathrm{R}\rangle},\beta)}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\frac{J(\beta)}{\rho_\beta(\beta)}\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+\rho_\beta(\beta)\right)\\
&= -\frac{J(\beta)}{\rho_\beta(\beta)}\left(1+\frac{g\epsilon_s(\eta^{-1}(y,\beta)-e_r)-\epsilon_w}{G_1(\eta^{-1}(y,\beta),\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\left(\frac{\partial \epsilon_y}{\partial y}+\rho_\beta(\beta)\right)\\
&+ \frac{J(\beta)}{\rho_\beta(\beta)}\left(1+\frac{g\epsilon_s(v_{peak}-e_r)-\epsilon_w}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+\rho_\beta(\beta)\right)\\
&= \hat{L}_1(\epsilon_y,\epsilon_s,\epsilon_w)\\
\epsilon_s' &= -\frac{\epsilon_s}{\tau_s}-\frac{\lambda_s\overline{\langle \mathrm{R}\rangle}}{\tau_s}+\frac{\lambda_s}{\tau_s}\int_\beta \frac{J(\beta)}{\rho_\beta(\beta)}\left(1+\frac{g\epsilon_s(v_{peak}-e_r)-\epsilon_w}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+\rho_\beta(\beta)\right)d\beta\\
\epsilon_s' &= -\frac{\epsilon_s}{\tau_s}\\
&+ \frac{\lambda_s}{\tau_s}\int_\beta\left\{\frac{J(\beta)}{\rho_\beta(\beta)}\left(\frac{g\epsilon_s(v_{peak}-e_r)-\epsilon_w}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+\rho_\beta(\beta)\right)+\frac{J(\beta)}{\rho_\beta(\beta)}\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}\right\}d\beta\\
&= \hat{L}_2(\epsilon_y,\epsilon_s,\epsilon_w)\\
\epsilon_w' &= -\frac{\epsilon_w}{\tau_w}\\
&+ \frac{\lambda_w}{\tau_w}\int_\beta\left\{\frac{J(\beta)}{\rho_\beta(\beta)}\left(\frac{g\epsilon_s(v_{peak}-e_r)-\epsilon_w}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+\rho_\beta(\beta)\right)+\frac{J(\beta)}{\rho_\beta(\beta)}\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}\right\}d\beta\\
&= \hat{L}_3(\epsilon_y,\epsilon_s,\epsilon_w)
\end{aligned}
$$

As in section 4.4.2, we will analyze the operator

$$
\hat{L}(\epsilon_y(y,\beta,t),\epsilon_s,\epsilon_w)=\begin{pmatrix}\hat{L}_1(\epsilon_y(y,\beta),\epsilon_s,\epsilon_w)\\ \hat{L}_2(\epsilon_y(y,\beta),\epsilon_s,\epsilon_w)\\ \hat{L}_3(\epsilon_y(y,\beta),\epsilon_s,\epsilon_w)\end{pmatrix} \tag{5.67}
$$

We will again deal with the product of 3 Banach spaces. Defining $\Omega_1=\mathbb{L}_2([0,1]\times(-\infty,\infty))$ then the space we will consider is $\Omega_1\times\mathbb{R}^2$. As before, we need to resolve the Fréchet derivative at $\boldsymbol{\epsilon}=0$, It should be clear that the linearization at $\boldsymbol{\epsilon}=0$ with $\boldsymbol{h}=(h_1(y,\beta),h_2,h_3)$,

given by $\boldsymbol{D}\hat{\boldsymbol{L}}(\mathbf{0})\boldsymbol{h}$ is:

$$\begin{aligned}
\boldsymbol{D}_1\hat{\boldsymbol{L}}(\mathbf{0})\boldsymbol{h} &= -\frac{J(\beta)}{\rho_\beta(\beta)}\left(\frac{\partial h_1(y,\beta)}{\partial y} - \left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1}\right) \\
&- J(\beta)\left(\frac{gh_2(e_r-\eta^{-1}(y))-h_3}{G_1(\eta^{-1}(y),\lambda_s\overline{\langle \mathrm{R}\rangle},\lambda_w\overline{\langle \mathrm{R}\rangle},\beta)} - \frac{gh_2(e_r-\eta^{-1}(1))-h_3}{G_1(\eta^{-1}(1),\lambda_s\overline{\langle \mathrm{R}\rangle},\lambda_w\overline{\langle \mathrm{R}\rangle},\beta)}\right) \\
\boldsymbol{D}_2\hat{\boldsymbol{L}}(\mathbf{0})\boldsymbol{h} &= -\frac{h_2}{\tau_s} + \frac{\lambda_s}{\tau_s}\int_\beta\left\{\frac{J(\beta)}{\rho_\beta(\beta)}\left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1} + J(\beta)\left(\frac{gh_2(v_{peak}-e_r)-h_3}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\right\}d\beta \\
\boldsymbol{D}_3\hat{\boldsymbol{L}}(\mathbf{0})\boldsymbol{h} &= -\frac{h_3}{\tau_w} + \frac{\lambda_w}{\tau_w}\int_\beta\left\{\frac{J(\beta)}{\rho_\beta(\beta)}\left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1} + J(\beta)\left(\frac{gh_2(v_{peak}-e_r)-h_3}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\right\}d\beta
\end{aligned}$$

and the resulting eigenvalue problem, $\boldsymbol{D}\hat{L}(\mathbf{0})\boldsymbol{h} = \mu\boldsymbol{h}$ yields the following:

$$\begin{aligned}
\mu h_1(y,\beta) &= -\frac{J(\beta)}{\rho_\beta(\beta)}\left(\frac{\partial h_1(y,\beta)}{\partial y} - \left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1}\right) \\
&- J(\beta)\left(\frac{gh_2(e_r-\eta^{-1}(y))-h_3}{G_1(\eta^{-1}(y),\lambda_s\overline{\langle \mathrm{R}\rangle},\lambda_w\overline{\langle \mathrm{R}\rangle},\beta)} - \frac{gh_2(e_r-\eta^{-1}(1))-h_3}{G_1(\eta^{-1}(1),\lambda_s\overline{\langle \mathrm{R}\rangle},\lambda_w\overline{\langle \mathrm{R}\rangle},\beta)}\right) \\
\mu h_2 &= -\frac{h_2}{\tau_s} + \frac{\lambda_s}{\tau_s}\int_\beta\left\{\frac{J(\beta)}{\rho_\beta(\beta)}\left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1} + J(\beta)\left(\frac{gh_2(v_{peak}-e_r)-h_3}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\right\}d\beta \\
\mu h_3 &= -\frac{h_3}{\tau_w} + \frac{\lambda_w}{\tau_w}\int_\beta\left\{\frac{J(\beta)}{\rho_\beta(\beta)}\left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1} + J(\beta)\left(\frac{gh_2(v_{peak}-e_r)-h_3}{G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle},\beta)}\right)\right\}d\beta
\end{aligned}$$

Using our functions $A(y,\beta)$ and $B(y,\beta)$ as in Chapter 2, only they now also depend on β, then once again we have the following:

$$\begin{aligned}
\mu h_1(y,\beta) &= -\frac{J(\beta)}{\rho_\beta(\beta)}\left(\frac{\partial h_1(y,\beta)}{\partial y} - \left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1}\right) \\
&- J(\beta)(h_2A(y,\beta)+h_3B(y,\beta)-h_2A(1,\beta)-h_3B(1,\beta)) \qquad (5.68) \\
\mu h_2 &= -\frac{h_2}{\tau_s} + \frac{\lambda_s}{\tau_s}\int_\beta\left\{\frac{J(\beta)}{\rho_\beta(\beta)}\left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1} + J(\beta)(h_2A(1,\beta)+h_3B(1,\beta))\right\}d\beta \qquad (5.69) \\
\mu h_3 &= -\frac{h_3}{\tau_w} + \frac{\lambda_w}{\tau_w}\int_\beta\left\{\frac{J(\beta)}{\rho_\beta(\beta)}\left.\frac{\partial h_1(y,\beta)}{\partial y}\right|_{y=1} + J(\beta)(h_2A(1,\beta)+h_3B(1,\beta))\right\}d\beta \qquad (5.70)
\end{aligned}$$

One can solve (5.68) as in Chapter 2, applying the conditions $h_1(1,\beta) = h_1(0,\beta) = 0$ to yield the following equation for $\left.\frac{\partial h_1}{\partial y}\right|_{y=1}$

$$\begin{aligned}\left.\frac{\partial h_1}{\partial y}\right|_{y=1} &= \left(\int_0^1 \exp\left(\mu y' \frac{\rho_\beta(\beta)}{J(\beta)}\right)\rho_\beta(\beta)(h_2 A(y') + h_3 B(y'))\,dy'\right)\left(\exp\left(\mu\frac{\rho_\beta(\beta)}{J(\beta)}\right)-1\right)^{-1}\frac{\rho_\beta(\beta)\mu}{J(\beta)} \\ &- \rho_\beta(\beta)(h_2 A(1) + h_3 B(1))\end{aligned}$$

Substituting this into equations (5.69)-(5.70) yields the following:

$$\begin{aligned}\mu h_2 &= -\frac{h_2}{\tau_s} + \frac{\mu\lambda_s}{\tau_s}\int_\beta\left\{\left(\int_0^1 \exp\left(\mu y'\frac{\rho_\beta(\beta)}{J(\beta)}\right)\frac{\rho_\beta(\beta)(h_2A(y')+h_3B(y'))}{\left(\exp\left(\mu\frac{\rho_\beta(\beta)}{J(\beta)}\right)-1\right)}\,dy'\right)\right\}d\beta \\ \mu h_3 &= -\frac{h_3}{\tau_w} + \frac{\mu\lambda_w}{\tau_w}\int_\beta\left\{\left(\int_0^1 \exp\left(\mu y'\frac{\rho_\beta(\beta)}{J(\beta)}\right)\frac{\rho_\beta(\beta)(h_2A(y')+h_3B(y'))}{\left(\exp\left(\mu\frac{\rho_\beta(\beta)}{J(\beta)}\right)-1\right)}\,dy'\right)\right\}d\beta\end{aligned}$$

Now, as we again want non-trivial eigenspaces, that is we want to remove the solution $h_2 = h_3 = 0$, then we require the determinant of the resulting linear system to be non-zero. This yields the following spectral equation

$$\begin{aligned}0 &= \left(\mu+\frac{1}{\tau_s}\right)\left(\mu+\frac{1}{\tau_w}\right) - \frac{\mu\lambda_w}{\tau_w}\left(\mu+\frac{1}{\tau_s}\right)\int_\beta \hat{B}(\mu,\beta)\left(\exp\left(\mu\frac{\rho_\beta(\beta)}{J(\beta)}\right)-1\right)^{-1}\rho_\beta(\beta)\,d\beta \\ &- \mu\frac{\lambda_s}{\tau_s}\left(\mu+\frac{1}{\tau_w}\right)\int_\beta \hat{A}(\mu,\beta)\left(\exp\left(\mu\frac{\rho_\beta(\beta)}{J(\beta)}\right)-1\right)^{-1}\rho_\beta(\beta)\,d\beta\end{aligned}$$

where $\hat{A}(\mu,\beta)$ and $\hat{B}(\mu,\beta)$ are defined as in Chapter 2, only with the β dependence made explicit. We will again look for perturbation solutions to this spectral equation in the mean-field limit, $\tau_w^{-1} = \epsilon$, $\tau_s^{-1} = \epsilon/\gamma$.

$$\begin{aligned}0 &= \left(\mu+\frac{\epsilon}{\gamma}\right)(\mu+\epsilon) - \epsilon\mu\lambda_w\left(\mu+\frac{\epsilon}{\gamma}\right)\int_\beta \hat{B}(\mu,\beta)\left(\exp\left(\mu\frac{\rho_\beta(\beta)}{J(\beta)}\right)-1\right)^{-1}\rho_\beta(\beta)\,d\beta \\ &- \epsilon\mu\frac{\lambda_s}{\gamma}(\mu+\epsilon)\int_\beta \hat{A}(\mu,\beta)\left(\exp\left(\mu\frac{\rho_\beta(\beta)}{J(\beta)}\right)-1\right)^{-1}\rho_\beta(\beta)\,d\beta\end{aligned}$$

The order $\epsilon = 0$ solution immediately implies that to lowest order in ϵ, all the eigenvalues vanish. We can thus look for perturbation solutions of the form $\mu = \mu_1\epsilon$. This yields the following

$$\begin{aligned}0 &= \epsilon^2\left(\mu_1+\frac{1}{\gamma}\right)(\mu+\mu_1) - \epsilon^2\lambda_w\left(\mu_1+\frac{1}{\gamma}\right)\int_\beta \hat{B}(0,\beta)J(\beta)d\,\beta \\ &- \epsilon^2\frac{\lambda_s}{\gamma}(\mu_1+1)\int_\beta \hat{A}(0,\beta)J(\beta)d\beta + O(\epsilon^3)\end{aligned} \tag{5.71}$$

Now, we will need to compute the quantities $\hat{B}(0,\beta)$ and $\hat{A}(0,\beta)$. This yields the following

$$\begin{aligned}\hat{A}(0,\beta) &= \int_0^1 \frac{g(e_r-\eta^{-1}(y,\beta)}{G_1(\eta^{-1}(y,\beta),\lambda_w\overline{\langle R\rangle},\lambda_s\overline{\langle R\rangle},\beta)}\,dy' \\ &= \frac{J(\beta)}{\rho_\beta(\beta)}\int_{v_{reset}}^{v_{peak}} \frac{g(e_r-v)}{G_1(v,\lambda_w\,\overline{\langle \mathrm{R}\rangle},\lambda_s\,\overline{\langle \mathrm{R}\rangle},\beta)^2}\,dv\end{aligned}$$

.

This implies that

$$\begin{aligned}&\Rightarrow \int_\beta \hat{A}(0,\beta)J(\beta)\,d\beta = \int_\beta \frac{J(\beta)^2}{\rho_\beta(\beta)}\int_{v_{reset}}^{v_{peak}} \frac{g(e_r-v)}{G_1(v,\lambda_w\,\overline{\langle \mathrm{R}\rangle},\lambda_s\,\overline{\langle \mathrm{R}\rangle},\beta)}\,dv d\beta \\ &= \int_\beta \rho_\beta(\beta)\left\{\left(\int_{v_{reset}}^{v_{peak}} \frac{dv'}{G_1(v,\lambda_w\,\overline{\langle \mathrm{R}\rangle},\lambda_s\,\overline{\langle \mathrm{R}\rangle},\beta)}\right)^{-2}\int_{v_{reset}}^{v_{peak}} \frac{g(e_r-v)}{G_1(v,\lambda_w\,\overline{\langle \mathrm{R}\rangle},\lambda_s\,\overline{\langle \mathrm{R}\rangle},\beta)^2}\,dv'\right\}d\beta \\ &= \left.\frac{\partial\langle R_i(t)\rangle(s,w)}{\partial s}\right|_{(\lambda_s R,\lambda_w R)}\end{aligned}$$

with a similar computation holding for $\hat{B}$:

$$\int_\beta \hat{B}(0,\beta)J(\beta) \;=\; \left.\frac{\partial\langle R_i(t)\rangle(s,w)}{\partial w}\right|_{\lambda_s R,\lambda_w R}.$$

Thus, we have the following to lowest order:

$$\begin{aligned}0 \;&=\; \epsilon^2\left(\mu_1+\frac{1}{\gamma}\right)(\mu_1+1) \\ &-\; \epsilon^2\frac{\lambda_s}{\gamma}(\mu_1+1)\left.\frac{\partial\langle R_i(t)\rangle(s,w)}{\partial s}\right|_{\lambda_w\,\overline{\langle \mathrm{R}\rangle},\lambda_s\,\overline{\langle \mathrm{R}\rangle}} - \epsilon^2\lambda_w\left(\mu_1+\frac{1}{\gamma}\right)\left.\frac{\partial\langle R_i(t)\rangle(s,w)}{\partial w}\right|_{\lambda_w\,\overline{\langle \mathrm{R}\rangle},\lambda_s\,\overline{\langle \mathrm{R}\rangle}} + O(\epsilon^3) = 0\end{aligned}$$

but this is the eigenvalue equation for MFII to leading order in ϵ with the substitution $\mu_1\epsilon = \mu$. Thus, the mean-field eigenvalues for the asynchronous steady state(s) are to leading order identical to the mean-field eigenvalues. While there is an infinite dimensional set of eigenvalues as in the homogeneous network case, they are much more difficult to solve for analytically. However, in the section 5.1.5, we can still resolve the fast system to obtain some information about the validity of the mean-field system.

5.1.5 The Fast System

Consider the fast system:

$$\frac{\partial\rho_V(v,t|\beta)}{\partial t}\rho_\beta(\beta) \;=\; -\frac{\partial}{\partial v}\left(G_1(v,s,w,\beta)\rho_V(v|t,\beta)\rho_\beta(\beta)\right) \tag{5.72}$$

$$s' \;=\; w' = 0, \quad s = s_0, \quad w = w_0 \tag{5.73}$$

A spectral analysis of this system will reveal the surprising fact that heterogeneity does not formally stabilize the asynchronous steady state any more then homogeneity does in the sense that the eigenvalues for the fast system are all pure imaginary. Once again if we convert to a PDE for the distribution function, and apply the Abbott/Vreeswijk transform we arrive at the following:

$$\frac{\partial P}{\partial t} = -\frac{J(\beta)}{\rho_\beta(\beta)}\left(\frac{\partial P}{\partial y} - \left.\frac{\partial P}{\partial y}\right|_{y=1}\right) \tag{5.74}$$

where $P(1,t,\beta) = \rho_\beta(\beta)$ and $P(0,t,\beta) = 0$. After using the substitution $\epsilon(y,\beta,t) = P(y,\beta,t) - y\rho_\beta(\beta)$, and eliminating any like terms, this immediately yields the following eigenvalue problem:

$$\mu(\beta)h(y,\beta) = -\frac{J(\beta)}{\rho_\beta(\beta)}\left(h'(y,\beta) - h'(1,\beta)\right) \tag{5.75}$$

The eigenvalues $\mu(\beta)$ can be determined by forcing the condition that $h(y,\beta) \neq 0$:

$$\mu = 2n\pi i\frac{J(\beta)}{\rho_\beta(\beta)}, \quad n = 1,2,\ldots \tag{5.76}$$

$$h(y) = \left(1 - \exp(-2n\pi i y)\right)\frac{1}{2n\pi i} \tag{5.77}$$

Since the system is linear and the eigenvalues are all pure imaginary, the steady state is stable but not asymptotically stable. To get more information, the solution for $\epsilon(y,t)$:

$$\epsilon(y,t) = -\sum_{n=1}^{\infty}\frac{A_n(\beta)}{2n\pi i}\exp\left(2n\pi i\frac{J(\beta)}{\rho_\beta(\beta)}t\right)\left(1 - \exp(-2n\pi i y)\right) + C.C. \tag{5.78}$$

where $C.C.$ denotes the complex conjugate of the first term in the equation. This yields the exact solution for $\rho(v,\beta,t)$ when one undoes all the substitutions:

$$\begin{aligned}\rho(v,\beta,t) &= \frac{J(\beta)}{G_1(v,\beta)} \\ &- \frac{J(\beta)}{\rho_\beta(\beta)G_1(v,\beta)}\left[\sum_{n=1}^{\infty}A_n(\beta)\exp\left(2n\pi i\left(\frac{J(\beta)}{\rho_\beta(\beta)}t - \eta(v,\beta)\right)\right) + C.C.\right]\end{aligned} \tag{5.79}$$

Writing $A_n(\beta) = a_n(\beta) + ib_n(\beta)$ we obtain

$$\begin{aligned}\rho(v,\beta,t) &= \frac{J(\beta)}{G_1(v,\beta)} + \frac{J(\beta)}{\rho_\beta(\beta)G_1(v,\beta)}\sum_{n=1}^{\infty}2a_n(\beta)\cos\left(2n\pi\left(\frac{J(\beta)}{\rho_\beta(\beta)}t - \eta(v,\beta)\right)\right) \\ &- \frac{J(\beta)}{\rho_\beta(\beta)G_1(v,\beta)}\sum_{n=1}^{\infty}2b_n(\beta)\sin\left(2n\pi\left(\frac{J(\beta)}{\rho_\beta(\beta)}t - \eta(v,\beta)\right)\right)\end{aligned} \tag{5.80}$$

If we recall our definition of $J(\beta)$ as

$$J(\beta) = \rho_\beta(\beta)\left(\int_{v_{reset}}^{v_{peak}} \frac{dv}{G(v',\beta)}dv'\right)^{-1} = \rho_\beta(\beta)\langle R|\beta\rangle$$

in addition to our definition of $\eta(v,\beta)$, then we can immediately see that (5.79) satisfies the original fast system, in addition to all the requisite boundary conditions:

$$\begin{aligned}
\int_{v_{reset}}^{v_{peak}} \rho(v,\beta,t) &= \int_{v_{reset}}^{v_{peak}} \frac{J(\beta)}{G_1(v',\beta)}\,dv' \\
&+ \sum_{n=1}^{\infty} A_n(\beta)\exp\left(2n\pi i\frac{J(\beta)}{\rho_\beta(\beta)}t\right)\int_{v_{reset}}^{v_{peak}} \frac{\partial\eta(v,\beta)}{\partial v}\exp(-2n\pi\eta(v,\beta))\,dv + C.C. \\
&= \rho_\beta(\beta) - \sum_{n=1}^{\infty}\frac{A_n(\beta)}{4n^2\pi^2}\exp\left(2n\pi i\frac{J(\beta)}{\rho_\beta(\beta)}t\right)(e^{2n\pi i} - e^0) + C.C. \\
&= \rho_\beta(\beta) \\
\int_\beta\int_v \rho(v,\beta,t)\,dvd\beta &= \int_v\int_\beta \rho(v,\beta,t)\,d\beta dv = 1 \\
J(v_{peak},\beta,t) &= J(v_{reset},\beta,t) = J(\beta) + \frac{J(\beta)}{\rho_\beta(\beta)}\left(\sum_{n=1}^{\infty} A_n(\beta)\exp\left(2n\pi i\frac{J(\beta)}{\rho_\beta(\beta)}t\right) + C.C.\right)
\end{aligned}$$

Now, one has to specify the coefficients of $A_n(\beta)$. In particular, expanding the initial density distribution $\rho_0(v,\beta) = \rho(v,\beta,0)$ as

$$\rho_0(v,\beta) = \frac{J(\beta)}{\rho_\beta(\beta)G_1(v,\beta)}\left(\rho_\beta(\beta) - \sum_{n=1}^{\infty} 2a_n(\beta)\cos(2n\pi\eta(v,\beta)) + 2b_n(\beta)\sin(2n\pi\eta(v,\beta))\right) \tag{5.81}$$

The coefficients can immediately be determined through the standard properties of a Fourier series when one uses the basis functions $\cos(2n\pi\eta(v,\beta))$ and $\sin(2n\pi\eta(v,\beta))$. This yields the following formulas for the coefficients

$$a_n(\beta) = \int_{v_{reset}}^{v_{peak}} \frac{J(\beta)}{\rho_\beta(\beta)G_1(v,\beta)}\cos(2n\pi\eta(v,\beta))\rho_0(v,\beta)\,dv \tag{5.82}$$

$$b_n(\beta) = \int_{v_{reset}}^{v_{peak}} \frac{J(\beta)}{\rho_\beta(\beta)G_1(v,\beta)}\sin(2n\pi\eta(v,\beta)))\rho_0(v,\beta)\,dv \tag{5.83}$$

As the solution is now fully specified, we can begin to understand the effects of heterogeneity on this system. First, lets define the same order parameter as in chapter 4:

$$r(t) = \frac{1}{N}\sum_{k=1}^{N}\exp(2\pi i y_k) = \frac{1}{N}\sum_{i=1}^{N}\exp(2\pi i\eta(v_k,\beta_k)) \tag{5.84}$$

$$\to \int_\beta\int_{v_{reset}}^{v_{peak}} \rho(v,\beta,t)\exp(2n\pi i\eta(v,\beta))\,dvd\beta \quad \text{as} \quad N\to\infty \tag{5.85}$$

Evaluation of this integral yields

$$r(t) = \int_{\beta} (a_1(\beta) + b_1(\beta)) \exp(2\pi i \langle R_i(t)|\beta\rangle t)\, d\beta \tag{5.86}$$

Now, if we can prove that

$$\lim_{t\to\infty} r(t) = 0$$

then we can make some claim as to the stability of the asynchronous state. This is not entirely true, as $r(t) = 0$ only implies that the oscillators are evenly distributed around the unit circle, as opposed to uniformly distributed across the unit circle. For example, we can have two synchronized subpopulations that move in anti-phase with one another and thus yield $r(t) = 0$. We can rewrite the integral as

$$I(x) = \int_{\beta_1}^{\beta_2} \gamma(\beta) \exp(f(\beta)xi)\, d\beta$$

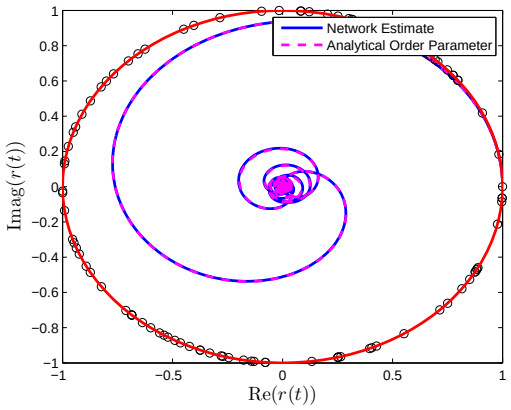
and thus we want the asymptotic behavior of $I(x)$ as $x \to \infty$. We can use the method of stationary phase which yields the asymptotic behavior of these kinds of integrals:

$$I(x) \sim \frac{\gamma(\beta)}{ixf'(\beta)} \exp\left(ixf(\beta)\right)\Big|_{\beta_1}^{\beta_2} \tag{5.87}$$

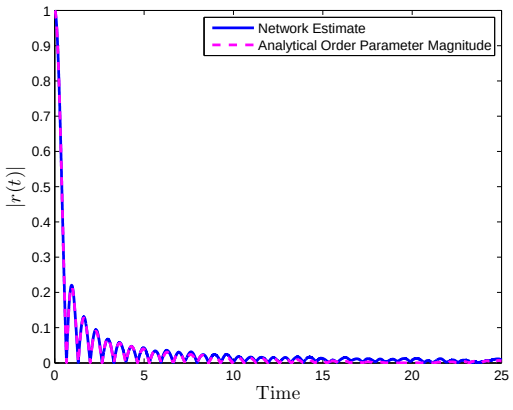
$$= \frac{1}{2\pi t}\left((a_1(\beta) + b_1(\beta))\left(\frac{\partial\langle R|\beta\rangle}{\partial\beta}\right)^{-1} \exp(2\pi i t\langle R|\beta\rangle)\right)\Big|_{\beta_1}^{\beta_2} \tag{5.88}$$

Thus, provided that all of these quantities are finite, then $r(t) \to 0$ as $t \to \infty$ as $|r(t)| \propto t^{-1}$. This is shown in figure 5.6 for a network of uncoupled quadratic integrate-and-fire neurons where we compare (5.86) to the numerically computed order parameter. Note that this is significantly slower convergence than a standard hyperbolic system which decays like $\exp(\lambda t)$ for $\lambda < 0$. Thus, even in the uncoupled state, the asynchronous steady state is stable in the sense that the order parameter decays asymptotically. Note that due to the method of stationary phase, even if the derivative(s) of $\langle R|\beta\rangle$ vanish for some $\beta^* \in [\beta_1, \beta_2]$ up to order p , then $|r(t)| \propto t^{-1/p}$ and thus $|r(t)| \to 0$ as $t \to \infty$.

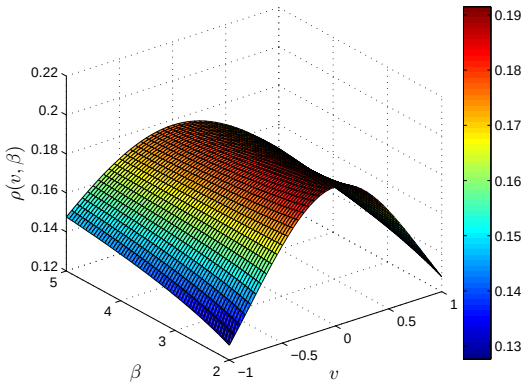
Given our work in Chapter 2, our analysis of the fast system with heterogeneity, and our eigenvalues results in section 5.1.4, we can immediately conclude that when the amount of heterogeneity is small, the eigenvalues determined by stability analysis of the steady state(s) in MFII are to leading order in ϵ identical to those of the associated moment closure reduced population density equation. Furthermore, they are the dominant eigenvalues that determine stability of the asynchronous steady state(s).



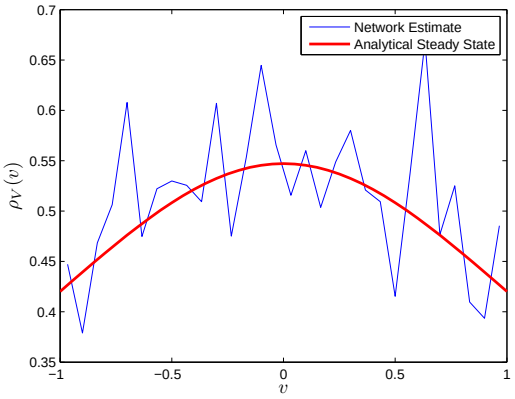
(a) Order parameter $r(t)$



(b) $|r(t)|$



(c) Steady state density, $\rho(v, \beta)$



(d) Steady state marginal density, $\rho_V(v)$

Figure 5.5: A network of 10^5 heterogeneous and uncoupled quadratic integrate-and-fire neurons is simulated. The heterogeneous parameter is the applied current to each neuron, and is drawn from a uniform distribution. The v variables for the neurons are all initialized at $v = v_{reset}$. In figure 5.5(a), The order parameter is computed using the analytical formula (5.86) (magenta dashed line) in addition to being estimated from the network (blue), (5.84). The computed phase of a subset of 100 oscillators is also plotted, demonstrating their uniform distribution along the unit circle at steady state. In figure 5.5(b), the order parameter for the network and the analytical prediction both decay, indicating asymptotic stability of the asynchronous steady state. The steady state bivariate density is also plotted in figure 5.5(c), in addition to the predicted (red) and measured (blue) marginal density in v is plotted in 5.5(d).

5.2 Networks with Noise

We will again consider our standard set of models. However, we will now add noise. We will specifically consider white noise:

$$\begin{aligned}
\dot{v}_i &= F(v_i) - w_i + gs(e_r - v_i) + I + \zeta_i = G_1(v_i, w_i, s) + \zeta_i && (5.89)\\
\dot{w}_i &= a(bv_i - w_i) = G_2(v_i, w_i) && (5.90)\\
\dot{s} &= -\frac{s}{\tau_s} + \frac{s_{jump}}{N}\sum_{j=1}^{N}\sum_{t<t_{j,k}} \delta(t - t_{j,k}) && (5.91)\\
v_i(t^-) &= v_{peak}, \rightarrow \begin{cases} v_i(t^+) = v_{reset} \\ w_i(t^+) = w_i(t^-) + w_{jump} \end{cases} && (5.92)
\end{aligned}$$

where ζ_i is a white noise process with mean 0 and standard deviation σ. The variable s is given as before.

5.2.1 The Population Density Equation

In the large network limit, one can again rigorously derive a population density equation for the network of neurons. The evolution equation for $\rho(v, w, t)$ is:

$$\frac{\partial \rho(v, w, t)}{\partial t} = -\nabla \cdot \boldsymbol{J}(v, w, s, t) \tag{5.93}$$

where

$$\begin{aligned}
\boldsymbol{J}(v, w, s, t) &= \begin{pmatrix} J^V(v, w, s, t) \\ J^W(v, w, t) \end{pmatrix} && (5.94)\\
J^V(v, w, s, t) &= G_1(v, w, s)\rho(v, w, t) - \frac{\sigma^2}{2}\frac{\partial \rho(v, w, t)}{\partial v} && (5.95)\\
J^W(v, w, t) &= G_2(v, w)\rho(v, w, t). && (5.96)
\end{aligned}$$

The boundary conditions on the flux are as before:

$$\begin{aligned}
J^V(v_{peak}, w, s, t) &= \lim_{v \to v^+_{reset}} J^V(v, w + w_{jump}, s, t) - \lim_{v \to v^-_{reset}} J^V(v, w + w_{jump}, s, t) && (5.97)\\
J^W|_{\partial W} &= 0. && (5.98)
\end{aligned}$$

This yields a discontinuous flux term due to the reset. However, if we force $v \in [v_{reset}, v_{peak}]$ by implementing a reflecting boundary on the neurons when $v = v_{reset}$ in addition to the typical reset at $v = v_{peak}$, we can simply rewrite the boundary condition as

$$J^V(v_{peak}, w, s, t) = J^V(v_{reset}, w + w_{jump}, s, t), \tag{5.99}$$

Note that we have an abuse of notation here as we have used J^V to denote the v component of the flux for both the half-infinity domain $(-\infty, v_{peak}]$ and the compact domain $[v_{reset}, v_{peak}]$. Furthermore, numerical simulation of the population density equation requires a restriction in the domain which we choose to be $[v_{reset}, v_{peak}]$ for convenience. However, for the purposes of the mean-field derivation, both cases can be considered.

In the large network limit, as in section 3.4, one can show that the equation for $s(t)$ converges to the ODE:

$$\dot{s} = -\frac{s}{\tau_s} + s_{jump} \int_W J^V(v_{peak}, w, s, t)\, dw \tag{5.100}$$

where the integral term is the network averaged firing rate $\langle R_i(t) \rangle$. To summarize, we have the following PDE/ODE coupled system:

$$\begin{aligned}
\frac{\partial \rho(v,w,t)}{\partial t} &= -\frac{\partial}{\partial v}\left(G_1(v,w,s)\rho(v,w,t) - \frac{\sigma^2}{2}\frac{\partial \rho(v,w,t)}{\partial v} \right) && (5.101)\\
&- \frac{\partial}{\partial w}\left(G_2(v,w)\rho(v,w,t) \right) && (5.102)\\
\dot{s} &= -\frac{s}{\tau_s} + s_{jump} \int_W J^V(v_{peak}, w, s, t)\, dw && (5.103)
\end{aligned}$$

subject to the boundary conditions (5.98)-(5.99) and initial conditions on s and ρ. In the noiseless case ($\sigma = 0$), this is enough conditions. In the noisy case, however, the order of the PDE is increased and another boundary condition is necessary. We will discuss this issue at length during the mean-field derivation as the boundary condition comes up when applying the separation of time scales.

The population density is equivalent to the marginal voltage density multiplied by the conditional w density:

$$\rho(v,w,t) = \rho_W(w|v,t)\rho_V(v,t). \tag{5.104}$$

Substituting this into equation (5.101), integrating with respect to w and using the boundary condition (5.98), we arrive at the one-dimensional PDE:

$$\begin{aligned}
\frac{\partial \rho_V(v,t)}{\partial t} &= -\frac{\partial}{\partial v}\left[\rho_V(v,t)\left(F(v) - \langle w|v \rangle + I + gs(e_r - v) \right) - \frac{\sigma^2}{2}\frac{\partial \rho_V(v,t)}{\partial v} \right] && (5.105)\\
&:= -\frac{\partial J(v, \langle w|v \rangle, s, t)}{\partial v}. && (5.106)
\end{aligned}$$

where the flux, J, has been redefined and $\langle w|v \rangle$ is the conditional mean of w given v. Note that we have used the fact that $G_1(v,w,s) = F(v) - w + gs(e_r - v) + I$ is affine in w.

Additionally, the equation for s becomes

$$\begin{aligned}\dot{s} &= -\frac{s}{\tau_s} + s_{jump} J(v_{peak}, \langle w|v_{peak}\rangle, s, t) \\ &= -\frac{s}{\tau_s} + s_{jump}\left(G_1(v_{peak}, \langle w|v_{peak}\rangle, s(t))\rho_V(v_{peak}, t) - \frac{\sigma^2}{2}\left.\frac{\partial \rho_V(v,t)}{\partial v}\right|_{v_{peak}}\right) \\ .\end{aligned} \tag{5.107}$$

Integration with respect to w is also needed to derive a new boundary condition on $\rho_V(v,t)$. Starting with the right-hand side of (5.99):

$$\begin{aligned}\int_W J^V(v_{reset}, w + w_{jump}, s, t)\, dw &= \int_W G_1(v_{reset}, s, w + w_{jump})\rho_W(w + w_{jump}|v_{reset}, t)\rho_V(v_{reset}, t)\, dw \\ &- \frac{\sigma^2}{2}\int_W \rho_V(v_{reset}, t)\left.\frac{\partial \rho_W(w + w_{jump}|v, t)}{\partial v}\right|_{v_{reset}} dw \\ &- \frac{\sigma^2}{2}\int_W \rho_W(w + w_{jump}|v_{reset}, t)\left.\frac{\partial \rho_V(v,t)}{\partial v}\right|_{v_{reset}} dw.\end{aligned}$$

Note that $\rho_W(w + w_{jump}|v)$ is merely the conditional density in w shifted by w_{jump} to the left. Thus, since we are still integrating over the entire w domain, we have the following:

$$\int_W \rho_W(w + w_{jump}|v_{reset}, t)\rho_V(v_{reset}, t)G(v_{reset}, s, w + w_{jump})\, dw = G(v_{reset}, s, \langle w|v_{reset}\rangle)\rho_V(v_{reset}, t)$$

$$\int_W \rho_V(v_{reset}, t)\left.\frac{\partial \rho_W(w + w_{jump}|v, t)}{\partial v}\right|_{v=v_{reset}} dw = \rho_V(v_{reset}, t)\frac{\partial}{\partial v}\left(\int_W \rho_W(w + w_{jump}|v, t)\, dw\right)\Bigg|_{v_{reset}} = 0$$

$$\int_W \rho_W(w + w_{jump}|v_{reset}, t)\left.\frac{\partial \rho_V(v,t)}{\partial v}\right|_{v=v_{reset}} dw = \left.\frac{\partial \rho_V(v,t)}{\partial v}\right|_{v_{reset}}$$

It follows that

$$\begin{aligned}\int_W J^V(v_{reset}, w + w_{jump}, s, t) &= G(v_{reset}, s, \langle w|v_{reset}\rangle)\rho_V(v_{reset}, t) - \frac{\sigma^2}{2}\left.\frac{\partial \rho_V(v,t)}{\partial v}\right|_{v_{reset}} \\ &= J(v_{reset}, s, \langle w|v_{reset}\rangle, t)\end{aligned}$$

Similar integration steps show that

$$\int_W J^V(v_{peak}, w, s, t) = G(v_{peak}, s, \langle w|v_{peak}\rangle)\rho_V(v_{peak}, t) - \frac{\sigma^2}{2}\left.\frac{\partial \rho_V(v,t)}{\partial v}\right|_{v_{peak}} \tag{5.108}$$

$$= J(v_{peak}, s, \langle w|v_{peak}\rangle, t) \tag{5.109}$$

and the boundary condition becomes

$$J(v_{reset}, s, \langle w|v_{reset}\rangle, t) = J(v_{peak}, s, \langle w|v_{peak}\rangle, t) \tag{5.110}$$

So far every step applied has been exact and no approximation has been made. However, without a PDE for $\langle w|v\rangle$, one cannot solve the PDE (5.105) for ρ_V. One could attempt to derive an equation for $\langle w|v\rangle$, however this will result in an equation which depends on $\langle w^2|v\rangle$ (see subsection 2.2). An approximation is necessary to end the dependence of the equations on these higher order moments, i.e., to close the system. One approach is to use a moment closure assumption, i.e., to impose a relationship between the higher moments and the lower moments. As in the noiseless network, we can consider two moment-closure assumptions

The simplest way to deal with the dependence of the PDE (5.105) on $\langle w|v\rangle$ is to use a standard first order moment closure assumption:

$$\langle w|v\rangle = \langle w\rangle. \tag{5.111}$$

Alternatively, one can make the higher order moment closure assumption

$$\langle w^2|v\rangle - \langle w|v\rangle^2 = \sigma^2_{w|v} = 0 \tag{5.112}$$

which yields a closed PDE/ODE system for $\langle w|v\rangle$, $\rho_V(v,t)$ and s.

First order moment closure can be applied to equations (5.106)-(5.107) to yield the following:

$$\frac{\partial \rho_V(v,t)}{\partial t} = -\frac{\partial}{\partial v}\left(\rho_V(v,t)G_1(v,\langle w\rangle,s) - \frac{\sigma^2}{2}\frac{\partial \rho_V(v,t)}{\partial v}\right) = -\frac{\partial}{\partial v}\left(J_V(v,\langle w\rangle,s,t)\right) \tag{5.113}$$

$$s' = -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s}J_V(v_{peak},\langle w\rangle,s,t) \tag{5.114}$$

$$\langle w\rangle' = G_1(\langle v\rangle,\langle w\rangle) + w_{jump}J(v_{peak},s,\langle w\rangle,t) \tag{5.115}$$

where the equation for $\langle w\rangle$ is derived in the next section.

As in Chapter 4, there are a few possible routes we can take from here. The first route is to simply use (5.113)-(5.115) and numerically solve the system. We can also proceed analytically by deriving the mean-field system. In the next section, we will look at our higher order moment closure assumption. While we can derive a partial differential equation for $\langle w|v\rangle$, unlike in the scalar case, it is in general not as well posed.

5.2.2 Second Order Moment Closure

To derive an equation for $\langle w|v\rangle$, we will need to multiply equation (3.60) by w and integrate with respect to w. This gives:

$$\begin{aligned}\frac{\partial}{\partial t}\int_W w\rho(v,w,t)\,dw &= -\frac{\partial}{\partial v}\int_W w\left(G_1(v,s,w)\rho(v,w,t) - \frac{\sigma^2}{2}\frac{\partial \rho(v,w,t)}{\partial v}\right)dw \\ &- \left(wJ^W\Big|_{\partial W} - \int_W G_2(v,w)\rho(v,w,t)\,dw\right)\end{aligned}$$

Using equation 5.106 gives

$$\begin{aligned}
\frac{\partial}{\partial t}\left(\rho_V(v,t)\langle w|v\rangle\right) &= -\frac{\partial}{\partial v}\left[\langle w|v\rangle\left(F(v)+gs(er-v)+I\right)\rho_V(v,t)-\langle w^2|v\rangle\rho_V(v,t)-\frac{\sigma^2}{2}\frac{\partial\langle w|v\rangle\rho_V}{\partial v}\right] \\
&+ G_2(v,\langle w|v\rangle)\rho_V(v,t) \\
&= -\frac{\partial}{\partial v}\left[\langle w|v\rangle G_1(v,\langle w|v\rangle,s)\rho_V(v,t)-\sigma^2_{w|v}\rho_V(v,t)-\frac{\sigma^2}{2}\frac{\partial\langle w|v\rangle\rho_V}{\partial v}\right] \\
&+ G_2(v,\langle w|v\rangle)\rho_V(v,t) \\
&= -\frac{\partial}{\partial v}\left[\langle w|v\rangle J(v,\langle w|v\rangle,s,t)\right]+\frac{\partial}{\partial v}\left(\sigma^2_{w|v}\rho_V(v,t)\right)+\frac{\sigma^2}{2}\frac{\partial}{\partial v}\left(\frac{\partial\langle w|v\rangle}{\partial v}\rho_V(v,t)\right) \\
&+ G_2(v,\langle w|v\rangle)\rho_V(v,t) \qquad (5.116)
\end{aligned}$$

where we have assume that $J^W = 0$ on ∂W. Recall that

$$\frac{\partial\rho_V(v,t)}{\partial t} = -\frac{\partial}{\partial v}\left(J(v,\langle w|v\rangle,s,t)\right) \qquad (5.117)$$

and thus

$$\begin{aligned}
\langle w|v\rangle\left(\frac{\partial\rho_V(v,t)}{\partial t}+\frac{\partial}{\partial v}(J(v,s,\langle w|v\rangle,t)\right)+\rho_V(v,t)\frac{\partial\langle w|v\rangle}{\partial t} &= -J(v,s,\langle w|v\rangle,t)\frac{\partial\langle w|v\rangle}{\partial v}+\frac{\partial}{\partial v}\left(\sigma^2_{w|v}\rho_V(v,t)\right) \\
&+\frac{\sigma^2}{2}\frac{\partial}{\partial v}\left(\frac{\partial\langle w|v\rangle}{\partial v}\rho_V(v,t)\right) \\
&+G_2(v,\langle w|v\rangle)\rho_V(v,t)
\end{aligned}$$

and as the first term is 0 due to (5.117), we have the following second order moment closure equation

$$\begin{aligned}
\rho_V(v,t)\frac{\partial\langle w|v\rangle}{\partial t} &= -J(v,\langle w|v\rangle,s,t)\frac{\partial\langle w|v\rangle}{\partial v}+\frac{\partial}{\partial v}\left(\sigma^2_{w|v}\rho_V(v,t)\right)+\frac{\sigma^2}{2}\frac{\partial}{\partial v}\left(\frac{\partial\langle w|v\rangle}{\partial v}\rho_V(v,t)\right) \\
&+ G_2(v,\langle w|v\rangle)\rho_V(v,t). \qquad (5.118)
\end{aligned}$$

Notice that with $\sigma^2_{w|v} = \langle w^2|v\rangle - \langle w|v\rangle$ we have an explicit dependence on $\langle w^2|v\rangle$, and we need a partial differential equation for the second conditional moment. While equation (5.118) is exact, to break the dependence and close the system we assume that $\sigma^2_{w|v} = 0$, which yields the following system of equations:

$$\frac{\partial\rho_V(v,t)}{\partial t} = -\frac{\partial}{\partial v}\left(J(v,\langle w|v\rangle,s,t)\right) \qquad (5.119)$$

$$\rho_V(v,t)\frac{\partial\langle w|v\rangle}{\partial t} = -J(v,\langle w|v\rangle,s,t)\frac{\partial\langle w|v\rangle}{\partial v}+\frac{\sigma^2}{2}\frac{\partial}{\partial v}\left(\frac{\partial\langle w|v\rangle}{\partial v}\rho_V(v,t)\right)+G_2(v,\langle w|v\rangle)\rho_V(v,t) \qquad (5.120)$$

$$s' = -\frac{s}{\tau_s}+\frac{\lambda_s}{\tau_s}J(v_{peak},s,\langle w|v_{peak}\rangle,t) \qquad (5.121)$$

The only caveat that remains is that we require a boundary condition on $\langle w|v\rangle$. To arrive at this, first note that if we integrate (5.116) with respect to v we arrive at an equation for $\langle w\rangle$:

$$\begin{aligned}
\langle w\rangle' &= G_2(\langle v\rangle,\langle w\rangle) - [\langle w|v_{peak}\rangle J(v_{peak},\langle w|v_{peak}\rangle,s,t) - \langle w|v_{reset}\rangle J(v_{reset},\langle w|v_{reset}\rangle,s,t)] \\
&+ \sigma^2_{w|v}\rho_V(V,t)\Big|_{\partial V} + \frac{\sigma^2}{2}\frac{\partial\langle w|v\rangle}{\partial v}\rho_V(v,t)\Big|_{\partial V} \\
&= G_2(\langle v\rangle,\langle w\rangle) + [\langle w|v_{reset}\rangle - \langle w|v_{peak}\rangle]\, J(v_{peak},s,\langle w|v_{peak}\rangle,t) \\
&+ \sigma^2_{w|v}\rho_V(V,t)\Big|_{\partial V} + \frac{\sigma^2}{2}\frac{\partial\langle w|v\rangle}{\partial v}\rho_V(v,t)\Big|_{\partial V}
\end{aligned}$$

As equation (5.122) was arrived at by integrating with respect to w then v, it should equal the equation obtained if we integrate with respect to v first, then w:

$$\begin{aligned}
\langle w\rangle' &= G_2(\langle v\rangle,\langle w\rangle) - \int_W\int_V w\frac{\partial J^V}{\partial v}\,dvdw \\
&= G_2(\langle v\rangle,\langle w\rangle) - \int_W w\int_V \frac{\partial J^V}{\partial v}\,dvdw \\
&= G_2(\langle v\rangle,\langle w\rangle) - \int_W w(J^V(v_{peak},s,w,t) - J^V(v_{reset},s,w,t))\,dw \\
&= G_2(\langle v\rangle,\langle w\rangle) - \int_W w(J^V(v_{peak},s,w,t) - J^V(v_{peak},s,w-w_{jump},t))\,dw \\
&= G_2(\langle v\rangle,\langle w\rangle) + w_{jump}J(v_{peak},s,\langle w|v_{peak}\rangle,t) + O(w^2_{jump})
\end{aligned}$$

where we have assumed that $J^V = 0$ on ∂W. Subtracting (5.122) from (5.122) yields

$$\begin{aligned}
0 &= [\langle w|v_{reset}\rangle - \langle w|v_{peak}\rangle - w_{jump}]J(v_{peak},\langle w|v_{peak}\rangle,s,t\rangle) + \sigma^2_{w|v}\rho_V(V,t)\Big|_{\partial V} && (5.122)\\
&+ \frac{\sigma^2}{2}\frac{\partial\langle w|v\rangle}{\partial v}\rho_V(v,t)\Big|_{\partial V} + O(w^2_{jump}) && (5.123)
\end{aligned}$$

If we now employ our moment closure assumption, we can eliminate the term involving $\sigma^2_{w|v}$

$$0 = [\langle w|v_{reset}\rangle - \langle w|v_{peak}\rangle - w_{jump}]J(v_{peak},\langle w|v_{peak},s,t\rangle) + \frac{\sigma^2}{2}\frac{\partial\langle w|v\rangle}{\partial v}\rho_V(v,t)\Big|_{\partial V} + O(w^2_{jump}) \tag{5.124}$$

Now, if $\sigma = O(w_{jump})$, then the leading order boundary condition becomes

$$\langle w|v_{reset}\rangle = \langle w|v_{peak}\rangle + w_{jump} \tag{5.125}$$

Additionally, we have numerically found that when the higher order moment closure assumption is invalid (when the contribution of the term is of $O(w_{jump})$, then we have

$$\langle w|v_{reset}\rangle \leq \langle w|v_{peak}\rangle + w_{jump}$$

This is illustrated in Figure 5.6, where we show for two examples that $\langle w \rangle$ and $\langle w|v \rangle$ lie in the same interval which has length w_{jump}.

The equations (5.119)-(5.121) in addition to the boundary condition (5.125) constitute the second order moment closure system with noise. Unfortunately, due to the $\rho_V(v,t)$ coefficient in front of equation (5.120), the equation is not necessarily well posed. As such, we will again primarily consider the first order moment closure system, and its mean-field reduction.

The Mean-Field Reduction

The coupled system of one PDE and two ordinary differential equations (5.113)-(5.115) derived using the first order moment closure is one step removed from our mean-field approximation. Once again, the final step in the derivation is the separation of time scales and resolving the slow system. The steady state of (5.113) must satisfy the following ordinary differential equation:

$$0 = -\frac{\partial}{\partial v}\left[G_1(v, \langle w \rangle, s)\rho_V(v) - \frac{\sigma^2}{2}\frac{\partial \rho_V(v)}{\partial v}\right] = -\frac{\partial J(v, \langle w \rangle, s)}{\partial v}. \tag{5.126}$$

It is clear from this equation that the boundary condition (5.110) is automatically satisfied at steady state as the solution for $J(v, \langle w \rangle, s)$ is independent of v. Thus an alternate boundary condition will be needed. The typical boundary conditions applied are:

$$J(v, \langle w \rangle, s) = J(v_{peak}, \langle w \rangle, s) = \langle R_i(t) \rangle_\sigma \tag{5.127}$$

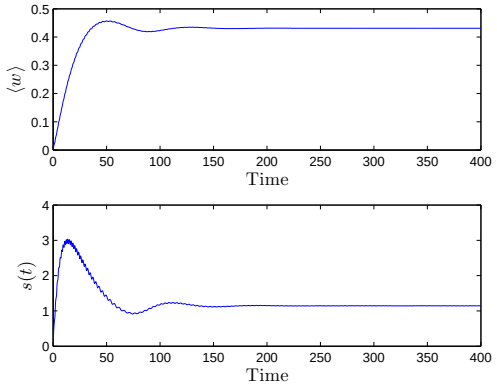
$$\rho_V(v_{peak}; \sigma) = 0 \tag{5.128}$$

when we consider the interval $[v_{reset}, v_{peak}]$. On the extended interval, the boundary condition (5.127) becomes:

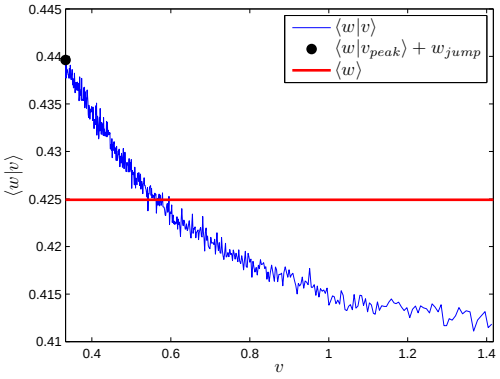
$$J(v, \langle w \rangle, s) = \begin{cases} \langle R_i(t) \rangle_\sigma & v_{reset} \leq v \leq v_{peak} \\ 0 & v < v_{reset} \end{cases} \tag{5.129}$$

These boundary conditions have been previously used in [66] in their analysis of the leaky integrate-and-fire models with white noise, and in [2]. We note that in these two papers the justification for the absorbing boundary condition appears to be different. In [66], the justification is that $\rho_V(v;\sigma) = 0$ for $v > v_{peak}$ and thus for continuity and integrability reasons, the authors set $\rho(v_{peak};\sigma) = 0$. In [2], the authors state that $\rho(v_{peak};\sigma) = 0$ as all the firing is due to noise, and thus the deterministic component of the flux should not contribute anything.

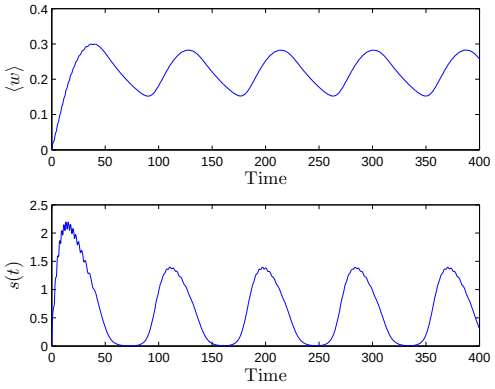
Here, we will derive the solution in some detail. This will allow us to offer an alternative justification for the boundary condition (5.128) and to investigate the limiting behaviour



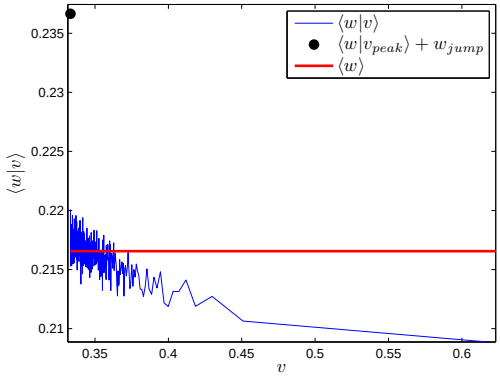
(a) Tonic Firing $s(t)$ and $\langle w \rangle$



(b) $\langle w|v \rangle$ at $t = 200$



(c) Bursting $s(t)$ and $\langle w \rangle$



(d) $\langle w|v \rangle$ at $t = 150$.

Figure 5.6: The first conditional moment $\langle w|v \rangle$ is computed by sorting the w_i as a increasing function of v_i and then averaging locally the v_i and w_i. A network of 50,000 neurons was simulated using the chattering neuron (CH) parameter sets in Table 5.2 in either the tonic firing (a)-(b) ($g = 0.33$, $I = 0.29$, $\sigma = 0.05$) or the bursting regions (c)-(d) ($g = 0.33$, $I = 0.11$, $\sigma = 0.05$). Note that $\langle w|v_{peak} \rangle + w_{jump}$ is plotted at $v = v_{reset}$ (black dot in (b),(d)) to demonstrate the validity of the boundary condition in the tonic firing region. The red line is $\langle w \rangle$. In both the tonic firing and bursting regions, $\langle w|v \rangle$ is a monotonically decreasing function of v with a narrow range. When the network is bursting, $\langle w|v_{reset} \rangle = \langle w|v_{peak} \rangle + w_{jump}$ during the active portion of the bursts, and $\langle w|v_{reset} \rangle < \langle w|v_{peak} \rangle + w_{jump}$ during the quiescent periods.

of $\rho(v_{peak};\sigma)$ and $\langle R_i(t)\rangle_\sigma$ as $\sigma \to 0$. Solving equation (5.126) for $\rho_V(v;\sigma)$ and using the boundary condition (5.127) yields:

$$\rho(v;\sigma) = -\frac{\langle R_i(t)\rangle_\sigma}{\sigma^2}\int_{v_{reset}}^{v} \exp\left(-\frac{2}{\sigma^2}(M(v')-M(v))\right)dv' + D\exp\left(\frac{2}{\sigma^2}M(v)\right)$$

where $M(v)$ is an anti-derivative of $G_1(v,\langle w\rangle,s)$ with respect to v and $D = \rho(v_{reset})\exp(\frac{2}{\sigma^2}M(v_{reset}))$. To simplify the notation we will supress the dependence of $M(v)$ on s and $\langle w\rangle$, until it is necessary. Before proceeding further, we use Laplace's method for integrals [17] to shed some insight into the asymptotic behavior of $\rho(v;\sigma)$. This requires looking at two separate cases, when $I - I^*(s,w) > 0$ and when $I - I^*(s,w) \leq 0$. To do this, we will need to determine the asymptotics of

$$I_1(\sigma,v) = \frac{2}{\sigma^2}\int_{v_{reset}}^{v} \exp\left(-\frac{2}{\sigma^2}(M(v')-M(v))\right)dv' \tag{5.130}$$

$$I_2(\sigma,v) = \frac{2}{\sigma^2}\int_{v}^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}(M(v')-M(v))\right)dv' \tag{5.131}$$

We will primarily consider the case that $I - I^*(s,w) > 0$, with a similar approach being applicable to $I - I^*(s,w) < 0$. If $I - I^*(s,w) > 0$, then we know that $G_1(v,s,w) > 0$ and if we define the quantity

$$\phi(v',v) = M(v) - M(v')$$

then

$$\frac{\partial\phi(v',v)}{\partial v'} = -G_1(v',s,w) < 0.$$

In this case the quantity inside the exponential, $-\frac{2}{\sigma^2}(M(v')-M(v))$ is a decreasing function of v' globally on the interval $[v_{reset}, v_{peak}]$. Defining $x = \frac{2}{\sigma^2}$, we can rewrite the integrals in the Laplace form and find the asymptotic behaviours:

$$I_1(x,v) = x\int_{v_{reset}}^{v} \exp\left(x\phi(v',v)\right)dv' \sim -\exp(x\phi(v_{reset},v))\left(\frac{\partial\phi(v',v)}{\partial v'}\bigg|_{v'=v_{reset}}\right)^{-1} \tag{5.132}$$

$$= \exp\left(\frac{2}{\sigma^2}(M(v)-M(v_{reset}))\right)\frac{1}{G_1(v_{reset},s,w)} \tag{5.133}$$

$$I_2(x,v) = x\int_{v}^{v_{peak}} \exp(x\phi(v',v))\,dv' \sim -\exp(x\phi(v,v))\left(\frac{\partial\phi(v',v)}{\partial v'}\bigg|_{v'=v}\right)^{-1} \tag{5.134}$$

$$= \frac{1}{G_1(v,s,w)} \tag{5.135}$$

as $x \to \infty$. Thus, as $x \to \infty$, $I_1(x,v)$ diverges while $I_2(x,v)$ converges to $1/G_1(v,s,w)$.

