UNIVERSITY OF CALIFORNIA SAN DIEGO

Linear-Threshold Network Dynamics: Properties and Applications to Dynamical Brain Behaviors

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy

in

Engineering Sciences (Mechanical Engineering)

by

Michael Patrick Durrell McCreesh

Committee in charge:

Professor Jorge Cortés, Chair
Professor Nikolay Atanasov
Professor Boris Kramer
Professor Miroslav Krstic
Professor Erfan Nozari

2024

The Dissertation of Michael Patrick Durrell McCreesh is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisor, Professor Jorge Cortés, for his support throughout the last five years. His interest in researching such a wide variety of topics allowed me to pursue topics of individual interest. Further, I am very thankful for the amount of time and help that he provided for every project, which both improved the quality of the work and demonstrated how to perform quality research.

I would also like to thank my committee members Professor Nikolay Atanasov, Professor Boris Kramer, Professor Miroslav Krstic, and Professor Erfan Nozari for their time and feedback on the work included in this dissertation.

I am incredibly grateful to Professor Bahman Gharesifard for being an incredible mentor throughout my time at Queen's University, both as an undergraduate and Masters student. His guidance and advice were invaluable to completing this degree.

The five years spent in San Diego while completing this degree have provided some of the best experiences of my life. For this, I would like to thank my friends, whether from triathlon, school, or meeting somewhere along the way, it would not have been as fun without you all.

To my family, thank you for your continued support, and giving me the confidence to move across the continent to pursue and complete this degree.

Finally, and most importantly, to Cleo. Without your love and belief in me, no matter the circumstances, I could not have done any of this. I look forward to all of our adventures to come.

Chapter 4, in full, is a reprint of the material [1] where it appears as "Sufficient Conditions for Oscillations in Competitive Linear-Threshold Brain Networks" by Michael McCreesh, Tommaso Menara and Jorge Cortés in IEEE Control Systems Letters. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in full, is a reprint of the material [2] where it appears as "Selective Inhibition and Recruitment in Linear-Threshold Thalamocortical Networks" by Michael McCreesh and Jorge Cortés in IEEE Transactions on Control of Network Systems. The dissertation author was the primary investigator and author of this paper.

Chapter 6, in full, is a reprint of material submitted for publication where it may appear as "Control of Linear-Threshold Brain Networks via Reservoir Computing" by Michael McCreesh and Jorge Cortés. The dissertation author was the primary investigator and author of this paper.

VITA

2017	Bachelor of Applied Science in Mathematics and Engineering (Computing and Communications) Queen's University
2019	Master of Applied Science in Mathematics and Engineering Queen's University
2024	Doctor of Philosophy in Engineering Sciences (Mechanical Engineering) University of California San Diego

PUBLICATIONS

1. **M. McCreesh**, and J. Cortés, "Control of Linear-Threshold Brain Networks via Reservoir Computing." Submitted.
2. **M. McCreesh**, E. Nozari and J. Cortés, "Linear-Threshold Rate Dynamics: Properties and Applications to Brain Phenomena," in preparation.
3. **M. McCreesh** and J. Cortés, "Selective Inhibition and Recruitment in Linear-Threshold Thalamocortical Networks", *IEEE Transactions on Control of Network Systems*, vol. 11, issue. 1, pp. 375-388, 2024.
4. **M. McCreesh**, T. Menara and J. Cortés, "Sufficient Conditions for Oscillations in Competitive Linear-Threshold Brain Networks", *IEEE Control Systems Letters*, vol. 7, pp. 2886-2891, 2023.

ABSTRACT OF THE DISSERTATION

Linear-Threshold Network Dynamics: Properties and Applications to Dynamical Brain Behaviors

by

Michael Patrick Durrell McCreesh

Doctor of Philosophy in Engineering Sciences (Mechanical Engineering)

University of California San Diego, 2024

Professor Jorge Cortés, Chair

The brain, composed of billions of interconnected neurons, forms a complex network that exhibits an incredibly wide array of behaviors. Treating this structure as a dynamical system provides a multitude of tools to use to model and understand the relationship between structure and function in the brain. As subnetworks exist at all levels in the brain, ranging from networks of individual neurons to networks of entire regions, a diverse set of models, each with different properties have been considered to study different dynamic brain behaviors. One such class of models are firing rate models, which monitor the average spike rate of populations of neurons. A particular firing rate model is the linear-threshold network model, which exhibits a wide range of

rich behaviors based on the underlying network structure and inputs. This ability to exhibit a variety of behaviors motivates the use of this model to study a variety of dynamical behaviors observed in the brain.

This dissertation considers three problems within the realm of modeling dynamical brain behaviors with the linear-threshold model. First, motivated by the appearance of oscillatory behavior when observing brain activity, we study the existence of oscillations in the linear-threshold dynamics. In order to provide sufficient conditions for oscillations in specific network topologies we also provide conditions for the stability of equilibrium points that maintain a specific support. Second, we discuss the dynamical brain behavior of selective inhibition and recruitment. We consider thalamocortical networks with both hierarchical and star-connected topologies, and focus on how the inclusion of the thalamus can improve the stabilizability properties of the linear-threshold dynamics relevant to the application. We finish by investigating the problem of reference tracking for the linear-threshold dynamics, which can be used to frame many brain behaviors. We approach this both analytically and with a data-driven approach to better match observations on how the brain processes information.

Chapter 1

Introduction

All cognitive and physical functions performed by the human body, both conscious and unconscious, involve communication from and decisions made by the brain. As such, understanding how the structure and activity in the brain result in function is a question of great relevance. This question defines the field of neuroscience, but such an understanding of how the brain works is relevant as well to engineers and scientists working to treat and cure brain disorders. Computational modeling has long been an approach to gain such an understanding.

The interconnected neurons that make up the brain interact in countless ways forming a highly complex dynamical system. The interactions in the system are dependent on both the neuronal dynamics and the changing topological structure and play out across varying spatial and temporal scales. Spatially, these range from the level of individual neurons, known as the microscale, to considering populations of neurons, known as the mesoscale, and looking at the interactions between large regions of the brain, known as the macroscale. Temporally, the dynamics of neurons vary based on their spatial location, with neurons at the periphery of the nervous system acting on a significantly faster timescale than those involved with higher-level processing. These dynamical intricacies make it a natural choice to apply system-theoretic techniques to analyze and understand computational models of the brain.

The system-theoretic approach to computational models of the brain ranges from analyzing networks at the scale of individual neurons to networks composed of entire brain regions.

With such diverse scales there is an equally diverse range of models that can be used to provide effective computational models, ranging from partial differential equations at the the neuronal level to linear models at the level of entire regions. Depending on the function that one is interested in, every model has its benefits and drawbacks. As such, the use of any one model is not sufficient to gain a full understanding of the brain.

At the individual level neurons communicate through spiking, where an electrical impulse is sent between neurons corresponding to a change in the membrane potential of a neuron. As such, many computational models of the brain are based on studying signals composed of a sequence of spikes. In this dissertation, instead of looking directly at the spike patterns of individual neurons, we consider the rate of spiking of populations of neurons exhibiting similar levels and types of activity. This is known as a firing rate model, and allows for a more streamlined mathematical description, but still maintains a tight connection with biological descriptions of neuron function. In particular we will consider the linear-threshold network model. The goal of this work is first to understand properties of the linear-threshold dynamics before using these properties to model neural behavior with a linear-threshold network.

1.1 Literature Review

Computational neuroscience, and in particular the modeling of neurons with dynamical systems traces back to the early 20th century with the integrate and fire model [3] in 1907. Many models followed, including the Hodgkin and Huxley model [4], the FitzHugh-Nagumo model [5, 6], and the Morris-Lecar model [7] at the scale of individual neurons. It later became established that some information processing in the brain is performed through activity at the population level rather than solely at the level of individual neurons [8, 9], leading to extensive work on modeling the behavior of populations of neurons [10, 11, 12, 13, 14, 15, 16, 17]. One of the most famous of these is the Wilson-Cowan model [18, 19]. Further models have been considered at the level of entire brain regions, with linear models proving to be as accurate as

many nonlinears. A comparison of such models can be found in [20].

Concurrently the study of control theory was vastly expanding its reach to many applications, with automobiles [21, 22], aircraft [23, 24], and robotics [25, 26, 27] being a small set of examples. Advances were made in many areas, such as linear and nonlinear systems, optimal and geometric control, controllability and observability, and many more. We direct the reader to [28] for a collection of papers outlining some of the most important advances in the 20th century.

However, it was not until the latter part of the 20th century that control theory techniques were applied to the variety of nonlinear dynamical models proposed for neuronal behavior [29, 30]. Over the last 30 years the number of works studying computational neuroscience from a controls perspective has grown massively, see e.g. [31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48]. Neural processes and behavior in the brain that have been studied using control theory include visual processing [49, 50], voluntary movement [51, 52], and pathological behavior due to disorders such as Parkinson's or epilepsy [53, 54, 55, 56]

The combination of using control theory techniques to study neuronal behavior has provided a variety of benefits to the field of study. By themselves, the neuronal dynamics provide descriptive models of the activity in the brain, helping us understand the past behavior. However, to enhance our understanding of the brain, we are interested in prescriptive models. The application of control theory techniques provide these models, where we can predict and explain future behavior based on expected or desired stimuli, provided as inputs or perturbations in the neural dynamics [57, 34].

In addition, a control-theoretic analysis of brain networks provides an avenue to gain a better understanding of the structure of the brain from a dynamical systems perspective. By investigating properties of the network structure such as robustness and efficiency [58], we can aim to uncover the benefits of the network architectures found in the brain and try to understand why the brain has developed in the way it has. In addition to understanding the properties of existing brain networks we can also apply the tools from control theory to learn how the networks can be modified to fix damage or improve performance [48].

One model that has been used when applying control-theoretic techniques to neural models in the linear-threshold network model. The linear-threshold network dynamics has been used to study a variety of brain processes and functions, ranging from the retinal behavior of a crab [59], to memory [60, 61], seizures [62, 63], and goal-driven selective attention [64, 65, 2]. Properties of controlling networks governed by the linear-threshold dynamics have been studied using both model-based [65, 66, 1] and data-based [67] control.

The dynamics has piecewise-affine nonlinearities, which generalize sigmoidal nonlinearities [68], and in addition to being consistent with empirical descriptions of neural physiology, include the celebrated Wilson-Cowan model [18, 69] as a particular case. Further, these dynamics are able to generate a rich suite of behaviors that allow for the modeling a wide variety of the dynamical activity seen in the brain. These include unique and multiple equilibria [70, 61], mono- and multi-stability [71, 64, 72, 73], and oscillatory and chaotic behavior [1, 66, 63, 74].

In this dissertation we further extend on the application of control-theoretic techniques to the linear-threshold dynamics as a model for neural behaviors.

1.2 Organization and Contributions

The overall contribution of this dissertation is twofold. First we aim to understand the properties of the linear-threshold network (LTN) dynamics from a system-theoretic perspective. Second, we explore how the dynamics and its properties can be used as a firing rate model for the brain, with the goal of providing a model of different neural behaviors.

The dissertation is structured as follows. Chapters 2 and 3 provide general background information relevant to the results provided in the following three chapters. Each of Chapters 4, 5, and 6 provide the technical results of this dissertation. These chapters each include a literature review specific to their topics and a further description of their contributions. An overview of the contributions of each chapter is as follows.

Chapter 2: In this chapter we provide the neuroscientific background for the develop-

ment of system-theoretic network models of the brain. In particular we discuss the construction of the firing rate models discussed in this dissertation. We include a discussion of different activation functions used in firing rate models, including the linear-threshold function used throughout this work.

Chapter 3: We introduce the notations used throughout this dissertation and a review of relevant mathematical concepts, beginning with an overview of control theory and tools from singular perturbation theory. In addition we include preliminary properties and results on the LTN dynamics considered throughout the dissertation.

Chapter 4: We consider system-theoretic properties of a network governed by the LTN dynamics. By considering the possible behaviors of specific sets of nodes, where the remaining components of the network have zero activity we first provide necessary and sufficient conditions for the existence of stable equilibria with a specific support on a general network topology. We then provide sufficient conditions for the appearance of oscillatory behavior, either periodic or chaotic, on a subset of nodes in *competitive* network topologies.

Chapter 5: We consider the use of LTN dynamics to study *goal-driven selective attention* (GDSA) in thalamocortical brain networks. Our first contribution is an analysis of the mechanisms involved in selective inhibition and recruitment in hierarchical thalamocortical networks from a controls perspective. Using singular perturbation theory for non-smooth differential equations, and leveraging the piecewise-affine nature of the LTN dynamics, we provide feedback-feedforward control laws that provide selective inhibition and recruitment of the network to an equilibrium trajectory.

For our second contribution we consider selective inhibition and recruitment in star-connected thalamocortical networks. For star-connected networks with a temporal hierarchy we again leverage singular perturbation theory, while for networks without such a hierarchy we generalize results on the stability of slowly-varying systems. For both networks we illustrate the benefits of the thalamocortical interconnections in relation to metrics such as failsafe mechanisms, control magnitude and network performance.

Chapter 6: In this chapter we consider the problem of reference tracking for the LTN dynamics. Our first contribution is providing a sufficient condition such that the dynamics can converge to a reference trajectory dependent on the properties of the synaptic weight matrix and reference signal.

Our second contribution involves the use of a data-driven control method to achieve reference tracking in the LTN dynamics which provides two benefits. First, it bypasses difficulties in determining if a network satisfies the previously supplied sufficient condition. Second, as our dynamics represent a brain model, using a data-driven control that is determined in an online fashion is more realistic. In particular we modify the machine learning frameworks of Reservoir Computing and Next-Generation Reservoir Computing to construct controllers to be applied to networks governed by the LTN dynamics. Simulations are provided illustrating the success of these controllers in the context of GDSA and seizure rejection through neuromodulation.

Chapter 7: We provide a summary of the contributions of this dissertation and discuss directions for future work.

Chapter 2

Computational Brain Models

Computational modeling of the brain is particularly challenging due, in part, to the different scales of information in which one can phrase and approach the problem. At the smallest level, or the "microscale", the brain is composed of billions of neurons whose dynamics can be measured at the individual level through their voltages. At the opposite end of the spectrum, the "macroscale", the brain can be divided into large regions each composed of tens of millions of neurons with "activity" patterns recorded using imaging modalities such as functional magnetic resonance imaging (fMRI) or electroencephalography (EEG). Various levels also fall in-between the two extremes, often referred to as the "mesoscale". Each scale is the host to different network structures, elemental components, and connectivity patterns. This heterogeneity of spatial scales alone makes it infeasible to study the brain and all its functions and emerging phenomena using a single computational model.

Both microscale models, illustrating the voltage dynamics of individual neurons, and macroscale models, showing functional connectivity between regions, have been the subject of much research using computational models. Details on microscale models can be found in [4, 75, 76, 68] and references therein, and for macroscale models, we direct the reader to [20, 77] and references within. Interestingly, while the microscale dynamics at the neuronal level must be nonlinear, at the macroscale this is not necessarily the case. Despite the frequent assumption that accurate brain models must be nonlinear [78, 79], the recent comparison [20]

of a large variety of linear and nonlinear macroscopic models did not find any advantage in the latter.

In this work we discuss computational brain models at the mesoscale, describing the interaction between populations of neurons each having similar function and statistical properties. Two main types of mesoscale models are local field potential models (LFPs) and firing rate models. LFP models are based upon measuring the electric potential in the extracellular space around neuron populations, while firing rate models measure the average firing rate of all the neurons within a population [68]. Both LFPs and firing rate models have been used extensively for studying brain function [80, 81, 82, 83] and, in their simplest forms, can be transformed into each other through an affine transformation [71]. In this work, we use firing rate models for our study of selective inhibition and recruitment in thalamocortical networks.

In addition to these spatial scales of information in the brain, there exist vast temporal differences between brain regions. Each region in the brain operates on its own different timescale, which potentially vary greatly, even to the point of the activity in one area appearing constant relative to that in another. As such, when considering a model at any scale of spatial information it is necessary to also account for the differing timescales between regions, by encoding a timescale into the model dynamics of each region.

Regardless of the scale of the brain network model, they all have basic graph-theoretic elements in common. A brain network is modeled as a collection of nodes, with nodes representing populations of neurons and edges representing the interconnections between these populations. These populations of neurons could be as small as individual neurons or as large as entire brain regions, depending on the scale of the model. Each of these nodes has its own defining properties (e.g., consisting of excitatory or inhibitory neurons) and they are connected to form a network structure as shown in Figure 2.1. Specific model aspects (functional forms, parametric constraints, etc.) are then defined in accordance with the information scale of the model. Our next section describes the particular assumptions made to describe firing rate models. The reader familiar with these models can safely skip this discussion.

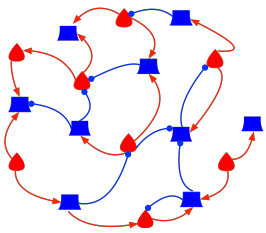


Figure 2.1. Graph-theoretic model of a brain network. Excitatory neurons and connections are shown in red, while inhibitory neurons and populations are shown in blue.

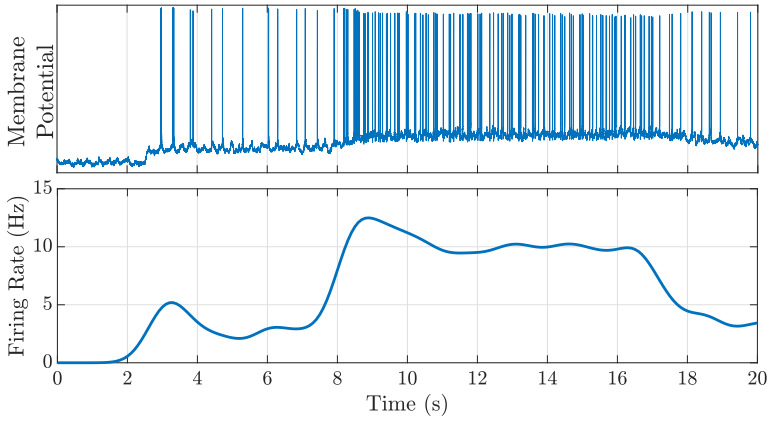


Figure 2.2. An intracellular recording showing a spike train as is used for communication between neurons (top) and the corresponding firing rate (bottom). This is estimated by binning spikes in 100^{ms} bins and smoothing using a Gaussian window with 500^{ms} standard deviation [64, 84, 85].

2.1 Firing Rate Models

In this section we outline the construction of firing rate models in the brain, as per [68, § 7]. At the level of neurons, brain dynamics consist of a series of spikes, corresponding to action potentials, being transmitted between neurons, see Figure 2.2. The spike train is transmitted from one neuron to another at a *synapse*, and as such the two neurons are referred to as the *pre-synaptic* and *post-synaptic* neurons. The sequence of spikes (both input and output signals) transmitted between neurons is defined by a neural response function, $\rho(t)$, modeled as an impulse train of the form $\rho(t) = \sum_k \delta(t - t_k)$, with δ denoting the Dirac delta function.

In many areas of the brain, the spike trains defined by the neural response function appear to be highly random, and observations have little trial-to-trial reproducibility, which makes

accurate spike train models difficult to construct. Replacing the neural response function with the average firing rate provides more trial-to-trial reproducibility (see Figure 2.2), along with providing some other benefits. First, spiking models can only accurately predict spike trains sequences if all inputs into a neuron are known. Given the complexity of the brain, knowing this is highly unlikely. Second, the probability of any two randomly selected neurons being connected is low. Hence, the construction of a network model that has a high degree of connectivity while maintaining this property requires using a large number of nodes. Therefore it is standard practice to instead model a single node in a network as the average response of a population of neurons. This allows for a less sparse network model. In this case, it is difficult to describe what the average response of the population would be when using spike trains. The use of firing rates instead allows us to specify the average response simply as the average firing rate of the neurons within the population.

We next explain how the firing rate model is constructed. First, we determine how the total synaptic input of a neuron is dependent on the firing rates of its pre-synaptic afferents. Consider a pair of pre- and post-synaptic neurons, with firing rates given by $x_{\text{pre}}(t)$ and $x_{\text{post}}(t)$. Then, the firing rate of the pre-synaptic neuron generates the synaptic input into the post-synaptic neuron in the form of an electrical current, denoted $I_{\text{post}}(t)$. Assuming the synapse has fast dynamics, $I_{\text{post}}(t)$ is approximately proportional to $x_{\text{pre}}(t)$ with proportionality constant $w_{\text{post,pre}}$, where $w_{\text{post,pre}}$ is known as the synaptic weight. The pre-synaptic neuron is *excitatory* if $w_{\text{post,pre}} > 0$ and is *inhibitory* if $w_{\text{post,pre}} < 0$. As such, an excitatory neuron increases the activity of its out-neighbors while an inhibitory neuron decreases it. We note that excitation and inhibition is a property of neurons, rather than synapses, so a neuron either excites or inhibits all of its out-neighbors, but not a combination (this is known as *Dale's law*). The synaptic current of a neuron that receives multiple synaptic inputs follows a superposition law, with

$$I_{\text{post}}(t) = \sum_j w_{\text{post},j} x_j(t), \tag{2.1}$$

where the sum is taken over the neurons providing inputs. For simplicity we assume that $I_{\text{post}}(t)$ is measured in Hertz in order to match the units of firing rates. This is accomplished through an implicit multiplicative constant converting it from the traditional current unit of amperes, which also renders the synaptic weights $w_{\text{post},j}$ dimensionless constants.

Next, we model how the firing rate of the post-synaptic neuron depends on the synaptic input as $x_{\text{post}}(t) = F(I_{\text{post}}(t))$, where $F(\cdot)$ is a typically nonlinear "activation" function. For an individual neuron the firing rate dynamics is then described by as

$$\tau \dot{x}_{\text{post}}(t) = -\mathbf{x}_{\text{post}}(t) + F(I_{\text{post}}(t)), \tag{2.2}$$

where τ is a timescale constant indicating a "lag" between the change in the synaptic input and the change in the firing rate.

While this derivation of a firing rate model as described above uses individual neurons, it is common to replace the individual neurons with populations of neurons with similar activation patterns, resulting in a firing rate model at the mesoscale. In this case, the firing rates x_{pre} and x_{post} represent the average firing rate of the population of neurons. Finally, to move from the dynamics of a single pair of neurons (or populations of neurons) to the dynamics of a brain network, we take the following steps. We consider a network with n nodes and let $\mathbf{x} \in \mathbb{R}^n$ represent the firing rates of the nodes. Then, combining the synaptic weights $w_{i,j}$ into a matrix $\mathbf{W}$ and using (2.1) and (2.2), we obtain the network firing rate dynamics

$$\tau \dot{\mathbf{x}} = -\mathbf{x} + F(\mathbf{W}\mathbf{x} + \mathbf{d}(t)), \tag{2.3}$$

where the activation function is applied component-wise. The term $\mathbf{d}(t)$ is added to the synaptic input to model external inputs to the network, such as un-modeled background activity, external inputs, or non-zero thresholds. What completes the firing rate model, and distinguishes between them, is the choice of activation function. A variety of different functions have been used in the

literature for studying firing rate models, with different benefits and drawbacks to each.

When considering activation functions for the firing rate model they must all satisfy a set of base criteria in order to be a realistic representation of neuron firing rates. First, since the unit of measurement in the model is frequency (Hz) it is necessary that the domain of the dynamics is nonnegative, which is guaranteed by requiring that the activation function maps from $\mathbb{R}$ to $\mathbb{R}_{\geq 0}$ for an individual neuron (population of neurons). Second, due to the refractory period after an action potential, neurons have a maximum firing rate that can differ between types of neurons [86]. As such the firing rate model needs to exhibit bounded activity, either through having an upper bound on the activation function or other properties that bound the activty of the dynamics. Beyond these base requirements other properties of different activation functions can be beneficial, such as differentiability and computation complexity, and there is wide flexibility in choice, ranging from highly nonlinear to linear [20]. For firing rate models in particular, three common activation functions are the sigmoid [17], the threshold-linear function (also known as rectified linear unit or ReLU) [87, 73] and the linear-threshold model [17, 64]. A comparison of the properties of models is provided in Table 2.1 with graphs of the functions in Figure 2.3.

Table 2.1. A comparison of three commonly used activation functions: sigmoidal, threshold linear and linear threshold. Images of the functions are shown in Figure 2.3.

Property	**Sigmoidal**	**Threshold-Linear**	**Linear-Threshold**
Lower Bounded	Yes	Yes	Yes
Upper Bounded	Yes	No	Yes
Differentiable	Yes	Piecewise	Piecewise
Nonlinearity Type	Exponential	Piecewise-affine	Piecewise-affine
Computational Complexity	High	Low	Low

In this work we consider the linear-threshold activation function, which gives the linear-threshold network (LTN) firing rate model, governed by the following dynamics

$$\tau \dot{\mathbf{x}} = -\mathbf{x} + [\mathbf{W}\mathbf{x} + \mathbf{d}(t)]_{\mathbf{0}}^{\mathbf{m}}. \tag{2.4}$$

We note that the vector of firing rate upper bounds $\mathbf{m}$ can be modeled as either finite

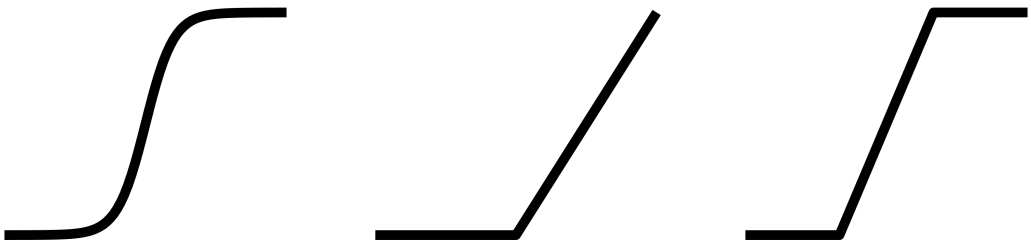


Figure 2.3. The sigmoidal (left), threshold-linear (middle), and linear-threshold (right) activation functions. All have been used extensively in brain modeling and their necessary properties are compared in Table 2.1.

or infinite. In the case of an infinite upper bound, the linear-threshold model reduces to the threshold-linear model. While the linear-threshold activation function is more biologically accurate, if it is assumed that the activity of the network is not approaching the firing rate threshold, then using a threshold-linear network (TLN) model instead can be convenient for network analysis.

2.2 Modeling Different Brain Regions

Since distinct functions utilize different areas in the brain, it is important for models to accommodate different structures to describe multiple phenomena. A majority of the literature on brain networks studies the cortex and cortical networks [80, 88, 89, 90] due to its role in higher-level processes in the brain, including memory and attention [91]. However, most cortical regions have inputs from subcortical areas, such as the thalamus, that play critical roles in the many functions undertaken by the cortex [92]. With the distinct regions having such a variety of functions, modeling them identically can be an overly simplistic assumption. In order to account for the differences in properties between them, we can provide restrictions at various levels (functional forms, hyper-parameters, or parameters) on the models used for different regions. Assuming a homogeneous use of the firing rate model with a linear-threshold functional form as derived in Section 2.1, the heterogeneity in properties of different regions can be encoded into the structure of the synaptic weight matrices making up the model. Here we will discuss

constraints on cortical and thalamic regions in the brain.

The cortex is composed of a mix of excitatory and inhibitory neuron populations and, while excitatory neurons significantly outnumber inhibitory neurons, both play important roles in the transmission and processing of information [89]. As such, we allow our firing rate model for cortical regions to be composed of populations of excitatory and inhibitory neurons with arbitrary numbers and connectivity patterns. The only restriction on the synaptic weight matrices of cortical regions is based on the fact that populations of neurons care either excitatory or inhibitory for all their outgoing connections. This is reflected in the matrices such that each column has either nonnegative or nonpositive values [68].

The *thalamus* connects with cortical regions through a series of parallel pathways, with most thalamic nuclei projecting to a unique cortical population [93]. However, lateral connections within the thalamus (including both excitatory and inhibitory populations, the latter of which lying primarily in the thalamic reticular nucleus) construct the transthalamic pathways between cortical regions that can lie in different places within a hierarchical structure in the cortex [92, 94]. Experimental observations indicate that along these pathways one of the mechanisms through which the thalamus and cortex interact is feedforward inhibition mediated by local interneurons [95, 96, 97]. In particular, these observations show that the cortex receives excitatory thalamic input but is inhibited due to connections between the thalamic input and inhibitory interneurons both for first-order sensory thalamic nuclei [98, 99, 100] and higher-order thalamic nuclei [97].

Given the complexity of thalamic structure, we make the following simplifying assumptions towards its computational modeling. First, we allow for each thalamic nucleus to project to any cortical population, potentially mediated in reality through the lateral connections within the thalamus and its internal dynamics. Second, we model the projections (outgoing connections) of the thalamus onto the cortical regions as being strictly inhibitory, mimicking the above-cited experimental observations of feedforward inhibition of the cortex by the thalamus while also simplifying the model. We note that the connections back from the cortical regions to the

thalamus are allowed to be both excitatory and inhibitory. Finally, the internal dynamics of the thalamus are restricted only such that each column has a nonpositive or nonnegative sign, similarly to the cortical regions[1].

2.3 Control Mechanisms in Brain Models

Brain models in general, and the simplified and tractable form of firing rate models in particular, are natural pathways to the study of control mechanisms of and for the brain. Akin to engineered systems, the types of control utilized in the brain can be (roughly) separated into feedback and feedforward. Feedback control operates off of circuits where the populations providing the control input are directly stimulated by the populations within the network. In these circuits, the magnitude of the control input is directly dependent on the activity level in the network. On the other hand, feedforward control is based upon input received from populations of neurons that are further from the network and is not dependent upon current activity levels within the network.

Feedback control is a mechanism based upon the interaction of two neuronal populations that form a closed loop. While feedback exists across the brain, a large component of feedback control occurs in local feedback loops. In the feedback circuit, the first population stimulates a second "control" population, which in return stimulates the first population in order to control its dynamics. As the "control" neuron population can be either excitatory or inhibitory, both excitatory or inhibitory feedback control exists within the brain. However, despite the existence of more excitatory than inhibitory neurons, the inhibitory neurons frequently exhibit higher firing rates and are able to influence the firing rates of other neuronal populations more than excitatory populations can, cf. [89]. As such, inhibitory feedback control is more common than excitatory feedback [101]. Figure 2.4(left) illustrates a standard inhibitory feedback loop.

Feedforward control is more often studied in the context of (potentially unidirectional)

[1]We note that while in our models we restrict the sign pattern of the synaptic weight matrices for biological reasons, the mathematical results provided are valid for matrices with arbitrary sign patterns.

Figure 2.4. Feedback and feedforward mechanisms of control within brain networks. The left panel shows an inhibitory feedback loop, where an excitatory signal (red) from the main neuronal population (grey) stimulates an inhibitory interneuron (blue), which in turn inhibits the main population. The right panel shows the feedforward inhibition mechanism, where an excitatory signal from a separate neuronal population stimulates an inhibitory interneuron as well as the main population. The interneuron then inhibits the main population, typically resulting in its net inhibition.

non-local connections between neuronal populations. In the case of cortical populations, for instance, they receive afferents from subcortical nuclei (i.e., the thalamus) as well as cortical populations in distant regions. These long-distance connections may not form clear feedback loops, but instead provide a feedforward control input that can modify the dynamics of the receiving neuronal population. While long-range connections in the brain are almost universally excitatory, they can indeed induce feedforward inhibition by exciting inhibitory "interneurons" (neurons with only local output connections), which in turn inhibit their downstream neuronal population. If this two-hop inhibition is stronger than the direct excitatory afferent received by the downstream population, as is commonly the case, a net feedforward inhibition would occur [101]. Figure 2.4(right) illustrates the feedforward inhibition mechanism.

Chapter 3

Preliminaries

In this chapter we provide the mathematical preliminaries for the results in this dissertation. Section 3.1 provides the notation for this work and Section 3.2 provides relevant matrix analysis results. In Section 3.3 a brief introduction to control theory is provided and Section 3.4 discusses singular perturbation theory. Sections 3.5 provides a discussion of basic properties of the LTN dynamics. The reader familiar with these topics can safely skip any or all of these sections.

3.1 Notation

We introduce here the notations that will be used in the remainder of this dissertation. We let $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{n\times m}$, denote reals, real-valued vectors and real-valued matrices, respectively. Vectors and matrices are identified by bold-faced letters. For vectors (matrices) $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ (resp. $\mathbb{R}^{n\times m}$), $\mathbf{x} \leq \mathbf{y}$ is the component-wise comparison (analogously with $<, >, \geq$). For a vector $\mathbf{x}$ and set of indices $\alpha \subseteq \{1, \ldots, N\}$ we denote by $\mathbf{x}_\alpha$ the vector composed of the elements of $\mathbf{x}$ by the indices in α. For a set of indices $\alpha \subseteq \{1, \ldots, N\}$, we denote by $\bar{\alpha}$ the complement of α, that is $\{1, \ldots, N\} \backslash \alpha$. For a vector $\mathbf{x} \in \mathbb{R}^n$ we refer to the set of non-zero components as the support of $\mathbf{x}$, and denote it by $\text{supp}(\mathbf{x})$.

For a matrix $\mathbf{W} \in \mathbb{R}^{n\times n}$ we denote its eigenvalues by $\lambda_1(\mathbf{W}) \geq \lambda_2(\mathbf{W}) \geq \cdots \geq \lambda_n(\mathbf{W})$. When relevant we will denote $\lambda_1(\mathbf{W})$ by $\lambda_{\max}(\mathbf{W})$ and $\lambda_n(\mathbf{W})$ by $\lambda_{\min}(\mathbf{W})$. For a matrix $\mathbf{W}$ and

two sets of indices, α_1, α_2 we let $\mathbf{W}_{\alpha_1\alpha_2}$ the submatrix defined by the rows indexed α_1 and columns indexed by α_2. If $\alpha_1 = \alpha_2 = \alpha$ we will denote this principal submatrix by $\mathbf{W}_\alpha$. For a 2×2-partitioned block matrix

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}^{00} & \mathbf{W}^{01} \\ \mathbf{W}^{10} & \mathbf{W}^{11} \end{bmatrix},$$

we use the notation $\mathbf{W}^{\ell,\text{all}} = \begin{bmatrix} \mathbf{W}^{\ell 0} & \mathbf{W}^{\ell 1} \end{bmatrix}$ and $\mathbf{W}^{\text{all},\ell} = \begin{bmatrix} (\mathbf{W}^{0\ell})^\top & (\mathbf{W}^{1\ell})^\top \end{bmatrix}^\top$ for $\ell \in \{0,1\}$. The identity matrix of dimension n is $\mathbf{I}_n$. $\mathbf{0}_n$ and $\mathbf{1}_n$ denote the n-vector of zeros and the n-vector of ones, respectively. When clear from the context, we omit the dimensional subindex for the identity or zero matrices. For $\mathbf{W} \in \mathbb{R}^{n\times n}$, we denote its element-wise absolute value, spectral radius, and induced 2-norm by $|\mathbf{W}|$, $\rho(\mathbf{W})$, and $\|\mathbf{W}\|$, resp. Similarly, we let $\|\mathbf{x}\|$ denote the 2-norm of a vector $\mathbf{x} \in \mathbb{R}^n$.

For $x \in \mathbb{R}$ and $m \in \mathbb{R}_{>0}$, $[x]_0^m$ denotes $\min\{\max\{x, 0\}, m\}$. For $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{m} \in \mathbb{R}^n_{>0}$, this operation is done component-wise as $[\mathbf{x}]_{\mathbf{0}}^{\mathbf{m}} = [[x_1]_0^{m_1}, \ldots, [x_n]_0^{m_n}]$. For two vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$, $\mathbf{x} \oplus \mathbf{y} \in \mathbb{R}^{n+m}$ is the concatenation of vectors $\mathbf{x}$ and $\mathbf{y}$. For scalar $r > 0$ and vector $\mathbf{x} \in \mathbb{R}^n$ let $B(r, \mathbf{x})$ denote the ball of radius r centered at $\mathbf{x}$. If $\mathbf{x} = \mathbf{0}$ we will denote the ball centered at the origin by B_r.

3.2 Matrix Analysis

Due to the interplay between dynamics and topological structure when studying network systems such as brain models, a variety of matrix classes are useful for characterizing properties and behavior. When studying linear-threshold brain models the following matrix classes are of interest.

Definition 3.2.1. (Matrix Classes): *A matrix* $\mathbf{W} \in \mathbb{R}^{n\times n}$ *is*

- *a* $\mathcal{P}$-matrix *if all principal minors of* $\mathbf{W}$ *are positive (denoted* $\mathbf{W} \in \mathcal{P}$*);*

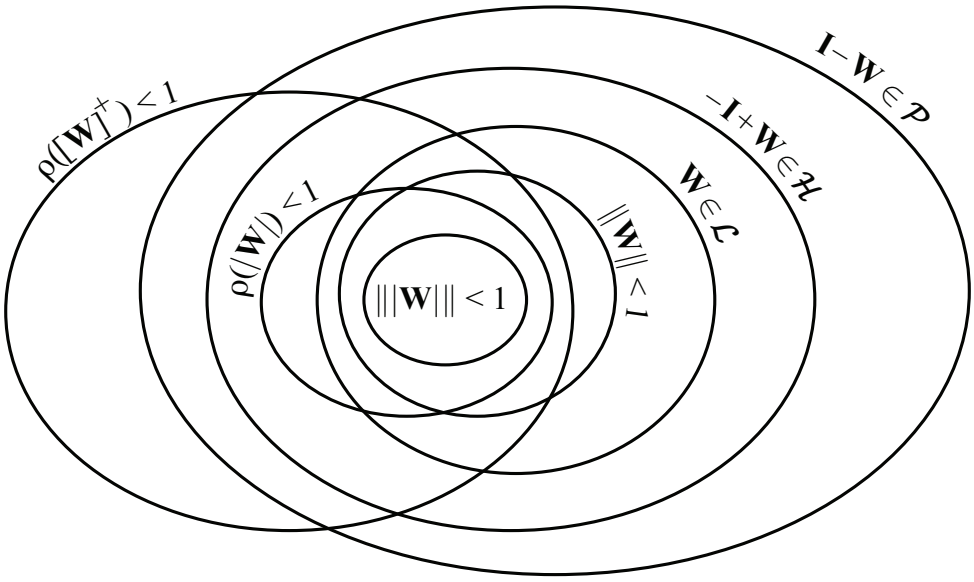


Figure 3.1. Inclusions between the matrix classes from Lemma 3.2.2.

- totally Hurwitz *if all principal submatrices of* $\mathbf{W}$ *are Hurwitz (denoted* $\mathbf{W}\in\mathcal{H}$*);*
- totally $\mathcal{L}$-stable *if there exists* $P>0$ *such that* $(-\mathbf{I}+\mathbf{W}^\top\Sigma)+(-\mathbf{I}+\Sigma\mathbf{W})<0$ *for* $\Sigma=\text{diag}(\sigma)$ *and all* $\sigma=\{0,1\}^n$ *(denoted* $\mathbf{W}\in\mathcal{L}$*);*
- absolutely Schur stable *if* $\rho(|\mathbf{W}|)<1$.

These matrix classes have a variety of relationships, which are given next.

Lemma 3.2.2. (Inclusions Between Matrix Classes [64, Lemma II.3]): *For matrix* $\mathbf{W}\in\mathbb{R}^{n\times n}$ *the following hold:*

- *if* $\mathbf{W}$ *is absolutely Schur stable then* $-\mathbf{I}+\mathbf{W}\in\mathcal{H}$*;*
- *it* $\|\mathbf{W}\|<1$ *then* $\mathbf{W}\in\mathcal{L}$*;*
- *if* $\mathbf{W}\in\mathcal{L}$ *then* $-\mathbf{I}+\mathbf{W}\in\mathcal{H}$*;*
- *if* $-\mathbf{I}+\mathbf{W}\in\mathcal{H}$ *then* $\mathbf{I}-\mathbf{W}\in\mathcal{P}$.

In Chapter 4 we study competitive networks which are defined by Z-matrices, defined as follows.

Definition 3.2.3. (Z- and M-Matrices): *A matrix* $\mathbf{W} \in \mathbb{R}^{n\times n}$ *is called a Z-matrix if for all* $i, j \in \{1,\dots,n\}$ *with* $i \neq j$, $\mathbf{W}_{ij} \leq 0$. *If additionally* $\mathrm{Re}(\lambda_i) \geq 0$ *for all* $i \in \{1,\dots,n\}$, $\mathbf{W}$ *is called a M-matrix.*

In studying oscillations we are interested in checking the instability of all 2×2 principal submatrices of a given matrix. Therefore, it is desirable to have a simple condition to check for instability of a 2×2 matrix. The *determinant condition for instability* reads as follows: "given $\mathbf{W} \in \mathbb{R}^{2\times 2}$, if $\mathrm{Tr}(\mathbf{W}) \leq 0$, $\mathbf{W}$ is unstable if and only if $\det(\mathbf{W}) < 0$; if $\mathrm{Tr}(\mathbf{W}) > 0$, then $\mathbf{W}$ is unstable". The next result details when the determinant condition for instability holds for Z-matrices.