This would seem to imply that if $\langle R_i(t)\rangle_\sigma$ is convergent in the $\sigma \to 0$ limit, the density function contains a divergent term as if $G_1(v,\langle w\rangle,s) > 0$, then $M(v) > M(v_{reset})$ and

the first term diverges exponentially fast as $\sigma \to 0$. Thus, to obtain a convergent density function, we need to remove the first term in the integral.

Rewriting the density as:

$$\begin{aligned}\rho(v;\sigma) &= \frac{2\langle R_i(t)\rangle_\sigma}{\sigma^2}\int_v^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}(M(v') - M(v))\right) dv' \\ &+ \left[D - \frac{2\langle R_i(t)\rangle_\sigma}{\sigma^2}\int_{v_{reset}}^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}M(v')\right) dv'\right] \exp\left(\frac{2}{\sigma^2}M(v)\right),\end{aligned}$$

then since we are still free to specify a boundary condition, we may choose D (and hence $\rho(v_{reset})$) to eliminate the divergent term, yielding:

$$\rho(v;\sigma) = \frac{2\langle R_i(t)\rangle_\sigma}{\sigma^2}\int_v^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}(M(v') - M(v))\right) dv' \tag{5.136}$$

Note that this choice of D is equivalent to applying the boundary condition (5.128). Thus, the boundary condition can be seen as a regularity condition requiring the density $\rho(v;\sigma)$ be well behaved in the small noise limit.

As in the noiseless case, applying the normalization condition on $\rho(v;\sigma)$ yields an expression for the firing rate:

$$\langle R_i(t)\rangle_\sigma = \left(\frac{2}{\sigma^2}\int_{v_{reset}}^{v_{peak}}\int_{v'}^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}(M(v',\langle w\rangle,s) - M(v,\langle w\rangle,s))\right) dv'dv\right)^{-1}. \tag{5.137}$$

This leads to the following mean-field system for the network:

$$\dot{s} = -\frac{s}{\tau_s} + s_{jump}\langle R_i(t)\rangle_\sigma \tag{5.138}$$

$$\dot{\langle w\rangle} = G_2(\langle v\rangle,\langle w\rangle) + w_{jump}\langle R_i(t)\rangle_\sigma \tag{5.139}$$

Using the expansions of the integrals given above shows that the solution has the following asymptotic behaviour

$$\begin{aligned}\rho(v;\sigma) &\sim \rho_0(v) = \frac{\langle R_i(t)\rangle_0}{G_1(v,\langle w\rangle,s)} \quad \sigma \to 0 \\ \langle R_i(t)\rangle_\sigma &\sim \langle R_i(t)\rangle_0 \quad \sigma \to 0\end{aligned}$$

for $I > I^*(\langle w\rangle, s)$. This implies that, in the tonic firing region of the parameter space, as $\sigma \to 0$, the firing rate converges to the noiseless value, which in turn implies that the mean-field equations converge to the noiseless mean-field equations derived in Chapter 2 of this thesis. The convergence of the density is more delicate. The steady state densities are shown for both the simulated networks and the analytical solution in Figures 5.7(a) and 5.7(b) respectively. Note that the firing rate and density at steady state are related by

$$\rho_0(v)G_1(v,\langle w\rangle,s) = \langle R_i(t)\rangle_0 \quad I > I^*(\langle w\rangle, s)$$

for the noiseless network. Since $\langle R_i(t)\rangle_0 > 0$ when $I > I^*(\langle w\rangle, s)$, the boundary condition (5.128) leads to an inconsistency at $v = v_{peak}$. Thus the convergence of the density is only pointwise and for $v_{reset} \leq v < v_{peak}$. An example of this is shown in Figure 5.7(c).

This inconsistency can be dealt with by noting that $\rho(v_{peak};\sigma) = 0$ is a sufficient, but not necessary condition for $\rho(v;\sigma)$ to converge to $\rho_0(v)$ for $v \neq v_{peak}$. In fact it can be weakened to yield convergence even at v_{peak}. Specifically, making the following choice for D

$$D = \frac{2\langle R_i(t)\rangle_\sigma}{\sigma^2}\int_{v_{reset}}^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}M(v')\right)\,dv' + \exp\left(-\frac{2}{\sigma^2}M(v_{peak})\right)\rho_0(v_{peak})$$

one can show that the term

$$\exp\left(\frac{2}{\sigma^2}(M(v) - M(v_{peak}))\right)\rho_0(v_{peak})$$

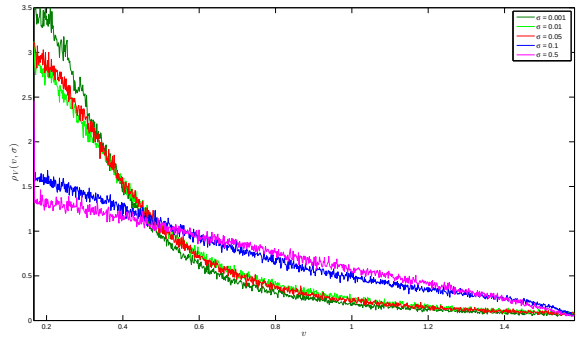
added to the density converges to $\rho_0(v_{peak})$ if $v = v_{peak}$ and 0 otherwise. Thus, the criteria $\rho(v_{peak};\sigma) = 0$ is not necessary even for convergence at $v = v_{peak}$ as $\sigma \to 0$.

A similar approach when $I < I^*(\langle w\rangle, s)$ can demonstrate the same convergence. Thus, solutions of the non-smooth noiseless mean-field system could be used as order zero solutions in a weak noise perturbation expansion of solutions of the mean-field system above.

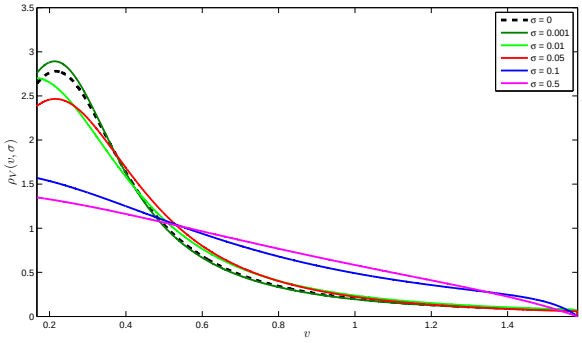
5.2.3 Numerical Simulations

In this section, we compare simulations of the PDE system (5.113)–(5.115), the mean-field system (5.138)-(5.139) and of the full network (5.89)-(5.90) with 10,000 neurons.

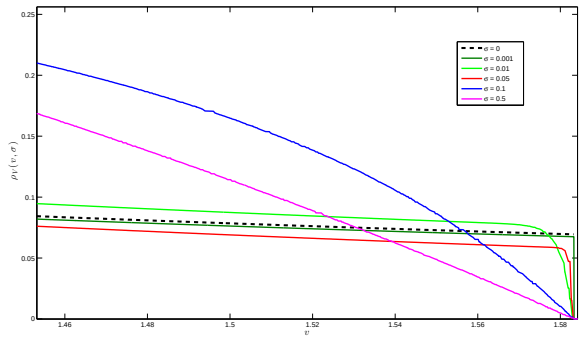
We begin by considering different parameter sets for the Izhikevich model, taken from [92], which were fit to data for various neuron types. We use parameter sets for the CA1 pyramidal neuron, the intrinsically bursting neuron (IB), the chattering neuron (CH), and the rapidly spiking neuron (RS). The parameter values are given in Table I. As illustrated in Figure 5.6 for the chattering neuron, when these neurons are connected with excitatory coupling, the networks can exhibit both tonic firing and network induced bursting with or without noise. We will focus on the situation where the networks are bursting as this is where the mean-field systems can lose accuracy. The results of simulations using the intrinsically bursting and chattering neuron parameter values are shown in Figure 7.1. In the bursting region, the frequency error present in the mean-field system is dramatically reduced in the moment-closure reduced PDE, as shown in Figure 7.1. Similar results were found for the CA1 and rapidly spiking parameter values (not shown). This demonstrates that the bulk of the frequency error in the mean-field system is actually due to the separation of time scales approximation. Thus, the PDE system is superior to the mean-field system in predicting the steady state and dynamics for the actual network.



(a) $\rho_V(v;\sigma)$ for various σ from simulations of a network with 50,000 neurons



(b) $\rho_V(v;\sigma)$ from eq. (5.136)



(c) A close up of the convergence of the density $\rho_V(v;\sigma)$ as $\sigma \to 0$ near v_{peak}

Figure 5.7: (a) A coupled network of 50,000 Izhikevich neurons was simulated until steady state and the steady state density $\rho_V(v;\sigma)$ was determined by using a normalized histogram. (b) The solution for the steady-state density was found analytically using eq. (5.136). (c) The nature of the convergence of the density $\rho(v;\sigma)$ to $\rho_0(v)$, the analytical solution to the steady-state density without noise. The density function $\rho(v;\sigma)$ only converges pointwise to $\rho_0(v)$ on $[v_{reset}, v_{peak})$, with the derivative becoming unbounded at $v = v_{peak}$. The parameters are the rapid spiking (RS) parameter set in Table 5.2., with g, I chosen such that the steady state of the network was tonically firing.

To quantify the amount of synchrony in the network, one can use an order parameter defined by :

$$r(t) = \frac{1}{N}\sum_{k=1}^{N} \exp\left(2\pi i \left[\frac{v_k - v_{reset}}{v_{peak} - v_{reset}}\right]\right) \tag{5.140}$$

which has been done for example in [2]. If $|r(t)| = 1$, then the neurons are perfectly synchronized across the network, while if $|r(t)| = 0$, they are evenly distributed around the unit cycle. Note that the order parameter used here is different from those used previously for practical reasons, it is significantly easier to compute.

As shown in Figures 7.1(b) and 7.1(d), the first order moment closure equation provides a great deal more information about synchrony than the mean-field system.

In addition to the plain Izhikevich model derived from topological normal form theory, various modifications have been suggested to make model better fit the spiking dynamics and spike profiles for different neurons. For example, the model can be fit to a fast spiking inhibitory interneuron via the following (see page 299 of [92])

$$\dot{w} = \begin{cases} a((v - v_b)^3 - w) & \text{if} \quad v \geq v_b \\ -aw & \text{if} \quad v < v_b \end{cases} \tag{5.141}$$

Additionally, it is possible to fit sharper spike upstrokes present in actual neurons via the following adjustment:

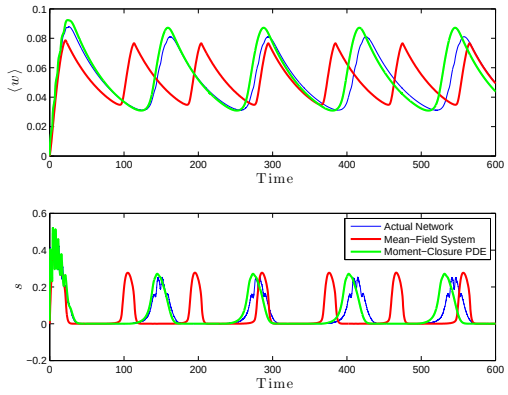
$$\dot{v} = k(v)v(v - \alpha) - w + gs(e_r - v) + I$$

where

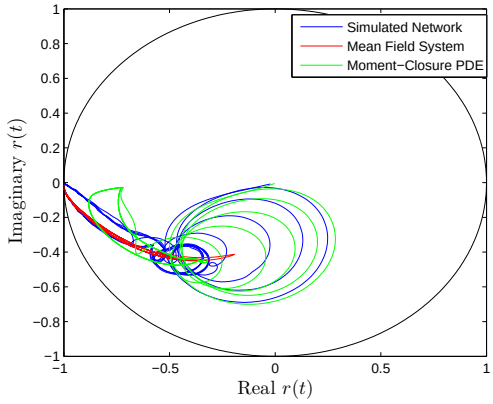
$$k(v) = \begin{cases} k_{min} & \text{if} \quad v \leq \alpha \\ 1 & \text{if} \quad v > \alpha \end{cases} \tag{5.142}$$

This has been done for a hippocampal CA3 pyramidal neuron in [49] in addition to other examples in [92]. The parameter values for these models are given in Table 5.2.

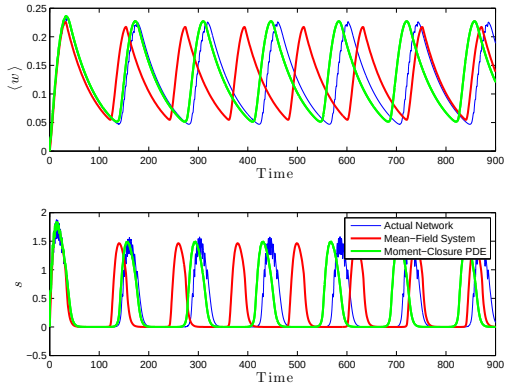
For both of these modified Izhikevich models, one can derive the corresponding moment-closure reduced PDE and mean-field system. Comparisons of simulations of these systems with those of the full network are shown in Figures 5.8 and 5.9. It is clear that in both cases, the PDE substantially outperforms the mean-field system, both in reproducing network behaviour and capturing synchrony levels.



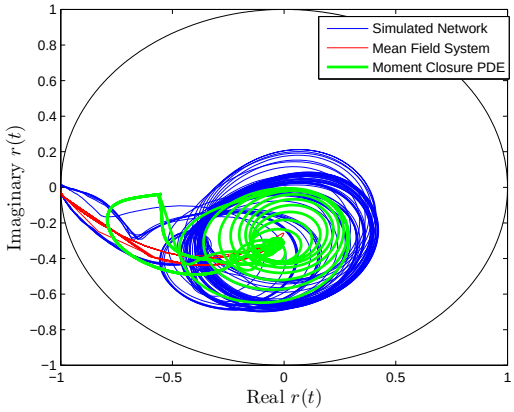
(a) Izhikevich Model, IB Parameters



(b) Order Parameter, IB

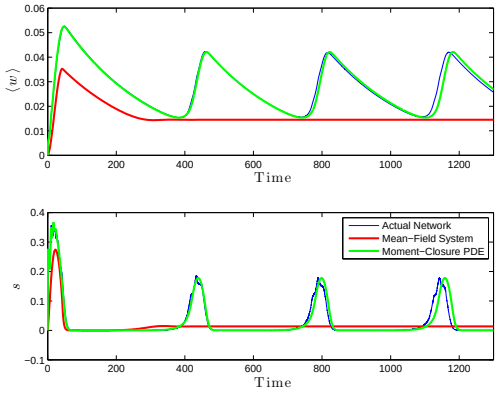


(c) Izhikevich Model, CH Parameters

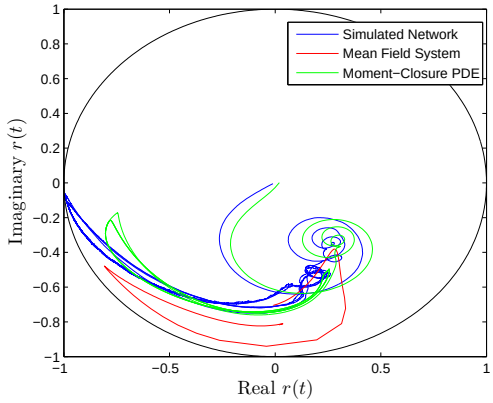


(d) Order Parameter, CH

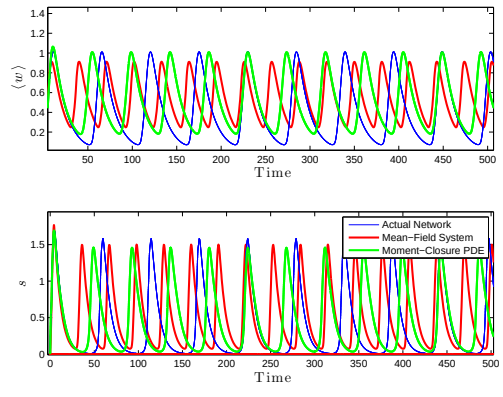
Figure 5.8: Comparison of direct numerical simulations of large coupled networks of Izhikevich neurons with noise, the mean field system and the moment closure PDE system. The direct simulations are shown in blue, while the mean-field system is shown in red, and the first order moment closure PDE is shown in green. (a),(c) Network mean variables; (b),(d) order parameter as defined in eq. (5.140). The PDE system has substantially less frequency error than the mean-field system and gives a better representation of the amount of synchronization in the network. The parameter sets are those of an intrinsically bursting neuron (a),(b) and a chattering neuron (c),(d). The values can be found in table 5.2. The standard deviation for the noise is $\sigma = 0.02$ for the intrinsically bursting network, and $\sigma = 0.014$ for the chattering neuron network with the other parameters being $g = 0.33, I = 0.037$ and $g = 0.56$ and $I = 0.055$, respectively.



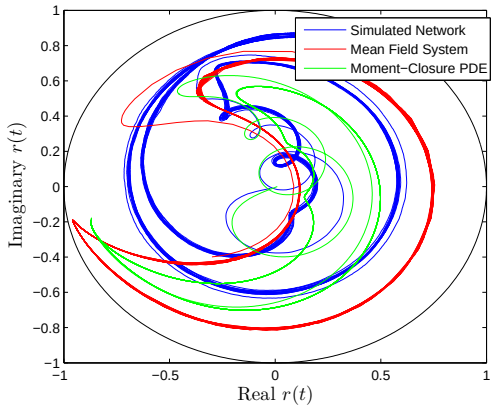
(a) Izhikevich Model with k-switching



(b) Order Parameter, k-switching



(c) Izhikevich Model, FS Parameters



(d) Order Parameter, FS

Figure 5.9: Comparison of direct simulations of large coupled networks of networks of Izhikevich neurons with noise, the mean field system and the moment closure PDE. (a) the model with k-switching, defined by eq. (5.142), to accurately represent spike half-widths. (c) the model for fast spiking interneurons which has nonlinear w dynamics given by eq. (5.141). The standard deviation of the noise is $\sigma = 0.1$ for the fast spiking network, with $g = 1.81$ and $I = 0.0661$ with the parameter $v_b = 0$. For the k-switching network, the parameter values used were $\sigma = 0.032$, $I = 0.0189$, $g = 0.7692$ in addition to $k_{min} = 0.03$. The other parameters can be found in Table 5.2 The direct simulations are shown in blue, while the mean-field system is shown in red, and the first order moment closure PDE is shown in green. As with the plain Izhikevich model, the PDE has substantially less frequency error than the mean-field system. The order parameter for the networks, as defined by (5.140), is shown in (b), (d). While not perfect, the moment-closure reduced PDE provides substantially more information about network synchrony than the mean-field system.

Parameter Set	CA1	CH	IB	RS	FS	KS
α	0.25	0.33	0.4	0.33	0.18	0.72
v_{reset}	0.25	0.33	0.25	0.17	0.18	0.154
v_{peak}	1.67	1.42	1.67	1.58	1.45	1.462
w_{jump}	0.028	0.028	0.019	0.04	0	0.012
$1/\tau_W$	0.033	0.017	0.017	0.07	0.2	0.005
b	0.017	0.011	0.056	-0.048	1.38	-0.003

Table 5.2: The dimensionless parameters for the fitted Izhikevich models used in network, mean-field, and population density simulations throughout the text. The models are CA1 pyramidal cell (CA1), chattering neuron (CH), intrinsically bursting neuron (IB), rapid spiking neuron (RS), fast spiking (FS) and k-switching (KS). The corresponding dimensional parameters can be found in [92]. The values of the following parameters were the same for all simulations: $\tau_s = 1.5, e_r = 1$, and $s_{jump} = 1$. The parameters g and I and σ vary, and are treated as bifurcation parameters.

5.2.4 Linear Stability Analysis for the White-Noise System

As in chapter 2, we will derive and analyze the full spectrum of eigenvalues of the population density equation in the mean-field limit. Unfortunately, the eigenvalue problem is significantly more challenging than in previous cases. However, we will outline how one could solve the eigenvalue problem in the system with noise and illustrate an important point about previous attempts at solving for the eigenvalue spectrum.

The PDE/ODE system we will consider is given by:

$$\frac{\partial}{\partial t}\rho_V(v,t) = -\frac{\partial}{\partial v}\left[G_1(v,\langle w\rangle,s)\rho_V(v,t) - \frac{\sigma^2}{2}\frac{\partial \rho_V(v,t)}{\partial v}\right] \tag{5.143}$$

$$s' = -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s}J(v_{peak},\langle w\rangle,s,t) \tag{5.144}$$

$$\langle w\rangle' = -\frac{\langle w\rangle}{\tau_w} + \frac{\lambda_w}{\tau_w}J(v_{peak},\langle w\rangle,s,t) \tag{5.145}$$

We will follow the same general scheme as in chapter 2.

The steady states are given by the following set of equations:

$$\rho_V(v,t) = \frac{2\,\overline{\langle \mathrm{R}\rangle}}{\sigma^2}\int_v^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}(M(v',s,w) - M(v,s,w))\right)\,dv' \tag{5.146}$$

$$s = \lambda_s\overline{\langle R\rangle} \tag{5.147}$$

$$w = \lambda_w\overline{\langle R\rangle} \tag{5.148}$$

$$\overline{\langle \mathrm{R}\rangle} = \frac{\sigma^2}{2}\left(\int_{v_{reset}}^{v_{peak}}\int_{v'}^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}(M(v'',\lambda_s\,\overline{\langle \mathrm{R}\rangle},\lambda_w\,\overline{\langle \mathrm{R}\rangle}) - M(v',\lambda_s\,\overline{\langle \mathrm{R}\rangle},\lambda_w\,\overline{\langle \mathrm{R}\rangle})\right)\,dv'dv''\right)^{-1} \tag{5.149}$$

As before, we will have to proceed with the Abbott-Vreeswijk transform. The purpose of the transform is to take the steady state density to a constant uniform distribution. Thus, the appropriate transform is given by

$$\begin{aligned} y &= \eta(v) = \frac{2\,\overline{\langle \mathrm{R} \rangle}}{\sigma^2} \int_{v_{reset}}^{v} \int_{v'}^{v_{peak}} \exp\left(-\frac{2}{\sigma^2} (M(v'', \lambda_s \overline{\langle \mathrm{R} \rangle}, \lambda_w, \overline{\langle \mathrm{R} \rangle}) - M(v', \lambda_s \overline{\langle \mathrm{R} \rangle}, \lambda_w \overline{\langle \mathrm{R} \rangle})) \right) dv'' dv' \\ &= \int_{v_{reset}}^{v} Q(v')\, dv' \end{aligned}$$

where

$$Q(v) = \int_{v}^{v_{peak}} \frac{2\,\overline{\langle \mathrm{R} \rangle}}{\sigma^2} \exp\left(-\frac{2}{\sigma^2} (M(v', \lambda_s \overline{\langle \mathrm{R} \rangle}, \lambda_w, \overline{\langle \mathrm{R} \rangle}) - M(v, \lambda_s \overline{\langle \mathrm{R} \rangle}, \lambda_w \overline{\langle \mathrm{R} \rangle})) \right) dv'$$

It should be clear that this transformation is well defined, $\eta(v_{peak}) = 1$, $\eta(v_{reset}) = 0$, and that the transformation is invertible as $\eta'(v) = Q(v) > 0$ for $v \in [v_{reset}, v_{peak})$. Also note that $Q(v_{peak}) = 0$, which will be used later to simplify some of the equations. We integrate (5.113) once to yield a partial differential equation for the distribution function:

$$\begin{aligned} \frac{\partial P_V(v,t)}{\partial t} &= -G_1(v,w,s) \frac{\partial P_V(v,t)}{\partial v} + \frac{\sigma^2}{2} \frac{\partial^2 P_V(v,t)}{\partial v^2} \\ &+ \left[G_1(v,w,s) \frac{\partial P_V(v,t)}{\partial v} \right] \Bigg|_{v=v_{peak}} - \frac{\sigma^2}{2} \frac{\partial^2 P_V(v,t)}{\partial v^2} \Bigg|_{v=v_{peak}} \end{aligned} \tag{5.150}$$

Now, we will need to apply the Abbott-Vreeswijk transform. The chain rule provides the necessary partial derivatives

$$\begin{aligned} \hat{P}(y,t) &= P_V(\eta^{-1}(y), t) \\ \frac{\partial P_V(v,t)}{\partial t} &= \frac{\partial \hat{P}(y,t)}{\partial t} \\ \frac{\partial P_V(v,t)}{\partial v} &= \frac{\partial \hat{P}(y,t)}{\partial y} Q(\eta^{-1}(y)) \\ \frac{\partial^2 P_V(v,t)}{\partial v^2} &= \frac{\partial^2 \hat{P}(y,t)}{\partial y^2} Q(\eta^{-1}(y))^2 + \frac{\partial \hat{P}(y,t)}{\partial y} Q'(\eta^{-1}(y)) \\ &= \frac{\partial^2 \hat{P}(y,t)}{\partial y^2} Q(\eta^{-1}(y))^2 + \frac{2}{\sigma^2} \frac{\partial \hat{P}(y,t)}{\partial y} \left(Q(\eta^{-1}(y)) G_1(\eta^{-1}(y), \lambda_w \overline{\langle \mathrm{R} \rangle}, \lambda_s \overline{\langle \mathrm{R} \rangle}) - \overline{\langle \mathrm{R} \rangle} \right) \end{aligned}$$

substituting these derivatives back into (5.150), while dropping the hats for convenience

yields the following

$$\begin{aligned}
\frac{\partial P(y,t)}{\partial t} &= -\frac{\partial P}{\partial y}Q(\eta^{-1}(y))\left(G_1(\eta^{-1}(y),\langle w\rangle,s)-G_1(\eta^{-1}(y),\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle})\right)-\overline{\langle \mathrm{R}\rangle}\frac{\partial P}{\partial y}\\
&+ \frac{\sigma^2}{2}Q(\eta^{-1}(y))^2\frac{\partial^2 P}{\partial y^2}+\left.\frac{\partial P}{\partial y}\right|_{y=1}Q(v_{peak})\left(G_1(v_{peak},\langle w\rangle,s)-G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle})\right)\\
&+ \overline{\langle \mathrm{R}\rangle}\left.\frac{\partial P}{\partial y}\right|_{y=1}-\frac{\sigma^2}{2}Q(v_{peak})^2\left.\frac{\partial^2 P}{\partial y^2}\right|_{y=1}\\
s' &= -\frac{s}{\tau_s}+\frac{\lambda_s}{\tau_s}\left(\left.\frac{\partial P}{\partial y}\right|_{y=1}Q(v_{peak})\left(G_1(v_{peak},\langle w\rangle,s)-G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle})\right)+\overline{\langle \mathrm{R}\rangle}\left.\frac{\partial P}{\partial y}\right|_{y=1}\right)\\
&- \frac{\lambda_s}{\tau_s}\left(\frac{\sigma^2}{2}Q(v_{peak})^2\left.\frac{\partial^2 P}{\partial y^2}\right|_{y=1}\right)\\
w' &= -\frac{w}{\tau_w}+\frac{\lambda_w}{\tau_w}\left(\left.\frac{\partial P}{\partial y}\right|_{y=1}Q(v_{peak})\left(G_1(v_{peak},\langle w\rangle,s)-G_1(v_{peak},\lambda_w\overline{\langle \mathrm{R}\rangle},\lambda_s\overline{\langle \mathrm{R}\rangle})\right)+\overline{\langle \mathrm{R}\rangle}\left.\frac{\partial P}{\partial y}\right|_{y=1}\right)\\
&- \frac{\lambda_w}{\tau_w}\left(\frac{\sigma^2}{2}Q(v_{peak})^2\left.\frac{\partial^2 P}{\partial y^2}\right|_{y=1}\right)
\end{aligned}$$

The next step is to consider perturbations off the steady states:

$$\begin{aligned}
\epsilon_s(t) &= s(t)-\lambda_s\overline{\langle \mathrm{R}\rangle} && (5.151)\\
\epsilon_w(t) &= \langle w\rangle-\lambda_w\overline{\langle \mathrm{R}\rangle} && (5.152)\\
\epsilon_y(t,y) &= P(y,t)-y,\quad \epsilon_y(0,t)=\epsilon_y(1,t)=0 && (5.153)
\end{aligned}$$

Applying these transformations yields the following

$$\begin{aligned}
\frac{\partial \epsilon_y(y,t)}{\partial t} &= -\left(\frac{\partial \epsilon_y}{\partial y}+1\right)Q(\eta^{-1}(y))(g\epsilon_s(e_r-\eta^{-1}(y))-\epsilon_w)-\overline{\langle \mathrm{R}\rangle}\left(\frac{\partial \epsilon_y}{\partial y}+1\right)+\frac{\sigma^2}{2}Q(\eta^{-1}(y))^2\frac{\partial^2\epsilon_y}{\partial y^2}\\
&+\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+1\right)Q(v_{peak})(g\epsilon_s(e_r-v_{peak})-\epsilon_w)+\overline{\langle \mathrm{R}\rangle}\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+1\right)-\frac{\sigma^2}{2}Q(v_{peak})^2\left.\frac{\partial^2\epsilon_y}{\partial y^2}\right|_{y=1}\\
&= L_1^\sigma(\epsilon(y,t),\epsilon_s(t),\epsilon_w(t))
\end{aligned}$$

with the other two components of the operator given by

$$\begin{aligned}
\epsilon_s' &= -\frac{\epsilon_s}{\tau_s}+\frac{\lambda_s}{\tau_s}\left(\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+1\right)Q(v_{peak})(g\epsilon_s(e_r-v_{peak})-\epsilon_w)+\overline{\langle \mathrm{R}\rangle}\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}-\frac{\sigma^2}{2}Q(v_{peak})^2\left.\frac{\partial^2\epsilon_y}{\partial y}\right|_{y=1}\right)\\
&= L_2^\sigma(\epsilon(y,t),\epsilon_s(t),\epsilon_w(t))\\
\epsilon_w' &= -\frac{\epsilon_w}{\tau_w}+\frac{\lambda_w}{\tau_w}\left(\left(\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}+1\right)Q(v_{peak})(g\epsilon_s(e_r-v_{peak})-\epsilon_w)+\overline{\langle \mathrm{R}\rangle}\left.\frac{\partial \epsilon_y}{\partial y}\right|_{y=1}-\frac{\sigma^2}{2}Q(v_{peak})^2\left.\frac{\partial^2\epsilon_y}{\partial y}\right|_{y=1}\right)\\
&= L_3^\sigma(\epsilon(y,t),\epsilon_s(t),\epsilon_w(t)).
\end{aligned}$$

We will analyze the operator $L^\sigma(\epsilon(y,t),\epsilon_s(t),\epsilon_w(t))$ on the same Banach space as the operator L in Chapter 4. As we have centered all the steady states to $\epsilon_y(y,t)=0$, $\epsilon_s(t)=\epsilon_w(t)=0$, we need to determine the linearization of this operator about $\mathbf{0}$. It should again be clear that the linearization at 0, $DL^\sigma(\mathbf{0})\boldsymbol{h}$ is given by

$$\begin{aligned}
(\boldsymbol{DL}^\sigma)_1 &= -Q(\eta^{-1}(y))(gh_2(e_r-\eta^{-1}(y))-h_3)-\overline{\langle \mathrm{R}\rangle}\frac{\partial h_1}{\partial y}+\frac{\sigma^2}{2}Q(\eta^{-1}(y))^2\frac{\partial^2 h_1}{\partial y^2}\\
&+ Q(v_{peak})(gh_2(e_r-v_{peak})-h_3)+\overline{\langle \mathrm{R}\rangle}\frac{\partial h_1}{\partial y}\bigg|_{y=1}-\frac{\sigma^2}{2}Q(v_{peak})^2\frac{\partial^2 h_1}{\partial y^2}\bigg|_{y=1}\\
(\boldsymbol{DL}^\sigma)_2 &= -\frac{h_2}{\tau_s}+\frac{\lambda_s}{\tau_s}\left(Q(v_{peak})(gh_2(e_r-v_{peak})-h_3)+\overline{\langle \mathrm{R}\rangle}\frac{\partial h_1}{\partial y}\bigg|_{y=1}-\frac{\sigma^2}{2}Q(v_{peak})^2\frac{\partial^2 h_1}{\partial y^2}\bigg|_{y=1}\right)\\
(\boldsymbol{DL}^\sigma)_3 &= -\frac{h_3}{\tau_w}+\frac{\lambda_w}{\tau_w}\left(Q(v_{peak})(gh_2(e_r-v_{peak})-h_3)+\overline{\langle \mathrm{R}\rangle}\frac{\partial h_1}{\partial y}\bigg|_{y=1}-\frac{\sigma^2}{2}Q(v_{peak})^2\frac{\partial^2 h_1}{\partial y^2}\bigg|_{y=1}\right).
\end{aligned}$$

The spectrum of eigenvalues is determined by solving the equation

$$\boldsymbol{DL}^\sigma\boldsymbol{h}=\mu\boldsymbol{h} \tag{5.154}$$

which yields the following system eigenvalue problem

$$\begin{aligned}
\mu h_1(y) &= -Q(\eta^{-1}(y))(gh_2(e_r-\eta^{-1}(y))-h_3)-\overline{\langle \mathrm{R}\rangle}\frac{\partial h_1}{\partial y}+\frac{\sigma^2}{2}Q(\eta^{-1}(y))^2\frac{\partial^2 h_1}{\partial y^2}\\
&+ \overline{\langle \mathrm{R}\rangle}\frac{\partial h_1}{\partial y}\bigg|_{y=1} \qquad (5.155)\\
\mu h_2 &= -\frac{h_2}{\tau_s}+\frac{\lambda_s}{\tau_s}\left(\overline{\langle \mathrm{R}\rangle}\frac{\partial h_1}{\partial y}\bigg|_{y=1}\right)\\
\mu h_3 &= -\frac{h_3}{\tau_w}+\frac{\lambda_w}{\tau_w}\left(\overline{\langle \mathrm{R}\rangle}\frac{\partial h_1}{\partial y}\bigg|_{y=1}\right).
\end{aligned}$$

where we have used $Q(v_{peak})=0$ to simplify the eigenvalue problem. Now, in order to actually proceed in solving the full eigenvalue problem, one has to be very careful and break the process down into steps.

Step 1. First, one has to solve the second order, inhomogeneous differential equation (5.155) by first solving the homogeneous equation by using a WKB approximation assuming that $\sigma\to 0$. This is valid based on the $\sigma^2/2$ parameter in the leading order derivative and the asymptotics on $Q(\eta^{-1}(y))$ as $\sigma\to 0$.

Step 2. Use the homogeneous WKB solution to approximate an inhomogeneous WKB solution via variation of parameters

Step 3. Finally this solution can be used to resolve the spectral equation and proceed as before

We leave this for future work. Also note that even the fast system cannot be solved without employing a WKB approximation, as the second order nature of equation (5.155) is still retained when $h_2 = h_3 = 0$.

It is worth noting that whatever the final solution is, it is in no way related to the spectrum of eigenvalues found in [2]. In [2], the authors consider models of the form

$$\dot{x}_i = F(x_i) + G(x_i)s(t) \tag{5.156}$$

where $x_i \in [x_{reset}, x_{peak}]$ is the voltage like variable and $s(t)$ is the coupling function which is of the simple-exponential type. In order to perform their analysis, the authors use the Abbott/Vreeswijk transform:

$$y_i = \eta(x_i) = \int_{x_{reset}}^{x_i} \frac{Edx}{F(x) + G(x)E} \tag{5.157}$$

where E is the steady state firing rate which satisfies

$$E = \left(\int_{x_{reset}}^{x_{peak}} \frac{dx}{F(x) + G(x)E} \right)^{-1}$$

When the transform (5.157) is applied to a population of neurons without noise, it yields the following

$$y_i = E + \Gamma(y_i)\epsilon(t) \tag{5.158}$$

where $\epsilon(t) = s(t) - E$ and

$$\Gamma(y) = \frac{EG(\eta^{-1}(y))}{F(\eta^{-1}(y)) + EG(\eta^{-1}(y))} \tag{5.159}$$

While equation (5.158) is equivalent to (5.156) when the Abbott-Vreeswijk transform is defined, the following two systems are not equivalent:

$$\dot{x}_i = F(x_i) + G(x_i)s(t) + \zeta_i \tag{5.160}$$

$$\dot{y}_i = E + \Gamma(y_i)\epsilon(t) + \zeta_i \tag{5.161}$$

where ζ_i is a white noise process. The authors conduct their analysis of the latter system, which is entirely analytically tractable, however it does not correspond to any neural system with a realistic source of noise as undoing the transform yields

$$\dot{x} = F(x_i) + G(x_i)s(t) + A(x_i)\zeta_i, \quad A(x) = \frac{F(x) + G(x)E}{E} \tag{5.162}$$

which has a multiplicative noise amplitude. Furthermore, the amplitude of the noise increases as the neuron reaches the upstroke of the action potential. This will likely have some very bizarre effects on spiking as the integrate-and-fire models that actually incorporate action potential generation do so with nonlinear systems with finite escape time, and thus $A(x) \to \infty$ as x becomes large.

5.3 Sparsity

We will again start with our general network:

$$v_i' = F(v_i) - w_i + g_i(t)(e_r - v_i) + I \tag{5.163}$$
$$w_i' = a(bv_i - w_i) \tag{5.164}$$
$$g_i(t) = \sum_{i=1}^{N} \bar{g}_{ij} s_j(t) \tag{5.165}$$

where $s_j(t)$ is any particular synaptic coupling function that neuron j makes to all its downstream targets and $\bar{g}_{ij}$ is the maximal conductance associated with that synapse. Sparsely coupled networks are those for which $\bar{g}_{ij} = 0$ for the majority of the neurons in the network. We will primarily consider Erdos-Reyni [150] networks, where

$$\bar{g}_{ij} = \begin{cases} \bar{g} & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases} \tag{5.166}$$

Thus, if $p \ll 1$, then $\bar{g}_{ij}$ is a sparse matrix where each element is a Bernoulli random-variable. With this simplification, the conductance variable becomes

$$g_i(t) = \sum_{i=1}^{N} \bar{g}_{ij} s_j(t) = \bar{g} \sum_{j'=1}^{N_i} s_{\sigma_i(j')}(t) \tag{5.167}$$

where $\sigma_i(j')$ is the index for the non-zero N_i connections made to neuron i. As g_{ij} is a Bernoulli random variable for $i, j = 1, 2, \ldots N$, then

$$N_i \sim \mathbb{B}(p, N)$$

where $\mathbb{B}(p, N)$ is a Binomially distributed random variable with parameters N and p. Note that unlike the all-to-all coupled cases before, we have not normalized by the number of incoming connections

5.3.1 The Limit of Vanishing Connections

Given equation (5.167) for $g_i(t)$, consider the following:

$$g_i(t) = \bar{g} \sum_{j'=1}^{N_i} s_{\sigma_i(j')}(t) = (\bar{g} N_i) \left(\frac{1}{N_i} \sum_{j'=1}^{N_i} s_{\sigma_i(j')}(t) \right) \tag{5.168}$$

If $Np \gg 1$ and $N(1-p) \gg 1$, then we should naturally expect

$$\frac{1}{N_i} \sum_{j'=1}^{N_i} s_{\sigma_i(j')}(t) \to \langle s \rangle, \quad N \to \infty$$

as for each neuron i, the number of connections becomes large enough, that once normalized, converges to the mean of s_i.

The remaining term in front is $\bar{g}N_i$. Given that $N_i \sim \mathbb{B}(p, N)$, then it is natural to consider the substitution

$$\bar{g} = \frac{g^*}{Np}. \tag{5.169}$$

The reason for this is that the distribution of N_i becomes more narrowly centered around the mean Np. This is due to the fact that the standard deviation of the binomial distribution is $\sqrt{Np(1-p)}$ which implies that the coefficient of variation goes to 0 in this limit:

$$CV = \frac{\sqrt{Np(1-p)}}{Np} = \frac{\sqrt{1-p}}{\sqrt{Np}}.$$

To avoid network saturation as $N \to \infty$, we require the unitary synaptic conductance to scale like $1/N$ and consider as our new parameter g^*. then we arrive at

$$g_i(t) = g^* \left(\frac{N_i}{Np}\right) \left(\frac{1}{N_i} \sum_{j'=1}^{N_i} s_{\sigma_i(j')}(t)\right) \to g^* \langle s \rangle \quad \text{as} \quad N \to \infty \tag{5.170}$$

and thus provided that Np and $N(1-p)$ are both large, the quantity $g^* = \bar{g}Np$ is invariant in the sense that two different large networks with identical g^* should display the same macroscopic behavior. Note that in this limit, the mean-field system is identical to the all-to-all coupled network mean-field system with $g = g^*$ and thus no further work is necessary. Other limits can be considered such as $p \propto N^{-1}$, for example, where the networks sparsity scales with its size. I will leave these for future work.

We can now use much of the mean-field infrastructure we have set up to analyze a large realistic network of model neurons.

5.4 Case Study: Theta Rhythms in the Hippocampus

In this section, we will leverage the speed in computing a mean-field system versus a network simulation in order to do a multi-dimensional parameter search. Recently, Izhikevich models have been fit to a variety of hippocampal CA1 pyramidal neurons [62]. The pyramidal neurons and the subsequent model fits display varying degrees of spike frequency adaptation. As we have observed slow frequency bursting when adapting neurons are coupled together, it is natural to consider adaptation induced bursting as a mechanism for theta oscillations in the hippocampus. However, CA1 networks are sparsely coupled, with the probability of any two neurons being connected to one another being typically around $p = 0.01$. Given this low degree of connectivity, it is uncertain whether the network bursting can still be sustained.

Thus, we will use a mean-field system to determine if bursting still occurs with this low degree of connectivity, and how robust it is with regards to perturbations in the parameters. The speed gains in simulating a mean-field system vs the actual network allow us to conduct a numerical assault by simulating a large number of mean-field systems over parameter meshes.

The Izhikevich models used here are of the k-switching type, as given by equation (5.142). The coupling is a Kinetic Type-I synapse, and thus we use the double-exponential approximation to approximate the Kinetic synapse. The networks we consider are sparse, with $N = 30,000$ and $p = 0.01$ fixed, Thus the free parameter is the unitary synaptic conductance, $\bar{g} = g^*/300$. The mean-field system is given by the following system of equations:

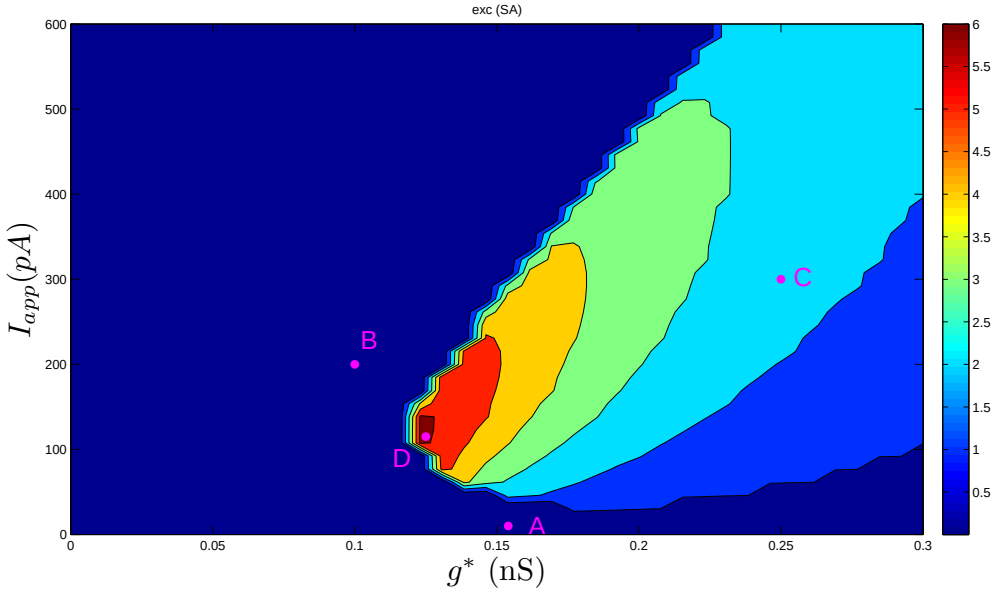
$$\begin{aligned}
\langle w\rangle' &= a(b\langle v\rangle - v_r - \langle w\rangle) + d\langle R_i(t)\rangle && (5.171)\\
\langle s\rangle' &= -\frac{\langle s\rangle}{\tau_R} + \langle h\rangle && (5.172)\\
\langle h\rangle' &= -\frac{\langle h\rangle}{\tau_D} + \frac{A}{\tau_R\tau_D}\langle R_i(t)\rangle && (5.173)\\
\langle R_i(t)\rangle &= \int_I \langle R_i(t)|I\rangle \rho_I(I)\, dI \\
\langle R_i(t)|I\rangle &= \begin{cases} \left(\int_{v_{reset}}^{v_{peak}} \frac{C_m}{k(V)(V-v_r)(V-v_t)-\langle w\rangle - g^*\langle s\rangle(V-E_r)}\right)^{-1} & I > I^*(\langle w\rangle, \langle s\rangle) \\ 0 & I < I^*(\langle w\rangle, \langle s\rangle) \end{cases} \\
\rho_V(v|I) &= \begin{cases} \frac{\langle R|I\rangle C_m}{k(V)(V-v_r)(V-v_t)-\langle w\rangle - g^*\langle s\rangle(V-E_r)} & I > I^*(\langle w\rangle, s) \\ \delta(v - v_-(I)) & I < I^*(\langle w\rangle, s) \end{cases} && (5.174)\\
\langle v\rangle &= \int_I \int_{v_{reset}}^{v_{peak}} v\rho_I(I)\rho_V(v|I)\, dv dI && (5.175)\\
I - I^*(\langle w\rangle, s) &= \min_{V\in[v_{reset}, v_{peak}]} k(V)(V-v_r)(V-V_t) + I - \langle w\rangle + g^*\langle s\rangle
\end{aligned}$$

where the various parameters/models are in their dimensional forms.

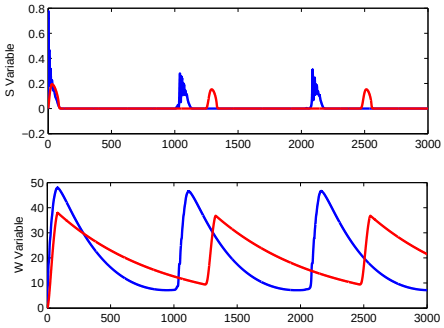
The parameters we consider are g^*, μ_I, σ_I, and τ_D the applied current to each neuron is a normally distributed random variable, I_i with mean μ_I and standard deviation σ_I. Finer meshes are run over the μ_I and g^* variables (40×40) , while coarser meshes are run over τ_D and σ_I (8×4). For example, the mean-field system is run over a 40×40 mesh, with $\sigma_I = 5$ pA and $\tau_D = 5$ ms as shown in figure 5.10. The mean-field system is compared to a simulated network with 10,000 neurons in 5.10. While there is a frequency error in the bursting, the mean-field system is able to predict both the onset of bursting in the parameter space, and general qualitative trends in the burst frequency and amplitude. The larger parameter search can be organized into parameter maps, as done in Figure 5.11 for a network of strongly adapting neurons and in 5.12 for a network of weakly adapting neurons. We have used these maps to guide numerical simulations of spiking networks with identical parameters and have found that the mean-field system can adequately predict both the

location of bursting in the parameter space, and the trends in the burst frequency, as shown in Figure 5.13. Thus, given this correspondence and the resulting maps in figure 5.11, it appears that low frequency bursting induced by spike frequency adaptation is robust to sparsity and heterogeneity. Without a corresponding mean-field system, this result would require possibly hundreds of direct network simulations. This would be unfortunate, given the length of a single 30,000 neuron network simulation being on the order of hours.

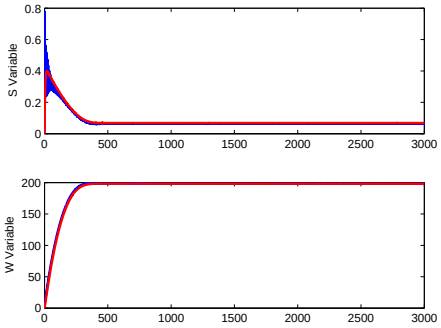
Through our mean-field facilitated parameter search, network simulations in figure 5.13 demonstrate prominent theta oscillations in the lower theta range (2-4 Hz). Furthermore, we have found that the bursting is robust with regards to sparse coupling and heterogeneity. Thus, adaptation induced bursting may be a potential mechanism for the presence of theta oscillations in hippocampal area CA1.



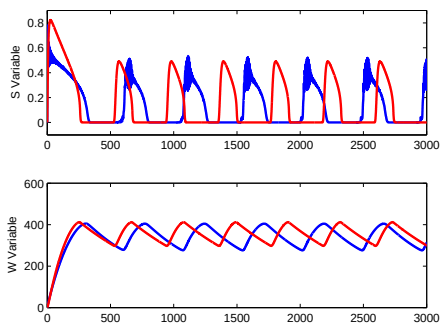
(a) A 40×40 mesh of 3 second long mean-field simulations of the network.



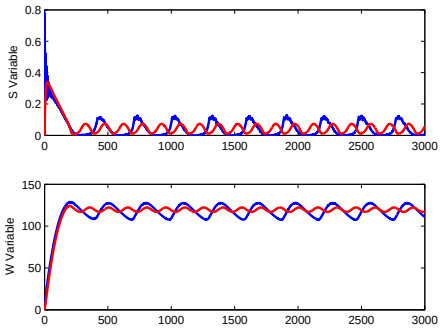
Point A: $g^* = 0.15$ nS, pA, $I = 10$ pA



Point B: $g^* = 0.1$ nS, pA, $I = 200$ pA



Point C: $g^* = 0.25$ nS, pA, $I = 300$ pA



Point D: $g^* = 0.13$ nS, pA, $I = 110$ pA

Figure 5.10: A network of 10,000 neurons with 1% sparsity is simulated for 3 seconds of time (blue), in addition to the mean-field system outlined in the text (red) for a variety of different parameters. The parameters are chosen from a 40×40 mesh (in figure 5.10(a)) of mean-field simulations that are run over a period of time.

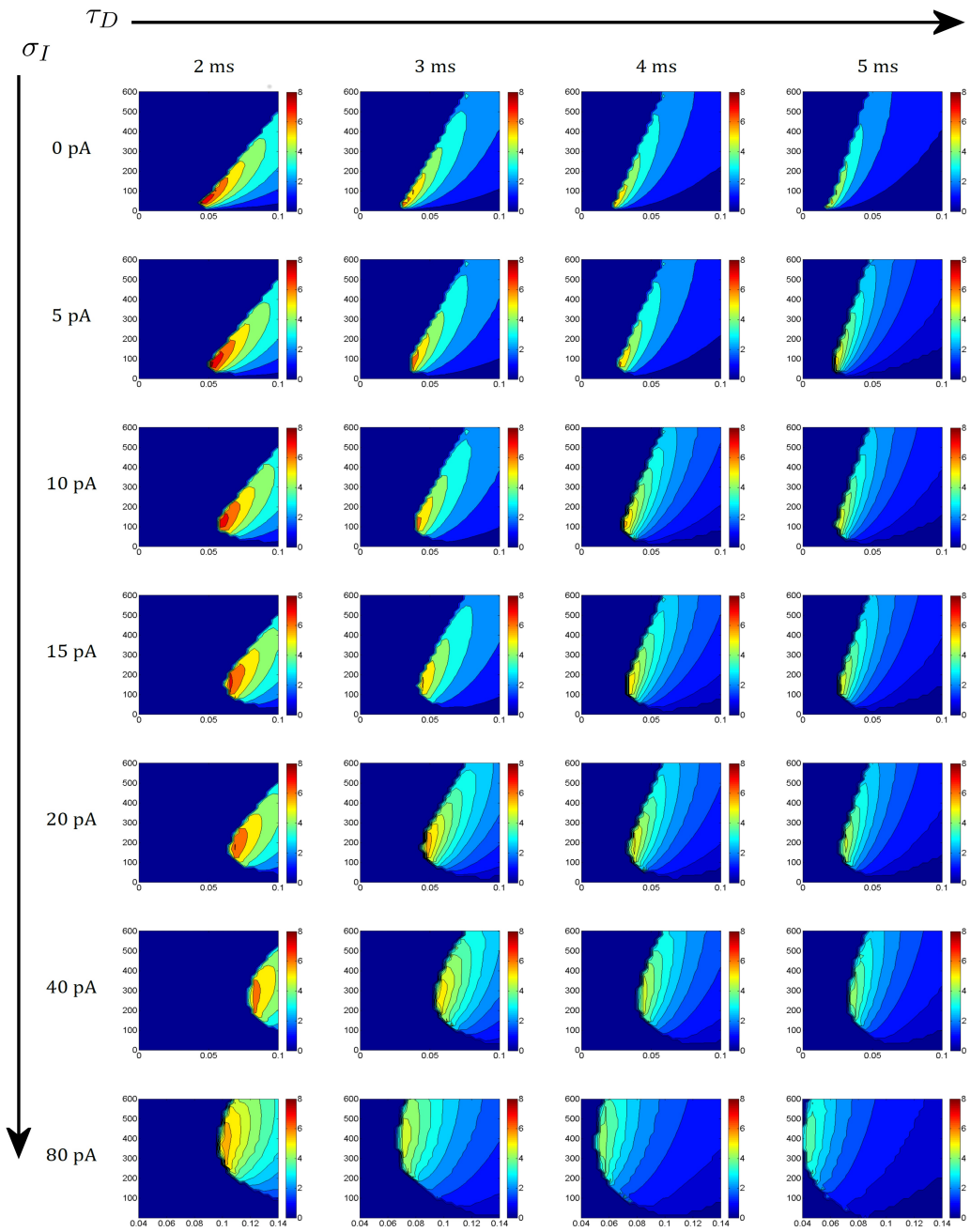


Figure 5.11: Shown above is a mean-field parameter map where the neurons are Izhikevich neurons that are fitted to strongly adapting hippocampal CA1 pyramidal neurons. Each panel is a 40×40 grid of simulations of the mean-field system (5.171)-(5.173). The burst frequency is computed using the peakfinder function from the matlab file exchange and filled, coloured contours of the burst frequency are plotted. Bluer colours indicate lower frequencies. The mean-field system predicts a large region of theta oscillations even with low connectivity.

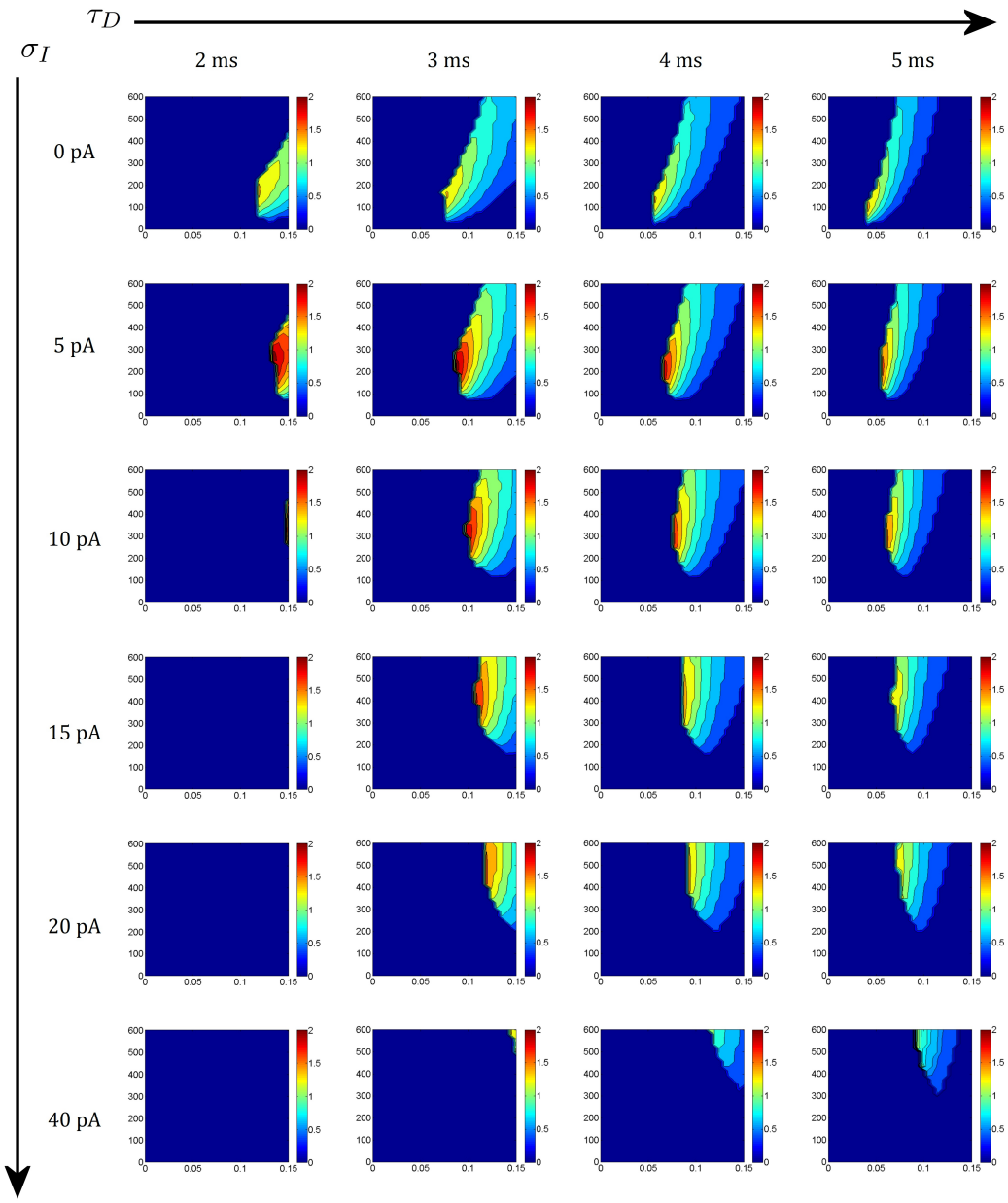
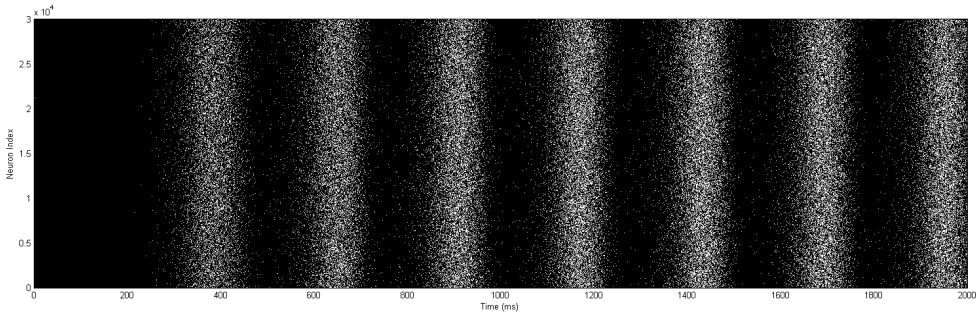
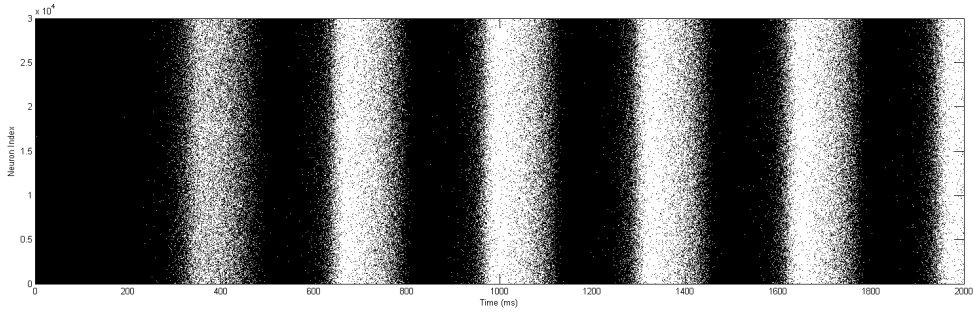


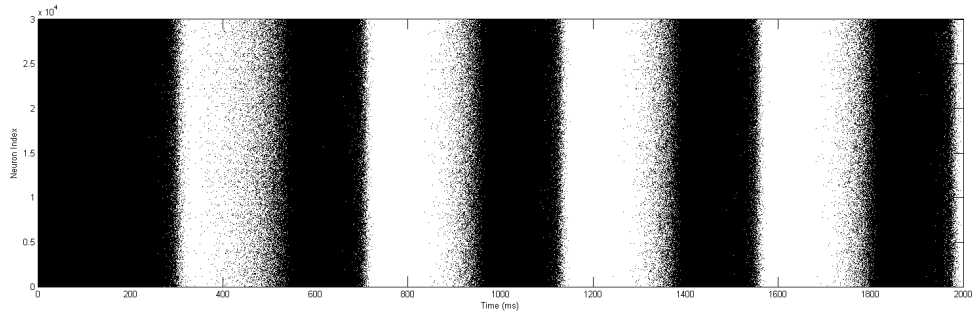
Figure 5.12: Shown above is a mean-field parameter map where the neurons are Izhikevich neurons that are fitted to weakly adapting hippocampal CA1 pyramidal neurons. Each panel is a 40×40 grid of simulations of the mean-field system (5.171)-(5.173). The burst frequency is computed using the peakfinder function from the matlab file exchange and filled, coloured contours of the burst frequency are plotted. Bluer colours indicate lower frequencies. The mean-field system predicts bursting frequencies that are outside of the theta range which eventually disappear for increasing amounts of neuronal heterogeneity.



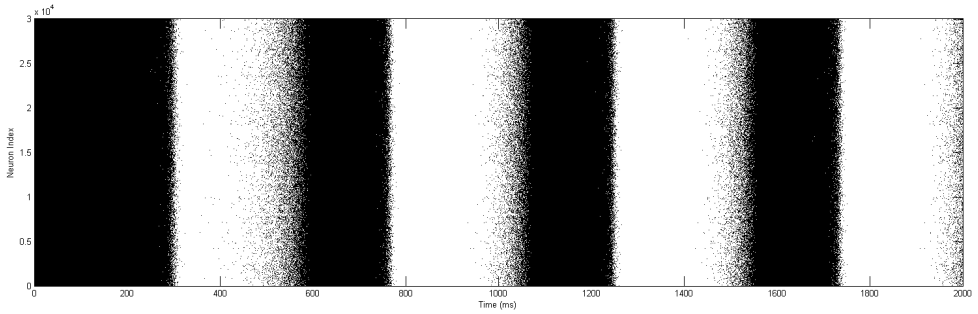
(a) $g = 0.02$ nS, $\mu_I = 250$



$g = 0.025$ nS, $\mu_I = 250$



$g = 0.03$ nS, $\mu_I = 250$



$g = 0.035$ nS, $\mu_I = 250$

Figure 5.13: A network of 30,000 neurons with 1% sparsity is simulated for 2 seconds of time using parameters determined by the mean-field map in figure 5.11. The g, I parameters were determined as in the sub-caption, while $\sigma_I = 10$ pA and $\tau_D = 3$ ms. The onset of bursting and the frequency trend as g is increased show broad agreement with the mean-field map in figure 5.11.

Chapter 6

Bifurcation Analysis of the Mean-Field System for a Homogeneous Network of Neurons

Here we explore, both analytically and numerically, many of the smooth and non-smooth bifurcations and phenomena that occur in the mean field system of equations from [134]. The primary mean field system we use is that of the Izhikevich model coupled in a homogeneous network with simple-exponential coupling. The smooth bifurcation analysis of this system is performed in section 6.2 while non-smooth bifurcations are explored in 6.3. We modify the parameters slightly as the neuronal model used in [49] was an alteration of the default Izhikevich model to better fit the action potential half-width observed in the data. We use this model primarily for two reasons: it is the most analytically tractable and the parameters have been fit to neuronal data. However, as we will see, many of the non-smooth bifurcations are present in the other models in the general class of two-dimensional adapting integrate and fire neurons. Moreover, the bifurcations are also present in the direct network simulations, as shown in 6.4. Whenever possible we present our results in terms of this general class. Finally, we will also numerically analyze the effects of noise and heterogeneity on the mean-field systems for the corresponding networks by using the numerical bifurcation software such as MATCONT or XPPAUT [47, 56]. This is done in section 6.5

6.1 The Network Equations

The network equations we consider are

$$\dot{v}_i = F(v_i) - w + I + gs(t)(e_r - v_i) = G_1(v_i, w_i, s) \tag{6.1}$$

$$\dot{w}_i = \frac{1}{\tau_w}(bv_i - w_i) = G_2(v_i, w_i) \tag{6.2}$$

$$\dot{s} = -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s N}\sum_{j=1}^{N}\sum_{t_{j,k}<t}\delta(t - t_{j,k}) \tag{6.3}$$

$$v_i(t^-_{spike}) = v_{peak} \rightarrow \begin{array}{rcl} v_i(t^+_{spike}) & = & v_{reset} \\ w_i(t^+_{spike}) & = & w_i(t^-_{spike}) + w_{jump}. \end{array} \tag{6.4}$$

where $\lambda_s = \tau_s s_{jump}$, $\lambda_w = \tau_w w_{jump}$. The specific forms of $F(v)$ we consider are:

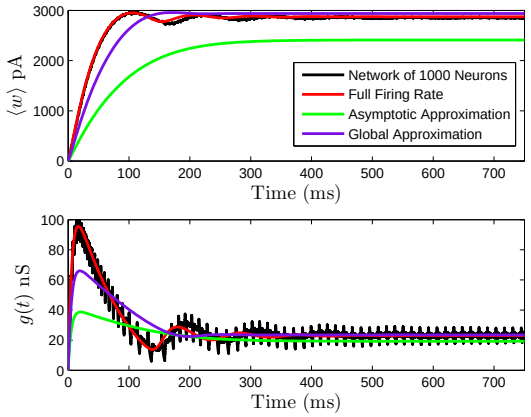
$$\begin{aligned} F(v) &= v(v-\alpha) \quad \text{(Izhikevich Model)} \\ F(v) &= e^v - v \quad \text{(Adaptive Exponential Model)} \\ F(v) &= v^4 - \frac{2v}{\tau_w} \quad \text{(Quartic Model)} \end{aligned}$$

These forms can be arrived at through a suitable non-dimensionalization of the original equations for these models [177]. Note that the non-dimensionalization for the Izhikevich model differs from the one used by [177] and is from [134]. We will use identical assumptions on the general form of $F(v)$ as in [177]:

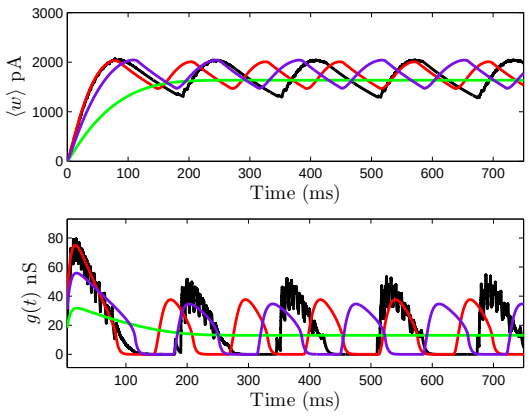
$$\lim_{v\to\infty} F'(v) = \infty, \quad \lim_{v\to-\infty} F'(v) < 0, \quad F''(v) > 0. \tag{6.5}$$

These networks often display bursting, an oscillatory behaviour where the individual neurons alternate between firing and quiescence [49, 134]. The other common behaviour is tonic firing, where the neurons all fire at a constant rate. The transition between these two behaviours is a bifurcation of the full network. Examples of this transition for networks of 1000 neurons with all-to-all coupling and parameters as in Table 6.1 are shown in Figure 7.1. In Figure 7.1(a) and 7.1(c) the neurons in the network fire spikes, and the mean-adaptation variable, w, and the synaptic coupling variable, s both converge to a stable steady state. In Figure 7.1(b) and 6.1(d) the neurons fire synchronized bursts, and the pair of variables (w, s) converge to a steady state limit cycle, representing the oscillation between firing and quiescence that the individual neurons undergo. This occurs as the current I is decreased.

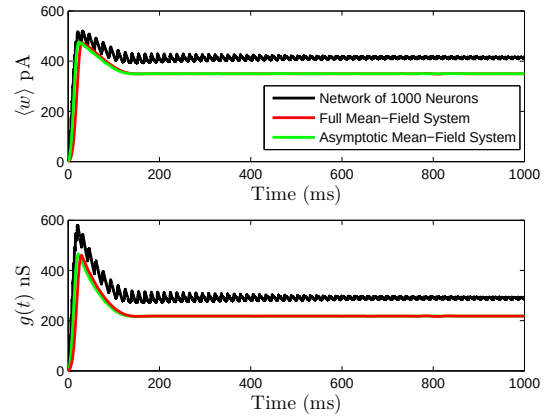
We will conduct our analysis on the mean-field system as it is greatly predictive of the behavior of the full network, as shown in Figure 7.1. The mean-field system for the



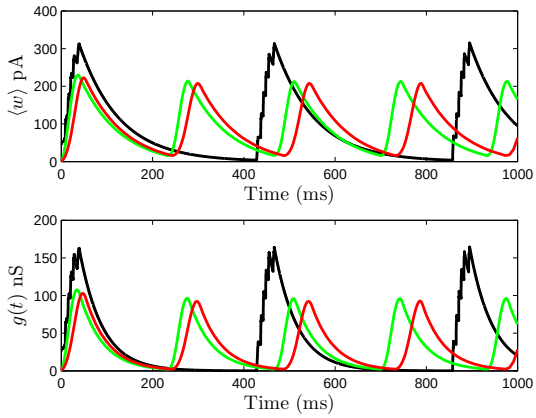
(a) Izhikevich, Tonic Firing, $I = 0.33$, $g = 0.61$ (dimensionless)



(b) Izhikevich, Bursting, $I = 0.24$, $g = 0.61$ (dimensionless)



(c) Adaptive Exponential, Tonic Firing, $I = 4.25$, $g = 930$ (dimensionless)



(d) Adaptive Exponential, Bursting, $I = 4.25$, $g = 465$ (dimensionless)

Figure 6.1: Network simulations consisting of 1000 neurons (black) versus the mean-field system of equations (6.6)-(6.8) (red), the system of equations with the asymptotically simplified firing rate (6.13) (green) for a network of Izhikevich neurons (a),(b) and Adaptive Exponential network of neurons (c),(d). The asymptotic firing rate, $\langle R_i(t) \rangle \sim \sqrt{I - I^*(s,w)}\sqrt{F''(v^*(s))}$, is a good approximation for the network of adaptive exponential integrate and fire neuron away from the switching manifold $I - I^*(s,w) = 0$. This is not the case for the network of Izhikevich neuron, which is better approximated when a global approximation, $k\sqrt{I - I^*(s,w)}\sqrt{F''(v^*(s))}$, to the firing rate is used (green). The parameter k in this global approximation is fitted to approximate the full firing rate for a large set of (s,w).

Parameter	Izhikevich Network ([49])	AdEx Network ([23])	QIF Network ([177])
g	0-4	0-1000	0-40
I	0-0.4	-1 - 12	0-40
τ_s	2.6	2.06	2
τ_w	130	3.63	50
$\hat{s}$	0.8	0.5	1
$\hat{w}$	0.0189	21.92	0.36
e_r	1	5	2
α	0.62	N/A	N/A
v_{peak}	1.46	65	10
v_{reset}	0.15	-1.25	0

Table 6.1: Parameters for various network types and the mean field systems. Note that the parameters above are dimensionless, while in some of the cited sources they are in dimensional form only. The non-dimensionalization for the AdEx and quartic neuron can be found in [177] while the non-dimensionalization for the Izhikevich model can be found in [134]

network (6.1)-(6.4) is given by the equations:

$$s' = -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s}\langle R_i(t)\rangle \tag{6.6}$$

$$w' = -\frac{w}{\tau_w} + \frac{\lambda_w}{\tau_w}\langle R_i(t)\rangle \tag{6.7}$$

$$\langle R_i(t)\rangle = \begin{cases} \left[\int_V \frac{dv}{G_1(v,s,w)}\right]^{-1} & \text{if} \quad H(w,s) > 0 \\ 0 & \text{if} \quad H(w,s) \leq 0 \end{cases}. \tag{6.8}$$

$$H(s,w) = \min_{v\in[v_{reset},v_{peak}]} G_1(v,s,w) \tag{6.9}$$

Here w and s correspond to the mean network adaptation and global synaptic coupling variable. Note that we have omitted the $\langle\rangle$ brackets denoting the average value of w present in [134] and elsewhere in the thesis for simplicity and clarity. Additionally, we have set $b = 0$ as one can show with suitable non-dimensionalization that it is small [134] and has been found to be unimportant with regards to the dynamics in standard parameter regimes in [85]. The function $\langle R_i(t)\rangle$ is the instantaneous network averaged firing rate, as a function of s and w. The function H determines when the integral in (6.8) is well-defined. It defines the *switching manifold* of the non-smooth system (6.6)–(6.8).

One can derive an expression for the switching manifold equation by determining when the denominator in (6.8) first becomes zero somewhere in the s, w phase space. We must find the minimum of $G(v, s, w)$, for $v \in [v_{reset}, v_{peak}]$, regarding s and w as fixed parameters. For the general class of models studied in [177], the function $F(v)$ is assumed to be strictly convex, that is $F''(v) > 0$. It follows that $G(v, s, w)$ is also strictly convex as a function

of v, and hence its minimum on $[v_{reset}, v_{peak}]$ occurs at a critical point. The critical points (as a function of v) are given by solving

$$\begin{aligned} \frac{\partial G}{\partial v} &= F'(v) - gs = 0 \Rightarrow \\ F'(v^*(s)) &= gs \end{aligned} \tag{6.10}$$

Thus $v^*(s)$ is the value of v at which G has a minimum. It may be the case that $v^* \notin [v_{reset}, v_{peak}]$. We will primarily ignore this particular case as typical parameter values usually ensure that $v^* \in [v_{reset}, v_{peak}]$.

The minimum value defines the function H

$$H(s,w) = G(v^*(s), s, w) = F(v^*(s)) - w + gs(e_r - v^*(s)) + I, \tag{6.11}$$

and the switching manifold equation

$$0 = H(s,w) = I - w + F(v^*(s)) + gs(e_r - v^*(s)) = I - I^*(s,w).$$

This latter expression is useful as we can think of I^* as an $s-$ and $w-$ dependent rheobase current. Anywhere in the phase space where $I - I^*(s,w) > 0$ the network is firing with mean firing rate given by

$$\langle R_i(t) \rangle = \left[\int_V \frac{dv}{F(v) - w + I + gs(e_r - v)} \right]^{-1}. \tag{6.12}$$

Anywhere that $I - I^*(s,w) \leq 0$ the network is quiescent and the mean firing rate is 0.

There are two important facts to note before we proceed further. First of all, $I^*(0,0) = I_{rh}$, the rheobase current for the uncoupled, non-adapting neuron, which is governed by the equation

$$\dot{v} = F(v) + I.$$

Based on the assumptions made on $F(v)$ in (6.5), this model neuron has a type-I firing profile. Additionally, given that $I^*(0,0) = -F(v^*(0))$, we have the following

$$F(v^*(0)) = -I_{rh}.$$

These two facts will prove important for our later analysis.

To conclude we display some expressions for specific models. The rheobase currents are given by

$$\begin{aligned} I^*(s,w) &= w - gse_r + \frac{(\alpha + gs)^2}{4} \quad \text{(Izhikevich)} \\ I^*(s,w) &= w - gse_r + (1+gs)(\log(1+gs) - 1) \quad \text{(AdEx)} \\ I^*(s,w) &= w - gse_r + 3\left(\frac{gs + 2a}{4}\right)^{\frac{4}{3}} \quad \text{(Quartic)} \end{aligned}$$

with corresponding minimum values of G:

$$\begin{aligned} v^*(s) &= \frac{\alpha + gs}{2} \\ v^*(s) &= \log(1+gs) \\ v^*(s) &= \left(\frac{gs+2a}{4}\right)^{1/3} \end{aligned}$$

For the Izhikevich model, the mean firing rate can be evaluated analytically:

$$\langle R_i(t)\rangle = \frac{\sqrt{I - I^*(s,w)}}{\arctan\left(\frac{v_{peak}-\frac{1}{2}(\alpha+gs)}{\sqrt{I-I^*(s,w)}}\right) - \arctan\left(\frac{v_{reset}-\frac{1}{2}(\alpha+gs)}{\sqrt{I-I^*(s,w)}}\right)} \quad \text{(Izhikevich mean firing rate)}$$

For the other models, the firing rate must be evaluated numerically. This can be done by integrating equation (6.12) over $[v_{reset}, v_{peak}]$ treating w, and s as fixed parameters at each time step. This approach can be used to numerically analyze the bifurcation types of these equations using numerical bifurcation software, such as MATCONT or XPPaut [47, 56]. However, the numerical integration method should be of high enough order accuracy for the numerical continuation results to be trusted.

Given the mean field system described above, one should consider whether numerical bifurcation or analytical approaches should be taken. Numerical bifurcation analysis can yield results which are accurate throughout the phase space, but require choosing a particular model and determining which parameters to fix and which to vary. Analytical methods can yield model independent results and give insight into the role of various parameters in system behaviour, however, they are often restricted to particular regions of the parameter space and/or phase space, as we shall see.

6.2 Smooth Bifurcations

In order to proceed analytically, we need to sacrifice some of the complexity of the original mean field system. In particular, as the usual formulas of $\langle R_i(t)\rangle$ are difficult to deal with analytically, we need to approximate the firing rate with a simpler alternative. This was done in Chapter 2, where we showed the following:

$$\langle R_i(t)\rangle \sim \frac{1}{\sqrt{2}\pi}\sqrt{F''(v^*(s))}\sqrt{I - I^*(s,w)} \tag{6.13}$$

This reduction is valid when $I - I^*(s,w)$ is small, that is when the dynamics are near the region in the (s,w) plane where they become non-smooth. We note that a similar approximation appears in [54, 61] and is effectively due to the topological normal form for Type-I firing being a particular quadratic-integrate-and-fire/theta neuron [60]. However, to

the best of our knowledge the non-smooth nature of those equations has not been explored. For example, equations (3.6)-(3.7) in [54] are similar to ours however the interpretation for those particular equations was for the firing rate of an E/I coupled pair of neurons.

For the Izhikevich neuron, equation 6.13 becomes

$$\langle R_i(t)\rangle \sim \frac{1}{\pi}\sqrt{I - I^*(s,w)} \quad \text{as} \quad I - I^*(s,w) \to 0 \tag{6.14}$$

while for the AdEx, we have

$$\langle R_i(t)\rangle \sim \frac{1}{\pi}\sqrt{\frac{1+gs}{2}}\sqrt{I - I^*(s,w)} \quad \text{as} \quad I - I^*(s,w) \to 0 \tag{6.15}$$

where the switching manifolds differ depending on the neuron model. One can see the validity of the approximations in Figure 7.1 which compares the behaviour of full network simulations with that of the mean-field model using the true firing rate and asymptotic approximation. The validity is further studied in Figure 7.2 which demonstrates that the difference between (6.12) and (6.13) is $o(\sqrt{I - I^*(s,w)})$ pointwise in (s,w) for the AdEX and Izhikevich systems.

The approximation (6.13) turns out to yield a system that is tractable to analysis of both the smooth and non-smooth bifurcations, and shows considerable accuracy when compared with both the actual network and the original mean field system in the vicinity of the switching manifold (see Figure 7.1). Additionally, when the system is not near the switching manifold, one can use an ad-hoc global approximation that still preserves the bifurcation structure near the switching manifold. For example, one can use

$$\langle R_i(t)\rangle = k\sqrt{F''(v^*(s))}\sqrt{I - I^*(s,w)} \tag{6.16}$$

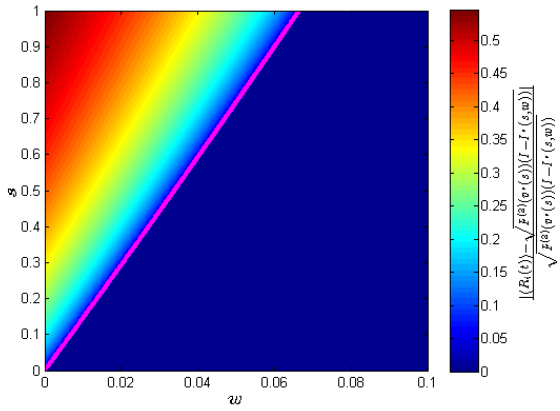
where k is fit to globally approximate the original firing rate. This can be done for example by plotting $\langle R_i(t)\rangle$ vs. $\sqrt{I - I^*(s(t),w(t))}$ for a small number of trajectories in the parameter plane and performing a linear fit. This is done for the Izhikevich network in Figure 7.1 with $k = 1/2$. One could use this approximation, for example, to conduct a rapid parameter search that is more global than the local analysis we present below. Alternative approaches for a global fit are also possible. The advantage of the approximation (6.16) is that is preserves the location of the co-dimension 2 non-smooth bifurcation points we describe later in this chapter.

With the simplification (6.13), the approximate mean field system that we analyze is given by

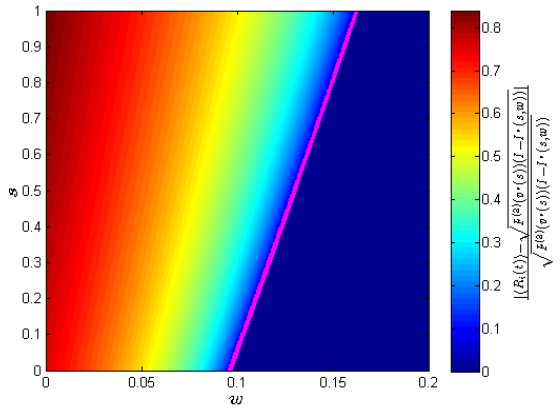
$$\dot{s} = f(s,w) = -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s}\langle R_i(t)\rangle \tag{6.17}$$

$$\dot{w} = g(s,w) = -\frac{w}{\tau_w} + \frac{\lambda_w}{\tau_w}\langle R_i(t)\rangle \tag{6.18}$$

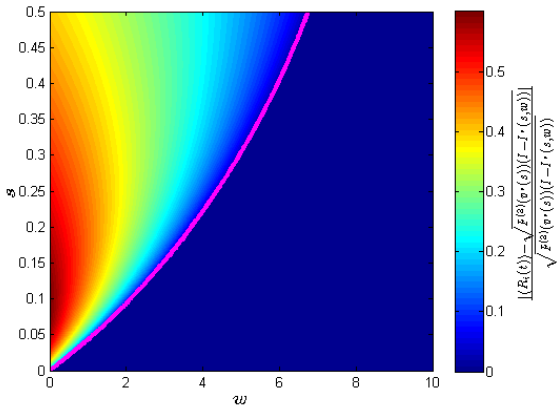
$$\langle R_i(t)\rangle = \begin{cases} \sqrt{F''(v^*(s))}\sqrt{I - I^*(s,w)} & I - I^*(s,w) > 0 \\ 0 & I - I^*(s,w) \le 0 \end{cases} \tag{6.19}$$



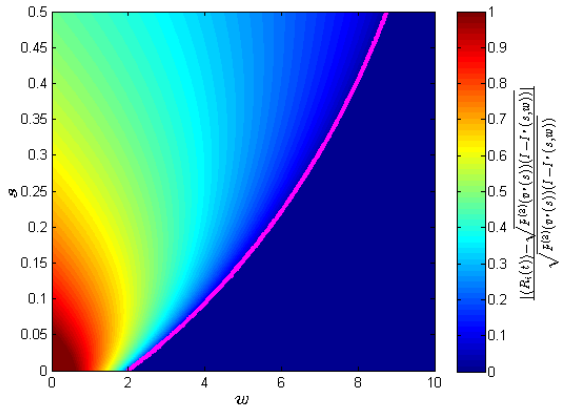
(a) Izhikevich Firing Rate, $I = I_{rh}$



(b) Izhikevich Firing Rate, $I > I_{rh}$]



(c) Adaptive Exponential Firing Rate, $I = I_{rh}$



(d) Adaptive Exponential Firing Rate, $I > I_{rh}$

Figure 6.2: Variation with (s, w) of the relative error in using the leading order asymptotic expansion, $\langle R_i(t) \rangle \sim \sqrt{F''(v^*(s))}\sqrt{I - I^*(s,w)}$, for the firing rate. There is always a neighborhood in the vicinity of the switching manifold, $I - I^*(s,w) = 0$, (magenta curve) where the firing rate is well approximated by the asymptotic expansion. This is the neighbourhood where the non-smooth bifurcations occur and to which we restrict our analysis.

where the switching manifold varies depending on which neuron model is used. We have absorbed any constant terms into λ_s and λ_w. In the following, we will refer to equations (6.17)-(6.19) as the reduced mean field system. The system is smooth and has derivatives of all orders everywhere except on the switching manifold, i.e., when $I - I^*(s,w) = 0$. On the switching manifold, the system is continuous but not differentiable. Thus, this is a piecewise-smooth continuous (PWSC) system. Equivalently, it has a uniform order of discontinuity of 1 [48]. Note that the derivatives of the vector field diverge at the switching manifold. Thus one cannot simply unfold the system around a boundary equilibrium bifurcation points as is often done in the literature [164], at least not without including the relevant square-root terms.

In the following sections we will carry out a detailed bifurcation study of the mean field system (6.17)-(6.18). Before proceeding we consider when our results will give information about the coupled PDE-ODE system (4.118)-(4.120) and hence about the original large network model. It is straightforward to see that an equilibrium point of the mean field system is a leading order approximation (as $I - I^*(s,w) \to 0$) of an equilibrium point of the full mean-field system (6.6)-(6.8). Further, any equilibrium point of the full mean-field system corresponds to an $O(1)$ (w.r.t ϵ) approximation of a steady state solution of the original ODE-PDE system given by (4.118)-(4.120). We showed in chapter 2 that the real parts of the eigenvalues were all negative provided that $I - I^*(\lambda_s \overline{\langle R \rangle}, \lambda_w \overline{\langle R \rangle})$ was small and $(g - \frac{\eta\gamma}{e_r - v^*(0)})\lambda_s(e_r - v^*(0)) > 0$ for $\tau_s, \tau_w \gg 1$. Thus, the stability analysis conducted here can be informally thought of as a bifurcation analysis of the moment-closure reduced population density equation provided that these conditions hold. A more formal approach would require the infinite-dimensional center manifold theorem.

6.2.1 Existence and Linear Stability of Equilibria

The equilibria of the mean field equations (6.17)-(6.18) depend on the sign of $I - I^*(s,w)$.

If $I - I^*(s,w) \leq 0$ then the only equilibrium point is the trivial solution, $e_0 = (0,0)$, which is a stable node. This equilibrium corresponds to all the neurons being quiescent, $\langle R_i(t) \rangle = 0$, thus we will refer to it as the non-firing solution. It will only exist when the origin of the phase space lies in the region where $I - I^*(0,0) \leq 0$, which corresponds to $I \leq I_{rh}$. Alternatively, in the language of non-smooth dynamical systems theory, e_0 is virtual if $I > I_{rh}$ and real if $I \leq I_{rh}$ [48].

If $I - I^*(s,w) > 0$, nontrivial equilibria (s,w) may exist. If they do then s, w must satisfy

$$s = \lambda_s \sqrt{F''(v^*(s))} \sqrt{I - I^*(s,w)} \tag{6.20}$$

$$w = \lambda_w \sqrt{F''(v^*(s))} \sqrt{I - I^*(s,w)}. \tag{6.21}$$

Equations (6.20) and (6.21) yield the following relationship

$$w = \frac{\lambda_w}{\lambda_s} s = \eta s. \tag{6.22}$$

Thus the equilibria are given by $(s, \eta s)$ where s satisfies the nonlinear equation

$$\frac{s}{\sqrt{F''(v^*(s))}} = \lambda_s \sqrt{I - I^*(s, \eta s)}. \tag{6.23}$$

We will label solutions to equation 6.23 as $\bar{s}$, and the full steady states as $e = (\bar{s}(g, I), \eta \bar{s}(g, I))$. Note that equation (6.20) implies that $\bar{s} = \lambda_s \sqrt{I - I^*(\bar{s}, \eta \bar{s})} \sqrt{F''(v^*(\bar{s}))} \geq 0$. Thus for an equilibrium to be a valid, it must satisfy $s \geq 0$ (which implies $w \geq 0$).

The equilibrium condition (6.23) for the quartic and AdEx models yield nonlinear equations without analytic closed form solutions. However, one can apply a power series (assuming that $\bar{s}$ is small) to come up with an approximation to the steady solutions. This series is justified by the fact that when one considers the steady state conditions on $\bar{s}$ and $\bar{w}$, $\bar{s}$ being small is equivalent to the equilibrium point being in the vicinity of the switching manifold. Note that v^* is actually a function of gs, as opposed to just s, as it is given by solving the algebraic equation (6.10). Thus, we can write down the following expansions for $v^*(s)$ and $F(v^*(s))$

$$\begin{aligned} v^*(s) &= v^*(0) + v^{*\prime}(0) gs + O((gs)^2) && (6.24) \\ F(v^*(s)) &= F(v^*(0)) + v^{*\prime}(0) \frac{(gs)^2}{2} + O((gs)^3) \\ &= -I_{rh} + v^{*\prime}(0) \frac{(gs)^2}{2} + O((gs)^3) && (6.25) \end{aligned}$$

where (6.25) can be derived using the relationship (6.10). Using these expansions in (6.23), we arrive at the following equation for the equilibria:

$$\begin{aligned} \frac{s^2}{\lambda_s^2 F''(v^*(0))} &= I - I_{rh} - v^{*\prime}(0) \frac{(gs)^2}{2} + gs(e_r - v^*(0)) - \eta s \\ 0 &= s^2 \left(\frac{1}{F''(v^*(0)) \lambda_s^2} + \frac{g^2}{2F''(v^*(0))} \right) + s\left(\eta - g(e_r - v^*(0))\right) + I_{rh} - I + O(s^3) \\ 0 &= A_2(g) s^2 + A_1(g) s + A_0 + O(s^3) \end{aligned}$$

Neglecting the higher order terms, this equation yields two solution branches:

$$\bar{s}_\pm = -\frac{A_1(g)}{2A_2(g)} \pm \sqrt{\frac{A_1(g)^2}{4A_2(g)^2} - \frac{A_0}{A_2(g)}}.$$

We will denote the corresponding equilibria as $\bar{e}_\pm = (\bar{s}_\pm, \bar{w}_\pm) = (\bar{s}_\pm, \eta \bar{s}_\pm)$. Defining the new parameters

$$\begin{aligned} \tilde{I} &= -\frac{A_0}{A_2(g)} = \frac{I - I_{rh}}{A_2(g)} && (6.26) \\ \beta &= -\frac{A_1(g)}{2A_2(g)} = \frac{g(e_r - v^*(0)) - \eta}{2A_2(g)} && (6.27) \\ &= \frac{(e_r - v^*(0))}{2A_2(g)} \left(g - \frac{\eta}{e_r - v^*(0)} \right) = M(g)(g - g^*) && (6.28) \end{aligned}$$

the s variable of the solution branches may be written

$$\bar{s}_{\pm}(\beta, \tilde{I}) = \beta \pm \sqrt{\beta^2 + \tilde{I}}. \tag{6.29}$$

Note that $A_2(g) > 0$. Further, since $v^*(0)$ is the minimum of $F(v)$ and the reversal potential for an excitatory synapse is above the resting membrane potential, v_r, we have $e_r > v_r > v^*(0)$. It follows that $M(g)$ is a strictly positive function.

The expressions above give the simplest approximation for the nontrivial equilibria of the two dimensional integrate and fire models. In fact, the approximation is exact in the case of the Izhikevich model as all higher order terms in the expansions (6.24)-(6.25) vanish. The exact expressions for the Izhikevich model are given by:

$$\bar{s}_{\pm}(g, I) = \frac{-(\eta - g(e_r - \frac{\alpha}{2})) \pm \sqrt{(\eta - g(e_r - \frac{\alpha}{2}))^2 + 4(I - \frac{\alpha^2}{4})(\frac{1}{\lambda_s^2} + \frac{g^2}{4})}}{\frac{1}{\lambda_s^2} + \frac{g^2}{4}}$$

Introducing the parameters

$$\tilde{I} = \frac{(I - \frac{\alpha^2}{4})}{\frac{1}{\lambda_s^2} + \frac{g^2}{4}} \tag{6.30}$$

$$\beta = -\frac{(\eta - g(e_r - \frac{\alpha}{2}))}{2(\frac{1}{\lambda_s^2} + \frac{g^2}{4})} \tag{6.31}$$

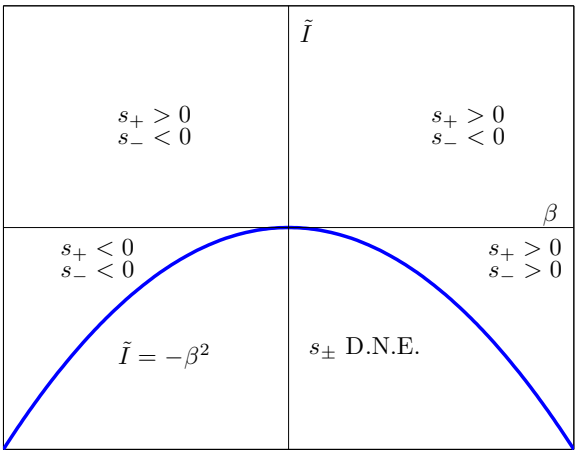
the steady states can again be written in the form (6.29). Note, as a check of consistency, that $I_{rh} = \frac{\alpha^2}{4}$, and $v^*(0) = \frac{\alpha}{2}$ for the Izhikevich model.

Based on the form (6.29) and the fact that $A_2(g) > 0$, it is straightforward to show the signs of $s_{\pm}$ are as shown in Figure 7.3(a). Since we require the equilibrium solutions to be positive, $e_{\pm}$ will have different regions of existence depending on the values of β and $\tilde{I}$. In particular, both equilibrium points exist when $I < I_{rh}$ and $g > \frac{\eta}{e_r - v^*(0)}$ in a wedge shaped region given by $\beta^2 + \tilde{I} > 0$. Only e_+ exists when $I > I_{rh}$. Neither solution exists in other parts of the parameter space. The regions of existence of $e_{\pm}$ and the non-firing solution are shown for the Izhikevich model in Figure 7.3(b).

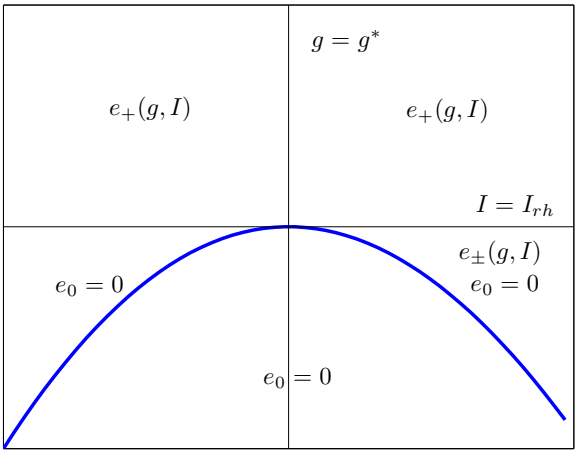
Away from the switching manifold, we can analyze the smooth bifurcations of the equilibria via linearization. The non-firing solution does not undergo any smooth bifurcations, as when it exists it lies in the region of phase space governed by the equations

$$s' = -\frac{s}{\tau_s}$$
$$w' = -\frac{w}{\tau_w}$$

Thus the non-firing solution is asymptotically stable when it exists and does not lie on the switching manifold, i.e., for $I < I_{rh}$. The nontrivial equilibria $e_{\pm}$ exist in the region of the



(a) $(\beta, \tilde{I})$ plane



(b) (g, I) plane

Figure 6.3: The existence of equilibria for the mean field system. (a) The sign of the s component of the nontrivial equilibria, in the $(\beta, \tilde{I})$ parameter plane. s_+ is positive in the first two quadrants and in a narrow wedge-shaped region in the fourth quadrant. s_- is also positive in this wedge-shaped region. (b) The existence of the trivial and nontrivial equilibria for the Izhikevich model in the (g, I) parameter space. The nontrivial equilibrium $e_+(g, I)$ only exists in the region $I > \alpha^2/4$, and for $I < \alpha^2/4$ in the wedge shaped region of the fourth quadrant indicated. The nontrivial equilibrium $e_-(g, I)$ only exists in this wedge shaped region. The trivial (non-firing) equilibrium e_0 only exists for $I \leq \alpha^2/4$.

phase space corresponding to $I - I^*(s,w) > 0$. In this case the Jacobian of the reduced mean-field system becomes

$$J(s) = \begin{pmatrix} -\frac{1}{\tau_s} + \frac{\lambda_s}{\tau_s}\left(\frac{g\lambda_s(e_r - v^*(s))F''(v^*(s))}{2s} + \frac{s}{2F''(v^*(s))\lambda_s}F'''(v^*(s))v^{*\prime}(s)\right) & -\frac{\lambda_s^2 F''(v^*(s))}{2s\tau_s} \\ \frac{\lambda_w}{\tau_w}\left(\frac{g\lambda_s(e_r - v^*(s))F''(v^*(s))}{2s} + \frac{s}{2F''(v^*(s))\lambda_s}F'''(v^*(s))v^{*\prime}(s)\right) & -\frac{1}{\tau_w} - \frac{\lambda_w\lambda_s F''(v^*(s))}{2s\tau_w} \end{pmatrix}$$

after one takes into account the steady state condition $sF''(v^*(s))/\lambda_s = \sqrt{I - I^*(s,\eta s)}$. The trace and determinant are given by the following:

$$\begin{aligned} \mathrm{Tr}(J) &= -\frac{F''(v^*(s))}{s}\left[\left(\frac{1}{\tau_s} + \frac{1}{\tau_w}\right)\frac{s}{F''(v^*(s))} - \frac{\lambda_s^2}{2\tau_s}\left(g(e_r - v^*(s)) + \frac{s^2 v^{*\prime}(s)F'''(v^*(s)))}{\lambda_s^2 F''(v^*(s))^2} - \eta\frac{\tau_s}{\tau_w}\right)\right] \\ \mathrm{Det}(J) &= \frac{F''(v^*(s))}{s\tau_s\tau_w}\left[\frac{s}{F'''(v^*(s))} + \frac{\lambda_w\lambda_s}{2} - \frac{\lambda_s^2}{2}(g(e_r - v^*(s)) - \frac{s^2 v^{*\prime}(s)F'''(v^*(s)))}{2F''(v^*(s))^2}\right] \end{aligned}$$

We can now discuss the stability of each equilibrium in its region of existence. To begin we use the expansions (6.24)-(6.25) in the determinant:

$$\begin{aligned} \mathrm{Det}(J) &= \frac{F''(v^*(s))}{s\tau_s\tau_w}\left[\frac{s}{F''(v^*(s))} + \frac{\lambda_w\lambda_s}{2} - \frac{\lambda_s^2}{2}(g(e_r - v^*(s)) - \frac{s^2 v^{*\prime}(s)F'''(v^*(s)))}{\lambda_s^2 F''(v^*(s))^2}\right] \\ &= \frac{F''(v^*(s))}{s\tau_s\tau_w}\left(\frac{1}{F''(v^*(0))} + \frac{g^2\lambda_s^2}{2F''(v^*(0))}\right)\left(s - M(g)(g - g^*) + O(s^2)\right) \\ M(g) &= \frac{e_r - v^*(0)}{2A_2(g)} \\ g^* &= \frac{\eta}{e_r - v^*(0)}. \end{aligned}$$

Substituting the equilibrium values of s and using the definition (6.27) of β, we obtain

$$\det(J)|_{\bar{s}_\pm} = \frac{F''(v^*(\bar{s}))A_2(g)\lambda_s^2}{\bar{s}\tau_s\tau_w}\left(\pm\sqrt{\beta^2 + \tilde{I}} + O(\bar{s}^2)\right). \tag{6.32}$$

Since the sign of $\frac{A_2(g)\lambda_s^2 F''(v^*(s))}{\tau_s\tau_w}$ is strictly positive, and the equilibria are only defined when $\bar{s} = \beta \pm \sqrt{\beta^2 + \tilde{I}} \geq 0$, we can immediately conclude that

$$\begin{aligned} \det(J)|_{s_+} &\geq 0 \\ \det(J)|_{s_-} &\leq 0 \end{aligned}$$

for small $s_\pm$. This implies that the equilibrium $\bar{e}_-$ is always an unstable saddle. The equilibrium $\bar{e}_+$, however, can be a node or a focus and its stability is determined by the trace. We will discuss this further in section 6.2.3. Note that these results are only valid when $\bar{s}_\pm$ is small for the QIF and AdEx models, but are globally valid for the Izhikevich model.

We can use the equations for the trace and determinant to formulate necessary conditions for the equilibria to display certain smooth bifurcations. In particular, $\det(J) = 0$ and $\mathrm{tr}(J) \neq 0$ are necessary conditions for an equilibrium to undergo a saddle-node bifurcation, while $\det(J) > 0$ and $\mathrm{tr}(J) = 0$ are necessary conditions for a Hopf bifurcation. Having both $\det(J) = 0$ and $\mathrm{tr}(J) = 0$, is a necessary condition for a Bogdanov-Takens bifurcation. Of course, to determine if these bifurcations actually occur requires checking additional genericity conditions. In the following section, we check these conditions where possible.

6.2.2 The Saddle Node Bifurcation Condition

As described above, necessary conditions for a saddle-node bifurcation are $\det(J) = 0$ and $\mathrm{tr}(J) \neq 0$. It is easy to see from (6.32) that the first condition is satisfied for both $e_\pm$ when $\beta^2 + \tilde{I} = 0$. It can be shown that the second condition is satisfied except at isolated points in the (g, I) parameter space as determined in section 6.2.3. In the following we will assume that we exclude these points.

To pursue this further, we study the existence of the equilibria. From the previous subsection, we know that $e_\pm$ both exist if $\beta^2 + \tilde{I} > 0$ and neither exists if $\beta^2 + \tilde{I} < 0$. When $\beta^2 + \tilde{I} = 0$, the two equilibria collapse into a single equilibrium, with $s = \beta$. We thus conclude that $\tilde{I} = -\beta^2$ corresponds to a two-parameter curve of saddle-node bifurcation. Rewriting this in terms of the original parameters yields the two-parameter bifurcation curve in terms of (g, I) :

$$I = I_{rh} - A_2(g)M(g)^2(g - g^*)^2 + O((g - g^*)^2) \stackrel{def}{=} I_{SN}(g). \tag{6.33}$$

Thus, for fixed g, $I_{SN}(g)$ is the value of the current that corresponds to a saddle-node bifurcation point.

There are a few things to note about I_{SN}. First, since $A_2(g)$ is a strictly positive function, $I_{SN}(g) \leq I_{rh}$ with $I_{SN}(g) = I_{rh}$ only if $g = g^*$. Also, this curve is only defined for $g \geq g^*$ for $g - g^*$ not too large. To see this, note that the saddle-node equilibrium, given by $s_{SN} = \beta = M(g)(g - g^*)$ only exists if $\beta > 0$. Since $M(g) > 0$, s_{SN} only exists if $g \geq g^*$. We shall see later that $g = g^*$ actually corresponds to a non-smooth co-dimension two bifurcation point. Finally, as $A_2(g)$ is a strictly positive function, one can show by using Sotomayor's theorem [143] that the saddle-node bifurcation is generic for the network of Izhikevich neurons provided that $g > g^*$. This is shown in the appendix to this chapter.

6.2.3 The Andronov-Hopf Bifurcation Condition

From the analysis of subsection 6.2.1, we know that only e_+ may undergo a Hopf bifurcation and that $\det(J)|_{s_+} > 0$ if $\beta^2 + \tilde{I} \neq 0$. We thus conclude that the determinant condition

for the Hopf bifurcation is given by $I \neq I_{SN}(g)$. To determine a necessary condition for the Hopf bifurcation, we begin by using the expansions (6.24)-(6.25)

$$\begin{aligned} Tr(J) &= -\left(\frac{1}{\tau_s}+\frac{1}{\tau_w}\right)+\frac{\lambda_s^2 F''(v^*(s))}{2s\tau_s}\left(g(e_r-v^*(s))+\frac{s^2 v^{*\prime}(s)F'''(v^*(s)))}{\lambda_s^2 F''(v^*(s))^2}-\eta\frac{\tau_s}{\tau_w}\right) \quad (6.34)\\ &= -\frac{F''(v^*(s))}{\tau_s s}\left(\frac{1}{F''(v^*(0))}+\frac{\gamma}{F''(v^*(0))}+\frac{g^2\lambda_s^2}{2F''(v^*(0))}\right)\left[s-N(g)(g-\bar{g})+O(s^2)\right] \\ N(g) &= \frac{\lambda_s^2(e_r-v^*(0))}{2}\left(\frac{1}{F''(v^*(0))}+\frac{\gamma}{F''(v^*(0))}+\frac{g^2\lambda_s^2}{2F''(v^*(0))}\right)^{-1} \quad (6.35)\\ \bar{g} &= \frac{\eta}{e_r-v^*(0)}\frac{\tau_s}{\tau_w} \quad (6.36)\end{aligned}$$

Note that the first term is strictly negative and $N(g)$ is a strictly positive function. Setting the trace to zero and using equations (6.26)-(6.29) which define $\tilde{I}, \beta$ and $\bar{s}_+$ yields

$$\begin{aligned} \bar{s}_+ &= N(g)(g-\bar{g})+O((g-\bar{g})^2) \\ &\Rightarrow M(g)^2(g-g^*)^2+\frac{I-I_{rh}}{A_2(g)} = (N(g)(g-\bar{g})-M(g)(g-g^*))^2 \end{aligned}$$

to lowest order in $\bar{s}$. Solving for I gives

$$I = I_{rh}+A_2(g)\left[N(g)^2(g-\bar{g})^2-2M(g)N(g)(g-\bar{g})(g-g^*)\right]+O((g-\bar{g})^2) \stackrel{def}{=} I_{AH}(g) \quad (6.37)$$

We thus conclude that if $I = I_{AH}(g)$ and $I \neq I_{SN}(g)$ then the equilibrium $\bar{s}_+$ has a pair of pure imaginary eigenvalues.

Recall that $N(g)$, $M(g)$, and $A_2(g)$, are positive functions. Further, it is easy to check that $N(g) < M(g)$. This leads to several observations. First, since the third equation in the sequence above can only be satisfied if $N(g)(g-\bar{g}) > M(g)(g-g^*)$, it follows that if $g^* \leq \bar{g}$ then no Hopf bifurcation occurs. Second, from the first equation in the sequence above we must have $g \geq \bar{g}$ in order for the equilibrium s_+ to exist at the Hopf bifurcation. When $g = \bar{g}$, $s_+ = 0$ and $I_{AH} = I_{rh}$. We shall see later that the point $I = I_{rh}$, $g = \bar{g}$ is a codimension-2 non-smooth bifurcation point. Finally, if $\bar{g} \leq g \leq g^*$, then $I_{AH} \geq I_{rh}$ with $I_{AH} = I_{rh}$ only if $g = \bar{g}$. If $g > g^*$, then it is possible for $I = I_{AH}(g)$ to intersect $I = I_{rh}$. We denote by $\hat{g}$ the value of g at the intersection point, if it exists.

We can now determine the stability of the equilibrium e_+ by studying the trace equation (6.34). Since the first term in this equation is strictly negative wherever it is defined (when $\bar{e}_+$ exists), the sign of the trace is determined by $\bar{s}_+ - N(g)(g-\bar{g})$. Since $\bar{s}_+$ and $N(g)$ are positive, it follows from the discussion above that when $g^* \leq \bar{g}$ the trace is negative, and hence $\bar{e}_+$ is asymptotically stable, wherever it exists. If $\bar{g} \leq g^*$ then the trace is negative (and $\bar{e}_+$ is asymptotically stable) if $g < \bar{g}$ or $g > \bar{g}$ and $I > I_{AH}(g)$. The trace is positive (and $\bar{e}_+$ is unstable) if $g > \bar{g}$ and $I < I_{AH}$. Note that if I is sufficiently close to I_{AH} then $\bar{e}_+$ will have a pair of complex conjugate eigenvalues.

In summary, for fixed g with $g > \bar{g}$ and $\bar{g} < g^*$, the equilibrium e_+ undergoes a Hopf bifurcation at $I = I_{AH}(g)$ if $I \neq I_{SN}(g)$. Further, we can now state completely the conditions for the saddle-node bifurcation: for fixed g with $g > g^*$, the equilibria e_+ and e_- undergo a saddle-node bifurcation when $I = I_{SN}(g)$ if $I \neq I_{AH}(g)$.

One can compute the first Lypaunov coefficient for the Izhikevich model in the limit that $\gamma \ll 1$ or $\gamma \gg 1$. This is done in the appendix, and is not included here due to the complexity. However, provided that $g > \bar{g}$ and $e_r - v^*(0) > 0$, the term is positive to the first two orders in both asymptotic limits of γ. Thus, under the long time scale conditions, $\tau_s, \tau_w \gg 1$, the bifurcation will be a subcritical bifurcation.

6.2.4 The Bogdanov-Takens Bifurcation Condition

Recall that necessary conditions for a Bogdanov-Takens bifurcation are $\det(J) = 0$ and $\mathrm{tr}(J) = 0$. Thus, from the analysis of the last two subsections, Bogdanov-Takens bifurcations (if they exist) will occur at intersection point(s) of the curves of saddle node and Hopf bifurcations in the g, I parameter space, i.e., at values of g such that $I_{AH}(g) = I_{SN}(g)$, with $g > \max(g^*, \bar{g})$. Using the expressions for these curves gives

$$\begin{aligned} I_{rh} - A_2(g)M(g)^2(g-g^*)^2 &= I_{rh} + A_2(g)\left[N(g)^2(g-\bar{g})^2 - 2M(g)N(g)(g-\bar{g})](g-g^*)\right] \\ 0 &= A_2(g)\left(N(g)(g-\bar{g}) - M(g)(g-g^*)\right)^2 \\ &\Updownarrow \\ 0 &= N(g)(g-\bar{g}) - M(g)(g-g^*) \end{aligned}$$

This latter equation may be simplified to a quadratic in g:

$$g^2 g^*(\gamma - 1)\frac{\lambda_s^2}{2} + g\gamma - g^*$$

Immediately one can see that a potential solution is $\gamma = 1$, $g = g^* = \bar{g}$. However, this is a single root which lies on the switching manifold. For $\gamma \neq 1$, we have the following two pairs of Bogdanov-Takens points:

$$g_{BT\pm} = \frac{2g^*}{\gamma \mp \sqrt{\gamma^2 - 2g^*\lambda_s^2(1-\gamma)}} \tag{6.38}$$

which yields the following leading order asymptotics provided that $1 \gg 1 - \gamma \geq 0$:

$$g_{BT+} = \frac{2}{\lambda_s^2(1-\gamma)} + O(1) \tag{6.39}$$

$$g_{BT-} = g^* + O(1-\gamma). \tag{6.40}$$

Note that g_{BT+} diverges as $1 - \gamma \to 0^+$ while $g_{BT-} \to g^*$, which given the previous work implies a collision of the Bogdanov-Takens equilibrium point with the switching manifold.

However, we note that these are only valid provided that $\gamma^2 - 2g^*\lambda_s^2(1-\gamma) > 0$. For the parameter sets we have looked at, we have not found a Bogdanov-Takens point as this quantity is negative. Thus, we leave the analysis of the Bogdanov-Takens point and its interaction with the switching manifold for future work.

Figure 6.4(a) shows the smooth bifurcations for the mean field system corresponding to a network of Izhikevich neurons with the parameter values from [49]. Note that the Hopf-bifurcation for both the full and reduced mean field systems corresponds closely to the onset of bursting in the actual network, as noted in [134]. For these parameter values, $\tau_w \gg \tau_s$ and no Bogdanov-Takens' points are observed. A series of corresponding phase portraits with nullclines and the switching manifold is shown in figure 6.5. Figure 6.4(b) shows the smooth bifurcations for the mean field systems corresponding to a network of AdEx neurons. In all figures the bifurcation curves derived from the small s expansions, i.e., equations (6.33) and (6.37), are compared with curves for the full mean field model generated numerically in MATCONT [47].

6.3 Non-Smooth Bifurcations

To study the non-smooth bifurcations for the mean field system (6.6)–(6.8), we will use the terminology and bifurcation classification for piecewise smooth continuous systems proposed in [48]. We note that some care must be used when applying these ideas to our system. Letting $x = [s, w]^T$ and recalling the definition (6.11) of the switching manifold, our system may be written in the general form used by [48]:

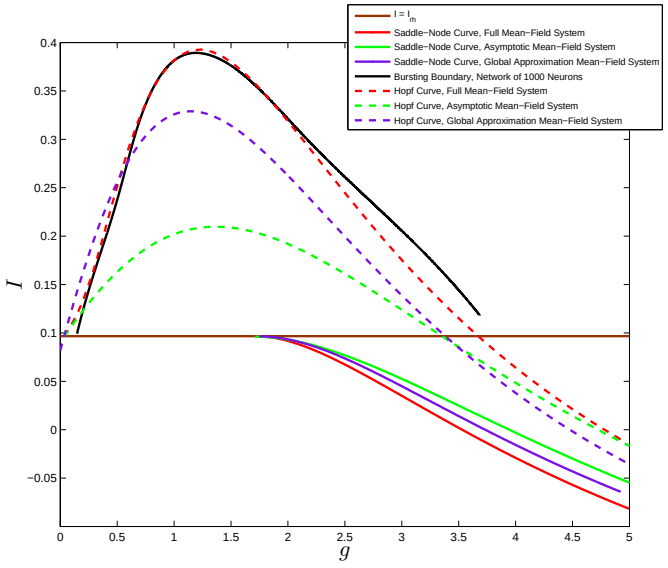
$$\dot{x} = \begin{cases} F_1(x, I), & \text{if } H(x, I) < 0 \\ F_2(x, I), & \text{if } H(x, I) > 0 \end{cases}$$

where

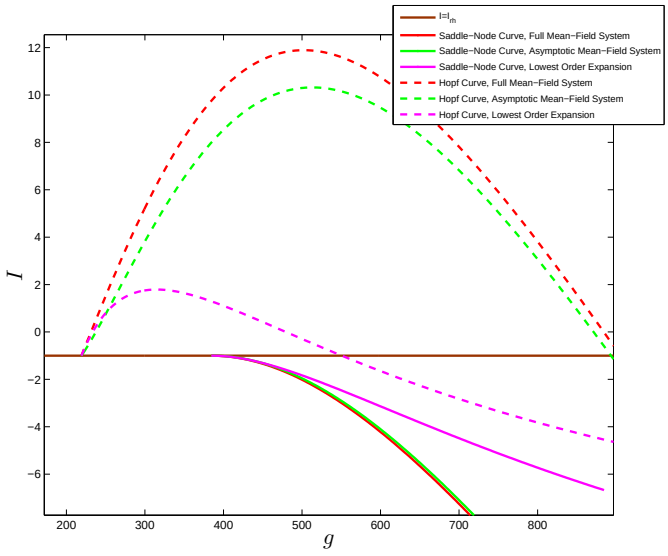
$$\begin{aligned} F_1(x, I) &= \begin{pmatrix} -\frac{s}{\tau_s} \\ -\frac{w}{\tau_w} \end{pmatrix} \\ F_2(x, I) &= F_1(x, I) + \sqrt{H(x, I)} \begin{pmatrix} \frac{\lambda_s}{\tau_s} \\ \frac{\lambda_w}{\tau_w} \end{pmatrix} \end{aligned}$$

However, F_2 is only defined for $H(x, I) > 0$. In contrast, the work of [48] assumes that both F_1 and F_2 are defined throughout the phase space. Nevertheless, we are able to classify a number of bifurcations in our system by analogy with the results in [48]. Additionally, due to the presence of the square-root, the Jacobian diverges in the vicinity of boundary equilibrium bifurcations. This means one cannot simply unfold the system, for example by reducing it to observer canonical form [164].