Lemma 3.2.4. (Requirements for Determinant Condition for Instability for Z-Matrices): *Let* $\mathbf{W} \in \mathbb{R}^{2\times 2}$ *be a Z-matrix. The determinant condition for instability holds for* $-\mathbf{I}+\mathbf{W}$ *in the following cases:*

1. *If* $w_{11} = w_{22} = 0$ *and* $\frac{1}{|w_{12}|} < |w_{21}|$*;*
2. *If one or both of* $w_{12}, w_{21} = 0$*, then either* w_{11} *or* $w_{22} > 1$*;*
3. *If neither of the preceding cases hold and* $\mathrm{Tr}(-\mathbf{I}+\mathbf{W}) \leq 0$*, then* $(-1+w_{11})(-1+w_{22}) - w_{12}w_{21} \leq 0$.

The proof follows directly from the equation for the determinant of a 2×2-matrix.

3.3 Introduction to Control Theory

To keep this dissertation as a self-contained exposition, in the section we give an overview of the basics of dynamical systems and control theory tools that will be used going forward. A reader familiar with control-theoretic tools will have no issues with moving past this section. For the reader without a background in control further information, on nonlinear systems in particular, can be found in [102], among other textbooks.

The overarching goal of control theory (in continuous-time) is to study the properties of the dynamical system

$$\dot{\mathbf{x}} = F(\mathbf{x}, \mathbf{u}), \tag{3.1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the system state and $\mathbf{u} \in \mathbb{R}^m$ is the control, and design the control signal such that the system exhibits a desired behavior. Typically we study the autonomous system $\dot{\mathbf{x}} = F(\mathbf{x}, \mathbf{0})$ before investigating how the behavior can be modified through the choice of a control signal $\mathbf{u}$. When analyzing dynamical systems, one of the most studied characteristics are the equilibria of the system.

Definition 3.3.1. (Equilibrium Point [102]): *A point* $\mathbf{x}^*$ *is an equilibrium point of the autonomous dynamical system* $\dot{\mathbf{x}} = F(\mathbf{x})$ *if* $F(\mathbf{x}^*) = \mathbf{0}$.

When studying the equilibrium points of the system we are interested in how the system behaves in an area near the point. If the system stays near the equilibrium point it is referred to as *stable*, otherwise it is *unstable*. A variety of stability concepts are defined as follows.

Definition 3.3.2. (Stability of a Dynamical System [102]): *Consider the autonomous dynamical system* $\dot{\mathbf{x}} = F(\mathbf{x})$. *An equilibrium point* $\mathbf{x}^*$ *is*

- locally stable *if, for each* $\epsilon > 0$*, there exists* $\delta = \delta(\epsilon) > 0$ *such that*

$$\|\mathbf{x}(0) - \mathbf{x}^*\| < \delta \Rightarrow \|\mathbf{x}(t) - \mathbf{x}^*\| < \epsilon,$$

  for all $t \geq 0$.

- asymptotically stable *if it is stable and* δ *can be chosen such that*

$$\|\mathbf{x}(0) - \mathbf{x}^*\| < \delta \Rightarrow \lim_{t \to \infty} \mathbf{x}(t) = \mathbf{x}^*.$$

Without loss of generality we can assume that $\mathbf{0}$ *is an equilibrium point of the system.*

- exponentially stable *if it is asymptotically stable and there exist* $\alpha,\beta,\delta > 0$ *such that*

$$||\mathbf{x}(0) - \mathbf{x}^*|| < \delta \Rightarrow ||\mathbf{x}(t) - \mathbf{x}^*|| \leq \alpha||\mathbf{x}(0) - \mathbf{x}^*||e^{-\beta t},$$

  for all $t \geq 0$.

- globally asymptotically (exponentially) stable *if it asymptotically (exponentially) stable for all* $\delta > 0$.

A variety of tests and conditions exist to determine if the system (3.1) satisfies any of the stability concepts in Definition 3.3.2. The applicability of these tests are dependent on the type of system being considered. One such test is Lyapunov's method, which is utilized in some proofs in this work. We outline this method as follows.

Definition 3.3.3. (Lyapunov Function [102]): *Consider the autonomous system* $\dot{\mathbf{x}} = F(\mathbf{x})$ *and without loss of generality assume that* $\mathbf{x} = \mathbf{0}$ *is an equilibrium point. A continuously differentiable function* $V : \mathbb{R}^n \to \mathbb{R}$ *is a Lyapunov function for the system if:*

i) $V(\mathbf{0}) = 0$*;*

ii) $V(\mathbf{x}) > 0$ *for all* $\mathbf{x} \in \mathbf{U}\backslash\{\mathbf{0}\}$*, where* $\mathbf{U} \subset \mathbb{R}^n$ *is a neighborhood including the origin;*

iii) $\dot{V}(\mathbf{x}) = \frac{\partial V}{\partial x}\frac{dx}{dt} \leq 0$ *for all* $\mathbf{x} \in \mathbf{U}$.

With this definition in hand we provide Lyapunov's second method for stability.

Theorem 3.3.4. (Lyapunov's Second Method [102]): *Consider the autonomous dynamical system* $\dot{\mathbf{x}} = F(\mathbf{x})$ *and assume without loss of generality that* $\mathbf{x} = \mathbf{0}$ *is an equilibrium point. If there exists a Lyapunov function V for the system on a neighborhood* $\mathbf{U}$*, then the equilibrium point* $\mathbf{x} = \mathbf{0}$ *is locally stable. Further, if*

$$\dot{V} < 0$$

for all $\mathbf{x} \in \mathbf{U} \backslash \{\mathbf{0}\}$*, the equilibrium point is locally asymptotically stable.*

3.4 Singular Perturbation Theory

Here we consider the stability properties of the singularly perturbed system

$$\dot{\mathbf{x}} = f(t, x, z, \epsilon) \tag{3.2a}$$

$$\epsilon \dot{\mathbf{z}} = g(t, x, z, \epsilon), \tag{3.2b}$$

where $f, g : [t_0, T] \times D_x \times D_z \times [0, \epsilon_0] \to D_x \times D_z$ with $D_x \subset \mathbb{R}^n$ and $D_z \subset \mathbb{R}^m$. Assume that f and g are locally Lipschitz in an domain that includes the origin and that the origin is an equilibrium point. Here the parameter ϵ defines the timescale difference between the component of the network defined by (3.2a) and the component defined by (3.2b). If ϵ is small then the dynamics in $\mathbf{z}$ are faster than those in $\mathbf{x}$. As such we refer to (3.2b) as the 'fast' dynamics, and (3.2a) as the 'slow' dynamics. We are interested in studying the stability of the system (3.2) as $\epsilon \to 0$, which gives a separation of timescales between the components in the dynamics.

In order to study the stability of the full system we define two additional systems when $\epsilon = 0$. First, when $\epsilon = 0$ equation (3.2b) reduces to $0 = g(t, \mathbf{x}, \mathbf{z}, 0)$. Defining the root of this equation to be $\mathbf{z} = h(t, \mathbf{x})$ we define the reduced-order model (ROM)

$$\dot{\mathbf{x}} = f(t, \mathbf{x}, h(t, \mathbf{x}), 0), \tag{3.3}$$

of which we can study the stability independently of the variable $\mathbf{z}$. Using this model and the change of variable $\mathbf{y} = \mathbf{z} - h(t, \mathbf{x})$ we can rewrite the original system (3.2) as

$$\dot{\mathbf{x}} = f(t, \mathbf{x}, \mathbf{y} + h(t, \mathbf{x}), \epsilon) \tag{3.4a}$$

$$\epsilon \dot{\mathbf{y}} = g(t, \mathbf{x}, \mathbf{y} + h(t, \mathbf{x}), \epsilon) - \epsilon \frac{\partial h(t, \mathbf{x})}{\partial t} - \epsilon \frac{\partial h(t, \mathbf{x})}{\partial \mathbf{x}} f(t, \mathbf{x}, \mathbf{y} + h(t, \mathbf{x}), \epsilon) \tag{3.4b}$$

Noting that (3.4b) is not necessarily zero when $\epsilon = 0$, we let $\tau = (t - t_0)/\epsilon$ and define the boundary-layer model (BLM)

$$\frac{d\mathbf{y}}{d\tau} = g(t_0, \mathbf{x}_0, \mathbf{y} + h(t_0, \mathbf{x}_0), 0). \tag{3.5}$$

Theorem 3.4.1. (Tikhonov Stability Theorem [102, Theorem 11.4]): *Consider the singularly perturbed system* (3.2)*. Assume that the following are satisfied for all* $(t, \mathbf{x}, \epsilon) \in [0, \infty) \times B_r \times [0, \epsilon_0]$

- $f(t, \mathbf{0}, \mathbf{0}, \epsilon) = \mathbf{0}$ *and* $g(t, \mathbf{0}, \mathbf{0}, \epsilon) = \mathbf{0}$*.*
- *The equation* $\mathbf{0} = g(t, \mathbf{x}, \mathbf{z}, \mathbf{0})$ *has an isolated root* $\mathbf{z} = h(t, \mathbf{x})$ *such that* $h(t, \mathbf{0}) = \mathbf{0}$*.*
- *The functions* f, g, h *and their partial derivatives up to the second order are bounded for* $\mathbf{z} - h(t, \mathbf{x}) \in B_\rho$ *for* $\rho > 0$*.*
- *The origin of the ROM* (3.3) *is exponentially stable.*
- *The origin of the BLM* (3.5) *is exponentially stable, uniformly in* $(t, \mathbf{x})$*.*

Then, there exists $\epsilon^* > 0$ *such that for all* $\epsilon < \epsilon^*$*, the singularly perturbed system* (3.2) *is exponentially stable to the origin.*

Later we will use a generalization of this result for studying the stability of hierarchical brain networks.

3.5 Preliminary Results on Linear-Threshold Networks

In this section we provide preliminary results regarding the linear-threshold brain model that we consider throughout this dissertation. We consider the linear-threshold network (LTN) dynamics

$$\tau \dot{\mathbf{x}} = -\mathbf{x} + [\mathbf{W}\mathbf{x} + \mathbf{d}(t)]_{\mathbf{0}}^{\mathbf{m}}, \tag{3.6}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the network state, $\mathbf{W} \in \mathbb{R}^{n\times n}$ is the synaptic weight matrix and $\mathbf{d}(t) \in \mathbb{R}^n$ represents any inputs or noise in the network. Here τ is a biological constant defining the timescale of the network. For the following results we will assume that $\mathbf{d}(t)$ is equal to a constant $\mathbf{d}$. The LTN dynamics can be written as a state-dependent xwitched affine system, with regions defined by the term $[\mathbf{W}\mathbf{x}+\mathbf{d}]_{\mathbf{0}}^{\mathbf{m}}$. The system has 3^n switching regions represented by switching variable $\sigma = \{0, \ell, s\}^n$. The regions are defined

$$\Omega_\sigma = \left\{ \mathbf{x} \;\middle|\; \begin{array}{l} (\mathbf{W}\mathbf{x}+\mathbf{d})_i \in (-\infty, 0] \;\forall i \text{ s.t. } \sigma_i = 0, \\ (\mathbf{W}\mathbf{x}+\mathbf{d})_i \in (0, \mathbf{m}_i) \;\forall i \text{ s.t. } \sigma_i = \ell, \text{ and} \\ (\mathbf{W}\mathbf{x}+\mathbf{d})_i \in [\mathbf{m}_i, \infty] \;\forall i \text{ s.t. } \sigma_i = s \end{array} \right.$$

The threshold term in (3.6) can then be expressed over each region using diagomal matrices Σ^ℓ and Σ^s. These are defined, for each $q \in \{\ell, s\}$, as follows: $\Sigma^q_{ii} = 1$ if $\sigma_i = q$ and $\Sigma^q_{ii} = 0$ otherwise. Using these matrices provides the piecewise-affine form of the dynamics (3.6) being defined as

$$\tau\dot{\mathbf{x}} = (-\mathbf{I} + \Sigma^\ell\mathbf{W})\mathbf{x} + \Sigma^\ell\mathbf{d} + \Sigma^s\mathbf{m}, \qquad \mathbf{x} \in \Omega_\sigma. \tag{3.7}$$

Throughout this work we will make the following assumption on the synaptic weight matrices used with the LTN dynamics.

Assumption 3.5.1. *Assume the synaptic weight matrix* $\mathbf{W}$ *satisfies*

1. $\det(\mathbf{W}) \neq 0$;
2. $\det(-\mathbf{I} + \Sigma^\ell\mathbf{W}) \neq 0$ *for each* Σ^ℓ *corresponding to a switching region* Ω_σ.

In practice this is not a restrictive assumption as the set of matrices that violate it has measure zero. Instead it guarantees the right-hand side of (3.7) has unique solutions which allows for the existence of well-defined equilibria. We now provide conditions on the synaptic weight

matrix guaranteeing the existence and uniqueness of equilibria for the LTN dynamics (3.6) for all inputs $\mathbf{d}$.

Theorem 3.5.2. (Existence and Uniqueness of Equilibria in LTN Dynamics [64, Theorem IV.1]): *Consider the LTN network* (3.6)*. The dynamics has a unique equilibrium for all inputs* $\mathbf{d} \in \mathbb{R}^n$ *if and only if* $\mathbf{I} - \mathbf{W} \in \mathcal{P}$.

We note that this result is equivalent for the threshold-linear dynamics that arise when $\mathbf{m} = \infty\mathbf{1}$. The matrix classes introduced in Definition 3.2.1 also provide structural conditions on the synaptic weight matrix for stability of the LTN dynamics (3.6).

Theorem 3.5.3. (Stability of LTN Dynamics [64, Theorem IV.8]): *Consider the LTN network dynamics* (3.6) *and assume* $\mathbf{W}$ *satisfies Assumption 3.5.1. Then the following hold.*

- *If* $\mathbf{W} \in \mathcal{L}$*, then for all* $\mathbf{d} \in \mathbb{R}^n$*, the network is globally exponentially stable to a unique equilibrium* $\mathbf{x}^*$*;*
- *If the network is locally asymptotically stable to a unique equilibrium* $\mathbf{x}^*$ *for all* $\mathbf{d} \in \mathbb{R}^n$*, then* $-\mathbf{I} + \mathbf{W} \in \mathcal{H}$.

This result provides a sufficient condition for global exponential stability and necessary condition for local stability. It is further hypothesized that the necessary condition is also sufficient for global exponential stability [64]. While Theorems 3.5.2 and 3.5.3 give a variety of necessary and sufficient conditions for both existence and uniqueness of equilibria and stability, they are dependent on determining if the synaptic weight matrix is in the classes $\mathcal{P}, \mathcal{H}$ and $\mathcal{L}$. These are difficult conditions to check, particularly as the scale of the network increases. For this reason the matrix inclusions of Lemma 3.2.2 are useful for providing more conservative, but easily computable conditions.

Chapter 4

Conditions for Stability and Oscillations in Linear-Threshold Networks

Oscillatory behavior is one of the most commonly observed phenomena in the brain, appearing in both healthy and pathological states. Within healthy activity, oscillations are linked to phenomena such as cognition [103] and consciousness [104], while also appearing in pathological behaviour such as epileptic seizures and Parkinson's disease [54]. The presence of such oscillations in neural activity motivates the study of the dynamical mechanisms behind their emergence. In contrast, stability in dynamical models of the brain is equally important to understand. Behaviors such as memory [105, 106], selective attention [64, 65] and vision [107] can all be modeled through varying forms of stability in the dynamical model. Further, direct approaches for analysis of oscillations are limited, particularly in higher than 2 dimensions. As such the analysis of oscillations is commonly approached using the proxy of lack of stable equilibria in the network.

Motivated by this, in this chapter we aim to study the oscillatory properties of both the linear-threshold and threshold-linear dynamics, and additionally provide results on the stability of the dynamics beyond those provided in Chapter 3.

4.1 Literature Review

Within the study of oscillatory behavior of the LTN and TLN dynamics, the literature is largely divided based upon network structure. Two of the main structures are excitatory-inhibitory networks and competitive networks (in the latter, all interneuronal interactions are inhibitory). While oscillatory behavior in excitatory-inhibitory networks has been studied extensively [66, 108, 109] using both threshold-linear and linear-threshold dynamics, for competitive networks the literature is largely restricted to threshold-linear dynamics. The works [74, 110] provide both analytic and graph-theoretic conditions for oscillations in a general competitive network governed by threshold-linear dynamics. The works [111, 112, 113] study combinatorial threshold-linear networks, a specific form of competitive networks, and provide conditions related to the existence of dynamic attractors, including limit cycles, and both quasi-periodic and chaotic attractors. These studies explicitly rule out node self-excitation and restrict their attention to all-to-all connectivity structures. This is a major difference with respect to this chapter, where we allow for self-excitation and consider arbitrary network structures.

4.2 Contributions

We study firing rate brain network models described by linear-threshold and threshold-linear dynamics as derived in Chapter 2. Our first contribution pertains to the existence of asymptotically stable equilibria that have non-zero activity on only a given subset of nodes in the network for a general linear-threshold dynamics. We provide a characterization of this in terms of the network structure. We build on this result in our second contribution, which characterizes the emergence of oscillatory behavior in competitive linear-threshold and threshold-linear dynamics. We provide sufficient conditions on the synaptic structure and the input that ensure the network does not have stable equilibria (a fact we use as a proxy for the existence of oscillations).

4.3 Problem Formulation

We consider the linear-threshold and threshold-linear dynamics, given by

$$\tau\dot{\mathbf{x}} = -\mathbf{x} + [\mathbf{W}\mathbf{x} + \mathbf{u}]_{\mathbf{0}}^{\mathbf{m}} \tag{4.1}$$

$$\tau\dot{\mathbf{x}} = -\mathbf{x} + [\mathbf{W}\mathbf{x} + \mathbf{u}]_{+}, \tag{4.2}$$

respectively. In this chapter we assume that the populations of neurons have similar timescales and take $\tau = 1$ without loss of generality.

In firing rate models such as linear-threshold or threshold-linear networks, the network structure determined by the synaptic weight matrix $\mathbf{W}$ is classified based upon the properties of the interneuron connections. An interneuron connection is *excitatory* if the corresponding element in the synaptic weight matrix is positive, and is *inhibitory* if the entry is negative. The diagonal elements in the synaptic weight matrix represent the impact a node's activity has on itself, which we refer to as *self-excitatory* if the matrix value is positive, and *self-inhibitory* if the matrix value is negative. A particular class of networks to which we pay attention are *competitive* networks, which represent inhibition-based competition between brain regions, a widely-observed phenomenon [114].

Definition 4.3.1 (Competitive Network)**.** *Consider a linear-threshold (resp., threshold-linear) network defined by synaptic weight matrix* $\mathbf{W}$*. The network is competitive if* $\mathbf{W}$ *is a Z-matrix (i.e., all interneuron connections are inhibitory) and the nodes are either self-excitatory or not self-connected (*$w_{ii} \geq 0$ *for all i).*

This definition generalizes the standard definition of competitive network, e.g., [110], which requires all diagonal elements to be zero. As oscillations in the brain are widely associated with inhibition [115], our goal is to determine conditions under which oscillations arise in competitive brain networks. We note that while these networks no longer satisfy Dale's Law, we make the same assumption as in [116], in which the network describes effective connections

between excitatory nodes, and inhibition in the network is mediated by intermediate connections with interneurons.

The notion of neural oscillation we consider here goes beyond periodic trajectories to also include chaotic behavior, as chaotic trajectories are of significant interest in computational neuroscience [117]. Formally, we say a trajectory $\mathbf{x}(t)$ of the LTN (4.1) or TLN (4.2) dynamics is oscillatory if it does not converge asymptotically to an equilibrium.

In this chapter, we use the lack of stable equilibria (LoSE) as a proxy for the existence of oscillations. This is because this criterion is widely applicable, whereas analytic tools for directly studying oscillations (such as Poincarè-Bendixson theory [118]) are limited to 2-dimensional systems or ones whose behavior can be confined to two dimensions. For the LTN dynamics, this proxy has been shown to be tight [66]. We formalize the problem considered as follows.

Problem 4.3.2. *Consider a competitive LTN (resp. TLN) with synaptic weight matrix* $\mathbf{W}$. *Determine conditions on the structure of* $\mathbf{W}$ *and the input vector* $\mathbf{u}$ *such that the network has no stable equilibria.*

4.4 Stable Equilibria in LTN and TLN Networks

This section studies the conditions for the existence of stable equilibria in a general network topology as a precursor to our focus in Section 4.5 on the study of oscillations in competitive networks. Given the dynamics of LTN (4.1) and TLN (4.2) networks, it is clear that the location and stability of the equilibria depend upon the specific input. This brings up two important observations when characterizing them: (i) stability statements could be made for all possible inputs, several inputs, or just one input. Here, we focus on the latter; (ii) rather than the specific location of the equilibria, we focus on its support. This means that we consider equivalence classes of equilibria, as the same set of nodes could correspond to many different actual equilibria.

The following definition makes this precise for TLN networks.

Definition 4.4.1. (Stable Node Set in TLN Dynamics [74]): *Consider a network governed by the threshold-linear dynamics* (4.2)*. A non-empty subset of nodes* $\sigma \subseteq \{1,\dots,N\}$ *is* stable *if there exists an asymptotically stable equilibrium point* $\mathbf{x}^*$ *such that supp*$(\mathbf{x}^*) = \sigma$ *for at least one input* $\mathbf{u} \in \mathbb{R}^n$.

For LTN dynamics, since they are guaranteed to have bounded trajectories, this definition becomes trivial: for any subset σ, there always exists an input $\mathbf{u} \in \mathbb{R}^n$ such that the point $(\mathbf{0}, \mathbf{m}_\sigma)$ is an asymptotically stable equilibrium point (since, for $\mathbf{u}$ with $\mathbf{u}_\sigma$ large enough and $\mathbf{u}_{\bar{\sigma}}$ small enough, the dynamics reduces to $\dot{\mathbf{x}}_\sigma = -\mathbf{x}_\sigma + \mathbf{m}_\sigma$ and $\dot{\mathbf{x}}_{\bar{\sigma}} = -\mathbf{x}_{\bar{\sigma}}$). In order to extend our treatment of stable node sets for LTN networks, we first consider the support of a bounded vector.

Definition 4.4.2. (Support of a Bounded Vector): *Let* $\mathbf{x} \in [\mathbf{0}, \mathbf{m}] \subset \mathbb{R}^n$*, with* $\mathbf{m} \in \mathbb{R}^n_{>0}$*. The support of* $\mathbf{x}$ *is the set* $\sigma = (\sigma_{\mathbf{m}}, \sigma_{\mathring{\mathbf{m}}}) \subseteq \{1,\dots,N\}$*, where* $\mathbf{x}_i = \mathbf{m}_i$ *for all* $i \in \sigma_{\mathbf{m}}$*,* $\mathbf{x}_i \in (0, \mathbf{m}_i)$ *for all* $i \in \sigma_{\mathring{\mathbf{m}}}$ *and* $\mathbf{x}_i = 0$ *for all* $i \in \bar{\sigma}$.

The synaptic weight matrix can be block-partitioned according to the the support σ as

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{\bar{\sigma}} & \mathbf{W}_{\bar{\sigma}\sigma_{\mathbf{m}}} & \mathbf{W}_{\bar{\sigma}\sigma_{\mathring{\mathbf{m}}}} \\ \mathbf{W}_{\sigma_{\mathbf{m}}\bar{\sigma}} & \mathbf{W}_{\sigma_{\mathbf{m}}} & \mathbf{W}_{\sigma_{\mathbf{m}}\sigma_{\mathring{\mathbf{m}}}} \\ \mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}\bar{\sigma}} & \mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}\sigma_{\mathbf{m}}} & \mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}} \end{bmatrix}, \tag{4.3}$$

where $(\bar{\sigma}, \sigma_{\mathbf{m}}, \sigma_{\mathring{\mathbf{m}}}) = \{1,\dots,N\}$. Next, we have the following notion of stability of node sets in LTN networks.

Definition 4.4.3. (Non-trivially Stable Node Set in LTN Dynamics): *Consider a network defined by the linear-threshold dynamics* (4.1) *with synaptic weight matrix* $\mathbf{W}$ *and upper bound* $\mathbf{m}$*. A non-empty subset of nodes* $\sigma = (\sigma_{\mathbf{m}}, \sigma_{\mathring{\mathbf{m}}}) \subseteq \{1,\dots,N\}$ *is* non-trivially stable *if there exists an asymptotically stable equilibrium point* $\mathbf{x}^*$ *for the dynamics with supp*$(\mathbf{x}^*) = \sigma = (\sigma_{\mathbf{m}}, \sigma_{\mathring{\mathbf{m}}})$ *for at least one input* $\mathbf{u} \in \mathbb{R}^n$ *and either* $\sigma_{\mathring{\mathbf{m}}} \neq \emptyset$ *or there exists* $i \in \sigma_{\mathbf{m}}$ *such that* $(\mathbf{W}\mathbf{x}^* + \mathbf{u})_i = \mathbf{m}_i$.

The key part of Definition 4.4.3 is that existence of the stable equilibrium cannot be

guaranteed solely on the basis of forced saturation by the input. By requiring that either: one of the non-zero components in the equilibrium point is not saturated ($\sigma_{\mathring{\mathbf{m}}} \neq \emptyset$); or, if it is at the saturation value, it is not over-saturated ($(\mathbf{W}\mathbf{x}^* + \mathbf{u})_i = \mathbf{m}_i$), it guarantees that the equilibrium is dependent on the structure of the network and the dynamic behavior, rather than the input. This definition, when applied to a TLN network reduces to Definition 4.4.1.

We next give a condition for the existence of non-trivially stable equilibria.

Theorem 4.4.4. (Existence of Non-trivially Stable Node Set): *Consider a network defined by either LTN or TLN dynamics with synaptic weight matrix* $\mathbf{W}$ *and upper bound* $\mathbf{m}$*. A subset of nodes* $\sigma = (\sigma_{\mathbf{m}}, \sigma_{\mathring{\mathbf{m}}})$ *is non-trivially stable with associated equilibrium* $\mathbf{x}^*$*, with* $\mathbf{x}_i^* \in (0, \mathbf{m}_i)$ *for all* $i \in \sigma_{\mathring{\mathbf{m}}}$ *and* $\mathbf{x}_i^* = \mathbf{m}_i$ *for all* $i \in \sigma_{\mathbf{m}}$ *if and only if the matrix* $(-\mathbf{I} + \mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}}$ *is stable.*

Proof of Theorem 4.4.4. For TLN dynamics, the result corresponds to [116, Theorem 1.2]. Hence, we focus on LTN dynamics. First suppose that $(-\mathbf{I} + \mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}}$ is stable. Let $\mathbf{x}^*_\sigma$ be the vector with support σ such that $\mathbf{x}^*_{\sigma_{\mathbf{m}}} = \mathbf{m}_{\sigma_{\mathbf{m}}}$ and $\mathbf{x}^*_{\sigma_{\mathring{\mathbf{m}}}} = \alpha \mathbf{m}_{\sigma_{\mathring{\mathbf{m}}}}$, where $\alpha \in (0,1)$ is arbitrary. Define $\mathbf{u}_{\sigma_{\mathring{\mathbf{m}}}} = \alpha(\mathbf{I} - \mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}} \mathbf{m}_{\sigma_{\mathring{\mathbf{m}}}}$. Choose $\mathbf{u}_{\bar{\sigma}}$ such that $\mathbf{u}_{\bar{\sigma}} < -\mathbf{W}_{\bar{\sigma}\sigma}\mathbf{x}^*_\sigma$ and $\mathbf{u}_{\sigma_{\mathbf{m}}}$ such that $\mathbf{u}_{\sigma_{\mathbf{m}}} > -\mathbf{W}_{\sigma_{\mathbf{m}}\sigma}\mathbf{x}^*_\sigma + \mathbf{m}_{\sigma_{\mathbf{m}}}$. With this choice of $\mathbf{u}$, the vector $\mathbf{x}^*_\sigma$ satisfies $(\mathbf{I} - \mathbf{W})_\sigma \mathbf{x}^*_\sigma = \mathbf{u}_\sigma$ and is therefore an equilibrium. We next prove it is stable.

To do so, consider the following change of variables. Define $(\mathbf{q}, \mathbf{y}, \mathbf{z}) = \mathbf{x} - \mathbf{x}^*$, with $\mathbf{q} = (\mathbf{x} - \mathbf{x}^*)_{\bar{\sigma}} \in \mathbb{R}^{|\bar{\sigma}|}$, $\mathbf{y} = (\mathbf{x} - \mathbf{x}^*)_{\sigma_{\mathbf{m}}} \in \mathbb{R}^{|\sigma_{\mathbf{m}}|}$ and $\mathbf{z} = (\mathbf{x} - \mathbf{x}^*)_{\sigma_{\mathring{\mathbf{m}}}} \in \mathbb{R}^{|\sigma_{\mathring{\mathbf{m}}}|}$. These variables represent the components of the dynamics corresponding to where the equilibrium is equal to 0, are on the boundary $\mathbf{m}$, and are in the interior of $[\mathbf{0}, \mathbf{m}]$, respectively. This change of variables shifts the equilibrium $\mathbf{x}^*$ to the origin, and the system becomes

$$\dot{\mathbf{q}} = -\mathbf{q} + [\mathbf{W}_{\bar{\sigma}}\mathbf{q} + \mathbf{W}_{\bar{\sigma}\sigma_{\mathbf{m}}}\mathbf{y} + \mathbf{W}_{\bar{\sigma}\sigma_{\mathring{\mathbf{m}}}}\mathbf{z} + (\mathbf{W}_{\bar{\sigma}\sigma}\mathbf{x}^*_\sigma + \mathbf{u}_{\bar{\sigma}})]_{\mathbf{0}}^{\mathbf{m}_{\bar{\sigma}}} \tag{4.4a}$$

$$\dot{\mathbf{y}} = -(\mathbf{y} + \mathbf{m}_{\sigma_{\mathbf{m}}}) + [\mathbf{W}_{\sigma_{\mathbf{m}}\bar{\sigma}}\mathbf{q} + \mathbf{W}_{\sigma_{\mathbf{m}}}\mathbf{y} + \mathbf{W}_{\sigma_{\mathbf{m}}\sigma_{\mathring{\mathbf{m}}}}\mathbf{z} + (\mathbf{W}_{\sigma_{\mathbf{m}}\sigma}\mathbf{x}^*_\sigma + \mathbf{u}_{\sigma_{\mathbf{m}}})]_{\mathbf{0}}^{\mathbf{m}_{\sigma_{\mathbf{m}}}} \tag{4.4b}$$

$$\dot{\mathbf{z}} = -(\mathbf{z} + \mathbf{x}^*_{\sigma_{\mathring{\mathbf{m}}}}) + [\mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}\bar{\sigma}}\mathbf{q} + \mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}\sigma_{\mathbf{m}}}\mathbf{y} + \mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}}\mathbf{z} + (\mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}\sigma}\mathbf{x}^*_\sigma + \mathbf{u}_{\sigma_{\mathring{\mathbf{m}}}})]_{\mathbf{0}}^{\mathbf{m}_{\sigma_{\mathring{\mathbf{m}}}}}. \tag{4.4c}$$

Note that with our choice of $\mathbf{u}$ above, the constant terms satisfy $\mathbf{W}_{\bar{\sigma}\sigma}\mathbf{x}^*_{\sigma} + \mathbf{u}_{\bar{\sigma}} < \mathbf{0}$, $\mathbf{W}_{\sigma_{\mathbf{m}}}\mathbf{x}^*_{\sigma_{\mathbf{m}}} + \mathbf{u}_{\sigma_{\mathbf{m}}} > \mathbf{m}_{\sigma_{\mathbf{m}}}$ and $\mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}}\mathbf{x}^*_{\sigma_{\mathring{\mathbf{m}}}} + \mathbf{u}_{\sigma_{\mathring{\mathbf{m}}}} \in (\mathbf{0}, \mathbf{m}_{\sigma_{\mathring{\mathbf{m}}}})$. It follows that in a neighborhood of the origin, the sign of the threshold terms are determined solely by the sign of the constant term. The behavior of the system (4.4) is determined by the linear system $\frac{d}{dt}[\mathbf{q}, \mathbf{y}, \mathbf{z}] = \mathbf{W}[\mathbf{q}, \mathbf{y}, \mathbf{z}]^\top$ where $\mathbf{W}$ takes the form (4.3), and in particular is lower triangular with diagonal elements $-\mathbf{I}_{\bar{\sigma}}, -\mathbf{I}_{\sigma_{\mathbf{m}}}$ and $(-\mathbf{I} + \mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}}$. Then, since $(-\mathbf{I} + \mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}}$ is stable, the equilibrium point is stable, and therefore the subset of nodes σ is non-trivially stable.

Now, suppose that $(-\mathbf{I} + \mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}}$ is not stable. We reason by contradiction. Assume $\sigma = (\sigma_{\mathbf{m}}, \sigma_{\mathring{\mathbf{m}}})$ is a non-trivially stable node set. This means that there exists an input $\mathbf{u}$ such that $\mathbf{x}^*$ is a stable equilibrium point with $\mathbf{W}_{\bar{\sigma}\sigma}\mathbf{x}^*_{\sigma} + \mathbf{u}_{\bar{\sigma}} \leq \mathbf{0}$, $\mathbf{W}_{\sigma_{\mathbf{m}}}\mathbf{x}^*_{\sigma_{\mathbf{m}}} + \mathbf{u}_{\sigma_{\mathbf{m}}} \geq \mathbf{m}_{\sigma_{\mathbf{m}}}$ and $\mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}}\mathbf{x}^*_{\sigma_{\mathring{\mathbf{m}}}} + \mathbf{u}_{\sigma_{\mathring{\mathbf{m}}}} \in (\mathbf{0}, \mathbf{m}_{\sigma_{\mathring{\mathbf{m}}}})$. Now, since $\mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}}\mathbf{x}^*_{\sigma_{\mathring{\mathbf{m}}}} + \mathbf{u}_{\sigma_{\mathring{\mathbf{m}}}} \in (\mathbf{0}, \mathbf{m}_{\sigma_{\mathring{\mathbf{m}}}})$, in a neighborhood of $\mathbf{x}^*$, the component of the dynamics $\mathbf{z}$ acts linearly as $\dot{\mathbf{z}} = -(\mathbf{I} + \mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}}\mathbf{z} + \mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}\bar{\sigma}}\mathbf{q} + \mathbf{W}_{\sigma_{\mathring{\mathbf{m}}}\sigma_{\mathbf{m}}}\mathbf{y}$. Since $(-\mathbf{I} + \mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}}$ is not stable, it then follows that $\mathbf{x}^*$ is not a stable equilibrium, providing a contradiction. This completes the proof. □

The characterization in Theorem 4.4.4 for the existence of a stable equilibrium for an arbitrary node set under LTN and TLN dynamics is useful for identifying and building networks that possess such equilibria. Conversely, it can also be used for the opposite purpose: identify and build networks that do not. The latter is aligned with seeing the LoSE as a proxy for the existence of oscillatory or chaotic behavior. As such, in the ensuing discussion we focus on identifying conditions on the network structure and the input that ensure that the characterization of Theorem 4.4.4 is *not* satisfied.

4.5 Oscillations in Competitive Networks

In this section, we focus on competitive networks and provide conditions on the structure of the synaptic weight matrix $\mathbf{W}$ and the input $\mathbf{u}$ such that a competitive linear-threshold or threshold-linear network lacks stable equilibria, thus satisfying our criteria for enabling

oscillations. We tackle the problem of LoSE by classifying equilibria by their support: equilibria supported on two or more nodes in the interior of $[\mathbf{0}, \mathbf{m}]$; equilibria supported on a single node; and equilibria with components lying on the boundary $\mathbf{m}$. We then provide conditions such that all equilibria in each class is not stable.

Theorem 4.5.1. (Oscillations in Competitive Networks with LTN Dynamics): *Consider a network defined by the LTN dynamics* (4.1) *with synaptic weight matrix* $\mathbf{W}$*, upper bound* $\mathbf{m}$*, and constant input* $\mathbf{u} \in \mathbb{R}^n$*. Let* $\mathbf{W}$ *be a Z-matrix with at least two diagonal elements, indexed* i_1, i_2*, such that* $w_{i_k i_k} < 1$ *and* $2|-1+w_{i_k i_k}| > \rho(-\mathbf{I}+\mathbf{W})$*, for* $k \in \{1,2\}$*. The following statements hold:*

1. *There are no stable equilibria* $\mathbf{x}^*$ *with* $|supp(\mathbf{x}^*)| \geq 2$ *and* $\mathbf{x}_i^* \in (\mathbf{0}, \mathbf{m})$ *for all* $i \in supp(\mathbf{x}^*)$ *if all* 2×2 *principal submatrices of* $-\mathbf{I}+\mathbf{W}$ *are unstable;*

2. *There are no equilibria* $\mathbf{x}^*$ *with* $|supp(\mathbf{x}^*)| = 1$ *and*

   - $\mathbf{x}_i^* = \mathbf{m}_i$ *if, for each* $i \in \{1,\dots,N\}$*, there exists* k *such that* $\mathbf{u}_k > -w_{ki}\mathbf{m}_i$*;*
   - $\mathbf{x}_i^* \in (0, \mathbf{m}_i)$ *if, for each* $i \in \{1,\dots,N\}$ *with* $w_{ii} \neq 1$*, there exists* k *such that* $\text{sign}(\mathbf{u}_i)\frac{\mathbf{u}_k}{\mathbf{u}_i} > -\text{sign}(\mathbf{u}_i)\frac{w_{ki}}{w_{ii}-1}$ *and for each* $i \in \{1,\dots,N\}$ *with* $w_{ii} = 1$ *either* $\mathbf{u}_i \neq 0$ *or* $\exists k \neq i$ *such that* $\mathbf{u}_k > -w_{ki}\mathbf{m}_i$*;*

3. *Consider a node set* $\sigma = (\sigma_{\mathbf{m}}, \sigma_{\mathring{\mathbf{m}}})$ *with* $|\sigma| \geq 2$ *and* $|\sigma_{\mathbf{m}}| \geq 1$*. The following hold:*

   - *If* $|\sigma_{\mathring{\mathbf{m}}}| \geq 2$*, then there do not exist any stable equilibria* $\mathbf{x}^*$ *with support* σ *if all* 2×2 *principal submatrices of* $(-\mathbf{I}+\mathbf{W})_{\sigma_{\mathring{\mathbf{m}}}}$ *are unstable;*
   - *If* $|\sigma| = |\sigma_{\mathbf{m}}| \geq 2$*, then there do not exist any equilibria* $\mathbf{x}^*$ *with support* σ^* *if there exists* $i \in \sigma_{\mathbf{m}}$ *such that* $\mathbf{u}_i < \mathbf{m}_i - \sum_{j\in\sigma_m} w_{ij}\mathbf{m}_j$*;*
   - *If* $\sigma_{\mathring{\mathbf{m}}} = \{i\}$*, then there do not exist any equilibria* $\mathbf{x}^*$ *with support* σ *if* $\mathbf{u}$ *is such that one or more of the following conditions hold:*

     (a) There exists $i \in \sigma_{\mathring{\mathbf{m}}}$ *such that:*

     i. If $w_{ii} < 1$*, then* $\mathbf{u}_i \notin (-\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_i, \mathbf{m}_i(1-w_{ii}) - \sum_{j\in\sigma_m} w_{ij}\mathbf{m}_i)$*, or*

ii. *If* $w_{ii} > 1$, *then* $\mathbf{u}_i \notin (\mathbf{m}_i(1-w_{ii}) - \sum_{j\in\sigma_m} w_{ij}\mathbf{m}_i, -\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_i)$, *or*

iii. *If* $w_{ii} = 1$, *then* $\mathbf{u}_i \neq -\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_i$.

(b) There exists $k \in \sigma_{\mathbf{m}}$ *such that if* $w_{ii} \neq 1$

$$\mathbf{u}_k < \mathbf{m}_k - \sum_{j\in\sigma_{\mathbf{m}}} w_{kj}\mathbf{m}_k - \Big(\frac{w_{ki}}{1-w_{ii}}\Big)\Big(\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_i + \mathbf{u}_i\Big),$$

or if $w_{ii} = 1$

$$\mathbf{u}_k \leq \mathbf{m}_k - \sum_{j\in\sigma_{\mathbf{m}}} w_{kj}\mathbf{m}_j.$$

(c) There exists $l \in \bar{\sigma}$ *such that if* $w_{ii} \neq 1$

$$\mathbf{u}_l > -\sum_{j\in\sigma_{\mathbf{m}}} w_{lj}\mathbf{m}_l - \Big(\frac{w_{li}}{1-w_{ii}}\Big)\Big(\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_i + \mathbf{u}_i\Big),$$

or if $w_{ii} = 1$

$$\mathbf{u}_l \geq -\sum_{j\in\sigma_{\mathbf{m}}} w_{lj}\mathbf{m}_j - w_{li}\mathbf{m}_i.$$

To prove this statement, the following result guaranteeing LoSE supported on multiple nodes is useful.

Lemma 4.5.2. (Conditions for Unstable Equilibria Supported on Multiple Nodes): *Consider a network defined by the LTN dynamics* (4.1) *with a Z-matrix* $\mathbf{W}$. *If all* 2×2 *principal submatrices of* $-\mathbf{I}+\mathbf{W}$ *are unstable and there exist* i_1, i_2 *with* $w_{i_k i_k} < 1$ *and* $2|-1+w_{i_k i_k}| > \rho(-\mathbf{I}+\mathbf{W})$ *for* $k \in \{1,2\}$, *then the network has no stable equilibria supported on more than one node in the interior of* $[\mathbf{0}, \mathbf{m}]$.

This result is a generalization of [74, Corollary 4.4] to the case of synaptic weight matrices with non-zero elements in the diagonal. The next result is useful in our proof of

Lemma 4.5.2.

Lemma 4.5.3. *Let* $\mathbf{W} \in \mathbb{R}^{n\times n}$ *be a stable Z-matrix with two or more negative diagonal elements* $w_{i_1i_1} < w_{i_2i_2} < 0$ *such that* $2|w_{i_ki_k}| > \rho(\mathbf{W})$ *for* $k \in \{1,2\}$. *Then* $\mathbf{W}$ *has a stable* 2×2 *principal submatrix.*

Proof of Lemma 4.5.3. As $\mathbf{W}$ is a Z-matrix, we can write it as $\mathbf{W} = \alpha\mathbf{I} - \mathbf{P}$, where $\alpha = \max_i\{w_{ii}\}$ and $\mathbf{P}$ is a non-negative matrix. Since $\mathbf{W}$ is a stable Z-matrix, $\rho(\mathbf{P}) > \alpha$ and $\rho(\mathbf{W}) = \rho(\mathbf{P}) - \alpha$.

Without loss of generality, assume that the two smallest diagonal elements are $w_{11} < w_{22} < 0$. We then claim the submatrix $\mathbf{W}_{12}$ is stable. We can write this matrix to be

$$\mathbf{W}_{12} = \begin{bmatrix} w_{11} & -\bar{w}_{12} \\ -\bar{w}_{21} & w_{22} \end{bmatrix} = \alpha\mathbf{I} - \mathbf{P}_{12},$$

where $\bar{w}_{12}, \bar{w}_{21} \geq 0$. Since $\mathbf{P}$ is a non-negative matrix, and $\mathbf{P}_{12}$ is a principal submatrix, we have $\rho(\mathbf{P}_{12}) < \rho(\mathbf{P})$ [119, Corollary 8.1.20] and therefore $\rho(\mathbf{W}) > \rho(\mathbf{W}_{12})$. On the other hand, note that since $\mathbf{W}_{12}$ is a Z-matrix, for $\gamma \geq \rho(\mathbf{W}_{12})$, we have $\mathbf{W}_{12} + \gamma\mathbf{I}$ is a M-matrix. Since $2|w_{22}| > \rho(\mathbf{W}_{12})$, the matrix $\mathbf{B}_{12} = \mathbf{W}_{12} + \gamma\mathbf{I}$ with $\gamma = 2|w_{22}|$, given as follows,

$$\mathbf{B}_{12} = \begin{bmatrix} w_{11} + 2|w_{22}| & -\bar{w}_{12} \\ -\bar{w}_{21} & |w_{22}| \end{bmatrix},$$

is a M-matrix. Since $w_{11} < w_{22}$, it follows that $|w_{11}| > w_{11} + 2|w_{22}|$. Therefore the matrix $\tilde{\mathbf{B}}_{12}$ defined by

$$\tilde{\mathbf{B}}_{12} = \begin{bmatrix} |w_{11}| & -\bar{w}_{12} \\ -\bar{w}_{21} & |w_{22}| \end{bmatrix},$$

satisfies $\tilde{\mathbf{B}}_{12} \geq \mathbf{B}_{12}$ and, by [120, Theorem 4.6], is a M-matrix. Then since $\tilde{\mathbf{B}}_{12} \geq \mathbf{W}_{12}$ and the diagonal elements of $\tilde{\mathbf{B}}_{12}$ are equal to the absolute value of the diagonal elements of $\mathbf{W}_{12}$,

by [121, Section 2.5, Problem 34b)], the eigenvalues of $\mathbf{W}_{12}$ satisfy $\lambda(\mathbf{W}_{12}) < -\lambda_{\min}(\tilde{\mathbf{B}}_{12}) < 0$, proving $\mathbf{W}_{12}$ is stable. □

We are now ready to prove Lemma 4.5.2.

Proof of Lemma 4.5.2. We recall from Theorem 4.4.4 that a subset of nodes $\sigma \subseteq \{1,\dots,N\}$ supports a stable equilibrium in the interior of the range $[\mathbf{0},\mathbf{m}]$ iff the matrix $(-\mathbf{I}+\mathbf{W})_\sigma$ is stable. Since all 2×2 principal submatrices of $-\mathbf{I}+\mathbf{W}$ are unstable, it is immediate that there are no stable equilibria supported on subsets of nodes σ with $|\sigma| = 2$ taking values in the interior of $[\mathbf{0},\mathbf{m}]$. It remains to be shown that this holds for all subsets of nodes $\sigma \subseteq \{1,\dots,N\}$ with $|\sigma| \geq 3$. We reason by contradiction. Suppose $(-\mathbf{I}+\mathbf{W})_\sigma$ is stable. Then, since there exist i_1, i_2 with $w_{i_k i_k} < 1$ and $2|-1+w_{i_k i_k}| > \rho(-\mathbf{I}+\mathbf{W})$ for $k \in \{1,2\}$, we can invoke Lemma 4.5.3 to ensure that $-\mathbf{I}+\mathbf{W}$ has exists a stable 2×2 principal submatrix. This contradicts our assumption that all its 2×2 principal submatrices are unstable, and therefore $(-\mathbf{I}+\mathbf{W})_\sigma$ cannot be stable, implying that σ, with $|\sigma| \geq 3$, does not support a stable equilibrium taking values in the interior of $[\mathbf{0},\mathbf{m}]$. □

We are now ready to prove Theorem 4.5.1.

Proof of Theorem 4.5.1. We proceed by deriving conditions so that each classification of equilibria by their support contains no stable equilibria.

Statement 1): It directly follows from Lemma 4.5.2: if all 2×2 principal submatrices of $-\mathbf{I}+\mathbf{W}$ are unstable[1], then there exist no stable equilibria supported on more than two nodes in the interior of $[\mathbf{0},\mathbf{m}]$.

Statement 2): Without loss of generality, assume $\mathbf{x}^*$ is a potential equilibrium point supported only on node i. Then the equilibrium equations are $\mathbf{x}_i^* = [w_{ii}\mathbf{x}_i^* + \mathbf{u}_i]_0^{\mathbf{m}_i}$ and $0 = [w_{ki}\mathbf{x}_i^* + \mathbf{u}_k]_0^{\mathbf{m}_k}$, for all $k \neq i$. These conditions are satisfied iff $w_{ki}\mathbf{x}_i^* + \mathbf{u}_k \leq 0$ for all $k \neq i$. There are three possible cases for the remaining equation:

- $w_{ii}\mathbf{x}_i^* + \mathbf{u}_i \geq \mathbf{m}_i$, which implies $\mathbf{x}_i^* = \mathbf{m}_i$;

[1] Lemma 3.2.4 provides detailed expressions for this condition to hold.

- $w_{ii}\mathbf{x}_i^* + \mathbf{u}_i \in (\mathbf{0}, \mathbf{m}_i)$, which implies $\mathbf{x}_i^* = \frac{\mathbf{u}_i}{1-w_{ii}}$ unless $w_{ii} = 1$, in which case $\mathbf{u}_i = 0$ and $\mathbf{x}_i^*$ can take any value within the interval;

- $w_{ii}\mathbf{x}_i^* + \mathbf{u}_i \leq 0$, which implies $\mathbf{x}_i^* = 0$, contradicting the assumption that $\mathbf{x}^*$ is supported on node i.

Therefore, we need to consider only the first two cases. For the case in which $\mathbf{x}_i^* = \mathbf{m}_i$, $\mathbf{x}^*$ is an equilibrium iff, for all $k \neq i$, it holds that $w_{ki}\mathbf{m}_i + \mathbf{u}_k \leq 0$. As such there is no equilibrium supported on a single node at the boundary if for all $i \in \{1, \ldots, N\}$ there exists $k \neq i$ such that $\mathbf{u}_k > -w_{ki}\mathbf{m}_i$. For the case where $\mathbf{x}_i^* \in (0, \mathbf{m}_i)$, first consider when $w_{ii} = 1$. In this case, $\mathbf{x}^*$ is an equilibrium for any $\mathbf{x}_i^* \in (0, \mathbf{m}_i)$ iff $\mathbf{u}_i = 0$ and $w_{ki}\mathbf{x}_i^* + \mathbf{u}_k \leq 0$ for all $k \neq i$. Rearranging these conditions, we get that there is no equilibrium supported on a single node $\mathbf{x}_i^* \in (0, \mathbf{m}_i)$ if either $\mathbf{u}_i \neq 0$ or there exists $k \neq i$ such that $\mathbf{u}_k > -w_{ki}\mathbf{m}_i$. If $w_{ii} \neq 1$, $\mathbf{x}^*$ is an equilibrium iff for all $k \neq i$

$$w_{ki}\left(\frac{\mathbf{u}_i}{1-w_{ii}}\right) + \mathbf{u}_k \leq 0.$$

Rearranging this expression, we get that there is no equilibrium supported on a single node with $\mathbf{x}_i^* \in (0, \mathbf{m}_i)$ if, for each $i \in \{1, \ldots, N\}$ with $w_{ii} \neq 1$, there exists $k \neq i$ such that

$$\text{sign}(\mathbf{u}_i)\frac{\mathbf{u}_k}{\mathbf{u}_i} > -\text{sign}(\mathbf{u}_i)\frac{w_{ki}}{1-w_{ii}}.$$

Statement 3): Equilibria supported on two or more nodes, with at least one taking values on the boundary $\mathbf{m}$, come in the following three cases based upon the structure of the node set $\sigma = (\sigma_{\mathbf{m}}, \sigma_{\mathring{\mathbf{m}}})$. The cases are: a) $|\sigma_{\mathbf{m}}| \geq 1, |\sigma_{\mathring{\mathbf{m}}}| \geq 2$; b) $|\sigma| = |\sigma_{\mathbf{m}}| \geq 2$; c) $|\sigma_{\mathbf{m}}| \geq 1$ and $|\sigma_{\mathring{\mathbf{m}}}| = 1$. Consequently,

a) If $|\sigma_{\mathring{\mathbf{m}}}| \geq 2$, these are equilibria supported on two or more nodes in the interior of $[\mathbf{0}, \mathbf{m}]$. By Lemma 4.5.2, if all 2×2 principal submatrices of $(-\mathbf{I} + \mathbf{W})$ are unstable, no such equilibrium is stable.

b) If $|\sigma| = |\sigma_{\mathbf{m}}| \geq 2$, then a point $\mathbf{x}^*$ with support σ is an equilibrium if and only if $\mathbf{m}_i \leq \sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{x}_j^* + \mathbf{u}_i$ for all $i \in \sigma_{\mathbf{m}}$. Rearranging this expression, we have that there is no equilibrium with this structure if there exists $i \in \sigma_m$ such that $\mathbf{u}_i < \mathbf{m}_i - \sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_j$;

c) If $|\sigma_{\mathring{\mathbf{m}}}| = 1$ and $|\sigma_{\mathbf{m}}| \geq 1$, let $\mathbf{x}^*$ be a candidate point with node i supported on $(0, \mathbf{m}_i)$ and let us identify the interval of inputs $\mathbf{u}$ that would actually make $\mathbf{x}^*$ an equilibrium. For $\mathbf{x}^*$ to be an equilibrium, the following must hold:

$$\mathbf{x}_i^* = \sum_{j\in\sigma} w_{ij}\mathbf{x}_j^* + \mathbf{u}_i, \tag{4.5a}$$

$$\mathbf{m}_k \leq \sum_{j\in\sigma} w_{kj}\mathbf{x}_j^* + \mathbf{u}_k, \qquad \forall k \in \sigma_{\mathbf{m}}, \tag{4.5b}$$

$$0 \geq \sum_{j\in\sigma} w_{lj}\mathbf{x}_j^* + \mathbf{u}_l, \qquad \forall l \in \bar{\sigma}. \tag{4.5c}$$

From (4.5a), we get

$$(1 - w_{ii})\mathbf{x}_i^* = \sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_j + \mathbf{u}_i. \tag{4.6}$$

To enforce the constraint that $\mathbf{x}_i^* \in (0, \mathbf{m}_i)$, the input $\mathbf{u}_i$ must satisfy the following:

- If $w_{ii} < 1$, then $\mathbf{u}_i \in (-\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_j, \mathbf{m}_i(1 - w_{ii}) - \sum_{j\in\sigma_m} w_{ij}\mathbf{m}_j)$;
- If $w_{ii} > 1$, then $\mathbf{u}_i \in (\mathbf{m}_i(1 - w_{ii}) - \sum_{j\in\sigma_m} w_{ij}\mathbf{m}_j, -\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_j)$;
- If $w_{ii} = 1$, then $\mathbf{u}_i = -\sum_{j\in\sigma_m} w_{ij}\mathbf{m}_j$.

First suppose $w_{ii} \neq 1$, and considering (4.5b), by rearranging and substituting (4.6), we get

$$\mathbf{u}_k \geq \mathbf{m}_k - \sum_{j\in\sigma_{\mathbf{m}}} w_{kj}\mathbf{m}_j - \Big(\frac{w_{ki}}{1 - w_{ii}}\Big)\Big(\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_j + \mathbf{u}_i\Big),$$

must be satisfied for all $k \in \sigma_{\mathbf{m}}$ in order for $\mathbf{x}^*$ to be an equilibrium. Similarly, from (4.5c),

$$\mathbf{u}_l \leq -\sum_{j\in\sigma_{\mathbf{m}}} w_{lj}\mathbf{m}_j - \Big(\frac{w_{li}}{1-w_{ii}}\Big)\Big(\sum_{j\in\sigma_{\mathbf{m}}} w_{ij}\mathbf{m}_j + \mathbf{u}_i\Big),$$

must be satisfied for all $l \in \bar{\sigma}$ for $\mathbf{x}^*$ to be an equilibrium. If instead $w_{ii} = 1$, by considering (4.5b) and noting from (4.6) that all $\mathbf{x}_i^* \in (0, \mathbf{m}_i)$ are possible equilibrium values, we get that there exists an equilibrium iff $\mathbf{u}_k > \mathbf{m}_k - \sum_{j\in\sigma_{\mathbf{m}}} w_{kj}\mathbf{m}_j$ is satisfied for all $k \in \sigma_{\mathbf{m}}$. In the same fashion, from (4.5c), there exists an equilibrium iff $\mathbf{u}_l < -\sum_{j\in\sigma_{\mathbf{m}}} w_{lj}\mathbf{m}_j - w_{li}\mathbf{m}_i$ holds for all $l \in \bar{\sigma}$. Taking the complement of this set of conditions on the input provides the conditions on equilibria for the case $|\sigma_{\mathring{\mathbf{m}}}| = 1$ and $|\sigma_{\mathbf{m}}| \geq 1$ given in the statement. □

Theorem 4.5.1 is noteworthy in that it provides quantitative conditions for the existence of oscillations in competitive LTNs. Note that these conditions depend both on the network structure and the inputs. This has physiological relevance in two ways. First, the structural conditions, in particular the pairwise instability of the nodes, show that a small portion can pull the network into a stable equilibrium rather than exhibiting oscillatory behavior. This aligns with observations [122] made of brain injuries, where a small injury can lead to significant behavioral changes. Second, the requirements on the inputs show that only the right ones lead to the emergence of oscillatory behavior for a given network structure. Given that inputs could come from other brain regions or external sources, this illustrates that the behavior of a brain network is highly dependent on its connections with other parts of the nervous system. Further, the dependence on the input opens the door to exciting design possibilities related to the robustness (or lack thereof) of oscillatory behavior: as an example, for a given input (resp. network structure), one might consider how to modify the network structure (resp. input) such that oscillatory behavior is maintained, gained, or lost.

Remark 4.5.4. (Oscillations in Competitive Networks with TLN Dynamics): The result in Theorem 4.5.1 is also applicable to threshold-linear networks by taking $\mathbf{m} = \infty\mathbf{1}$. In this case,

some of the conditions can be discarded right away since no equilibria exist with components on an upper threshold: specifically, conditions 3), along with the first component of the condition 2) can be discarded as they become trivially satisfied. This gives rise to a generalization to arbitrary networks of [74, Theorem 2.2], which only considers the all-to-all case, with no self-loops, and positive inputs. •

Remark 4.5.5. (Assumptions on Self-Excitatory Connections): We believe the assumption in Theorem 4.5.1 requiring at least two diagonal elements to be small enough is not necessary. In simulations, we have found that not enforcing this constraint still provides oscillatory behavior provided the other conditions are satisfied. •

Example 4.5.6. (Oscillations in a Seven-Node Competitive network with LTN dynamics): *We consider a competitive LTN dynamics with $n = 7$ nodes that exhibits oscillatory behavior, as per the conditions identified in Theorem 4.5.1, and illustrate the impact of the inputs on network behavior. While Theorem 4.5.1 permits arbitrary inputs, in this example we consider only homogeneous inputs of the form $\mathbf{u} = u\mathbf{1}_7$. The synaptic weight matrix $\mathbf{W}$ is as follows.*

$$\mathbf{W} = \begin{bmatrix} 0 & -0.349 & -0.055 & -0.434 & -0.745 & -0.053 & -0.381 \\ -2.907 & 0 & -0.338 & 0 & -0.376 & -0.556 & -0.558 \\ -18.07 & -2.981 & 0 & -0.764 & -0.043 & -0.823 & -0.807 \\ -0.696 & -0.03 & -0.01 & 1.435 & -0.166 & -0.331 & -0.179 \\ -1.425 & -2.664 & -23.347 & -0.20 & 0 & -0.353 & -0.958 \\ -18.83 & -1.866 & -1.255 & -0.517 & -2.887 & 0 & -0.06 \\ -2.643 & -1.84 & -1.325 & -0.138 & -1.064 & -16.64 & 0 \end{bmatrix}$$

With such inputs, and according to Theorem 4.5.1, oscillations are possible when the network is subject to inputs in the interval $u \in (1.3914, 1.9742)$. Figure 4.1 illustrates the network behavior with different inputs, three inside the range and two outside. For those inside, the same three nodes fall into limit cycles and one node saturates, but the relative values of the limit cycles vary

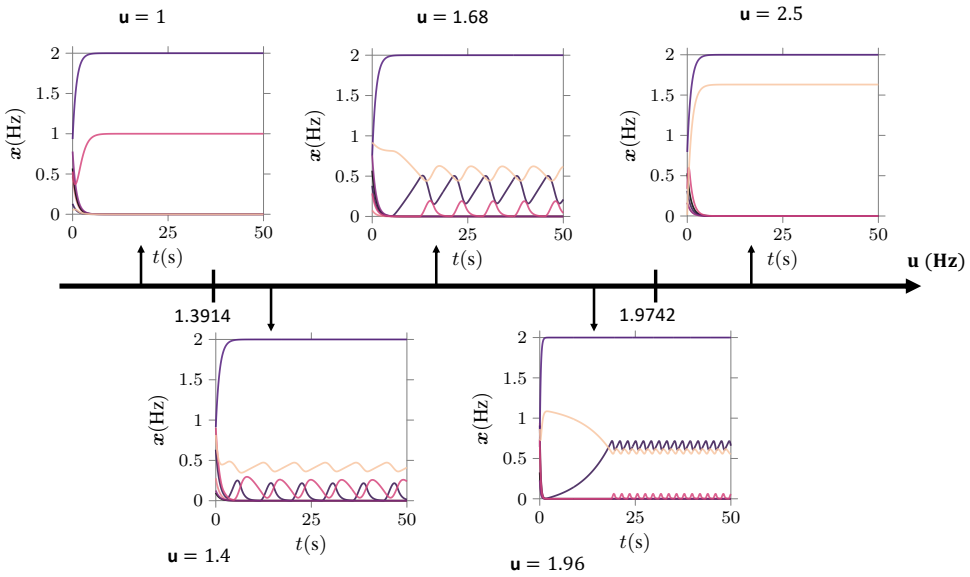


Figure 4.1. Oscillatory behavior of a 7-node competitive network with LTN dynamics subject to homogeneous inputs $\mathbf{u} = u\mathbf{1}_7$. According to Theorem 4.5.1, oscillations are possible in the input range $u \in (1.3914, 1.9742)$, while outside it, there exists a stable equilibrium. The panels vary based upon the system input and illustrate how the behavior changes. In panels with $u = 1$ and $u = 2.5$, the input lies outside of the oscillatory range and the dynamics settles to an equilibrium. In the other panels, the same three nodes exhibit oscillations while the remaining ones either settle to zero or saturate (albeit not shown here, heterogeneous inputs can make different set of nodes oscillate). The varying input values lead to settling into significantly different limit cycles.

significantly. For inputs outside the interval, the dynamics settles to a stable equilibrium. •

4.6 Conclusions

In this chapter we have studied both stable and oscillatory behaviors in the linear-threshold and threshold-linear networks. We have provided conditions characterizing the existence of a stable equilibrium supported on an arbitrary subset of nodes for linear-threshold networks. Using LoSE as a proxy for the presence of oscillations, we have characterized the emergence of oscillatory behavior in both linear-threshold and threshold-linear competitive networks, where all interneuron connections are inhibitory. Specifically, we have provided conditions on the structure of the network and the inputs such that the networks do not have stable equilibria. Future work will further explore the physiological interpretation of the conditions along with possible

additional requirements to be biologically plausible. We will also study dynamic attractors in linear-threshold competitive networks, analyze the robustness of oscillatory behavior to neuron addition and removal and its connection with neurogenesis in brain networks.

Acknowledgements

This chapter, in full, is a reprint of the material [1] where it appears as "Sufficient Conditions for Oscillations in Competitive Linear-Threshold Brain Networks" by Michael McCreesh, Tommaso Menara and Jorge Cortés in IEEE Control Systems Letters. The dissertation author was the primary investigator and author of this paper.

Chapter 5

Selective Attention in Thalamocortical Networks

The brain is a complex network composed of many subnetworks that perform a myriad of different functions. Communication of information between regions and its subsequent processing is one such function. Brain regions, such as the neocortex for example, have a hierarchical structure in which different cognitive levels operate on distinct timescales. Within this hierarchy, information travels from faster lower-level sensory brain regions to slower higher-level cognitive brain regions (*bottom-up* communication). Upon processing in the higher-level regions, information regarding decisions made by these regions is passed back down the hierarchy to perform some task (*top-down* communication). In this process, certain regions are selectively recruited to perform the given task, while other areas are selectively inhibited to ignore other inputs into the brain network.

Such hierarchies are not restricted to the neocortex, and neither top-down and bottom-up communications occur entirely inside the neocortex. In fact, most, if not all, direct corticocortical communications have a parallel transthalamic pathway upon which the information is transmitted and modulated [94]. Our goal here is to understand the role of transthalamic communication in enabling selective attention, with a view to characterizing its benefits. We seek to provide a dynamical explanation of this phenomena and validate the hypothesis that selective inhibition and recruitment are feasible in thalamocortical networks via feedback and feedforward mechanisms.

5.1 Literature Review

The hierarchical organization of the brain has been known for decades [123, 124] and has been extensively studied from different viewpoints [125, 126, 127]. The role of the communication between the thalamus and cortical regions in a thalamocortical hierarchy is a more recently studied problem. Historically, the thalamus has been viewed as a relay of sensory signals to the cortex. However, in recent literature, it has been shown to also play a further role in cognitive processes [128]. In particular, [94, 129] show that the thalamus transfers both sensory signals to the cortex using first-order relays, but that for most direct corticocortical connection, there exists a parallel transthalamic path made of higher-order thalamic relays. Further evidence is given that these paths operate using feedforward inhibitory control, mediated by interneurons, to communicate information from thalamic to cortical areas [130, 92, 95]. The works [131, 132] show that depending on the purpose (e.g., visual, auditory, somatomotor) of the hierarchical network, the thalamus connects to the hierarchy at different levels. In general, little theoretical understanding is available about the network properties of thalamocortical structures and their role in the hierarchical nature of the brain. To address this gap, we employ the linear-threshold network model (2.4) and build on the hierarchical selective recruitment framework introduced in [64, 65] for strictly cortical networks. Our results in this chapter expand the validity of this framework to a larger class of brain topologies that include the thalamus. We also rely on results from switched piecewise and affine systems [133] and singular perturbation theory [134, 135].

5.2 Contributions

We deal with thalamocortical brain networks where each brain region is modeled as a state-dependent switched system governed by a linear-threshold rate dynamics. Given our focus on selective attention, the neuronal populations in each region are divided into task-relevant and task-irrelevant nodes. Inspired by the types of pathways in thalamic circuitry, we consider two

interconnections topologies, multilayer hierarchical networks and star-connected networks. Our first contribution provides an analysis from a control perspective of the mechanisms involved in the selective inhibition and recruitment in hierarchical thalamocortical network topologies. Using singular perturbation theory for non-smooth differential equations, and the piecewise-affine nature of the linear-threshold dynamics, we provide conditions for the existence of feedback-feedforward control laws that achieve selective inhibition and recruitment of the network to a desired equilibrium trajectory. Our second contribution expands our results to star-connected thalamocortical networks, both with and without a temporal hierarchy between regions. For the latter class, we build on a generalization of stability results on slowly varying nonlinear systems to the case of exponential stability to provide conditions for the existence of a feedback-feedforward controller providing selective inhibition and recruitment. We achieve analogous results for the case of star-connected networks with a temporal hierarchy again using singular perturbation theory. Examples illustrate the beneficial role played by the thalamus interconnections in these networks for metrics such as failsafe mechanisms, control energy, and network performance. Taken together, our results provide a dynamical explanation from a systems and control perspective of the mechanisms involved in the emergence of selective attention in the brain and the role of the thalamus. Given the state-dependent switched nature of the dynamics, the results on stabilizability and distributed feedforward/feedback design are also of independent interest for the control of complex network systems.

5.3 Neuroscientific Background

Here we provide a summary of neuroscientific background beyond that found in Chapter 2 behind the modeling assumptions adopted in this chapter. The material provided here is focused on information related to our upcoming discussion of goal-driven selective attention in thalamocortical networks. It includes observations about brain organization, information transmission among brain regions, and the role that the thalamus is believed to play.

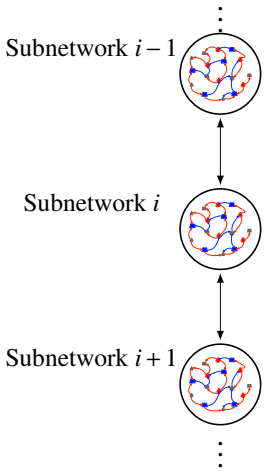
Task-Relevant and Task-Irrelevant Neuron Populations: In the nervous system, stimuli are represented using series of electric spikes generated by neurons that travel down nerve fibres [68]. A given stimulus is defined by a characteristic pattern of spikes traveling between neurons, which can also be represented as the firing rate of the neurons over time, as developed in Chapter 2. In such a representation, some neurons generate spikes (have a non-zero firing rate) during the transmission of the stimuli, known as being excited, while other neurons do not (have a firing rate of zero), which is referred to as being inhibited. We refer to the subset of neuron populations that are excited during the transmission of a stimuli as the 'task-relevant' nodes and the remaining populations as the 'task-irrelevant' nodes.

Information Pathways: In the brain, there exist information pathways between different spatial regions allowing for the transmission of stimuli between processing areas. The transmission of information can be seen as the activity in one region (that is, the firing rates of the neuron populations defined by the stimuli) driving the activity in the following region by exciting the task-relevant nodes and inhibiting the task-irrelevant nodes to propagate the stimuli (and any processing of it) through the pathway. This enables the brain to generate appropriate responses to the stimuli by propagating a response through the information pathway. This response, with its own set of task relevant/irrelevant nodes, selectively *recruits* (excites) the task-relevant nodes, and selectively *inhibits* the task-irrelevant nodes. Information pathways between brain regions form both spatial and temporal hierarchies, allowing for different levels of processing occurring in different regions [90]. The temporal timescale separation is directly related to the complexity of processing occurring in a given region. For a low-level sensory area, where the inputs are brief and are processed quickly, the timescales are fast. Meanwhile, further up the information pathway, in regions such as the prefrontal cortex, higher-level cognitive processes use more complex inputs from earlier in the pathway by integrating them over time, resulting in slower timescales [136]. In strictly cortical networks, cf. Fig. 5.1(a), selective inhibition and recruitment through these pathways has been studied [64, 65]. There, it is shown that achieving selective inhibition and recruitment in such a network is highly dependent on both the individual regions

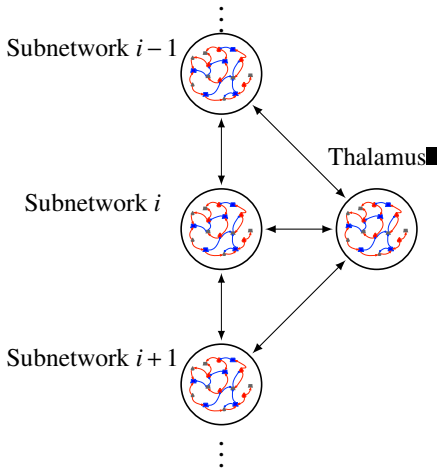
dynamics and its interconnections with neighboring regions. Our motivation here stems from the fact that elements of such networks also have multiple connections to the thalamus while maintaining a temporal hierarchy [94], and there does not exist an analysis from a systems and control viewpoint of selective inhibition and recruitment in network topologies with these additional information pathways.

Pathways between the Thalamus and Cortical Regions: In studying thalamocortical networks, we consider two topologies of interest, cf. Fig. 5.1(b),(c). Both share the common trait that each cortical layer is connected to the thalamus layer: however, connections between cortical layers differ in each case. These topologies are inspired by the fact that pathways in thalamic circuitry can be identified into two classes [137]. The first represents the role of the thalamus as a modulator of information being passed between cortical regions along a transthalamic route parallel to the existing corticocortical information pathways (higher-order relay) [94]. Since these cortical regions then form a temporal hierarchy, this leads to the hierarchical thalamocortical network shown in Fig. 5.1(a). The second class represents the thalamus as the main route for the transfer of information between two (or more) brain areas. In this case, the thalamus is relaying an input to the cortical regions (first-order relay) [94], which gives rise to the star-connected network shown in Fig. 5.1(b). The cortical regions to which information is being relayed can be parts of separate temporal hierarchies, however, as the regions in the network do not form a temporal hierarchy themselves, the timescales of the subnetworks are not directly related.

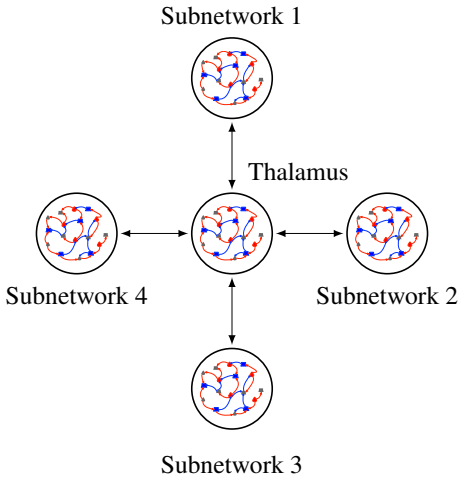
Role of the Thalamus: The thalamus is a component of many different brain networks and thus plays a role in a large number of functions. Traditionally, the thalamus has been considered primarily as being a sensory relay to the cortex, playing minimal other functional roles [138]. Despite being known since the late 19^{th} century to have additional functions such as a role in memory loss [139], up until recently the majority of the research into thalamic function has studied its function as a sensory relay [92, 140, 141]. This view was gradually changed by several works, including the pioneering work [142], establishing the thalamus as a heterogeneous structure only a small portion of whose nuclei play the role of a sensory relay.



(a) hierarchical cortical network



(b) hierarchical thalamocortical network



(c) star-connected thalamocortical network

Figure 5.1. Topologies for cortical and thalamocortical brain networks considered in this chapter. In (a), a strictly cortical network, each layer is connected only to the layers immediately above and below it. In (b), each layer is connected directly to the thalamus, as well as the layers directly above and below it. In (c), each layer is connected to the thalamus, but no direct connections exist between cortical regions. In both plots, task-relevant excitatory and inhibitory nodes are depicted in red and blue, resp., and task-irrelevant nodes are depicted in grey.

Through both the first-order and higher-order relays the thalamus can significantly increase the information contained in signals both being transmitted to and between cortical regions [94]. In hierarchical networks within the cortex, transthalamic pathways allow for layers near the top of the hierarchy to directly receive the outputs from the lower levels, in addition to the more processed inputs they receive from other cortical regions [92]. Thalamic signals into the higher cortical regions can also allow for receiving further details of motor signals, such as distinguishing between self-induced stimuli (such as those generated by eye movements) and stimuli from the external environment [94]. The thalamus also has a role in controlling recurrent cortical dynamics, as the cortical networks in Fig. 5.1a) are not able to self-sustain activity [130].

One mechanism for the interaction of thalamus with cortical regions is through feedforward inhibitory control [132, 95, 96]. As such, when studying thalamocortical networks, models need to take into account that many of the net impacts from the thalamus onto cortical regions are inhibitory, while returning connections are both excitatory and inhibitory.

5.4 Problem Setup

We start by providing details on the dynamic modeling of the thalamocortical network layers, then describe the effect that the interconnection topology has on the input to each layer, and finally formalize the problem under consideration.

5.4.1 Network Modeling

We consider a thalamocortical network $\mathcal{N}$ composed of N cortical layers and the thalamus. We use linear-threshold rate dynamics to model the evolution of each region in the network. These dynamics provide a mesoscale model of the evolution of the average firing rate of populations of neurons, rather than individual spike trains, by looking at the electrical currents flowing through synaptic connections, see [68]. The dynamics of cortical layer $\mathcal{N}_i$ composed of n_i nodes are

$$\tau_i \dot{\mathbf{x}}_i = -\mathbf{x}_i + [\mathbf{W}_i \mathbf{x}_i + \mathbf{d}_i(t)]_{\mathbf{0}}^{\mathbf{m}_i} \qquad \mathbf{0} \leq \mathbf{x}_i(0) \leq \mathbf{m}_i, \tag{5.1}$$

where $\mathbf{x}_i \in \mathbb{R}^{n_i}_{\geq 0}$ represents the state of the nodes within the layer, and each component of $\mathbf{x}_i$ represents a population of neurons with similar firing rate. The matrix $\mathbf{W}_i \in \mathbb{R}^{n_i \times n_i}$ is the synaptic connectivity between neuron populations within the layer and $\mathbf{d}_i(t) \in \mathbb{R}^{n_i}$ encapsulates the input into the model,

$$\mathbf{d}_i(t) = \mathbf{w}_i(t) + \mathbf{B}_i \mathbf{u}_i(t) + \mathbf{c}_i. \tag{5.2}$$

Here $\mathbf{w}_i(t)$ models the interconnections between layers, $\mathbf{u}_i(t)$ is the control used for the n_i nodes in $\mathcal{N}_i$, and $\mathbf{c}_i$ includes unmodeled background activity or external inputs. The firing rate threshold is $\mathbf{m}_i \in \mathbb{R}^{n_i}_{>0}$ and $\tau_i \in \mathbb{R}^{n_i}_{\geq 0}$ is the timescale.

In our study, the control $\mathbf{u}_i$ inhibits $r_i \leq n_i$ task-irrelevant nodes (drives their state to zero), while the remaining terms in (5.2) recruit the remaining $n_i - r_i$ nodes and determine the desired equilibrium or trajectory. To distinguish between task-relevant and task-irrelevant nodes, we use the following partition of network variables in each layer $\mathcal{N}_i$,

$$\mathbf{x}_i = \begin{bmatrix} \mathbf{x}_i^0 \\ \mathbf{x}_i^1 \end{bmatrix} \qquad \mathbf{W}_i = \begin{bmatrix} \mathbf{W}_i^{00} & \mathbf{W}_i^{01} \\ \mathbf{W}_i^{10} & \mathbf{W}_i^{11} \end{bmatrix} \tag{5.3a}$$

$$\mathbf{B}_i = \begin{bmatrix} \mathbf{B}_i^0 \\ \mathbf{0} \end{bmatrix} \qquad \mathbf{c}_i = \begin{bmatrix} \mathbf{c}_i^0 \\ \mathbf{c}_i^1 \end{bmatrix} \qquad \mathbf{m}_i = \begin{bmatrix} \mathbf{m}_i^0 \\ \mathbf{m}_i^1 \end{bmatrix}, \tag{5.3b}$$

where $\mathbf{x}_i^0 \in \mathbb{R}^{r_i}_{\geq 0}$ represent the task-irrelevant nodes, $\mathbf{x}_i^1 \in \mathbb{R}^{n_i - r_i}_{\geq 0}$ the task-relevant nodes and $\mathbf{B}_i \in \mathbb{R}^{n_i \times p_i}$ is such that the task-relevant nodes are not impacted by the control term $\mathbf{u}_i \in \mathbb{R}^{p_i}$. Throughout we shall assume that $p_i \geq r_i$, and that the matrices $\mathbf{B}_i^0$ have all full rank.

The thalamus layer $\mathcal{N}_T$ is modeled in the same way as the cortical layers with the difference appearing in the interconnection term $\mathbf{w}_T(t)$.

A final word about the brain mechanisms for the inhibition of regions [101]. Feedforward inhibition between two layers, $\mathcal{N}_i$ and $\mathcal{N}_j$, refers to when $\mathcal{N}_i$ sends an inhibitory signal to $\mathcal{N}_j$ to

achieve its goal activity pattern for $\mathcal{N}_j$ regardless of the current state of $\mathcal{N}_j$. In contrast, feedback inhibition refers to when the inhibition applied in a layer $\mathcal{N}_i$ is dependent upon the current activity level of the nodes desired to be inhibited [101]. We employ a combination of feedback and feedforward inhibition.

5.4.2 Interconnection Topology Among Network Layers

We detail here the dynamical interconnection between the network layers for the topologies in Fig. 5.1(b),(c).

Hierarchical Thalamocortical Networks

Consider the hierarchical thalamocortical network depicted in Fig. 5.1(a). The hierarchical structure is encoded by the ordered timescales of the layers: $\tau_1 \gg \tau_2 \gg \cdots \gg \tau_N$, prescribing progressively faster dynamics as one moves down the hierarchy. For generality, the timescale τ_T of the thalamus might fit anywhere within the hierarchy. For each cortical layer $\mathcal{N}_i$, the interconnection term $\mathbf{w}_i(t)$ in (5.2) takes the form

$$\mathbf{w}_i(t) = \mathbf{W}_{i,i-1}\mathbf{x}_{i-1}(t) + \mathbf{W}_{i,i+1}\mathbf{x}_{i+1}(t) + \mathbf{W}_{i,T}\mathbf{x}_T(t). \tag{5.4}$$

Here, the terms $\mathbf{W}_{i,i-1}, \mathbf{W}_{i,i+1}$, and $\mathbf{W}_{i,T}$ represent the weights of the synaptic connections between layers $\mathcal{N}_i$ and $\mathcal{N}_{i-1}$, $\mathcal{N}_{i+1}$ and $\mathcal{N}_T$, resp. It is important to note that since the thalamus impacts the cortical regions using feedforward inhibition, the interconnection matrix between the thalamus and the cortical layer satisfies $\mathbf{W}_{i,T} \leq 0$. Substituting (5.4) into the linear-threshold dynamics (5.1), we get the dynamics for a cortical layer $\mathcal{N}_i$

$$\tau_i \dot{\mathbf{x}}_i = -\mathbf{x}_i + [\mathbf{W}_{i,i}\mathbf{x}_i + \mathbf{W}_{i,i-1}\mathbf{x}_{i-1} + \mathbf{W}_{i,i+1}\mathbf{x}_{i+1} + \mathbf{W}_{i,T}\mathbf{x}_T + \mathbf{B}_i\mathbf{u}_i(t) + \mathbf{c}_i]_{\mathbf{0}}^{\mathbf{m}_i}, \tag{5.5}$$

for $i \in \{1, \ldots, N\}$. For consistency $\mathbf{W}_{1,0} = \mathbf{0} = \mathbf{W}_{N,N+1}$, and we assume that $r_1 = 0$, meaning no nodes are being inhibited in the top layer of the network.