We will supplement our analysis with numerical studies of our example systems. In particular, we will perform a detailed study of the mean field system corresponding to a network of Izhikevich neurons with parameters given in Table 7.1.

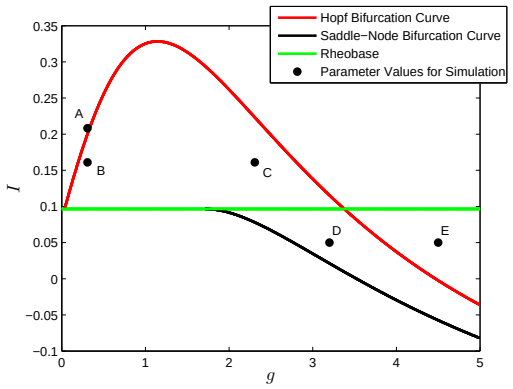


(a) Smooth bifurcations of the Izhikevich model

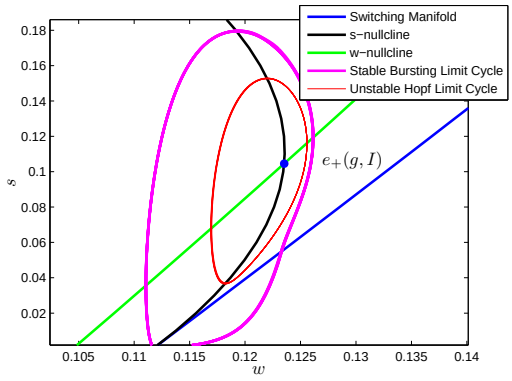


(b) Smooth bifurcations of the AdEx model

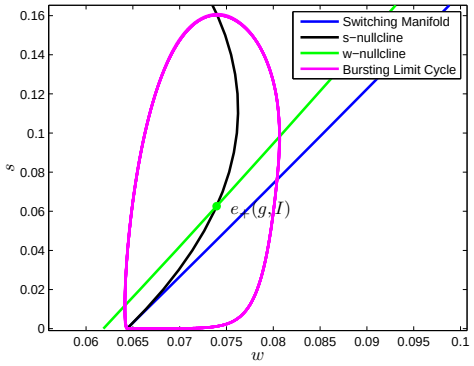
Figure 6.4: Comparison of various approximations of the two parameter bifurcation curves for the mean field equations of the Izhikevich model (a) and the AdEx Model (b). Shown are Hopf bifurcation curves (dashed lines), saddle-node bifurcation curves (dotted lines) computed for the full mean-field system (green), the reduced system (blue), and the lowest order approximation solution to the reduced system (purple). For the Izhikevich model, the bursting boundary is also determined for a network of 1000 neurons with the same parameters. The bifurcation curves for the saddle-node and Hopf bifurcations are computed using the MATLAB function *fsolve* on the determinant and trace equations of the Jacobian of the linearization. The red line corresponds to $I = I_{rh}$.



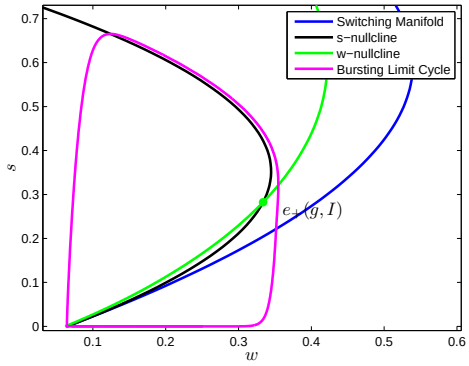
(a) Bifurcation Diagram



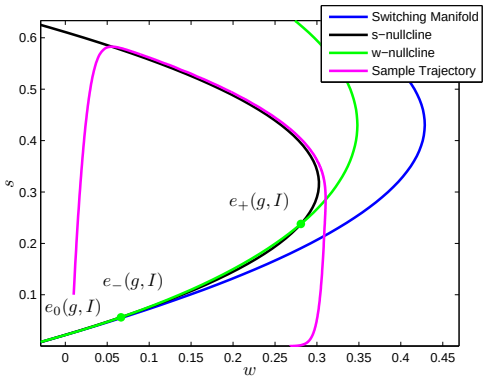
(b) Parameter set A



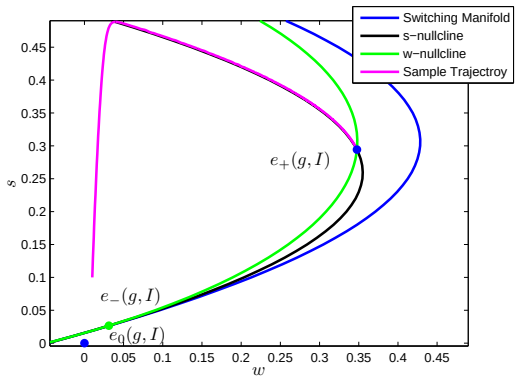
(c) Parameter set B



(d) Parameter set C



(e) Parameter set D



(f) Parameter set E

Figure 6.5: Shown above are a series of phase portraits with nullclines for the reduced mean-field equations for the Izhikevich network given by equations (6.17)-(6.19). Stable and unstable equilbiria are denoted by blue and green dots respectively and are labeled. The nullclines for the phase portraits in figures 6.5(b) to 6.5(e) are plotted in black (s-nullcline) and green (w-nullcline). The parameters for the phase portraits in subfigures 6.5(b)-6.5(e) are determined by the points A-E in figure 6.5(a), respectively.

6.3.1 Boundary Equilibrium Bifurcations ($I = I_{rh}$)

Recall that all the models we are considering have an equilibrium $e_0 = (0, 0)$ which exists (and is a stable node) if $I < I^*(0, 0) = I_{rh}$. When $I = I_{rh}$ this equilibrium lies on the switching manifold $I - I^*(s, w) = 0$. When $I > I_{rh}$, this equilibrium no longer exists as the origin is not an equilibrium of the part of the mean field system corresponding to $I - I^*(w, s) > 0$. In the terminology of non-smooth systems, the origin is a *virtual equilibrium* of the system for $I > I_{rh}$ and undergoes a *boundary equilibrium bifurcation* (BEB) when $I = I_{rh}$. The exact nature of this bifurcation depends on the value of g, in particular, its relationship to g^*, $\bar{g}$ and $\hat{g}$.

To determine the nature of the boundary equilibrium bifurcation, we begin by studying the nontrivial equilibria $e_\pm = (s_\pm, \eta s_\pm)$ when $I = I_{rh}$. Recalling the form (6.29) for $s_\pm$ and noting that $\tilde{I} = 0$ when $I = I_{rh}$, we find

$$s_+(\beta, 0) = \begin{cases} 0 & \beta < 0 \\ 2\beta & \beta \geq 0 \end{cases}$$

$$s_-(\beta, 0) = \begin{cases} 2\beta & \beta < 0 \\ 0 & \beta \geq 0 \end{cases}$$

Thus for $g < g^*$, e_+ collides with e_0 at $I = I_{rh}$, and for $g > g^*$, e_- collides with e_0.

Consider first the case $g^* < \bar{g}$ (which corresponds to $\tau_w < \tau_s$). In this case there is no Hopf bifurcation, so the results are straight forward. When $g < g^*$ e_+ is a sink which exists for $I > I_{rh}$. It collides with e_0 when $I = I_{rh}$ and ceases to exist when $I < I_{rh}$. Putting this together with the description of the existence and stability results for e_0, we conclude that, for this range of g values, the system undergoes a **persistence** BEB at $I = I_{rh}$. This will be either a focus/node or node/node persistence BEB depending on the classification of e_+. When $g > g^*$, recall that the equilibrium e_- is a saddle when it exists (for $I_{SN} < I < I_{rh}$). Since e_0 also exists for $I < I_{rh}$ and is a stable node, we conclude that for $g > g*$ there is a non-smooth saddle node BEB at $I = I_{rh}$.

Now consider the case $\bar{g} < g^*$. For $g < \bar{g}$, analysis similar to that above shows the system undergoes a **persistence BEB** at $I = I_{rh}$. Figure 6.6(a) shows this bifurcation for the mean field system corresponding to the Izhikevich network with parameters as in Table 7.1.

The situation for $\bar{g} < g < g^*$ is similar, except that e_+ is now an unstable focus for $I > I_{rh}$. Thus for this range of g values, there is a focus/node persistence BEB at $I = I_{rh}$. Since e_+ is a source and e_0 is a sink, we may expect (by analogy with the results in [48]) that a stable non-smooth limit cycle surrounding e_+ will be created as I increases through I_{rh}. Figure 6.6(b) confirms this for the mean field system corresponding to the Izhikevich network. Note that in this example, the amplitude of the limit cycle does not got to zero as I approaches I_{rh}. (See also Figure 6.8(c)) Further, the period of the limit cycle diverges

as $I \to I_{rh}^+$. See Figure 6.8(a). Thus the limit cycle appears to be created in homoclinic-like bifurcation as I increases through I_{rh}. We will thus refer to this as a **homoclinic persistence BEB**.

When $g > g^*$, analysis similar to that above shows that there is a non-smooth saddle node BEB at $I = I_{rh}$. Based on the analysis of the equilibrium points, there is no reason to expect anything more to occur with this bifurcation. However, our numerical examples show two cases. Figure 6.6(d) shows that a simple non-smooth saddle-node BEB occurs for the mean field system corresponding to the Izhikevich network with $g \gg \hat{g}$. Figure 6.6(c) shows the bifurcation for the same system with $g^* < g < \hat{g}$. In this case there is a non-smooth limit cycle for $I > I_{rh}$ that appears to be destroyed when $I = I_{rh}$. Thus this bifurcation appears to be a non-smooth version of the Saddle-node on an invariant circle (SNIC) bifurcation. We will refer to it as a **SNIC BEB**. The transition between the two types of BEBs that occur for $g > g^*$ will be discussed in a later section.

Based on our numerical results we hypothesize that a non-smooth limit cycle may be destroyed in a homoclinic-like bifurcation as I decreases through I_{rh}. We support this hypothesis in two ways.

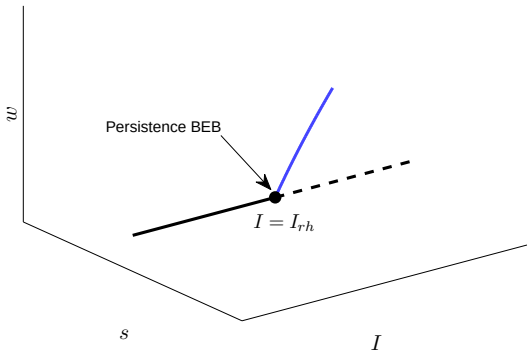
First, consider the vector field in the neighbourhood of the origin. Recall that the origin is always an attractor when it lies in the region where $H(s,w) < 0$. In the region where $H(s,w) > 0$, setting $I = I_{rh}$ and retaining only the highest order terms in s and w gives:

$$\begin{aligned}
s' &= -\frac{s}{\tau_s} + \frac{\lambda_s\sqrt{F''(v^*(s))}}{\tau_s}\sqrt{gs(e_r - v^*(0)) - v^{*\prime}(0)(gs)^2/2 - w} \\
&\approx \frac{\lambda_s\sqrt{F''(v^*(0))}}{\tau_s}\sqrt{gs(e_r - v^*(0)) - w} \\
w' &= -\frac{w}{\tau_w} + \frac{\lambda_w\sqrt{F''(v^*(s))}}{\tau_w}\sqrt{gs(e_r - v^*(0)) - v^{*\prime}(0)(gs)^2/2 - w} \\
&\approx \frac{\lambda_w\sqrt{F''(v^*(0))}}{\tau_w}\sqrt{gs(e_r - v^*(0)) - w}
\end{aligned}$$

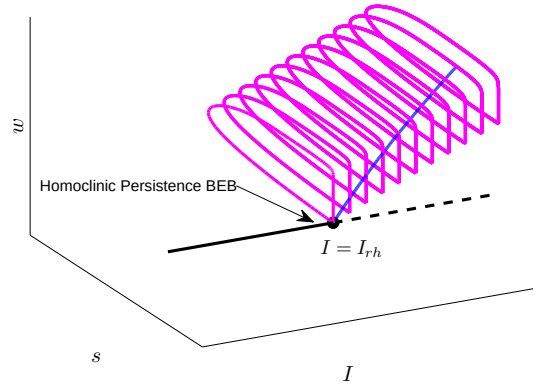
Thus, for $0 < s, w \ll 1$, and $I > I^*(s,w)$ the vector field points away from the origin and the boundary equilibrium $(0,0)$ is a repeller in this region. Since the boundary equilibrium point is as a repeller on one side of the switching manifold and an attractor on the other, it is possible for a non-smooth homoclinic orbit to this equilibrium point to exist when $I = I_{rh}$.

Second, we show that under certain parameter conditions, if a non-smooth limit cycle surrounds the equilibrium e_+, it must be destroyed when $I = I_{rh}$. To do this we show that trajectories that cross the switching manifold when $I = I_{rh}$ lie within the basin of attraction of the origin. Thus any non-smooth limit cycle must become homoclinic to the origin at $I = I_{rh}$. Note that if $I - I^*(s,w) < 0$, then

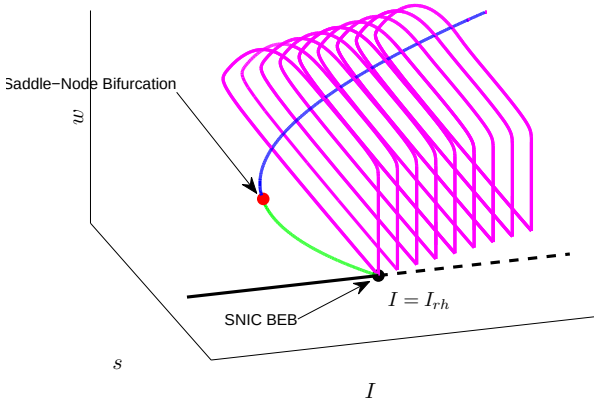
$$\frac{dw}{ds} = \frac{\tau_s}{\tau_w}\frac{w}{s} = \gamma\frac{w}{s}$$



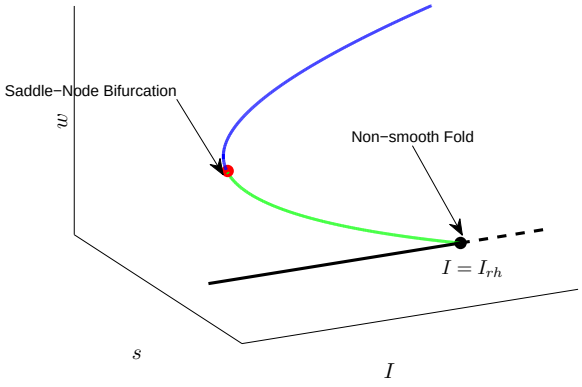
(a) $g < \bar{g}$



(b) $\bar{g} < g < g^*$



(c) $g^* < g < \hat{g}$



(d) $g \gg \hat{g}$

Figure 6.6: The four branches of boundary equilibrium bifurcations (BEB) that have been found in the mean field system for the Izhikevich network. In all figures, the equilibria are e_0 (black), e_+ (blue) and e_- (green), and solid lines indicate real equilibria, while dashed lines indicate virtual ones. The magenta lines are the non-smooth limit cycles determined via direct numerical integration. (a) the equilibrium e_+ collides with e_0 at $I = \frac{\alpha^2}{4}$. This results in the disappearance of e_+ for $I < \frac{\alpha^2}{4}$, while e_0 persists as a stable node. The situation is similar for (b), except that here the non-smooth limit cycle collides with the BEB equilibrium point in a kind of non-smooth homoclinic bifurcation. (d) the equilibrium e_- exists and is an unstable saddle for $I < \frac{\alpha^2}{4}$, as does the stable node e_0. These equilibria collide in a boundary equilibrium bifurcation at $I = \frac{\alpha^2}{4}$, and e_- is destroyed while e_0 becomes virtual. The bifurcation diagram in (c) is similar to that in (d) except for the emergence of a homoclinic limit cycle at the bifurcation point in a kind of non-smooth SNIC bifurcation.

and thus $w = Cs^\gamma$ for some constant C. Assuming that the trajectory starts with (s_0, w_0) on the switching manifold then $w = w_0 \left(\frac{s}{s_0}\right)^\gamma$ where $w_0 = gs_0(e_r - v^*(0)) - v^{*\prime}(0)\frac{(gs)^2}{2}$. Now suppose this trajectory crosses the switching manifold again at (s, w). Then

$$w_0 \left(\frac{s}{s_0}\right)^\gamma = gs(e_r - v^*(0)) - \frac{1}{F''(v^*(0))}\frac{(gs)^2}{2}$$

Clearly two solutions of this equation are (s_0, w_0) and $(0, 0)$. Dividing through by s and simplifying one obtains

$$(1 - ks_0)\frac{s^{\gamma-1}}{s_0^{\gamma-1}} = 1 - ks \tag{6.41}$$

where $k = \frac{g}{F''(v^*(0))2(e_r - v^*(0))}$.

If $\gamma > 1$ the left hand side of (6.41) is monotonically increasing while the right hand side is a line with negative slope. Hence (s_0, w_0) is the unique intersection point. This means every trajectory that enters the region $I - I^*(s, w) < 0$ when $I = I_{rh}$ is attracted to the origin. If $\gamma < 1$, the left side of (6.41) is now monotonically decreasing. Unless the line is tangent the curve at $(s_0, 1)$ there will always be another intersection point. Rearranging the equation shows that this intersection point will occur for $s < s_0$ if gs_0 is sufficiently small. For fixed g, this means that any trajectory that starts on the switching manifold at (w_0, s_0) with s_0 sufficiently small will be attracted to the origin. Thus all non-smooth limit cycles that are close enough to the origin for $I > I_{rh}$ will become homoclinic to the origin when $I = I_{rh}$.

Given how $g = \bar{g}$, and $g = g^*$ delimit the different types of BEB bifurcations, it should be clear that these special points represent higher codimension bifurcations along the $I = I_{rh}$ line. We shall explore these bifurcations further below, in addition to determining the geometrical meaning of these points.

6.3.2 Saddle-Node Boundary Equilibrium Bifurcation ($I = I_{rh}, g = g^*$)

From the results of the previous section, we can conclude that the point $I = I_{rh}, g = g^*$ is a special codimension-two bifurcation point where the boundary equilibrium bifurcation (BEB) changes from a persistence BEB to a non-smooth saddle-node. Note that the smooth branch of saddle-node bifurcations found earlier actually emanates out from the codimension-2 point (g^*, I_{rh}). We will show here that it does so in a highly non-generic way as the saddle-node equilibrium hits switching manifold tangentially at the BEB, and is the only equilibrium point that interacts with the switching manifold in this way.

We have seen that regardless of the parameter values, all the nontrivial equilibria lie on the curve $w = \eta s$. Thus as any parameter is varied the nontrivial equilibrium will follow this curve, which has slope

$$w'(s) = \eta. \tag{6.42}$$

Further, the only equilibrium that can be a boundary equilibrium point is $e_0 = (0, 0)$, the non-firing solution. Now the switching manifold can be written as

$$w(s) = I + F(v^*(s)) + gs(e_r - v^*(s)).$$

Thus, the slope of the switching manifold at the BEB is

$$w'(0) = g(e_r - v^*(0)). \tag{6.43}$$

Equating (6.42) and (6.43) shows that the nontrivial equilibrium undergoing the BEB will hit the switching manifold tangentially only if $g = g^* = \frac{\eta}{e_r - v^*(0)}$. From this it is straightforward to show that with $g = g^*$ fixed, the nontrivial equilibrium e_+ hits the switching manifold tangentially as $I \to I_{rh}$ and $s_+ \to 0$. More interesting is to consider what happens when g is varied. From our previous analysis we know that at the saddle-node bifurcation point, the saddle-node equilibrium, $e_{SN} = (s_S N(g), \eta s_S N(g))$ is defined by

$$s_{SN}(g) = M(g)(g - g^*) + O((g - g^*)^2)$$

Thus, as $g \to g^*$, $s_{SN}(g) \to 0$. This implies that the saddle-node equilibrium hits the switching manifold tangentially at $g = g^*$, $I = I_{rh}$.

In summary the point $g = g^*, I = I_{rh}$ is the collision between three branches of co-dimension-1 bifurcations: a pair of non-smooth boundary equilibrium bifurcations and a smooth branch of saddle-node bifurcations. The details of the BEB involved depend on the relationship between g^* and $\bar{g}$. If $g^* < \bar{g}$, the BEBs are simple: a simple node/focus or focus/focus persistence BEB occurs for $g < g^*$ and a non-smooth saddle-node BEB occurs for $g > g^*$. The case $\bar{g} < g^*$ is more complex due to the possible presence of limit cycles associated with the Hopf bifurcation. In the case we studied numerically and described in section 6.3.1, for $g < g^*$ we observe a homoclinic persistence BEB and for $g > g^*$ we observe a SNIC BEB.

While this bifurcation may be complicated, the bifurcation point can be determined analytically for all the models. It is shown in detail in figure 7.9(b). Of particular interest is the fact that associated with this point is a region in the $g > g^*$, $I < I_{rh}$ quadrant of the parameter space with bistability between firing and non-firing solutions. For parameter values in the region a brief stimulus could cause the network transition from quiescence to tonic firing.

6.3.3 Limit Cycle Grazing

The Andronov-Hopf bifurcation described in section 6.2.3 leads to the creation of a limit cycle. As I moves away from the bifurcation point, the amplitude of the limit cycle may increase enough that it hits the switching manifold tangentially, resulting in a grazing bifurcation. It is difficult to say much in general about the nature of this bifurcation,

however, analysis similar to that in the last section shows that if $I < I_{rh}$ then once a trajectory enters the region $I - I^*(s, w) < 0$, it cannot leave, but will be attracted to the origin. Thus we expect that if a grazing bifurcation occurs for $I < I_{rh}$ it will lead to the destruction of the limit cycle.

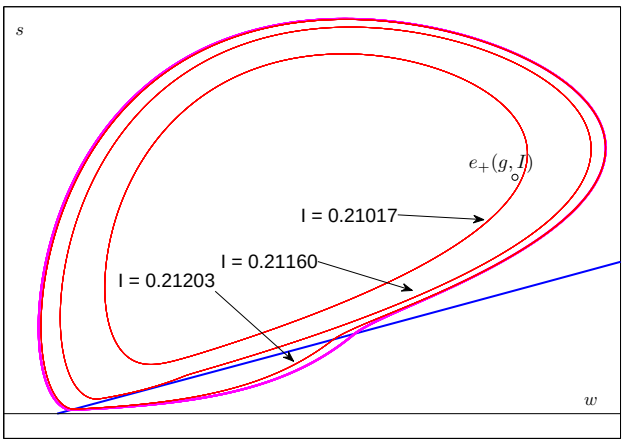
To gain more insight, we performed a numerical study of the mean field system corresponding to the Izhikevich network with parameter values as in Table 7.1. We first confirmed that the Hopf bifurcation is subcritical, using MATCONT and by numerically simulating the time reversed system. Additionally, the analytically determined first Lyapunov coefficient is positive to the first two orders in the ratio of the time constants for the Izhikevich model. We then showed that the unstable limit cycle generated by the Hopf can undergo two different types of grazing bifurcations, depending on the value I. For $I > I_{rh}$, the grazing bifurcation that occurs is a persistence type grazing, i.e., the unstable limit cycle generated via the subcritical Hopf bifurcation just becomes non-smooth after the grazing bifurcation. This is shown in Figure 6.7(a). Here, the limit cycle undergoes a grazing bifurcation at $I = 0.2680$, and it persists past it. Its amplitude rapidly increases past the grazing bifurcation, and it almost immediately undergoes a non-smooth saddle-node of limit cycles with a stable non-smooth limit cycle. For $I < I_{rh}$, the grazing bifurcation is a destruction type grazing as the limit cycle ceases to exist after the grazing for the reason discussed above. This is shown in Figure 6.7(b).

If the Hopf were supercritical we would expect to see the same two types of grazing bifurcations. The only difference would be that the grazing bifurcation would occur for $I < I_{AH}$ and we would not expect the saddle-node of limit cycles bifurcation to occur.

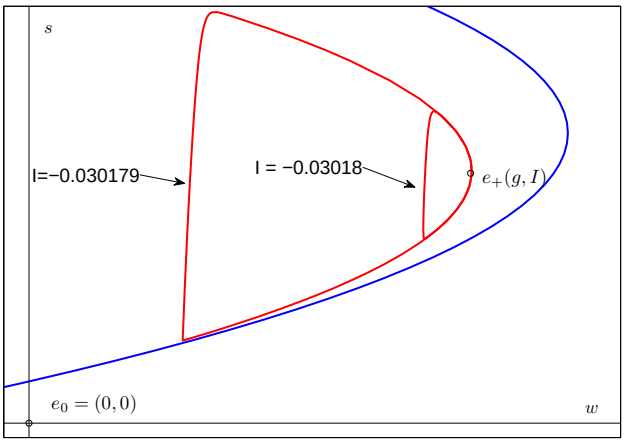
6.3.4 Hopf Boundary Equilibrium Bifurcation ($I = I_{rh}, g = \bar{g}$)

The analysis of section 6.3.1 showed that when $\bar{g} < g^*$ the point $I = I_{rh}, g = \bar{g}$ is a codimension-two bifurcation point where the boundary equilibrium changes from a simple focus/node persistence BEB to a homoclinic persistence BEB. Recall that the two parameter Hopf bifurcation curve is given by $I = I_{AH}(g)$ as defined in section 6.2.3. From the analysis in that section, the equilibrium point on the Hopf curve is $e_{AH} = (s_{AH}, \eta s_{AH})$ where $s_{AH}(g) = N(g)(g - \bar{g}) + O((g - \bar{g})^2)$. Setting $I = I_{AH}(g)$ we see that as $g \to \bar{g}$, $I_{AH} \to I_{rh}$ and $e_{AH} \to e_0$, that is the Hopf equilibrium point undergoes a BEB at $I = I_{rh}, g = \bar{g}$. We thus refer to this point as a Hopf boundary equilibrium bifurcation (Hopf BEB).

An alternative way to characterize the Hopf BEB is to fix $I = I_{rh}$ and let $g \to \bar{g}^+$. On



(a) Persistence grazing bifurcation and non-smooth saddle-node of limit cycles



(b) Destruction Grazing Bifurcation

Figure 6.7: Limit cycle grazing bifurcations for the Izhikevich system. (a) As I is increased above $I_{AH}(g)$, for fixed g, the unstable limit cycle (shown in red) generated by the subcritical Hopf bifurcation increases in amplitude. For large enough I, the limit cycle grazes the switching manifold (shown in blue). After the grazing, the limit cycle becomes non-smooth and subsequently collides with the non-smooth stable limit cycle (shown in pink). The two limit cycles annihilate each other in a non-smooth saddle node of limit cycles. Note that as I is varied, the switching manifold, the point e_+, and the unstable limit cycle all vary. However, aside from the unstable limit cycle, these other sets do not vary significantly. Thus, for clarity, we have only shown the switching manifold and stable non-smooth limit cycle for $I = 0.2690$, and e_+ for $I = 0.2604$. (b) For $I < I_{rh}$ the grazing bifurcation destroys the limit cycle.

$I = I_{rh}$, the mean field system for the Izhikevich network may be approximated as follows:

$$\begin{aligned} s' &= -\frac{s}{\tau_s} + \frac{\lambda_s\sqrt{F''(v^*(s))}}{\tau_s}\sqrt{gs\left(e_r - \frac{\alpha}{2}\right) - \frac{(gs)^2}{4} - w} \\ &\approx \frac{\lambda_s\sqrt{F''(v^*(0))}}{\tau_s}\sqrt{gs\left(e_r - \frac{\alpha}{2}\right) - w} \\ w' &= -\frac{w}{\tau_w} + \frac{\lambda_w\sqrt{F''(v^*(s))}}{\tau_w}\sqrt{gs\left(e_r - \frac{\alpha}{2}\right) - \frac{(gs)^2}{4} - w} \\ &\approx \frac{\lambda_w\sqrt{F''(v^*(0))}}{\tau_w}\sqrt{gs\left(e_r - \frac{\alpha}{2}\right) - w} \end{aligned}$$

for (s, w) in the vicinity of the origin. Thus, we have

$$\begin{aligned} \frac{dw}{ds} &= \eta\gamma + H.O.T. \\ \Rightarrow w &= \eta\gamma s + H.O.T. \end{aligned}$$

for the trajectory of the homoclinic limit cycle. Additionally, linearizing the switching manifold about the origin yields;

$$w = gs\left(e_r - \frac{\alpha}{2}\right)$$

Now, using these two equations we can solve for grazing bifurcations of the homoclinic limit cycle with the switching manifold at the origin. Solving the grazing condition $w'(0) = \eta\gamma$ yields

$$g = \frac{\eta\gamma}{(e_r - \alpha/2)} = \bar{g}$$

Thus, the Hopf BEB bifurcation can be seen as a grazing bifurcation which destroys the non-smooth homoclinic limit cycle.

Our analysis so far shows three branches of bifurcation emanating from this co-dimension-two point: two non-smooth BEB branches and a branch of Hopf bifurcation. As shown in Figure 6.6, for $g < \bar{g}$ there is a simple persistence BEB, while for $\bar{g} < g < g^*$ there is a homoclinic persistence BEB. We have numerically studied the bifurcations that occur in a neighbourhood of this point for the Izhikevich model and find that that two more branches of bifurcation appear to emanate from this point as we describe below.

Let g be fixed with $g > \bar{g}$ and consider the sequence of bifurcations involving limit cycles. At $I = I_{rh}$ a stable non-smooth limit cycle is created in a homoclinic persistence BEB, at $I = I_{AH} > I_{rh}$ an unstable smooth limit cycle is created in a subcritical Hopf bifurcation. As I increases the smooth limit cycle becomes non-smooth in a grazing bifurcation and

then is destroyed along with the stable non-smooth limit cycle in a saddle-node of limit cycles. We wish to determine how the grazing and saddle-node of limit cycles bifurcations behave near $g = \bar{g}$.

To do this we followed the stable non-smooth limit cycle along the Hopf bifurcation curve. Specifically, we numerically computed the amplitude and period of the limit cycle along the curve $(g, I_{AH}(g))$ in the (g, I) parameter space with $g \to \bar{g}$. The results are shown in Figure 6.9, specifically figure 6.9(a). The stable non-smooth limit cycle is computed using direct simulations of the ODE system, where the system is initialized exterior to the limit cycle in the phase plane which ensures convergence. From this figure, we can see that the amplitude of the stable non-smooth limit cycles goes to 0 as $g \to \bar{g}$. This implies that this limit cycle collapses to the origin $(0, 0)$. But as this bursting limit cycle is one part of the saddle-node of limit cycles bifurcation, then this bifurcation must also emerge from Hopf BEB. Since the grazing bifurcation lies between the saddle-node of limit cycles and the Hopf bifurcation, the persistence grazing bifurcation must also emerge from the point $g = \bar{g}, I = I_{rh}$. The entire sequence of bifurcations near the Hopf BEB is shown in figures 7.9(a) and 7.9(d).

6.3.5 A Co-dimension 3 Non-smooth Bifurcation

We briefly note that if $\tau_w = \tau_s$, then we have

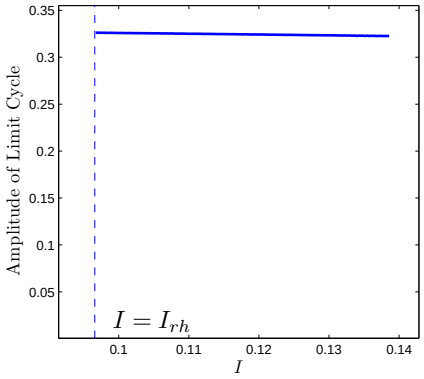
$$\bar{g} = g^*, \tag{6.44}$$

which means that the Hopf and saddle-node BEB points coincide in a non-smooth codimension-3 bifurcation point. This bifurcation point may be thought of as a Bogdanov-Takens equilibrium point lying on a switching manifold. However, we note that there is no Bogdanov-Takens bifurcation (or for that matter saddle-node or Hopf bifurcations) at this point in the classical sense, as the Jacobian of the system diverges, and hence the conditions associated with these different smooth bifurcations cannot be satisfied.

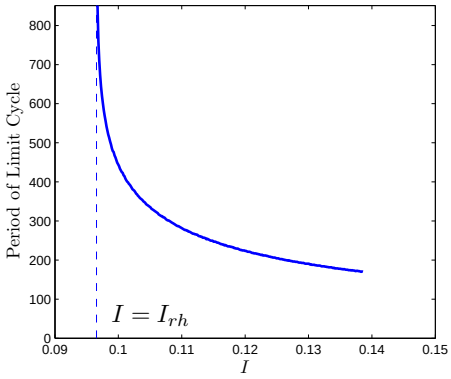
This point appears to act as an organizing center for the bifurcation diagram, with all the non-smooth bifurcations emanating from it. Due to the complexity of this point, we will leave its analysis for future work. However, it does illustrate how rich the non-smooth bifurcation sequence of this relatively simple PWSC system is.

6.3.6 A Global Co-dimension 2 Non-smooth Bifurcation

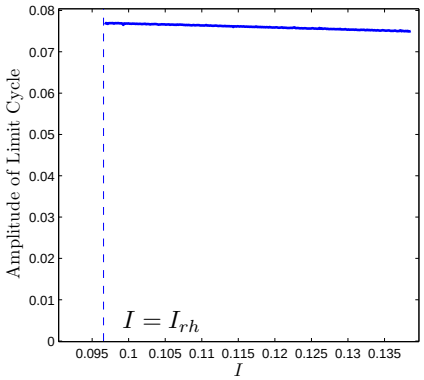
In addition to the two local non-smooth bifurcations that occur at $g = \bar{g}$, and $g = g^*$, there appears to be a global codimension-2 bifurcation that occurs for these mean field systems. Recall that there are two different types of grazing bifurcations, a destruction type (which occurs for $I < I_{rh}$) and a persistence type (which occurs for $I > I_{rh}$). These are shown in figure 6.7. Thus there is a co-dimension two point when the grazing bifurcation crosses



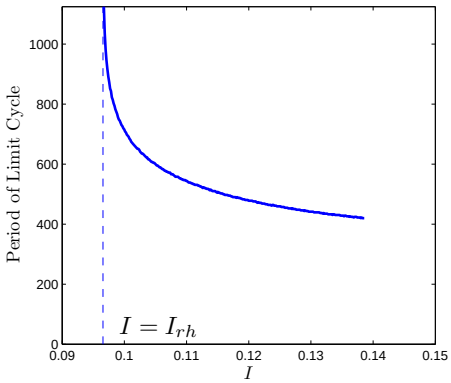
(a) Amplitude of the stable non-smooth limit cycle



(b) Period of the stable non-smooth limit cycle

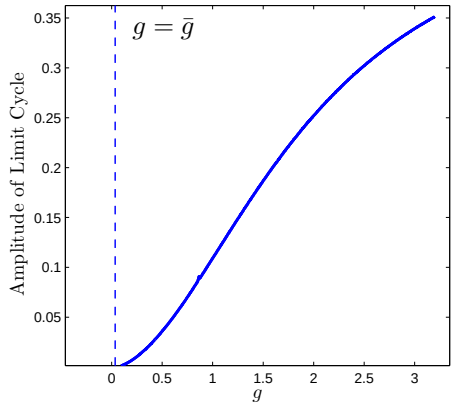


(c) Amplitude of the stable non-smooth limit cycle

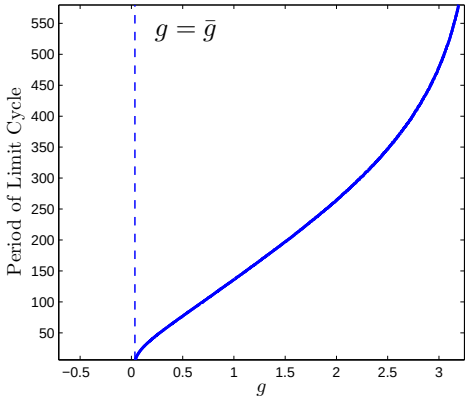


(d) Period of the stable non-smooth limit cycle

Figure 6.8: The amplitude ((a) and (c)) and period ((b) and (d)) of the bursting limit cycle in the Izhikevich system for fixed g with $\bar{g} < g < g^*$ (left column) and $g > g^*$ (right column), respectively, as $I \to I_{rh}$. These two quantities are resolved via direct numerical simulation of the limit cycle. Note the period diverges as $I \to I_{rh}$, while the amplitude is non-zero, indicative of a homoclinic limit cycle. The amplitude is computed as the difference between the maximum and minimum w component in the steady state limit cycle.



(a) Amplitude of the stable non-smooth limit cycle



(b) Period of the stable non-smooth limit cycle

Figure 6.9: The amplitude (a) and period (b) of the bursting limit cycle followed along the two-parameter Hopf bifurcation curve in the Izhikevich system. The Hopf bifurcation curve is entirely parameterized by g, in the (I, g) plane, and thus as we decrease g, we can compute the amplitude and period of the bursting limit cycle via direct numerical simulations. As can be seen, the amplitude decreases towards 0 as $g \to \frac{\hat{w}}{\hat{s}(e_r - \alpha/2)} = \bar{g}$, as does the period. As the bursting limit cycle is the exterior limit cycle in a non-smooth saddle node bifurcation of limit cycles, this bifurcation must also emanate from $\bar{g}$. Additionally, as the saddle-node of limit cycles occurs subsequent to a persistent grazing bifurcation of the unstable Hopf limit cycle, the grazing bifurcation must also emerge from this point. Also note that this is the only point in the parameter space where the homoclinic limit cycle generated does not have a divergent period as $I \to I_{rh}$. This is due to the fact that the homoclinic limit cycle has collapsed down to a point exactly at $g = \bar{g}$, and thus does not exist at this parameter value.

$I = I_{rh}$. As for the other co-dimension two points, one may expect there would be a change in the BEB bifurcations at this point. In the case we have studied numerically it appears that the BEB changes from SNIC type before this transition to a regular non-smooth fold after. This is shown in figure 7.9(c) and figure 7.9(d). Note that this transition occurs for $g > \hat{g}$, i.e., after the second intersection of the Hopf curve with $I = I_{rh}$. It also appears that the saddle-node of non-smooth limit cycles bifurcation emanates from this point. Note that this does not imply that there is a second impact with the Hopf equilibrium and the switching manifold, as $s_{AH}(g) = N(g)(g-\bar{g}) > 0$. This bifurcation results in the destruction of the homoclinic limit cycle that exists on $I = I_{rh}$, and it is very difficult to analyze, as it is a non-local co-dimension 2 non-smooth bifurcation. Geometrically, however it occurs when the unstable smooth limit cycle (generated via the Hopf bifurcation) grazes the switching manifold at $I = I_{rh}$. If the Hopf bifurcation were supercritical instead of subcritical we would expect a similar codimension two point to occur (if a grazing bifurcation occurred). However, it would occur for $g < \hat{g}$.

Again, due to the complexity of this particular bifurcation, further analysis is beyond the scope of this thesis, and we leave it for future work.

6.4 Non-Smooth Bifurcations Demonstrated in the Network Simulations

While the preceding analysis revealed a great deal of novelty and non-smooth bifurcations for the reduced mean field system, in order for the non-smooth analysis to be useful, it has to reflect the phenomena displayed by the actual network. Here, we will demonstrate many of the non-smooth bifurcations predicted in the analysis are present in a full network of neurons. We will primarily consider a network of Izhikevich neurons.

A difficulty is that one cannot easily expose unstable equilibria and limit cycles in the large network of neurons using numerical simulations. For example, the equilibrium point e_- is a saddle in the mean field, and short of somehow initializing the network of neurons on the stable manifold of the saddle, it cannot be resolved via direct simulations. However, the unstable node e_+ can be resolved by modifying the network as follows. Using the separation of time scales between the fast variable s, and the slow variable w, we replace the full network (6.1)–(6.4) by the following:

$$\begin{aligned}
\dot{v}_i &= v_i(v_i - \alpha) - w_i + gs(er - v_i) \\
\dot{w}_i &= a(bv_i - w) \\
s &= \frac{\bar{w}}{\eta} = \frac{1}{\eta}\left(\frac{1}{N}\sum_{i=1}^{N} w_i\right) \\
v_i(t^-_{spike}) &= v_{peak} \rightarrow \begin{array}{rcl} v_i(t^+_{spike}) &=& v_{reset} \\ w_i(t^+_{spike}) &=& w_i(t^-_{spike}) + \hat{w}, \end{array}
\end{aligned}$$

for $i = 1, 2, \ldots N$. Here the dynamics of s are replaced entirely by its steady state, large network solution: $\lambda_s \langle R \rangle \approx w \frac{\lambda_s}{\lambda_w} = w/\eta$, with w replaced by the finite mean $\bar{w}$. We will refer to this network of neurons as the slow network.

The mean field system corresponding to the slow network is simply a one-dimensional non-smooth ODE, given by:

$$\begin{aligned} \dot{w} &= -\frac{w}{\tau_w} + \frac{\lambda_w}{\tau_w}\langle R \rangle \\ \langle R \rangle &= \begin{cases} \sqrt{I - I^*(w/\eta, w)} & I \geq I^*(w/\eta, w) \\ 0 & I < I^*(w/\eta, w) \end{cases} \end{aligned}$$

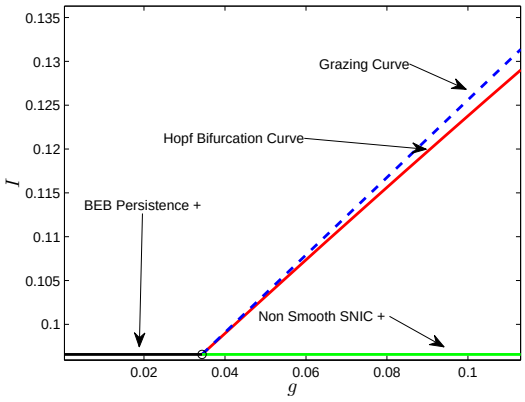
The mean field system for the slow network has the same steady states as the mean field system for the full network: the two firing solutions, $w_\pm$, and the non-firing solution, $w_0 = 0$, with regions of existence as for the full network (see Figure 7.3). However, being a one dimensional system, no Hopf bifurcations (and thus oscillations) are present in the mean field system for the slow network. Clearly the non-firing solution w_0 is always stable where it exists. The stability of the firing solutions is determined by the eigenvalue

$$\lambda(w_\pm) = -\frac{\lambda_w^2}{\tau_w} A_2(g) \left(1 - \frac{M(g)(g - g^*)}{M(g)(g - g^*) \pm \sqrt{M(g)^2 (g - g^*)^2 + \tilde{I}}} \right), \tag{6.45}$$

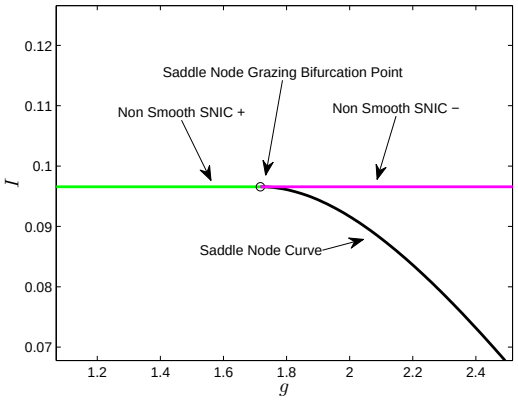
where the functions $A_2(g)$, and $M(g)$ are identical to those for the mean field system for the full network. It follows that w_+ is always stable and w_- is always unstable where they exist, and that they undergo saddle-node bifurcation at $I = I_{SN}$, as for the full network. In summary, we should expect bistability between w_+ and w_0 for $I_{SN} < I < I_{rh}$, but at most one stable state elsewhere in the parameter space. Indeed, if we simulate the slow network with a slowly varying current that either decreases from current values greater than I_{rh} or increases from current values less than I_{rh}, we get bistability for $g > g^*$ and none for $g < g^*$. This is shown in Figure 6.11.

Using the simulations of the slow network and the full network, we can piece together a pseudo-bifurcation diagram for the full network. This is shown in figure 6.12. The boundary equilibrium bifurcations that occur near the vicinity of g^* are also observed in the actual network. Given the similarities between the bifurcation diagram for the actual network, and that predicted by the non-smooth mean field equations, it would appear that in order to understand the bifurcations that occur in these networks, one has to consider non-smooth bifurcation theory.

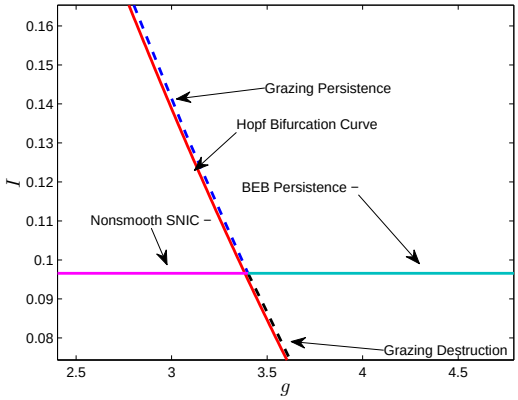
One might ask whether the non-smooth nature of the mean field system we studied is a direct result of the non-smooth nature of the integrate and fire neural models. This is not the case. The firing rate of any type I neuron in the vicinity of the saddle-node on an invariant circle bifurcation is proportional to $\sqrt{I - I_{rh}}$. Assuming that the dynamics of the neuron voltage is much faster then the dynamics of the all the other intrinsic and synaptic currents, the mean field model for such a type I neuron should also have a square-root



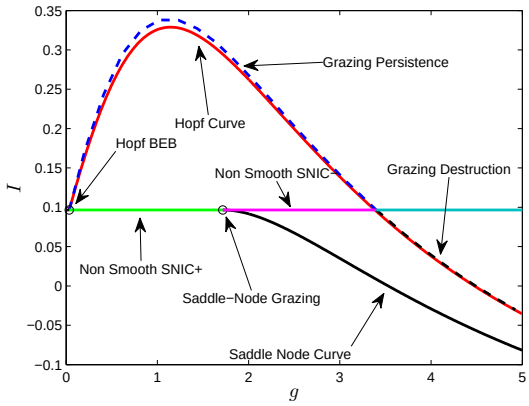
(a) Hopf BEB Point (Bottom Left Corner))



(b) Saddle-Node BEB (Center)

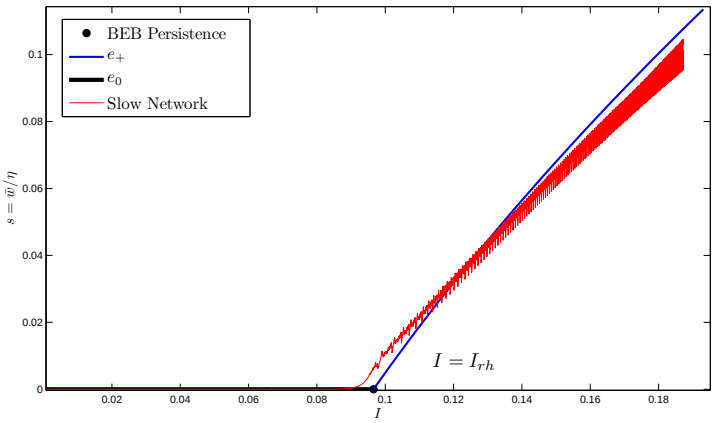


(c) Grazing Alternation (Bottom Right Corner)

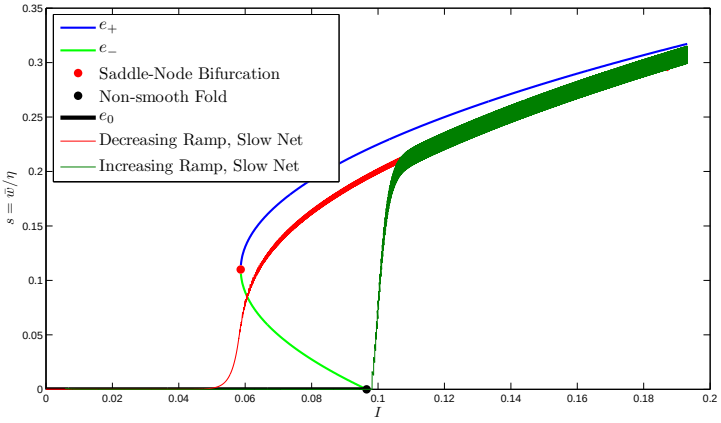


(d) Total Bifurcation Diagram

Figure 6.10: The entire bifurcation sequence for the Izhikevich model, including all known non-smooth and smooth bifurcation points. (d) is the entire diagram in the two-parameter space. (a), (b), and (c) are the bottom left, center, and bottom right regions, respectively. (a) The co-dimension two bifurcation point involving the collision of a branch of Hopf bifurcations with the switching manifold. This co-dimension two point also involves a collision with a branch of grazing bifurcations of the unstable limit cycle generated by the sub-critical Hopf, in addition to a branch of saddle-node of limit cycles (not shown for clarity). A non-smooth SNIC bifurcation, and BEB persistence bifurcation also collide simultaneously at the codimension two point $(\bar{g}, I_{rh})$. (b) The codimension two saddle-node grazing point, which occurs when a saddle-node bifurcation grazes a switching manifold. The saddle-node branch of bifurcations collides at the codimension-two point (g^*, I_{rh}) along with two branches of non-smooth SNIC bifurcations. (c) A global codimension-two point. This bifurcation point involves the switching of a grazing bifurcation in the unstable Hopf limit cycle from a persistence case, to a destruction case. The non-smooth SNIC bifurcation also collides with a branch of BEB persistence bifurcations for the equilibrium $e_-(g, I)$.

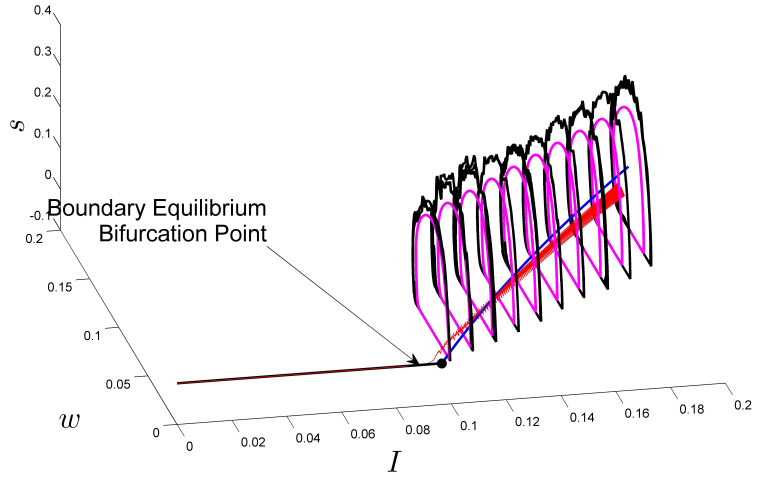


(a) $g = g^* - 1$

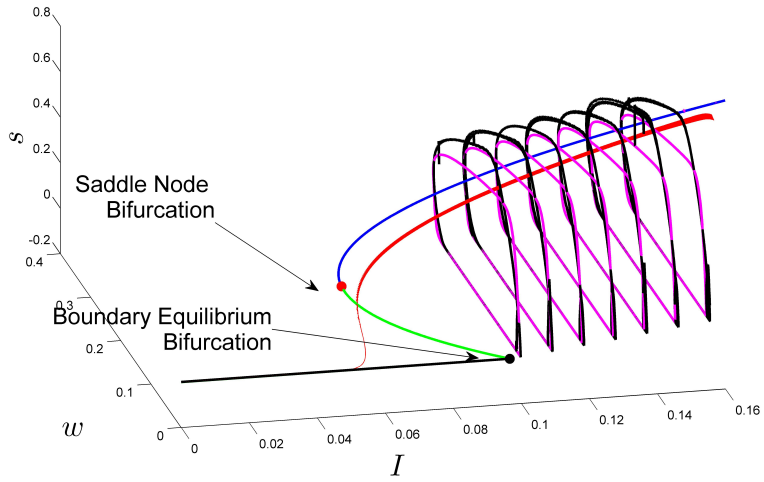


(b) $g = g^* + 1$

Figure 6.11: Comparison of the predictions of the mean field analysis and numerical simulations of a 100-neuron slow network with a slow current ramp for the Izhikevich system. The current is either descending (red) or ascending (green). (a) When $g < g^*$, as I is decreased the steady state solution for the network collides with the non-firing solution, as predicted by the mean field analysis. (b) When $g > g^*$, the descending current results in firing for $I < I_{rh}$, until the steady state falls off sharply near $I = I_{SN}$. The ascending current only results in firing when $I = I_{rh}$ is reached. This behaviour agrees with the mean field analysis which predicts there is bistability between these two stable states, with an unstable steady state separating them.



(a) $g = g^* - 1$



(b) $g = g^* + 1$

Figure 6.12: Using simulations of the slow network (red) to converge to the non-bursting steady state, and the full network (black) to converge to the stable bursting limit cycle, we can piece together a pseudo-bifurcation diagram for the full network of Izhikevich neurons that very closely mirrors the bifurcation diagram predicted from the non-smooth mean field equations. Indeed, it appears that the transitions that occur at $I = I_{rh}$ are well explained as non-smooth boundary equilibrium bifurcations of the mean variables of the full network. This suggests the existence of the co-dimension 2 non-smooth saddle-node BEB point for the mean variables of the actual network as well. Note that the limit cycles have been smoothed out for clarity, some of the high frequency oscillations due to synchrony in the peaks have been removed.

nonlinearity. For example, this is the case for the finite network model in the work of [54]. Thus, one has to consider non-smooth bifurcations and bifurcation analysis when working with mean field systems for type I neurons. The same is true of type 2 neurons, however the firing rate for these neurons changes discontinuously at I_{rh}, and thus it is likely that the mean field systems for type 2 neurons would be completely non-smooth, as opposed to piecewise smooth continuous.

6.5 Numerical Bifurcation Results for Networks with Noise

In this section we will numerically analyze the mean-field systems

$$\dot{s} = -\frac{s}{\tau_s} + \frac{\lambda_s}{\tau_s}\langle R_i(t)\rangle_\sigma \tag{6.46}$$

$$\dot{w} = -\frac{w}{\tau_w} + \frac{\lambda_w}{\tau_w}\langle R_i(t)\rangle_\sigma \tag{6.47}$$

where $\langle R_i(t)\rangle_\sigma$ is for a network with noise:

$$\langle R_i(t)\rangle_\sigma = \left(\frac{2}{\sigma^2}\int_{v_{reset}}^{v_{peak}}\int_{v'}^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}(M(v',w,s) - M(v,w,s))\right) dv'dv\right)^{-1} \tag{6.48}$$

$$M(v,w,s) = \int G(v',s,w)\, dv'. \tag{6.49}$$

The mean-field system for the network with noise is smooth to all orders, unlike the PWSC mean-field system for a noiseless homogeneous network. Hence the mean-field system for a network with noise acts as a regularization for the mean-field system without noise. We will again primarily consider the Izhikevich system, where

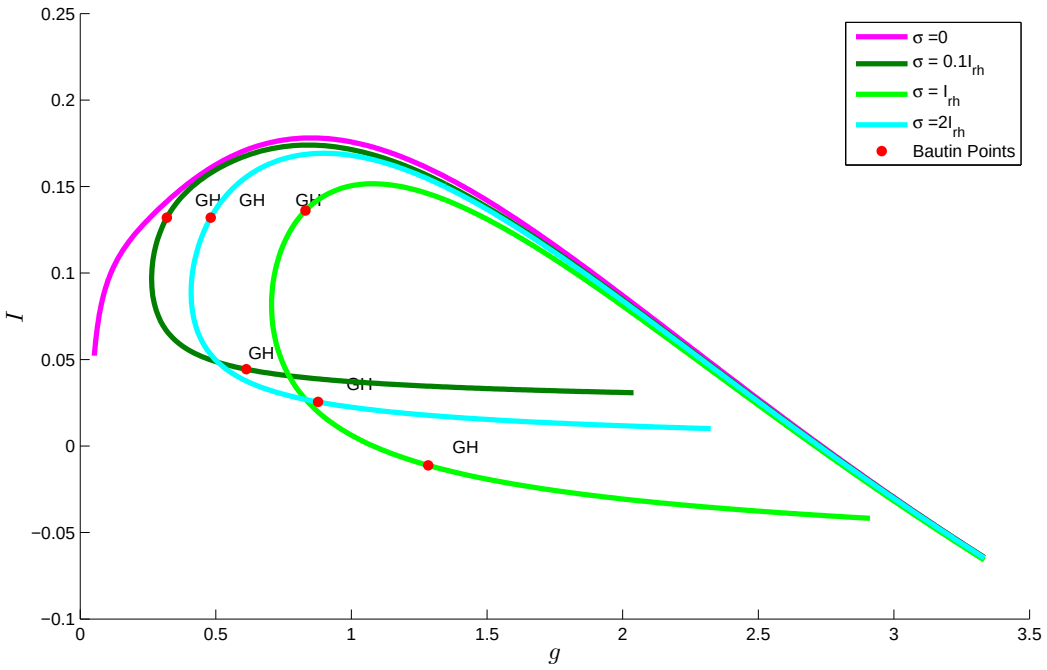
$$M(v,w,s) = \frac{v^3}{3} - \frac{(\alpha+gs)}{2}v^2 + v\left(I - \frac{(\alpha+gs)^2}{4} + gse_r - w\right)$$

First, we will see what happens to the limit cycle that emerges through a Hopf bifurcation. As it turns out, the Hopf bifurcation makes a change from a sub-critical Hopf bifurcation to a super-critical Hopf bifurcation in the low g region. This is shown in figure 6.13 for the mean-field system with noise.