For the thalamus layer, to reflect the different connectivity it has in the network, the interconnection term $\mathbf{w}_T(t)$ is

$$\mathbf{w}_T(t) = \sum_{i=1}^{N} \mathbf{W}_{T,i}\mathbf{x}_i(t). \tag{5.6}$$

Here $\mathbf{W}_{T,i}$ represents the weight of the synaptic connections between layers $\mathcal{N}_T$ and $\mathcal{N}_i$ for $i \in \{1,\dots,N\}$. Then, substituting (5.6) into (5.1), the dynamics for the thalamus layer is

$$\tau_T\dot{\mathbf{x}}_T = -\mathbf{x}_T + [\mathbf{W}_T\mathbf{x}_T + \sum_{i=1}^{N} \mathbf{W}_{T,i}\mathbf{x}_i + \mathbf{B}_T\mathbf{u}_T(t) + \mathbf{c}_T]_{\mathbf{0}}^{\mathbf{m}_T}. \tag{5.7}$$

We denote the timescale ratio between layers by $\epsilon = (\epsilon_1,\dots,\epsilon_N,\epsilon_T)$, where $\epsilon_i = \tau_i/\tau_{i-1}$ and $\epsilon_T = \tau_T/\min_{\tau_j>\tau_T}\tau_j$. For a subnetwork $\mathcal{N}_i$ such that the thalamus timescale fits in the hierarchy directly above it, i.e. $\tau_T > \tau_i$ but there does not exist j such that $\tau_T > \tau_j > \tau_i$, we let the timescale ratio between $\mathcal{N}_T$ and $\mathcal{N}_i$ be given by $\bar{\epsilon}_i = \tau_i/\tau_T$.

Star-Connected Thalamocortical Networks

Consider the star-connected thalamocortical network depicted in Fig. 5.1(b). In contrast to the hierarchical network, this topology does not form a hierarchical timescale, and as such there is no direct relationship satisfied by the timescales. The lack of an explicit relation encodes the thalamus' role as a sensory relay to multiple brain regions, each part of potentially unrelated temporal hierarchies. Without loss of generality, we assume subnetwork $\mathcal{N}_1$ represents a subcortical structure and provides the input to the network for the thalamus to relay to the other brain regions. As such, there are no nodes in $\mathcal{N}_1$ that are desired to be inhibited, meaning $r_1 = 0$. We model the subcortical input subnetwork with a linear-threshold dynamics, but without an independent control term, instead modeling input changes by allowing $\mathbf{c}_1$ to be time-varying.

That is

$$\tau_1 \dot{\mathbf{x}}_1 = -\mathbf{x}_1 + [\mathbf{W}_{1,1}\mathbf{x}_1 + \mathbf{W}_{1,T}\mathbf{x}_T + \mathbf{c}_1(t)]_{\mathbf{0}}^{\mathbf{m}_1}. \tag{5.8}$$

For the cortical regions $\mathcal{N}_i$, $i \in \{2, \ldots, N\}$, the interconnection term $\mathbf{w}_i(t)$ is given by $\mathbf{w}_i(t) = \mathbf{W}_{i,T}\mathbf{x}_T(t)$, and we recall that as the thalamus utilizes feedforward inhibition, $\mathbf{W}_{i,T} \leq 0$. Meanwhile, the interconnection of the thalamus with the cortical regions is defined by $\mathbf{w}_T(t) = \sum_{i=1}^{N} \mathbf{W}_{T,i}\mathbf{x}_i(t)$. Then, from (5.1), for $i \in \{2, \ldots, N\}$ the dynamics takes the form

$$\begin{aligned} \tau_i \dot{\mathbf{x}}_i &= -\mathbf{x}_i + [\mathbf{W}_{i,i}\mathbf{x}_i + \mathbf{W}_{i,T}\mathbf{x}_T + \mathbf{B}_i\mathbf{u}_i(t) + \mathbf{c}_i]_{\mathbf{0}}^{\mathbf{m}_i}, \\ \tau_T \dot{\mathbf{x}}_T &= -\mathbf{x}_T + [\mathbf{W}_T\mathbf{x}_T + \sum_{i=1}^{N} \mathbf{W}_{T,i}\mathbf{x}_i + \mathbf{B}_T\mathbf{u}_T(t) + \mathbf{c}_T]_{\mathbf{0}}^{\mathbf{m}_T}. \end{aligned} \tag{5.9}$$

5.4.3 Problem Statement

For a purely cortical hierarchical brain network, as in Fig. 5.1a), selective inhibition and recruitment can be achieved [64, 65] using a combination of feedback and feedforward control, dependent on the subnetwork dynamics satisfying a set of stability properties. However, with the thalamus mediating in the network topology and its impact on the dynamics of the individual subnetworks, such results are inapplicable for the hierarchical thalamocortical network topology considered here. In addition, for the star-connected thalamocortical topology, the results do not apply, as the assumption of a hierarchical relationship between subnetwork timescales no longer holds. These observations motivate our study of control mechanisms to achieve selective inhibition and recruitment for both hierarchical and star-connected thalamocortical networks, formalized next.

Problem 5.4.1. *Consider a thalamocortical network with the interconnection topologies described in Section 5.4.2. For each topology, determine conditions on the individual subnetworks $\mathcal{N}_i$, $\mathcal{N}_T$, $i \in \{1, \ldots, N\}$ and the connections between them such that a feedback-feedforward control*

$\mathbf{u}(t) = \mathbf{K}\mathbf{x}(t) + \bar{\mathbf{u}}(t)$ *exists making the network converge to an equilibrium trajectory* $\mathbf{x}^*(t)$ *of the form*

$$\mathbf{x}^*(t) = \begin{bmatrix} \mathbf{x}^{0^*}(t) \\ \mathbf{x}^{1^*}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{x}^{1^*}(t) \end{bmatrix}. \tag{5.10}$$

Further, we study differences between the hierarchical and star-connected topologies, particularly in how the latter can be considered as a failsafe topology for the former. In addition, we seek to evaluate how the addition of the thalamus subnetwork to strictly cortical networks results in improved performance. In particular we investigate how the thalamus layer impacts the control energy needed to achieve selective inhibition and the convergence speed of the network.

5.5 Hierarchical Thalamocortical Networks

In this section, we consider a hierarchical multilayer thalamocortical network, cf. Fig. 5.1(a), where the cortical layers are governed by (5.5) and the thalamus layer is governed by (5.7).

5.5.1 Equilibrium Maps for Individual Layers

We note that, for a general linear-threshold dynamics (5.1) with $\mathbf{W} \in \mathbb{R}^{n\times n}$, its equilibrium map $h : \mathbb{R}^n \to \mathbb{R}^n_{\geq 0}$

$$h(\mathbf{c}) = h_{\mathbf{W},\mathbf{m}}(\mathbf{c}) = \{\mathbf{x} \in \mathbb{R}^n_{\geq 0} \mid \mathbf{x} = [\mathbf{W}\mathbf{x} + \mathbf{c}]^{\mathbf{m}}_{\mathbf{0}}\}, \tag{5.11}$$

maps a constant input $\mathbf{c} \in \mathbb{R}^n$ to the set of equilibria of (5.1). For thalamocortical networks, the timescale of the thalamus layer impacts the specific form of the equilibrium maps. For simplicity, we assume the thalamus lies inside the hierarchy, i.e., there exist $a, b \in \{1, \ldots, N\}$ such that $b = a + 1$ and $\tau_a \gg \tau_T \gg \tau_b$. This choice results in a layer $\mathcal{N}_a$ with an equilibrium map different than any that appear in the cases when the thalamus timescale is on the boundary of the

hierarchy. Our results can also be stated for the latter case with appropriate adjustments to the derived control law for selective inhibition and recruitment.

The hierarchical nature of the topology plays an important role in defining these equilibrium maps. At the theoretical limit of timescale separation between two layers, the state of a given layer becomes constant at the timescale of layers lower in the hierarchy as well as being a static function of the above layer's states. As such, when defining the equilibrium maps for a given layer, the constant input into an equilibrium map can represent the state of layers higher in the hierarchy, given that they are relative constants at that level of the dynamics. The equilibrium maps for the layers fall into three categories: below or above the thalamus, and the thalamus itself. In all cases, the maps are defined recursively from the bottom to the top of the network. At each layer, the equilibrium map takes a constant input, representing the inputs from higher levels in the hierarchy along with any external inputs to the system, and outputs the set of equilibrium values.

We next give explicit expressions for the equilibrium maps of the task-relevant component of a layer in each of the categories, denoted $h_i^1 : \mathbb{R}^{n_i - r_i} \to \mathbb{R}_{\geq 0}^{n_i - r_i}$, by combining the hierarchical model described in Section 5.4.2 with (5.11).

Equilibrium Maps for Layers below Thalamus

We begin by considering layers below the thalamus, i.e., with $i \geq b$. Given values of $\mathbf{x}_T \in \mathbb{R}^{n_T}$ and $\mathbf{c}_{i+1}^1 \in \mathbb{R}^{n_{i+1} - r_{i+1}}$,

$$h_i^1(\mathbf{c}) = \{\mathbf{x}_i^1 \mid \mathbf{x}_i^1 = [\mathbf{W}_{i,i}^{11}\mathbf{x}_i^1 + \mathbf{W}_{i,i+1}^{11} h_{i+1}^1(\mathbf{W}_{i+1,i}^{11}\mathbf{x}_i^1 + \mathbf{W}_{i+1,T}^{11}\mathbf{x}_T^1 + \mathbf{c}_{i+1}^1) + \mathbf{c}]_{\mathbf{0}}^{\mathbf{m}_i^1}\}. \tag{5.12}$$

We note that since $\mathbf{W}_{N,N+1} = \mathbf{0}$ by convention, this recursion is well-defined for $\mathcal{N}_N$, the bottom layer in the network. In particular, for layer N, (5.12) reduces to (5.11), the standard equilibrium map for linear-threshold models.

Equilibrium Map for Thalamus

Since the thalamus is connected to all the cortical layers, the recursive definition of its equilibrium map is dependent on the equilibrium maps of all of the layers below it. Using the notation of (5.12) for writing the recursion then becomes intractable for large networks. Instead, we use the following notation. For layer $\mathcal{N}_i$, $i \in \{1,\dots,N\}$, in the thalamocortical hierarchy, let $\mathbf{x}^1_{i_e} \in \mathbb{R}^{n_i - r_i}_{\geq 0}$ be such that

$$\mathbf{x}^1_{i_e} \in h^1_i(\mathbf{W}^{11}_{i,T}\mathbf{x}^1_T + \mathbf{W}^{11}_{i,i-1}\mathbf{x}^1_{i-1} + \mathbf{c}^1_i).$$

Despite $\mathbf{x}_{i_e}$ being dependent on a set of inputs, we do not specify them: when this notation is used, it is implicitly assumed that the input values are given or determined in the recursion. Depending on the point in the recursion being considered, the inputs $\mathbf{x}_T$ and/or $\mathbf{x}_{i-1}$ will be replaced by their equilibrium values. Now, given values $\mathbf{c}^1_b,\dots,\mathbf{c}^1_N$, where $\mathbf{c}^1_i \in \mathbb{R}^{n_i - r_i}$, the thalamus equilibrium map $h^1_T : \mathbb{R}^{n_T - r_T} \to \mathbb{R}^{n_T - r_T}_{\geq 0}$ is

$$h^1_T(\mathbf{c}) = \{\mathbf{x}^1_T \mid \mathbf{x}^1_T = [\mathbf{W}^{11}_T\mathbf{x}^1_T + \sum_{j=b}^{N} \mathbf{W}^{11}_{T,j}\mathbf{x}^1_{j_e} + \mathbf{c}]^{\mathbf{m}^1_T}_{\mathbf{0}}\}. \tag{5.13}$$

Equilibrium Maps for Layers above Thalamus

We note that, in the recursive definition for the equilibrium maps above the thalamus, the inputs to the thalamus equilibrium map (5.13) will include equilibrium values of layers above the thalamus in the hierarchy, in addition to the equilibrium values from lower in the hierarchy. To distinguish which maps are inputting equilibrium maps in a condensed manner, we introduce the following notation. For $i \in \{1,\dots,a\}$, we let $\mathbf{x}^1_{T_{e(i)}} \in \mathbb{R}^{n_T - r_T}$ denote a value inside the thalamus equilibrium set satisfying

$$\mathbf{x}^1_{T_{e(i)}} \in h^1_T(\sum_{j=1}^{i} \mathbf{W}^{11}_{T,j}\mathbf{x}^1_j + \sum_{j=i+1}^{a} \mathbf{W}^{11}_{T,j}\mathbf{x}^1_{j_e} + \mathbf{c}^1_T).$$

Using this notation, we define the remaining equilibrium maps. We first provide the equilibrium maps for layers $\{\mathcal{N}_i\}_{i=1}^{a-1}$ and finish with the map for $\mathcal{N}_a$. For $i \in \{1,\dots,a-1\}$, given $\mathbf{c}_{i+1}^1 \in \mathbb{R}^{n_i-r_i}$, the equilibrium map $h_i^1 : \mathbb{R}^{n_i-r_i} \to \mathbb{R}_{\geq 0}^{n_i-r_i}$ is

$$h_i^1(\mathbf{c}) = \{\mathbf{x}_i^1 \mid \mathbf{x}_i^1 = [\mathbf{W}_{i,i}^{1,\text{all}}\mathbf{x}_i + \mathbf{W}_{i,i+1}h_{i+1}^1(\mathbf{W}_{i+1,i}^{1,\text{all}}\mathbf{x}_i + \mathbf{c}_{i+1}^1) + \mathbf{W}_{i,T}^{11}\mathbf{x}_{T_{e(i)}} + \mathbf{c}]_{\mathbf{0}}^{\mathbf{m}_i^1}\}. \tag{5.14}$$

The expression for the equilibrium map of the layer $\mathcal{N}_a$, which is directly above the thalamus, differs from the other layers due to the fact that it depends directly on a layer below the thalamus in addition to being dependent on the thalamus. Given $\mathbf{c}_b^1 \in \mathbb{R}^{n_b-r_b}$, the equilibrium map $h_a^1 : \mathbb{R}^{n_a-r_a} \to \mathbb{R}_{\geq 0}^{n_a-r_a}$ is

$$h_a^1(\mathbf{c}) = \{\mathbf{x}_a^1 \mid \mathbf{x}_a^1 = [\mathbf{W}_{a,a}^{11}\mathbf{x}_a^1 + \mathbf{W}_{a,b}^{11}h_b^1(\mathbf{W}_{b,a}^{11}\mathbf{x}_a^1 + \mathbf{W}_{b,T}^{11}\mathbf{x}_{T_{e(a)}}^1) + \mathbf{c}_b^1) + \mathbf{W}_{a,T}^{11}\mathbf{x}_{T_{e(a)}}^1 + \mathbf{c}]_{\mathbf{0}}^{\mathbf{m}_a^1}\}. \tag{5.15}$$

We conclude this section by establishing that all the equilibrium maps in the thalamo-cortical network (5.12)-(5.15) are piecewise-affine and use this fact to justify they are globally Lipschitz too. To begin, we note that since general linear-threshold dynamics are switched affine, their equilibrium map (5.11) can be written in a piecewise-affine form. In particular, this equilibrium map can be written as follows

$$h(\mathbf{c}) = \{\mathbf{F}_\sigma\mathbf{c} + \mathbf{f}_\sigma \mid \mathbf{G}_\sigma\mathbf{c} + \mathbf{g}_\sigma \geq \mathbf{0},\ \sigma \in \{0,\ell,s\}^n\}, \tag{5.16}$$

for some matrices and vectors of the form

$$\begin{aligned}
\mathbf{F}_\sigma &= (\mathbf{I} - \Sigma^\ell\mathbf{W})^{-1}, \qquad \mathbf{f}_\sigma = (\mathbf{I} - \Sigma^\ell\mathbf{W})^{-1}\Sigma^{\mathbf{s}}\mathbf{m}, \\
\mathbf{G}_\sigma &= \begin{bmatrix} \Sigma^\ell + \Sigma^s - \mathbf{I} & \Sigma^\ell & -\Sigma^\ell & \Sigma^s \end{bmatrix}^\top \mathbf{F}_\sigma, \\
\mathbf{g}_\sigma &= \begin{bmatrix} \mathbf{f}_\sigma^\top(\Sigma^\ell + \Sigma^s - \mathbf{I}) & \mathbf{f}_\sigma^\top\Sigma^\ell & (\mathbf{m} - \mathbf{f}_\sigma)^\top\Sigma^\ell & (\mathbf{f}_\sigma - \mathbf{m})^\top\Sigma^s \end{bmatrix}^\top,
\end{aligned}$$

where Σ^ℓ is a diagonal matrix with $\Sigma_{ii}^\ell = 1$ if $\sigma_i = \ell$ and zero otherwise, and Σ^s is defined

analogously. Now, since $\mathbf{W}_{N,N+1} = \mathbf{0}$, the equilibrium map for the bottom layer in the hierarchy, h_N^1, has the same form as (5.11), and hence can be written in the form (5.16). The following result generalizes [65, Lemma IV.1] and shows that all the equilibrium maps in the hierarchical thalamocortical network are piecewise-affine.

Lemma 5.5.1. (Piecewise-affinity of equilibrium maps in hierarchical thalamocortical linear-threshold models): *Let $h_i : \mathbb{R}^n \to \mathbb{R}^n$, $i \in \{1, \dots, k\}$, be piecewise-affine functions,*

$$h_i(\mathbf{c}) = \mathbf{F}_{\lambda_i}^i \mathbf{c} + \mathbf{f}_{\lambda_i}^i, \qquad \forall \mathbf{c} \in \Psi_{\lambda_i} \triangleq \{\mathbf{c} \mid \mathbf{G}_{\lambda_i}^i \mathbf{c} + \mathbf{g}_{\lambda_i}^i \geq \mathbf{0}\},$$

for $\lambda_i \in \Lambda_i$, where Λ_i is a finite index set such that $\bigcup_{\lambda_i \in \Lambda_i} \Psi_{\lambda_i} = \mathbb{R}^n$. Define $\Lambda = \Lambda_1 \times \Lambda_2 \times \cdots \times \Lambda_k$, $\lambda = (\lambda_1, \dots, \lambda_k)$ and $\Psi_\lambda = (\Psi_{\lambda_1}, \dots, \Psi_{\lambda_k})$. Given matrices $\mathbf{W}_1, \mathbf{W}_2^i, \mathbf{W}_3^i$ and vectors $\bar{\mathbf{c}}_i, \mathbf{c}'$ for all $i \in \{1, \dots, k\}$ assume

$$\mathbf{x} = \left[\mathbf{W}_1 \mathbf{x} + \sum_{i=1}^{k} \mathbf{W}_2^i h_i(\mathbf{W}_3^i \mathbf{x} + \bar{\mathbf{c}}_i) + \mathbf{c}' \right]_{\mathbf{0}}^{\mathbf{m}}, \tag{5.17}$$

is known to have a unique solution $\mathbf{x}' \in \mathbb{R}^{n'}$ for each $\mathbf{c}' \in \mathbb{R}^{n'}$, and let $h'(\mathbf{c}')$ be this solution. Then, there exists a finite index set Λ' and $\{(\mathbf{F}'_\lambda, \mathbf{f}'_\lambda, \mathbf{G}'_\lambda, \mathbf{g}'_\lambda)\}_{\lambda' \in \Lambda'}$ such that

$$h'(\mathbf{c}') = \mathbf{F}'_{\lambda'} \mathbf{c}' + \mathbf{f}'_{\lambda'}, \quad \forall \mathbf{c}' \in \Psi'_{\lambda'} \triangleq \{\mathbf{c}' \mid \mathbf{G}'_{\lambda'} \mathbf{c}' + \mathbf{g}'_{\lambda'} \geq \mathbf{0}\}$$

for $\lambda' \in \Lambda'$ and $\bigcup_{\lambda' \in \Lambda'} \Psi_{\lambda'} = \mathbb{R}^{n'}$.

The equilibrium maps for the hierarchical thalamocortical network satisfy Lemma 5.5.1 with $r = 1$, $r = N - b$ and $r = 2$ for the layers below the thalamus, the thalamus, and the layers above the thalamus, resp., and therefore the maps are piecewise-affine. Finally, the fact that these maps are globally Lipschitz follows from [65, Lemma IV.2]. This property is necessary to be able to apply later the generalization of Tikhonov's singular perturbation stability theorem to

non-smooth ODEs given in [135, Proposition 1][1] to take advantage of the timescale separation between the layers.

5.5.2 Stability Assumptions and Conditions

In the hierarchical thalamocortical model described in Section 5.4, only the task-irrelevant components of the dynamics are directly controlled over, cf. (5.3b). This means that assumptions on the stability of the task-relevant components of the dynamics are needed to guarantee their stability and recruitment to an equilibrium trajectory. In particular, we are interested in the reduced-order task-relevant dynamics, that is, the system dynamics in which the inputs from layers lower in the hierarchy have been replaced by their equilibrium values. Here, we provide details on these assumptions and identify sufficient conditions for them to hold.

Top Layer of the Hierarchy

The top layer $\mathcal{N}_1$ in the hierarchy does not have any nodes that are to be inhibited. In addition, its role is in driving the selective recruitment in the lower levels, rather than being recruited itself. As such, we only require that the trajectories of its dynamics are bounded. Formally, for all sets of constants $\mathbf{c}_1 \in \mathbb{R}^{n_1}$, $\mathbf{c}_i^1 \in \mathbb{R}^{n_i - r_i}$, $i \in \{2, \dots, N\}$, and $\mathbf{c}_T \in \mathbb{R}^{n_T - r_T}$, we assume

$$\tau_1 \dot{\bar{\mathbf{x}}}_1 = -\mathbf{x}_1 + [\mathbf{W}_{1,1}\bar{\mathbf{x}}_1 + \mathbf{W}_{1,2}^{11} h_2^1(\mathbf{W}_{2,1}\bar{\mathbf{x}}_1 + \mathbf{c}_2^1) + \mathbf{W}_{1,T}^{11}\mathbf{x}_{T_{e(1)}} + \mathbf{c}_1]_{\mathbf{0}}^{\mathbf{m}_1}, \tag{5.18}$$

has bounded solutions. We note that our earlier assumption that the thalamus is in the middle of the hierarchy makes the top layer a cortical one. If, instead, the thalamus was the top layer, one would instead assume here that its dynamics has bounded solutions, replacing (5.18) accordingly.

[1] To apply the result to a non-smooth ODE such as (5.1) we need to justify the following: 1) Lipschitzness of the dynamics uniformly in t, 2) Existence, uniqueness and Lipschitzness of the equilibrium map of the fast dynamics, 3) Lipschitzness and boundedness of the reduced-order model, 4) Asymptotic stability of the fast dynamics uniformly in t and the slow variable, and 5) Global attractivity of the fast dynamics for any fixed slow variable.

Lower Layers in the Hierarchy

In each of the layers below the top one, we seek to accomplish selective inhibition and recruitment by having the dynamics converge to a parameter-dependent equilibrium trajectory. For the task-irrelevant components, we aim to use a control law to stabilize them to zero. In such a case, an additional challenge is then to identify conditions on the interconnected network layers that ensure the coupled system displays the desired behavior. Due to the different reduced-order dynamics throughout the network, we provide the details for this assumption in four categories: layers above the thalamus, layer directly above the thalamus, thalamus, and layers below the thalamus.

Layers above the Thalamus

For each layer $\mathcal{N}_i$, with $i \in \{1, \ldots, a-1\}$, above the thalamus, the dynamics are directly dependent on two layers below it, and hence the reduced-order dynamics are dependent on two equilibrium maps. For achieving selective inhibition and recruitment, we assume that, for all sets of constants $\mathbf{c}_{i+1}^1 \in \mathbb{R}^{n_{i+1}-r_{i+1}}$, $\mathbf{c}_i^1 \in \mathbb{R}^{n_i-r_i}$, and $\mathbf{c}_j^1 \in \mathbb{R}^{n_j-r_j}$, $j \in \{i+2, \ldots, N\} \cup \{T\}$, the reduced-order dynamics

$$\tau_i \dot{\mathbf{x}}_i^1 = -\mathbf{x}_i^1 + [\mathbf{W}_{i,i}^{11}\mathbf{x}_i^1 + \mathbf{W}_{i,i+1}^{11} h_{i+1}^1(\mathbf{W}_{i+1,i}^{11}\mathbf{x}_i^1 + \mathbf{c}_{i+1}^1) + \mathbf{W}_{i,T}^{11}\mathbf{x}_{T_{e(i)}} + \mathbf{c}_i^1]_{\mathbf{0}}^{\mathbf{m}_i^1}, \tag{5.19}$$

are GES to an equilibrium trajectory defined by the constants.

Layer Directly above the Thalamus

For the layer $\mathcal{N}_a$ directly above the thalamus, while still dependent on two layers below it in the hierarchy, one of these layers is below the thalamus. In the reduced-order model, this changes the equilibrium maps. Thus, for all constants $\mathbf{c}_a^1 \in \mathbb{R}^{n_a-r_a}$, $\mathbf{c}_b^1 \in \mathbb{R}^{n_b-r_b}$, and $\mathbf{c}_i^1 \in \mathbb{R}^{n_i-r_i}$, $i \in \{b+1, \ldots, N\} \cup \{T\}$, we assume its reduced-order dynamics

$$\tau_a \dot{\mathbf{x}}_a^1 = -\mathbf{x}_a^1 + [\mathbf{W}_{a,a}^{11}\mathbf{x}_a^1 + \mathbf{W}_{a,b}^{11} h_b^1(\mathbf{W}_{b,a}^{11}\mathbf{x}_a^1 + \mathbf{W}_{b,T}\mathbf{x}_{T_{e(a)}} + \mathbf{c}_b^1) + \mathbf{W}_{a,T}^{11}\mathbf{x}_{T_{e(a)}} + \mathbf{c}_a^1]_{\mathbf{0}}^{\mathbf{m}_a^1}, \tag{5.20}$$

are GES to an equilibrium trajectory determined by the chosen constants.

Thalamus Layer

Since the thalamus layer depends on all the other layers in the network, its reduced-order dynamics are dependent on the $N - b$ equilibrium maps of all the layers (from $\mathcal{N}_b$ to $\mathcal{N}_N$) below the thalamus. For all constants $\mathbf{c}_b^1 \in \mathbb{R}^{n_b - r_b}, \ldots, \mathbf{c}_N^1 \in \mathbb{R}^{n_N - r_N}$ and $\mathbf{c}_T^1 \in \mathbb{R}^{n_T - r_T}$, we assume the reduced-order dynamics

$$\tau_T \dot{\mathbf{x}}_T^1 = -\mathbf{x}_T^1 + [\mathbf{W}_T^{11}\mathbf{x}_T^1 + \sum_{j=1}^{a} \mathbf{W}_{T,j}^{1,\text{all}}\mathbf{x}_j^1 + \sum_{j=b}^{N} \mathbf{W}_{T,j}^{11}\mathbf{x}_{j_e}^1 + \mathbf{c}_T^1]_{\mathbf{0}}^{\mathbf{m}_T^1}, \tag{5.21}$$

are GES to an equilibrium trajectory defined by the constants $\mathbf{c}_b^1, \ldots, \mathbf{c}_N^1$ and $\mathbf{c}_T^1$.

Layers below the Thalamus

Finally, the layers below the thalamus are directly dependent on only one layer below them in the hierarchy. For all constants $\mathbf{c}_{i+1}^1 \in \mathbb{R}^{n_{i+1} - r_{i+1}}$ and $\mathbf{c}_i^1 \in \mathbb{R}^{n_i - r_i}$, we assume its reduced-order dynamics

$$\tau_i \dot{\mathbf{x}}_i^1 = -\mathbf{x}_i^1 + [\mathbf{W}_{i,i}^{11}\mathbf{x}_i^1 + \mathbf{W}_{i,i+1}^{11} h_{i+1}^1(\mathbf{W}_{i+1,i}^{1,\text{all}}\mathbf{x}_i + \mathbf{W}_{i+1,T}^{1,\text{all}}\mathbf{x}_T + \mathbf{c}_{i+1}^1) + \mathbf{W}_{i,T}^{1,\text{all}}\mathbf{x}_T + \mathbf{c}_i^1]_{\mathbf{0}}^{\mathbf{m}_i^1}, \tag{5.22}$$

and we assume they are GES to an equilibrium trajectory dependent upon the constants $\mathbf{c}_{i+1}^1$ and $\mathbf{c}_i^1$.

The stability of task-relevant dynamics in specific regions (subnetworks) is a reasonable assumption, given that this phenomenon has been widely observed within the brain [143]. In our model above, we do not fully control the task-relevant components of these dynamics

since they are mediated through the interconnection with other subnetworks. is why we next turn our attention to identifying conditions under which the reduced-order dynamics for layers $\mathcal{N}_2, \ldots, \mathcal{N}_N, \mathcal{N}_T$ are indeed GES. The following result makes use of the fact that the equilibrium maps are piecewise-affine, cf. Lemma 5.5.1.

Lemma 5.5.2. (Sufficient condition for existence and uniqueness of equilibria and GES in multilayer linear-threshold networks with parallel connections): *Let $h_i : \mathbb{R}^n \to \mathbb{R}^n$, $i \in \{1, \ldots, K\}$, be piecewise-affine functions,*

$$h_i(\mathbf{c}) = \mathbf{F}^i_{\lambda_i}\mathbf{c} + \mathbf{f}^i_{\lambda_i}, \qquad \forall \mathbf{c} \in \Psi_{\lambda_i} \triangleq \{\mathbf{c} \mid \mathbf{G}^i_{\lambda_i}\mathbf{c} + \mathbf{g}^i_{\lambda_i}\}$$

for all $\lambda_i \in \Lambda_i$ where Λ_i is a finite index set such that $\bigcup_{\lambda_i \in \Lambda_i} \Psi_{\lambda_i} = \mathbb{R}^n$. Define $\bar{\mathbf{F}}_i \triangleq \max_{\lambda_i \in \Lambda_i} |\mathbf{F}^i_{\lambda_i}|$ as the matrix made of the entry-wise maximum of the elements in $\{|\mathbf{F}^i_{\lambda_i}|\}_{\lambda_i \in \Lambda_i}$. For $i \in \{1, \ldots, K\}$ let the matrices $\mathbf{W}^i_1, \mathbf{W}^i_2$, be arbitrary and also consider arbitrary matrix $\mathbf{W}$. Then, if $\rho(|\mathbf{W}| + \sum_{i=1}^K |\mathbf{W}^i_1|\bar{\mathbf{F}}_i|\mathbf{W}^i_2|) < 1$, the dynamics

$$\tau\dot{\mathbf{x}} = -\mathbf{x} + [\mathbf{W}\mathbf{x} + \sum_{i=1}^{K} \mathbf{W}^i_1 h_i(\mathbf{W}^i_2\mathbf{x} + \bar{\mathbf{c}}_i) + \mathbf{c}]^{\mathbf{m}}_{\mathbf{0}}$$

is GES to a unique for all constants $\bar{\mathbf{c}}_i, \mathbf{c}$.

Lemma 5.5.2 generalizes [65, Theorem IV.4] to thalamocortical networks and its proof follows a similar line of arguments. The application of this result to our setting results in, if

$$\begin{aligned}
\rho(|\mathbf{W}^{11}_{i,i}| + |\mathbf{W}^{11}_{i,i+1}|\bar{\mathbf{F}}_{i+1}|\mathbf{W}^{11}_{i+1,i}| + |\mathbf{W}^{11}_{i,T}|\bar{\mathbf{F}}_T|\mathbf{W}^{11}_{T,i}|) < 1 \qquad & \forall\ i \in \{2, \ldots, a-1\} \\
\rho(|\mathbf{W}^{11}_{i,i}| + |\mathbf{W}^{11}_{i,i+1}|\bar{\mathbf{F}}_{i+1}|\mathbf{W}^{11}_{i+1,i}|) < 1 \qquad & \forall\ i \in \{b, \ldots, N\} \\
\rho(|\mathbf{W}^{11}_{T}| + \sum_{i=b}^{N} |\mathbf{W}^{11}_{T,i}|\bar{\mathbf{F}}_i|\mathbf{W}^{11}_{i,T}|) < 1, &
\end{aligned}$$

then the dynamics (5.19)-(5.22) are GES to an equilibrium.

Remark 5.5.3. (Comparison with conditions for a strictly cortical hierarchical network): *Sufficient conditions for GES of the reduced-order dynamics of a strictly cortical network can be obtained from the above by allowing* $\mathbf{W}_{i,T}^{11} = 0$ *for all i, and then these conditions reduce to those found in [65], as expected. These sufficient conditions to guarantee GES, for layers above the thalamus, are harder to satisfy in the thalamocortical network than in a cortical one. However, as the conditions are only sufficient, this does not mean that it is in fact more difficult to achieve GES of the reduced-order dynamics in the thalamocortical case, as the proof above does not explicitly invoke the inhibitory nature of the thalamus.* •

5.5.3 Selective Inhibition and Recruitment

We are ready to illustrate how selective inhibition and recruitment can be achieved in the hierarchical thalamocortical network model. Here, we first formalize the concept mathematically and then provide a feedforward-feedback control that achieves it. Recall that selective inhibition corresponds to the task-irrelevant components of the network converging to zero, and recruitment corresponds to having the task-relevant components converge to an equilibrium trajectory. The timescale ratio between layers, ϵ, must approach zero to encode the separation of timescales observed in the brain. As such, we require convergence of the task-relevant components of the network to an equilibrium as this ratio approaches zero. Formally, selective inhibition and recruitment is achieved if the following equations are satisfied for any $0 < t_1 < t_2 < \infty$: first, for all layers $\mathcal{N}_i$, $i \in \{2, \ldots, N\} \cup \{T\}$, it holds that

$$\text{[inhibition]:} \quad \lim_{\epsilon \to 0} \sup_{t \in [t_1, t_2]} \|\mathbf{x}_i^0(t)\| = 0; \tag{5.23a}$$

second, for the top layer $\mathcal{N}_1$ in the hierarchy,

$$\text{[driving layer]:} \quad \lim_{\epsilon \to 0} \sup_{t \in [t_1, t_2]} \|\mathbf{x}_1^1(t) - \bar{\mathbf{x}}_1^1(t)\| = 0; \tag{5.23b}$$

for all layers $\{\mathcal{N}_i\}_{i=2}^{a}$ above the thalamus,

$$\text{[recruitment]:} \quad \lim_{\epsilon\to 0} \sup_{t\in[t_1,t_2]} \|\mathbf{x}_i^1(t) - h_i^1(\mathbf{W}_{i,i-1}^{11}\mathbf{x}_{i-1}^1(t) + \mathbf{c}_i^1)\| = 0; \tag{5.23c}$$

for the thalamus layer,

$$\text{[recruitment]:} \quad \lim_{\epsilon\to 0} \sup_{t\in[t_1,t_2]} \|\mathbf{x}_T^1(t) - x_{T_{e(a)}}\| = 0; \tag{5.23d}$$

and finally, for the layers $\{\mathcal{N}_i\}_{i=b}^{N}$ below the thalamus,

$$\text{[recruitment]:} \quad \lim_{\epsilon\to 0} \sup_{t\in[t_1,t_2]} \|\mathbf{x}_i^1(t) - h_i^1(\mathbf{W}_{i,i-1}^{1,\text{all}}\mathbf{x}_{i-1}(t) + \mathbf{W}_{i,T}^{1,\text{all}}\mathbf{x}_T(t) + \mathbf{c}_i^1)\| = 0. \tag{5.23e}$$

Intuitively, we note from (5.23) that achieving selective inhibition and recruitment means that one can make the error between the network trajectories and the equilibrium trajectories arbitrarily small if the timescale ratio is small enough. The following result shows that selective inhibition and recruitment can be achieved in the hierarchical thalamocortical network by means of a combination of feedforward and feedback control.

Theorem 5.5.4. (Selective inhibition and recruitment in hierarchical thalamocortical networks with Feedforward-Feedback Control): *Consider an N-layer thalamocortical network as shown in Fig. 5.1(a), with layer dynamics given by* (5.5) *and* (5.7)*. Without loss of generality, let* $\tau_1 \gg \tau_2 \gg \cdots \gg \tau_T \gg \cdots \gg \tau_N$ *and* $a, b \in \{1,\dots,N\}$ *such that* $\tau_a \gg \tau_T \gg \tau_b$*, with* $b = a+1$*. Assume the stability conditions* (5.18)-(5.22) *for the reduced-order subnetworks are satisfied. Then, for* $i \in \{1,\dots,N\}\cup\{T\}$ *and constants* $\mathbf{c}_i \in \mathbb{R}^{n_i}$ *and* $\mathbf{c}_T \in \mathbb{R}^{n_T}$*, there exist control laws* $\mathbf{u}_i(t) = \mathbf{K}_i\mathbf{x}_i(t) + \bar{\mathbf{u}}_i(t)$*, with* $\mathbf{K}_i \in \mathbb{R}^{p_i\times n_i}$ *and* $\bar{\mathbf{u}}_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}^{p_i}$*, such that the closed-loop system achieves selective inhibition and recruitment* (5.23)*.*

Proof of Theorem 5.5.4. We prove the result by constructing a control and iteratively applying the generalization of Tikhonov's theorem from [135]. Throughout the proof, we make

use of the following notation. For $i \leq a$, let $\mathbf{x}_{1:i} = \begin{bmatrix} \mathbf{x}_1^\top & \dots & \mathbf{x}_i^\top \end{bmatrix}^\top$ and, for $i \geq b$, let $\mathbf{x}_{1:T:i} = \begin{bmatrix} \mathbf{x}_1^\top \dots \mathbf{x}_a^\top & \mathbf{x}_T^\top & \mathbf{x}_b^\top & \dots & \mathbf{x}_i^\top \end{bmatrix}^\top$. We first define the control for layer $\mathcal{N}_N$. Let $\mathbf{K}_N$ and $\bar{\mathbf{u}}_N(t)$ be such that

$$\mathbf{B}_N^0 \mathbf{K}_N \leq -\mathbf{W}_{N,N}^{0,\text{all}}, \tag{5.24a}$$

$$\mathbf{B}_N^0 \bar{\mathbf{u}}_N(t) \leq -\mathbf{W}_{N,N-1}^{0,\text{all}} \mathbf{x}_{N-1}(t) - \mathbf{W}_{N,T}^{0,\text{all}} \mathbf{x}_T(t) - \mathbf{c}_N^0. \tag{5.24b}$$

The inequalities (5.24) can be satisfied due to our assumption that the matrices $\mathbf{B}_i^0$ have full rank and $p_i \geq r_i$ for all $i \in \{1, \dots, N\} \cup \{T\}$. Substituting (5.24) into the dynamics,

$$\begin{aligned}
\tau_1 \dot{\mathbf{x}}_1 &= -\mathbf{x}_1 + [\mathbf{W}_{1,1}\mathbf{x}_1 + \mathbf{W}_{1,2}\mathbf{x}_2 + \mathbf{W}_{1,T}\mathbf{x}_T + \mathbf{c}_1]_{\mathbf{0}}^{\mathbf{m}_1}, \\
&\vdots \\
\tau_T \dot{\mathbf{x}}_T &= -\mathbf{x}_T + \left[\mathbf{W}_T \mathbf{x}_T + \sum_{j=1}^{N} \mathbf{W}_{T,j}\mathbf{x}_j + \mathbf{B}_T \mathbf{u}_T + \mathbf{c}_T\right]_{\mathbf{0}}^{\mathbf{m}_T}, \\
&\vdots \\
\epsilon_N \tau_{N-1} \dot{\mathbf{x}}_N^0 &= -\mathbf{x}_N^0, \\
\epsilon_N \tau_{N-1} \dot{\mathbf{x}}_N^1 &= -\mathbf{x}_N^1 + [\mathbf{W}_{N,N}^{1,\text{all}} \mathbf{x}_N + \mathbf{W}_{N,N-1}^{1,\text{all}} \mathbf{x}_{N-1} + \mathbf{W}_{N,T}^{1,\text{all}} \mathbf{x}_T + \mathbf{c}_N^1]_{\mathbf{0}}^{\mathbf{m}_1^1}.
\end{aligned}$$

Taking $\epsilon_N \to 0$ then provides a separation of timescales between $\mathbf{x}_N$ and $\mathbf{x}_{1:T:N-1}$. Note that, by assumption, the reduced-order dynamics (5.22) is GES. Using then the fact that the cascaded interconnection of a GES system with an exponentially vanishing system is also GES, cf. [64, Lemma A.1], we deduce that, for any constants $\mathbf{x}_{N-1}$ and $\mathbf{x}_T$, $\mathbf{x}_N$ is GES to $(\mathbf{0}_{r_N}, h_N^1(\mathbf{W}_{N,N-1}^{1,\text{all}} \mathbf{x}_{N-1} + \mathbf{W}_{N,T}^{1,\text{all}} \mathbf{x}_T + \mathbf{c}_N^1))$. Recalling that, by Lemma 5.5.1 and [65, Lemma IV.2], the equilibrium maps h_i^1 are globally Lipschitz for all $i \in \{1, \dots, N\} \cup \{T\}$, and noting that the entire network is Lipschitz due to the Lipschitzness of the linear-threshold function $[\cdot]_{\mathbf{0}}^{\mathbf{m}}$, we can apply [135, Proposition 1],

giving for $0 < t_1 < t_2 < \infty$,

$$\begin{aligned}
&\lim_{\epsilon_1 \to 0} \sup_{[t_1,t_2]} \|\mathbf{x}_N^0(t)\| = 0, \\
&\lim_{\epsilon_1 \to 0} \sup_{[t_1,t_2]} \left\|\mathbf{x}_N^1(t) - h_N^1(\mathbf{W}_{N,N-1}^{1,\text{all}} \mathbf{x}_{N-1}(t) + \mathbf{W}_{N,T}^{1,\text{all}} \mathbf{x}_T(t) + \mathbf{c}_N^1)\right\| = 0, \\
&\lim_{\epsilon_1 \to 0} \sup_{[t_1,t_2]} \|\mathbf{x}_{1:T:N-1} - \mathbf{x}_{1:T:N-1}^{(1)}\| = 0,
\end{aligned} \tag{5.25}$$

where $\mathbf{x}_{1:T:N-1}^{(1)}$ represents the first-step reduced-order model coming from replacing $\mathbf{x}_N$ by its equilibrium value $(\mathbf{0}_{r_N}, h_N^1(\mathbf{W}_{N,N-1}^{1,\text{all}} \mathbf{x}_{N-1} + \mathbf{W}_{N,T}^{1,\text{all}} \mathbf{x}_T + \mathbf{c}_N^1))$. We continue the proof by iterating the above process for each network layer, utilizing constructed control laws and applying [135, Proposition 1] to the corresponding reduced-order models built from substitution of the equilibrium values.

We now construct control laws for layers $\mathcal{N}_i$ with $i \in \{b, \ldots, N-1\}$. We choose $\mathbf{K}_i$ and $\bar{\mathbf{u}}_i(t)$ such that

$$\mathbf{B}_i^0 \mathbf{K}_i \leq -|\mathbf{W}_{i,i}^{0,\text{all}}| - |\mathbf{W}_{i,i+1}^{01}| \bar{\mathbf{F}}_{i+1} |\mathbf{W}_{i+1,i}^{1,\text{all}}|, \tag{5.26a}$$

$$\mathbf{B}_i^0 \bar{\mathbf{u}}_i(t) \leq -\mathbf{W}_{i,i-1}^{0,\text{all}} \mathbf{x}_{i-1}(t) - \mathbf{W}_{i,T}^{0,\text{all}} \mathbf{x}_T(t) - \mathbf{c}_i^0 - |\mathbf{W}_{i,i+1}^{01}| \bar{\mathbf{F}}_{i+1} |\mathbf{W}_{i+1,T}^{1,\text{all}} \mathbf{x}_T(t) + \mathbf{c}_{i+1}^1|, \tag{5.26b}$$

in which $\bar{\mathbf{F}}_i \in \mathbb{R}^{(n_i - r_i) \times (n_i - r_i)}$ is the entry-wise maximal gain of the equilibrium map $h_i^1(\cdot)$ as defined in Lemma 5.5.2. We use the control law (5.26) to construct the reduced-order model, consider the timescale separation by letting $\epsilon_i \to 0$ for $i \in \{b, \ldots, N-1\}$, and finally apply [135, Proposition 1] to obtain

$$\begin{aligned}
&\lim_{\epsilon \to 0} \sup_{t \in [t_1,t_2]} \|\mathbf{x}_i^{(N-i)^0}(t)\| = 0, \\
&\lim_{\epsilon \to 0} \sup_{t \in [t_1,t_2]} \|\mathbf{x}_i^{(N-i)^1}(t) - h_i^1(\mathbf{W}_{i,i-1}^{1,\text{all}} \mathbf{x}_{i-1}^{(N-i)}(t) + \mathbf{W}_{i,T}^{1,\text{all}} \mathbf{x}_T^{(N-i)}(t) + \mathbf{c}_i^1)\| = 0, \\
&\lim_{\epsilon \to 0} \sup_{t \in [t_1,t_2]} \|\mathbf{x}_{1:T:i-1}^{(N-i)}(t) - \mathbf{x}_{1:T:i-1}^{(N-i+1)}(t)\| = 0,
\end{aligned}$$

for all $i \in \{b, \dots, N-1\}$. For the thalamus layer, we note that provided the initial conditions lie in $[\mathbf{0}, \mathbf{m}_i]$ for all $i \in \{1, \dots, N\} \cup \{T\}$, by the properties of the linear-threshold dynamics, we have that $\mathbf{x}_i(t) \leq \mathbf{m}_i$ for all $t \geq 0$. Utilizing these bounds, we define the control for the thalamus layer such that it satisfies

$$\mathbf{B}_T^0 \mathbf{K}_T \leq -|\mathbf{W}_T^{0,\text{all}}| - \sum_{j=b}^{N} |\mathbf{W}_{T,j}^{01}| \bar{\mathbf{F}}_j |\mathbf{W}_{j,T}^{0,\text{all}}|, \tag{5.27a}$$

$$\mathbf{B}_T^0 \bar{\mathbf{u}}_T(t) \leq \sum_{j=1}^{1} \mathbf{W}_{T,j}^{0,\text{all}} \mathbf{x}_j(t) - \mathbf{c}_T^1 - |\mathbf{W}_{T,b}^{0,\text{all}}| \bar{\mathbf{F}}_b |\mathbf{W}_{b,a}^{1,\text{all}} \mathbf{x}_a(t) + \mathbf{c}_b^1| - \sum_{j=b+1}^{N} |\mathbf{W}_{T,j}^{01}| \bar{F}_j |\mathbf{W}_{j,j-1}^{1,\text{all}} \mathbf{m}_{j-1} + \mathbf{c}_j^1|. \tag{5.27b}$$

Then, after constructing the reduced-order model using the control laws (5.27a) and (5.27b) and letting $\epsilon_T \to 0$ to create the timescale separation, we again apply [135, Proposition 1] to get

$$\begin{aligned}
&\lim_{\epsilon \to 0} \sup_{t \in [t_1, t_2]} \|\mathbf{x}_T^{(T)^0}(t)\| = 0, \\
&\lim_{\epsilon \to 0} \sup_{t \in [t_1, t_2]} \|\mathbf{x}_T^{(T)^1}(t) - h_T^1(\sum_{j=1}^{a} \mathbf{W}_{T,j}^{1,\text{all}} \mathbf{x}_j^{(T)}(t) + \mathbf{c}_T^1)\| = 0, \\
&\lim_{\epsilon \to 0} \sup_{t \in [t_1, t_2]} \|\mathbf{x}_{1:a}^{(T)}(t) - \mathbf{x}_{1:a}^{(N-b+1)}(t)\| = 0.
\end{aligned}$$

What remains is to consider the layers above the thalamus. These layers maintain the same form of control law except for the layer immediately above the thalamus. For layer $\mathcal{N}_a$ we define terms $\mathbf{K}_a$ and $\bar{\mathbf{u}}_a(t)$ such that

$$\mathbf{B}_a^0 \mathbf{K}_a \leq -|\mathbf{W}_{a,a}^{0,\text{all}}| - |\mathbf{W}_{a,b}^{01}| \bar{\mathbf{F}}_b |\mathbf{W}_{b,a}^{1,\text{all}}|, \tag{5.28a}$$

$$\mathbf{B}_a^0 \bar{\mathbf{u}}_a(t) \leq -\mathbf{W}_{a,a-1}^{0,\text{all}} \mathbf{x}_{a-1}(t) - |\mathbf{W}_{a,b}^{01}| \bar{\mathbf{F}}_b |\mathbf{W}_{b,T}^{1,\text{all}} \mathbf{x}_{T_{e(a)}} + \mathbf{c}_b^1|. \tag{5.28b}$$

Now, for layers $\mathcal{N}_i$, $i \in \{2, \dots, a-1\}$, we let the controls $\mathbf{K}_i$ and $\bar{\mathbf{u}}_i(t)$ be such that the following

hold

$$\mathbf{B}_i^0\mathbf{K}_i \leq -|\mathbf{W}_{i,i}^{0,\text{all}}| - |\mathbf{W}_{i,i+1}^{01}|\bar{\mathbf{F}}_{i+1}|\mathbf{W}_{i+1,i}^{1,\text{all}}|, \tag{5.29a}$$

$$\mathbf{B}_i^0\bar{\mathbf{u}}_i(t) \leq -\mathbf{W}_{i,i-1}^{0,\text{all}}\mathbf{x}_{i-1}(t) - |\mathbf{W}_{i,i+1}^{01}|\bar{\mathbf{F}}_{i+1}|\mathbf{c}_{i+1}^1|, \tag{5.29b}$$

Once the reduced-order models are constructed, letting $\epsilon_i \to 0$ for $i \in \{2,\dots,a\}$, applying [135, Proposition 1] gives

$$\lim_{\epsilon\to 0}\sup_{t\in[t_1,t_2]} \|\mathbf{x}_i^{(N-i)^0}(t)\| = 0,$$

$$\lim_{\epsilon\to 0}\sup_{t\in[t_2,t_2]} \|\mathbf{x}_i^{(N-i)^1}(t) - h_i^1(W_{i,i-1}^{1,\text{all}}\mathbf{x}_{i-1}^{(N-i)}(t) + \mathbf{c}_i^1)\| = 0,$$

$$\lim_{\epsilon\to 0}\sup_{t\in[t_1,t_2]} \|\mathbf{x}_{1:i-1}^{(N-i)}(t) - \mathbf{x}_{1:i-1}^{(N-i+1)}(t)\| = 0.$$

for $i \in \{2,\dots,a\}$. The equations (5.23) for selective inhibition and recruitment, are then obtained through repeated application of the triangle inequality, completing the proof. □

Fig. 5.2 illustrates Theorem 5.5.4. While the result is for the case in which the thalamus has a timescale inside the hierarchy, it still holds when this timescale is at the top or bottom of the hierarchy (the proof remains the same, with the appropriate modifications to the inequalities derived in (5.24)-(5.29)).

Remark 5.5.5. (Comparison with strictly cortical networks): *Regarding selective inhibition and recruitment, Theorem 5.5.4 is to multilayer thalamocortical networks what [65, Theorem IV.3] is to multilayer cortical networks. Despite the analytical similarities, the consideration of the thalamus provides a significant generalization from a biological perspective, as transthalamic connections exist in most brain networks [94]. Technically, the addition of the thalamus, while only adding a layer, significantly complicates the analysis due to its connection with all of the cortical layers. These connections result in every layer having connections from timescales not simply immediately above or below it in the hierarchy, which impacts the determination of*

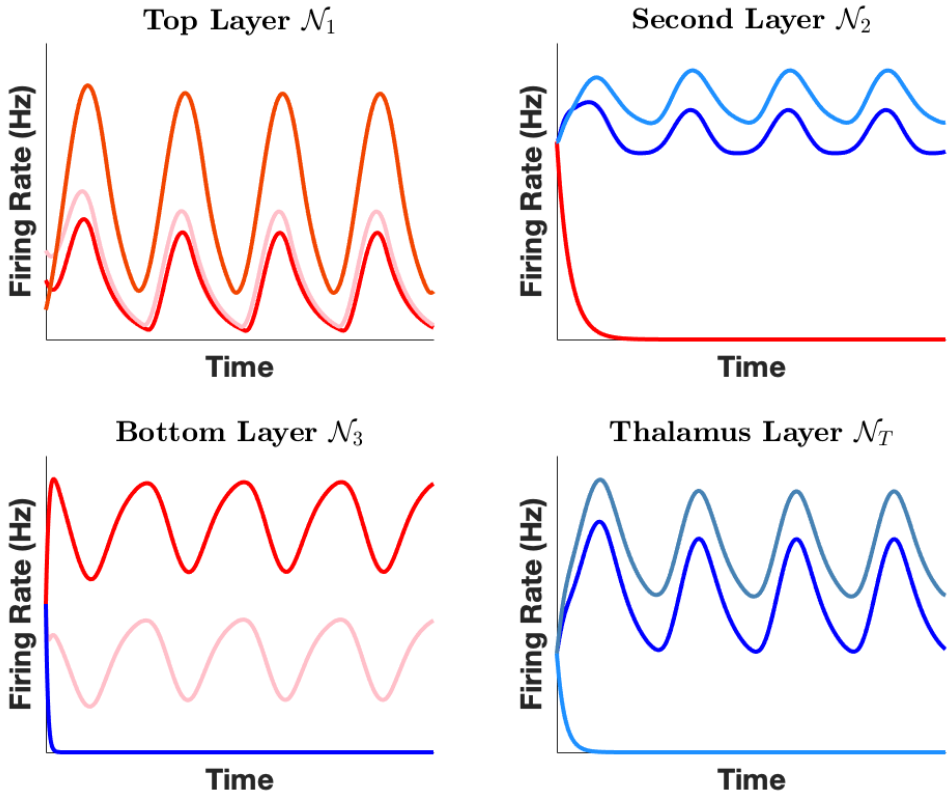


Figure 5.2. Trajectories of a three-layer hierarchical thalamocortical network under a periodic input. Red and blue lines correspond with excitatory and inhibitory nodes, resp. Each layer consists of three nodes with the top layer in the network containing only excitatory nodes, the thalamus being strictly inhibitory and layers two and three having both excitatory and inhibitory nodes. Using a control law designed through Theorem 5.5.4, in each layer other than the top layer, nodes are able to be either selectively inhibited to zero or recruited into a periodic equilibrium trajectory. Timescale ratios are $\epsilon_2 = 0.54$, $\epsilon_3 = 0.15$ and $\epsilon_T = 0.61$.

convergence to equilibrium values for layers above the thalamus. The control laws (5.24)-(5.29) *allow for smaller energy sufficient controls than for the strictly cortical networks in [65] due to the inhibitory properties of the thalamic connection matrices* $\mathbf{W}_{i,T}$*, cf. Section 5.7.* •

Remark 5.5.6. (Timescale separation): *The assumption of timescale separation between layers in Theorem 5.5.4 allows us to use singular perturbation theory and is justified by the different complexities involved in processing information along pathways in the brain, cf Section 5.3. Such temporal hierarchies naturally occur in many other network systems, e.g., social networks or electric circuits [134]. In practice, a ratio of timescales as small as* 0.5 *can provide selective*

recruitment within a small error bound, see e.g., [65, Example III.4] and our simulation results in Section 5.7. •

Remark 5.5.7. (Recruited Equilibrium Trajectory): *In the definition of selective recruitment* (5.23) *we see that the task-relevant component of the network gets recruited to a trajectory defined by the equilibrium maps* (5.12)- (5.14)*. Due to their recursively-defined nature, the actual values of these maps can only be selected through choice of the constant inputs* $\mathbf{c}_i^1$*, and it is difficult to choose any particular equilibrium trajectory in this manner. However for many applications, including selective attention, it is desirable to be able to converge to a given reference signal. This difficulty will be addressed in Chapter 6.* •

5.6 Star-connected Thalamocortical Networks

In this section we consider star-connected thalamocortical networks, cf. Fig. 5.1(b), where the cortical regions are each connected only to the thalamus and the dynamics are governed by (5.8)-(5.9). In this topology, there is no direct relationship between the timescales of each layer and as such, no hierarchical structure. Without timescale separation, the specifics of selective inhibition and recruitment, both in terms of equilibria and stability criteria, differ from the hierarchical case.

5.6.1 Equilibria and Stability Conditions

With the lack of timescale separation, the decomposition described in Section 5.5.1 of the network equilibrium map as a collection of equilibrium maps for each layer, with the state of layers higher in the hierarchy represented by a constant input, no longer holds. As such, equilibrium values must be determined concurrently for all the layers. Since the task-irrelevant components get selectively inhibited to zero, the equilibrium for the task-relevant components is

given by the solution $\mathbf{x}_i^{1^*}$, $i \in \{1,\dots,N\} \cup \{T\}$ to the system of equations:

$$\begin{aligned}\mathbf{x}_i^1 &= [\mathbf{W}_{i,i}^{1,\text{all}}\mathbf{x}_i + \mathbf{W}_{i,T}^{1,\text{all}}\mathbf{x}_T + \mathbf{c}_i^1]_{\mathbf{0}}^{\mathbf{m}_i^1}, \qquad i = 1,\dots,N,\\ \mathbf{x}_T^1 &= [\mathbf{W}_T^{1,\text{all}}\mathbf{x}_T + \sum_{i=1}^{N} \mathbf{W}_{T,i}^{1,\text{all}}\mathbf{x}_i + \mathbf{c}_T^1]_0^{\mathbf{m}_T^1}.\end{aligned} \tag{5.30}$$

The task-relevant equilibrium to which the system converges is dependent upon $\mathbf{c}_i \in \mathbb{R}^{n_i - r_i}$, $i \in \{1,\dots,N\}$ and $\mathbf{c}_T \in \mathbb{R}^{n_T - r_T}$, and so we represent it by $\mathbf{x}_i^{1^*}(\mathbf{c}_1,\dots,\mathbf{c}_N,\mathbf{c}_T)$. In keeping with the role of layer $\mathcal{N}_1$ as driving the selective recruitment in the other layers, rather than being recruited itself, we consider in what follows an input signal $\mathbf{c}_1 : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_i - r_i}$, rather than a constant, that gives rise to an equilibrium trajectory $\mathbf{x}_i^{1^*}(\mathbf{c}_1(t),\dots,\mathbf{c}_N,\mathbf{c}_T)$ for the dynamics.

As per the description of the star-connected thalamocortical network, cf Section 5.4, only the task-irrelevant component of the dynamics is directly controlled, and as such the ability to achieve selective inhibition and recruitment is dependent on the stability properties of the task-relevant components. To ensure this, we employ below the fact [64, Theorem IV.8] that, for a generic linear-threshold network mode $\tau\dot{\mathbf{x}} = -\mathbf{x} + [\mathbf{W}\mathbf{x} + \mathbf{c}]_{\mathbf{0}}^{\mathbf{m}}$, the condition $\rho(|\mathbf{W}|) < 1$ is sufficient to ensure that, for all $\mathbf{c} \in \mathbb{R}^n$, the dynamics is GES to an equilibrium.

5.6.2 Selective Inhibition and Recruitment

We are ready to formalize selective inhibition and recruitment for star-connected networks and provide conditions for its achievement. We recall that the subnetwork $\mathcal{N}_1$ corresponds with a subcortical region applying a sensory input signal to the thalamus to be relayed to the cortical regions. As such, we do not inhibit any components in this subnetwork and instead assume that it is stable to a trajectory $\bar{\mathbf{x}}_1(\mathbf{c}_1(t))$ dictated by its own input signal. For the remaining layers, we wish to inhibit the task-irrelevant components to zero and recruit the task-relevant components to the equilibria trajectory $\mathbf{x}_i^{1^*}(\mathbf{c}_1(t),\dots,\mathbf{c}_N,\mathbf{c}_T)$, $i \in \{2,\dots,N\} \cup \{T\}$. This can be formalized to

selective inhibition and recruitment is achieved if for the input layer $\mathcal{N}_1$,

$$\text{[driving layer]:} \quad \lim_{t\to\infty} \|\mathbf{x}_1(t) - \bar{\mathbf{x}}_1(\mathbf{c}_1(t))\| = 0; \tag{5.31a}$$

and for all layers $\{\mathcal{N}_i\}_{i=2}^N$ and $\mathcal{N}_T$,

$$\text{[inhibition]:} \quad \lim_{t\to\infty} \|\mathbf{x}_i^0(t)\| = 0; \tag{5.31b}$$

$$\text{[recruitment]:} \quad \lim_{t\to\infty} \|\mathbf{x}_i^1(t) - \mathbf{x}_i^{1^*}(\mathbf{c}_1(t), \ldots, \mathbf{c}_N, \mathbf{c}_T)\| = 0. \tag{5.31c}$$

We also employ a weaker notion, referred to as ϵ-selective inhibition and recruitment, which is met if there exists t^* such that the functions in (5.31) are all upper bounded by $\epsilon > 0$ for $t > t^*$.

For convenience, we also introduce the notation:

$$\bar{\mathbf{W}}^{11} = \begin{bmatrix} \mathbf{W}_{1,1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{W}_{1,T}^{11} \\ \mathbf{0} & \mathbf{W}_{2,2}^{11} & \cdots & \mathbf{0} & \mathbf{W}_{2,T}^{11} \\ \vdots & \cdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{W}_{N,N}^{11} & \mathbf{W}_{N,T}^{11} \\ \mathbf{W}_{T,1}^{11} & \mathbf{W}_{T,2}^{11} & \cdots & \mathbf{W}_{T,N}^{11} & \mathbf{W}_{T}^{11} \end{bmatrix}. \tag{5.32}$$

Note the Schur decomposition $\bar{\mathbf{W}}^{11} = \mathbf{Q}^\top(\mathbf{D}_{\bar{\mathbf{W}}^{11}} + \mathbf{N}_{\bar{\mathbf{W}}^{11}})\mathbf{Q}$, where $\mathbf{Q}$ is unitary, $\mathbf{D}_{\bar{\mathbf{W}}^{11}}$ is diagonal, and $\mathbf{N}_{\bar{\mathbf{W}}^{11}}$ is upper triangular with a zero diagonal [144]. The next result establishes conditions to achieve selective inhibition and recruitment in star-connected systems without a hierarchy of timescales.

Theorem 5.6.1. (Selective inhibition and recruitment of star-connected networks): *Consider an N-layer star-connected thalamocortical network as shown in Fig. 5.1(b), with layer dynamics given by* (5.8) *and* (5.9)*. Suppose the following hold for all values of* $\mathbf{c}_i \in \mathbb{R}^{n_i}$, $i \in \{1, \ldots, N\}$, *and* $\mathbf{c}_T \in \mathbb{R}^{n_T}$*:*

(i) The input layer $\mathcal{N}_1$ has no nodes to be inhibited, $\rho(|\mathbf{W}_{1,1}|) < 1$, and the input $\mathbf{c}_1(t)$ lies in a compact set and has a bounded rate derivative;

(ii) For each $i \in \{2,\dots,N\} \cup \{T\}$, the matrix $\mathbf{W}_{i,i}^{11}$ satisfies $\rho(|\mathbf{W}_{i,i}^{11}|) \leq \alpha_i$, with $\alpha_i < 1$;

(iii) If $\rho(\sum_{i=1}^{N} |\mathbf{W}_{i,T}^{11}\mathbf{W}_{T,i}^{11}|) \neq 0$, then $\alpha + \max(\delta, \delta^{1/p}) < 1$, where p is the dimension of $\mathbf{N}_{\bar{\mathbf{W}}^{11}}$ and

$$\alpha = \max_{i \in \{2,\dots,N\} \cup \{T\}} \{\alpha_i\}, \qquad \delta = \gamma \sum_{j=1}^{p-1} \|\mathbf{N}_{\bar{\mathbf{W}}^{11}}\|^j,$$
$$\gamma = \max\{\sum_{i=1}^{N-1} \mathbf{W}_{i,T}^{11}\mathbf{W}_{i,T}^{11^\top}, \sum_{i=1}^{N-1} \mathbf{W}_{T,i}^{11}\mathbf{W}_{T,i}^{11^\top}\}.$$

Then, there exist control laws $\mathbf{u}_i(t) = \mathbf{K}_i\mathbf{x}_i(t) + \bar{\mathbf{u}}_i(t)$, with $\mathbf{K}_i \in \mathbb{R}^{p_i \times n_i}$ and $\bar{\mathbf{u}}_i : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}^{p_i}$, and $\epsilon > 0$ such that the cortical and thalamic regions within the closed-loop system achieve $\epsilon-$selective inhibition and recruitment. Furthermore, if $\|\dot{\mathbf{c}}_1(t)\| \to 0$ as $t \to \infty$, then the network achieves selective inhibition and recruitment (5.31).

Proof of Theorem 5.6.1. First, for cortical region $\mathcal{N}_i$, $i \in \{2,\dots,N\}$, define the control laws $\mathbf{u}_i(t) = \mathbf{B}_i\mathbf{K}_i\mathbf{x}_i(t) + \mathbf{B}_i\bar{\mathbf{u}}(t)$ such that

$$\mathbf{B}_i\mathbf{K}_i \leq -\mathbf{W}_{i,i}^{0,\text{all}}, \tag{5.33a}$$
$$\mathbf{B}_i\bar{\mathbf{u}}_i(t) \leq -\mathbf{W}_{i,T}^{0,\text{all}}\mathbf{x}_T(t) - \mathbf{c}_i^0. \tag{5.33b}$$

In a similar fashion, define the control law for the thalamus, $\mathcal{N}_T$ by $\mathbf{u}_T(t) = \mathbf{B}_T\mathbf{K}_T\mathbf{x}_T(t) + \mathbf{B}_T\bar{\mathbf{u}}(t)$ such that it satisfies

$$\mathbf{B}_T\mathbf{K}_T \leq -\mathbf{W}_T^{0,\text{all}}, \tag{5.34a}$$
$$\mathbf{B}_T\bar{\mathbf{u}}_T(t) \leq -\mathbf{W}_{T,1}^{0,\text{all}}\mathbf{x}_1(t) - \mathbf{c}_T^0. \tag{5.34b}$$

Now, we permute the system variables and define corresponding timescale matrices as follows

$$\bar{\mathbf{x}} = \begin{bmatrix} \mathbf{x}^0 \\ \mathbf{x}^1 \end{bmatrix} \qquad \mathbf{x}^0 = \begin{bmatrix} \mathbf{x}_2^0 & \dots & \mathbf{x}_N^0 & \mathbf{x}_T^0 \end{bmatrix}^\top$$
$$\mathbf{x}^1 = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2^1 & \dots & \mathbf{x}_N^1 & \mathbf{x}_T^1 \end{bmatrix}^\top \qquad \tau^0 = \operatorname{diag}(\tau_2, \dots, \tau_N, \tau_T)$$
$$\tau^1 = \operatorname{diag}(\tau_1, \tau_2, \dots, \tau_N, \tau_T).$$

Substituting in the control laws (5.33) and (5.34), we have the following controlled system dynamics

$$\tau^0 \dot{\mathbf{x}}^0 = -\mathbf{x}^0 \tag{5.35a}$$
$$\tau^1 \dot{\mathbf{x}}^1 = -\mathbf{x}^1 + [\bar{\mathbf{W}}\bar{\mathbf{x}} + \mathbf{c}(t)]_{\mathbf{0}}^{\bar{\mathbf{m}}}, \tag{5.35b}$$

where $\bar{\mathbf{W}} = \begin{bmatrix} \bar{\mathbf{W}}^{10} & \bar{\mathbf{W}}^{11} \end{bmatrix}$, with $\bar{\mathbf{W}}^{11}$ as in (5.32), and $\mathbf{c}(t)$ the permutation of the signal $\mathbf{c}_1(t)$ and the constants $\mathbf{c}_i$, $i \in \{2, \dots, N\} \cup \{T\}$ corresponding to the permuted variables. Now, we consider a 'frozen' version of the dynamics (5.35), in which we fix $\mathbf{c}_1(t)$ to a constant $\bar{\mathbf{c}}_1$. By [64, Lemma A.1] the frozen version of the dynamics (5.35) is GES to an equilibrium $\mathbf{x}^*$, with $\mathbf{x}^0 \to \mathbf{0}$, if $\dot{\mathbf{y}} = -\mathbf{y} + [\bar{\mathbf{W}}^{11}\mathbf{y} + \mathbf{c}]_{\mathbf{0}}^{\bar{\mathbf{m}}}$ is GES to an equilibrium. By assumptions (ii) and (iii), along with [144, Theorem 2], we have that $\rho(\bar{\mathbf{W}}^{11}) < 1$ and therefore, the dynamics $\dot{\mathbf{y}} = -\mathbf{y} + [\bar{\mathbf{W}}^{11}\mathbf{y} + \mathbf{c}]_{\mathbf{0}}^{\bar{\mathbf{m}}}$ is GES to an equilibrium, cf. [64, Theorem IV.8]. Therefore the frozen version of (5.35) is GES to an equilibrium using the control laws defined above. From here, the combined application of Theorems 5.9.1 and 5.9.2 gives that selective inhibition and recruitment (5.31) is achieved if $\|\dot{\mathbf{c}}_1(t)\| \to 0$ and is ϵ-selectively inhibited and recruited if it is bounded but does not tend to zero. □

Note that in addition to requiring the stability of the task-relevant dynamics of each network layer when considered independently, Theorem 5.6.1 relies on two main assumptions. First, the result requires that the time derivative of the input is bounded. This is reasonable, as it corre-

sponds to the physical input to the network not changing instantaneously. Second, Theorem 5.6.1 requires that the magnitude of the combination of thalamocortical and corticothalamic interconnections does not exceed a certain stability margin. This aligns with neuroscientific observations: in fact, enhanced corticothalamic feedback may result in pathological behavior [93, 145]

Remark 5.6.2. (Remote synchronization in star-connected brain networks): *Remote synchronization is a phenomenon observed in the brain in which distant brain regions with similar structure synchronize their activity despite the lack of a direct link [146]. This should then naturally arise in star-connected networks if there is morphological symmetry between cortical regions, as this topology directly shows regions without direct links. From* (5.30)*, we note that if any two cortical regions $\mathcal{N}_i$ and $\mathcal{N}_j$ have identical task-relevant dynamics, i.e., $\mathbf{W}_{i,i}^{1,\text{all}} = \mathbf{W}_{j,j}^{1,\text{all}}$, $\mathbf{W}_{i,T}^{1,\text{all}} = \mathbf{W}_{j,T}^{1,\text{all}}$ and $\mathbf{c}_i^1 = \mathbf{c}_j^1$, it follows that the equilibrium points will satisfy $\mathbf{x}_i^{1^*} = \mathbf{x}_j^{1^*}$, meaning that remote synchronization is achieved provided that the conditions of Theorem 5.6.1 are satisfied. Remarkably, this conclusion seems to be independent of the particular dynamics of the individual layers. In fact, the work [147] studies remote synchronization in star-connected brain networks, cf. Fig. 5.1(b), with layer dynamics given by Kuramoto oscillator dynamics,*

$$\begin{aligned} \dot{\theta}_T &= \omega_0 + \sum_{i=1}^{N} K_i \sin(\theta_i - \theta_0 - \xi) \\ \dot{\theta}_i &= \omega + A_i \sin(\theta_0 - \theta_i - \xi), \qquad i = 1, \ldots, N. \end{aligned} \tag{5.36}$$

According to [147], the outer cortical regions in star-connected brain networks can remotely synchronize despite no direct links between the regions provided the network dynamics satisfy conditions that parallel those required for the star-connected linear-threshold networks studied above. In particular, to be able to achieve remote synchronization, the network weights must satisfy $A_i \geq (N-1)K_i$ for all $i \in \{1, \ldots, N\}$. This condition guarantees the existence of a locally stable equilibrium point, and is equivalent to the requirement of the matrices defining the task-relevant components of the linear-threshold network being individually stable. •

We conclude by noting that the thalamus, as a relay, can function as a failsafe for the hierarchical thalamocortical network, allowing for selective inhibition and recruitment even if corticocortical connections become damaged. In fact, the hierarchical topology where the matrices $\mathbf{W}_{i,i-1}, \mathbf{W}_{i,i+1}$ are equal to zero for all i reduces to the star-connected topology, while maintaining the timescale separation between layers. Therefore, the star-connected topology with a hierarchical timescale structure can be considered as a failsafe for the hierarchical network. The next result provides conditions for selective inhibition and recruitment for this topology.

Corollary 5.6.3. (Selective inhibition of a star-connected hierarchical thalamocortical network): *Consider a hierarchical star-connected thalamocortical network of the form shown in Fig. 5.1(b) with timescales* $\tau_1 \gg \tau_2 \gg \cdots \gg \tau_N$ *and layer dynamics given by* (5.5) *and* (5.7)*. Without loss of generality let* τ_T *be such that* $\tau_1 \gg \tau_T \gg \tau_N$ *and let* $a, b \in \{1, \ldots, N\}$ *such that* $\tau_a \gg \tau_T \gg \tau_b$ *and* $b = a + 1$*. Assume the stability assumptions* (5.18)-(5.22) *for the reduced-order subnetworks are satisfied. Then, for* $i \in \{1, \ldots, N\} \cup \{T\}$ *and constants* $\mathbf{c}_i \in \mathbb{R}^{n_i}$ *and* $\mathbf{c}_T \in \mathbb{R}^{n_T}$*, there exist control laws* $\mathbf{u}_i(t) = \mathbf{K}_i \mathbf{x}_i(t) + \bar{\mathbf{u}}_i(t)$*, with* $\mathbf{K}_i \in \mathbb{R}^{p_i \times n_i}$ *and* $\bar{\mathbf{u}}_i : \mathbb{R}_{\geq 0} \to \mathbb{R}^{p_i}_{\geq 0}$*, such that the closed-loop system achieves selective inhibition and recruitment* (5.23)*.*

The proof of the result is similar to that of Theorem 5.5.4, with differences occurring in the constructed control laws on the basis that $\mathbf{W}_{i,i+1} = \mathbf{W}_{i,i-1} = \mathbf{0}$ for all $i \in \{1, \ldots, N\}$. The loss of these connections plays a significant role in the form of the control. In particular, the amount of feedforward control coming from the thalamus to the cortical regions increases, due to the fact that direct feedforward control between cortical regions is not possible. Fig. 5.3 illustrates Corollary 5.6.3 on the star-connected network obtained by removing the direct connections between cortical regions in the network of Fig. 5.2.

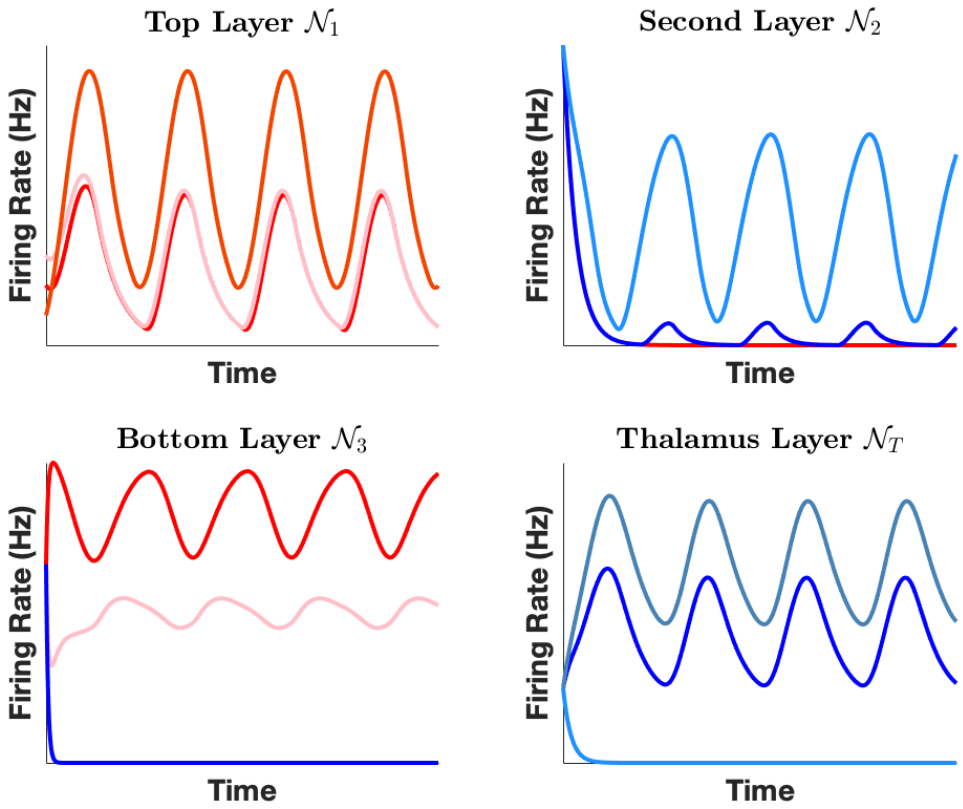


Figure 5.3. Illustration of the star-connected thalamocortical network as a safety net for the hierarchical network. The trajectories of the network in Fig. 5.2 with the connections between the cortical regions removed. The star-connected topology allows for successful selective inhibition of the desired set of nodes despite the lack of connections between cortical regions (cf. Corollary 5.6.3), providing a failsafe topology.

5.7 Quantitative Comparison of Cortical and Thalamocortical Networks

This section seeks to quantitatively illustrate ways in which the presence of the thalamus might have a beneficial effect in the behavior and performance of the dynamic models for brain networks adopted here. Fig. 5.3 has already illustrated the failsafe role played by the thalamus in hierarchical thalamocortical networks. Here we focus on two other beneficial impacts of the thalamus we observed in simulation: the control energy required to achieve selective inhibition and the convergence time in thalamocortical networks versus cortical ones.

Example 5.7.1. (Reduced average control energy in thalamocortical vs cortical networks):

We investigate the control energy required to achieve selective inhibition. Control energy here refers to the aggregate of the inputs at all layers integrated over time and averaged across trials. We consider hierarchical pairs of thalamocortical and cortical networks, where the latter is obtained by disconnecting the thalamus in the former. Fig. 5.4 shows that thalamocortical networks require a lower control energy to achieve selective inhibition in the cortical regions relative to the corresponding strictly cortical networks, matching the intuition that they are easier to selectively inhibit due to the thalamus impacting the cortical regions in an inhibitory fashion. •

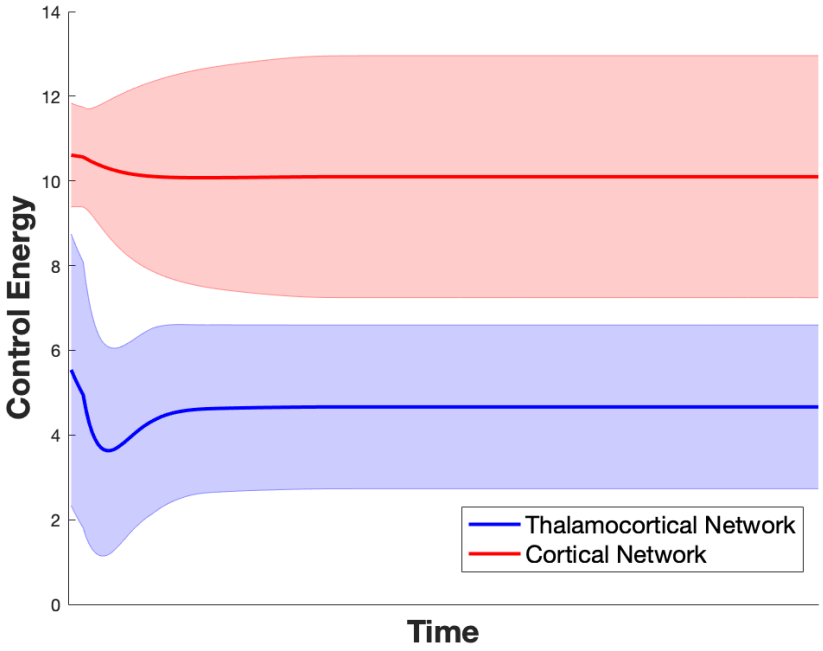


Figure 5.4. Comparison of average control energy for selective inhibition and recruitment between cortical and thalamocortical networks. The plot displays the control energy required to selectively inhibit the bottom cortical layer averaged over 100 random thalamocortical networks and the corresponding cortical network in which the thalamus is removed. Shaded regions correspond to the error bars. All networks are composed of two three-node cortical regions. The top cortical layer has two excitatory nodes and the bottom layer has only one. The thalamus is composed of two inhibitory nodes. Each thalamocortical and cortical network generated satisfies the assumptions of Theorem 5.5.4 and [65, Theorem IV.3], resp., along with biological sign constraints. To make the required control energy directly comparable, the thalamocortical and cortical networks are inhibiting the same set of nodes within the bottom cortical layer.