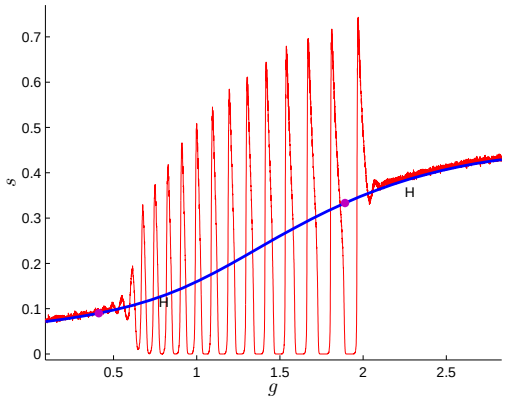
In figure 6.13, the bifurcation curves computed by MATCONT show two Bautin points that delineate three branches of Hopf bifurcations, two sub-critical branches and a super-critical branch. Additionally, it appears that as $\sigma \to 0$, in the limit of a noiseless system, the two points coalesce into a single point, probably at $(\bar{g}, I_{rh})$. This is somewhat unusual as the bifurcation point at $(\bar{g}, I_{rh})$ was defined by two-conditions, and hence corresponds to a co-dimension two bifurcation point. However, two-codimension two points colliding

with one another typically should result in a co-dimension 3 bifurcation. In the context of the mean-field system with noise, the third parameter can be thought of as σ with the bifurcation condition $\sigma = 0$, resulting in the mean-field system without noise. However, due to the incredible difficulty in explicitly computing the firing rate for the mean-field system with noise when $\sigma \ll 1$, it is difficult to determine numerically if these two points actually coalesce, and if this occurs at the Hopf boundary equilibrium bifurcation point. The mean-field system is qualitatively and quantitatively accurate for the behavior of the asynchronous steady state as shown in figures 6.13(b) and 6.13(c). The network considered here is an Izhikevich network with 3000 neurons using the intrinsically bursting parameter set, however the bifurcation structure is qualitatively similar across the parameter sets and models considered.

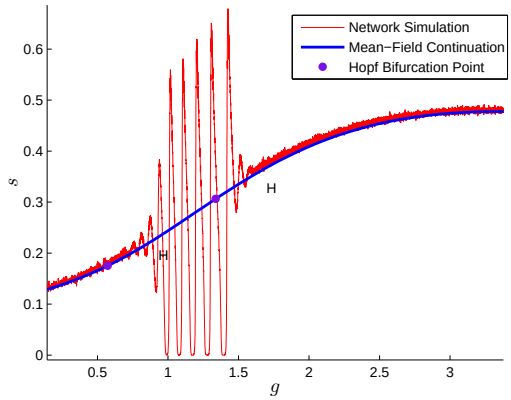
In addition to the boundary equilibrium Hopf bifurcation point, the boundary equilibrium saddle-node bifurcation point also becomes regularized as smooth bifurcation, in particular a cusp bifurcation point. This is shown in figure 6.14.



(a) Two parameter Hopf Continuation

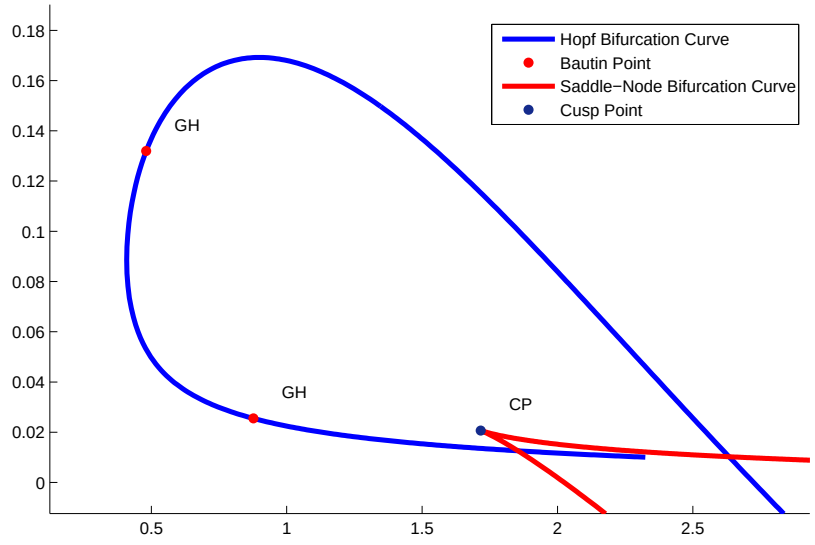


(b) $I = 0.15$

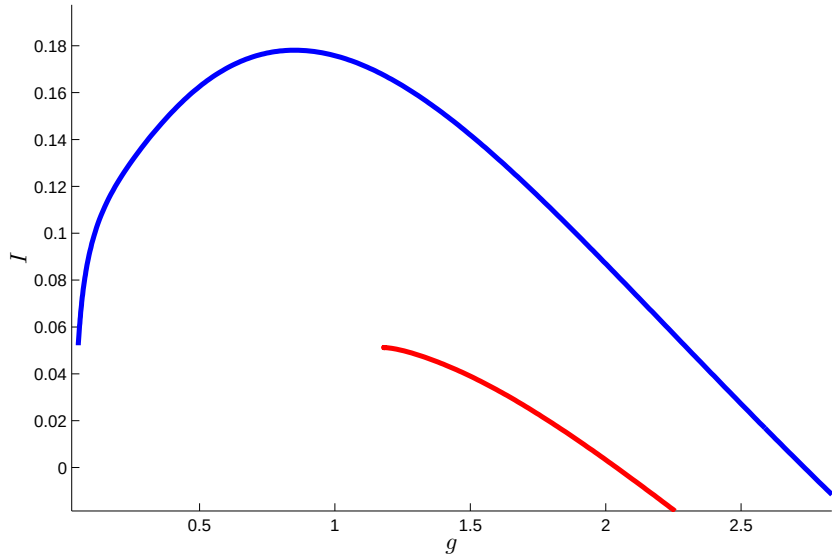


(c) $I = 0.1$

Figure 6.13: Shown above in figure 6.13(a) is the two parameter Hopf bifurcation curve for the mean-field system with noise and decreasing values of σ for the Izhikevich neurons with an intrinsically bursting parameter set [92]. The curves are computed by using the numerical bifurcation software MATCONT. It appears that the co-dimension 2 non-smooth bifurcation point at $g = \bar{g}$ for the noiseless system (magenta) is regularized as a pair of co-dimension 2 smooth Bautin points for the smooth mean-field system with noise. In the bottom panels, a network of 3000 Izhikevich neurons with the intrinsically bursting parameter set is simulated with a slowly time varying conductance, $g(t)$, and fixed I. The bifurcation points and steady state curve is determined through MATCONT. The locations where the network oscillations begin and end are well approximated by the Hopf bifurcation points.



(a) $\sigma = I_{rh}$



(b) $\sigma = 0$

Figure 6.14: Shown above are the two-parameter bifurcation curves for the mean-field system with noise ($\sigma = I_{rh}$, top) and without noise. The Hopf curves are plotted in blue and the saddle-node curves are plotted in red. The co-dimension two points $(\bar{g}, I_{rh})$ and (g^*, I_{rh}) appear to be regularized as a pair of Bautin points (also referred to as generalized Hopf bifurcation points, hence the GH labelling) and a cusp point (CP), respectively

Chapter 7

The Neural Engineering Framework, Inverse Mean-Field Theory, and Analytical Solutions to Weight Matrices for Networks of Neurons with Prescribed Dynamics

In previous chapters, we have analytically derived the macroscopic or mean-field dynamics for recurrently coupled networks of neurons. In this chapter, we will reverse our perspective and see if we can construct networks of neurons with prescribed macroscopic dynamics. It turns out that this problem is entirely tractable analytically, however for any prescribed set of macroscopic dynamics, the resulting network is non-unique.

One possible numerical solution to the problem of constructing networks with prescribed macroscopic dynamics is the Neural Engineering Framework (NEF) [52]. In the NEF approach, given a network of neurons with a source of heterogeneity, one can find a set of optimal linear weights, referred to as linear decoders, for their firing rates in such a way that the weighted linear sum of the firing rates optimally approximates any function of choice. This allows for specifying the network connectivity in such a way as to obtain arbitrary dynamics from the network(s) of neurons [51]. For example, the NEF has been used to develop a wide variety of models, including the most behaviorally sophisticated spiking neural model to date [53] as well as more specialized models of path integration [37], working memory [167], visual attention [19], motor control [46], various cognitive functions [15, 148], and many others.

However, the optimality requirement in the linear decoders introduces complications in the NEF approach. The optimal decoders are computed via least-squares optimization which is a computationally-intensive process; and yet very little information about the network can be determined once the optimal decoders have been obtained. Additionally,

one cannot determine how the distribution of heterogeneity in the neurons is related to the other distributions across the network, such as the distribution of connection weights.

Here, we will show that if one loosens the optimality requirement in the linear decoders, it is possible to obtain linear decoders that converge to any function of choice in the large network limit. Due to their form, we will refer to these decoders as scale-invariant linear decoders. These scale-invariant decoders have several advantages over optimal decoders, at the primary cost of a slower convergence rate in network size. However, using any gradient descent algorithm that does not directly compute the Hessian, one can decrease the error of the scale-invariant decoders very rapidly with very few iterations for any finite network size.

In section 7.1, we will introduce the Neural Engineering Framework and demonstrate that as the networks become arbitrarily large, the optimal decoders tend to an asymptotic limit in section 7.1.1. This will be our motivation for defining a scale-invariant decoder. In section 7.2 we will determine what this asymptotic limit is for the scalar case and for multivariable functions in section 7.2.4. In section 7.3 we will demonstrate how the decoders can yield the weights and simulate spiking networks with the specified dynamics by using these weights.

7.1 The Neural Engineering Framework

Suppose we knew the firing rate for a class of neurons, $f(I)$ as a function of the input current I. This is typically referred to as the FI curve. We can take any input variable x and linearly transform it into a current via $I = \alpha x + \beta$. If we allow α and β to be drawn from a random distribution, then we can generate a network of neurons with firing rates $f(\alpha_i x + \beta_i)$ where α_i, β_i are drawn from some specified probability distribution $\rho_{\alpha,\beta}(\alpha, \beta)$. As a function of x the curve $f(\alpha_i x + \beta_i)$ is typically referred to as the tuning curve of the neuron. The output of these neurons is the sum of their weighted firing rates:

$$\hat{g}_N(x) = \sum_{i=1}^{N} \phi_i f(\alpha_i x + \beta_i). \tag{7.1}$$

Thus, the network takes any input x belonging to the appropriate space, and transforms it into some function $\hat{g}_N(x)$. If for example we wanted to compute the function $g(x)$, we would need to pick ϕ_i such that $\hat{g}_N(x) \approx g(x)$. The ϕ_i are referred to as the linear decoders in the NEF approach [52]. They can be determined by minimizing the the following functional with respect to ϕ over some region X in x [52, 157]:

$$C(\phi) = \int_X (\hat{g}_N(x) - g(x))^2 \, dx + \lambda \sum_{i=1}^{N} \phi_i^2 = \int_X \left(\sum_{i=1}^{N} \phi_i f(\alpha_i x + \beta_i) - g(x) \right)^2 dx + \lambda \sum_{i=1}^{N} \phi_i^2 \tag{7.2}$$

where the first term in (7.2) corresponds to the error in the approximation and the second term penalizes large ϕ_i. Minimizing $C(\phi)$ for ϕ yields the following linear system of equations:

$$\begin{aligned} \phi^* &= A^{-1}\Gamma && (7.3)\\ A_{ij} &= \int_X f(\alpha_i x+\beta_i) f(\alpha_j x+\beta_j)\,dx + \delta_{ij}\lambda && (7.4)\\ \Gamma_j &= \int_X f(\alpha_i x+\beta_i) g(x)\,dx. && (7.5) \end{aligned}$$

Equations (7.3)-(7.5) correspond with standard function approximation, although the basis functions f are randomly drawn [18]. We will refer to the optimal decoders as ϕ^* and any other decoder as ϕ. There are various functions f that have appeared in the literature. These are derived from complicated neural models using topological normal form theory [60, 92], are experimental fits to data from real neurons or conductance based models, or are analytically derived from integrate and fire neurons. The general form of the integrate-and-fire models we will consider is given by

$$\begin{aligned} \dot{v} &= F(v)+I && (7.6)\\ v(t^-) &= v_{peak}, \quad \to v(t^+)=v_{reset} && (7.7)\\ & && (7.8) \end{aligned}$$

which has the firing rate

$$f(I) = \begin{cases} \left(\int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v)+I}\right)^{-1} & I>0 \\ 0 & I<0 \end{cases}. \qquad (7.9)$$

Specific examples that we have seen in previous chapters include:

$$\begin{aligned} F(v) &= -\frac{v}{\tau_v} \quad \text{(Leaky Integrate-and-Fire Model [1, 25, 105])} && (7.10)\\ F(v) &= v^2 \quad \text{(Quadratic Integrate-and-Fire Model) [90, 92]} && (7.11)\\ F(v) &= v^2, \quad v_{reset}=-\infty, \quad v_{peak}=\infty \quad \text{(Theta Model)[60]} && (7.12)\\ F(v) &= \exp(v)-v \quad \text{(Exponential-Integrate-and-Fire Model)[23, 130].} && (7.13) \end{aligned}$$

Other FI curves can be determined from the measured FI curves of more sophisticated conductance based models or experimental measurements. For example, the function

$$f(I) = \begin{cases} I+c & I>0 \\ 0 & I<0 \end{cases} \qquad (7.14)$$

can be fit to type-II firing rates when $c>0$, and can be shown to be the steady state firing rate for neurons that display spike frequency adaptation when $c=0$ [57]. Equation (7.14) has also been fit to conductance based models in (with $x=0$) [163] and adequately describes the $f(I)$ curves for many real cortical neurons [4, 11, 172].

As x is often thought of as a real world input variable in the NEF approach, the α_i, β_i distribution can only be known once one specifies a distribution of maximal firing rates, r_i^{max} and x-intercepts, a_i for the tuning curves. For the time being, we will restrict the variable x to the interval $[-1, 1]$. It can be rescaled to an arbitrary interval, so this is no loss of generality. We will show later how x can also be extended to a vector as in the original NEF framework [51, 52]. Once one specifies the distribution of (r_i^{max}, x_i), one obtains the transformation of random variables:

$$r_i^{max} = f(\alpha_i + \beta_i) \tag{7.15}$$

$$0 = \alpha_i a_i + \beta_i \tag{7.16}$$

where the maximal firing rate, r_i^{max} is achieved at $x = +1$. As the firing rate of neurons can either increase or decrease with respect to x, we can think of the neurons as two subpopulations: an ON and OFF. Solving equations (7.15)-(7.16) yields the transformation of random variables for the ON population.. The maximal firing rate for neurons in the ON population is reached at $x = 1$ [52]. Multiplying α_i by -1 yields the transformations for the OFF neurons:

$$\alpha_i = \pm \frac{f^{-1}(r_i^{max})}{1 - a_i} \tag{7.17}$$

$$\beta_i = -\frac{a_i f^{-1}(r_i^{max})}{1 - a_i} \tag{7.18}$$

where the $\pm$ indicates ON/OFF, respectively. We will treat a_i and r_i^{max} as our primary sources of heterogeneity in the case of approximating a function of a single variable and we will assume that the marginal densities are given by $\rho_a(a)$ and $\rho_r(r^{max})$. In this case, we can rewrite the sum (7.1) as

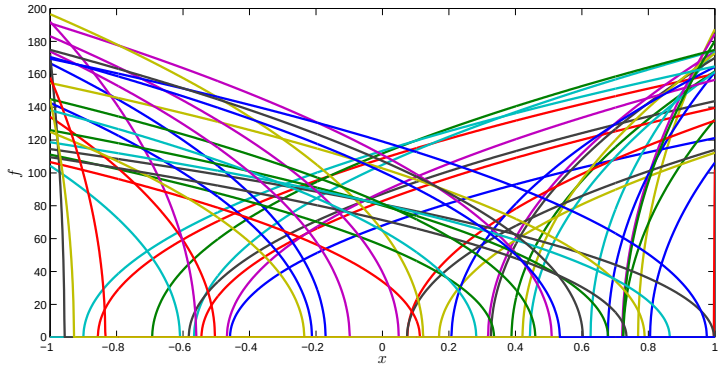
$$\hat{g}_N(x) = \sum_{i=1}^{N/2} \phi_i f\left(f^{-1}(r_i^{max})\left(\frac{x - a_i}{1 - a_i}\right)\right) + \sum_{i=1}^{N/2} \phi_i f\left(f^{-1}(r_i^{max})\left(\frac{-x - a_i}{1 - a_i}\right)\right) \tag{7.19}$$

where the first half represents the population of ON neurons and the second half of the sum represents the population of OFF neurons.

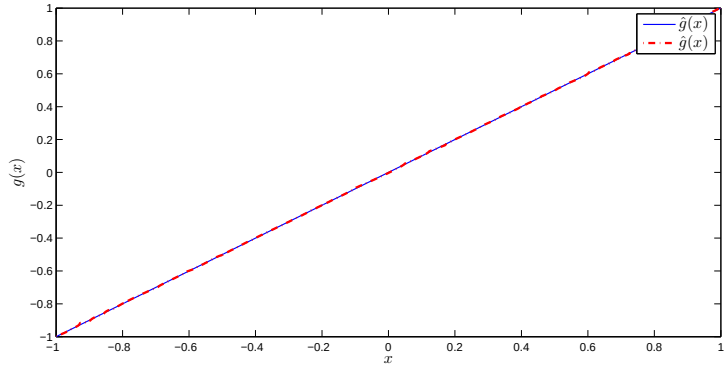
Suppose, for example, we wanted to approximate the function $g(x) = x$ using a population of 50 quadratic integrate and fire tuning curves with 25 ON and 25 OFF neurons. This is shown in figure 7.1 where the decoders are given by equation (7.5). Note that reasonable accuracy is achieved despite the small population of neurons.

So far this has been fairly standard function approximation with a non-orthogonal basis [18]. The difference in the NEF approach is that one uses these linear decoders obtained from the firing rate curves to design a network of spiking neurons and the function is represented in the output of the network simulation. For example, recall the differential equation for the quadratic integrate and fire model is given by

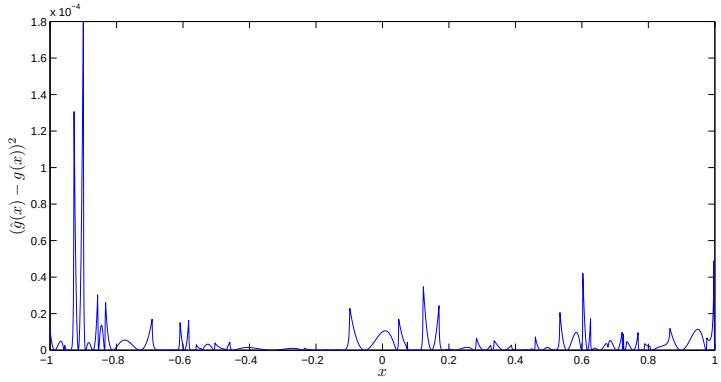
$$\dot{v}_i = v_i^2 + \alpha_i x + \beta_i \tag{7.20}$$



(a) $f(\alpha_i x + \beta_i)$ for $i = 1, 2, \ldots 50$



(b) $g(x)$



(c) $(\hat{g}(x) - g(x))^2$

Figure 7.1: Function approximation with neuronal tuning curves. Figure 7.1(a): The tuning curves for a population of 50 quadratic integrate and fire neurons with their intercepts and maximum firing rates drawn from independent uniform random variables on $[-1, 1]$ and $[100, 200]$. The maximum firing rate is in Hertz. Figure 7.1(b): The function $g(x) = x$, in addition to the approximation $\hat{g}(x)$ using the tuning curves from Figure 7.1(a). Figure 7.1(c): The squared error in the approximation $\hat{g}(x)$.

where if $v(t^-) = \infty$, $v(t^+) = -\infty$, or more informally, it has a peak at ∞ and is reset to $-\infty$. This can be written as the equivalent θ model with the transformation $v = \tan(\theta/2)$ yielding:

$$\dot{\theta}_i = 1 - \cos(\theta_i) + (1 + \cos(\theta_i))(\alpha_i x + \beta_i) \tag{7.21}$$

which produces a spike when $\theta(t^-) = \pi$ and is reset to $\theta(t^+) = -\pi$. Each of these differential equations generates a sequence of action potentials at specific spike times, t_{jk} where t_{jk} is the kth spike fired by the jth neuron. These spike times are then fed into a post-synaptic filter $s(t)$;

$$s'(t) = -\frac{s(t)}{\tau_s} + \frac{1}{\tau_s} \sum_{j=1}^{N} \sum_{t<t_{jk}} \phi_j \delta(t - t_{jk}). \tag{7.22}$$

The linear decoders, ϕ_j, are used to weight the spikes of their corresponding neurons. This post-synaptic filter equation can be explicitly integrated to yield:

$$s(t) = \sum_{j=1}^{N} \sum_{t<t_{jk}} \phi_j \exp\left(\frac{t_{jk} - t}{\tau_s}\right) = \sum_{j=1}^{N} \sum_{t<t_{jk}} \phi_j E(t - t_{jk}) \tag{7.23}$$

where $E(t) = \exp(-t/\tau_s)$. The integrated spike train for the jth neuron is approximately equal to its tuning curve, $f(\alpha_j x + \beta_j)$:

$$\int_0^t \sum_{t<t_{jk}} \delta(t - t_{jk})\, dt \approx \int_0^t f(\alpha_j x + \beta_j)\, dt,$$

provided that x varies on a suitably slow time scale [41, 52]. In this case, the dynamics in equation (7.22) are approximately given by

$$s' = -\frac{s}{\tau_s} + \frac{1}{\tau_s} \sum_{j=1}^{N} \phi_j f(\alpha_j x + \beta_j) \tag{7.24}$$

This allows one to approximate an arbitrary dynamical system [51]. For example, if we consider a recurrent network ($x = s$), then to approximate the dynamics $s' = G(s)$ we merely require

$$\sum_{i=1}^{N} \phi_i f(\alpha_i s + \beta_i) \approx s + \tau_s G(s) = \hat{g}_N(s) \tag{7.25}$$

and where the ϕ_i are given by equation (7.3). Returning to the neural equations, if we take $x = s$ and consider a recurrently coupled network of neurons then we have the following:

$$\begin{aligned} \dot{v}_i &= F(v_i) + \alpha_i s + \beta_i & (7.26)\\ &= F(v_i) + \alpha_i \sum_{j=1}^{N} \sum_{t<t_{jk}} \phi_j E(t - t_{jk}) + \beta_i & (7.27)\\ &= F(v_i) + \sum_{j=1}^{N} \sum_{t<t_{jk}} \omega_{ij} E(t - t_{jk}) + \beta_i & (7.28) \end{aligned}$$

where $\omega_{ij} = \alpha_i \phi_j$ is the NEF equation for the weight coupling neuron j to neuron i [51, 52] and the quantity

$$I_{syn,i} = \sum_{j=1}^{N} \sum_{t<t_{jk}} \omega_{ij} E(t - t_{jk}) + \beta_i$$

is the post-synaptic current going to the ith neuron.

For example, if we wanted the macroscopic dynamics to be exponential decay, $G(s) = ks$, then we require $\hat{g}_N(s) = s(1 + \tau_s k)$. We would obtain the ϕ_i by using equations (7.3)-(7.5) which yields the optimal decoders ϕ_i^* for $\hat{g}_N(x) = x(1+\tau_s k)$ and simulate our spiking network using the weights $\omega_{ij} = \alpha_i \phi_j^*$. This yields a recurrently coupled spiking neural network with macroscopic dynamics $s' = ks$.

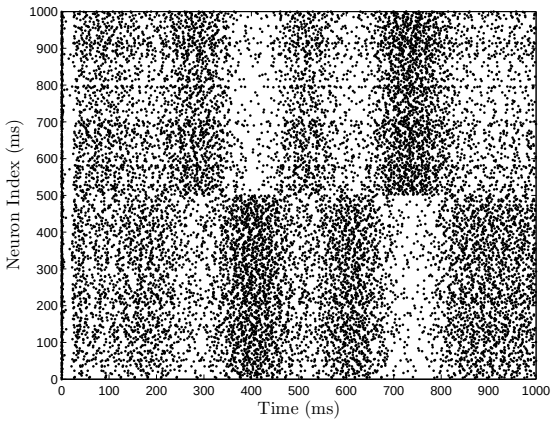
In addition to recurrent networks, one can also construct feedforward networks with the NEF approach. For example, we can also treat x as an input variable. This allows a network to represent an input variable x in terms of its spiking. If τ_s is not large, as in the case of excitatory synapses, then one can represent the input variable x as a postsynaptic current s:

$$s \approx \sum_{i=1}^{N} \phi_i f(\alpha_i x + \beta_i) = \hat{g}_N(x) \approx x \tag{7.29}$$

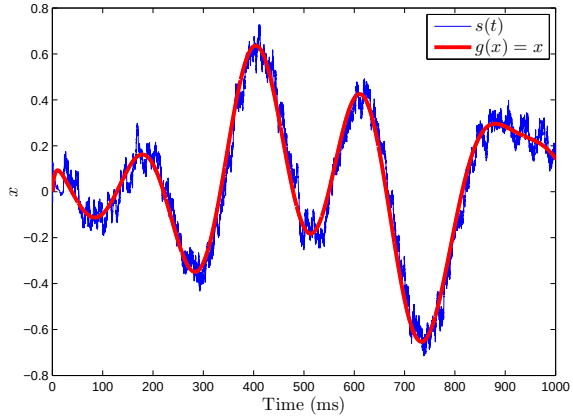
assuming that x varies on a suitably slow time scale (slower than τ_s). This is shown for example with networks of various sizes in Figure 7.2, with a synaptic time constant of $\tau_s = 5$ ms approximating the function $g(x) = x$. The network of differential equations for the neurons is simulated using equation (7.20). These neurons then generate a spike train which is weighted by the decoders. The weighted spike train is fed into the post-synaptic current variable $s(t)$, which acts as the approximation for $g(x) = x$. A time varying $x(t)$ is used that varies on a suitably slow time scale.

It is clear that given the fact that arbitrary functions or dynamics (via recurrent networks) can be computed, then one can generate multiple networks that perform different functions, and feed into one another. In this way, one could create large networks composed of interconnected subnetworks that perform functions such as controlling limbs, detecting objects, and performing tasks by using the mathematical approaches that already exist for accomplishing these feats and translating them into an equivalent neural network representation. This is the core idea in the NEF [52, 53].

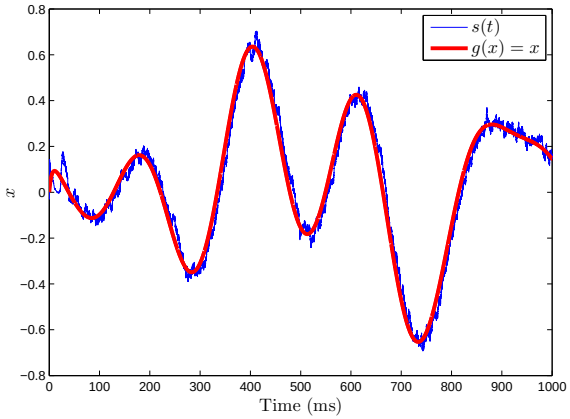
Although a network of $N = 100$ neural tuning curves $f(\alpha x_i + \beta_i)$ is sufficient for a good approximation of many functions, depending on the dynamics being computed, significantly more neurons are needed when we are generating actual spikes and use post-synaptic currents, as shown in figure 7.2. Hundreds, if not thousands, of neurons are necessary for adequate approximation when spikes are used. The network size becomes even larger when we want to perform complicated functions involving more than one variable x. As the decoders are determined by large matrix inversion (equation 7.3), this can take quite a while when dealing with more then 5000 neurons on a conventional computer.



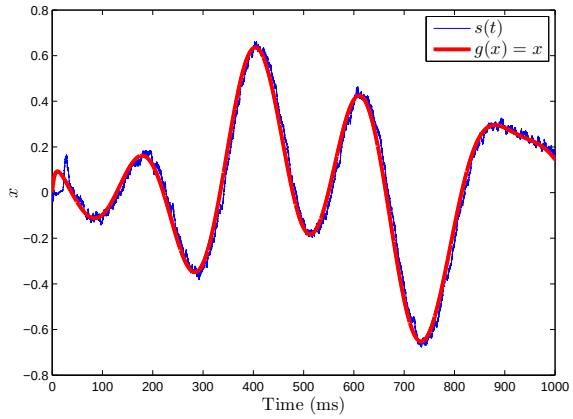
(a) Spike Raster, $N = 1000$



(b) $N = 1000$



(c) $N = 5000$



(d) $N = 10000$

Figure 7.2: Representation with a spiking neuronal network. The function $g(x) = x$ is approximated by networks of various sizes using equation (7.20). This generates a spike train, as shown in figure 7.2(a) for a network of $N = 10^3$ neurons. A time varying randomly generated signal (red) is fed into the network, and is computed via the synaptic current variable $s(t)$ using equation 7.22 (blue). As the network size increases, the approximation becomes better.

Furthermore, the smaller the synaptic time constant τ_s, the more neurons are required. This is due to the fact that equation (7.24) is effectively a kernel density estimator of the firing rate and when the bandwidth is too small, the resulting estimate is under-smoothed, thus requiring more neurons for a comparable degree of accuracy as that of a network with larger τ_s. Furthermore, as the NEF weights are numerically determined (via the NEF decoders), one cannot conduct any meaningful analysis of the behavior of large recurrent networks. Thus, an analytical solution to the NEF decoders (and thus weights) would allow a greater insight into large networks and may also facilitate faster numerics.

7.1.1 Decoder Asymptotics as $N \to \infty$.

In order to proceed analytically, we will first look at the behavior of the optimal decoders ϕ^* for large networks ($N \to \infty$). To facilitate plotting, let us consider the case where for an arbitrary neural model, $f^{-1}(r_i^{max}) = \pm 1 - a_i$, that is the maximim firing rate is given by $r_i^{max} = f(\pm 1 - a_i)$, which reduces the random variables associated with the heterogeneity to the intercept variable, a_i. Additionally, note that for Type-I neurons:

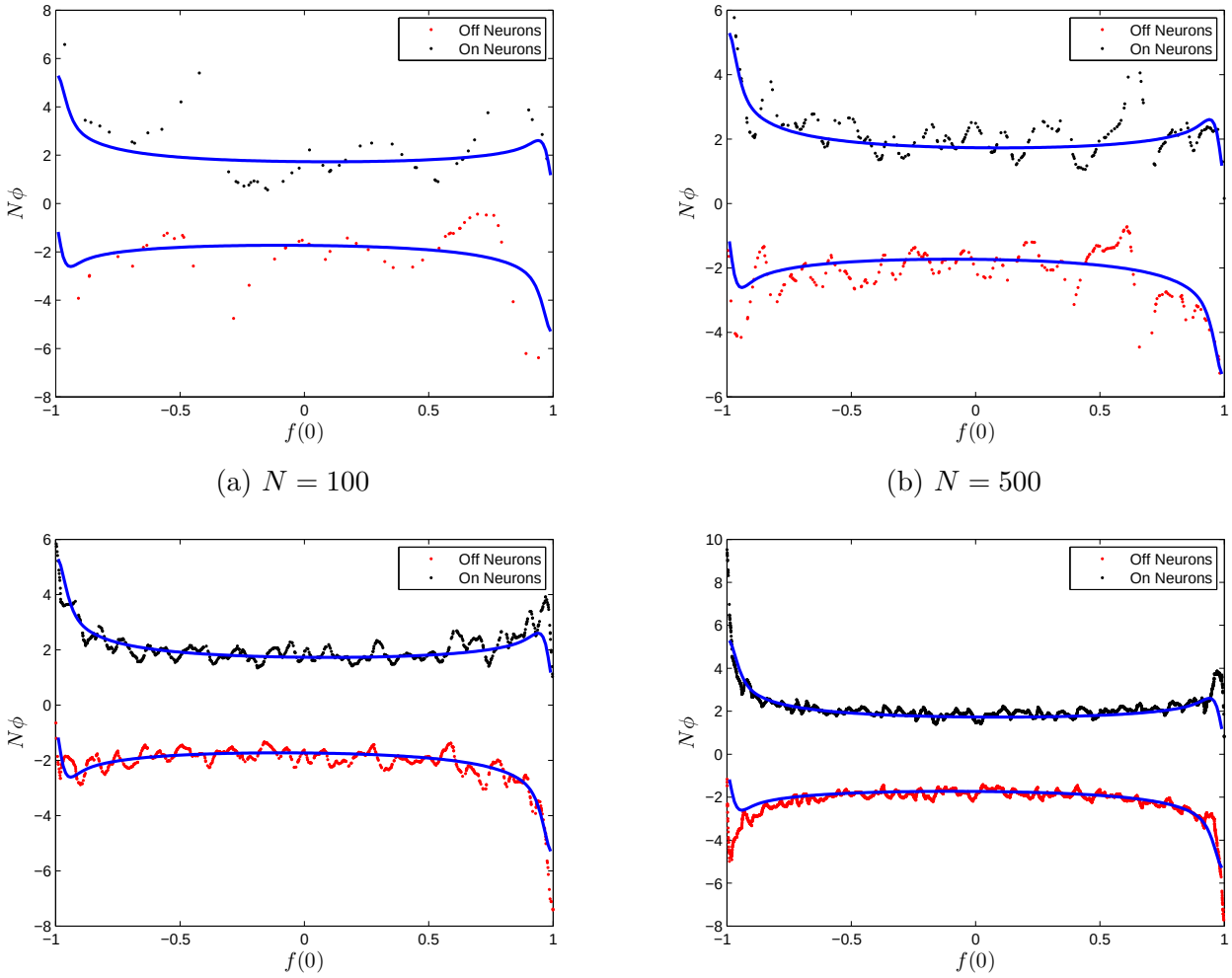
$$f\left(\frac{f^{-1}(r_i^{max})}{1-a_i}(\pm x - a_i)\right) = \frac{r_i^{max}}{\sqrt{1-a_i}}\sqrt{\pm x - a_i} \tag{7.30}$$

which immediately implies that we can absorb the quantitiy $\frac{r_i^{max}}{\sqrt{1-a_i}}$ into the decoder ϕ_i and rescale any solution we obtain by this quantity at the end. Recall that the $\pm$ correspond to the ON and OFF neurons respectively. With equation (7.30), the sum in (7.19) becomes:

$$\hat{g}_N(x) = \sum_{i=1}^{N/2} \phi_i^* f(x - a_i) + \sum_{i=1}^{N/2} \phi_i^* f(-x - a_i) \tag{7.31}$$

where the ϕ_i^* are determined by equation (7.5). If one were to plot the decoders for large N, then one can easily see that in the limit of large network size ($N \to \infty$), the individual decoders vanish ($\phi_i^* \to 0$, not shown here). However, for increasing N, it seems that the quantity $\gamma_i = N\phi_i^*/2$ converges to some non-zero value $\gamma(a_i)$ and thus it appears that γ_i converges to some function of the x-intercept, a, the source of heterogeneity for the neurons. This is shown in figure 7.3 for increasingly large networks. The quantity $N\phi_i^*/2$ is plotted versus a_i. The predicted surface for convergence, $\gamma(a)$, is also plotted which is determined by optimizing over a uniform mesh in the parameter space. We will refer to any γ_i that satisfies $\gamma_i = \phi_i N/2$ for some decoder ϕ_i as *scale invariant decoders* and $\gamma^{\pm}(a)$ as the *decoder surface*. We will not necessarily use the same decoders for the ON and OFF neurons hence the superscript on $\gamma(a)$. We will show in the subsequent sections how to determine the decoder surfaces for the type-I and type-II (approximate) firing rates for single variable and multivariable functions.

In order to determine the decoder surface analytically, we need to understand the behavior of the network as $N \to \infty$. Using the scale invariant decoders from equation



(a) $N = 100$ (b) $N = 500$

(c) $N = 2000$ (d) $N = 10000$

Figure 7.3: Convergence of the optimal decoders to an invariant surface. The function $g(x) = x$ is approximated by networks of various sizes using the firing rate curves for theta/Quadratic integrate-and-fire neurons. Plotted are the decoders ϕ_i scaled up by the network size N for ON (black dots) and OFF (red dots) neurons as a function of the intercept, $f(0)$ for QIF firing rate functions. The quantity $N\phi_i$ appears to converge as $N \to \infty$ to the blue curve.

(7.31):

$$\hat{g}_N(x) = \sum_{i=1}^{N/2} \phi_i f(x-a_i) + \sum_{i=1}^{N/2} \phi_i f(-x-a_i) \tag{7.32}$$

$$= \frac{2}{N}\sum_{i=1}^{N/2} \gamma^+(a_i) f(x-a_i) + \frac{2}{N}\sum_{i=1}^{N/2} \gamma^-(a_i) f(-x-a_i) \tag{7.33}$$

$$= \overline{\gamma_i f(x-a_i)} + \overline{\gamma_i f(-x-a_i)} \tag{7.34}$$

where the overline denotes the finite average over the inhomogeneity in the intercepts. As $N \to \infty$, our finite averages become expectations:

$$\overline{\gamma_i f(x-a_i)} \to E(\gamma^+(a_i) f(x-a_i)) = \int_{-1}^{x} \gamma^+(a)\rho_a(a) f(x-a)\, da \tag{7.35}$$

$$\overline{\gamma_i f(-x-a_i)} \to E(\gamma^-(a_i) f(-x-a_i)) = \int_{-1}^{-x} \gamma^-(a)\rho_a(a) f(-x-a)\, da \tag{7.36}$$

where $\rho_a(a)$ is the probability density describing the heterogeneity variable a_i for the neurons. Now, in order to do this, we would need to formally show that $\phi_i^* \to \gamma(a_i)/N$, where ϕ_i^* are the optimal decoders. However, this is unnecessary as we can regard $\phi_i = \gamma(a_i)/N$ as a suboptimal decoder and independent of the optimal decoders which are generated by minimizing the integral (7.2). Furthermore, as we shall show later, ϕ_i^* will not necessarily converge to $\gamma(a_i)$ as $\gamma(a_i)$ is non unique, where as ϕ_i^* is the optimal decoder which is unique due to the quadratic error surface in $C(\phi)$.

Note that $\gamma(a)\rho_a(a)$ appears as a product in the integral. These terms can be collapsed into a single function $\hat{P}^{\pm}(a) = \gamma^{\pm}(a)\rho_a(a)$. We will refer to these quantities as the *weighted decoders* of the ON/OFF neurons and use the weighted decoders to define the linear operators;

$$L_+(\hat{P}^+) = \int_{-1}^{x} \hat{P}^+(a) f(x-a)\, da = \hat{g}_+(x) \tag{7.37}$$

$$L_-(\hat{P}^-) = \int_{-1}^{-x} \hat{P}^-(a) f(-x-a)\, da = \hat{g}_-(x) \tag{7.38}$$

$$M(\hat{P}^+, \hat{P}^-) = L_+(\hat{P}^+) + L_-(\hat{P}^-) = \hat{g}(x). \tag{7.39}$$

which we will refer to as the *tuning curve transforms* (TCT). The TCTs map functions from the space of the variable(s) assigned to the heterogeneous parameters to the space of functions we are trying to approximate. Note that these operators are actually applied to different weighted decoders as the decoder surfaces are different for ON and OFF neurons. Furthermore, the density $\rho_a(a)$ need not be identical for both ON and OFF neurons. However in all numerical implementations, $\rho_a(a)$ will be identical for the sake of simplicity.

Suppose we could determine $\hat{P}(a)$ analytically. In this case, as $\gamma(a)\rho_a(a) = \hat{P}(a)$, whenever $\rho_a(a) \neq 0$, we can compute $\gamma(a) = \hat{P}(a)/\rho_a(a)$ and leave $\gamma(a)$ undefined otherwise (as there is no neuron that has parameter(s) in this region). Now, given the fact that we

obtain a linear operator as $N \to \infty$ case, the real problem becomes in finding the $(\hat{P}^+, \hat{P}^-)$ such that $M(\hat{P}^+, \hat{P}^-)$ maps to $g(x)$, the function we want to approximate. That is, we have to invert the operator M for these $\hat{P}$. If we know this $\hat{P}$, then as we presumably know the distribution of tuning curve intercepts, we can determine the decoders ϕ_i with:

$$\phi_{\pm}(a_i) = \frac{2\gamma^{\pm}(a_i)}{N} = \frac{2}{N}\frac{\hat{P}^{\pm}(a_i)}{\rho_a(a_i)}. \tag{7.40}$$

Thus the analytically determined scale-invariant decoders $\gamma(a_i)$ are effectively weights for a Monte-Carlo sum of the tuning curves.

7.2 Determining the Decoder Surface

We will explicitly invert the tuning curve transforms for single variable functions in section 7.2.1. The resulting equation for the weighted decoders is a convolution integral. In section 7.2.4 we will show that with a basis to basis mapping, one can also invert the tuning curve transforms for multi-variable functions.

7.2.1 Single Variable Functions

If we work with the operators L^+ and L^- separately, the problem becomes entirely tractable. One of the surprising things about the operators L_+ and L_- is that provided that the functions we are considering are constrained to a subset where $g_+(x)$ vanishes to first order at $x = -1$ and $g_-(x)$ vanishes to first order at $x = 1$, and are both smooth, then the operators are invertible analytically on this constrained subspace of functions using Laplace transforms. Additionally, by using both $P^+(a)$ and $P^-(a)$, one can compute any smooth function irrespective of the conditions at $x = \pm 1$. Furthermore, piecewise smooth continuous functions can also be computed (see Appendix). Closed form solutions do not exist for all neuronal firing rates as the Laplace transform cannot always be inverted explicitly. However, the type-I and type-II firing rate models do have analytically determined decoders.

Type-I Weighted Decoders

Here, we will make use of the Laplace transform to analytically determine the functions one can represent and numerically invert for the decoder surfaces in the NEF approach. In particular, consider the operator $L_+(\hat{P})$ which is given by

$$\hat{g}_+(x) = L_+(\hat{P}) = \int_{-1}^{x} \hat{P}(a)\sqrt{x-a}\, da \tag{7.41}$$

Using a series of substitutions we can write the following:

$$\hat{g}_+(x) = \int_0^{x+1} \hat{P}(b-1)\sqrt{x+1-b}\,db \tag{7.42}$$

$$= \int_0^{y} \hat{P}(b-1)\sqrt{y-b}\,db = \int_0^{y} Q(b)R(y-b)\,db = m(y) \tag{7.43}$$

where $Q(b) = \hat{P}(b-1) = \hat{P}(a)$ and $R(y-b) = \sqrt{y-b} = \sqrt{x+1-b} = \sqrt{x-a}$, and $m(y) = \hat{g}(y-1) = \hat{g}(x)$. Thus the expectation is a convolution when written in the appropriate variables:

$$m(z) = (Q \star R)(z) \tag{7.44}$$

where $\star$ denotes the convolution operator. Taking the Laplace transform yields:

$$\mathcal{L}(m(z)) = \mathcal{L}(Q(z))\mathcal{L}(R(z)) = \mathcal{L}(Q(z))\frac{2}{\sqrt{\pi}s^{3/2}} \tag{7.45}$$

$$\Rightarrow \mathcal{L}(Q(z)) = \frac{2}{\sqrt{\pi}}\mathcal{L}(m(z))s^{3/2} \tag{7.46}$$

$$= \mathcal{L}(m(z))\mathcal{L}(C(z)) \tag{7.47}$$

for some function $C(z)$. Unfortunately $\mathcal{L}(C(z)) = \frac{2}{\sqrt{\pi}}s^{3/2}$ is not invertible as we have:

$$\mathcal{L}(\sqrt{t}) = \frac{1}{2}\frac{\sqrt{\pi}}{s^{3/2}} \quad \mathcal{L}(f''(t)) = s^2\mathcal{L}(f(t)) + sf(0) + f'(0) \tag{7.48}$$

from the general properties of the Laplace transform. This implies that if $f(t) = -4\sqrt{t}$, $f''(t) = t^{-3/2}$ and the Laplace transform is undefined as $f'(0) \to \infty$. Suppose instead that we assume that $m(z)$ is twice differentiable and that $m(0) = 0$ and $m'(0) = 0$. Then we can write the following:

$$\mathcal{L}(Q(z)) = \frac{2}{\sqrt{\pi}}s^{3/2}\mathcal{L}(m(z)) \tag{7.49}$$

$$= \frac{2}{\sqrt{\pi}}s^{-1/2}\mathcal{L}(m''(z))$$

$$= \mathcal{L}\left(\frac{2}{\pi\sqrt{z}}\right)\mathcal{L}(m''(z))$$

$$\Rightarrow Q(z) = \int_0^{z} \frac{2}{\pi\sqrt{t}}m''(z-t)\,dt \tag{7.50}$$

This is how to find $Q(z)$ as a convolution of $t^{-1/2}$ with $m''(z)$, the second derivative of the function we wish to approximate. Based on the assumptions above, we require that $m(z) = bz^2 + O(z^3)$. Writing the convolution in terms of the original variables, we have the following:

$$\hat{P}(a) = \int_0^{a+1} \frac{2}{\pi\sqrt{t}}\hat{g}''(a-t)\,dt \tag{7.51}$$

This implies that we can approximate any function that vanishes to two orders at $x = -1$ using the following decoder for the ON neurons:

$$\phi_i^+(a_i) = \frac{2}{N}\frac{1}{\rho_a(a_i)}\int_0^{a_i+1}\frac{2}{\pi\sqrt{t}}\hat{g}''(a_i - t)\,dt \tag{7.52}$$

Additionally, if one repeats this process for a population of OFF neurons, then we can immediately write down the decoder for a function g that vanishes to two orders at $x = 1$:

$$\phi_i^-(a_i) = \frac{2}{N}\frac{1}{\rho_a(a_i)}\int_0^{a_i+1}\frac{2}{\pi\sqrt{t}}\hat{g}''(-a_i - t)\,dt \tag{7.53}$$

Given the constraints on the derivatives of $\hat{g}$, it is clear that using a population of ON and OFF neurons, the resulting approximant has the form:

$$\begin{aligned}\hat{g}(x) &= \hat{g}_+(x) + \hat{g}_-(x)\\ &= \frac{1}{2}\left(g(x) - g(-1) - (x+1)g'(-1)\right) + \frac{1}{2}\left(g(x) - g(1) - (x-1)g'(1)\right)\\ &= g(x) - A - Bx\end{aligned} \tag{7.54}$$

where $g(x)$ is the function we want to approximate. The remainder is a linear term, $A+Bx$. If we can approximate an arbitrary linear function with a population of type-1 neurons with heterogeneity, then we can accommodate the remainder term. Note the following:

$$\frac{2}{\pi}\int_{-1}^{\pm x}\frac{\sqrt{\pm x - a}}{\sqrt{a+1}}\,da = \frac{4}{\pi}\int_0^{\sqrt{\pm x+1}}\sqrt{x+1-u^2}\,du = 1 \pm x \tag{7.55}$$

Thus, with the functions $\hat{P}^+ = \frac{2C}{\pi\sqrt{1+a}}$ and $\hat{P}^- = \frac{2D}{\pi\sqrt{1-a}}$, for the ON and OFF neurons respectively, we can approximate

$$C(x+1) + D(1-x) = (C-D)x + (C+D) = Bx + A$$

with $C = (A+B)/2$ and $D = (A-B)/2$. Thus to approximate the function $g(x)$, we can use the following $\hat{P}$:

$$\hat{P}^+(a) = \frac{2}{\pi}\left(\frac{g(-1)+g(1)}{2} + g'(-1)\right)\frac{1}{\sqrt{1+a}} + \int_0^{a+1} g''(a-t)\frac{2}{\pi\sqrt{t}}\,dt \tag{7.56}$$

$$\hat{P}^-(a) = \frac{2}{\pi}\left(\frac{g(-1)+g(1)}{2} - g'(1)\right)\frac{1}{\sqrt{1+a}} + \int_0^{a+1} g''(-t-a)\frac{2}{\pi\sqrt{t}}\,dt \tag{7.57}$$

To remove the linear error, we used a linear combination of $1+x$ and $1-x$. This choice is not unique. For example, we could have used a linear combination

$$g_\pm(x) = \frac{(1 \pm x)^2}{1+x^2} \tag{7.58}$$

to eliminate x term and a separate linear combination of

$$g_\pm(x) = 1 \pm x^2 \tag{7.59}$$

to eliminate the constant term. The reason equations (7.58)-(7.59) are not used for type-I neurons is due to complexity in the associated weighted decoders. However, these forms are simpler for Type-II firing rates and hence they are used there.

Type-II Weighted Decoders

The same process can be carried out with the type-II firing rate form, which yields the operators:

$$L_{\pm}(\hat{P}^{\pm}) = \int_{-1}^{\pm x} \hat{P}^{\pm}(a)(\pm x - a + c)\, da \tag{7.60}$$

with the resulting values for $\hat{P}^{\pm}$ being

$$\hat{P}^{\pm}(a) = 2A + 2Ba\frac{(a^2-3)}{(a^2+1)^3} + g''(\pm a) \tag{7.61}$$

for $c = 0$. For $c > 0$, the inversion for functions that vanish to second order at $x = \pm 1$ yields

$$\hat{P}^{\pm}(a) = \frac{1}{c}\int_0^{a+1} \exp\left(-\frac{t}{c}\right) g''(\pm(a-t))\, dt \tag{7.62}$$

One can use the convolution (7.62) to compute the weighted decoders for

$$\hat{g}^{\pm}(x) = \frac{(1 \pm x)^2}{1 + x^2}.$$

The solution is lengthy and thus we do not include it here.

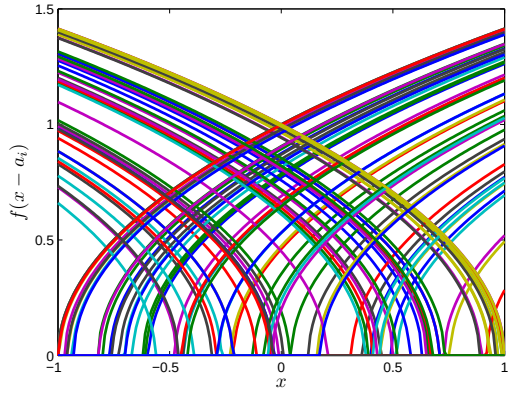
To demonstrate, we have approximated the function $g(x) = \sin(2\pi x)$ using 1,000 type I tuning curves, as shown in figure 7.4. An important thing to notice is the linearity in equations (7.56)-(7.57) and (7.39) in the target function g and thus linearity for the scale-invariant decoders $\gamma(a)$. Furthermore, due to the fact we have considered $\hat{P}^{\pm}(a)$ to be separate for the ON and OFF populations, our operator for determining $g(x)$ is $\hat{g}(x) = M(P^+, P^-) = L_+(P^+) + L_-(P^-)$. However, while our range in L_+ and L_- was constrained, it was not constrained enough to provide a unique solution to $M(P^+, P^-) = g(x)$. In particular, if we consider any function $\epsilon(x)$ that lies in both admissable spaces (vanishes to first orders at $x = 1$ and $x = -1$), then $\epsilon(x)$ can be represented by both populations with $\hat{P}^{\pm}_{\epsilon}(a)$ and thus

$$\tilde{P}^{+}_{g(x)+\epsilon(x)}(a) = \hat{P}^{+}_{g(x)}(a) + \hat{P}^{+}_{\epsilon(x)}(a) \tag{7.63}$$

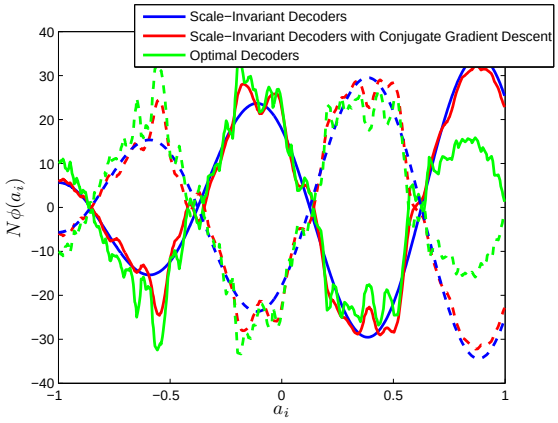
$$\tilde{P}^{-}_{g(x)-\epsilon(x)}(a) = \hat{P}^{-}_{g(x)}(a) - \hat{P}^{-}_{\epsilon(x)}(a) \tag{7.64}$$

are also valid solutions to $M(P^+, P^-) = g(x)$. One can interpret this as a degree of freedom in terms of the decoders (and thus the synaptic weights). For example, we can use $\epsilon(x)$ to minimize the expected squared error however this is a substantially more complicated process that is outside of the scope of this thesis. We will assume that $\epsilon(x) = 0$ unless otherwise specfied.

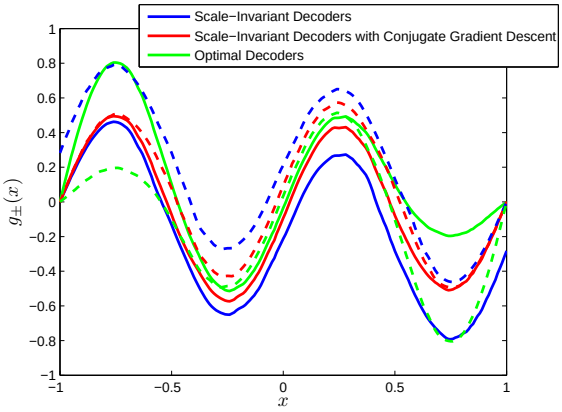
We should note that it is possible to numerically invert the Laplace transforms resulting from the derivation process for the other firing rate curves. However, for now we will primarily work with the type I and type II curves. Our numerics will also primarily consist of networks of theta neurons (type-I).



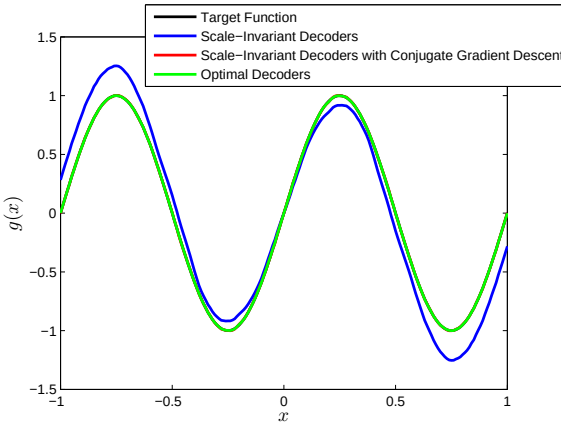
(a) Sample Tuning Curves



(b) On Neuron Population Decoder Surface



(c) Off Neuron Population Decoder Surface



(d) $g(x) = \sin(2\pi x)$ and its approximants

Figure 7.4: Comparison of the scale-invariant decoders and the optimal decoders. The function $g(x) = \sin(2\pi x)$ is approximated using populations of $N = 10^3$ ON and OFF neurons. Shown in figure 7.4(a) are 100 randomly drawn tuning curves from the population. Shown in figure 7.4(b) are the scale invariant decoders (blue), the scale invariant decoders with conjugate gradient descent fine-tuning (red) and the optimal decoders multiplied by N (green) for both the ON (solid) and OFF (dashed) groups of neurons. The different decoders correspond to different $g_{\pm}(x)$, shown in figure 7.4(c), which when summed together yield approximations to $g(x) = \sin(2\pi x)$ in figure 7.4(d). The conjugate gradient descent improves the approximation by 1-2 orders of magnitude (from 10^{-2} to 10^{-4}) while still maintaining a tight correlation with the scale-invariant decoders ($p > 0.95$). With a regularization parameter of $\lambda = 0.05$, the optimal decoders have an error on the order of 10^{-6}

7.2.2 Convergence Rate for Single-Variable Functions

With our decoder surfaces in hand, we can now proceed to determine the various convergence properties in the limit as $N \to \infty$. In particular, we have the following:

$$\begin{aligned} E((g(x) - \hat{g}_N(x))^2) &= \frac{1}{N}\left(\int_{-1}^{x} \gamma^+(a)^2 \rho_a(a) f(x-a)^2\, da - g_+(x)^2\right) \\ &+ \frac{1}{N}\left(\int_{-1}^{-x} \gamma^-(a)^2 \rho_a(a) f(-x-a)^2\, da - g_-(x)^2\right) \end{aligned} \tag{7.65}$$

which immediately implies that our approximant $\hat{g}_N(x)$ converges in mean-square to $g(x)$ pointwise in x provided that:

$$\int_{-1}^{1} \gamma^{\pm}(a)^2 \rho_a(a)\, da = \int_{-1}^{1} \frac{\hat{P}^{\pm}(a)^2}{\rho_a(a)}\, da < \infty \tag{7.66}$$

Letting $\gamma_N = (\gamma(a_i)/N, \ldots \gamma(a_N)/N)$, then we can also consider how the distribution of the quadratic cost function $C(\gamma_N)$ scales as $N \to \infty$; from equation (7.2):

$$C(\gamma_N) = \int_X \left(\frac{1}{N}\sum_{i=1}^{N} \gamma^{e_i}(a_i) f(e_i x - a_i) - g(x)\right)^2 dx + \frac{\lambda}{N^2}\sum_{i=1}^{N} \gamma(a_i)^2 \tag{7.67}$$

where e_i is the symbolic variable $\pm$ denoting the identity of a neuron as OFF/ON. One can show that provided that equation (7.66) holds, in addition to $\gamma(a)$ being a bounded function on $[-1, 1]$, then we have the following

$$E(C(\gamma_N)) \leq O(N^{-1}), \quad E((C(\gamma_N) - E(C(\gamma_N)))^2) \leq O(N^{-2}) \tag{7.68}$$

which is shown in the appendix. This implies that as $N \to \infty$, then $C(\gamma_N) \to 0$ in a mean-square sense. As the cost function is strictly positive, then we can interpret this as γ_N minimizing the cost asymptotically. Note that $\gamma(a)$ being bounded implies that the weights are bounded. Moreover, even if $\hat{P}(a)$ is unbounded, $\rho_a(a)$ can be selected such that $\gamma(a)$ is bounded.

For other neuronal models, one can merely use the maximum firing rates and intercepts to approximate their tuning curves with the type I standard form or the linear firing rate tuning curves. This will yield scale-invariant decoders that can be used on the tuning curves for the actual neuronal model with some degree of error. Additionally, gradient descent algorithms can be used to refine the scale-invariant decoders and that take into account the systematic error in using the type-I/type-II tuning curve approximation.

7.2.3 The Relationship Between the Neural Engineering Framework and Mean-Field Analysis

Prior to deriving the decoder surface, it is important to take stock and determine how our work so far related to previous chapters and mean-field analysis. By using a scale-invariant

decoder, the network equations can be rewritten as

$$\dot{v}_i = F(v_i) + \alpha_i \left(\frac{1}{N} \sum_{i=1}^{N} s_i \right) + \beta_i = F(v) + \alpha_i \langle s \rangle + \beta_i \tag{7.69}$$

$$s_i' = -\frac{s_i}{\tau_s} + \frac{1}{\tau_s} \sum_{t<t_{i,k}} \gamma(\alpha_i, \beta_i) \delta(t - t_{i,k}) \tag{7.70}$$

which in the limit that $N \to \infty$, yields the following population density equation

$$\frac{\partial \rho(v, \alpha, \beta, s, t)}{\partial t} = -\frac{\partial}{\partial v} \left(J^V(\alpha, \beta, s, v, t) \right) - \frac{\partial}{\partial s} \left(J^S(\alpha, \beta, s, v, t) \right) \tag{7.71}$$

$$J^V(\alpha, \beta, s, v, t) = \rho(v, s, \alpha, \beta, t)(F(v) - \alpha \langle s \rangle + \beta) \tag{7.72}$$

$$J^S(\alpha, \beta, s, v, t) = -\rho(v, s, \alpha, \beta, t) \left(\frac{s}{\tau_s} \right) \tag{7.73}$$

$$J^V(v_{peak}, \alpha, \beta, s, t) = J^V(v_{reset}, \alpha, \beta, s + \gamma(\alpha, \beta), t) \tag{7.74}$$

Integration of (7.71) with respect to s, while conditioning the density as

$$\rho_S(s|v, \alpha, \beta, t) \rho_V(v, \alpha, \beta, t)$$

In addition to application of the moment closure $\langle s|v, \alpha, \beta \rangle = \langle s \rangle$ yields the following system

$$\frac{\partial \rho_{V\alpha\beta}(v, \alpha, \beta, t)}{\partial t} = -\frac{\partial}{\partial v} \left(\rho_{V\alpha\beta}(v, \alpha, \beta, t)(F(v) - \alpha \langle s \rangle + \beta) \right) \tag{7.75}$$

$$= -\frac{\partial}{\partial v} J(v, \alpha, \beta, \langle s \rangle, t) \tag{7.76}$$

$$\langle s \rangle' = -\frac{\langle s \rangle}{\tau_s} + \frac{1}{\tau_s} \int_\alpha \int_\beta \gamma(\alpha, \beta) \tilde{J}(v_{peak}, \alpha, \beta, \langle s \rangle) \, d\alpha d\beta \tag{7.77}$$

where the flux has been redefined to

$$\tilde{J}(v, \alpha, \beta, \langle s \rangle, t) = \rho_{V\alpha\beta}(v, \alpha, \beta, t)(F(v) - \alpha \langle s \rangle + \beta)$$

The assumptions used to derive MFII from Chapter 3 immediately yields the following

$$\bar{J}(\alpha, \beta, \langle s \rangle) = \rho_{\alpha\beta}(\alpha, \beta) \left(\int_{v_{reset}}^{v_{peak}} \frac{dv}{F(v) + \alpha \langle s \rangle + \beta} \right)^{-1} \tag{7.78}$$

$$= \rho_{\alpha\beta}(\alpha, \beta) f(\alpha \langle s \rangle + \beta) \tag{7.79}$$

$$\langle s \rangle' = -\frac{\langle s \rangle}{\tau_s} + \frac{1}{\tau_s} \int_\alpha \int_\beta \gamma(\alpha, \beta) \rho_{\alpha\beta}(\alpha, \beta) f(\alpha \langle s \rangle + \beta) \, d\alpha d\beta \tag{7.80}$$

After dropping the brackets on $\langle s \rangle$, the resulting equation is precisely the resulting NEF equation (7.24) in the large network limit. However, with this perspective, even more information is obtained. In particular, we also know the voltage density:

$$\rho(v, \langle s \rangle) = \int_{\alpha,\beta} \frac{\bar{J}(\alpha, \beta, \langle s \rangle)}{F(v) + \alpha \langle s \rangle + \beta} \, d\alpha d\beta = \int_\alpha \int_\beta \frac{\rho_{\alpha\beta}(\alpha, \beta) f(\alpha \langle s \rangle + \beta)}{F(v) + \alpha \langle s \rangle + \beta} \, d\alpha d\beta \tag{7.81}$$

assuming that all the neurons in the population are in the firing regime. More complicated formulas hold for other regimes. Thus, with the same derivation procedure as MFII, one can obtain the NEF system of equations with a scale-invariant decoder.

7.2.4 Multivariable Functions

It is clear that in the preceding section, one could approximate any arbitrary single variable function using scale invariant decoders. The same can be said about multi-variable functions. We will first introduce linear encoding for multi-variable inputs. In the Neural Engineering Framework, it is assumed that the current input into each neuron takes the form:

$$I_i = \alpha_i \langle \boldsymbol{e}_i, \boldsymbol{x} \rangle + \beta_i$$

where $\boldsymbol{e}$ is the encoding unit vector that lies on the n-dimensional unit sphere and $\boldsymbol{x}$ lies in the interior; $\langle \boldsymbol{e}, \boldsymbol{x} \rangle$ is the standard dot-product. The maximum firing rate occurs when $\boldsymbol{x}$ and $\boldsymbol{e}$ are colinear. The vector $\boldsymbol{e}$ is also referred to as the preferred orientation vector. In this case, there are no ON and OFF neurons as they are effectively taken care of by the angle in between $\boldsymbol{x}$ and $\boldsymbol{e}$. Note that because the unit sphere in one-dimension is merely ± 1, we have a direct correspondence with the e_i from the single variable analysis in the previous section.

Once again, we can non-dimensionalize:

$$f(\alpha_i x + \beta_i) \quad = \quad f\left(\frac{f^{-1}(r_i^{max})}{1 - a_i} (\langle \boldsymbol{e}, \boldsymbol{x} \rangle - a_i) \right).$$

If $\boldsymbol{x}$ and $\boldsymbol{e}$ are colinear, then the maximum firing rate occurs when $\boldsymbol{e} = \boldsymbol{x}$, and the firing rate is zero when $\langle \boldsymbol{e}_i, \boldsymbol{x} \rangle = a_i$, the equation for the hyperplane with normal vector $\boldsymbol{e}$. To simplify the situation, we will again assume that $f^{-1}(r_i^{max}) = 1 - a_i$, to remove this term. As before, for type-I firing rates this occurs without any loss of generality.

While it may seem like this setup complicated matters somewhat, in the limit that $N \to \infty$, the end result is simpler than the single variable case as we can make use of the orthogonality of the trigonometric functions to derive an appropriate basis to basis mapping. Consider suboptimal decoders of the form $\phi = \frac{\gamma(\boldsymbol{e},a)}{N} = \frac{\gamma_e(\boldsymbol{e})\gamma_a(a)}{N}$ (some functions are resolvable without the separability assumption; see appendix). For a separable decoder we have:

$$\hat{g}_N(\boldsymbol{x}) = \frac{1}{N} \sum_{i=1}^{N} \gamma_e(\boldsymbol{e}_i)\gamma_a(a_i) f(\langle \boldsymbol{e}_i, \boldsymbol{x} \rangle - a_i)$$

which in the large network limit becomes:

$$\hat{g}_N(\boldsymbol{x}) \quad = \quad \int_{\|\boldsymbol{e}\|=1} \int_{-1}^{\langle \boldsymbol{e}, \boldsymbol{x} \rangle} \gamma_e(\boldsymbol{e})\gamma_a(a) f(\langle \boldsymbol{e}, \boldsymbol{x} \rangle - a) \rho_e(\boldsymbol{e}) \rho_a(\boldsymbol{a}), da \, d\boldsymbol{e} \tag{7.82}$$

$$\quad = \quad \int_{\|\boldsymbol{e}\|=1} \int_{-1}^{\langle \boldsymbol{e}, \boldsymbol{x} \rangle} \hat{P}_e(\boldsymbol{e}) \hat{P}_a(a) f(\langle \boldsymbol{e}, \boldsymbol{x} \rangle - a) \, da \, d\boldsymbol{e} \tag{7.83}$$

$\hat{g}_N(x)$	$\hat{P}(\theta)$	n
1	$\frac{1}{2\pi}$	1
x	$\frac{\cos(\theta)}{\pi}$	1
y	$\frac{\sin(\theta)}{\pi}$	1
x^2+y^2+2	$\frac{1}{\pi}$	2
$\frac{1}{2}(x^2-y^2)$	$\frac{sin(2\theta)}{\pi}$	2
xy	$\frac{\cos(2\theta)}{\pi}$	2

Table 7.1: The basis-to-basis mapping for a polar coordinate system for the first $n = 1, 2$

From our previous work, we know that the weighted decoder $\hat{P}_a(a)$ can be chosen such that:

$$\int_{-1}^{\langle \boldsymbol{e}, \boldsymbol{x} \rangle} \hat{P}_a(a) f(\langle \boldsymbol{e}, \boldsymbol{x} \rangle - a)\, da = (\langle \boldsymbol{e}, \boldsymbol{x} \rangle + 1)^n \tag{7.84}$$

The specific form of $\hat{P}_a(a)$ that performs this transformation varies from neural model to neural model. For the type I/type II firing rate, it is given by a recurrence relationship in terms of the binomial exponent n and is included in the appendix. With $\hat{P}_a(a)$ determined, the decoders $\hat{P}_e(\boldsymbol{e})$ are characterized by the integral equation:

$$\hat{g}_N(x) = \int_{\|\boldsymbol{e}\|=1} \hat{P}_e(\boldsymbol{e})(\langle \boldsymbol{e}, \boldsymbol{x} \rangle + 1)^n \, d\boldsymbol{e}$$

To proceed further, we will exploit the orthogonality of the Fourier series in a hyper-spherical coordinate system. For example, in two-dimensions we have:

$$\hat{g}_N(x) = \int_0^{2\pi} \hat{P}_\theta(\theta)(\cos(\theta)x + \sin(\theta)y + 1)^n \, d\theta \tag{7.85}$$

The second term in the integrand is a polynomial in $\cos(\theta)$ and $\sin(\theta)$. By DeMoivres formula, this can be expressed as a sum of a Fourier series with coefficients that depend on x and y where the series contains no $\cos(m\theta)$ or $\sin(m\theta)$ for $m > n$. Thus, we can extract out polynomial basis functions in the x and y using $\hat{P}(\theta) = \cos(m\theta)\sin(k\theta)$ for $m, k < n$. For example, the first few $\hat{P}(\theta)$ and the corresponding $\hat{g}_N(x)$ are shown in table 7.1.

As we can obtain a polynomial basis where the maximum polynomial power is arbitrary, we can approximate any arbitrary integrable function. In general, one uses a sequence of trigonometric bases in the heterogeneous space to yield a polynomial basis in the function approximation space. One may wonder if the non-uniqueness in the scalar case was due to the peculiarities of the unit sphere in 1-dimension (an isolated pair of points). This turns out not be the case. A counter example is shown in the appendix.

Convergence Rate for Multivariable Functions

One can again determine the order of convergence for the scale-invariant decoders for a multi-variable function. An application of the central limit theorem yields:

$$\begin{aligned} E(\hat{g}_N(\boldsymbol{x})) &= E\left(\gamma(\boldsymbol{e}_i)\gamma(\boldsymbol{a}_i)f(\langle \boldsymbol{e}_i, \boldsymbol{x}\rangle - a_i)\right) = g(x) && (7.86) \\ E((\hat{g}_N(x) - E(\hat{g}_N(x)))^2) &= E((\hat{g}_N(x) - g(x))^2) \\ &= \frac{1}{N}\left[E\left(\gamma(\boldsymbol{e}_i)^2\gamma(a_i)^2 f^2(\langle \boldsymbol{e}_i, \boldsymbol{x} - a_i)\right) - g(x)^2\right] && (7.87) \end{aligned}$$

and thus the expected square error converges like $1/N$ implying that $\hat{g}_N(x)$ converges to $g(x)$ in mean-squared. As the convergence rate is somewhat slow, it is natural to ask whether or not it is possible to improve the the expected squared error.

While there are many analytical paths one may take, we leave these approaches for future work. We will primarily use gradient descent variants that do not require computing the Hessian. If we knew the Hessian, then for a finite network we could immediately solve the system of equations (7.3)-(7.5) as the problem is entirely quadratic and can be resolved numerically with the Hessian matrix. However, solving the quadratic problem with the Hessian requires large matrix inversion, and this is simply not feasible for large networks. Thus, we can use Hessian-free gradient descent methods. For example, one can use various conjugate gradient type algorithms to improve the expected squared error significantly with only a few iterations, and no large matrix inversion. Additionally, one can use the methods of stochastic gradient descent, such as weight perturbation, and node perturbation [188]. We will primarily use conjugate-gradient descent implemented with the PCG function in matlab. The crucial thing about these approaches is that we can obtain substantial improvements to the expected squared error with only slight perturbations to the scale-invariant decoders, as we shall see when we look at specific examples.

7.3 Using the Decoders to Simulate Networks with Arbitrary Dynamics

To simulate networks with arbitrary dynamics, we can use the decoders derived in the previous sections along with neurons that correspond to the appropriate firing rates [52]. Suppose the variable $\boldsymbol{s}(t)$ represents a vector of decoded firing rates given by the following

equation:

$$s_i(t)' = -\frac{s_i}{\tau_s} + \frac{1}{\tau_s}\sum_{j=1}^{N}\phi_{ji}\sum_{t_{j,k}<t}\delta(t-t_{j,k}) \tag{7.88}$$

$$\boldsymbol{s} = -\frac{\boldsymbol{s}}{\tau_s} + \frac{1}{\tau_s}\sum_{j=1}^{N}\boldsymbol{\phi}_j\sum_{t_{j,k}<t}\delta(t-t_{j,k}) \tag{7.89}$$

$$\approx -\frac{\boldsymbol{s}}{\tau_s} + \frac{1}{\tau_s}\sum_{j=1}^{N_i}\boldsymbol{\phi}_j f(\langle \boldsymbol{e}, \boldsymbol{s}\rangle - a_i) \tag{7.90}$$

Where $\boldsymbol{\phi}_j$ is the decoder for the jth neuron. Equation (7.90) is referred to as the rate equation while equation (7.88) is the equation for s_i under neuronal spiking with a simulated spiking neuronal network. We will illustrate both systems in the following examples. Note that the rate equation is only predictive of the spiking equation under asynchronous spiking. When the synaptic time constant is too low, the neurons all synchronize, and the synchronous spiking destroys the dynamics predicted by the rate equation. The time constant used will be 50 ms unless otherwise stated.

Using the same procedure as before, by integrating the spiking equation for $\boldsymbol{s}(t)$ explicitly, one can derive the NEF equation for the synaptic weights:

$$\omega_{ij} = \alpha_i\langle \boldsymbol{e}_i, \boldsymbol{\phi}_j\rangle \tag{7.91}$$

where ω_{ij} is the synaptic weight for the post-synaptic neuron i and the presynaptic neuron j in a recurrent neuronal network and $\boldsymbol{\phi}_j$ is the scale invariant decoder for the function

$$G(\boldsymbol{s}) + \frac{\boldsymbol{s}}{\tau_s}$$

For a scale invariant decoder, this yields the following synaptic weight:

$$\omega_{ij} = \frac{1}{N\rho(a_j)\rho_{\boldsymbol{e}}(\boldsymbol{e}_j)}\frac{f^{-1}(r_i^{max})}{1-a_i}\left\langle \boldsymbol{e}_i, \hat{P}_{\boldsymbol{e}}(\boldsymbol{e}_j)\hat{P}_a(a_j)\right\rangle \tag{7.92}$$

Note that $\omega_{ij} = \Omega(r_i^{max}, a_i, a_j, \boldsymbol{e}_i\boldsymbol{e}_j)$, is merely a function of random variables for the presynaptic and post-synaptic neurons. Thus, instead of thinking of the weights as a matrix of numerical values, or as a direct graph, one may think of the weights as defining a hypersurface in a higher dimensional space. For example, the formula (7.92) describes a hypersurface with $2m+3$ dimensions where m is the dimension of the dynamics the network simulates.

In the following examples, we will generate networks with these analytically determined weights using scale-invariant decoders that display the prescribed dynamics. Additionally, we will assume that

$$r_i^{max} = M\sqrt{1-a_i} \tag{7.93}$$

$$\rho_a(a) = \frac{1}{2\sqrt{2}\sqrt{1+a}} \tag{7.94}$$

The variable M controls the maximum firing rate of the neurons, with the range of maximum firing rates being between $[0, \sqrt{2}M]$. We take M to be 60 Hz for all subsequent numerics, unless otherwise specified. Note that we need α_i to compute the weights. The α_i differs depending on whether or not we are simulating a scalar system or a multi-variable system. For a scalar system, $\alpha_i = M^2 e_i$ where $e_i = 1$ for ON neurons and -1 for OFF neurons. For a vector, $\alpha_i = M^2$. For multivariable dynamics we will assume uniform distributions in the hyperspherical coordinate systems. With the former assumption, the tuning curves for the neurons simply become $M\sqrt{\langle \boldsymbol{e}, \boldsymbol{x}\rangle - a}$ where the M can be absorbed into the decoder. We will generate networks of spiking theta neurons that simulate a neural integrator, a Van der Pol Oscillator, and the Lorenz system.

7.3.1 Example 1: Neural Integrator

A neural integrator is a recursively coupled neural network that integrates an incoming signal, $u(t)$. The coupling variable $s(t)$, will have dynamics given by

$$s'(t) = u(t) = -\frac{s}{\tau_s} + \frac{1}{\tau_s}\sum_{i=1}^{N} \phi_i \sqrt{e_i(s(t) + \tau_s u(t)) - a_i} \tag{7.95}$$

where $e_i = 1$ if neuron i is an ON neuron and -1 for OFF neurons. Note that we have scaled up $u(t)$ by τ_s as this allows us to write:

$$\tau_s u + s = \sum_{i=1}^{N} \phi_i \sqrt{e_i(s + \tau_s u) - a_i}. \tag{7.96}$$

Letting $z = \tau_s u + s$, then we require the scale invariant decoders such that $\hat{g}(z) = \sum_{i=1}^{N} \phi_i \sqrt{e_i(s + \tau_s u) - a_i} \approx z$. A set of $g_{\pm}(x)$ and the corresponding scale-invariant decoders is given by:

$$\hat{g}^{+}(x) = \frac{1}{2}(1 + x), \quad \hat{g}^{-}(x) = -\frac{1}{2}(1 - x) \tag{7.97}$$

$$\hat{P}^{+}(a) = \frac{2}{\pi}\frac{1}{\sqrt{1 + a}M}, \quad \hat{P}^{-}(a) = -\frac{2}{\pi}\frac{1}{\sqrt{1 + a}M} \tag{7.98}$$

$$\phi_i = \frac{\gamma^{\pm}(a)}{N} = e_i \frac{4\sqrt{2}}{NM\pi} \tag{7.99}$$

which yields $\hat{g}(z) = z$. From formula (7.17) for α_i and (7.94) for the density of a_i, and $\alpha_i = e_i M^2$, we have the neuronal weight

$$\omega_{ij} = \alpha_i \phi_j = e_i e_j \frac{4\sqrt{2}M}{N\pi}. \tag{7.100}$$

All the synaptic weights here are given by $4\sqrt{2}M/(N\pi)$ for ON/ON and OFF/OFF connections and $-4\sqrt{2}M/(N\pi)$ for ON/OFF and OFF/ON connections. Now, we will exploit symmetry and non-uniqueness to generate two more neuronal integrators with the

same initial distributions of heterogeneity $\rho_a(a)$. In particular, consider the function $\epsilon(x) = (1-x^2)^2$, this function lies in both function spaces for $g^{\pm}(x)$ as it vanishes to second order at both boundaries. Additionally, it can be computed using the following scale-invariant decoders:

$$\hat{P}^{\pm}_{\epsilon(x)}(a) = \frac{32\pi}{5M}(4a^2 - 2a - 1)\sqrt{a+1} \tag{7.101}$$

$$\phi_i = \frac{\gamma(a_i)}{N} = e_i \frac{64\sqrt{2}}{5M\pi N}(1+a_i)(4a_i^2 - 2a_i - 1) \tag{7.102}$$

which implies the following decoders for the ON/OFF population still give us $g(x) = x$

$$\phi_i = e_i \frac{4\sqrt{2}}{NM\pi} + \frac{64\sqrt{2}}{5M\pi N}(1+a_i)(4a_i^2 - 2a_i - 1) \tag{7.103}$$

which yields the weight matrix

$$\omega_{ij} = \alpha_i \phi_j = e_i e_j \frac{4\sqrt{2}M}{N\pi} + e_i e_j \frac{64M\sqrt{2}}{5\pi N}(1+a_j)(4a_j^2 - 2a_j - 1) \tag{7.104}$$

Additionally, we can exploit symmetry to derive yet another weight matrix with precisely the same network of neurons:

$$g_+(x) = \frac{1}{4}(1+x)^2, \quad g_-(x) = -\frac{1}{4}(1-x)^2 \tag{7.105}$$

$$\hat{P}^+(a) = \frac{1}{M\pi}\sqrt{1+a}, \quad \hat{P}^-(a) = -\frac{1}{M\pi}\sqrt{1+a} \tag{7.106}$$

$$\phi_i = \frac{\gamma(a_i)}{N} = e_i \frac{2\sqrt{2}(1+a_i)}{MN\pi} \tag{7.107}$$

$$\omega_{ij} = e_i e_j \frac{2\sqrt{2}(1+a_i)M}{N\pi} \tag{7.108}$$

and thus, we have the following three separate weight matrices

$$\omega_{ij} = \alpha_i \phi_j = e_i e_j \frac{4\sqrt{2}M}{N\pi} \tag{7.109}$$

$$\omega_{ij} = \alpha_i \phi_j = e_i e_j \frac{4\sqrt{2}M}{N\pi} + e_i e_j \frac{64M\sqrt{2}}{5\pi N}(1+a_j)(4a_j^2 - 2a_j - 1) \tag{7.110}$$

$$\omega_{ij} = e_i e_j \frac{2\sqrt{2}(1+a_j)M}{N\pi} \tag{7.111}$$

for $i, j = 1, 2, \ldots N$ that yield identical macroscopic dynamics from the same network of neurons as $N \to \infty$. Furthermore, while all the weights converge to 0 as $N \to \infty$, the scaled weights $N\omega_{ij}$ do not converge towards one another in the same limit. One important point is that none of the weights necessarily satisfy Dale's principle, that the action of neuron j on all its downstream targets is precisely the same, either excitatory or inhibitory. Or more precisely that

$$\text{sign } \omega_{ij} = \text{sign } \omega_{i'j}$$

for all $i, i', j = 1, 2, \ldots N$. Fortunately however, this issue has already been dealt with in the existing literature [142]. To summarize, one is able to take the weights generated by the NEF solution, and linearly transform them to yield a new network consisting of excitatory and inhibitory neurons (instead of ON/OFF) with weights that respect Dale's principle.

To summarize, the network equations are the following

$$\theta_i' = 1 - \cos(\theta_i) + \pi^2(1 + \cos(\theta_i))(\alpha_i(s + \tau_s u(t)) + \beta_i), \quad i = 1, 2, \ldots N \tag{7.112}$$

$$s' = -\frac{s}{\tau_s} + \frac{1}{\tau_s}\sum_{j=1}^{N}\sum_{t<t_{j,k}} \phi_j \delta(t - t_{j,k}) \tag{7.113}$$

where the ϕ_j are determined by either (7.99),(7.102),(7.107) or the optimal decoders. The weights for this particular network are given by equations (7.109),(7.110), or (7.111) or again, the NEF formula for weights with optimal decoders. The signal $u(t)$ is randomly generated and integrated by the network. Note that the factor of π^2 in front of the currents for each neuron removes the π^{-1} factor for the theta neuron firing rate that is typically present.

We have simulated 4 neural networks with 5000 neurons each. All the networks were generated with the same random sample drawn from the distribution (7.94) using the weights given by equations (7.109),(7.110),(7.111), in addition to the weights generated by determining the optimal decoders. The scale-invariant decoders that correspond to (7.109)-(7.111) were put through conjugate gradient fine-tuning with the final decoders being correlated to the initial decoders with a correlation coefficient greater than 0.98 in all cases. This lowered the root-mean-squared-error by 2-3 orders of magnitude with only slight perturbations off the scale-invariant decoding surface in each case. A sample set of tuning curves is shown in Figure 7.5(a) with the $\hat{g}^{\pm}(x)$ that correspond to the four different weight structures in Figure 7.5(b). The neural integrators are shown in figure 7.5(c)-7.5(f). The scale-invariant decoders are fine-tuned with conjugate gradient descent and in all cases the fine-tuned decoders are very highly correlated ($p > 0.98$) with the scale-invariant decoders indicating only small perturbations off the scale-invariant decoder surface with substantial improvements in the root mean-squared error (RMSE) in computing $\hat{g}(z) = z$. The RMSE was $O(10^{-5})$ with conjugate-gradient descent vs $O(10^{-2})$ without. The variable $\lambda = 0.01$ was taken in the conjugate gradient descent fine-tuning.

The synaptic weights that correspond to the integrators are shown in figure 7.6. The neurons have been sorted into ON/OFF populations and increasing a within a sub-population prior to plotting the weight matrix in the left column of figure 7.6. A sub-sample of 20 neurons is also selected (which is identical across the four integrator networks) and their weights are plotted in the right column of figure 7.6. For the optimal decoders, $g^{\pm}(x) \approx \pm(1 \pm x)/2$, which results in a weight structure similar to (7.109). The weights differ substantially however in comparison to equations (7.110) and (7.111) as the $\hat{g}^{\pm}$ differ substantially from $\pm(1 \pm x)/2$

This example illustrates that identical networks of neurons can have identical macroscopic dynamics with vastly different weight matrices. While this is not particularly surprising as going from a microscopic description (the individual weights) to a macroscopic description (the dynamics) of a dynamical system is seldom a unique process, the surprising thing is one can explore this issue analytically. An important point to note is that even though the weight matrices are non-unique, they all have the form

$$\omega_{ij} = \Omega(e_i, e_j, a_j) \tag{7.114}$$

and the weight matrices are nothing more than sample points drawn from different surfaces.

7.3.2 Example 2: Van der Pol Oscillator

The Van der Pol oscillator [180] is given by the dynamical system:

$$\dot{x} = \mu\left(x - \frac{1}{3}x^3 - y\right) = F(x, y) \tag{7.115}$$

$$\dot{y} = \frac{x}{\mu} = G(x) \tag{7.116}$$

Here, we will simulate the oscillator with large networks of neurons using the scale-invariant decoders with conjugate gradient descent fine tuning. As the decoding is linear, then from the above equations we only require the decoders for the functions $f(x, y) = x$, $f(x, y) = y$ and $f(x, y) = x^3$. We use a 2-dimensional spherical coordinate system for the encoding vectors $\boldsymbol{e} = (\cos(\theta_i), \sin(\theta_i))$ and assume uniform distributions in a, the intercepts for the tuning curves and θ. Note that

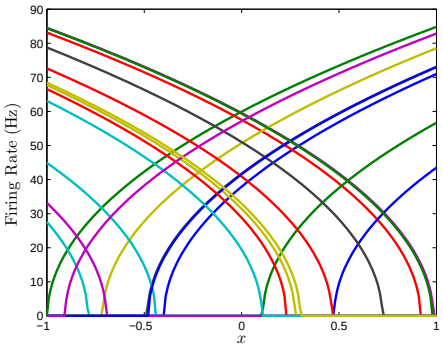
$$\begin{aligned}
x &= \int_0^{2\pi}\int_{-1}^{\cos(\theta)x+\sin(\theta)y} \frac{\cos(\theta)}{\pi}\frac{2}{\pi\sqrt{1+a}}\sqrt{\cos(\theta)x + \sin(\theta)y - a}\, da d\theta \\
y &= \int_0^{2\pi}\int_{-1}^{\cos(\theta)x+\sin(\theta)y} \frac{\sin(\theta)}{\pi}\frac{2}{\pi\sqrt{1+a}}\sqrt{\cos(\theta)x + \sin(\theta)y - a}\, da d\theta \\
x^3 + 3x &= \int_0^{2\pi}\int_{-1}^{\cos(\theta)x+\sin(\theta)y} \frac{\cos(3\theta) + \cos(\theta)}{\pi}\frac{16(\sqrt{1+a})^3}{\pi}\sqrt{\cos(\theta)x + \sin(\theta)y - a}\, da d\theta \\
&.
\end{aligned}$$

This immediately leads to following scale-invariant decoders for the sub-functions x, y, and x^3:

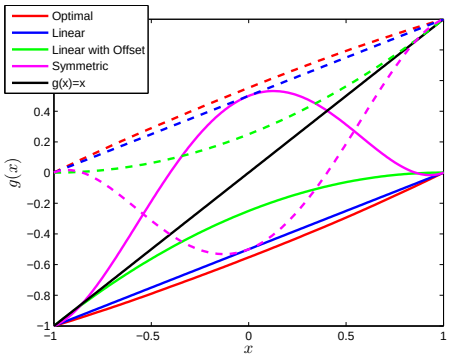
$$\phi_i^x = \frac{4\sqrt{2}\cos\theta_i}{MN\pi} \tag{7.117}$$

$$\phi_i^y = \frac{4\sqrt{2}\sin\theta_i}{MN\pi} \tag{7.118}$$

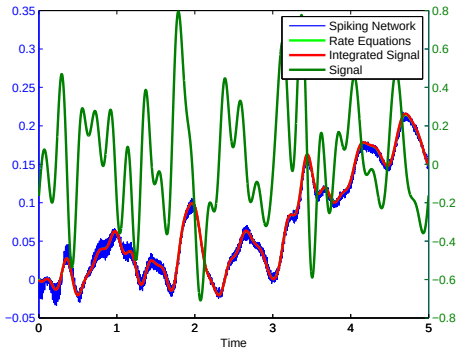
$$\phi_i^{x^3} = 32\sqrt{2}\left(\cos(3\theta_i) + \cos(\theta_i)\right)\frac{(1+a_i)^2}{\pi MN} - 3\phi_i^x \tag{7.119}$$



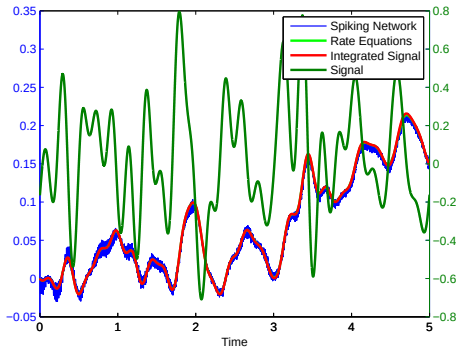
(a) 20 Neuron Subset of the 5000 Neuron Network



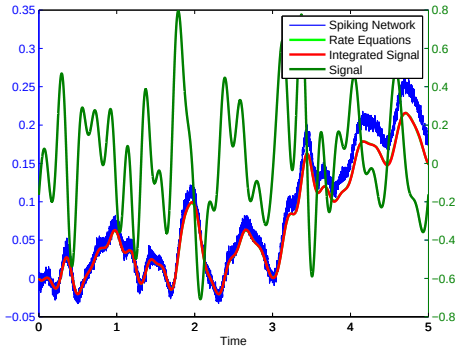
(b) $g_{\pm}(x)$ for the various networks



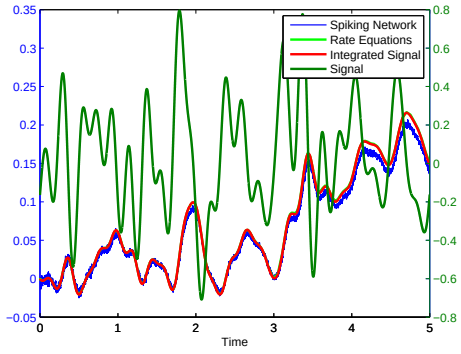
(c) Neural Integrator with optimal decoders and NEF weights



(d) Neural integrator with weights from equation 7.109

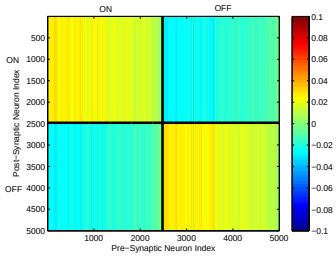


(e) Neural integrator with weights from equation 7.110

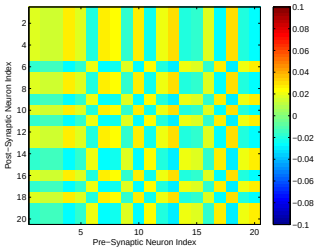


(f) Neural integrator with weights from equation 7.111

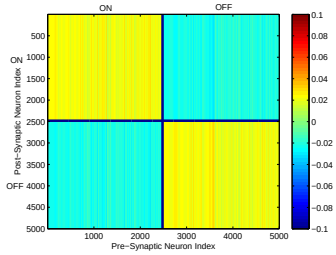
Figure 7.5: Neural integrators generated using the same initial heterogeneous network of $N = 5 \cdot 10^3$ Theta neurons. Shown in figure 7.5(a) is a subset of 20 tuning curves from the network. Shown in figure 7.5(b),7.5(c),7.5(d) are the integrators generated using the decoders from equations (7.99),(7.102) and (7.107) which results in weight matrices (7.109),(7.110), and (7.111). While the integrators in figure 7.5(c) and 7.5(d) perform equally well in integrating the signal, the integrator in figure 7.5(f) has some degree of error in the mean in exchange for reducing the higher order correlations.



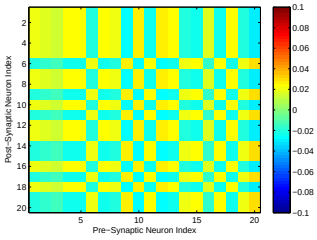
(a) Weight matrix with ϕ^*



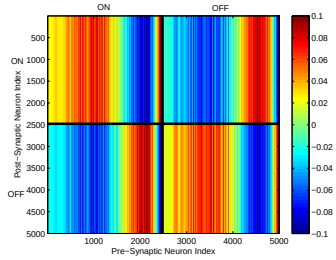
(b) Randomly drawn submatrix



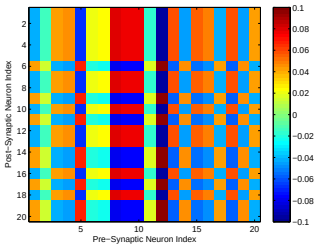
(c) Weight matrix from equation 7.109



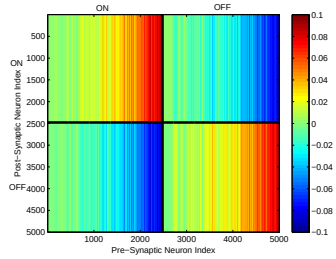
(d) Randomly drawn submatrix



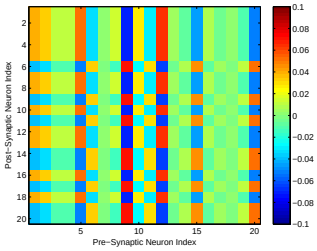
(e) Weight matrix from equation 7.110



(f) Randomly drawn submatrix



(g) Weight matrix from equation 7.111



(h) Randomly drawn submatrix

Figure 7.6: Different connectivity weights can generate identical macroscopic dynamics from identical neuronal populations. Shown above are the weight matrices generated for the neural integrators in figure 7.5. All four weight matrices generate neural integrators with varying levels of performance. The underlying network is unchanged as the neurons have the same distribution of heterogeneity with an identical realization from the distribution. The only difference between the networks is the synaptic weight matrix used to couple the neurons together.

which yields the decoders for $F(x,y)$ and $G(x,y)$:

$$\phi_i^F = \phi_i^x + \tau_s \mu \left(\phi^x - \frac{1}{3}\phi^{x^3} - \phi^y \right) \tag{7.120}$$

$$\phi_i^G = \phi_i^y + \tau_s \frac{\phi^x}{\mu} \tag{7.121}$$

Thus, the weights are given by

$$\omega_{ij} = \Omega(\theta_i, \theta_j, a_j) = M^2 \cos(\theta_i)\phi_j^F + M^2 \sin(\theta_i)\phi_j^G \tag{7.122}$$

To summarize, the network equations are given by the following

$$\theta_i' = 1 - \cos(\theta_i) + \pi^2(1 + \cos(\theta_i))(\alpha_i \langle \boldsymbol{e}_i, \boldsymbol{s} \rangle + \beta_i) \tag{7.123}$$

$$\boldsymbol{s}' = -\frac{\boldsymbol{s}}{\tau_s} + \frac{1}{\tau_s} \sum_{j=1}^{N} \sum_{t<t_{j,k}} \boldsymbol{\phi}_j \delta(t - t_{j,k}) \tag{7.124}$$

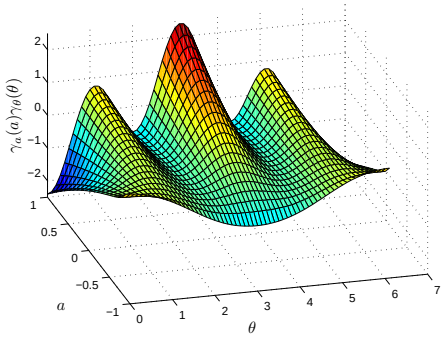
$$\boldsymbol{\phi}_j = \begin{pmatrix} \phi_i^F \\ \phi_i^G \end{pmatrix} \tag{7.125}$$

Equations (7.123)-(7.125) are for the spiking neural network. The rate equations are immediately given by

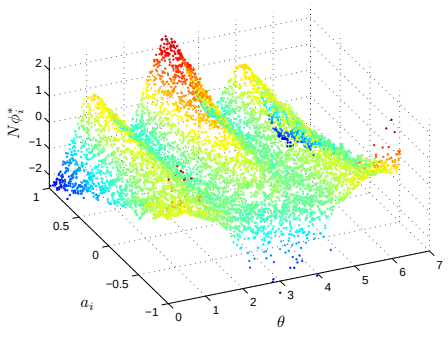
$$\boldsymbol{s}' = -\frac{\boldsymbol{s}}{\tau_s} + \frac{1}{\tau_s} \sum_{j=1}^{N} \boldsymbol{\phi}_j \sqrt{\alpha_i \langle \boldsymbol{e}_i, \boldsymbol{s} \rangle + \beta_i} \tag{7.126}$$

Both the rate equations and the spiking neural network are simulated using the scale invariant decoders to weight the firing rates/spikes. To more explicitly show the effects of the conjugate-gradient descent fine tuning, we have plotted the scale-invariant decoder surfaces $N\phi^F$ and $N\phi^G$ in figure 7.7, in addition to the decoders after conjugate gradient descent fine tuning. The surfaces are used to initialize conjugate gradient descent fine-tuning for a network of 10,000 neurons. If we plot the fine-tuned decoders $N\bar{\phi}_i$ (shown in figure 7.7(b) and 7.7(d), they are slightly perturbed off of these surfaces as shown by the strong linear relationship in the scale-invariant surface and the conjugate gradient descent optimized decoders $N\bar{\phi}_i$ in figure 7.7(e)-7.7(f). The correlation coefficient was greater then 0.95 in both cases, while the mean squared error was reduced from $O(10^{-2})$ to $O(10^{-6})$. The parameter μ for the Van der Pol Oscillator was taken to be 5.

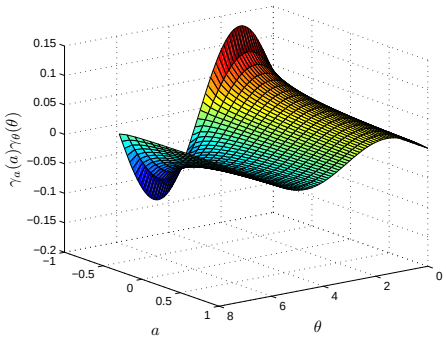
We have simulated the Van der Pol Oscillator, as shown in Figure 7.8 in both the relaxation ($\mu = 5$) and harmonic ($\mu = 0.7$) oscillator regimes. In both cases, we have excellent correspondence with the network and the actual oscillator. The synaptic weight matrices are also shown in figure 7.8 for a subset of neurons. Like the neural integrator, the weights again lie on a surface, only the surface is 3-dimensional.



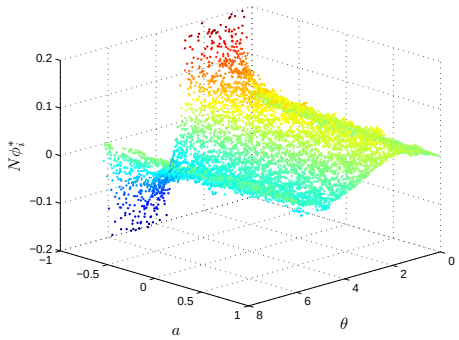
(a) The scale-invariant decoder surface for the x component of the Van der Pol Oscillator



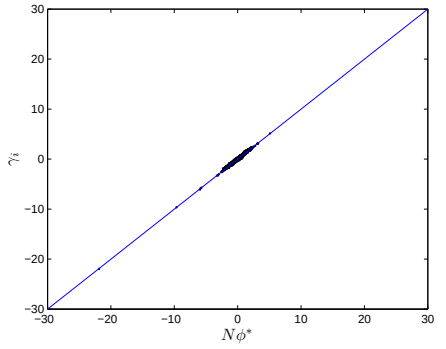
(b) The scale-invariant decoder surface for the x component of the Van der Pol Oscillator



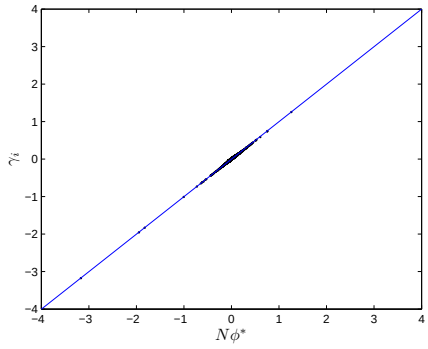
(c) $N\phi_i^*$ for the x component of the Van der Pol Oscillator



(d) $N\phi_i^*$ for the y component of the Van der Pol Oscillator

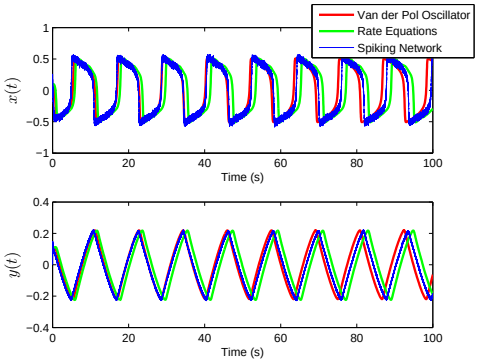


(e) $N\phi_i^*$ vs $\gamma_a(a_i)\gamma_\theta(\theta_i)$ for the x component of the Van der Pol Oscillator

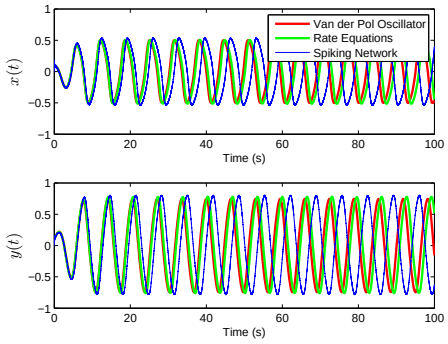


(f) $N\phi_i^*$ vs $\gamma_a(a_i)\gamma_\theta(\theta_i)$ for the x component of the Van der Pol Oscillator

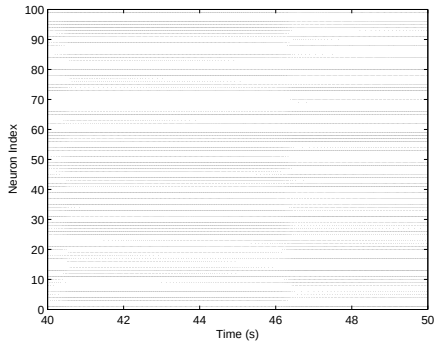
Figure 7.7: The Van der Pol Oscillator is approximated by using a scale-invariant decoder surface. The two functions $F(x, y)$ and $G(x, y)$ that are responsible for the dynamics of the Van der Pol Oscillator have scale-invariant decoder surfaces given by equations (7.120)-(7.121). The equations for the scale-invariant decoder surfaces are plotted in 3D in figures 7.7(a) and 7.7(c), respectively.



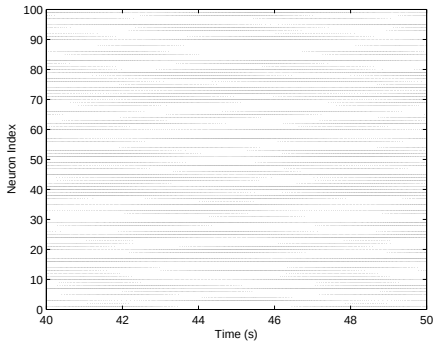
(a) The Van der Pol oscillator with $\mu = 5$



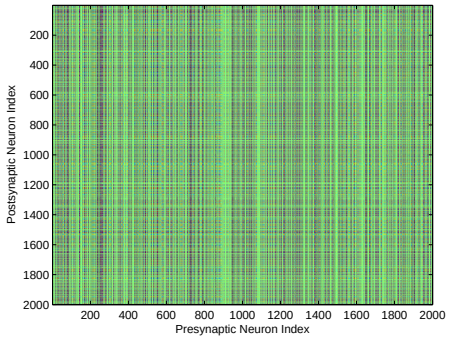
(b) The Van der Pol oscillator with $\mu = 0.7$



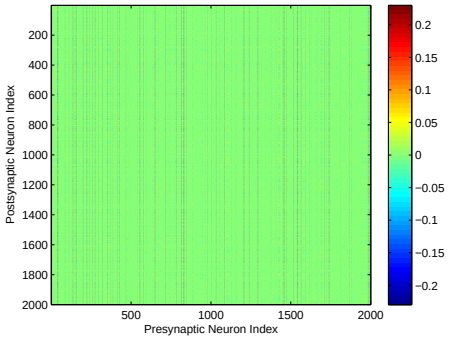
(c) Spike raster plot for 100 randomly selected neurons



(d) Spike raster plot for 100 randomly selected neurons



(e) Weight Matrix for 100 randomly selected neurons



(f) Weight Matrix for 100 randomly selected neurons

Figure 7.8: The Van der Pol oscillator is simulated using a network of 10^4 theta neurons with scale-invariant decoders after conjugate-gradient descent fine tuning in the relaxation oscillator regime ($\mu = 5$, left column) and in the harmonic oscillator regime ($\mu = 0.7$, right column). Shown in the top row is the comparison between the oscillator (red), the rate equations (green), and the spiking network (blue). Shown in the middle is the spike raster plot for a 10 second interval of both networks. The last row consists of a computed weight matrix for 2000 randomly selected neurons in each network

7.3.3 Example 3: Lorenz Attractor

The Lorenz system is given by the equations

$$\begin{aligned} \dot{x} &= \sigma(y-x) = F(x,y) && (7.127)\\ \dot{y} &= x(\rho - z) - y = G(x,y,z) && (7.128)\\ \dot{z} &= xy - \beta z = H(x,y,z) && (7.129) \end{aligned}$$

and is known to exhibit chaotic behavior for specific values of σ, ρ and β [108]. In order to approximate the Lorenz system, we require the decoders for the functions x, y, z, xz and xy. With a 3-dimensional spherical coordinate system, the encoding vectors $\boldsymbol{e}$ are given by $\boldsymbol{e} = (\sin(\theta)\cos(\psi), \sin(\theta)\sin(\psi), \cos(\theta))$ where $\theta \in [0, \pi]$ $\psi \in [0, 2\pi]$. The decoders as a function of (ψ, θ, a) are given by:

$$\begin{aligned} \phi_i^x &= \frac{4\sqrt{2}\cos(\psi_i)}{NM\pi}\\ \phi_i^y &= \frac{4\sqrt{2}\sin(\psi_i)}{NM\pi}\\ \phi_i^z &= \frac{8\sqrt{2}\cos(\theta_i)}{NM\pi}\\ \phi_i^{xz} &= \frac{24\sqrt{2}\cos(\theta_i - \psi_i)(1+a_i)}{2NM} - \frac{3\pi}{4}\phi_i^y\\ \phi_i^{xy} &= \frac{64\sqrt{2}\sin(2\psi)(1+a_i)}{NM\pi} \end{aligned}$$

which yields the decoders for F, G, H as:

$$\begin{aligned} \phi_i^F &= \sigma(\phi_i^y - \phi_i^x) && (7.130)\\ \phi_i^G &= \phi_i^x \rho - \phi_i^{xz} - \phi_i^y && (7.131)\\ \phi_i^H &= \phi_i^{xy} - \beta\phi_i^z && (7.132) \end{aligned}$$

To summarize, the network equations are given by the following

$$\begin{aligned} \theta_i' &= 1 - \cos(\theta_i) + \pi^2(1+\cos(\theta_i))(\alpha_i\langle \boldsymbol{e}_i, \boldsymbol{s}\rangle + \beta_i) && (7.133)\\ \boldsymbol{s}' &= -\frac{\boldsymbol{s}}{\tau_s} + \frac{1}{\tau_s}\sum_{j=1}^{N}\sum_{t<t_{j,k}} \boldsymbol{\phi}_j \delta(t - t_{j,k}) && (7.134)\\ \boldsymbol{\phi}_j &= \begin{pmatrix} \phi_i^F \\ \phi_i^G \\ \phi_i^H \end{pmatrix} && (7.135) \end{aligned}$$

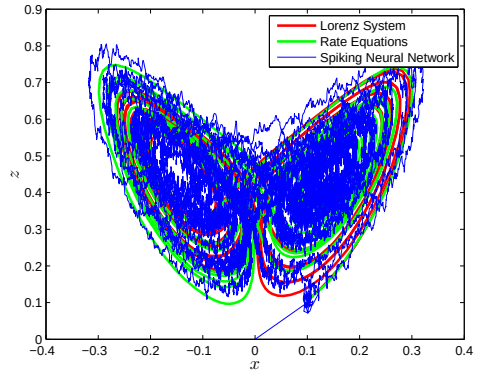
Equations (7.123)-(7.125) are for the spiking neural network. The rate equations are immediately given by

$$\boldsymbol{s}' = -\frac{\boldsymbol{s}}{\tau_s} + \frac{1}{\tau_s}\sum_{j=1}^{N} \boldsymbol{\phi}_j \sqrt{\alpha_i \langle \boldsymbol{e}_i, \boldsymbol{s} \rangle + \beta_i} \tag{7.136}$$

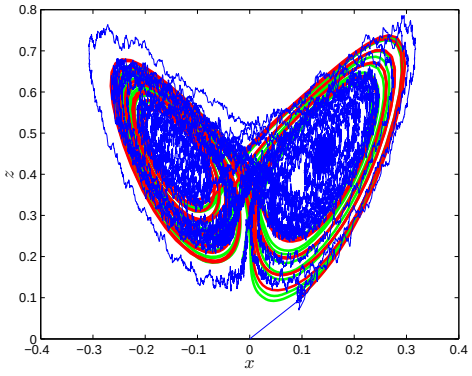
The strange attractor generated by the Lorenz system and the neural rate equations using the decoders from (7.130)-(7.132) are shown in figure 7.9. The chaotic behavior and the strange attractor is preserved when one uses a spiking neuronal network with the decoder weighting the spikes. Note that a great many neurons are required to adequately visualize the strange attractor, however the chaotic behavior is present even for smaller networks. We have also plotted the location of neural spiking with regards to the strange attractor. The neurons tend to spike more in specific regions of the strange attractor in accordance with their preferred orientation vectors and their a_i parameters. The weights are again given by the NEF formula:

$$\omega_{ij} = \Omega(\theta_i, \theta_j, \psi_i, \psi_j, a_j) = M^2 \sin(\theta_i)\cos(\psi_i)\phi_j^F + M^2 \sin(\theta_i)\sin(\psi_i)\phi_j^G + M^2 \cos(\theta_i)\phi_j^H \tag{7.137}$$

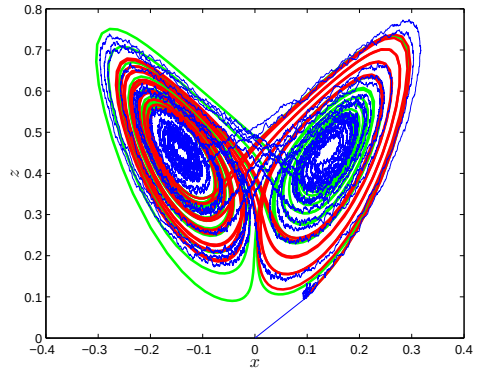
which defines a 5 dimensional surface.



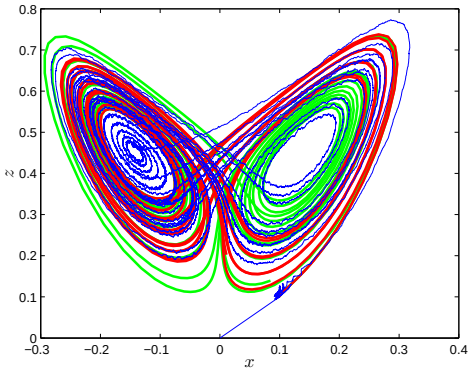
(a) $N = 15,000$



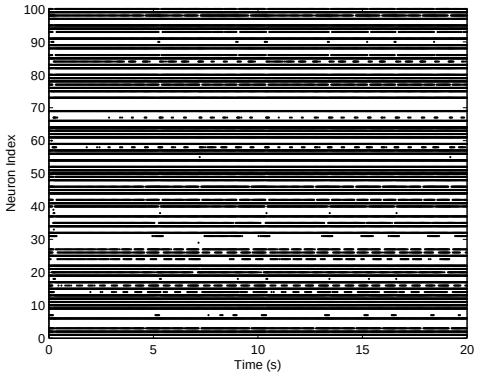
(b) $N = 30,000$



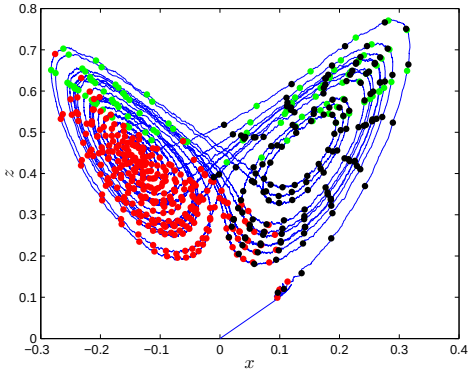
(c) $N = 200,000$



(d) $N = 1,000,000$



(e) Spike Raster Plot



(f) $N = 1,000,000$

Figure 7.9: The Lorenz attractor generated with a spiking neuronal network (blue), the neural rate equations (green) and integrating the Lorenz system (red) with neural networks of increasing size in figures 7.9(a)-7.9(d). The decoders used are scale-invariant with conjugate gradient descent fine-tuning. The synaptic time constant was $\tau_s = 50$ ms. Shown in figure 7.9(e) is the spike raster plot for 100 neurons chosen at random from the $N = 10^6$ neuron simulation. The location of the spikes with regards to the strange attractor for three randomly selected neurons are shown in figure 7.9(f).

Chapter 8

Conclusions, Related Approaches, and Future Directions

In this chapter, we will discuss related approaches to large spiking neural networks, whether it be in mean-field analysis of these networks (section 8.1) or in the synthesis of large spiking networks with specified dynamics (section 8.3). In section 8.2, I will outline some future directions for mean-field analysis. The same will be done for scale-invariant decoding and the Neural Engineering Framework in 8.4.

8.1 Related Mean-Field Approaches

The construction of mean-field systems began with the Wilson-Cowan equations which are discussed in section 8.1.1. These were a set of equations used to phenomenologically model key averages in a network of excitatory and inhibitory neurons [190]. More recently, other approaches have emerged that are more constructive or bottom-up in nature. As in the work done in this chapter, these approaches start with a spiking neural network and consider various kinds of averaging procedures. These will be discussed in section 8.1.2.

8.1.1 The Wilson-Cowan Equations

The Wilson-Cowan equations were originally derived in [190] and describe the macroscopic behavior of a network of excitatory and inhibitory neurons. Wilson and Cowan define the two state variables for their resulting equations as

$$E(t) \quad : \quad \text{The proportion of excitatory cells firing per unit time at the instant } t \tag{8.1}$$
$$I(t) \quad : \quad \text{The proportion of inhibitory cells firing per unit time at the instant } t \tag{8.2}$$

Wilson and Cowan proceed in deriving their system of differential equations by looking at these quantities at time $t + \tau$ for a short time after. They assume that the proportion

of cells that are firing at $t+\tau$ are the proportion that are not refractory due to previous firing multiplied by the proportion of neurons that have received enough excitation to reach threshold. The proportion that is not refractory is immediately given by

$$p_E(t) = 1 - \int_{t-r}^{t} E(t')\,dt', \quad p_I(t) = \quad 1 - \int_{t-r}^{t} I(t')\,dt' \tag{8.3}$$

for each population. Then the authors seek out to determine what proportion of neurons have received enough activation to fire. Assuming that the neurons have thresholds that are heterogeneously distributed with distribution $\rho_\theta(\theta)$ and that all neurons within a population receive the same excitation, then what Wilson and Cowan define as the population response function, $s(x)$ will take the form

$$s(x) = \int_0^x \rho_\theta(\theta)\,d\theta \tag{8.4}$$

Alternatively one could assume that the neurons are homogeneous, but they all receive a different amount of afferent synapses such that the input x is weighted by the afferent synaptic weight w_i for each neuron i. In order for a neuron to be excited, we must have $\theta = w_i x$ which yields another possible population response

$$s(x) = \int_{\theta/x} \rho_w(w)\,dw \tag{8.5}$$

The resulting population response functions in either case are sigmoidal when the distributions are unimodal, which justifies the use of sigmoids in artificial neural network theory. Finally, the authors assume that the average level of excitation is given by

$$x_i(t) \quad = \quad \int_{-\infty}^{t} \alpha(t-t')(c_{Ei}E(t') - c_{Ii}I(t') + P_i(t'))\,dt' \quad i = E, I \tag{8.6}$$

Wilson and Cowan note that the the right hand side of the expressions (8.8)-(8.6) only contains integrations versions of $E(t)$ and $I(t)$, and thus use this as a justification to consider only the course grained versions of $I(t)$ and $E(t)$. In particular, the authors use a kind of averaging to replace the variables $I(t)$ and $E(t)$ with their coarse-grained versions by using the following:

$$\bar{z}(t) = \frac{1}{T}\int_t^{t+T} z(t')\,dt', \tag{8.7}$$

which is a time average. This results in the following system of differential equations

$$\tau\frac{d\bar{E}}{dt} \quad = \quad -\bar{E} + (1 - r\bar{E})s(kc_{EE}\bar{E} - kc_{EI}\bar{I} + kP_E(t)) \tag{8.8}$$

$$\tau\frac{d\bar{I}}{dt} \quad = \quad -\bar{I} + (1 - r\bar{I})s(kc_{IE}\bar{E} - kc_{II}\bar{I} + kP_I(t)) \tag{8.9}$$

which are the classical Wilson-Cowan equations.