Example 5.7.2. (Convergence time of thalamocortical and cortical networks): *We consider the speed at which thalamocortical networks converge to an equilibrium as another metric to evaluate the role of the thalamus. We compare the convergence time for a cortical network with that of a thalamocortical network maintaining the same cortical regions. While performing the comparison is interesting as a function of multiple network parameters (e.g., network size, layer size, ratio of excitatory-inhibitory nodes), we focus specifically on the thalamus and, in particular, on varying its timescale with respect to the cortical regions. Thalamocortical networks with varying timescales are of particular interest, as even restricting only to the visual thalamus, the thalamus operates at both slow and fast timescales [148]. Fig. 5.5 shows that thalamocortical networks have faster average convergence time, with the margin between the two networks decreasing as the timescale τ_T increases. This validates the beneficial role played by the thalamus, with faster thalamus dynamics (smaller τ_T) helping the cortical regions converge faster, leading to overall decreased convergence time.* •

5.8 Conclusions

In this chapter we have investigated the properties of both multilayer hierarchical and star-connected thalamocortical brain networks modeled with linear-threshold dynamics. Our primary motivation was understanding the role played by the thalamus in achieving selective inhibition and recruitment of neural populations. For both types of interconnection topologies, we have described how the equilibria at each layer depends on the rest of the network and identified suitable stability conditions. For hierarchical networks, these take the form of GES requirements of the reduced-order dynamics of individual layers. For star-connected thalamocortical network without a hierarchy of timescales, these take the form of stability of the task-relevant dynamics of each layer when considered independently and the magnitude of the combination of thalamocortical and corticothalamic interconnections not exceeding a certain stability margin. We note that, while we have shown these properties for thalamocortical networks using the linear-threshold

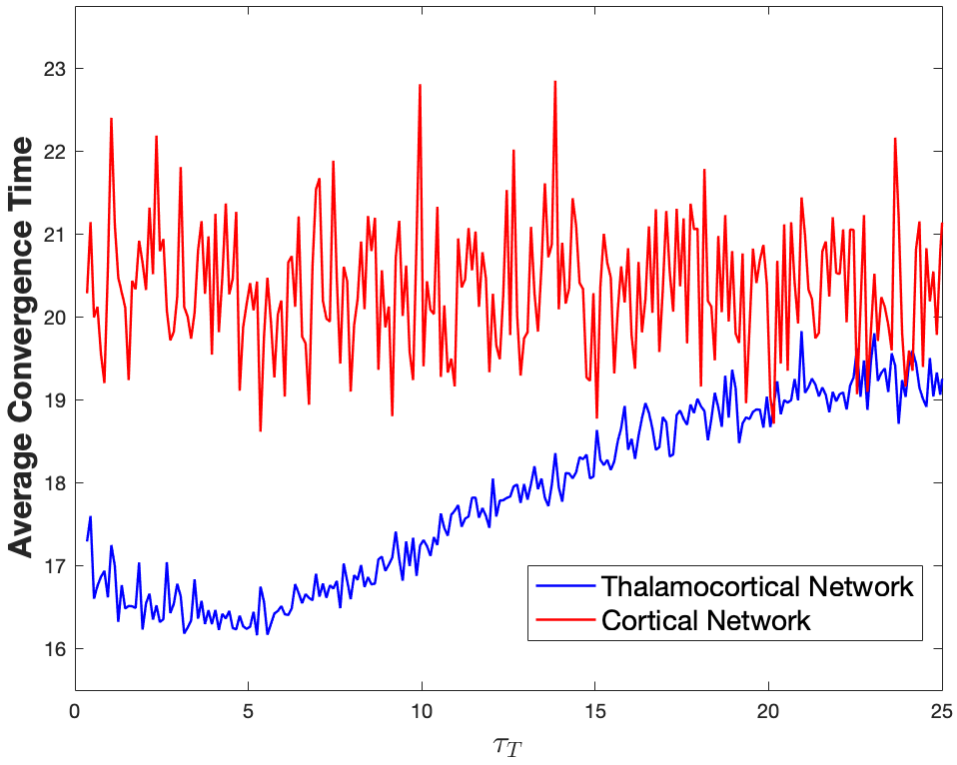


Figure 5.5. Comparison of convergence time between the cortical and thalamocortical networks with changing timescales. The graph illustrates the average convergence time to the equilibrium for 30 randomly generated cortical and thalamocortical networks with 20 layers, each containing two excitatory and two inhibitory nodes. For each simulation the initial condition is set to be a uniform distance away from the equilibrium and the set of nodes to be inhibited is randomly selected. The average convergence time for nodes within the thalamocortical network is lower than for the cortical network. As the timescale τ_T of the thalamus increases, the margin of improvement decreases.

dynamics considered throughout this dissertation, we expect them to also hold for other rate dynamics models, as they also share the underlying hierarchical and star-connected topologies, which are key to the results.

One of the difficulties encountered in selective attention and recruitment as treated in this chapter is the determination of the equilibrium signal to which the task-relevant component of the network is recruited. In the linear-threshold network model this is dependent on the recursively-defined piecewise-affine equilibrium maps, parameterized by the input term. With this formulation it is difficult to recruit the network to a specific desired equilibrium. In the next chapter we will aim to address this difficulty by considering the problem of reference tracking with the linear-threshold dynamics.

Future work will seek to analytically characterize the robustness and performance of thalamocortical networks, study the role of the thalamus in other cognitive tasks beyond selective attention (e.g., sleep consciousness, oscillations, and learning), and explore the impact of the addition and deletion of neuronal populations in the performance and expressivity of brain networks.

5.9 Chapter Appendix

For completeness, here we include two results on the stability of slowly varying nonlinear systems to a continuum equilibria, generalized from [149, Theorem 3.1] to the case of exponential stability. Let

$$\dot{\mathbf{x}} = f(\mathbf{x}(t), \mathbf{d}(t)), \tag{5.37}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ and $\mathbf{d}(t) \in \mathcal{D}$, and $\mathcal{D}$ is a compact subset of $\mathbb{R}^m$. We assume f is continuous on $\mathbb{R}^n \times \mathcal{D}$ and is locally Lipschitz in both $\mathbf{x}$ and $\mathbf{d}$. We further consider the 'frozen' version of the

system (5.37) with a fixed parameter $\mathbf{d}$,

$$\dot{\mathbf{x}} = f(\mathbf{x}(t), \mathbf{d}), \quad \mathbf{d} \in \mathcal{D}. \tag{5.38}$$

We denote the solution to (5.38) for each initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$ and $\mathbf{d} \in \mathcal{D}$ by $\mathbf{x}(t, \mathbf{x}_0, \mathbf{d})$. Let $\mathcal{A}$ be a forward invariant set of the system (5.38).

Theorem 5.9.1. (Exponential Stability of Slowly Varying Nonlinear Systems): *Consider the nonlinear system* (5.37) *and assume there exists a continuously differentiable function* $V : \mathbb{R}^n \times \mathcal{D} \to \mathbb{R}$ *such that*

$$k_1 |\mathbf{x}|_{\mathbf{A}}^2 \leq V(\mathbf{x}, \mathbf{d}) \leq k_2 |\mathbf{x}|_{\mathbf{A}}^2 \tag{5.39a}$$

$$\frac{\partial V(\mathbf{x}, \mathbf{d})}{\partial \mathbf{x}} f(\mathbf{x}, \mathbf{d}) \leq -k_3 |\mathbf{x}|_{\mathbf{A}}^2 \tag{5.39b}$$

$$\left| \frac{\partial V(\mathbf{x}, \mathbf{d})}{\partial \mathbf{d}} \right| \leq k, \tag{5.39c}$$

for all $\mathbf{x} \in \mathbb{R}^n$ *and* $\mathbf{d} \in \mathcal{D}$*, with* k_1, k_2, k_3 *and* k *nonnegative constants. If* $\|\dot{d}(t)\|$ *is uniformly bounded in time, then there exist constants* γ, λ *and* T *such that*

$$|\mathbf{x}(t, \mathbf{x}_0, \mathbf{d})|_{\mathbf{A}} \leq \gamma \|\mathbf{x}_0\| e^{-\lambda(t - t_0)} \quad \forall\ t_0 \leq t \leq t_0 + T \tag{5.40a}$$

$$|\mathbf{x}(t, \mathbf{x}_0, \mathbf{d})|_{\mathbf{A}} \leq \frac{k_3}{k_2} |\mathbf{x}_0|_{\mathbf{A}} \quad \forall\ t \geq t_0, \tag{5.40b}$$

If $\lim_{t \to \infty} \|\dot{\mathbf{d}}(t)\| = 0$*, then the system is exponentially stable.*

The proof follows a similar line of reasoning as in [149, Theorem 3.1] with slight modifications in the use of stability results from [102] to account for exponential stability. We now provide a converse Lyapunov result, modified from [149, Theorem 3.2], complementary to the above result.

Theorem 5.9.2. (Converse Lyapunov Theorem for Slowly Varying Nonlinear Systems): *If,*

for every fixed $\mathbf{d} \in \mathcal{D}$*, the nonlinear system* (5.37) *is GES to an equilibrium, then there exists a Lyapunov function that satisfies* (5.39a)*,* (5.39b) *and* (5.39c).

The proof of this result is identical to [149, Theorem 3.2], with the assumption of GES giving the desired function form in the final step.

Acknowledgements

This chapter, in full, is a reprint of the material [2] where it appears as "Selective Inhibition and Recruitment in Linear-Threshold Thalamocortical Networks" by Michael McCreesh and Jorge Cortés in IEEE Transactions on Control of Network Systems. The dissertation author was the primary investigator and author of this paper.

Chapter 6

Reference Tracking in Linear-Threshold Networks via Reservoir Computing

When a function is performed by any part of the human body, it is associated with patterns of activity in the brain. Examples include actions such as eye movements, where gaze direction is associated with particular cognitive processes [150], recollection of memories, where stored memories are associated with specific sequences of activity [151], and motion planning and movement [152]. In dealing with network models for brain dynamics, it is therefore important to ascertain their ability to exhibit a given activity pattern at a certain time. For mathematical models of neuronal activity, such as Wilson-Cowan [18], Hodgkin-Huxley [4], integrate and fire [153], sigmoidal or threshold-based firing rate models [68], among others which vary significantly in scale and properties [76, 20], this ability translates into the problem of reference tracking.

The classical approach to solving the reference tracking problem for a controlled dynamical system is the computation of an explicit input signal (that might be state-dependent) such that the desired activity pattern is exhibited by the network model. In the present context, this is undesirable for various reasons. First, the brain functions by processing the constant stream of information in the form of electrical signals it receives to achieve the desired activity pattern rather than a priori computing explicit control inputs. Second, the actual computation of these explicit expressions requires precise knowledge of the strength of synaptic interconnections in the brain, which is both impractical and challenging. Furthermore, the required conditions on

network structure for these inputs to work become increasingly difficult to check with the scale of the model.

An alternative approach to tackle both of these problems is using learning methods. Neural networks in various forms have increasingly been used to model the way that the brain learns from, and adapts to, its surroundings and stimuli [154, 155, 156, 157]. Machine learning techniques allow for the online determination of controls that result in the desired activity based on data, rather than computing explicit expressions, matching the functioning of the brain by constantly processing information, and also removes the requirement for exact knowledge of the synaptic connections. These observations motivate our study here of two brain functions that can be modeled with reference tracking using learning methods: selective inhibition/recruitment and seizure rejection.

6.1 Literature Review

To learn controls in an online fashion that solve the reference tracking problem for a brain network with a known reference, we employ machine learning techniques, which allow to consider large-scale complex networks, such as those seen in the brain [158, 159, 160]. In particular, we use reservoir computing [161, 162], which is based on training an output vector (the *readout*) from a Recurrent Neural Network (RNN), rather than its internal weights.

Reservoir computing has a variety of advantages over classical gradient-descent based RNN training, the main one being that the training process is faster and computationally inexpensive compared to standard RNN training [163]. Further, reservoir computing displays certain parallels with biological activity. First, it gives an interpretation as to how arbitrary cortical circuits without supervised adaptation are able to perform purposeful computations [164]. Second, the ability of a single reservoir to perform multiple tasks by training multiple output vectors corresponding with each task mimics neural circuits' ability to have multiple purposes [163]. Due to these parallels, reservoir computing has been used to study a variety of dynamical brain behaviors,

including connectivity and memory tasks [165], cortical dynamics in monkeys [166] and seizure detection [167]. Reservoir computing has also been used as a controller for dynamical systems, in which the reservoir output is used to determine an input such that the system converges to a desired equilibrium trajectory [168]. This has led to the framework of next-generation reservoir computing (NG-RC), proposed in [169, 170] for both prediction and control. In Section 6.5 we provide an overview of the RC and NG-RC frameworks, but for more details we refer the reader to [171, 163, 172, 169] and the references therein.

6.2 Contributions

Our first contribution relates to the ability to achieve reference tracking for the LTN dynamics. We begin by providing conditions for both existence and uniqueness of equilibrium trajectories for the dynamics. Using this fact we illustrate that under certain conditions on the reference signal and the network structure, reference tracking can be achieved for LTN networks and provide examples of controls that result in tracking. Our second contribution relates to achieving reference tracking using the data-driven machine learning frameworks. Considering the problem of reference tracking in the context of selective inhibition and recruitment and seizure rejection, we use examples to show that the reservoir computing and next-generation reservoir computing frameworks can be used as controllers to achieve reference tracking. Our contributions expand on state of the art results by considering convergence to equilibrium trajectories rather than constant equilibrium points and applying a machine learning approach.

6.3 Problem Formulation

We consider a brain model governed by the linear-threshold dynamics

$$\tau\dot{\mathbf{x}} = -\mathbf{x} + [\mathbf{W}\mathbf{x} + \mathbf{c}(t)]_{\mathbf{0}}^{\mathbf{m}}, \tag{6.1}$$

as developed in Chapter 2, where each component of $\mathbf{x} \in \mathbb{R}^n$ represents the firing rate of a population of neurons, $\mathbf{W} \in \mathbb{R}^{n \times n}$ represents the synaptic connectivity between populations, and $\mathbf{c}(t)$ captures any inputs to the network. These inputs could be internal, representing connections from neurons not in the model, or external, such as either invasive or non-invasive neurostimulation signals. The constant $\mathbf{m}$ provides a threshold on the firing rate of the neurons, and τ is a diagonal matrix defining the timescale of the dynamics for each population.

Recall from Section 3.5 that the LTN dynamics can be represented as a state-dependent switched affine system, which has switching regions defined by the threshold term $[\mathbf{W}\mathbf{x} + \mathbf{c}(t)]_{\mathbf{0}}^{\mathbf{m}}$. The dynamics has 3^N switching regions, each defined by a switching variable $\sigma \in \{0, \ell, s\}^N$, and are defined as follows

$$\begin{aligned} \Omega_\sigma = \{\mathbf{x} \mid\ & (\mathbf{W}\mathbf{x} + \mathbf{c})_i \in (-\infty, 0] \forall i \text{ s.t. } \sigma_i = 0, \\ & (\mathbf{W}\mathbf{x} + \mathbf{c})_i \in (0, \mathbf{m}_i) \forall i \text{ s.t. } \sigma_i = \ell, \text{ and} \\ & (\mathbf{W}\mathbf{x} + \mathbf{c})_i \in [\mathbf{m}_i, \infty] \forall i \text{ s.t. } \sigma_i = s\}. \end{aligned} \tag{6.2}$$

The threshold term can then be expressed over each of these regions using diagonal matrices Σ^ℓ and Σ^s. These are defined, for $q \in \{\ell, s\}$, as follows: $\Sigma^q_{ii} = 1$ if $\sigma_i = q$ and $\Sigma^q_{ii} = 0$ otherwise. This leads to the piecewise-affine form of the dynamics (6.1) being defined as

$$\tau \dot{\mathbf{x}}(t) = (-\mathbf{I} + \Sigma^\ell \mathbf{W})\mathbf{x}(t) + \Sigma^\ell \mathbf{c}(t) + \Sigma^s \mathbf{m}, \quad \mathbf{x} \in \Omega_\sigma. \tag{6.3}$$

We recall from Section 3.5 that one can identify conditions on the synaptic weight matrix $\mathbf{W}$ that ensure existence and uniqueness of equilibrium, and stability properties of the dynamics. Of interest in this chapter are the following two results:

i) if $\mathbf{W}$ satisfies $\mathbf{I} - \mathbf{W} \in \mathcal{P}$, then for each constant input $\mathbf{c}(t) = \mathbf{c}$ the dynamics (6.1) has a unique equilibrium;

ii) if $\mathbf{W} \in \mathcal{L}$, then for all constant inputs $\mathbf{c}$, the dynamics is globally exponentially stable to a unique equilibrium.

Beyond having the network converge to stable equilibria, for many applications it is instead desirable to have the network track a particular reference trajectory. This requires non-constant inputs $\mathbf{c}(t)$, designed in a way that makes the network activity converge to the desired trajectory. Due to the complexity of brain networks and the fact that their structure is not precisely known, designing such a controller analytically is challenging. This motivates the use of a data-driven learning techniques to determining appropriate control signals to achieve trajectory tracking. In this chapter we use reservoir computing, which we discuss in detail in Section 6.5, to determine controls. We formalize the problem as follows.

Problem 6.3.1. *Consider a network defined by the LTN dynamics* (6.1) *with state* $\mathbf{x}(t) \in \mathbb{R}^n$ *and reference signal* $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$. *Determine a control signal* $\mathbf{c}^* : \mathbb{R} \to \mathbb{R}^n$ *such that the network converges to the reference signal, i.e.,*

$$\lim_{t\to\infty} ||\mathbf{x}(t) - \mathbf{r}(t)|| = 0.$$

6.4 Linear Threshold Networks and Reference Tracking

In this section we show that, given a reference trajectory, there exists a control that makes the LTN dynamics track it asymptotically. This result sets the basis for our forthcoming use of reservoir computing techniques to synthesize the controller.

Given a reference signal $\mathbf{r}$, consider the error with the system state, $\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{r}(t)$. The corresponding LTN error dynamics is given by

$$\tau\dot{\mathbf{e}}(t) = -\mathbf{e}(t) + [\mathbf{W}\mathbf{e}(t) + \mathbf{W}\mathbf{r}(t) + \mathbf{c}(t)]_{\mathbf{0}}^{\mathbf{m}} - \mathbf{r}(t) - \tau\dot{\mathbf{r}}(t). \tag{6.4}$$

To show that reference tracking is achieved, we show that the error dynamics is globally

exponentially stable (GES) to the origin. We first provide conditions for the existence and uniqueness of an equilibrium trajectory for the dynamics (6.4), i.e., a trajectory that, at each $t \in \mathbb{R}_{\geq 0}$, specifies an equilibrium $\mathbf{e}^*(t)$.

Theorem 6.4.1. (Existence and Uniqueness of Equilibrium Trajectories in LTN Error Dynamics): *Consider the LTN error dynamics given by* (6.4). *If* $\mathbf{I} - \mathbf{W} \in \mathcal{P}$ *then the dynamics* (6.4) *has a unique equilibrium trajectory.*

Proof of Theorem 6.4.1. We illustrate that a unique equilibrium trajectory exists by showing that, for each time $t \in \mathbb{R}_{\geq 0}$, a unique equilibrium exists. This generates the unique equilibrium trajectory as t evolves. For fixed t, let $\mathbf{d}_1 = \mathbf{W}\mathbf{r}(t) + \mathbf{c}(t)$ and $\mathbf{d}_2 = \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t)$. This leaves us with the time-invariant dynamics

$$\tau\dot{\mathbf{e}}(t) = -\mathbf{e}(t) + [\mathbf{W}\mathbf{e}(t) + \mathbf{d}_1]_{\mathbf{0}}^{\mathbf{m}} - \mathbf{d}_2.$$

To show existence and uniqueness of an equilibrium, we note that this dynamics can be written in a similar piecewise-affine form to the original LTN dynamics (6.3),

$$\tau\dot{\mathbf{e}}(t) = (-\mathbf{I} + \Sigma^{\ell}\mathbf{W})\mathbf{e}(t) + \Sigma^{\ell}\mathbf{d}_1 + \Sigma^{s}\mathbf{m} - \mathbf{d}_2,$$

defined over the same switching regions given in (6.2) and with the same system matrix $(-\mathbf{I} + \Sigma^{\ell}\mathbf{W})$. Then, by following the proof of [64, Theorem IV.1], we have that this dynamics has a unique equilibrium for any values of $\mathbf{d}_1, \mathbf{d}_2$ if and only if $\mathbf{I} - \mathbf{W} \in \mathcal{P}$. Now, to construct the equilibrium trajectory we can denote an equilibrium point corresponding to $\mathbf{d}_1, \mathbf{d}_2$ by $\mathbf{e}^*(\mathbf{d}_1, \mathbf{d}_2)$. Substituting in the expressions for $\mathbf{d}_1, \mathbf{d}_2$ gives $\mathbf{e}^*(\mathbf{W}\mathbf{r}(t) + \mathbf{c}(t), \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t))$ as the equilibrium point for a time t. Allowing t to vary in $\mathbb{R}_{\geq 0}$ we have the trajectory $\{\mathbf{e}^*(\mathbf{W}\mathbf{r}(t) + \mathbf{c}(t), \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t))\}_{t \in \mathbb{R}_{\geq 0}}$, composed of the equilibrium points at each time. □

The equilibrium trajectory corresponds to the reference signal $\mathbf{r}$ being tracked, possibly with a steady-state error. We next provide conditions to ensure this steady-state error is zero. We

note that to do so, $\mathbf{r}(t)$ must lie in $[\mathbf{0}, \mathbf{m}]$ for all t due to the bounding of the dynamics. The next result identifies an additional condition.

Theorem 6.4.2. (Reference Tracking for LTN Dynamics): *Consider the LTN dynamics* (6.1) *and let* $\mathbf{r} : \mathbb{R} \to [\mathbf{0}, \mathbf{m}]$ *be a desired reference trajectory that satisfies* $\mathbf{r}(t) + \tau\dot{\mathbf{r}}(t) \in [\mathbf{0}, \mathbf{m}]$ *for all* $t \in \mathbb{R}$. *Then, if* $\mathbf{W} \in \mathcal{L}$, *under the input* $\mathbf{c}_{\text{ol}}(t) = -\mathbf{W}\mathbf{r}(t) + \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t)$, *the dynamics is GES to the reference trajectory* $\mathbf{r}$.

Proof of Theorem 6.4.2. We prove the result by showing that with the input $\mathbf{c}_{\text{ol}}(t) = -\mathbf{W}\mathbf{r}(t) + \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t)$ the error dynamics (6.4) is GES to the origin. With this control, the dynamics (6.4) becomes

$$\tau\dot{\mathbf{e}}(t) = -\mathbf{e}(t) + [\mathbf{W}\mathbf{e}(t) + \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t)]_{\mathbf{0}}^{\mathbf{m}} - \mathbf{r}(t) - \tau\dot{\mathbf{r}}(t)$$

Since $\mathbf{r} + \tau\dot{\mathbf{r}} \in [\mathbf{0}, \mathbf{m}]$, it is immediate that $\mathbf{e}(t) = \mathbf{0}$ is an equilibrium for these dynamics for all t. Then, since $\mathbf{W} \in \mathcal{L}$ implies $\mathbf{I} - \mathbf{W} \in \mathcal{P}$, by Theorem 6.4.1, this is the unique equilibrium trajectory. To conclude the proof, we show that, under the assumption that $\mathbf{W} \in \mathcal{L}$, the origin is globally asymptotically stable under this dynamics, following a similar argument to [64, Theorem IV.8].

Let $\mathbf{W}_i$ denote the i'th row of $\mathbf{W}$. After some manipulations, we can rewrite the error dynamics as

$$\tau\dot{\mathbf{e}}(t) = (-\mathbf{I} + \mathbf{M}(t)\mathbf{W})\mathbf{e}(t),$$

where $\mathbf{M}(t)$ is a diagonal matrix given by

$$\mathbf{M}_{ii}(t) = \begin{cases} \frac{[\mathbf{W}_i\mathbf{e} + (\mathbf{r} + \tau\dot{\mathbf{r}})_i]_{\mathbf{0}}^{\mathbf{m}_i} - [(\mathbf{r} + \tau\dot{\mathbf{r}})_i]_{\mathbf{0}}^{\mathbf{m}_i}}{\mathbf{W}_i\mathbf{e}} & \text{if } \mathbf{W}_i\mathbf{e} \neq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Since the threshold operator is monotonically increasing and Lipschitz with constant one, we

have $\mathbf{M}_{ii}(t) \in [0,1]$. Therefore, $\mathbf{M}(t)$ lies in the convex hull of $\{\Sigma^\ell(\sigma)\}_{\sigma\in\{0,\ell\}^n}$ for all t. Thus, there exists a time-varying convex combination $(\alpha_\sigma(t))_{\sigma\in\{0,\ell\}^n}$ such that

$$\mathbf{M}(t) = \sum_{\sigma\in\{0,\ell\}^n} \alpha_\sigma(t)\Sigma^\ell, \quad t \geq 0.$$

Now, since $\mathbf{W} \in \mathcal{L}$, there exists a matrix $\mathbf{P} > 0$ and a scalar $\lambda > 0$ such that

$$(-\mathbf{I} + \mathbf{W}^\top\Sigma^\ell)\mathbf{P} + \mathbf{P}(-\mathbf{I} + \Sigma^\ell\mathbf{W}) \leq -\lambda\mathbf{I}, \quad \forall\sigma \in \{0,\ell\}^n.$$

Consider then the candidate Lyapunov function $V(\mathbf{e}) = \mathbf{e}^\top\mathbf{P}\mathbf{e}$. Its derivative along the dynamics satisfies

$$\begin{aligned}
\tau\dot{V}(\mathbf{e}(t)) &= \mathbf{e}^\top[(-\mathbf{I} + \mathbf{W}^\top\mathbf{M}(t))\mathbf{P} + \mathbf{P}(-\mathbf{I} + \mathbf{M}(t)\mathbf{W})]\mathbf{e} \\
&= \mathbf{e}^\top\Big[\sum_{\sigma\in\{0,\ell\}^n} \alpha_\sigma(t)[(-\mathbf{I} + \mathbf{W}^\top\Sigma^\ell)\mathbf{P} + \mathbf{P}(-\mathbf{I} + \Sigma^\ell\mathbf{W})]\Big]\mathbf{e} \\
&\leq -\lambda\|\mathbf{e}\|^2 \leq -\frac{\lambda}{\rho(\mathbf{P})}V(\mathbf{e}(t)),
\end{aligned}$$

which ensures the origin is GES under (6.4). □

Remark 6.4.3. (Lack of Uniqueness of Input Map for Reference Tracking): *We note that, while the input* $\mathbf{c}_{\text{ol}}(t)$ *in Theorem 6.4.2 is sufficient for ensuring GES of the origin for the error dynamics, and hence reference tracking, it is not unique. For example, due to the thresholding of the LTN dynamics, any (constant) reference trajectory that lies on the lower or upper threshold can be converged to using a higher magnitude control (either positive or negative, depending on the network and which threshold) than the one defined in* $\mathbf{c}_{\text{ol}}(t)$ *in order to drive the signal to the boundary.* •

The open-loop control in Theorem 6.4.2 requires the network to satisfy $\mathbf{W} \in \mathcal{L}$ to achieve reference tracking. However, this condition is not always met, as there exist many brain networks

that are not stable on their own, see e.g., [64, 173]. Our next result shows that tracking can still be achieved with unstable synaptic weight matrices using feedback control of the form $\mathbf{c}(t) = \mathbf{K}\mathbf{x}(t) + \mathbf{d}(t)$, where $\mathbf{K} \in \mathbb{R}^{n\times n}$ is the feedback gain. Note that $\mathbf{K}$ may have rows entirely composed of zeros, corresponding to nodes in the network that cannot be directly impacted by feedback. With this control, the effective network is described by $\mathbf{W} + \mathbf{K}$, and the next result shows that as long as this matrix belongs to $\mathcal{L}$, reference tracking for (6.1) can be achieved.

Theorem 6.4.4. (Reference Tracking for LTN Dynamics with Closed-Loop Control): *Consider the LTN dynamics* (6.1) *and let* $\mathbf{r} : \mathbb{R}_{>0} \to [\mathbf{0}, \mathbf{m}]$ *be a desired reference trajectory that satisfies* $\mathbf{r}(t) + \tau\dot{\mathbf{r}}(t) \in [\mathbf{0}, \mathbf{m}]$ *for all* $t \in \mathbb{R}_{>0}$. *Let* $\mathbf{K} \in \mathbb{R}^{n\times n}$ *such that* $\mathbf{W} + \mathbf{K} \in \mathcal{L}$. *Then, under the input* $\mathbf{c}_{\text{cl}}(t) = \mathbf{K}\mathbf{x}(t) - (\mathbf{W} + \mathbf{K})\mathbf{r}(t) + \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t)$, *the dynamics is GES to the reference signal* $\mathbf{r}$.

Proof of Theorem 6.4.4. With a feedback-feedforward control of the form $\mathbf{c}_{\text{cl}}(t) = \mathbf{K}\mathbf{x}(t) + \mathbf{d}(t)$, the error dynamics (6.4) can be written as

$$\tau\dot{\mathbf{e}}(t) = -\mathbf{e}(t) + [(\mathbf{W} + \mathbf{K})\mathbf{e}(t) + (\mathbf{W} + \mathbf{K})\mathbf{r}(t) + \mathbf{d}(t)]_{\mathbf{0}}^{\mathbf{m}} - \mathbf{r}(t) - \tau\dot{\mathbf{r}}(t). \tag{6.5}$$

In turn, this corresponds to the error dynamics (6.4) with the synaptic weight matrix $\mathbf{W} + \mathbf{K} = \tilde{\mathbf{W}}$. By Theorem 6.4.2, if $\tilde{\mathbf{W}} \in \mathcal{L}$, the dynamics is GES to the origin with the input $\mathbf{d}(t) = -\tilde{\mathbf{W}}\mathbf{r}(t) + \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t)$. Therefore the LTN dynamics (6.1) achieves reference tracking with the input $\mathbf{c}_{\text{cl}}(t)$. □

Remark 6.4.5. (Reference Tracking in Under-actuated LTN Systems): *In the results above, we have assumed that the network is fully actuated for the purpose of determining a control such that reference tracking can be achieved. For systems that are not fully actuated, i.e., the control is of the form* $\mathbf{B}\mathbf{c}(t)$, *where* $\mathbf{B}$ *is a diagonal matrix with zeros indicating components of the network that cannot be actuated, it is generally not possible to achieve reference tracking in all components. If* $\mathbf{W} \in \mathcal{L}$, *then the non-actuated component still converges to an equilibrium trajectory, but this depends on the interconnection with the other network components rather*

than the desired reference signal. The only case when the under-actuated component i will converge to the reference trajectory is if $[\mathbf{W}_i\mathbf{r}(t)]_{\mathbf{0}}^{\mathbf{m}} = \mathbf{r}(t) + \tau\dot{\mathbf{r}}(t)$ *for all* $t \in \mathbb{R}_{\geq 0}$. *We note that the actuated components of the network will still converge to their desired reference under the controls of Theorems 6.4.2 or 6.4.4.* •

Having established reference tracking under the LTN dynamics with both open-loop (cf. Theorem 6.4.2) and closed-loop (cf. Theorem 6.4.4) controls, we note that the implementation of the corresponding inputs, either as inputs from other neuronal populations or through neurostimulation, is challenging, and becomes increasingly unrealistic with larger network dimensions, characteristic of brain modeling. First, evaluating the controls provided is difficult because it requires full knowledge of the synaptic weight matrix $\mathbf{W}$. Due to both the complexity of the brain and the difficulty in measuring the individual impact of individual neuron populations, particularly those deep in the brain [174], this is unrealistic. Beyond knowledge of its structure, the results require that $\mathbf{W}$ (or $\mathbf{W} + \mathbf{K}$) are in $\mathcal{L}$. This becomes computationally difficult to check as the network scales up, and the known sufficient conditions to ensure being in $\mathcal{L}$ become increasingly conservative. As such, for large networks, it is preferable to divide the network into layers based on the timescales, where each layer has an individual constant τ_i. In this case, dependent on the timescales, the results of Theorems 6.4.2 and 6.4.4 are no longer directly applicable due to the activity of the interconnection terms affecting the dynamics of each layer. In such a case, one can use a singular perturbation argument [65], but the recursively-defined equilibrium maps involved make the determination of the exact control signals particularly challenging.

Second, explicitly computing a direct control signal is not a realistic representation of the way the brain is believed to operate. However, implementing a data-driven approach, where all that is needed is the desired reference signal and the control is determined through internal dynamics, seems more realistic and not as computationally expensive. In the following section we provide an explanation of the reservoir computing and next-generation reservoir computing

frameworks. We demonstrate both can be used to achieve reference tracking in linear-threshold networks later in Sections 6.6 and 6.7.

6.5 Reservoir Computing

Here we provide an overview of the mathematical basis of the reservoir computing framework, which we later apply to control synthesis for reference tracking in LTN networks. We begin by discussing reservoir computing for predicting system outputs and for control before considering the same problems using next-generation reservoir computing. We finish by discussing parameter selection and comparing the two frameworks.

6.5.1 Reservoir Computing for Prediction

We first overview the basic structure of a reservoir computer for predicting the outputs of an unknown system. Assume we have a dynamical system defined by

$$\mathbf{y}(t) = f(\mathbf{c}(t)) \tag{6.6}$$

where $\mathbf{y} \in \mathbb{R}^m$ is the system output, $\mathbf{c} \in C \subseteq \mathbb{R}^n$ is the input and the driving function $f : \mathbb{R}^n \to \mathbb{R}^m$ for the system is unknown. We define a reservoir as a dynamical system

$$\dot{\mathbf{x}}(t) = F(\mathbf{J}\mathbf{x}(t) + \mathbf{J}_{\text{in}}\mathbf{c}(t)), \tag{6.7}$$

where $\mathbf{x} \in X \subseteq \mathbb{R}^N$ is the internal state of the reservoir, $\mathbf{J} \in \mathbb{R}^{N\times N}$ is a matrix that provides the reservoir structure, and $\mathbf{J}_{\text{in}} \in \mathbb{R}^{N\times n}$ in an input matrix. We assume that $N \gg n$ and $F : X \times C \to X$ is a nonlinear activation function. The goal of the reservoir computing framework is to use the activity of the high-dimensional reservoir dynamical system (6.7) to estimate the outputs of the unknown system (6.6).

The output estimate is defined

$$\tilde{\mathbf{y}}(t) = \mathbf{J}_{\text{out}}\mathbf{x}(t),$$

where $\mathbf{J}_{\text{out}} \in \mathbb{R}^{m\times N}$ is an output vector trained to achieve an accurate estimate of the unknown system. The key component for the reservoir computing framework to be able to achieve this goal is the *echo state property* [163]. This property is dependent on the activation function F and the reservoir matrix $\mathbf{J}$. For the activation function, we require [175] that $F : X\times C \to X$ is defined on compact sets X and C. The compactness of the state space X is given for the most commonly used activation functions in machine learning, such as tanh or the logistic sigmoid function, and importantly, also for the linear-threshold function $[\cdot]_{\mathbf{0}}^{\mathbf{m}}$. We assume the compactness of the input set C, which is realistic in most applications.

To define the echo state property, we introduce the following notations. We will denote the reservoir dynamical system (6.7), which combines the activation function, F, and the reservoir matrix, $\mathbf{J}$, by $\mathbf{F}(\mathbf{x}(t), \mathbf{c}(t)) = F(\mathbf{J}\mathbf{x}(t) + \mathbf{J}_{\text{in}}\mathbf{c}(t))$. Let $X^{+\infty} \coloneqq \{\mathbf{x}^{+\infty} = \{\mathbf{x}(t)\}_{t=0}^{\infty} \mid \mathbf{x}(t) \in X,\ \forall t \geq 0\}$ and $C^{+\infty} \coloneqq \{\mathbf{c}^{+\infty} = \{\mathbf{c}(t)\}_{t=0}^{\infty} \mid \mathbf{c}(t) \in C,\ \forall t \geq 0\}$ denote sets of right infinite state and input sequences. A right infinite state sequence $\mathbf{x}^{+\infty}$ is *compatible* with input state sequence $\mathbf{c}^{+\infty}$ when $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x}(t), \mathbf{c}(t))$ for all $t \geq 0$.

Definition 6.5.1. (Echo State Property [175]): *A reservoir* $\mathbf{F} : X\times C \to X$ *defined on compact sets* X *and* C *satisfies the echo state property with respect to* C *if and only if for any right infinite input sequence* $\mathbf{c}^{+\infty} \in C^{+\infty}$ *and any two right infinite state vector sequences* $\mathbf{x}_1^{+\infty}, \mathbf{x}_2^{+\infty} \in X^{+\infty}$ *compatible with* $\mathbf{c}^{+\infty}$*, there exists a sequence* $\{\delta_t\}_{t=0}^{\infty}$ *such that* $\|\mathbf{x}_1(t) - \mathbf{x}_2(t)\| \leq \delta_t$*, where* $\lim_{t\to\infty} \delta_t = 0$.

The echo state property relates to the asymptotic convergence of the state of the reservoir, which is influenced by a driving input. It can be thought of as the concept of fading memory, in that trajectories of the reservoir should converge to the same point given the same input, regardless of the prior history of the reservoir. Necessary and sufficient conditions for the echo

state property to hold are dependent on both the reservoir matrix, $\mathbf{J}$, and the activation function, F. Regardless of the activation function, guaranteeing the echo state property depends on the stability of the reservoir matrix[1].

What remains is to train the reservoir computing framework to get an accurate estimate $\tilde{\mathbf{y}}$ of the output $\mathbf{y}$. The defining feature of reservoir computing is that we only train the output vector $\mathbf{J}_{\text{out}}$, rather than the internal weights of the reservoir. In order to train the reservoir output matrix, we take a training input $\{\mathbf{c}_T(t)\}$ and drive the system (6.6) using this input, getting a driven training output timeseries $\{\mathbf{y}_T(t)\}$. Then, using the training signal $\{\mathbf{c}_T(t)\}$, we drive the reservoir (6.7) to get a timeseries of driven reservoir states $\{\mathbf{x}_T(t)\}$. Compiling the set of training outputs and driven reservoir states into matrices $\mathbf{Y}_T$ and $\mathbf{X}_T$, we compute the output matrix $\mathbf{J}_{\text{out}}$ using a Tikhonov regularization,

$$\text{argmin}_{\mathbf{J}_{\text{out}}} ||\mathbf{Y}_T - \mathbf{J}_{\text{out}}\mathbf{X}_T||^2 + ||\beta\mathbf{J}_{\text{out}}||^2, \tag{6.8}$$

where $\beta > 0$ is a regularization parameter. The performance of the reservoir computer for predicting outputs can then be measured by using the trained system to provide a prediction of a second set of arbitrary inputs and comparing the accuracy with the true outputs for this set. The main parameter for how well the reservoir computer works is the size of the reservoir, N. As N increases, the higher-dimensional reservoir can then exhibit an increasingly large number of possible behaviors, especially relatively to the size of the system being predicted. Then, in the process of training the system, we are able to relate these behaviors to the system trajectories through the linear output operator $\mathbf{J}_{\text{out}}$. Based on this, with a randomly defined reservoir, by increasing N sufficiently, the reservoir computer can achieve good prediction of the system outputs. Parameters such as the reservoir weight matrix $\mathbf{J}$, the input matrix $\mathbf{J}_{\text{in}}$, the training input, and the regularization parameter β can also impact the quality of the output prediction.

[1]Much of the reservoir computing literature is in discrete time, which also impacts sufficient conditions for the echo state property. In the discrete-time case with $F(\cdot) = \tanh(\cdot)$, a sufficient condition for the echo state property is that $\mathbf{J}$ is diagonally Schur stable.

6.5.2 Reservoir Computing as a Controller

We are interested in applying a reservoir computer as a controller for a dynamical system, with the goal of controlling the system to a given reference trajectory [168]. Consider a dynamical system

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \mathbf{c}(t)) \qquad \mathbf{y}(t) = g(\mathbf{x}(t)), \tag{6.9}$$

where $\mathbf{x}(t), \mathbf{c}(t) \in \mathbb{R}^n$ are the state variables and inputs, respectively, and $\mathbf{y}(t) \in \mathbb{R}^m$ is the output of the system. The functions f and g define the system evolution and measurable outputs, and potentially are unknown. Let $\mathbf{r}(t) \in \mathbb{R}^m$ denote a desired reference trajectory for the output. The reservoir dynamics is defined as

$$\dot{\mathbf{u}}(t) = F(\mathbf{J}\mathbf{u}(t) + \mathbf{J}_{\text{in}}\mathbf{y}(t) + \mathbf{J}_{\text{ref}}\mathbf{r}(t+\delta)), \tag{6.10}$$

where $\mathbf{u}(t) \in \mathbb{R}^N$ is the reservoir state, $\mathbf{J} \in \mathbb{R}^{N\times N}$ defines the internal reservoir connections, $\mathbf{J}_{\text{in}} \in \mathbb{R}^{N\times m}$ scales the system output $\mathbf{y}(t)$ into the reservoir, $\mathbf{J}_{\text{ref}} \in \mathbb{R}^{N\times m}$ scales the reference signal, and $\delta > 0$ dictates how far ahead we provide a desired reference value. The function F is the activation function, and we assume that F and $\mathbf{J}$ are chosen such that the reservoir has the echo state property. We then connect the output of the reservoir dynamics (6.10) with the input of the system (6.9) by defining

$$\mathbf{c}(t) = \mathbf{J}_{\text{out}}\mathbf{u}(t).$$

Figure 6.1 illustrates this setup.

To train the reservoir computer, we use an open-loop version of the schematic shown in Figure 6.1 with a training input and a delay on the output. Here, instead of inputting the reference trajectory into the reservoir, we use the current system output, $\mathbf{y}(t)$, along with the future output,

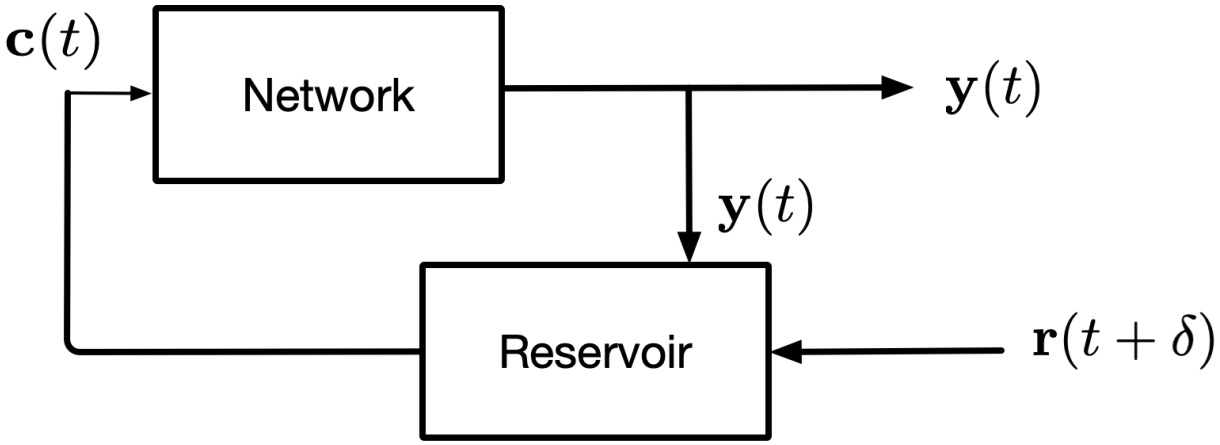


Figure 6.1. A reservoir computer steers the dynamical system (6.9) to a reference signal $\mathbf{r}(t)$.