The Wilson-Cowan equations have generated a tremendous amount of anaylsis, predictions, and results for computational neuroscience due to their small, coupled autonomous ODE nature. Their simple nature make them tractable to bifurcation analysis both numerically and analytically [40, 78]. While phenomenlogical in nature, they can still be used to predict the behavior of biological neural networks, such as the rat whisker barrel, the somatosensory network that receives inputs from the rat's whiskers [146] or bursting behavior in the chick spinal cord [175]. Additionally, one can easily extend the Wilson-Cowan equations to spatially distributed networks [58, 59].

The mean-field systems derived here are closely related to the Wilson-Cowan approach both in the final underlying structure, and in spirit. In particular, the firing rate that emerges in the population density equations is also commonly referred to as a population activity [61, 70] and is similar definition to the Wilson-Cowan definition of the population response, which is the proportion of of cells in a subpopulation that respond to a given level of excitation per unit time. The population activity is defined as

$$A(t) = \frac{1}{\Delta_t}\frac{1}{N}\left(\sum_{t<t_{j,k}} \delta(t - t_{j,k})\right)$$

and is equal to the flux across the voltage threshold/peak as $\Delta_t \to 0$ and $N \to \infty$.

However, there are a few crucial differences. All the parameters in the original model neurons still appear in the final set of equations as we derive the resulting mean-field equations from a population density intermediate. This even includes heterogeneous parameter distributions or networks with noise. As we have already stated, the assumptions on the derivation of the Wilson-Cowan is that the population response function, $s(x)$ is determined by some sort of heterogeneity in the population, either in the thresholds or the weights. However, we have seen that for cases of extreme heterogeneity, only MFIII is valid as the original network of neurons contains a period-doubling bifurcation in the corresponding macroscopic variables. This is not possible with a smooth, two-dimensional system of differential equations which would most certainly be the result of using a Wilson-Cowan system with a network of adapting neurons as the only quantities one can think of are the average amounts of adaptation and the proportion of active excitatory neurons. For extreme heterogeneity, the correct Wilson-Cowan equations need to consider conditional moments of the adaptation variable or possibly higher order unconditioned moments.

8.1.2 Other Constructive Approaches

We distinguish our approach from a Wilson-Cowan approach by its bottom-up nature. We start with a network of neurons, and derive the subsequent equations. This approach is becoming increasingly common in the literature and here we compare the work of other authors in this regard. The majority of these approaches use dynamical systems averaging to obtain their final mean-field system of equations. This is different from the coarse

time graining in [190] but related. In dynamical systems averaging, one can convert a non-autonomous system of the form

$$\dot{x} = \epsilon f(t, x, \epsilon)$$

into an autonomous system of the form

$$\dot{\hat{x}} = \epsilon \frac{1}{T} \int_0^T f(\tau, \hat{x}, 0)\, d\tau$$

where $\hat{x}$ corresponds to the averaged variable [158].

One of the earliest methods in deriving rate equations from spiking neural networks is the work of Ermentrout (1994), [54]. In this paper, Ermentrout uses dynamical systems averaging in the limit of slow synapses to derive the rate equations for a network of neurons undergoing a type-I bifurcation with conductance based coupling. The equations governing the synaptic gating variables were assumed to be slow, allowing one to use dynamical systems averaging to replace these equations with their temporal averages. For example, for a network of two Morris-Lecar neurons, one obtains the following equations for the resulting synaptic conductance:

$$\tau_e s_e' = -s_e + C_e \sqrt{(\gamma_{ee} s_e - g_e^*(\gamma_{ie} s_i, I_e))_+} \tag{8.10}$$

$$\tau_i s_i' = -s_i + C_i \sqrt{(\gamma_{ei} s_e - g_e^*(\gamma_{ii}, I_i))_+} \tag{8.11}$$

where $(x)_+$ is x if x is positive and 0 otherwise [54]. The quantity $g^*(g, I)$ is the Type-I bifurcation point for the single neuron model. One could obtain the same set of equations if one were to reduce the Morris-Lecar models in the vicinity of a type-I bifurcation to the theta-model (Type-I normal form), and subsequently using the population density approach assuming that one had a large, homogeneous population of neurons. The assumption here though is that instead of a pair of neurons, there are a pair of populations of identical neurons. The advantage to the population density approach over averaging is the ability to compute the steady state voltage density, and thus obtain any moment of the voltage variables when the number of neurons is large. Additionally, spectral analysis also determines when the resulting mean-field system is appropriate. Aside from this extra information obtained from the population density approach, averaging and using a population density equation with a separation of time scales are equivalent for a homogeneous network.

The work in [54] was followed up by [163] which derived a set of rate-based equations for networks of conductance based neurons, and determined how to compute the synaptic weights coupling the neurons together. The network considered is given by

$$CV_i' = g_L(E_L - V_i(t)) + \sum_{j=1}^{N} g_{ij}(t)(E_j - V_i(t)) + g_i^{inp}(E^{inp} - V_i) \tag{8.12}$$

$$g_{ij}' = -\frac{g_{ij}}{\tau_{ij}} + G_{ij} \sum_{t<t_{jk}} \delta(t - t_{jk}) \tag{8.13}$$

$$g^{inp\prime} = -\frac{g^{inp}}{\tau^{inp}} + G^{inp} R_i^{inp}(t) \tag{8.14}$$

where $R_i^{inp}(t)$ is a Poisson process. The authors assume that the firing rates are linear (rectified neurons), which is valid for adapting neurons at steady state [57]. The resulting rate equations are given by

$$f_i = \beta \left(\sum_{j=1}^{N} J_{ij} f_j + J_{inp} f_i^{inp} - T \right)_+ \tag{8.15}$$

$$J_{ij} = G_{ij} \tau_{ij} (E_j - E_L - V_C) \tag{8.16}$$

$$J^{inp} = G^{inp} \tau^{inp} (E^{inp} - E_L - V_C) \tag{8.17}$$

where J_{ij} is the resulting weight between neuron i and neuron j. Unlike the work in [54], or the work done in this thesis, these are not derived explicitly using either averaging or the population density equation.

La Camera et al. [102, 103] derived mean-field equations for an uncoupled network of linear integrate and fire neurons with the same adaptation model as us and containing synaptically filtered noise. The mean adaptation equations presented by La Camera et al. [102, 103] are the same as our differential equation for the first moment of the adaptation current w in our networks (compare Eq. (44) with Eq. (3.3) in [103]). Again, there is a different interpretation of what the network averaged firing rate is under the case of noise. We note that it should be possible to derive precisely the same set of equations as in [103] for our network of two-dimensional neurons with synaptically filtered noise, by using the first order moment-closure and separation of time scales together, in conjunction with a diffusion approximation. In fact, the first order moment closure approximation may be better under cases of low firing rates when one adds noise. The reason for this is that adding noise to the voltage variable of a neuron (either synaptically filtered or otherwise) should decouple it from the adaptation variable.

Vladimirski et al. ([186]) have derived mean field equations for a network of linear integrate and fire neurons with synaptic depression and input currents which are either heterogeneous and deterministic or homogeneous and noisy. For the heterogeneous network, they note that one cannot use a single variable for the average depression to predict the behavior of the network. However, for their homogeneous network, they derive a mean-field equation for their mean synaptic depression, in terms of their network averaged firing rate (Eq. (17)). This equation is analogous to the equation for the mean adaptation variable in our model and those described above. Unlike our mean-field model, the mean field model from [186] has no differential equation for the mean synaptic gating variable. They use a time scale separation argument to derive an expression for the average of this variable over one oscillation period. This is equivalent to applying a quasi-steady state approximation to the equation for the s variable in our model. Further, since Vladimirski et. al. represent depression using a kinetic model [45] the derivation of the mean field equation is done by time averaging the fast gating variable over an oscillation period, as opposed to the approach used by us [102, 103]. Finally, [186] note that adaptation and depression work through two different mechanisms, yet both can yield rhythmic bursting.

It would be interesting to derive the mean-field system for both depressing and adapting two-dimensional networks and compare how these two mechanisms differ analytically.

In the work of Nesse et. al (2008) [132] the authors derive a mean-field approach to a coupled network of linear integrate and fire neurons with noise and adaptation. This type of network displays bursting and is used as a model of the bursting present in the Pre-Botzinger complex [132]. The authors make the assumption that the characteristic membrane time constant, τ_m is much faster (smaller) then the adaptation and synaptic time constants and thus a separation of time scales is applicable, as in here. The neural equations are:

$$\tau_m \frac{dv_i}{dt} = -v_i + s - h_i + I_V + x_i$$

where $v_i \in [-1, 1]$ and h_i is the adaptation variable, which can either be a simple activity dependent adaptation variable given by

$$\tau_h \frac{dh_i}{dt} = -h_i + a_h(s + x_i)$$

or that of a more complicated two dimensional gating current. The neurons are all to all coupled with α synapses, with $\alpha = \frac{1}{\tau_s}$ and all receive the common bias current I_v. Noise is filtered through an α synapse as well, to simulate stochastic spike trains from a Poisson distribution. The equations are given by

$$\begin{aligned} \tau_x \frac{dx_i}{dt} &= -x_i + y_i \\ \tau_x \frac{dy_i}{dt} &= -y_i + a_x \mu_i(t) \end{aligned}$$

The authors subsequently derive a master equation for y, and then apply the diffusion approximation to yield the Fokker-Plank equation. The resulting PDE is given by

$$\tau_x \frac{\partial q(y,t)}{\partial t} = -\frac{\partial}{\partial y}\left[(va_x - y)q(y,t)\right] + \frac{va_x^2}{2}\frac{\partial^2}{\partial y^2}\left[q(y,t)\right]$$

The authors subsequently argue that given this concatenation of the Kramers-Moyal expansion, one can approximate the synaptically filtered noise (after rescaling) as an Ornstein-Uhlenbeck process given by

$$\tau_x \frac{dx_i}{dt} = -x_i + y_i \tag{8.18}$$

$$\tau_x \frac{dy_i}{dt} = -y_i + \sigma\sqrt{\tau_x}\eta_i(t) \tag{8.19}$$

where $\eta_i(t)$ is a standard white noise process. To derive their mean field equations, the authors make the assumption that $u_i = s - h_i + x_i + I_v$ is a slowly varying in time relative

to v_i for each individual neuron. In which case, the firing rate can be computed for each neuron as

$$f(u_i) = \begin{cases} \frac{1}{\tau_r + \log\left(\frac{u_i+1}{u_i-1}\right)} & u_i > 1 \\ 0 & u_i < 1 \end{cases}$$

They then subsequently use averaging to write

$$\frac{1}{N}\sum_{i=1}^{N}\sum_{j\in S_i}\delta(t-t_j^i) \to \frac{1}{N}\sum_{i=1}^{N} f(u_i)$$

and the resulting set of mean-field equations is

$$\tau_h \frac{dh}{dt} = -h + a_h s \tag{8.20}$$

$$\tau_s \frac{ds}{dt} = -s + w \tag{8.21}$$

$$\tau_s \frac{dw}{dt} = -w + a_s \langle f \rangle \tag{8.22}$$

$\langle f \rangle$ is given by

$$\langle f \rangle = \int f(s - h + I_v + x)\rho_x(x)\, dx \tag{8.23}$$

where $\rho(x)$ is the steady state distribution for the Ornstein-Uhlenbeck random variable x_i (8.18). A similar derivation was used to derive a set of mean field equations for the more complicated adaptation current [132]. This is also a different approach from our sequence of steps for deriving a mean-field system. One thing to immediately note about (8.20)-(8.22) is that had we considered the same network, only with heterogeneity instead of coloured noise, we could have derived equations (8.20)-(8.22) by using a population density equation followed by a separation of time scales. In fact, the resulting equations (8.20)-(8.22) are MFII for the corresponding heterogeneous network with heterogeneous (and static) currents x_i with distribution $\rho_x(x)$. As the noise was twice synaptically filtered with a time constant that was five times greater then the membrane potential time constant, the temporal correlation was sufficiently large in the $x_i(t)$ variables so that they could effectively be considered as static heterogeneity and thus the accuracy in equation (8.20)-(8.22). There are alternative approaches to deriving mean-field systems with synaptically filtered noise [110].

Mean-field analysis on a network of Izhikevich neurons was also performed by Visser and Van Gils (2014) [185]. The approach they take is different from the population density derivations considered here and is based on phase plane analysis methods. While these phase plane constructions can account for some behaviors better then the systems derived in this thesis under a separation of time scales, as the derivation approach is more ad-hoc then a separation of time scales, it might become difficult to extend their derivation to other cases (heterogeneity, noise etc.).

8.1.3 Firing Rate Equations

Prior to outlining some future directions for mean-field analysis, we will consider one other final set of approaches that are prevalent in the literature: firing rate equations. Firing rate equations describe the dynamics of the firing rate as a set of closed differential equation. Technically, they should be derived under the assumption that the time scale of the voltage is much slower then the time scale of the synapses [41]. We have already seen a set of firing rate equations emerge in the section 8.1.1, the Wilson-Cowan equations [190].

More recently, Schaffer and colleagues have suggested a set of firing rate equations that are derived from eigenvalue expansions of the population density equation [159]. In particular, consider the population density equation for the network of integrate-and-fire neurons with dynamics $\dot{v} = G(v)$:

$$\frac{\partial \rho(v,t)}{\partial t} = -\frac{\partial}{\partial v} J(v,t) = -\frac{\partial}{\partial v}(G(v)\rho(v,t)) \tag{8.24}$$

$$\langle R_i(t) \rangle = J(v_{peak}, t) = G(v_{peak})\rho(v_{peak}, t) \tag{8.25}$$

In the case of the equation (8.24), the resulting operator is entirely linear and one can use an eigenvalue decomposition and can be solved using the Abbott-Vreeswjik transform. One obtains the solution to $\rho(v,t)$ as:

$$\rho(v,t) = \rho_\infty(v) + \sum_{n=1}^{\infty} \rho_n(v) a_n(t), \quad a_n'(t) = \lambda_n a_n(t) \tag{8.26}$$

where $\rho_n(v)$ are the eigenvectors associated with the eigenvalues λ_n. Then the flux is given by

$$J(v,t) = \rho_\infty(v)G(v) + G(v)\sum_{n=1}^{\infty} \rho_n(v) a_n(t) \tag{8.27}$$

$$= \langle R_i(t) \rangle_\infty + G(v)\sum_{n=1}^{\infty} \rho_n(v) a_n(t) \tag{8.28}$$

where $\langle R_i(t) \rangle_\infty$ is the steady state firing rate/flux. Suppose we take the eigenvalues and reorder them so that the $n = 1$ eigenvalue is the dominant eigenvalue. Then we can write down the following

$$\langle R_i(t) \rangle = J(v_{peak}, t) = \langle R_i(t) \rangle_\infty + G(v_{peak}) \sum_{n=1}^{\infty} \rho_n(v_{peak}) a_n(t) \tag{8.29}$$

$$\approx \langle R_i(t) \rangle_\infty + G(v_{peak})\rho_1(v_{peak}) a_n(t) \tag{8.30}$$

$$\langle R_i(t) \rangle' = G(v_{peak})\rho_1(v_{peak})\lambda_1 a_1(t) = \lambda_1(\langle R_i(t) \rangle - \langle R_i(t) \rangle_\infty) \tag{8.31}$$

which is the reduction suggested in [159] (see Materials and Methods in [159] for more detail). However, this approach immediately fails to yield a one-dimensional closed form

autonomous differential equation when one considers more eigenvalues then just the dominant one. Furthermore, since all the eigenvalues and eigenvectors can be determined explicitly for a variety of integrate and fire models, why not simply use all of them and the explicit solution to $\langle R_i(t) \rangle$ given by equation (8.28)?

Similar firing rate equations were suggested in [76] in their study of networks of excitatory and inhibitory quadratic integrate-and-fire neurons with heterogeneity with double exponential coupling (equations (5.5) and (5.6)). Under the limit of slow coupling for their synapses, the authors also use a perturbation approach to resolve their resulting spectral equation. The spectral equation simplifies and the authors note that the spectral equation is identical to that of a two-dimensional system of firing rate equations. There does not seem to be a derivation of the slow system of firing rate equations, and it appears that the authors concluded the validity of their firing rate equations from the correspondence of characteristic polynomials that emerge from their perturbation solution and their suggested firing rate equations. If this is the case, this poses a significant problem as characteristic polynomials do not uniquely determine the dynamical systems. Two completely different sets of firing rate equations could have identical characteristic polynomials. Furthermore, in the limit of slow synapses one should expect that the resulting low dimensional dynamics should be equations for these synapses, as opposed to the firing rates.

A more careful construction of firing rate models appears to occur in [2] where the resulting rate equations still have the original dynamical variables in the coupling, in addition to a firing rate equation with an undetermined time constant. The authors suggest some caution in using eigenvalue matching for generating these types of equations, especially for large n eigenvalues.

8.2 Conclusions and Future Directions for Mean-Field Analysis

In this thesis we have successfully derived mean-field systems for large networks of integrate-and-fire neurons with adaptation and conductance based coupling by using moment-closure reduced population density equations and time scale-separation. We have made these reductions for networks that are all-to-all coupled, or random sparse coupling. Additionally, these reductions have been performed with the presence of noise or heterogeneity in the neurons. The resulting systems are low dimensional and amenable to bifurcation analysis either directly or numerically and can predict the behavior of the original network with both qualitative and quantitative accuracy.

8.2.1 Firing Rate Equations via an Alternate Time Scale-Separation

Following [41], a set of firing rate equations should emerge in the limit that the synapses are much faster then the neural dynamics. Consider a network of integrate and fire neurons

with conductance based, all-to-all coupling, $\dot{v}_i = G(v_i, s)$ where s is given by simple exponential coupling with time scale $\tau_s = \epsilon$. Then the resulting population density equation should be

$$\frac{\partial \rho_V(v,t)}{\partial t} = -\frac{\partial}{\partial v}(G(v,s)\,\rho_V(v,t)) \tag{8.32}$$

$$\epsilon s' = -s + \lambda_s G(v_{peak}, s(t))\rho_V(v_{peak}, t) \tag{8.33}$$

Then the system (8.32)-(8.33) is the slow system and the change of time scale $\tau = \epsilon^{-1}t$ yields the fast system

$$\frac{\partial \rho_V(v,t)}{\partial \tau} = -\epsilon\frac{\partial}{\partial v}(G(v,s)\rho_V(v,t)) \tag{8.34}$$

$$\dot{s} = -s + \lambda_s G(v_{peak}, s)\rho_V(v_{peak}, t) \tag{8.35}$$

If we work with the slow system, we substitute the steady state value of s and the resulting equation is

$$\frac{\partial \rho_V(v,t)}{\partial t} = -\frac{\partial}{\partial v}(G(v, \lambda_s J(v_{peak}, t))\rho_V(v,t)) \tag{8.36}$$

and the flux becomes implicitly defined with

$$J(v_{peak}, t) = G(v_{peak}, \lambda_s J(v_{peak}, t))\rho_V(v_{peak}, t) \tag{8.37}$$

With some clever analysis one might be able to resolve a firing rate equation, and consider what conditions the equation is a valid descriptor of the network activity. The analysis of this other time scale separation and the derivation of the resulting firing rate equations will be left for future work.

8.2.2 Geometric Singular Perturbation Approaches

We have seen that for physiologically reasonable values of the time scale parameters, the voltage and the synapses often function on different time scales. We have primarily considered slow systems here as approximate solutions to the moment-closure reduced population density equation. However, both in the case of homogeneous networks and heterogeneous networks, the fast system has been analytically tractable. Thus, one possible extension to the work done in this thesis is the applicability of various techniques in singular perturbation theory to obtain better approximations by using both the fast and slow system solutions. Geometric singular perturbation approaches have already been used to analyze infinite-dimensional dynamical systems with a separation of time scales [122]. A multiple-time scale expansion should be applicable to the population density equations we consider, in addition to other techniques in geometric singular perturbation theory [94]. We leave this for future work.

8.2.3 Bifurcation Analysis of the Mean-Field System with Noise

We have numerically conducted some preliminary bifurcation analysis of the mean-field system with noise in section 6.5. We have seen that the mean-field system with noise can be considered a regularization of the system without noise. As the system without noise is analytically tractable, and one can obtain higher order expansions to the firing rate using Laplace's method [17], then one might be able to analytically resolve the behavior of the mean-field system with noise and determine how it behaves as a regularization. Of particular interest is the co-dimension 3 bifurcation where two Bautin bifurcation branches emerge from the point $(\sigma = 0, g = \bar{g}, I = I_{rh})$. Unfortunately, we could only investigate this issue numerically. With higher order expansions of the firing rate it might be possible to investigate this bifurcation point analytically. We leave this for future work.

8.3 Related Approaches to the Neural Engineering Framework

While there are many examples in the literature of spiking neural networks created to accomplish very specific tasks, the vast majority of these are ad-hoc without a formal unifying theory. The only constructive method in the literature for generating a spiking neural network with arbitrary dynamics is the predictive coding approach of Boerlin and colleagues [20]. However the work in [20] only considers linear dynamics. Despite this constraint, we will consider the work done in [20] in some detail.

The appropriate point of comparison between the NEF approach and the predictive coding approach from [20] is how these two methods consider the role of optimization. In the NEF approach, an initial heterogeneous network is considered with current based coupling and optimal decoders for the spikes of each neuron are determined. In the predictive coding approach however, the optimization is carried much further in the sense that there is no initial network at all. The authors assume, as in the NEF approach that the spike trains of an as of yet undetermined network is weighted by the decoders:

$$\dot{\hat{\boldsymbol{s}}} = -\frac{\hat{\boldsymbol{s}}}{\tau_s} + \sum_{j=1}^{N} \sum_{t<t_{j,k}} \boldsymbol{\phi}_j \delta(t - t_{j,k}) \tag{8.38}$$

which will act as an approximation for the linear dynamical system

$$\dot{\boldsymbol{s}} = \boldsymbol{A}\boldsymbol{s} + \boldsymbol{c}(t) \tag{8.39}$$

The firing rate of the neurons is then determined by the following unweighted sum:

$$\dot{r}_i = -\frac{r_i}{\tau_s} + \sum_{t<t_{j,k}} \delta(t - t_{j,k}), \quad i = 1, 2, \ldots N \tag{8.40}$$

Additionally, define $\boldsymbol{o}(t)$ as a vector of delta trains, so that

$$\boldsymbol{r}(t) = -\frac{\boldsymbol{r}}{\tau_s} + \boldsymbol{o}(t)$$

Then the authors subsequently determine the network that minimizes the following functional:

$$E(t) = \int_0^t \left(\|\boldsymbol{s}(t') - \hat{\boldsymbol{s}}(t')\|_{L_2}^2 + \nu \|\boldsymbol{r}(t')\|_{L_1} + \mu \|\boldsymbol{r}(t')\|_{L_2}^2 \right) dt' \tag{8.41}$$

where L_2 and L_1 denote the respective norms. The first term minimizes the error, the second term minimizes the spiking, while the third term distributes the spiking equitably to neurons in the network. Thus, the predictive coding approach optimizes in time versus the NEF approach which optimizes in the phase space. It is likely that these two approaches are related when $\nu = \mu = 0$, and we leave much of this analysis for future work. The authors immediately realize that the functional (8.41) is unworkable aside from a greedy optimization approach for each neuron. They implement that via the criterion

$$E(t|\text{neuron k spikes}) < E(t|\text{neuron k is silent}) \tag{8.42}$$

which based on the explicit integrability of the synapses/filters considered, allows them to simplify condition (8.42). The authors look at a short time into the future, $t + \epsilon$ to see the affects of the spiking of neuron k, this yields the following sets of equations

$$\begin{aligned}
E(t|\text{neuron k spikes}) &= \int_t^{t+\epsilon} \|\boldsymbol{s}(t') - \hat{\boldsymbol{s}}(t') - \boldsymbol{\phi}_k \exp(-(t'-t)/\tau_s)\|^2 \, dt' && (8.43)\\
&+ \int_t^{t+\epsilon} \nu \|\boldsymbol{r}(t') + \boldsymbol{e}_k \exp(-(t'-t)/\tau_s)\|_1 && (8.44)\\
&+ \int_t^{t+\epsilon} \|\boldsymbol{r}(t') + \boldsymbol{e}_k \exp(-(t'-t)/\tau_s)\|_{L_2}^2 \, dt' && (8.45)\\
E(t|\text{neuron k is silent}) &= \int_0^{t+\epsilon} \|\boldsymbol{s}(t') - \hat{\boldsymbol{s}}(t')\|^2 \, dt' && (8.46)\\
&+ \int_t^{t+\epsilon} \nu \|\boldsymbol{r}(t')\|_1 && (8.47)\\
&+ \int_t^{t+\epsilon} \|\boldsymbol{r}(t')\|_{L_2}^2 \, dt' && (8.48)
\end{aligned}$$

where we can define

$$h(t) = \begin{cases} \exp(-t/\tau_s) & t > 0 \\ 0 & t < 0 \end{cases}. \tag{8.49}$$

One can simplify the condition (8.42) to yield the following

$$\int_t^{t+\epsilon} \left\{ 2h(t-t')\boldsymbol{\phi}_k^T \left[\boldsymbol{s}(t') - \hat{\boldsymbol{s}}(t') \right] - 2\mu \exp(-(t'-t)/\tau_s)\tau_s^{-1}(\boldsymbol{e}_k^T \boldsymbol{r}(t')) \right\} dt' \tag{8.50}$$

$$> \int_t^{t+\epsilon} \left\{ 2h(t-t')\boldsymbol{\phi}_k^T \boldsymbol{\phi}_k + \nu h(t-t')\tau_s^{-1} + \mu h(t-t')^2 \tau_s^{-2} \right\} dt' \tag{8.51}$$

The authors subsequently use both (8.49) in addition to assuming that ϵ is small, $\epsilon \ll \tau_s$ to remove the time integral completely. In this case, $h(t) \approx 1$, and they arrive at the following equation

$$\phi_k^T(\boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t)) - \mu\tau_s^{-1}\boldsymbol{e}_k^T\boldsymbol{r}(t) > \frac{1}{2}\left(\boldsymbol{\phi}_k^T\boldsymbol{\phi}_k + \nu\tau_s^{-1} + \mu\tau_s^{-2}\right) \tag{8.52}$$

The authors then immediately note that the right hand side of (8.52) is time varying, while the left hand side is a fixed constant. Thus, they cleverly suggest to interpret the right hand side as a neural voltage, and the left hand side as the threshold. This yields the following

$$\begin{aligned} \boldsymbol{v}(t) &= \boldsymbol{\phi}^T(\boldsymbol{s}(t) - \hat{\boldsymbol{s}}(t)) - \mu\tau_s^{-1}\boldsymbol{r}(t) && (8.53)\\ s(t) &= (\boldsymbol{\phi}\boldsymbol{\phi}^T)^{-1}\boldsymbol{\phi}\boldsymbol{v}(t) + \mu\tau_s^{-1}(\boldsymbol{\phi}\boldsymbol{\phi}^T)^{-1}\boldsymbol{\phi}\boldsymbol{r}(t) + \hat{\boldsymbol{s}}(t) && (8.54)\\ &= \boldsymbol{L}\boldsymbol{v}(t) + \mu\tau_s^{-1}\boldsymbol{L}\boldsymbol{r}(t) + \hat{\boldsymbol{s}}(\boldsymbol{t}). && (8.55) \end{aligned}$$

Differentiating equation (8.53) immediately yields the dynamics for the neurons

$$\boldsymbol{v}'(t) = \boldsymbol{\phi}^T(\boldsymbol{s}'(\boldsymbol{t}) - \hat{\boldsymbol{s}}'(\boldsymbol{t})) - \mu\tau_s^{-1}\boldsymbol{r}'(t) \tag{8.56}$$

Take note of this equation for later, it is of extreme importance for the generalizability of this approach. We can use equation (8.55), in addition to equation (8.39) to eliminate $s(t)$:

$$\begin{aligned} \dot{\boldsymbol{v}} &= \boldsymbol{L}^{\boldsymbol{T}}\boldsymbol{A}\boldsymbol{L}\boldsymbol{v} + \left(\tau_s\boldsymbol{\phi}^T A\boldsymbol{\phi} + \boldsymbol{\phi}^T\boldsymbol{\phi} + \mu\tau_s^{-1}\boldsymbol{\phi}^T AL + \mu\tau_s^{-2}\right)\boldsymbol{r}(t) && (8.57)\\ &- \left(\boldsymbol{\phi}^T\boldsymbol{\phi} + \mu\tau_s^{-2}I_N\right)\boldsymbol{o}(t) + \boldsymbol{\phi}^T\boldsymbol{c}(t) && (8.58) \end{aligned}$$

which along with the threshold given by the right hand side of (8.52) yields a leaky-integrate-and-fire neuron. The authors impose a couple of constraints in the large network limit. In particular, they force the average firing rates of the neurons and the read outs to be constant in the limit of large networks. This immediately forces the decoders $\boldsymbol{\phi} \propto N^{-1}$ which also implies that $\boldsymbol{L} \propto N^{-1}$. But then the very term that makes the integrate and fire neuron (8.58) a LIF model scales like N^{-2}. This would not be too much trouble, if all the terms in the dynamics scale like N^{-2} and the threshold in the neurons had a similar scaling. While the threshold for the neurons scales similarly, the coefficient in front of the firing rate scales like N^{-1}. Thus, the neuronal dynamics for networks that are large are effectively that of a perfect integrate-and-fire neuron, that has no voltage dynamics and merely integrates the current. This model has a unique advantage over any other integrate-and-fire model in the sense that any noise present does not effect the tuning curve. Furthermore, their optimization weights μ and ν are determined so that they do not dominate the dynamics. As the dominant term in the dynamics is N^{-1}, that immediately forces $\mu \propto N^{-1}$ and the authors suggest $\mu \propto N^{-2}$ for better accuracy. The connectivities for the network are determined by

$$\begin{aligned} \boldsymbol{\Omega}^s &= \boldsymbol{\phi}^T(\boldsymbol{A} + \tau_s^{-1})\boldsymbol{\phi} && (8.59)\\ \boldsymbol{\Omega}^f &= \boldsymbol{\phi}^T\boldsymbol{\phi} + \mu\tau_s^{-2}\boldsymbol{I}_N && (8.60)\\ \boldsymbol{W}(t) &= \boldsymbol{\Omega}^s h(t) - \boldsymbol{\Omega}^f\delta(t) && (8.61) \end{aligned}$$

and thus we require a set of slow filtered connections with weights determined by $\mathbf{\Omega}^s$ and a set of fast instantaneous connections with magnitude $\mathbf{\Omega}^f$.

The general mathematical approach, while very elegant, simply places far too much on optimization to generate a biologically plausible network for a few reasons. First, this approach cannot be generalized to nonlinear dynamics. Using equation (8.56) for some nonlinear dynamics $\boldsymbol{s} = \boldsymbol{G}(\boldsymbol{s})$ yields

$$\dot{\boldsymbol{v}}' = \boldsymbol{\phi}^T(\boldsymbol{G}(\boldsymbol{s}) - \hat{\boldsymbol{s}}'(t)) - \mu\tau_s^{-1}\boldsymbol{r}'(t) \tag{8.62}$$

$$\dot{\boldsymbol{v}}' = \boldsymbol{\phi}^T\left\{\boldsymbol{G}\left[(\boldsymbol{\phi}\boldsymbol{\phi}^T)^{-1}\boldsymbol{\phi}\boldsymbol{v}(t) + \mu\tau_s^{-1}(\boldsymbol{\phi}\boldsymbol{\phi}^T)^{-1}\boldsymbol{\phi}\boldsymbol{r}(t) + \hat{\boldsymbol{s}}(\boldsymbol{t})\right] - \hat{\boldsymbol{s}}'(\boldsymbol{t})\right\} - \mu\tau_s^{-1}\boldsymbol{r}'(t) \tag{8.63}$$

and while equation (8.63) is a closed system of differential equations for the voltage, the dynamics are incredibly complicated and the nonlinear dynamics of the macroscopic equations determine the nonlinear dynamics of the individual neurons, not the other way around. Furthermore, even in the case of linear dynamics, the differential equations defining the neurons effectively just integrate all their incoming inputs, as in the perfect integrate-and-fire neuron, a very pathological neuron model [66]. Finally the synaptic weights feature instantaneous delta coupling, which is not biologically plausible.

8.4 Conclusions and Future Directions for the Neural Engineering Framework and Scale-Invariant Linear Decoding

We have used the Neural Engineering Framework to demonstrate how weight matrices that couple recurrent networks of type-I or type-II neurons have the form

$$\omega_{ij} = \Omega(\beta_i, \beta_j) \tag{8.64}$$

where β_i and β_j are the heterogeneous parameters for neuron i and neuron j in the network allow the network to have arbitrary dynamics. Additionally, we have shown that these weight matrices are non-unique by using the common range in the tuning curve transforms. Here, we will discuss various extensions to the NEF approach.

8.4.1 Other Kinds of Networks

The networks considered were networks of non-adapting theta neurons with current based coupling. One immediate extension is to conductance based coupling that we intend to leave for future work. Additionally, to the best of our knowledge, the NEF approach has not been analytically extended to other types of synaptic coupling such as kinetic coupling or double-exponential coupling. These types of coupling functions add their own dynamics

to the mix and thus act as constraints on the set of dynamics a network can perform. For example, if we consider a network of theta neurons with current based coupling and a double exponential synapse, we have the following equation:

$$\begin{aligned} s' &= -\frac{s}{\tau_R} + h && (8.65)\\ h' &= -\frac{h}{\tau_D} + \int_{\alpha,\beta} \rho_{\alpha,\beta}(\alpha,\beta)\gamma(\alpha,\beta)\sqrt{\alpha s + \beta}\, d\alpha\, d\beta && (8.66)\\ &= -\frac{h}{\tau_D} + \hat{g}(s) && (8.67) \end{aligned}$$

and thus any recurrent network coupled with a double-exponential synapse has to have second-order dynamics. Adaptation presents a similar problem where MFII yields the following macroscopic dynamics with a scale-invariant decoder

$$\begin{aligned} s' &= -\frac{s}{\tau_s} + \int_{\alpha,\beta} \rho_{\alpha,\beta}(\alpha,\beta)\gamma(\alpha,\beta)\sqrt{\alpha s + \beta - w}\, d\alpha\, d\beta && (8.68)\\ &= -\frac{s}{\tau_s} + X(s,w) && (8.69)\\ w' &= -aw + \int_{\alpha,\beta} \rho_{\alpha,\beta}(\alpha,\beta)\sqrt{\alpha s + \beta - w}\, d\alpha\, d\beta && (8.70)\\ &= -aw + Y(s,w) && (8.71) \end{aligned}$$

In general, the way to proceed with other kinds of networks is to first derive the mean-field system for the network to determine what kind of dynamics the network can represent recurrently. Then one can either use scale-invariant decoders or optimal decoders to select a particular set of dynamics out of the subset of admissable dynamical systems. We leave this generalization for future work.

8.4.2 Homogeneous Networks

In this section, we will show that the NEF approach can easily be extended to homogeneous networks. We leave much of the numerics and interpretation on rate coding for future work. Consider a network of neurons with firing rate $f(I)$ and current based, simple exponential coupling. Then the current arriving to each neuron, s_i, is determined by

$$s_i(t) = \sum_{j=1}^{N} \sum_{t<t_{j,k}} \omega_{ij} E(t - t_{j,k}) \tag{8.72}$$

where ω_{ij} are the synaptic weights. The NEF approach forces this current to approximate the following equation:

$$s_i = \sum_{j=1}^{N} \sum_{t<t_{j,k}} \omega_{ij} E(t - t_{j,k}) \approx \alpha_i s \tag{8.73}$$

where the dynamics of s are given by

$$s' = G(s). \tag{8.74}$$

These are the macroscopic dynamics we want the recurrent network to perform. The NEF approach does this by forcing $\omega_{ij} = \alpha_i \phi_j^*$ where ϕ_j^* are the optimal decoders. Suppose however we forced the following relationship instead:

$$s_i = \sum_{j=1}^{N} \sum_{t<t_{j,k}} \omega_{ij} E(t - t_{j,k}) \approx \alpha_i s + \beta_i \tag{8.75}$$

in which case all the formerly intrinsic neuronal heterogeneity is now extrinsic heterogeneity in the connections. If we differentiate equation (8.75) we arrive at the following equation:

$$s_i' = -\frac{s_i}{\tau_s} + \frac{1}{\tau_s} \sum_{j=1}^{N} \omega_{ij} f(s_j) \approx \alpha_i G(s) \tag{8.76}$$

$$\Rightarrow \quad -\frac{\alpha_i s + \beta_i}{\tau_s} + \frac{1}{\tau_s} \sum_{j=1}^{N} \omega_{ij} f(\alpha_j s + \beta_i) \approx \alpha_i G(s) \tag{8.77}$$

and thus if we minimize the following functional

$$C_i(\boldsymbol{\omega}_i) = \int_X \left(\sum_{j=1}^{N} \omega_{ij} f(\alpha_j x + \beta_j) - \tau_s \alpha_i G(x) - \alpha_i x - \beta_i \right)^2 dx + \lambda \sum_{j=1}^{N} \omega_{ij}^2 \quad i = 1, 2, \ldots N \tag{8.78}$$

where $\boldsymbol{\omega}_i$ is the ith row of the weight matrix, the resulting macroscopic dynamics are given by $\dot{s} = G(s)$, only the current due to coupling for each neuron is approximately the quantity $\alpha_i s + \beta_i$. Alternatively, the integral quantity can be minimized by using a scale-invariant approach to the weights. Instead of solving a single optimization problem for the decoders, a homogeneous network requires solving N optimization problems for each row of the weight matrix. We leave this extension and the numerical implementation for future work. We also note that this network allows for a direct comparison between intrinsic and extrinsic sources of heterogeneity.

8.4.3 Weight Matrix Analysis

In the previous section, we considered more general weight matrices and showed that you could use them to construct a homogeneous network with arbitrary recurrent dynamics. A more interesting question that occurs after considering other kinds of weight matrices is the following: For a recurrent network of neurons with rate coding, what is the family of weight matrices that result in macroscopic dynamics $\dot{s} = F(s)$?. We have already shown the non-uniqueness that emerges by considering tuning-curve transforms and their mutual range in the function space. It turns out that this is only the tip of the iceberg.

Consider the following equation for the current arrive at neuron i due to synaptic coupling:

$$s_i = \sum_{j=1}^{N} (\omega_{ij} + \epsilon_{ij}) E(t - t_{j,k}) \tag{8.79}$$

then as in the previous section, if we want the macroscopic dynamics $\dot{s} = G(s)$, then we require the following

$$\sum_{j=1}^{N} (\omega_{ij} + \epsilon_{ij}) f(\alpha_j x + \beta_i) \approx \hat{g}_i(x) \tag{8.80}$$

for $\hat{g}_i(s) = \tau_s \alpha_i G(x) + x\beta_i$. Note if the weights ω_{ij}, $j = 1, 2, \ldots N$ are used to approximate $\hat{g}_i(x)$ either optimally or with a scale-invariant solution, then this leaves

$$\sum_{j=1}^{N} \epsilon_{ij} f(\alpha_j x + \beta_i) \approx 0 \tag{8.81}$$

which in the large network limit yields:

$$L^+(\hat{P}_i^{\ +}) + L^-(\hat{P}_i^{\ -}) \;\; = \;\; 0 \tag{8.82}$$

but we have already seen that as these two linear operators share a common range, we can simply take functions, $\epsilon_i(x)$ that lie in both ranges like so:

$$L^+(\hat{P}_i^{\ +}) \;\; = \;\; \epsilon_i(x) \tag{8.83}$$

$$L^-(\hat{P}_j^{\ -}) \;\; = \;\; -\epsilon_i(x) \tag{8.84}$$

while this may seem like a repeat of the work done in section 7.2, their is one crucial difference. The perturbations can vary with the row index of the weight matrix. This substantially increases the space of matrices over which our network can display the macroscopic dynamics we want. One natural question that emerges is whether or not this larger set of weight matrices covers all possible weight matrices that yield the prescribed macroscopic dynamics for the network. We leave this question for future work.

8.4.4 The Population Density Equations and Stability of the NEF solution

To the best of our knowledge, no one has explored the applicability of population density methods to the Neural Engineering Framework. If population density equations were to be applied to NEF generated networks, one could obtain significantly more information about the network, and not just at steady state values. Additionally, one could reproduce the spectral analysis conducted here with heterogeneous networks and one might able

to determine under what conditions the macroscopic network equations are stable as a function of the distribution of heterogeneous parameters, or the synaptic time constants coupling the networks together.

For theta-neurons specifically, we may be able to use some previous work with population density equations. Networks of heterogeneous theta oscillators have been extensively analyzed in [14, 109, 171] by using the Ott-Antonsen Anzats initially applied to networks of Kuramoto Oscillators [115, 141]. Additionally, one of the weight solutions for a network with one-dimensional dynamics that arises from the scale-invariant decoders sets all the weights to $\pm\omega$ by setting the density to $\rho \propto |\hat{P}(a)|$ where the constant of proportionality normalizes $|\hat{P}(a)|$. Given that, it may be possible to apply some of the existing literature on the stability analysis of networks of heterogeneous theta neurons to this network.

8.4.5 Relationships Between the Analytical and Optimal Decoders

The optimal decoders appear to have an asymptotic weighted decoder, the product of the density function multiplied by the scaled optimal decoder, $N\phi_i^*\rho_i^*$, when one computes this quantity numerically after large-matrix inversion (see Figure 7.2). For the optimal decoders, the weighted decoder has high frequency oscillations that are related to the idiosyncracies of the particular sample of random neurons generated. These seem to attenuate with increasing network size, and regularization parameter λ. These high frequency oscillations are for example eliminated when the neurons are drawn from a grid. Indeed, when the neurons are drawn from a grid, the optimal decoders for a much larger network can be approximated by simply interpolating between the decoder values for the smaller grid network, and rescale the interpolated decoders in accordance with the network size.

Thus, one may ask is the weighted decoder generated by the optimal decoders 1) convergent as $N \to \infty$, and 2) does it converge to any particular weighted decoder in the set defined by the requirement that $\hat{g}(x) = M(P^+, P^-) = L_+(P^+) + L_-(P^-) = g(x)$? We have seen numerically that the quantity $N\rho_a(a_i)\phi_i^*$ does converge to some surface $P^\pm$ that varies depending on the identity of neuron i as an ON/OFF neuron. The likely candidate for the specific $\hat{P}^\pm$ in the set defined by $M(P^+, P^-)$ is the surface that minimizes equation (7.65). However, the relationship between the optimal decoders, and any particular scale-invariant decoder as $N \to \infty$ is outside of the scope of this thesis and we leave it for future work.

APPENDICES

A Code for Mean-Field System Integration

A.1 Mean-Field System for a Homogeneous Network

In this section, we will include a pair of ODE file that can be used to simulate a mean-field system. One file will be for the network with neurons that have

$$\dot{v} = F(v)$$

for arbitrary $F(v)$ as the dynamics, while the second will be for an Izhikevich network. The integrals

$$\langle R_i(t)\rangle^{-1} = \int_{v_{reset}}^{v_{peak}} \frac{1}{F(v) + I - \langle w\rangle + gs(e_r - v)}\, dv \tag{85}$$

$$\langle v\rangle = \int_{v_{reset}}^{v_{peak}} \frac{v\langle R_i(t)\rangle}{F(v) + I - \langle w\rangle + gs(e_r - v)}\, dv \tag{86}$$

have to be computed at each time step in the integration scheme, and thus slows down integration somewhat when this has to be numerically done (as in the AdEx or other models) versus when there is an explcit expression for this integral (such as the Izhikevich). One can use any method of numerical integration to compute (85) such as the trapezoidal method, Gaussian Quadrature, Simpsons method, or Monte-Carlo methods. We have primarily used the MATLAB *ode45* function as our numerical integrator ODE [116, 161] with the trapezoidal method to compute (85). For the Izhikevich neuron, both the (85) and (86) can be computed explicitly. Equation (86) is given by

$$\langle v\rangle = \begin{cases} \frac{1}{2}(\alpha + gs) + \frac{\langle R_i(t)\rangle}{2} \log\left(\frac{(v_{peak} - (\alpha+gs)/2)^2 + I - I^*(s,w)}{(v_{reset} - (\alpha+gs)/2)^2 + I - I^*(s,w)}\right) & I - I^*(s,w) > 0 \\ \frac{1}{2}(\alpha + gs) - \sqrt{I^*(s,w) - I} & I - I^*(s,w) < 0 \end{cases} \tag{87}$$

```
function dy = ONEIZNETWORKQSSA(v,g,I,er,ts,tw,sjump,wjump,b,t,y)
s = y(1);
w = y(2);
%Need to define F according to which neuronal model you are using, for
%example:
%F = exp(v)-v  %AdEX
vreset = v(1); vpeak = v(end);
G = F+g*s*(er-v)+I-w;
H = min(F+g*s*(er-v)+I-w);
if H>0;
J = trapz(v,1./(G));
R = 1/J;
meanv = trapz(v,v*R./(G));
else R = 0;
    vminus = fsolve(@(x) F(x)+g*s*(er-x)+I-w,vreset);
    %The expression for vminus varies depending on the neuronal model used,
    %see Thesis for a list.  It can also be computed numerically however using the
    %fsolve function, initialzied in the correct region.
    meanv = vminus;
end
dy(1) = -s/ts + sjump*R;
dy(2) = b*meanv/tw-w/tw + wjump*R;
end
```

Figure 1: Code to compute the right hand side of the mean-field system for an arbitrary neuronal mode, $\dot{v} = F(v)$. By using ODE45 to numerically integrate the autonomous, non-smooth system of equations, one can simulate the mean-field system

```
function dy = ONEIZNETWORKQSSA(vreset,vpeak,g,I,er,ts,tw,sjump,wjump,b,t,y)
s = y(1);
w = y(2);
H = I-0.25*(alpha+g*s).^2+g*s*er-w;
if H>0;
R = sqrt(H)/(atan( (vpeak-(alpha+g*s)*0.5)/sqrt(H))-
atan( (vreset-(alpha+g*s)*0.5)/sqrt(H)));
meanv = 0.5*(alpha+g*s)+0.5*R*log( ((vpeak-0.5*(alpha+g*s))^2 + H)
./((vreset-0.5*(alpha+g*s))^2 + H));
else
   R = 0;
   meanv = 0.5*(alpha+g*s)-sqrt(-H);
end

dy(1) = -s/ts + sjump*R;
dy(2) = b*meanv/tw-w/tw + wjump*R;
end
```

Figure 2: Code to compute the right hand side of the mean-field system for the Izhikevich model explicitly, without numerical integration to resolve $\langle R_i(t)\rangle$ and $\langle v\rangle$

A.2 MFII and MFIII for a Heterogeneous Network

To compute the mean-field system with heterogeneity, one has to compute the integrals

$$\langle R_i(t)\rangle = \int_{\boldsymbol{\beta}} \rho_\beta(\beta)\langle R_i(t)|\beta\rangle \, d\beta \tag{88}$$

$$\langle R_i(t)|\beta\rangle^{-1} = \int_{v_{reset}}^{v_{peak}} \frac{1}{G_1(v,w,s,\beta)} \, dv \tag{89}$$

at each time step in the numerical integration process. One can again use numerical integration to compute these two integrals by defining a 2D mesh using the meshgrid function in matlab. Note that a similar set of equations have to be computed in order to evaluate

$$\langle v\rangle = \int_{\beta} \rho_\beta(\beta)\langle v|\beta\rangle \, d\beta \tag{90}$$

$$\langle v|\beta\rangle = \int_{v_{reset}}^{v_{peak}} \frac{\langle R|\beta}{G_1(v,\langle w\rangle,s,\beta)} \, dv \tag{91}$$

Example code is shown below for MFII. A slight modification to the code allows us to simulate MFIII as well.

```
function dy = MFII(alpha,g,I,er,vpeak,vreset,ts,tw,sjump,wjump,mu1,std1,t,y)
s = y(1);
w = y(2);
M = 25000; %Grid Size I
m = 200; %Grid size in v
minpar = mu1 - 5*std1;  %Set the range for integration
maxpar = mu1 + 5*std1;
dp = (maxpar-minpar)/M; %Grid Step Size
I =minpar + cumsum(ones(M,1))*dp; %Grid of heterogeneous parameter
rho = normpdf(I,mu1,std1) %Density Function

H = I + g*er*s - w - ((alpha + g*s).^2)/4; %The switching manifold
x = (vpeak-0.5*(alpha+g*s))./sqrt(H);
z = (vreset-0.5*(alpha+g*s) )./sqrt(H);
RI = real((H>0).*sqrt(H)./(atan(x)-atan(z))); %Compute <R|\beta> explicitly
R = simpsons((RI.*rho),M,minpar,maxpar); %Use simpsons method integration.

%% For non-izhikevich,LIF neurons, can use the following piece of code,
%% with something similar for <v>
% v = vreset+(vpeak-vreset)*(1:1:m)/m;
%[VX,IX]=meshgrid(v,I);
% HX = IX + g*er*s - w - ((alpha + g*s).^2)/4
%RI = (H>0)./trapz(v,(HX>0)./(F(VX)+g*s*(er-VX)+I-W))
%R = trapz(I,RI.*rho);
dydt=[-s/ts + sjump*R; -w/tw + wjump*R];
end
```

Figure 3: The code to run MFII for a network of Izhikevich neurons

```
function dy = MFII(alpha,g,I,er,vpeak,vreset,ts,tw,sjump,wjump,mu1,std1,t,y)
s = y(1);
w = y(2:m+1);
M = 25000; %Grid Size I
m = 200; %Grid size in v
minpar = mu1 - 5*std1;  %Set the range for integration
maxpar = mu1 + 5*std1;
dp = (maxpar-minpar)/M; %Grid Step Size
I =minpar + cumsum(ones(M,1))*dp; %Grid of heterogeneous parameter
rho = normpdf(I,mu1,std1); %Density Function

H = I + g*er*s - w - ((alpha + g*s).^2)/4; %The switching manifold
x = (vpeak-0.5*(alpha+g*s))./sqrt(H);
z = (vreset-0.5*(alpha+g*s) )./sqrt(H);
RI = real((H>0).*sqrt(H)./(atan(x)-atan(z))); %Compute <R|\beta> explicitly
R = simpsons((RI.*rho),M,minpar,maxpar); %Use integration.

%% For non-izhikevich,LIF neurons, can use the following piece of code,
%% with something similar for <v>
% v = vreset+(vpeak-vreset)*(1:1:m)/m;
%[VX,IX]=meshgrid(v,I);
% HX = IX + g*er*s - w - ((alpha + g*s).^2)/4
%RI = (H>0)./trapz(v,(HX>0)./(F(VX)+g*s*(er-VX)+I-W))
%R = trapz(I,RI.*rho);
dydt=[-s/ts + sjump*R; -w/tw + wjump*RI];
end
```

Figure 4: The code to run MFIII for a network of Izhikevich neurons

A.3 Mean-Field System for a Homogeneous Network with Noise

In order to numerically simulate the mean-field system for a network with noise, one has to compute the integral:

$$\langle R_i(t)\rangle_\sigma^{-1} = \frac{2}{\sigma^2}\int_{v_{reset}}^{v_{peak}}\int_{v}^{v_{peak}} \exp\left(-\frac{2}{\sigma^2}(M(v',\langle w\rangle,s) - M(v,\langle w\rangle,s))\right) dv'dv \tag{92}$$

as a function of s and $\langle w\rangle$ at each time step. This requires numerically computing a double integral over a triangular region in the v plane. As $\sigma \to 0$, the exponential term inside the integral often becomes difficult to work with due to the $\frac{1}{\sigma^2}$. However, by using the substituion $v' = v + \frac{\sigma^2}{2}z$, one arrives at the integral:

$$\begin{aligned}\langle R_i(t)\rangle_\sigma^{-1} &= \int_{v_{reset}}^{v_{peak}}\int_0^{\frac{2}{\sigma^2}(v_{peak}-v)} \exp\left(-\frac{2}{\sigma^2}\left[M\left(v+\frac{\sigma^2 z}{2},\langle w\rangle,s\right) - M(v,\langle w\rangle,s)\right]\right) dzdv \\ &= \int_{v_{reset}}^{v_{peak}}\int_0^{\frac{2}{\sigma^2}(v_{peak}-v)} \exp\left(-\left[\sum_{i=1}^{\infty}\frac{\partial^i M(v,\langle w\rangle,s)}{\partial v^i} z^i\left(\frac{\sigma^2}{2}\right)^{i-1}\right]\right) dzdv\end{aligned}$$

which can be converted into a rectangular region of integration by using the Heaviside function:

$$\int_{v_{reset}}^{v_{peak}}\int_0^{\frac{2}{\sigma^2}(v_{peak}-v_{reset})} \exp\left(-\left[\sum_{i=1}^{\infty}\frac{\partial^i M(v,\langle w\rangle,s)}{\partial v^i} z^i\left(\frac{\sigma^2}{2}\right)^{i-1}\right]\right) H\left(\frac{2}{\sigma^2}(v_{peak}-v) - z\right) dzdv$$

Note that the term inside the exponential no longer has a $\frac{2}{\sigma^2}$ term which yields numerical difficulties in the $\sigma \to 0$ limit. While the bounds of the integral now diverge as the upper bound now has a $\frac{\sigma^2}{2}$, the integrand converges to zero for large z exponentially fast. For the Izhikevich and quartic integrate and fire models, there is only a finite number of terms in the sum, as $F(v)$ and thus $M(v)$ is a polynomial in v. For other models, one can take a finite number of terms to approximate the firing rate. The remaining integral can be simply computed with the two-dimensional trapezoidal method over a rectangular region. The Matlab function trapz is used to compute the integral at each time step over a two-dimensional finite mesh in the v' and z variables. This is used for both direct simulation of the mean-field system and numerical bifurcation analysis of the system in MATCONT. Note that this implementation is similar to the one suggested in [24] only we compute the firing rate at each time step, as there does not appear to be much computational overhead in this approach versus using a function table, as first suggested in [24].

```
function dy = MFII(alpha,g,I,b,er,vpeak,vreset,ts,tw,sjump,wjump,sigma,x,z,t,y)
Irh = alpha^2/4;
s = y(1);
w = y(2);
epsilon = 0.5*sigma^2;
vmax = vpeak;
vmin = vreset;
%% Trapezoidal method is faster, but less accurate.
This is for the Izhikevich neuron.
%  df = x.^2-(alpha+g*s)*x - w + I + g*s*er;
%  d2f = 2*x - (alpha+g*s);
%  d3f = 2;
% R = 1/(trapz(x1,trapz(z1,exp(-z.*df - (z.^2).*d2f*epsilon/2
%%- (z.^3)*d3f*(epsilon^2)/6))));
% meanv = trapz(x1,R*x1.*trapz(z1,exp(-z.*df
%%- (z.^2).*d2f*epsilon/2 - (z.^3)*d3f*(epsilon^2)/6)));
R = 1/quad2d(@(x,z) exp(-z.*(x.^2-(alpha+g*s)*x - w + I + g*s*er)
- (z.^2).*(2*x - (alpha+g*s))*epsilon/2
- (z.^3)*(2)*(epsilon^2)/6),vreset,vpeak,0,1000);
meanv = quad2d(@(x,z) R*x.*exp(-z.*(x.^2-(alpha+g*s)*x - w + I + g*s*er)
 - (z.^2).*(2*x - (alpha+g*s))*epsilon/2
 - (z.^3)*(2)*(epsilon^2)/6),vreset,vpeak,0,1000);
dy(1) = -y(1)/ts + sjump*R;
dy(2) = -y(2)/tw + b*meanv/tw +  wjump*R;

end
```

B Genericity Conditions

Here, I will derive the genericity conditions for the smooth bifurcations displayed by the Izhikevich system (Hopf and Saddle-Node Bifurcations). As a first step, we need to rewrite the expression $\sqrt{I - I^*(s,w)}$. In particular, note that

$$\begin{aligned}
\sqrt{I - I^*(s,w)} &= \sqrt{\frac{1}{\lambda_s^2}s^2 + s\eta - w - (A_2(g)s^2 + A_1(g)s + A_0)} && (93)\\
&= \frac{1}{\lambda_s}\sqrt{s^2 + s\eta - w - \lambda_s^2 A_2(g)\left(\left(s + \frac{A_1(g)}{2A_2(g)}\right)^2 - \left(\frac{A_1(g)^2}{4A_2(g)^2} - A_0\right)\right)} && (94)\\
&= \frac{1}{\lambda_s}\sqrt{s^2 + s\eta - w - \lambda_s^2 A_2(g)\left((s-\beta)^2 - (\beta^2 + \tilde{I})\right)} && (95)
\end{aligned}$$

and the dynamics are thus given by

$$\begin{aligned}
s' &= -\frac{s}{\tau_s} + \frac{1}{\tau_s}\sqrt{s^2 + s\lambda_w\lambda_s - w\lambda_s^2 - \lambda_s^2 A_2(g)\left((s-\beta)^2 - (\beta^2 + \tilde{I})\right)} && (96)\\
w' &= -\frac{w}{\tau_w} - \frac{\eta}{\tau_w}\sqrt{s^2 + s\lambda_w\lambda_s - w\lambda_s^2 - \lambda_s^2 A_2(g)\left((s-\beta)^2 - (\beta^2 + \tilde{I})\right)} && (97)
\end{aligned}$$

B.1 Saddle-Node Genericity Conditions

Note that

$$\beta(g) = -\frac{A1(g)}{2A_2(g)} = \frac{g(er - v^*(0)) - \eta}{\frac{1}{\lambda_s^2} + \frac{v^{*\prime}(0)g^2}{2}} = M(g)(g - g^*)$$

and this transformation is locally invertible near g^* as $\beta(g^*) = M(g^*) > 0$ and so we are fine taking the parameter transformation $(g, I) \to (\beta, \tilde{I})$ locally near g^* and we can write down $A_2(g) = \tilde{A}_2(\beta)$ and again, we will drop the tilde for convenience on I and $A_2(\beta)$. We will treat I as our bifurcation parameter with the saddle-node bifurcation occuring when $I = -\beta^2$. We have used Maple to aid in computing the two non-degeneracy conditions for Sotomayors theroem. First, we need the Jacobian of (96)-(97)

Now, the Jacobian evaluated at $I = -\beta^2$, $(s,w) = (\beta, \beta\frac{\lambda_w}{\lambda_s})$

$$A = DF_{\pm}(0,0) = \begin{pmatrix} \frac{1}{2}\frac{\lambda_w\lambda_s}{\tau_s\beta} & -\frac{1}{2}\frac{\lambda_s^2}{\tau_s\beta} \\ \frac{\eta}{2}\frac{\lambda_s\lambda_w + 2\beta}{\tau_w\beta} & -\frac{1}{2}\frac{\eta\lambda_s^2 + 2\beta}{\tau_w\beta} \end{pmatrix} \quad (98)$$

The eigenvectors v and w associated with the eigenvalue 0 of A and A^T are given by:

$$v = \begin{pmatrix} \frac{\lambda_s}{\lambda_w} \\ 1 \end{pmatrix} \tag{99}$$

$$w = \begin{pmatrix} -\frac{(\lambda_s\lambda_w+2\beta)}{\lambda_s^2}\frac{\tau_s}{\tau_w} \\ 1 \end{pmatrix} \tag{100}$$

The first of Sotomayors criterions is that $w^T f_I(s,w)\Big|_{(\beta,\eta\beta,-\beta^2)} \neq 0$. To compute it, we require $F_I(s,w)$ which is given by the matrix

$$\frac{\partial F}{\partial I} = \begin{pmatrix} \frac{1}{2}\frac{\lambda_s^2 A_2(\beta)}{\tau_s\beta} \\ \frac{1}{2}\frac{\lambda_w\lambda_s A_2(\beta)}{\tau_w\beta} \end{pmatrix} \tag{101}$$

As a note, throughout all the derivations the term $|\beta|$ shows up, but as we are only considering $|\beta| > 0$, $g > g^*$ then we are fine just writing it as β. This gives us the first of Sotomayors criteria:

$$w^T F_I(x_0,\mu_0) = -\frac{1}{2}\frac{(\lambda_s\lambda_w+2\beta)}{\lambda_s^2}\frac{\tau_s}{\tau_w}\frac{\lambda_s^2 A_2(\beta)}{\tau_s\beta} + \frac{1}{2}\frac{\lambda_w\lambda_s A_2(\beta)}{\tau_w\beta} \tag{102}$$

$$= -\frac{1}{2}\frac{\lambda_s\lambda_w A_2(\beta)}{\tau_w\beta} - \frac{A_2(\beta)}{\tau_w} + \frac{1}{2}\frac{\lambda_w\lambda_s A_2(\beta)}{\tau_w\beta} \tag{103}$$

$$= -\frac{A_2(\beta)}{\tau_w} \tag{104}$$

Now, fortunately $A_2(\beta) > 0$,and so the first of Sotomayors condition is satisfied. For the second condition, we have

$$w^T\frac{\partial^2 F}{\partial x^2}(v,v) = w^T\begin{pmatrix} v^T\frac{\partial^2 F_1}{\partial x^2}v \\ v^T\frac{\partial^2 F_2}{\partial x^2}v \end{pmatrix}$$

The two necessary Hessian matrices evaluated at the saddle-node point are

$$\mathrm{Hess}(F_1) = \begin{pmatrix} -\frac{1}{4}\frac{\lambda_s(4A_2(\beta)\beta^2\lambda_s+\lambda_s\lambda_w^2+4\lambda_w\beta)}{\tau_s\beta^3} & \frac{1}{4}\frac{\lambda_s^2(\lambda_s\lambda_w+2\beta)}{\tau_s\beta^3} \\ \frac{1}{4}\frac{\lambda_s^2(\lambda_s\lambda_w+2\beta)}{\tau_s\beta^3} & -\frac{1}{4}\frac{\lambda_s^4}{\tau_s\beta^3} \end{pmatrix} \tag{105}$$

$$\mathrm{Hess}(F_2) = \begin{pmatrix} -\frac{1}{4}\frac{\lambda_w(4A_2(\beta)\beta^2\lambda_s+\lambda_s\lambda_w^2+4\lambda_w\beta)}{\tau_w\beta^3} & \frac{1}{4}\frac{\lambda_s\lambda_w(\lambda_s\lambda_w+2\beta)}{\tau_s\beta^3} \\ \frac{1}{4}\frac{\lambda_s\lambda_w(\lambda_s\lambda_w+2\beta)}{\tau_w\beta^3} & -\frac{1}{4}\frac{\lambda_w\lambda_s^3}{\tau_w\beta^3} \end{pmatrix} \tag{106}$$

$$\tag{107}$$

which when evaluated with their respective quadratic forms in v we have

$$\begin{pmatrix} -\frac{\lambda_s^4 A_2(\beta)}{\beta\lambda_w^2\tau_s} & -\frac{\lambda_s^3 A_2(\beta)}{\beta\tau_w\lambda_w} \end{pmatrix}$$

and this finally yields the the second Sotomayor condition

$$w^T \frac{\partial^2 F}{\partial x^2}(v,v) = -\frac{(\lambda_s\lambda_w + 2\beta)}{\lambda_s^2}\frac{\tau_s}{\tau_w}\cdot\left(-\frac{\lambda_s^4 A_2(\beta)}{\beta\lambda_w^2\tau_s}\right) - \frac{\lambda_s^3 A_2(\beta)}{\beta\tau_w\lambda_w} \tag{108}$$

$$= \frac{\lambda_s^3 A_2(\beta)}{\tau_w\beta\lambda_w} + 2\frac{\lambda_s^2}{\lambda_w^2\tau_w}A_2(\beta) - \frac{\lambda_s^3 A_2(\beta)}{\beta\tau_w\lambda_w} \tag{109}$$

$$= 2\frac{\lambda_s^2}{\lambda_w^2\tau_w}A_2(\beta) \tag{110}$$

which, again, thanks to $A_2(\beta) > 0$, we are fine here. Thus, as both of Sotomayors condition are strictly non-zero, the saddle-node bifurcation is generic provided that $\beta > 0$ (as other wise none of these matrices are defined).

B.2 Hopf Genericity Conditions

We have primarily used Maple to derive the first Lyapunov coefficient for the branch of Hopf Bifurcations for the Izhikebich model. Unfortunately, the end result of computing the Lyapunov coefficient using the formulas from [74] is too lengthy and complicated to determine the sign directly. As opposed to directly determining its sign, we will work in asymptotic regimes. Additionally, we will return to the original parameter set We will primarily work with $\gamma = \tau_s/\tau_w$ as our perturbation parameter. We will see that for both $\gamma \to 0$ and $\gamma \to \infty$, $l_1(0) > 0$ if $e_r > v^*(0)$ and if $z = g - \bar{g} > 0$. The asymptotic relationships are:

$$l_1(0) = \frac{\sqrt{2}}{16}\frac{(z^2+\lambda_s^2+2)^{5/2}(z^2\lambda_s\lambda_w + z\lambda_s(er - v^*(0)) + \lambda_w)}{\lambda_s\sqrt{(er-v^*(0))z}^3\lambda_w^{3/2}}\frac{1}{\gamma^{5/2}} \tag{111}$$

$$+ \frac{3}{16}\frac{(z^2/2+\lambda_s^2+2)^3}{\lambda_s^2 z^2(er-v^*(0))\lambda_w^2}\frac{1}{\gamma^2} \quad \gamma \to 0 \tag{112}$$

$$l_1(0) = \frac{3}{16}\frac{\lambda_w^2}{(er-v^*(0))^6\lambda_s^2 z^2}\gamma^3 \tag{113}$$

$$+ \frac{3}{2}\frac{(4(e_r - v^*(0))^2 + \lambda_w^2/2 + 2z\lambda_s\lambda_w(er - v^*(0)))}{(er-v^*(0))^6 z^2\lambda_s^2}\gamma^2 + O(\gamma) \quad \gamma \to \infty \tag{114}$$

Thus, to both leading orders, the first Lyapunov coefficients can be seen to be positive provided that $er - v^*(0) > 0$, and $z = g - \bar{g} > 0$.

C The First Two Moments of $C(\gamma_N)$

In this section, we will prove the first two moments of $C(\gamma_N)$ have the following asymptotics

$$E(C(\gamma_N)) \leq O(N^-2), \quad E(C(\gamma_N)^2) \leq O(N^{-2}) \tag{115}$$

which immediately implies that

$$E((C(\gamma_N) - E(C(\gamma_N)))^2) \leq O(N^{-2})$$

and that as $N \to \infty$, $C(\gamma_N) \to 0$ in a mean-squared. We will perform this proof with $C(\gamma_N)$ for only the population of ON neurons, with a similar result holding for the population of OFF neurons and both populations together.

C.1 The First Moment of $C(\gamma_N)$

Recall that

$$C(\gamma_N) = \int_{-1}^{1} \left(\frac{1}{N} \sum_{i=1}^{N} \gamma(a_i) \sqrt{x - a_i} - g^+(x) \right)^2 dx + \frac{\lambda}{N} \sum_{i=1}^{N} \frac{\gamma(a_i)^2}{N} \tag{116}$$

We will now write down the following

$$e_N(x) = \frac{1}{N} \sum_{i=1}^{N} \gamma(a_i) \sqrt{x - a_i} - g^+(x) \tag{117}$$

Based on equation (7.65), then we know that

$$E(e_N(x)) = 0, \quad E(e_N(x)^2) = \frac{1}{N} e(x) \tag{118}$$

where

$$e(x) = \int_{-1}^{x} \gamma^+(a)^2 \rho_a(a) f(x-a)^2 \, da - g^+(x)^2$$

The condition that

$$\int_{-1}^{1} \gamma^+(a)^2 \rho_a(a) < \infty$$

implies that $e(x)$ is bounded if $f(x-a) < F$ for some F:

$$e(x) \leq \int_{-1}^{x} \gamma^+(a) f(x-a)^2 \, da \leq F^2 \int_{-1}^{x} \gamma^+(a)^2 \rho_a(a) \, da \leq F^2 E(\gamma(a_i)^2)$$

Technically this is only true if $f(x-a)$ is bounded, but for the physical reason that neurons have finite firing rates, we never have to consider this case. Now, we can return to the quantity $e_N(x)$, and note that as it is a linear combination of continuous functions in x, it is also bounded. Additionally, as we are on the closed interval $[-1, 1]$, it attains its upper bound at some point (possibly non-unique), x^* such that $e_N(x) < e_N(x^*)$. Thus,

$$\begin{aligned} C(\gamma_N) &= \int_{-1}^{1} e_N(x)^2 \, dx + \frac{\lambda}{N} \sum_{i=1}^{N} \frac{\gamma(a_i)^2}{N} && (119) \\ &\leq 2 e_N(x^*)^2 + \frac{\lambda}{N} \sum_{i=1}^{N} \frac{\gamma(a_i)^2}{N} && (120) \\ E(C(\gamma_N)) &\leq 2E(e_N(x^*)^2) + \frac{\lambda}{N} E(\gamma(a_i)^2) = N^{-1}(2F^2 + \lambda) E(\gamma(a_i)^2) && (121) \end{aligned}$$

C.2 The Second Moment of $C(\gamma_N)$

The second moment is be computed by taking the expectation of $C(\gamma_N)^2$. Note the following:

$$\begin{aligned} C(\gamma_N)^2 &= \left(\int_{-1}^{1} e_N(x)^2\,dx + \frac{\lambda}{N}\sum_{i=1}^{N}\frac{\gamma(a_i)^2}{N}\right)^2 && (122)\\ &= \left[\int_{-1}^{1} e_N(x)^2\,dx\right]^2 + 2\int_{-1}^{1} e_N(x)\,dx\left(\frac{\lambda}{N}\sum_{i=1}^{N}\frac{\gamma(a_i)^2}{N}\right) + \frac{\lambda^2}{N^4}\left(\sum_{i=1}^{N}\gamma(a_i)^4 + \sum_{i\neq j}\gamma(a_i)^2\gamma(a_j)^2\right) \\ &\leq 4e_N(x^*)^4 + 4e_N(x^*)^2\frac{\lambda}{N}\left(\sum_{i=1}^{N}\frac{\gamma(a_i)^2}{N}\right) + \frac{\lambda^2}{N^4}\left(\sum_{i=1}^{N}\gamma(a_i)^4 + \sum_{i\neq j}\gamma(a_i)^2\gamma(a_j)^2\right) && (123)\\ &\leq 4e_N(x^*)^4 + \frac{4\lambda}{N}e_N(x^*)^2\gamma^2 + \frac{\lambda^2}{N^4}\left(\sum_{i=1}^{N}\gamma(a_i)^4 + \sum_{i\neq j}\gamma(a_i)^2\gamma(a_j)^2\right) && (124)\end{aligned}$$

where $\gamma(a) < \gamma, a \in [-1,1]$. Taking the expectation yields

$$\begin{aligned} E(C(\gamma_N)^2) &\leq 4E(e_N(x^*)^4) + \frac{4\lambda\gamma^2}{N}E(e_N(x^*)^2) + \frac{\lambda^2}{N^3}\left(E(\gamma(a_i)^4) + \frac{N-1}{N}E(\gamma(a_i)^2)^2\right) \\ &\leq 4E(e_N(x^*)^4) + \frac{4F^2\lambda\gamma^2}{N^2}E(\gamma(a_i)^2) + \frac{\lambda^2}{N^3}\left(E(\gamma(a_i)^4) + \frac{N-1}{N}E(\gamma(a_i)^2)^2\right) \end{aligned} \quad (125)$$

Now, as $e_N(x^*)$ is a normally distributed random variable asymptotically (by the central limit theorem) with mean 0 and $\sigma^2 = e(x^*)/N$. We have

$$E(e_N(x^*)^4) \sim 3E(e_N(x^*)^2)^2 \leq 3\left(\frac{F^2}{N}E(\gamma(a_i)^2)\right)^2 \quad (126)$$

based on the unique property of the normal distribution that all higher order moments are defined by the first two moments. Finally we have

$$\begin{aligned} E(C(\gamma_N)^2) &\leq \frac{12F^4}{N^2}E(\gamma(a_i)^2)^2 + \frac{4F^2\lambda\gamma^2}{N^2}E(\gamma(a_i)^2) + \frac{\lambda^2}{N^3}\left(E(\gamma(a_i)^4) + \frac{N-1}{N}E(\gamma(a_i)^2)^2\right) \\ &\sim O(N^{-2}) \end{aligned} \quad (127)$$

An important thing to note is that the assumption on $\gamma(a)$ being bounded immediately yields the existence of $E(\gamma(a_i)^k)$ for all k, while we only need the existence of the moment for $k = 4$ at most in order to prove that $E(C(\gamma_N)^2) \leq O(N^{-2})$, and thus there is some room to loosen the conditions on this proof.

D Piecewise Smooth Continuous Dynamics with Scale-Invariant Decoders

In this section, we will show how we can approximate the piecewise defined function

$$h(x) = \begin{cases} h_L(x) & x < b \\ h_R(x) & x > b \end{cases}. \tag{128}$$

Consider a population of ON neurons with the following weighted decoder:

$$\hat{P}(a) = \begin{cases} P_L(a) & a < 0 \\ P_R(a) & a > 0 \end{cases}$$

which yields

$$\hat{g}(x) = \begin{cases} \int_{-1}^{x} \hat{P}_L(a)\sqrt{x-a}\,da & -1 < x < 0 \\ \int_{-1}^{0} \hat{P}_L(a)\sqrt{x-a}\,da + \int_{0}^{x} \hat{P}_R(a)\sqrt{x-a} & 0 < x < 1 \end{cases}$$

$\hat{g}(x)$ is continuous at $x = 0$, and so we cannot compute a non-smooth function analytically with this type of weighted decoder. If we attempt to use a δ pulse for $\hat{P}(a)$, $\hat{P}(a) = \delta(a)\gamma(a)$ then all we arrive at is

$$\hat{g}(x) = \gamma(0)\sqrt{x} = k\sqrt{x}$$

and so we can get the $\sqrt{x}$ function. From the preceding argument however, it should be clear that we can get an arbitrary piecewise smooth continuous function. In fact, one can be constructed rather easily. Consider the two operators

$$\hat{g}_R^+(x) = \int_0^x \hat{P}_R^+(a)\sqrt{x-a} \quad 0 < x < 1 \tag{129}$$

$$\hat{g}_L^-(x) = \int_0^{-x} \hat{P}_L^-(a)\sqrt{-x-a} \quad -1 < x < 0 \tag{130}$$

The first thing to realize is that with these two operators, we can approximate any function that vanishes to first order at $x = 0$. For example, we have

$$\hat{P}_R^+(a) = \frac{2}{\pi}\frac{g_R'(0)}{\sqrt{a}} + \int_0^a \frac{2}{\pi\sqrt{t}} g_R{}''(a-t)\,dt \tag{131}$$

$$\hat{P}_L^-(a) = \frac{2}{\pi}\frac{g_L'(0)}{\sqrt{a}} + \int_0^a \frac{2}{\pi\sqrt{t}} g_L{}''(t-a)\,dt \tag{132}$$

Note that $g_R^+(x)$ and $g_L^-(x)$ both equal 0 when $x = 0$ and thus we can compute a piecewise smooth continuous function that is 0 on the switching boundary. To compute a function that is non-zero on the switching boundary, we can merely use the following:

$$\hat{P}^{+}(a) = \begin{cases} \frac{1}{\pi}\frac{g(0)}{\sqrt{1+a}} & a<0 \\ \frac{1}{\pi}\frac{g(0)}{\sqrt{1+a}} + \frac{2}{\pi}\frac{g_R'(0)}{\sqrt{a}} + \int_0^a \frac{2}{\pi\sqrt{t}} g_R{}''(a-t)\,dt & 0<a<1 \end{cases} \tag{133}$$

$$\hat{P}^{-}(a) = \begin{cases} \frac{1}{\pi}\frac{g(0)}{\sqrt{1+a}} & a<0 \\ \frac{1}{\pi}\frac{g(0)}{\sqrt{1+a}} + \frac{2}{\pi}\frac{g_L'(0)}{\sqrt{a}} + \int_0^a \frac{2}{\pi\sqrt{t}} g_L{}''(t-a)\,dt & 0<a<1 \end{cases} \tag{134}$$

where $g(0) = g_L(0) = g_R(0)$. This allows us to compute an arbitrary piecewise smooth-continuous function. One can also approximate arbitrary non-smooth functions. To do this, we will take any sigmoid that converges to the heaviside function. For example, we can take the error function

$$g(x) = \frac{1}{2} + \frac{1}{2}\text{erf}(kx) \quad k \gg 1 \tag{135}$$

and compute the resulting $\hat{P}^{\pm}(a)$ as a standard function. Linearity in our operators implies that with the $\hat{P}^{\pm}$ given by equations (133)-(134) allows us to merely add on the $\hat{P}^{\pm}$ for the function $g(x)$.

E Non-Uniqueness of Multi-variable Scale-Invariant Decoders

Consider the two-dimensional function $g(x,y) = xy$. With a two-dimensional polar coordinate system, we will assume all the decoding vectors are evenly distributed as follows:

$$\rho_\theta(\theta) = \frac{1}{4}\delta\left(\theta - \frac{\pi}{4}\right) + \frac{1}{4}\delta\left(\theta - \frac{3\pi}{4}\right) + \frac{1}{4}\delta\left(\theta - \frac{5\pi}{4}\right) + \frac{1}{4}\delta\left(\theta - \frac{7\pi}{4}\right)$$

and thus, $\hat{P}(\theta)$ must also be δ valued:

$$\hat{P}_\theta(\theta) = \frac{A}{4}\delta\left(\theta - \frac{\pi}{4}\right) + \frac{B}{4}\delta\left(\theta - \frac{3\pi}{4}\right) + \frac{C}{4}\delta\left(\theta - \frac{5\pi}{4}\right) + \frac{D}{4}\delta\left(\theta - \frac{7\pi}{4}\right)$$

where A, B, C, and D are yet to be determined and so we have the following:

$$\begin{aligned} \int_0^{2\pi}\int_{-1}^{\cos(\theta)x+\sin(\theta)y} & \hat{P}_a(a)\hat{P}_\theta(\theta)\sqrt{\cos(\theta)x+\sin(\theta)y-a}\,da\,d\theta \\ &= \frac{A}{4}\int_{-1}^{x/\sqrt{2}+y/\sqrt{2}+1} \hat{P}_a(a)\sqrt{\frac{x}{\sqrt{2}}+\frac{y}{\sqrt{2}}-a}\,da \\ &+\frac{B}{4}\int_{-1}^{-x/\sqrt{2}+y/\sqrt{2}+1} \hat{P}_a(a)\sqrt{-\frac{x}{\sqrt{2}}+\frac{y}{\sqrt{2}}-a}\,da \\ &+\frac{C}{4}\int_{-1}^{-x/\sqrt{2}-y/\sqrt{2}+1} \hat{P}_a(a)\sqrt{-\frac{x}{\sqrt{2}}-\frac{y}{\sqrt{2}}-a}\,da \\ &+\frac{D}{4}\int_{-1}^{x/\sqrt{2}-y/\sqrt{2}+1} \hat{P}_a(a)\sqrt{\frac{x}{\sqrt{2}}-\frac{y}{\sqrt{2}}-a}\,da \end{aligned} \tag{136}$$

Now, from our previous work, we know that if we take

$$\hat{P}_a(a) = \frac{8}{\pi}\sqrt{a+1},$$

then we have the following:

$$\begin{aligned}
\hat{g}(x,y) &= \frac{A}{4}\left(\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} + 1\right)^2 + \frac{B}{4}\left(-\frac{x}{\sqrt{2}} + \frac{y}{\sqrt{2}} + 1\right)^2 + \frac{C}{4}\left(-\frac{x}{\sqrt{2}} - \frac{y}{\sqrt{2}} + 1\right)^2 \\
&+ \frac{D}{4}\left(\frac{x}{\sqrt{2}} - \frac{y}{\sqrt{2}} + 1\right)^2 \\
&= \frac{x^2}{8}(A+B+C+D) + \frac{y^2}{8}(A+B+C+D) + \frac{xy}{4}(A-B+C-D) \\
&+ \frac{\sqrt{2}x}{4}(A-B-C+D) + \frac{\sqrt{2}y}{4}(A+B-C-D) + \frac{1}{4}(A+B+C+D)
\end{aligned}$$

now, while we have 6 terms in the polynomial, three of the coefficients that are unnecessary are identical. This yields the four linear equations:

$$\begin{aligned}
A+B+C+D &= 0 \\
A-B-C+D &= 0 \\
A+B-C-D &= 0 \\
A-B+C-D &= 4
\end{aligned}$$

which yields $\hat{g}(x,y) = xy$ when $A = C, B = D, A = -B = 1$. And thus, the resulting scale invariant decoders are given by

$$\gamma(a) = \frac{8}{\pi}\sqrt{a+1} \tag{137}$$

for all 4 discrete sub-populations. This is one particular solution, however consider equation (136) but instead let us use a different $\hat{P}$ for all the discrete subpopulations. defining $z_1 = (x+y)/\sqrt{2}4$ and $z_2 = (x-y)/\sqrt{2}$, then we have $g(x,y) = h(z_1,z_2) = z_1^2 - z_2^2$ and

$$\hat{h}(z_1,z_2) = L_{z_1}^+(\hat{P}_{z_1}^+(a)) + L_{z_1}^-(\hat{P}_{z_1}^-(a)) + L_{z_2}^+(\hat{P}_{z_2}^+(a)) + L_{z_2}^-(\hat{P}_{z_2}^-(a)) \tag{138}$$

where the subscripts denote the independent variable in the function range. Given our previous discussion of non-uniqueness in the single variable case, it should be clear that we can use functions $\epsilon(z_1)$ and $\epsilon(z_2)$ where $\epsilon(z_1)$ is in the function range of both $L_{z_1}^{\pm}$ with an identical condition holding for $\epsilon(z_2)$ to create non-unique $\hat{P}^{\pm}$.

F The Recurrence Relationship for $(\langle e, x\rangle + 1)^n$

Now, in order to be able to approximate any arbitrary function, we will have to use a basis to basis mapping. The first thing to note is the following recurrence relationship

$$\begin{aligned}
A_n(x) &= \int_{-1}^{x} \frac{(1+a)^n}{2\sqrt{1+a}}\sqrt{x-a}\,da \\
&= \int_0^{\sqrt{x+1}} u^{2n}\sqrt{x+1-u^2}\,du \\
&= -\frac{1}{3}u^{2n-1}((x+1)-u^2)^{3/2}\Big|_0^{\sqrt{x+1}} + \int_0^{\sqrt{x+1}} \frac{2n-1}{3}(x+1-u^2)u^{2n-2}\sqrt{x+1-u^2}\,du \\
&= \frac{2n-1}{3}(x+1)\int_0^{\sqrt{x+1}} u^{2(n-1)}\sqrt{x+1-u^2}\,du - \frac{2n-1}{3}\int_0^{\sqrt{x+1}} u^{2n}\sqrt{x+1-u^2}\,du \\
&= \frac{2n-1}{3}(x+1)A_{n-1}(x) - \frac{2n-1}{3}A_n(x) \\
A_n(x) &= (x+1)\left(\frac{2n-1}{2n+2}\right)A_{n-1}(x)
\end{aligned}$$

for $n > 2$ Also note that

$$A_0(x) = \int_{-1}^{x} \frac{1}{2\sqrt{1+a}}\sqrt{x-a}\,da = \frac{\pi}{4}(x+1)$$

and thus if we initialize the recurrence with $B_0(x) = \frac{4}{\pi}A_0(x)$ we obtain the following:

$$\begin{aligned}
B_0(x) &= (x+1) \\
B_1(x) &= \frac{1}{4}(x+1)^2 \\
B_2(x) &= \frac{1}{8}(x+1)^3 \\
B_n(x) &= \frac{(2n-1)(2n-3)\ldots 5\cdot 3\cdot 1}{2^n(n+1)!}(x+1)^n
\end{aligned}$$

Now, the reason we are doing this is because it should be clear that we can approximate for arbitrary n a function of the form $(x+1)^n$. Then the integral (7.83) simplifies to:

$$\hat{g}(x) = \int_{\|\boldsymbol{e}\|=1} \hat{P}_e(\boldsymbol{e})(\langle \boldsymbol{e}, \boldsymbol{x}\rangle + 1)^n\,d\boldsymbol{e}$$

References

[1] L.F. Abbott. Lapicque's introduction of the integrate-and-fire model neuron (1907). *Brain research bulletin*, 50(5):303–304, 1999.

[2] L.F. Abbott and C. van Vreeswijk. Asynchronous states in networks of pulse-coupled oscillators. *Learning and Memory*, 48(2):1483–1490, 1993.

[3] L.F. Abbott, J.A. Varela, K. Sen, and S.B. Nelson. Synaptic depression and cortical gain control. *Science*, 275(5297):221–224, 1997.

[4] B. Ahmed, J.C. Anderson, R.J. Douglas, K.A. Martin, and D. Whitteridge. Estimates of the net excitatory currents evoked by visual stimulation of identified neurons in cat visual cortex. *Cerebral cortex*, 8(5):462–476, 1998.

[5] R. Ananthanarayanan and D.S. Modha. Anatomy of a cortical simulator. In *Proceedings of the 2007 ACM/IEEE conference on Supercomputing*, page 3. ACM, 2007.

[6] P. Andersen, R. Morris, D. Amaral, T. Bliss, and J. O'Keefe. *The hippocampus book*. Oxford University Press, USA, 2006.

[7] A.A. Andronov. *Theory of bifurcations of dynamic systems on a plane*, volume 554. Israel Program for Scientific Translations;[available from the US Dept. of Commerce, National Technical Information Service, Springfield, Va.], 1971.

[8] F. Apfaltrer, C. Ly, and D. Tranchina. Population density methods for stochastic neurons with realistic synaptic kinetics: Firing rate dynamics and fast computational methods. *Network: Computation in Neural Systems*, 17(4):373–418, 2006.

[9] M. Augustin, J. Ladenbauer, and K. Obermayer. How adaptation shapes spike rate oscillations in recurrent neuronal networks. *Frontiers in computational neuroscience*, 7, 2013.

[10] M.C. Avery and J.L. Krichmar. Improper activation of d1 and d2 receptors leads to excess noise in prefrontal cortex. *Frontiers in computational neuroscience*, 9, 2015.

[11] R. Azouz, C.M. Gray, L.G. Nowak, and D.A. McCormick. Physiological properties of inhibitory interneurons in cat striate cortex. *Cerebral Cortex*, 7(6):534–545, 1997.

[12] L. Badel, S. Lefort, R. Brette, C.C.H Petersen, W. Gerstner, and M.J.E. Richardson. Dynamic iv curves are reliable predictors of naturalistic pyramidal-neuron voltage traces. *Journal of Neurophysiology*, 99(2):656–666, 2008.

[13] S. Banach. Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales. *Fund. Math*, 3(1):133–181, 1922.

[14] E. Barreto, B. Hunt, E. Ott, and P. So. Synchronization in networks of networks: The onset of coherent collective behavior in systems of interacting populations of heterogeneous oscillators. *Physical Review E*, 77(3):036107, 2008.

[15] T. Bekolay, M. Laubach, and C. Eliasmith. A spiking neural integrator model of the adaptive control of action by the medial prefrontal cortex. *The Journal of Neuroscience*, 34(5):1892–1902, 2014.

[16] R. Ben-Yishai, R.L Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844–3848, 1995.

[17] C.M. Bender and S.A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers I*. Springer Science & Business Media, 1999.

[18] C.M. Bishop. *Neural networks for pattern recognition*. Oxford university press, 1995.

[19] B. Bobier, T.C. Stewart, and C. Eliasmith. A unifying mechanistic model of selective attention in spiking neurons. 2014.

[20] M.M. Boerlin, C.K. Machens, and S. Deneve. Predictive coding of dynamical variables in balanced spiking networks. *PLoS Comput Biol*, 9(11):e1003258, 2013.

[21] C. Börgers and N. Kopell. Synchronization in networks of excitatory and inhibitory neurons with sparse, random connectivity. *Neural computation*, 15(3):509–538, 2003.

[22] R. Brette. Exact simulation of integrate-and-fire models with synaptic conductances. *Neural Computation*, 18(8):2004–2027, 2006.

[23] R. Brette and W. Gerstner. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *Journal of Neurophysiology*, 94(5):3637–3642, 2005.

[24] N. Brunel and P.E. Latham. Firing rate of the noisy quadratic integrate-and-fire neuron. *Neural Computation*, 15(10):2281–2306, 2003.

[25] N. Brunel and M.C.W. Van Rossum. Lapicque's 1907 paper: from frogs to integrate-and-fire. *Biological cybernetics*, 97(5-6):337–339, 2007.

[26] A.N. Burkitt, H. Meffin, and D.B. Grayden. Study of neuronal gain in a conductance-based leaky integrate-and-fire neuron model with balanced excitatory and inhibitory synaptic input. *Biological cybernetics*, 89(2):119–125, 2003.

[27] R.J. Butera, J. Rinzel, and J.C. Smith. Models of respiratory rhythm generation in the pre-bötzinger complex. i. bursting pacemaker neurons. *Journal of neurophysiology*, 82(1):382–397, 1999.

[28] R.J. Butera, J. Rinzel, and J.C. Smith. Models of respiratory rhythm generation in the pre-bötzinger complex. ii. populations of coupled pacemaker neurons. *Journal of Neurophysiology*, 82(1):398–415, 1999.

[29] G. Buzsáki. Large-scale recording of neuronal ensembles. *Nature neuroscience*, 7(5):446–451, 2004.

[30] C.A. Buzzi, P.R. da Silva, and M.A. Teixeira. A singular approach to discontinuous vector fields on the plane. *Journal of Differential Equations*, 231(2):633–655, 2006.

[31] S.A. Campbell. Advanced ordinary differential equations. Course Notes for AMATH751, 2011.

[32] J. Carr. *Applications of centre manifold theory*, volume 35. Springer Science & Business Media, 2012.

[33] F.S. Chance, L.F. Abbott, and A.D. Reyes. Gain modulation from background synaptic input. *Neuron*, 35(4):773–782, 2002.

[34] C. Clopath, L. Büsing, E. Vasilaki, and W. Gerstner. Connectivity reflects coding: a model of voltage-based stdp with homeostasis. *Nature neuroscience*, 13(3):344–352, 2010.

[35] C. Clopath and W. Gerstner. Voltage and spike timing interact in stdp–a unified model. *Frontiers in synaptic neuroscience*, 2, 2010.

[36] A. Colombo, M. Di Bernardo, S.J. Hogan, and M.R. Jeffrey. Bifurcations of piecewise smooth flows: Perspectives, methodologies and open problems. *Physica D: Nonlinear Phenomena*, 241(22):1845–1860, 2012.

[37] J. Conklin and C. Eliasmith. A controlled attractor network model of path integration in the rat. *Journal of computational neuroscience*, 18(2):183–203, 2005.

[38] J.A. Connor and C.F. Stevens. Prediction of repetitive firing behaviour from voltage clamp data on an isolated neurone soma. *The Journal of Physiology*, 213(1):31–53, 1971.

[39] R.M. Corless, G.H. Gonnet, D.E.G. Hare, D.J. Jeffrey, and D.E. Knuth. On the lambertw function. *Advances in Computational mathematics*, 5(1):329–359, 1996.

[40] R. Curtu and G.B. Ermentrout. Oscillations in a refractory neural net. *Journal of mathematical biology*, 43(1):81–100, 2001.

[41] P. Dayan and LF Abbott. Theoretical neuroscience: computational and mathematical modeling of neural systems. *Journal of Cognitive Neuroscience*, 15(1):154–155, 2003.

[42] F. Della Rossa and F. Dercole. The transition from persistence to nonsmooth-fold scenarios in relay control system. In *Proceedings of the 18th IFAC World Congress, Milano*, 2011.

[43] A. Destexhe, Z.F. Mainen, and T.J. Sejnowski. Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *Journal of computational neuroscience*, 1(3):195–230, 1994.

[44] A. Destexhe, Z.F. Mainen, and T.J. Sejnowski. Kinetic models of synaptic transmission. *Methods in neuronal modeling*, 2:1–25, 1998.

[45] A. Destexhe, Z.F. Mainen, and T.J. Sejnowski. *Methods in Neuronal Modelling*. MIT–Press, Cambridge–MA, 1998.

[46] T. DeWolf and C. Eliasmith. The neural optimal control hierarchy for motor control. *Journal of neural engineering*, 8(6):065009, 2011.

[47] A. Dhooge, W. Govaerts, and Y.A. Kuznetsov. Matcont: a matlab package for numerical bifurcation analysis of odes. *ACM Transactions on Mathematical Software (TOMS)*, 29(2):141–164, 2003.

[48] M. Di Bernardo, C.J. Budd, A.R. Champneys, P. Kowalczyk, A.B. Nordmark, G.O Tost, and P.T. Piiroinen. Bifurcations in nonsmooth dynamical systems. *SIAM review*, pages 629–701, 2008.

[49] M. Dur-e Ahmad, W. Nicola, S.A. Campbell, and F.K. Skinner. Network bursting using experimentally constrained single compartment ca3 hippocampal neuron models with adaptation. *Journal of computational neuroscience*, 33(1):21–40, 2012.

[50] G.M. Edelman and E.M. Izhikevich. Large-scale model of mammalian thalamocortical systems, year=2007, volume=5, number=9, pages= 3593 - 3598,. *PNAS*.

[51] C. Eliasmith. A unified approach to building and controlling spiking attractor networks. *Neural Computation*.

[52] C. Eliasmith and C.H. Anderson. *Neural engineering: Computation, representation, and dynamics in neurobiological systems*. MIT press, 2004.

[53] C. Eliasmith, T.C. Stewart, X. Choo, T. Bekolay, T. DeWolf, C. Tang, and D. Rasmussen. A large-scale model of the functioning brain. *Science*.

[54] B. Ermentrout. Reduction of conductance-based models with slow synapses to neural nets. *Neural Computation*, 6(4):679–695, 1994.

[55] G.B. Ermentrout. Type i membranes, phase resetting curves, and synchrony. *Neural computation*, 8(5):979–1001, 1996.

[56] G.B. Ermentrout. *Simulating, analyzing, and animating dynamical systems: a guide to XPPAUT for researchers and students*, volume 14. Siam, 2002.

[57] G.B. Ermentrout. Linearization of f-i curves by adaptation. *Neural Computation*, 10:1721–1729, 2006.

[58] G.B. Ermentrout and J.D. Cowan. Temporal oscillations in neuronal nets. *Journal of mathematical biology*, 7(3):265–280, 1979.

[59] G.B. Ermentrout and J.D. Cowan. Large scale spatially organized activity in neural nets. *SIAM Journal on Applied Mathematics*, 38(1):1–21, 1980.

[60] G.B. Ermentrout and N. Kopell. Parabolic bursting in an excitable system coupled with a slow oscillation. *SIAM Journal on Applied Mathematics*, 46:233–253, 1986.

[61] G.B. Ermentrout and D.H. Terman. *Mathematical foundations of neuroscience*, volume 35. Springer Science & Business Media, 2010.

[62] K.A. Ferguson, C.Y.L. Huh, B. Amilhon, S. Williams, and F.K. Skinner. Simple, biologically-constrained ca1 pyramidal cell models using an intact, whole hippocampus context. *F1000Research*, 3, 2014.

[63] K.A. Ferguson, F. NJap, W. Nicola, F.K. Skinner, and S.A Campbell. Examining the limits of cellular adaptation bursting mechanisms in biologically-based excitatory networks of the hippocampus. *Journal of Computational Neuroscience (To appear 2015)*.

[64] A.F. Filippov. Differential equations with discontinuous right-hand side. *Matematicheskii sbornik*, 93(1):99–128, 1960.

[65] R. FitzHugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal*, 1(6):445, 1961.

[66] N. Fourcaud and N. Brunel. Dynamics of the firing probability of noisy integrate-and-fire neurons. *Neural computation*, 14(9):2057–2110, 2002.

[67] E. Freire, E. Ponce, F. Rodrigo, and F. Torres. Bifurcation sets of continuous piecewise linear systems with two zones. *International Journal of Bifurcation and Chaos*, 8(11):2073–2097, 1998.

[68] A.P. Georgopoulos, J.T. Lurito, M. Petrides, A.B. Schwartz, and J.T. Massey. Mental rotation of the neuronal population vector. *Biology and Computation: A Physicist's Choice*, (3):183, 1994.

[69] A.P. Georgopoulos, A.B. Schwartz, and R.E. Kettner. Neuronal population coding of movement direction. *Science*, 233(4771):1416–1419, 1986.

[70] W. Gerstner and W.M. Kistler. *Spiking Neuron Models*. Cambridge University Press, 2002.

[71] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep sparse rectifier neural networks. In *International Conference on Artificial Intelligence and Statistics*, pages 315–323, 2011.

[72] B.F. Grewe, D. Langer, H. Kasper, B. Kampa, and F. Helmchen. High-speed in vivo calcium imaging reveals neuronal network activity with near-millisecond precision. *Nature methods*, 7(5):399–405, 2010.

[73] D.M. Grobman. Homeomorphism of systems of differential equations. *Doklady Akademii Nauk SSSR*, 128(5):880–881, 1959.

[74] J. Guckenheimer and P. Holmes. *Nonlinear oscillations, dynamical systems, and bifurcations of vector fields*, volume 42. Springer Science & Business Media, 1983.

[75] B.S. Gutkin and G.B. Ermentrout. Dynamics of membrane excitability determine interspike interval variability: a link between spike generation mechanisms and cortical spike train statistics. *Neural Computation*, 10(5):1047–1065, 1998.

[76] D. Hansel and G. Mato. Asynchronous states and the emergence of synchrony in large networks of interacting excitatory and inhibitory neurons. *Neural Computation*, 15(1):1–56, 2003.

[77] D. Hansel, G. Mato, C. Meunier, and L Neltner. On numerical simulations of integrate-and-fire neural networks. *Neural Computation*, 10(2):467–483, 1998.

[78] J. Harris and B. Ermentrout. Bifurcations in the wilson–cowan equations with nonsmooth firing rate. *SIAM Journal on Applied Dynamical Systems*, 14(1):43–72, 2015.

[79] P.M. Harrison. *Experimentally verified reduced models of neocortical pyramidal cells*. PhD thesis, University of Warwick, 2014.

[80] P.M Harrison, L. Badel, M.J. Wall, and M.J.E. Richardson. Experimentally verified parameter sets for modelling heterogeneous neocortical pyramidal-cell populations. *PLoS Comput Biol*, 11(8):e1004165, 2015.

[81] P. Hartman. On local homeomorphisms of euclidean spaces. *Bol. Soc. Mat. Mexicana (2)*, 5:220–241, 1960.

[82] S.S. Haykin. *Neural networks and learning machines*, volume 3. Pearson Education Upper Saddle River, 2009.

[83] S. Herculano-Houzel. The human brain in numbers: a linearly scaled-up primate brain. *Frontiers in human neuroscience*, 3, 2009.

[84] G. Hermann and J. Touboul. Heterogeneous connections induce oscillations in large-scale networks. *Physical review letters*, 109(1):018702, 2012.

[85] L. Hertäg, D.l Durstewitz, and N. Brunel. Analytical approximations of the firing rate of an adaptive exponential integrate-and-fire neuron in the presence of synaptic noise. *Frontiers in computational neuroscience*, 8, 2014.

[86] J.L. Hindmarsh and R.M. Rose. A model of neuronal bursting using three coupled first order differential equations. *Proceedings of the Royal Society of London B: Biological Sciences*, 221(1222):87–102, 1984.

[87] A.L. Hodgkin and A.F. Huxley. The dual effect of membrane potential on sodium conductance in the giant axon of loligo. *The Journal of physiology*, 116(4):497–506, 1952.

[88] J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, 79(8):2554–2558, 1982.

[89] E.M. Izhikevich. Simple model of spiking neurons. *Neural Networks, IEEE Transactions on*, 14(6):1569 – 1572, nov. 2003.

[90] E.M. Izhikevich. Simple model of spiking neurons. *Neural Networks, IEEE Transactions on*, 14(6):1569 – 1572, nov. 2003.

[91] E.M. Izhikevich. Which model to use for cortical spiking neurons? *IEEE transactions on neural networks*, 15(5):1063–1070, 2004.

[92] E.M. Izhikevich. *Dynamical systems in neuroscience*. MIT press, 2007.

[93] J. Jost and X. Li-Jost. *Calculus of variations*, volume 64. Cambridge University Press, 1998.

[94] JK Kevorkian and Julian D Cole. *Multiple scale and singular perturbation methods*, volume 114. Springer Science & Business Media, 2012.

[95] Z.P. Kilpatrick, B. Ermentrout, and B. Doiron. Optimizing working memory with heterogeneity of recurrent cortical excitation. *The Journal of Neuroscience*, 33(48):18999–19011, 2013.

[96] K.K. Kirli, G.B. Ermentrout, and R.Y. Cho. Computational study of nmda conductance and cortical oscillations in schizophrenia. *Frontiers in computational neuroscience*, 8, 2014.

[97] B.W. Knight. Dynamics of encoding in neuron populations: Some general mathematical features. *Neural Comput.*, 12:473–518, March 2000.

[98] R. Kobayashi, Y. Tsubo, and S. Shinomoto. Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. *Frontiers in computational neuroscience*, 3, 2009.

[99] A. Kolmogorov and S.V. Fomin. *Introductory real analysis*. Courier Corporation, 2012.

[100] Y.A. Kuznetsov. *Elements of applied bifurcation theory*, volume 112. Springer Science & Business Media, 2013.

[101] Y.A. Kuznetsov, S. Rinaldi, and A. Gragnani. One-parameter bifurcations in planar filippov systems. *International Journal of Bifurcation and chaos*, 13(08):2157–2188, 2003.

[102] G. La Camera, M. Giugliano, W. Senn, and S. Fusi. The response of cortical neurons to in vivo-like input current: theory and experiment. *Biological cybernetics*, 99(4-5):279–301, 2008.

[103] G. La Camera, A. Rauch, H. Lüscher, W. Senn, and S. Fusi. Minimal models of adapted neuronal response to in vivo–like input currents. *Neural computation*, 16(10):2101–2124, 2004.

[104] J. Ladenbauer, M. Augustin, L. Shiau, and K. Obermayer. Impact of adaptation currents on synchronization of coupled exponential integrate-and-fire neurons. *PLoS Comput Biol*, 8(4):e1002478–e1002478, 2012.

[105] L. Lapicque. Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation. *J. Physiol. Pathol. Gen*, 9(1):620–635, 1907.

[106] P.E. Latham, B.J. Richmond, P.G. Nelson, and S. Nirenberg. Intrinsic dynamics in neuronal networks. i. theory. *Journal of Neurophysiology*, 83(2):808–827, 2000.

[107] R.I. Leine and H. Nijmeijer. *Dynamics and bifurcations of non-smooth mechanical systems*, volume 18. Springer Science & Business Media, 2013.

[108] E.N. Lorenz. Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141, 1963.

[109] T.B. Luke, E. Barreto, and P. So. Complete classification of the macroscopic behavior of a heterogeneous network of theta neurons. *Neural computation*, 25(12):3207–3234, 2013.

[110] C. Ly. A principled dimension-reduction method for the population density approach to modeling networks of neurons with synaptic dynamics. *Neural computation*, 25(10):2682–2708, 2013.

[111] C. Ly and G.B. Ermentrout. Analysis of recurrent networks of pulse-coupled noisy neural oscillators. *SIAM Journal on Applied Dynamical Systems*, 9(1):113–137, 2010.

[112] C. Ly and G.B. Ermentrout. Analytic approximations of statistical quantities and response of noisy oscillators. *Physica D: Nonlinear Phenomena*, 240(8):719–731, 2011.

[113] C. Ly and D. Tranchina. Critical Analysis of Dimension Reduction by a Moment Closure Method in a Population Density Approach to Neural Network Modeling. *Neural Computation*, 19:2032–2092, 2007.

[114] H. Markram. The blue brain project. *Nature Reviews Neuroscience*, 7(2):153–160, 2006.

[115] E.A. Martens, E. Barreto, S.H. Strogatz, E. Ott, P. So, and T.M. Antonsen. Exact results for the kuramoto model with a bimodal frequency distribution. *Physical Review E*, 79(2):026204, 2009.

[116] MATLAB. *version 7.10.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts, 2014.

[117] D.A. McCormick, Y. Shu, and Y. Yu. Neurophysiology: Hodgkin and huxley model—still standing? *Nature*, 445(7123):E1–E2, 2007.

[118] W.S. McCulloch and W. Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[119] Sam McKennoch, Thomas Voegtlin, and Linda Bushnell. Spike-timing error backpropagation in theta neuron networks. *Neural computation*, 21(1):9–45, 2009.

[120] J.F. Mejias and A. Longtin. Optimal heterogeneity for coding in spiking neural networks. *Physical Review Letters*, 108(22):228102, 2012.

[121] J.F. Mejias, A. Payeur, E. Selin, L. Maler, and A. Longtin. Subtractive, divisive and non-monotonic gain control in feedforward nets linearized by noise and delays. *Frontiers in computational neuroscience*, 8, 2014.

[122] G. Menon and G. Haller. Infinite dimensional geometric singular perturbation theory for the maxwell–bloch equations. *SIAM journal on mathematical analysis*, 33(2):315–346, 2001.

[123] R.E. Mirollo and S.H. Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 50(6):1645–1662, 1990.

[124] J. Modolo and A. Beuter. Linking brain dynamics, neural mechanisms, and deep brain stimulation in parkinson's disease: an integrated perspective. *Medical engineering & physics*, 31(6):615–623, 2009.

[125] J. Modolo, A. Garenne, J. Henry, and A. Beuter. Development and validation of a neural population model based on the dynamics of a discontinuous membrane potential neuron model. *Journal of integrative Neuroscience*, 6(04):625–655, 2007.

[126] J. Modolo, J. Henry, and A. Beuter. Dynamics of the subthalamo-pallidal complex in parkinson's disease during deep brain stimulation. *Journal of biological physics*, 34(3-4):251–266, 2008.

[127] C. Morris and H. Lecar. Voltage oscillations in the barnacle giant muscle fiber. *Biophysical journal*, 35(1):193, 1981.

[128] J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962.

[129] R. Naud and W. Gerstner. The performance (and limits) of simple neuron models: Generalizations of the leaky integrate-and-fire model. In *Computational Systems Neurobiology*, pages 163–192. Springer, 2012.

[130] R. Naud, N. Marcille, C. Clopath, and W. Gerstner. Firing patterns in the adaptive exponential integrate-and-fire model. *Biological Cybernetics*, 99:335–347, 2008.

[131] B. Naundorf, F. Wolf, and M. Volgushev. Unique features of action potential initiation in cortical neurons. *Nature*, 440(7087):1060–1063, 2006.

[132] W.H. Nesse, A. Borisyuk, and P.C. Bressloff. Fluctuation-driven rhythmogenesis in an excitatory neuronal network with slow adaptation. *Journal of computational neuroscience*, 25(2):317–333, 2008.

[133] W. Nicola and S.A. Campbell. Non-smooth bifurcations of mean field systems of two-dimensional integrate and fire neurons. *SIAM J. Dynamical Systems (To appear 2015)*.

[134] W. Nicola and S.A. Campbell. Bifurcations of large networks of two-dimensional integrate and fire neurons. *Journal of computational neuroscience*, 35(1):87–108, 2013.

[135] W. Nicola and S.A. Campbell. Mean-field models for heterogeneous networks of two-dimensional integrate and fire neurons. *Frontiers in computational neuroscience*, 7, 2013.

[136] W. Nicola, C. Ly, and S.A. Campbell. One-dimensional population density approaches to recurrently coupled networks of neurons with noise. *SIAM J. Applied. Mathematics (To appear 2015)*.

[137] D.Q. Nykamp and D. Tranchina. A population density approach that facilitates large-scale modeling of neural networks: Analysis and an application to orientation tuning. *Journal of Computational Neuroscience*, 8:19–50, 2000.

[138] A. Omurtag, B.W. Knight, and L. Sirovich. On the simulation of large populations of neurons. *Journal of Computational Neuroscience*, 8:51–63, 2000.

[139] J. Orchard, H. Yang, and X. Ji. Does the entorhinal cortex use the fourier transform? *Frontiers in computational neuroscience*, 7, 2013.

[140] R. Oşan, R. Curtu, J. Rubin, and G.B. Ermentrout. Multiple-spike waves in a one-dimensional integrate-and-fire neural network. *Journal of mathematical biology*, 48(3):243–274, 2004.

[141] E. Ott and T.M. Antonsen. Long time evolution of phase oscillator systems. *Chaos: An interdisciplinary journal of nonlinear science*, 19(2):023117, 2009.

[142] C. Parisien, C.H. Anderson, and C. Eliasmith. Solving the problem of negative synaptic weights in cortical models. *Neural computation*, 20(6):1473–1494, 2008.

[143] M.M. Peixoto. *Dynamical systems*. Academic Press New York, 1973.

[144] W. Penfield and H. Jasper. Epilepsy and the functional anatomy of the human brain. 1954.

[145] L. Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media, 2013.

[146] D.J. Pinto, J.C. Brumberg, D.J. Simons, G.B. Ermentrout, and R. Traub. A quantitative population model of whisker barrels: re-examining the wilson-cowan equations. *Journal of computational neuroscience*, 3(3):247–264, 1996.

[147] A.A. Prinz, C.P. Billimoria, and E. Marder. Alternative to hand-tuning conductance-based models: construction and analysis of databases of model neurons. *Journal of Neurophysiology*, 90(6):3998–4015, 2003.

[148] D. Rasmussen and C. Eliasmith. A spiking neural model applied to the study of human performance and cognitive decline on raven's advanced progressive matrices. *Intelligence*, 42:53–82, 2014.

[149] D. Reid, H. Tawfik, A.J. Hussain, and H. Al-Askar. Forecasting weather signals using a polychronous spiking neural network. In *Intelligent Computing Theories and Methodologies*, pages 116–123. Springer, 2015.

[150] A Renyi and P Erdos. On random graphs. *Publicationes Mathematicae*, 6(290-297):5, 1959.

[151] Alfred Renyi. *Probability theory. 1970*. 1970.

[152] C. Rossant, D.F.M. Goodman, B. Fontaine, J. Platkiewicz, A.K. Magnusson, and R. Brette. Fitting neuron models to spike trains. *Frontiers in neuroscience*, 5:9, 2011.

[153] C. Rossant, D.F.M Goodman, J. Platkiewicz, and R Brette. Automatic fitting of spiking neuron models to electrophysiological recordings. *Frontiers in neuroinformatics*, 4, 2010.

[154] M. Rudolph and A. Destexhe. Analytical integrate-and-fire neuron models with conductance-based dynamics for event-driven simulation strategies. *Neural computation*, 18(9):2146–2210, 2006.

[155] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning representations by back-propagating errors. *Cognitive modeling*, 5:3, 1988.

[156] S. Sadeh, C. Clopath, and S. Rotter. Processing of feature selectivity in cortical networks with specific connectivity. *PloS one*, 10(6):e0127547, 2015.

[157] E. Salinas and L.F. Abbott. Vector reconstruction from firing rates. *Journal of Computational Neuroscience*.

[158] J.A. Sanders, F. Verhulst, and J.A. Murdock. *Averaging methods in nonlinear dynamical systems*, volume 2. Springer, 2007.

[159] E.S. Schaffer, S. Ostojic, and L.F. Abbott. A complex-valued firing-rate model that approximates the dynamics of spiking networks. 2013.

[160] A. Schüz and G. Palm. Density of neurons and synapses in the cerebral cortex of the mouse. *Journal of Comparative Neurology*, 286(4):442–455, 1989.

[161] L.F. Shampine and M.W. Reichelt. The matlab ode suite. *SIAM journal on scientific computing*, 18(1):1–22, 1997.

[162] M.J. Shelley and L. Tao. Efficient and accurate time-stepping schemes for integrate-and-fire neuronal networks. *Journal of Computational Neuroscience*, 11(2):111–119, 2001.

[163] O. Shriki, D. Hansel, and H. Sompolinsky. Rate models for conductance-based cortical neuronal networks. *Neural Computation*, 15:1809–1841, 2006.

[164] D.J.W. Simpson. *Bifurcations in piecewise-smooth continuous systems*. World Scientific, 2010.

[165] D.J.W. Simpson and J.D. Meiss. Andronov–hopf bifurcations in planar, piecewise-smooth, continuous flows. *Physics Letters A*, 371(3):213–220, 2007.

[166] D.J.W. Simpson and J.D. Meiss. Aspects of bifurcation theory for piecewise-smooth, continuous systems. *Physica D: Nonlinear Phenomena*, 241(22):1861–1868, 2012.

[167] R. Singh and C. Eliasmith. Higher-dimensional neurons explain the tuning and dynamics of working memory cells. *The journal of neuroscience*, 26(14):3667–3678, 2006.

[168] L. Sirovich, A. Omurtag, and B. W. Knight. Dynamics of neuronal populations: The equilibrium solution. *SIAM Journal on Applied Mathematics*, 60(6):pp. 2009–2028, 2000.

[169] L. Sirovich, A Omurtag, and K. Lubliner. Dynamics of neural populations: Stability and synchrony. *Network: Comput. Neural Syst*, 17:3–29, 2006.

[170] J.C. Skou. The influence of some cations on an adenosine triphosphatase from peripheral nerves. *Biochimica et biophysica acta*, 23:394–401, 1957.

[171] P. So, T.B. Luke, and E. Barreto. Networks of theta neurons with time-varying excitability: Macroscopic chaos, multistability, and final-state uncertainty. *Physica D: Nonlinear Phenomena*, 267:16–26, 2014.

[172] C.E. Stafstrom, P.C. Schwindt, and W.E. Crill. Repetitive firing in layer v neurons from cat neocortex in vitro. *Journal of Neurophysiology*, 52(2):264–277, 1984.

[173] S.H. Strogatz. From kuramoto to crawford: exploring the onset of synchronization in populations of coupled oscillators. *Physica D: Nonlinear Phenomena*, 143(1):1–20, 2000.

[174] S.H. Strogatz and R.E. Mirollo. Stability of incoherence in a population of coupled oscillators. *Journal of Statistical Physics*, 63:613–635, 1991.

[175] J. Tabak, W. Senn, M.J. O'Donovan, and J. Rinzel. Modeling of spontaneous activity in developing spinal cord using activity-dependent depression in an excitatory network. *The Journal of Neuroscience*, 20(8):3041–3056, 2000.

[176] A. Tonnelier, H. Belmabrouk, and D. Martinez. Event-driven simulations of nonlinear integrate-and-fire neurons. *Neural Computation*, 19(12):3226–3238, 2007.

[177] J. Touboul. Bifurcation analysis of a general class of nonlinear integrate-and-fire neurons. *SIAM Journal on Applied Mathematics*, 68(4):1045–1079, 2008.

[178] A. Treves. Mean-field analysis of neuronal spike dynamics. *Network: Computation in Neural Systems*, 4(3):259–284, 1993.

[179] B.P. Tripp and J. Orchard. Population coding in sparsely connected networks of noisy neurons. *Frontiers in computational neuroscience*, 6, 2012.

[180] B. Van der Pol. Lxxxviii. on "relaxation-oscillations". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):978–992, 1926.

[181] C. van Vreeswijk. Partial synchronization in populations of pulse-coupled oscillators. *Phys. Rev. E*, 54:5522–5537, Nov 1996.

[182] C van Vreeswijk. Partial synchronization in populations of pulse-coupled oscillators. *Physical Review E*, 54(5):5522, 1996.

[183] C. van Vreeswijk, L.F. Abbott, and G.B. Ermentrout. When inhibition not excitation synchronizes neural firing. *Journal of Computational Neuroscience*, 1:313–321, 1994.

[184] C. van Vreeswijk and D. Hansel. Patterns of synchrony in neural networks with spike adaptation. *Neural Computation*, 13(5):959–992, 2001.

[185] S. Visser and S.A. Van Gils. Lumping izhikevich neurons. *EPJ Nonlinear Biomedical Physics*, 2(1):1–17, 2014.

[186] B.B. Vladimirski, J. Tabak, M.J. O'Donovan, and J. Rinzel. Episodic activity in a heterogeneous excitatory network, from spiking neurons to mean field. *Journal of computational neuroscience*, 25(1):39–63, 2008.

[187] X. Wang and G. Buzsáki. Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *The journal of Neuroscience*, 16(20):6402–6413, 1996.

[188] Xie X. Werfel, J. and H.S. Seung. Learning curves for stochastic gradient descent in linear feedforward networks. *Neural Computation*.

[189] S. Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*, volume 2. Springer Science & Business Media, 2003.

[190] H.R. Wilson and J.D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical journal*, 12(1):1, 1972.

[191] E.A. Zilli and M.E. Hasselmo. Coupled noisy spiking neurons as velocity-controlled oscillators in a model of grid cell spatial firing. *The Journal of neuroscience*, 30(41):13850–13860, 2010.