$\mathbf{y}(t+\delta)$, which is directly computable from the system and training input. The future output acts in the role of the reference signal, with the goal of determining how the training signal moves the system from $\mathbf{y}(t)$ to $\mathbf{y}(t+\delta)$. The reservoir training dynamics is then

$$\dot{\mathbf{u}}(t) = F(\mathbf{J}\mathbf{u}(t) + \mathbf{J}_{\text{in}}\mathbf{y}(t) + \mathbf{J}_{\text{ref}}\mathbf{y}(t+\delta)).$$

Figure 6.2 illustrates this training setup. The output vector for the reservoir computer, $\mathbf{J}_{\text{out}}$, is then

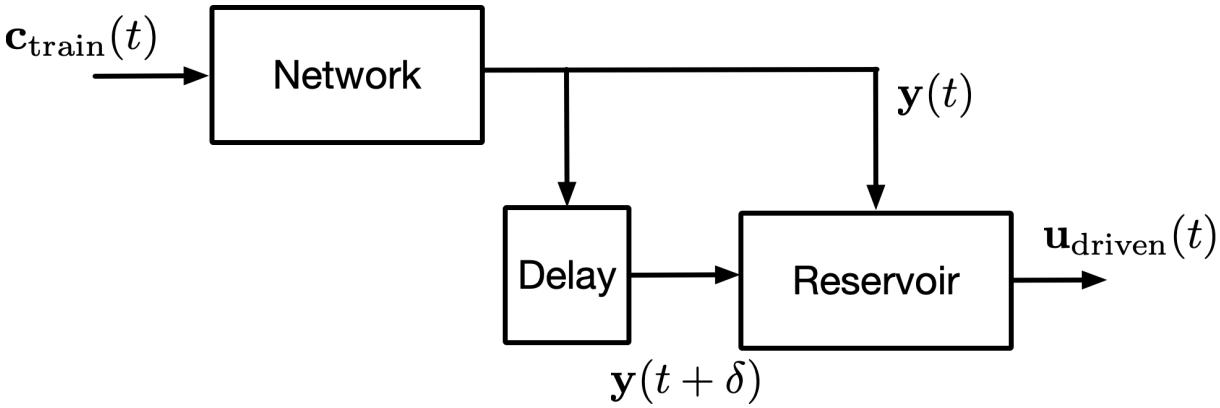


Figure 6.2. Schematic for training a reservoir computer to steer the dynamical system (6.9) to a desired reference trajectory.

trained using the Tikhonov regularization (6.8) to minimize the difference between $\mathbf{c}_{\text{train}}$ and the reservoir prediction, $\mathbf{J}_{\text{out}}\mathbf{u}_{\text{driven}}$. This setup can be extended by providing different information to

the reservoir, either about the reference signal $\mathbf{r}(t)$, or the system output, $\mathbf{y}(t)$. A common choice is providing the reservoir with $\dot{\mathbf{y}}(t)$. However, this is typically discretely estimated from $\mathbf{y}(t)$ and as such it is just a transform from providing $\mathbf{y}(t)$ [168]. Instead, in our treatment, we provide additional information regarding the reference signal, in particular $\dot{\mathbf{r}}(t)$, which we can compute since it is known in advance. This leads to the final schematic for the reservoir controller in Figure 6.3.

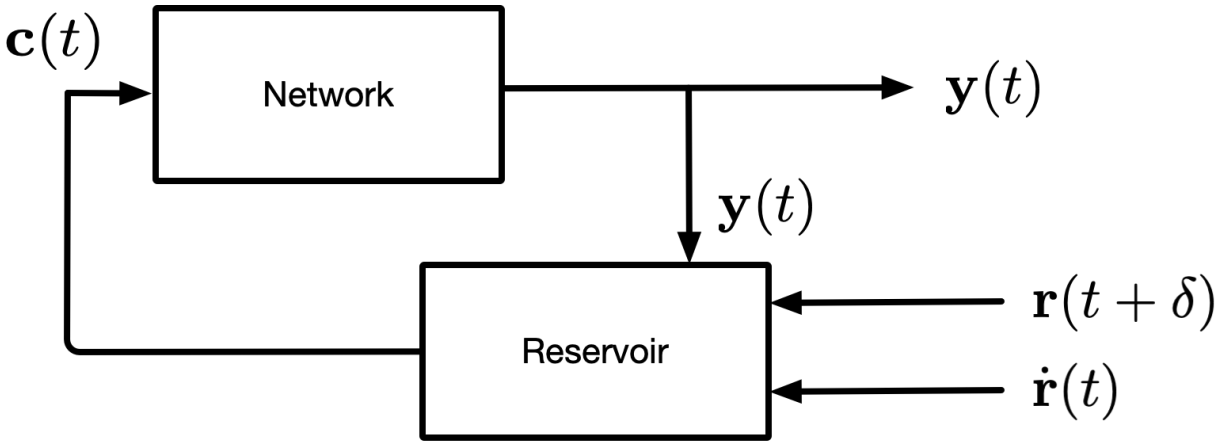


Figure 6.3. Schematic for using a reservoir computer as a controller incorporating additional information on the reference signal $\mathbf{r}(t)$.

6.5.3 Next-Generation Reservoir Computing

The observation, cf. [176], that the underlying equations of the reservoir computing framework have similarities with nonlinear vector autoregression (NVAR) and dynamic mode decomposition led to the construction of the next-generation reservoir computing (NG-RC) framework [169] based on NVAR. We explain this framework below, both for prediction and control.

NG-RC for Prediction

Similarly to the classical RC, we consider the dynamical system (6.6). Instead of defining a reservoir, we define two feature vectors, $O_{\text{lin},t}$ and $O_{\text{nl},t}$, defined as follows: $O_{\text{lin},t} = \mathbf{c}(t) \oplus \mathbf{c}(t-i_1) \oplus \cdots \oplus \mathbf{c}(t-i_p)$ is the linear feature vector defined for p discretely sampled prior

points of the system inputs. $O_{\text{nl},t}$ is a nonlinear feature vector that is an arbitrary nonlinear function of the vector $O_{\text{lin},t}$. Any nonlinear function can be chosen, albeit it is common [169] to employ the monomials up to some order k. These feature vectors are then used to predict the outputs of the unknown system (6.6) in place of the activity of the reservoir (6.9).

The two feature vectors are then concatenated, commonly with an additional constant $d \in \mathbb{R}$, to give $O_{\text{total},i} = d \oplus O_{\text{lin},i} \oplus O_{\text{nl},i}$ and the predicted output is defined by $\tilde{\mathbf{y}}(t) = \mathbf{J}_{\text{out}} O_{\text{total},t}$. The output vector $\mathbf{J}_{\text{out}}$ is then determined with a Tikhonov regularization as

$$\text{argmin}_{\mathbf{J}_{\text{out}}} \|\mathbf{y}(t) - \mathbf{J}_{\text{out}} O_{\text{total},t}\|^2 + \|\beta \mathbf{J}_{\text{out}}\|^2, \tag{6.11}$$

which requires running the system with a minimum of $i_p + 1$ inputs in order to fully determine the vector $O_{\text{total},t}$. This setup, with one prior timestep and quadratic monomials, is shown in Figure 6.4.

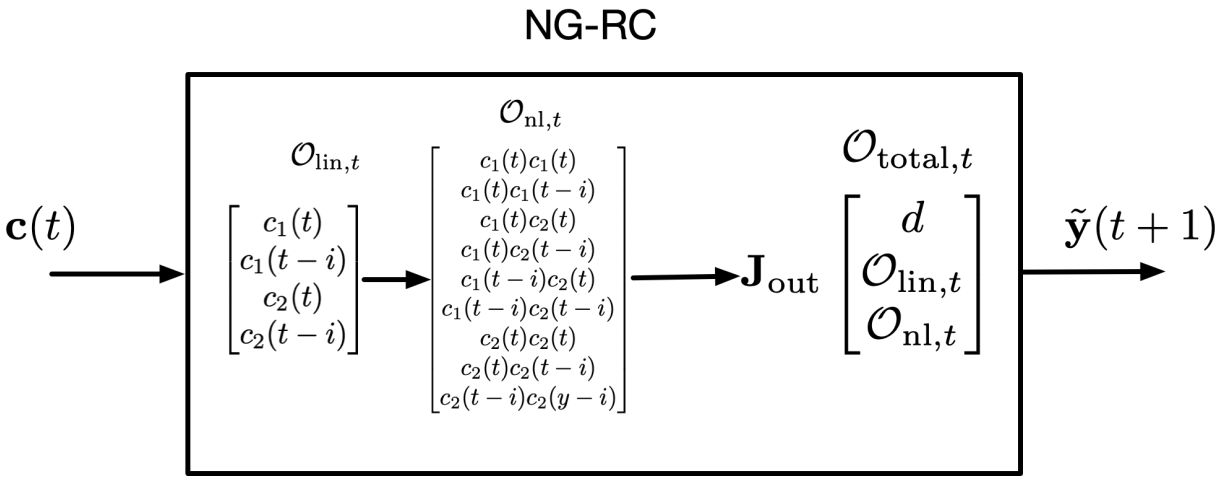


Figure 6.4. Schematic for using next-generation reservoir computing to predict system outputs with quadratic monomials used for the nonlinearity. The matrix $\mathbf{J}_{\text{out}}$ is trained with a known input/output sequence $\{(\mathbf{c}(t), \mathbf{y}(t))\}$ with a minimum length of two points, before being able to predict future outputs based on the input.

NG-RC for Control

Similarly to the RC, we are interested in applying the NG-RC as a controller to bring a dynamical system to a desired reference trajectory. To do so, we use an open-loop training period followed by a closed-loop control period. To construct the NG-RC controller we follow the approach in [170] and consider an extension to employ deep learning which we leverage in our simulations.

We consider the control system given by (6.9). We define linear and nonlinear feature vectors $O_{\text{lin},t}$ and $O_{\text{nl},i}$. Unlike NG-RC for prediction, $O_{\text{lin},t}$ is not dependent on prior time steps and instead $O_{\text{lin},t} = \mathbf{y}(t)$, the observable outputs. The nonlinear feature vector is then a function of the outputs, where again we typically use polynomial expressions.

What differs between the prediction and control setups is the definition of the total feature vector, $O_{\text{total},t}$. Rather than just combining the linear and nonlinear feature vectors with a constant, we also include the system input, giving $O_{\text{total},t} = \mathbf{c}(t) \oplus d \oplus O_{\text{lin},t} \oplus O_{\text{nl},t}$. With this feature vector, Tikhonov regularization (6.11) is used to learn the output vector $\mathbf{J}_{\text{out}}$ in an open-loop fashion such that the prediction is given by $\hat{\mathbf{y}}(t) = \mathbf{J}_{\text{out}} O_{\text{total},t}$, as in the top schematic of Figure 6.5.

Then, for the purpose of defining a control, we split the output and feature vector in the following way.

$$\mathbf{J}_{\text{out}} O_{\text{total},t} = \mathbf{J}_{\text{out}}^{X} \begin{bmatrix} d \\ O_{\text{lin},t} \\ O_{\text{nl},t} \end{bmatrix} + \mathbf{J}_{\text{out}}^{\mathbf{c}} \mathbf{c}(t) = \mathbf{J}_{\text{out}}^{X} O_{X,t} + \mathbf{J}_{\text{out}}^{\mathbf{c}} \mathbf{c}(t).$$

From here, a closed-loop system is created where the matrices $\mathbf{J}_{\text{out}}^{X}$ and $\mathbf{J}_{\text{out}}^{\mathbf{c}}$, along with the reference signal $\mathbf{r}(t)$, are used to define the control term. With the error term $\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{r}(t)$, the control is defined by

$$\mathbf{c}(t) = (\mathbf{J}_{\text{out}}^{\mathbf{c}})^{-1} [\mathbf{r}(t+\delta) - \mathbf{J}_{\text{out}}^{X} O_{X,t} + \mathbf{K}\mathbf{e}(t)], \tag{6.12}$$

where $\mathbf{K}$ is a proportional control matrix determined by trial-and-error to optimize performance. The closed-loop control period is shown in the bottom schematic of Figure 6.5. This control is derived in [170] for a discrete-time system with the control input entering linearly, i.e., of the form $\mathbf{y}(t+1) = \mathbf{F}(X(t)) + \mathbf{B}\mathbf{c}(t)$. While the LTN dynamics is not of this form, our simulations in Sections 6.6 and 6.7 show that this control still achieves satisfactory performance.

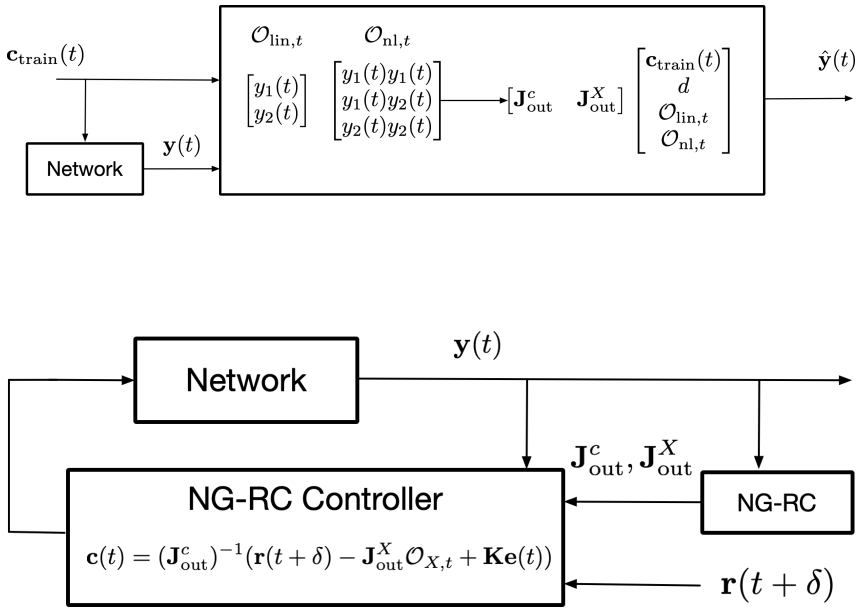


Figure 6.5. Schematics of the training and control phases of a next-generation reservoir computer controller. The top schematic shows the open-loop learning component of using the NG-RC as a controller. The closed-loop control component is shown in the second schematic, with the control given as in (6.12).

We can also extend this approach to construct a controller out of multiple NG-RCs rather than just one. In this approach each successive NG-RC is used to minimize the error remaining from the control determined by the prior NG-RC. The additional NG-RC layers are added by training output vectors $\mathbf{J}_{\text{out},i}$ for each layer with the feature vector $O_{\text{total},t}$ such that the output prediction is

$$\tilde{\mathbf{y}}(t) = \sum_{i=1}^{M} \mathbf{J}_{\text{out},i} O_{\text{total},t},$$

where M is the number of layers. The layers are trained successively, with each layer being

dependent on the output of the layers before. Each successive layer is trained using Tikhonov regularization on the error between the true output and the predicted output by the sum of the layers before. Formally, if we denote $\tilde{\mathbf{y}}_{1:i}(t)$ to be the predicted output from the first i layers, the output vector for layer $i+1$ is determined by the following

$$\mathbf{J}_{\text{out}}^{i+1} = \text{argmin} \|(\tilde{\mathbf{y}}_{1:i}(t) - \mathbf{y}(t)) - \mathbf{J}_{\text{out}}^{i+1} O_{\text{total},t}\|^2 + \|\beta \mathbf{J}_{\text{out}}^{i+1}\|^2.$$

The control signal is then given by

$$\mathbf{c}(t) = (\sum_{i=1}^{m} \mathbf{J}_{\text{out},i}^{c})^{-1} (\mathbf{r}(t+\delta) + \sum_{j=1}^{m} (\mathbf{J}_{\text{out},j}^{X} O_{X,t}) + \mathbf{K}\mathbf{e}(t)).$$

e note that this process could be modified to use different feature vectors across the multiple NG-RC's. The process for computing the output vectors would be the same, with some differences appearing in the determination of the final control signal.

Remark 6.5.2. (Comparison of Parameter Selection in RC and NG-RC): *The selection of parameters is of paramount importance to the performance of the RC and NG-RC algorithms. One of the benefits of RC compared to traditional machine learning algorithms is that there are significantly fewer parameters to optimize [171]. This is even more so for NG-RC, cf. [169]. In comparing RC and NG-RC, one of the biggest differences is in the number of parameters that need to be selected, and the difficulty in their selection. In RC, the parameters to be selected are the reservoir matrix,* $\mathbf{J}$*, the activation function,* $\mathbf{F}$*, the input vector* $\mathbf{J}_{in}$*, the regularization parameter,* β *and the training signal,* $\mathbf{c}_{train}$*. Of these, the most important is the selection of the reservoir matrix, which is typically done randomly, and its relation with the activation function to guarantee the echo state property. Despite significant research into choices of the reservoir [171, 177, 178, 172] an optimal choice is not known. In NG-RC for control, the parameters to be tuned are the nonlinear feature vector,* $O_{nl,t}$*, the proportional control vector* $\mathbf{K}$*, and the training signal,* $\mathbf{c}_{train}$*. In NG-RC for prediction, one needs to additionally choose the*

linear feature vector, $O_{lin,t}$. •

Remark 6.5.3. (System Output for Reference Tracking): *We note that in in this section we have defined the reservoir computing framework for an arbitrary system output* $\mathbf{y}(t)$*, while in Problem 6.3.1 we aim for reference tracking of the system state. As such, in the following applications we consider the specific case of* $\mathbf{y}(t) = \mathbf{x}(t)$*, which matches the problem description.*

•

6.6 Application to Selective Inhibition and Recruitment

In this section, we provide our first illustration of the use of RC and NG-RC controllers to achieve reference tracking for linear-threshold networks. We consider the problem of selective inhibition and recruitment in cortical networks, and illustrate the recruitment of a subset of a network to a chosen reference signal. We do this with both the RC and NG-RC controllers and include a comparison of their performance. We first give a brief review of the selective inhibition and recruitment framework from Chapter 5.

Selective inhibition and recruitment is the problem of reacting to stimuli by inhibiting task-irrelevant neuron populations to zero, while recruiting the remaining task-relevant neuron populations to a particular activity pattern. Due to the hierarchical nature of the brain [179], this problem has been studied in networks composed of subnetworks operating at different timescales, and leads to the following formalization of the network structure. Consider a network composed of N subnetworks, each of composed of n_i nodes and with corresponding timescale τ_i. We construct a hierarchy by organizing the subnetworks such that $\tau_N \leq \tau_{N-1} \leq \cdots \leq \tau_1$ and each subnetwork is connected only to the subnetworks directly above and below it in the hierarchy, cf. Figure 6.6. Subnetworks at the bottom of the network, with fast timescales, represent regions in the brain that operate quickly, such as sensory areas, while the top of the network represents regions such as the prefrontal cortex, which operate relatively slowly.

The problem of selective inhibition and recruitment is then formalized as follows. For

each subnetwork let $\mathbf{x}_i^1 \in \mathbb{R}^{r_i}$ denote the set of task-relevant nodes and $\mathbf{x}_i^0 \in \mathbb{R}^{n_i - r_i}$ denote the set of task-irrelevant nodes. Then, determine a control $\mathbf{c}_i^* : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_i}$ such that the task-relevant nodes are recruited to a non-zero equilibrium trajectory $\mathbf{r}_i^* : \mathbb{R}_{\geq 0} \to \mathbb{R}^{n_i}$, i.e., $\mathbf{x}_i^1 = \mathbf{r}_i^*(t)$ and the task-irrelevant nodes are inhibited to zero, i.e., $\mathbf{x}_i^0(t) = \mathbf{0}$.

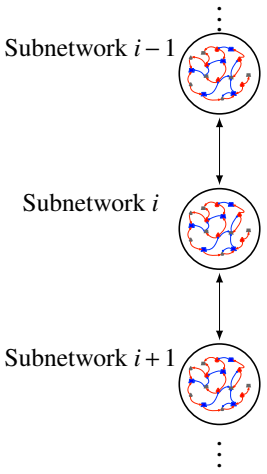


Figure 6.6. A hierarchical cortical brain network as considered for selective inhibition and recruitment. Nodes are divided between excitatory (red), inhibitory (blue), and those to be inhibited (gray). Edge colors match the direction of the node from which they originate, indicating that they provide either an excitatory (red) or inhibitory (blue) connection, while gray labels indicate that their node of origin has been inhibited and the connection provides no activity.

This problem has been addressed in [65, 2] for constant, recursively-defined equilibrium trajectories $\mathbf{r}^*(t) = \mathbf{r}$. However, this involves the explicit computation of a control with complexity that increases significantly with scale based on the size and hierarchical nature of the network. Here, instead, we show how the problem can be solved in a data-driven way using the RC and NG-RC frameworks. Once the readout has been learned, the reservoir computer determines a control that produces the desired reference tracking behavior.

6.6.1 Setup

We consider a network composed of three subnetworks, each one being an excitatory-inhibitory pair. We then aim to recruit one node in each network, while inhibiting the other to zero using reservoir controllers. In this example, the controllers are representing the impact of other

neuronal populations, outside the explicitly modeled ones, that impact behavior. The subnetworks considered are defined by the following randomly generated synaptic weight matrices:

$$\mathbf{W}_1 = \begin{bmatrix} 0.0112 & -0.9903 \\ 0.4101 & -0.5115 \end{bmatrix}, \quad \mathbf{W}_2 = \begin{bmatrix} 0.4614 & -0.7342 \\ 0.0950 & -0.5115 \end{bmatrix},$$
$$\mathbf{W}_3 = \begin{bmatrix} 0.1136 & -0.2110 \\ 0.7732 & -0.0800 \end{bmatrix}.$$

These networks are combined with the timescales $\tau_1 = 4$, $\tau_2 = 1$ and $\tau_3 = 1/3$ to create a hierarchy. The interconnections between the networks are also randomly generated, cf. Section 6.6.3.

When tuning the RC and NG-RC controllers, as the size of the network increases, the number of parameters and difficulty in tuning them all concurrently increases drastically. As such, instead of training an RC or NG-RC to control the entire network at once, we begin by training each subnetwork individually. Following the determination of a controller for each individual layer, we train a second RC on the error remaining from using the first controller on the interconnected network.

We use a randomly generated 100-node reservoir, and provide inputs related to both the reference signal and its derivative. For the NG-RC, we use a nonlinear feature vector composed of the unique quadratic monomials, and the constant term in the feature vector $O_{\text{total},t}$ is equal to 0.5. The training signals for the networks are created by sampling a $\mathcal{N}(0, 0.1)$ distribution.

In the following sections we illustrate that the with this reservoir and NG-RC the proposed controllers provide successful convergence to provided reference signals. The plots in Figures 6.7-6.11 show this visually, with details and error values in the text.

6.6.2 Individual Layers

We begin by training each excitatory-inhibitory pair individually to track a given reference signal without the interconnections in the network. While training both RC and NG-RC, we

tune the regularization parameter β (RC and NG-RC) and feedback vector $\mathbf{K}$ (NG-RC) for each layer. Figures 6.7, 6.8, and 6.9 show plots of both the RC and NG-RC converging to the desired reference signal. In each plot, the system evolves without control until $t = 25$ when the controller is turned on. Control parameters are provided in Table 6.1, along with the root-mean-square error (RMSE) between the reference and actual signal, calculated from the point the controller is turned on. Further parameters related to the training and control of the systems are discussed in Section 6.6.4.

For the top layer, we recruit the excitatory node to the reference signal $\mathbf{r}^*(t) = \sin(\frac{\pi t}{100}) + 2$, while the inhibitory node is inhibited to zero, cf. Figure 6.7. Here we see that both the NG-RC and RC controllers track the reference signal successfully, though the RC works slightly better in terms of tracking, with a RMSE of 0.0293 versus a RMSE of 0.0401 for the NG-RC.

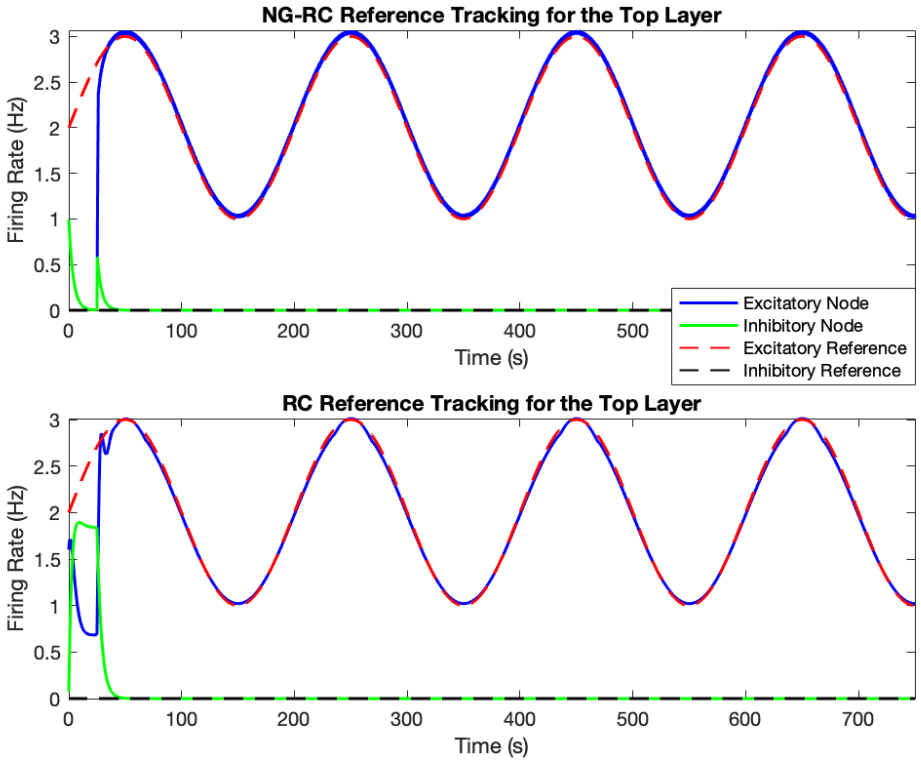


Figure 6.7. Selective inhibition and recruitment without considering interconnections between the layers using NG-RC and RC controllers for the subnetwork in the top layer. Control parameters and the RMSE between the system and reference signal are given in Table 6.1.

For the middle layer, we again recruit the excitatory node, this time to the reference

Table 6.1. The parameters and errors for the NG-RC (left) and RC (right) controllers for each network layer. The NG-RC parameters are the regularization parameter β and the proportional control $\mathbf{K}$. For the RC we consider only the regularization parameter. For both controllers, the RMSE is between the reference signal and the network state, including both the excitatory and inhibitory nodes.

NG-RC				RC		
Layer	β	**K**	**RMSE**	**Layer**	β	**RMSE**
Top	0.5	-5	0.0401	Top	0.3	0.0293
Middle	1.2	-1	0.0805	Middle	0.5	0.0255
Bottom	0.7	-0.1	0.0752	Bottom	0.1	0.0835

signal $\mathbf{r}^*(t) = \sin(\frac{2\pi t}{100}) + 2$, cf. Figure 6.8. Again both the NG-RC and RC controllers result in the tracking of the reference signal, though for this network and timescale, the NG-RC fluctuates rapidly around the desired values rather than following it exactly. This is reflected in the RMSE, where the NG-RC takes a value of 0.0805 versus a value of 0.0255 for the RC controller.

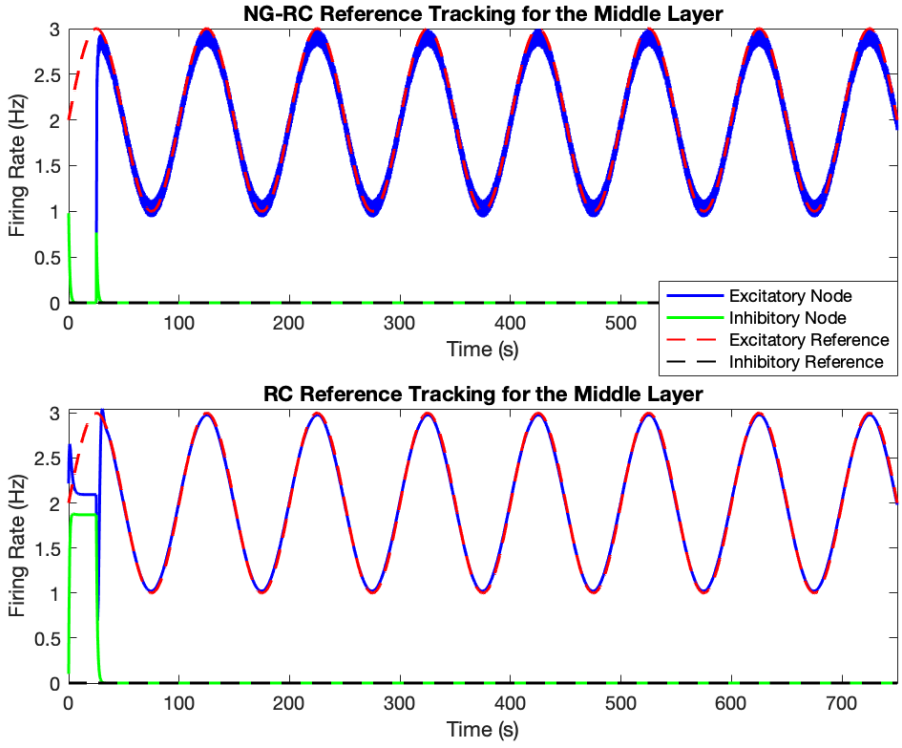


Figure 6.8. Selective inhibition and recruitment without considering interconnections between the layers using NG-RC and RC controllers for the subnetwork in the middle layer. Control parameters and the RMSE between the system and reference signal are given in Table 6.1.

For the bottom layer, we recruit the inhibitory node to the triangle wave with frequency $\frac{1}{100}$ Hz and amplitude 1, centered at 2, while inhibiting the excitatory node to zero, cf. Figure 6.9. Both the NG-RC and RC controllers result in effective recruitment to the desired signal, with the NG-RC (resp. RC) controlled system lying slightly above (resp. below) the reference signal. Here the RMSE for the NG-RC is slightly lower than for the RC, at 0.0752 versus 0.0835.

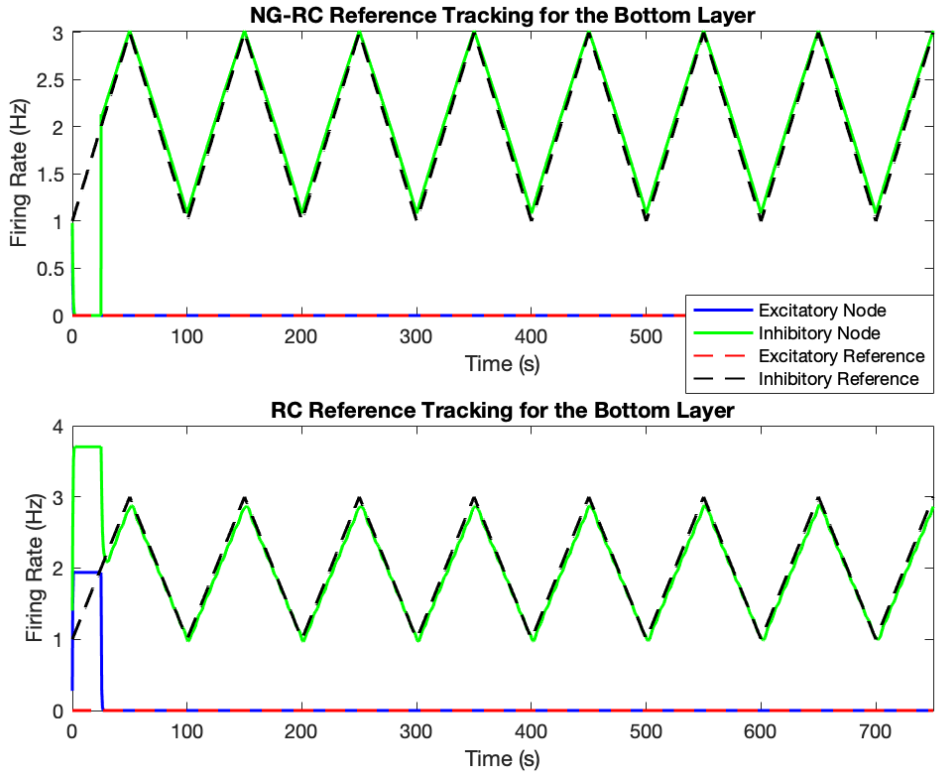


Figure 6.9. Selective inhibition and recruitment without considering interconnections between the layers using NG-RC and RC controllers for the subnetwork in the bottom layer. Control parameters and the RMSE between the system and reference signal are given in Table 6.1.

6.6.3 Interconnected Network

We now consider an interconnected system defined by the subnetworks $\mathbf{W}_1, \mathbf{W}_2$, and $\mathbf{W}_3$. Due to the difficulty in tuning the parameters as the network size increases, we do this by using a multi-layer approach, as described for the NG-RC in Section 6.5. We use the reservoir controller as determined for the individual layers before training a second layer to cover the error

introduced by the interconnection. Here we show that as the magnitude of the interconnection weight increases it becomes more difficult to control the overall network, and also illustrate that the NG-RC controller is more robust to changes in the weights of the network interconnections than the RC controller.

We consider a network $\mathbf{W}$ defined by $\mathbf{W}_{\text{layers}} + \gamma \mathbf{W}_{\text{connections}}$, with

$$\mathbf{W}_{\text{layers}} = \begin{bmatrix} \mathbf{W}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{W}_3 \end{bmatrix},$$

$$\mathbf{W}_{\text{connections}} = \begin{bmatrix} \mathbf{0} & \mathbf{W}_{12} & \mathbf{0} \\ \mathbf{W}_{21} & \mathbf{0} & \mathbf{W}_{23} \\ \mathbf{0} & \mathbf{W}_{32} & \mathbf{0} \end{bmatrix},$$

where $\mathbf{W}_{12}, \mathbf{W}_{21}, \mathbf{W}_{23}$ and $\mathbf{W}_{32}$ are interconnection matrices between the layers, while $\gamma \in \mathbb{R}_{\geq 0}$ weights the connection strength. The interconnection matrices are randomly generated and scaled such that $\|\mathbf{W}_{\text{connections}}\| \approx 0.01$. In this way, small γ values give a network that is close to having no connections.

For a interconnection weight $\gamma = 20$, Figure 6.10 shows that the two-layer NG-RC controller tracks the desired reference signals for each layer. This occurs after re-tuning weights, in particular the regularization parameter for the deep layer and the feedback parameter $\mathbf{K}$. For the same interconnection weight, Figure 6.11 shows the performance of the two-layer RC controller, exhibiting general recruitment of the network to the reference signal, but with worse performance than both the individual-layer recruitment and the NG-RC controller. In particular, we note that recruitment is not achieved as well for the middle layer, with the network moving both above and below the reference at different points.

To directly compare the performance of the NG-RC and RC controllers, Figure 6.12 plots the RMSE of the signals with the references for both controllers as the interconnection weight

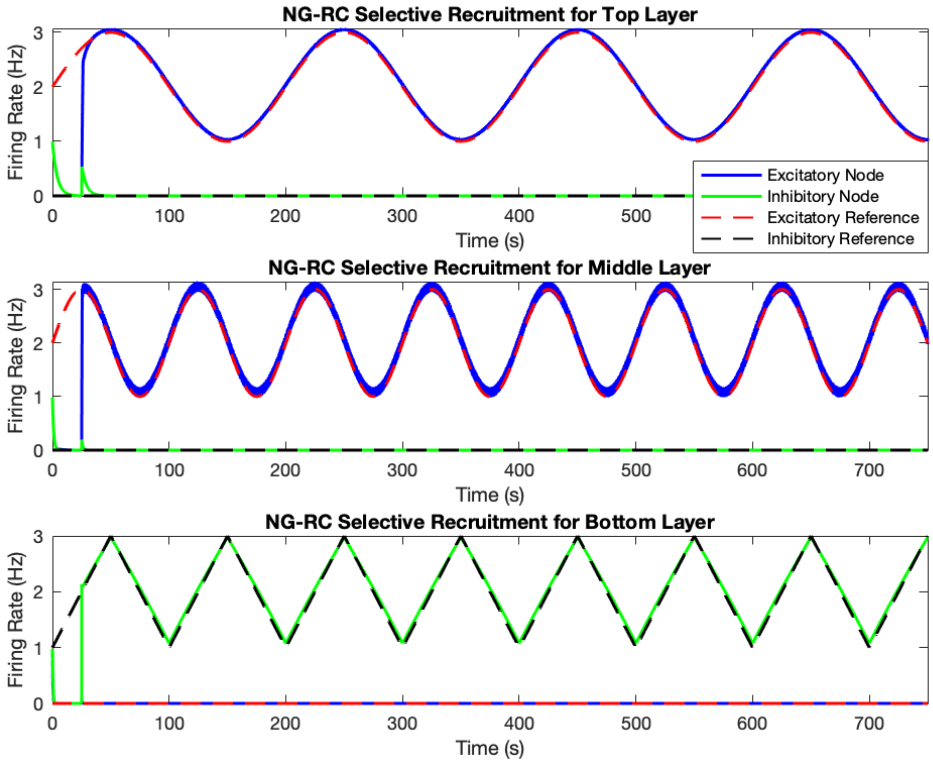


Figure 6.10. Selective inhibition and recruitment for all three layers in the interconnected network with $\gamma = 20$ using the two-layer NG-RC controller. With this level of interconnection, the controller provides performance similar to the single-layer recruitment, with RMSE errors for each layer being $0.0433, 0.0819$, and 0.0572 for the top, middle, and bottom layers, resp.

γ increases. The NG-RC controller is significantly more robust to increasing interconnection weights compared to the RC controller, which quickly moves away from satisfactory recruitment of the network to the reference trajectory. One explanation for this improved robustness by the NG-RC controller is in the determination of the control input. For the RC controller, the control input is determined strictly from the output of the reservoir. Meanwhile, the NG-RC controller additionally reconsiders the error between the reference signal and the system state, and modifies the control accordingly using the parameter $\mathbf{K}$. Therefore, despite the increasing magnitude of interconnections adding additional error to the attempted tracking, the NG-RC controller is able to control these errors for longer due to its proportional control term.

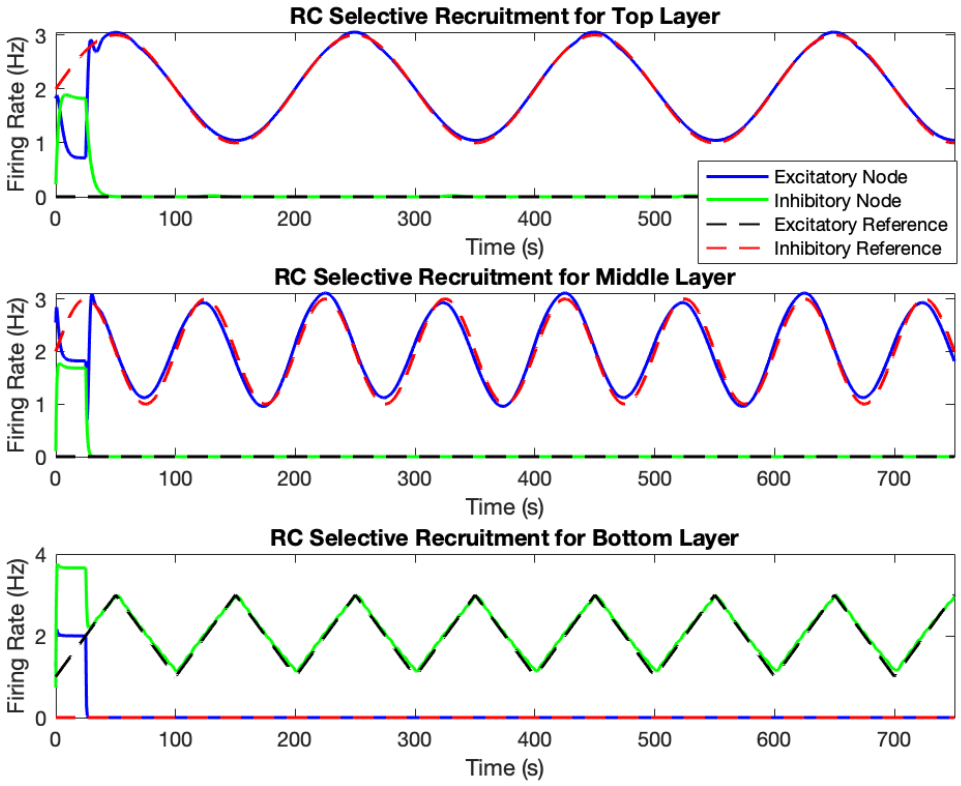


Figure 6.11. Selective inhibition and recruitment for all three layers in the interconnected network with $\gamma = 20$ using the two-layer RC controller. With this level of interconnection, the controller provides similar levels of performance to the individual networks for the top and bottom layers, while some additional error is introduced in the middle layer. The errors for the top, middle, and bottom layers are $0.0308, 0.1222$ and 0.0716, resp.

6.6.4 Comparison between the RC and NG-RC Frameworks

From the results and plots above, we see that the RC and NG-RC frameworks both have scenarios where they more successfully selectively recruit the system to a reference. In particular, for smaller networks, the RC framework produces a similar, or slightly improved, quality recruitment. Meanwhile, the NG-RC controller is more robust to increasing the magnitude of interconnections between layers. However, these comparisons are made only on the error between the reference signal and predicted signal. Depending on the situation, further metrics may be important in comparing the two controllers, such as training signal length, training time, and control signal and magnitude. In Table 6.2, we compare the training parameters and

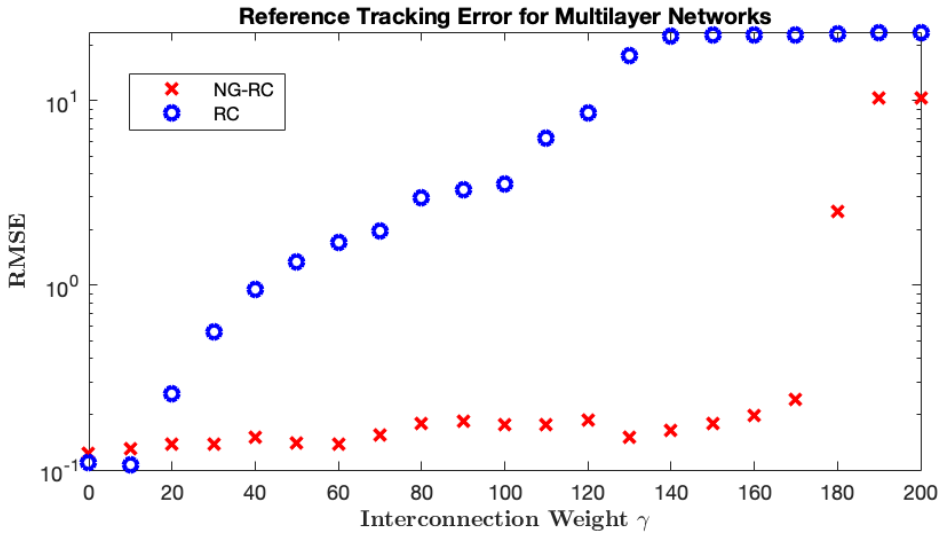


Figure 6.12. Comparison of recruitment error between the NG-RC and RC controllers as the magnitude of interconnections in the multilayer network increases. The NG-RC controller maintains a smaller error for significantly larger interconnections than the RC controller, before both become unsuccessful with large interconnections.

times, listed for the single-layer NG-RC controller, the single-layer RC controller, the multilayer NG-RC, and multilayer RC controller[2].

Table 6.2. Comparison of training parameters and times.

Network	Signal Length	Training Time	Control Time
NG-RC Single	500	0.0093	1.1543
NG-RC Multi	500	0.0655	1.6158
RC Single	80000	4.9564	6.0179
RC Multi	80000	26.18	18.17

Table 6.2 shows that the NG-RC framework allows for a much shorter training signal, and results in much faster times, both for the learning and controlling portion of the simulation. This aligns with the discussion of the frameworks in Section 6.5 and illustrates that, if training time is important when using these frameworks, NG-RC performs significantly better. We note that the training signal lengths for both frameworks was determined after experimentation, with the lengths chosen to be the minimum lengths ensuring that performance in terms of recruitment

[2]Simulations were all computed in MATLAB r2023a on a 2019 MacBook Pro with a 2.8GHz Quad-Core Intel Core i7 processor.

error were acceptable.

Figure 6.13 shows the control signal being generated by the RC and NG-RC controllers for the bottom layer without any interconnections, and compares it with the one obtained in Theorem 6.4.2. It is clear that while the two frameworks use similar controls to achieve selective recruitment, for selective inhibition they use significantly different controls in terms of magnitude (this is consistent with our observation in Remark 6.4.3). In particular, both frameworks apply a control to the node being inhibited that is significantly higher magnitude (more inhibitory) than the analytically determined control from Theorem 6.4.2. Further, the NG-RC control is significantly higher magnitude than the RC control. Due to the observed difference in generated controls, depending on physical system constraints (such as those on control magnitude), the RC controller could be preferred, despite the longer training time.

6.7 Application to Seizure Rejection

In our second application of reference tracking for LTN dynamics using reservoir computing, we consider the problem of epileptic seizure rejection. Epilepsy is a disease which impacts 50 million people worldwide and up to 30 percent of those have drug-resistant epilepsy [180], which instead can be treated with neuromodulation [181]. Data-driven methods have been been used in seizure detection [182] to predict seizure activity based on electroencephalogram (EEG) data. During epileptic seizures, brain activity becomes highly synchronized in a pathological manner, which results in the seizure symptoms. A key problem in epilepsy research is the detection of seizures in the early stages before symptoms begin to appear. This is desirable as then an interjection could be made in order to prevent the remainder of a seizure, and ideally, prevent most or all of the associated symptoms. Here we wish to look at how the reservoir computer controller network design, representing an external neurostimulation device, can be used to apply the control action to prevent seizure behavior upon detection through having the network track a desired ‘safe’ signal.

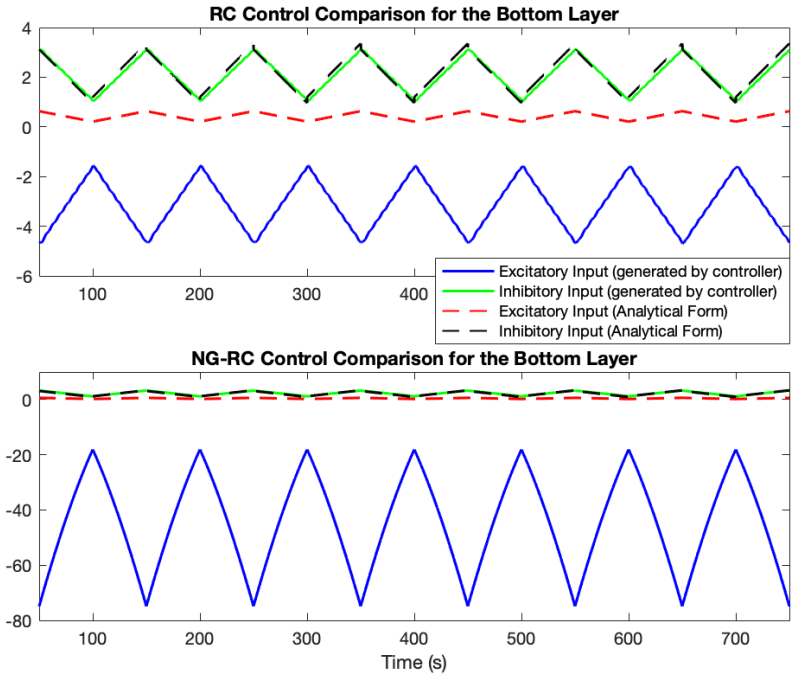


Figure 6.13. Comparison of control signals from the RC and NG-RC controllers with the analytically determined control from Theorem 6.4.2. Both the RC (top) and NG-RC (bottom) controllers generate control inputs that match the analytically determined control in Theorem 6.4.2 for the recruited node, in this case the inhibitory node. Meanwhile, the excitatory node is inhibited to zero using excessive inhibition, with significantly more inhibition being used by the NG-RC controller. Both result in high-performance tracking.

We first train a LTN model of a brain network with input and output restrictions to track EEG data that includes seizures. Then, we use the reservoir controller so that, when a seizure is predicted, the network activity is brought to a desired pattern that does not exhibit seizure symptoms. In this section we utilize an arbitrary 'safe' signal that does not predict seizure activity using a synchrony-based approach. However, in clinical application the desired reference trajectory during intervention would be carefully determined by physicians to avoid any other possible pathological behaviors. Following the intervention we then allow the brain network to return to normal function without additional input from the controller.

6.7.1 Overview of the Approach

We consider EEG data taken from the "CHB-MIT Scalp EEG Database" [183, 184]. In the data, the seizure locations are noted and are also predicted in [185]. The seizure prediction method in [185] is based on the synchrony measure weighted phase lag index (WPLI) [186], which we also use here. Other seizure prediction approaches are discussed in [187].

WPLI is defined to measure synchrony between two signals, and is used in particular to compare electrophysiological signals. In [185], this is used comparing channels of an EEG for the purpose of seizure prediction. While their prediction technique involves further analysis on top of computing the WPLI, what is important to note is that the metric attains a high value (excessive synchrony) before a seizure. In our seizure rejection approach, we aim to have the network track a signal that reduces the WPLI between the brain regions as determined by the EEG data, in order to move away from seizure activity. In particular, whenever the WPLI is computed and determined to be above threshold $m_{\text{intervene}}$, we will intervene with a control term, computed using a RC, for a pre-determined amount of time $t_{\text{intervene}}$, that drives the network to a safe activity pattern, r_{safe}. After this, the intervention will be stopped and only begin again if the next computation of the WPLI is above the threshold. This procedure is summarized in Algorithm 1.

Algorithm 1. Seizure Rejection with RC Controller

Input: $T, m_{\text{intervene}}, t_0, t_{\text{intervene}}, r_{\text{safe}}$

Train RC on brain network model
Initialize control $v(t_0) = 0$
Initialize counter $k = 1$
while $t \geq t_0$ **do**
  if $t = kT + t_0$ for $k \in \mathbb{Z}$ and not intervening **then**
    Compute and update WPLI
    $k = k + 1$
  end if
  if WPLI $> m_{\text{intervene}}$ and $t < (k-1)T + t_{\text{intervene}}$ **then**
    Compute control $v(t)$ with RC to drive system to r_{safe}
    Propagate brain network model with control $v(t)$
  else
    Do not intervene and set $v(t) = 0$
    Propagate brain network model with control $v(t)$
  end if
end while

6.7.2 Weighted Phase Lag Index and Data Processing

The WPLI measures synchrony between two signals based on the instantaneous phase of the two signals over a given time window. We compute the instantaneous phase through the Hilbert transform of a signal [188]. In particular, the instantaneous phase of a signal $\mathbf{x}(t)$ with Hilbert transform $\hat{\mathbf{x}}(t)$ is given by

$$\phi(t) = \arctan\left(\frac{\hat{\mathbf{x}}(t)}{\mathbf{x}(t)}\right).$$

For a time window Δ_t containing N points, the WPLI between signals $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ is then defined as:

$$\text{WPLI}_{\Delta_t} = \frac{|\frac{1}{N}\sum_{p=1}^{N} \sin(\phi_1(t) - \phi_2(t))|}{\frac{1}{N}\sum_{p=1}^{N} |\sin(\phi_1(t) - \phi_2(t))|}. \tag{6.13}$$

The WPLI takes values in the interval $[0, 1]$, with low values corresponding to no coupling between the signals or a phase difference equal to 0 (mod π), while stronger phase locking gives

higher values for the WPLI. If the signal is phase-locked with a non-zero difference, then the WPLI is equal to 1.

Using the WPLI to compare to EEG signals on raw data fails to capture any trends due to the artifacts and noise characteristic of EEG. We therefore take two steps to process the raw signal and reduce noise. First, the data is filtered using a passband filter on a specified, typically small, band. These are typically experimentally determined and we use, following [185], the range of $8-13$ Hz. Second, the signal is differentiated with respect to time and the absolute value is computed. This is done to flatten the basic noise and emphasize the peaks of the signal [189].

6.7.3 Reproducing the EEG Data

As seizure activity is based on synchrony between brain regions, and the WPLI is computed between two signals, we consider a network with two outputs, each representing a channel of the EEG. In particular, we consider an 6-node brain network governed by linear-threshold dynamics, with two outputs. In addition, to match neurostimulation constraints due to implanting electrodes, we also limit the inputs to two of the nodes. Here, we consider the second seizure of patient 3 in the MIT EEG database, which is clinically found to occur at time $t = 730$ seconds within the EEG file and lasts 65 seconds. Following [185], we let the network outputs represent channels F4-C4 and T8-P8 of the EEG. We use a NG-RC setup as in Figure 6.5 to learn a control signal, $\mathbf{c} : \mathbb{R} \to \mathbb{R}^2$, such that each network mimics the EEG data of channels F4-C4 and T8-P8.

The synaptic weight matrix of the randomly generated network is given by

$$\mathbf{W} = \begin{bmatrix} 0.3711 & 0.0642 & 0.3530 & -0.2614 & -0.2079 & -0.0668 \\ 0.1369 & 0.03837 & 0.1442 & -0.0067 & -0.1887 & -0.3239 \\ 0.1001 & 0.1440 & 0.2700 & -0.1189 & -0.1174 & -0.0091 \\ 0.2363 & 0.0310 & 0.0015 & -0.2568 & -0.1436 & -0.1791 \\ 0.0542 & 0.0760 & 0.1080 & -0.557 & -0.0440 & -0.3020 \\ 0.2213 & 0.1887 & 0.1996 & -0.2418 & -0.3127 & -0.0345 \end{bmatrix}, \tag{6.14}$$

with inputs in nodes 1 and 3. The outputs are defined by

$$\begin{bmatrix} \mathbf{y}_{F4-C4} \\ \mathbf{y}_{T8-P8} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}, \tag{6.15}$$

and actuation is limited to the first and third nodes. The replication of the EEG data is done using an NG-RC controller with a 500-point training signal sampled from a $\mathcal{N}(0, 0.1)$ distribution. Figure 6.14 illustrates the replication of the EEG data, along with the original data, shown for a 9-second time period preceding the seizure by tracking the original data with the NG-RC controller. For the purposes of computing the WPLI and seizure rejection, we consider the data from $t = 490$ seconds to $t = 855$ seconds, which corresponds to four minutes before the seizure and one minute after the seizure. On this timescale, it is more difficult to visually observe the success of the replication of the EEG data using the NG-RC. Figure 6.15 shows the EEG replication over the extended time period along with the WPLI over this time. The WPLI computed from the original data is also included to illustrate accuracy of the replication. For both the replicated and original data, the WPLI is computed with a time window of 6 seconds and overlap of one second as in [185]. It is evident that as the seizure approaches the WPLI is increasing, reaching the maximum value of 1 prior to the seizure (which begins at $t = 730$), before dropping to a low level during and after the pathological behavior.

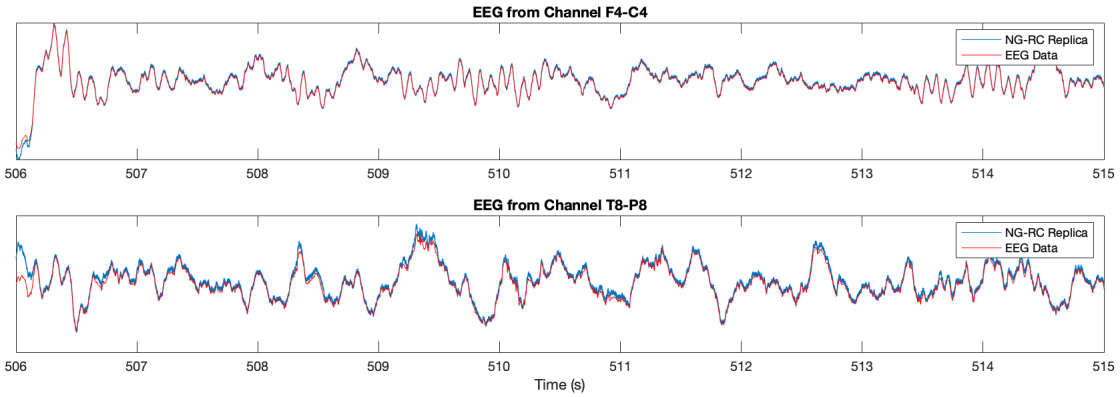


Figure 6.14. Illustration of the replication of the EEG data from the second seizure of patient 3 over the timeframe $t = (506, 515)$ seconds for channels F4-C4 and T8-P8 using a NG-RC with network matrix (6.14) and outputs (6.15). This precedes the beginning of the seizure at $t = 730$. The replication achieves a root-mean-square error (RMSE) of 0.0310 for the F4-C4 channel and 0.0316 for the T8P8 channel on this time interval.

6.7.4 Rejecting Seizure Behavior

Medical interventions will seek to prevent the seizure behavior displayed in Figure 6.15. In this section, we apply the seizure rejection method described in Algorithm 1 to accomplish this by interjecting in the model to keep the WPLI between the channels F4-C4 and T8-P8 below a threshold. To do so, we run the system with the base control $\mathbf{c}$ learned in the prior section using the NG-RC, so that the network exhibits the same behavior as the EEG data. Then, we interject the system with an additional control $\mathbf{v} : \mathbb{R} \to \mathbb{R}^2$ each time the computed WPLI reaches or exceeds a threshold of 0.8, so that the WPLI is reduced by utilizing the NG-RC controller to have the network track a pair of signals that do not exhibit excessive synchrony.

In our simulations we use a control (determined by the NG-RC controller) that modulates the network to a pre-determined safe signal for a period of 60 seconds, a duration based on the average length of a generalized tonic-clonic seizure [190]. We define our safe signal based on two sinusoids with different frequencies to guarantee a lack of synchrony and add white noise to match the noisy nature of EEG measurements. Figure 6.16 illustrates the application of Algorithm 1 to the EEG with a first rejection at $t = 625$, when the WPLI first meets the threshold of 0.8, and a second intervention at $t = 700$. We see from here on that, with the intervention, the

WPLI stops increasing towards the upper limit of 1 (which is the expected marker for leadup to seizure behavior [185]).

Remark 6.7.1. (Real-time Seizure Rejection): *In the above computations, we have used WPLI as a metric for seizure prediction, which depends on the instantaneous phase of the EEG signals as computed using the Hilbert transform. However, since the Hilbert transform is non-causal, this means that the above method cannot be applied in a real-time manner due to the requirement of future knowledge of the EEG signal for compuation of the phase. In order to address this problem and apply Algorithm 1 in real-time, we would need to be able to compute the phase in a causal manner. However, the problem of accurately computing the phase of a signal in real-time is difficult and current methods are not always accurate [191] depending on the signal. Various methods have been considered using filter models [192], machine learning techniques [193] and autoregressive models [194]. Due to differences in phase definition, along with the inherent noise in the EEG measurements, applying these approaches to the WPLI computation did not accurately reproduce the expected results. Our future work will aim to determine a method for causal instantaneous phase calculation that allows for the accurate implementation of Algorithm 1 in real-time.* •

6.8 Conclusions and Future Work

We have tackled the problem of control design for reference tracking in linear-threshold firing rate network models through the reservoir computing framework. We first formally designed explicit open- and closed-loop controllers that achieve reference tracking under suitable conditions on the synaptic connectivity. To overcome the difficulty of determining precisely the strength of interconnections in the brain, required by these controllers, and the fact that the identified conditions become increasingly difficult to check with network size, together with considerations of biological implausibility, we have proposed the use of reservoir computing to synthesize the control signals. We have shown how the RC and NG-RC frameworks can be

used as controllers for the problem of selective recruitment and inhibition of LTN networks, allowing for an arbitrarily chosen equilibrium trajectory. We have also used an NG-RC controller to replicate EEG data as well as reject epileptic seizure activity. Future work will apply the reservoir computing control framework to larger networks, with the limiting factor being the determination of optimal parameters for learning. Another direction for future work is the study of how the composition of the network and reservoir, and symmetries between them due to their structure as linear-threshold networks, could be exploited to achieve improved control performance. Finally, we will investigate how different reservoir or feature vector structures can improve performance and explore reference tracking with limited sensing.

Acknowledgements

This chapter, in full, is a reprint of material submitted for publication where it may appear as "Control of Linear-Threshold Brain Networks via Reservoir Computing" by Michael McCreesh and Jorge Cortés in IEEE Open Journal of Control Systems. The dissertation author was the primary investigator and author of this paper.

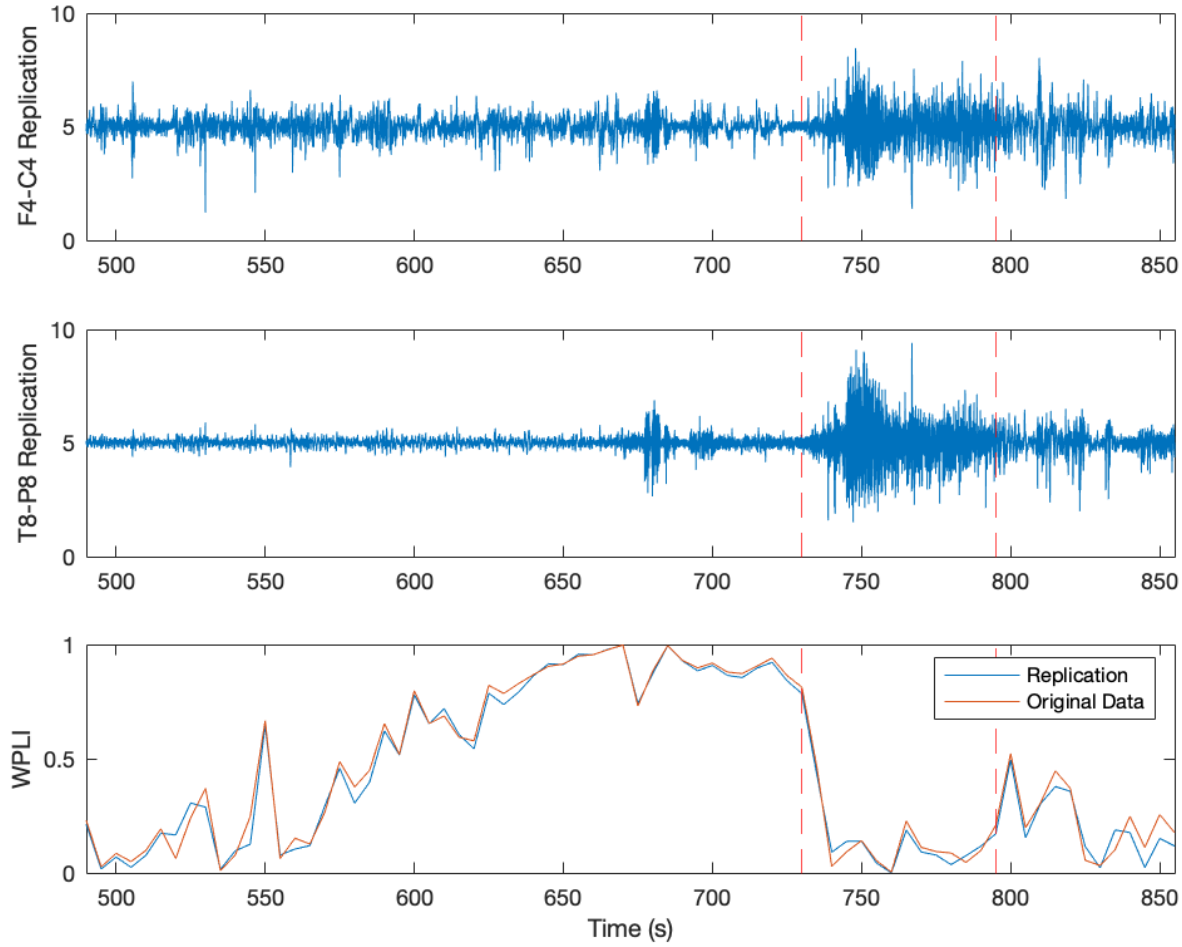


Figure 6.15. The replicated EEG from four minutes prior to and one minute after the second seizure of patient 3 ($t \in (490, 855)$). The first panel illustrates the activity of the F4-C4 electrode while the second panel is the T8-P8 electrode. Over this time frame the RMSE on the replication for the F4-C4 electrode is 0.0737, while for the T8-P8 electrode it is 0.0896. The third panel shows the WPLI for both the replicated EEG and the original data, which has a RMSE of 0.0474. We note that the WPLI increases and peaks prior to the seizure, the start and end of which are indicated by the red lines.

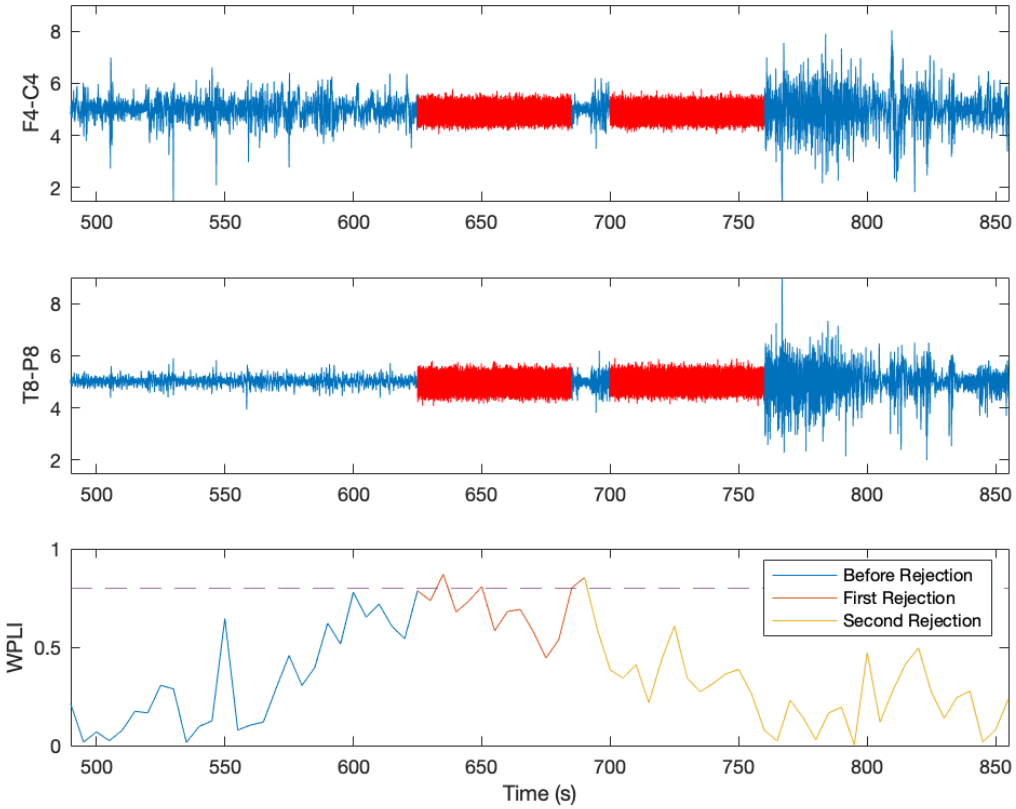


Figure 6.16. The EEG in channels F4-C4 and T8-P8 from $t = 490$ to $t = 855$, with two interventions, at $t = 625$ and at $t = 700$, when the WPLI reaches the threshold of 0.8. The controlled portions of the EEG are shown in red. We see that after interventions the WPLI moves below the threshold of 0.8, and does not exhibit the drastic drop in WPLI at the beginning of the seizure. During normal periods, a low-magnitude control (of average amplitude 100) is applied to track the base EEG activity, before an additional high-magnitude control (of average amplitude 4000) is applied during the seizure rejection periods to modulate to a safe signal. These controls are of high amplitude due to the high-frequency EEG activity requiring fast changes in value.

Chapter 7

Conclusions

7.1 Summary

In this dissertation we have studied the application of system-theoretic methods to brain networks. The use of analytic tools from fields outside of neuroscience, such as Granger causality from econometrics or entropy from information theory, have led to advances in our understanding of how the brain functions. Seeing the benefits of bringing in outside tools, along with improvements in systems and control theory tools, has led to the significant expansion in the use of system-theoretic methods to study brain networks. One method for modeling the brain is through firing rate models, in which the activity of the brain is measured by the average spiking rate of populations of neurons. We use the linear-threshold dynamics to provide such a model, due to the richness of behavior that it exhibits. We examine the properties of the dynamics along with its application to specific neural activities.

In Chapter 4, motivated by the appearance of oscillations in observed brain activity, we have studied the properties of oscillatory behavior in the linear-threshold dynamics. Prior works have studied the properties and emergence of oscillatory behavior in certain classes of the unbounded threshold-linear dynamics but these do not extend directly to the linear-threshold case. We limit our discussion to the case of competitive networks which, while being a restriction, can still be used to represent certain brain networks.

Due to the difficulties of direct analysis of oscillations in greater than two dimensions we

use the lack of stable equilibria as a proxy for oscillations, which then encompasses both the periodic and chaotic behavior that have been observed in the brain. Our approach characterizes the loss of stable equilibria in a case-by-case basis related to the support of a potential equilibrium point. This solution allows for the analysis of the oscillatory behavior of specific groups of nodes, as can be desired for individual applications. We provided necessary and sufficient conditions for the existence of a stable equilibrium point supported on a given set, along with sufficient conditions for the lack of stable equilibria with a given support.

In Chapter 5, we have considered the use of the linear-threshold network dynamics in modeling selective attention in thalamocortical brain networks. This problem has been studied prior when considering only hierarchical cortical networks, however such a topology misses components that are believed to play a significant role in communication between regions, in particular, the thalamus. Selective attention, and its subproblems of selective inhibition and recruitment, pertain to the stabilization of the network to specific equilibrium trajectories. Through tools provided by singular perturbation theory, and leveraging the piecewise-affine properties of the linear-threshold dynamics, we provided sufficient conditions such that in both hierarchical and star-connected thalamocortical network topologies, selective inhibition and recruitment can be achieved. In addition to being able to achieve this goal, we illustrated through simulation that the addition of the thalamus components results in achieving selective inhibition and recruitment both faster and with a lower control energy.

Motivated by difficulties in determining a control to achieve selective recruitment to an arbitrary equilibrium trajectory, in Chapter 6 we address the problem of reference tracking for networks governed by the linear-threshold dynamics. Based on the synaptic weight matrix, we provide a closed-form control to achieve exponential convergence to sufficiently well-behaved reference signals. However, since in the context of a brain network, the synaptic weight matrix is typically not fully known, the analytic approach is not directly applicable. To address this complication a data-driven approach was taken, utilizing both reservoir computing and next-generation reservoir computing, in order to achieve reference tracking. The effectiveness of this

approach was shown using two example applications. First, as per the original motivation of the work, it was shown that selective recruitment to an arbitrary reference signal is possible. Second, it was shown that with the data-driven controller a linear-threshold network can effectively replicate EEG signals and be utilized in the context of seizure rejection. Each of these studies leaves open avenues for future research, which we describe next.

7.2 Future Directions

The results obtained in this dissertation lay the groundwork for future research in a variety of directions. First, there are a variety of different properties of the linear-threshold dynamics that warrant further research. The results on oscillations in the linear-threshold dynamics in this work are focused on sufficient conditions for the existence of oscillations in competitive network topologies. Immediate future work is in order to extend these to necessary and sufficient conditions if possible. In addition, while competitive networks are observed in the brain and are related to a variety of behaviors, they do not explain all brain activity. As such it is of interest to explore oscillations in the LTN dynamics for other network topologies, such as the high-resolution models found using functional brain data from diagnostic tools including fMRI, EEG, and MEG. Further, while it is useful to understand when oscillations do exist, research into the properties of such oscillations is of interest. In particular, investigating conditions that lead to different types of dynamic attractors, along with the robustness of the oscillatory behavior to both network and input perturbations is a possible direction to explore. Understanding the properties of the oscillations that appear in the dynamics can allow for better modeling of behaviors related with oscillations, such as Parkinson's disease.

During our treatment of selective attention in thalamocortical networks, we illustrated how the inclusion of the thalamus in the topology can improve network performance in multiple metrics. However, while the thalamocortical topologies are a helpful extension on the original modeling of hierarchical selective recruitment, which considered only cortical networks, it does

not cover the extensive number of topologies that appear in brain networks. As such an interesting avenue for further research is into how different topologies can improve our modeling of selective attention. By using topologies that include additional brain regions to the cortex and thalamus, each of which will have their own properties, we can look to gain a further understanding into how different components of the brain impact selective attention.

In addition to investigating further topologies, further research into the control signal being applied is warranted. In the model-based control we allow for a non-zero control signal on all task-relevant components of the network, while in the data-driven formulation we place no restrictions. However, it is unlikely that this is the case in the brain, and as such research into the determination of satisfactory controls with input constraints is an interesting and important avenue of future study.

For the study of reference tracking in brain models, particularly with the linear-threshold dynamics, there are a variety of future avenues for research. First, we have provided results guaranteeing lossless tracking of a desired signal. However this requires assumptions on the behavior of the signal. An important extension is to loosen the requirements on the reference signal and provide a characterization of if reference tracking is still possible with a potential bounded steady state error.

Second, the data-driven approach we took to achieving reference tracking using reservoir computing has multiple available extensions. Direct avenues of further research include investigating properties of the reservoir computing framework itself, with an important direction being the determination of conditions that define a reservoir that will provide good performance. Both the reservoir computer and next-generation reservoir computer could also be further modified to include aspects of deep learning, in the hopes of improving performance related to error and convergence properties. Finally, testing the reservoir computing approach with a live system, rather than using a model constructed from historical data, is a particularly important direction for future research.

The systems and controls approach to modeling neural behavior also provides an avenue

to study a variety of problems other than those discussed in this dissertation. One particularly interesting problem that lends itself to using a system-theoretic approach is neurogenesis, or the addition of new neurons into a network. While not as common in adults as in the developing brain, it does occur in regions such as the hippocampus throughout our lifetime, which is involved in processes including learning and memory. This problem can be approached as the addition of new nodes into the network, and we can study how the properties of the new neurons and the original network interact to maintain, alter or improve behavior. An interesting subproblem is the design of new components that provide desired behavior when added to the network. While the addition of artificial neurons is likely beyond our current technical capacity, it is something that could be possible in the future and as such is an interesting problem to study. Conversely, one could study the opposite problem, that is the removal of neurons from the network due to death or damage, to better understand how individual neurons or populations impact overall behavior.

Bibliography

[1] M. McCreesh, T. Menara, and J. Cortés, "Sufficient conditions for oscillations in competitive linear-threshold brain networks," *IEEE Control Systems Letters*, vol. 7, pp. 2886–2891, 2023.

[2] M. McCreesh and J. Cortés, "Selective inhibition and recruitment in linear-threshold thalamocortical networks," *IEEE Transactions on Control of Network Systems*, vol. 11, no. 1, pp. 375–388, 2024.

[3] L. F. Abbott, "Lapicque's introduction of the integrate-and-fire model neuron (1907)," *Brain Research Bulletin*, vol. 50, no. 5-6, pp. 303–304, 1999.

[4] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of Physiology*, vol. 117, no. 4, pp. 500–544, 1952.

[5] R. FitzHugh, "Impulses and physiological states in theoretical models of nerve membrane," *Biophysical Journal*, vol. 1, no. 6, pp. 445–466, 1961.

[6] J. Nagumo, S. Arimoto, and S. Yoshizawa, "An active pulse transmission line simulating nerve axon," *Proceedings of the IRE*, vol. 50, no. 10, pp. 2061–2070, 1962.

[7] C. Morris and H. Lecar, "Voltage oscillations in the barnacle giant muscle fiber," *Biophysical Journal*, vol. 35, no. 1, pp. 193–213, 1981.

[8] J. S. Griffith, "A field theory of neural nets: I: Derivation of field equations," *The Bulletin of Mathematical Biophysics*, vol. 25, pp. 111–120, 1963.

[9] J. L. Feldman and J. D. Cowan, "Large-scale activity in neural nets I: Theory with application to motoneuron pool responses," *Biological Cybernetics*, vol. 17, no. 1, pp. 29–38, 1975.

[10] W. Lu and T. Chen, "New conditions on global stability of Cohen-Grossberg neural networks," *Neural Computation*, vol. 15, no. 5, pp. 1173–1189, 2003.

[11] D. Jaeger and R. Jung, *Encyclopedia of Computational Neuroscience*. Springer, 2015.

[12] J. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," *Proceedings of the National Academy of Sciences*, vol. 81, no. 10, pp. 3088–3092, 1984.

[13] J. J. Hopfield and D. W. Tank, "Computing with neural circuits: A model," *Science*, vol. 233, no. 4764, pp. 625–633, 1986.

[14] S. Grossberg, "Nonlinear neural networks: Principles, mechanisms, and architectures," *Neural networks*, vol. 1, no. 1, pp. 17–61, 1988.

[15] S. Grossberg, "Some networks that can learn, remember, and reproduce any number of complicated space-time patterns, I," *Journal of Mathematics and Mechanics*, vol. 19, no. 1, pp. 53–91, 1969.

[16] S. Coombes, "Waves, bumps, and patterns in neural field theories," *Biological Cybernetics*, vol. 93, pp. 91–108, 2005.

[17] P. C. Bressloff, "Spatiotemporal dynamics of continuum neural fields," *Journal of Physics A: Mathematical and Theoretical*, vol. 45, no. 3, p. 033001, 2011.

[18] H. R. Wilson and J. D. Cowan, "Excitatory and inhibitory interactions in localized populations of model neurons," *Biophysical Journal*, vol. 12, no. 1, pp. 1–24, 1972.

[19] H. R. Wilson and J. D. Cowan, "A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue," *Kybernetik*, vol. 13, no. 2, pp. 55–80, 1973.

[20] E. Nozari, M. A. Bertolero, J. Stiso, L. Caciagli, E. J. Cornblath, X. He, A. S. Mahadevan, G. J. Pappas, and D. S. Bassett, "Macroscopic resting-state brain dynamics are best described by linear models," *Nature Biomedical Engineering*, vol. 8, no. 1, pp. 68–84, 2024.

[21] L. M. Sweet, "Automotive applications of modern control theory," *SAE Transactions*, pp. 3022–3029, 1982.

[22] P. Ioannou, Z. Xu, S. Eckert, D. Clemons, and T. Sieja, "Intelligent cruise control: theory and experiment," in *IEEE Conf. on Decision and Control*, pp. 1885–1890, IEEE, 1993.

[23] B. L. Stevens, F. L. Lewis, and E. N. Johnson, *Aircraft Control and Simulation: Dynamics, Controls Design, and Autonomous Systems*. Wiley, 2015.

[24] J. H. Blakelock, *Automatic Control of Aircraft and Missiles*. John Wiley & Sons, 1991.

[25] T. Tarn, A. K. Bejczy, A. Isidori, and Y. Chen, "Nonlinear feedback in robot arm control," in *IEEE Conf. on Decision and Control*, pp. 736–751, IEEE, 1984.

[26] K. S. Fu, R. C. Gonzalez, C. S. G. Lee, and H. Freeman, *Robotics: control, sensing, vision, and intelligence*, vol. 1. McGraw-Hill New York, 1987.

[27] J. Slotine, "The robust control of robot manipulators," *The International Journal of Robotics Research*, vol. 4, no. 2, pp. 49–64, 1985.

[28] T. Basar, *Control theory: twenty-five seminal papers (1932-1981)*. 2001.

[29] S. Schiff, *Neural control engineering: the emerging intersection between control theory and neuroscience*. MIT Press, 2012.

[30] S. J. Schiff, "Kalman meets neuron: the emerging intersection of control theory with neuroscience," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 3318–3321, IEEE, 2009.

[31] S. Gu, F. Pasqualetti, M. Cieslak, Q. K. Telesford, A. B. Yu, A. E. Kahn, J. D. Medaglia, J. M. Vettel, M. B. Miller, S. T. Grafton, and D. S. Bassett, "Controllability of structural brain networks," *Nature Communications*, vol. 6, p. 8414, October 2015.

[32] M. S. Madhav and N. J. Cowan, "The synergy between neuroscience and control theory: the nervous system as inspiration for hard control challenges," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 3, pp. 243–267, 2020.

[33] S. Ching, M. Y. Liberman, J. J. Chemali, M. B. Westover, J. D. Kenny, K. Solt, P. L. Purdon, and E. N. Brown, "Real-time closed-loop control in a rodent model of medically induced coma using burst suppression," *Anesthesiology: The Journal of the American Society of Anesthesiologists*, vol. 119, no. 4, pp. 848–860, 2013.

[34] P. Srivastava, E. Nozari, J. Z. Kim, H. Ju, D. Zhou, C. Becker, F. Pasqualetti, G. J. Pappas, and D. S. Bassett, "Models of communication and control for brain networks: distinctions, convergence, and future outlook," *Network Neuroscience*, vol. 4, no. 4, pp. 1122–1159, 2020.

[35] G. Drion, T. O'Leary, J. Dethier, A. Franci, and R. Sepulchre, "Neuronal behaviors: a control perspective," in *IEEE Conf. on Decision and Control*, (Osaka, Japan), Dec. 2015. Tutorial.

[36] C. J. Honey, T. Thesen, T. H. Donner, L. J. Silbert, C. E. Carlson, O. Devinsky, W. K. Doyle, N. Rubin, D. J. Heeger, and U. Hasson, "Slow cortical dynamics and the accumulation of information over long timescales," *Neuron*, vol. 76, no. 2, pp. 423–434, 2012.

[37] Y. Nakahira, Q. Liu, T. J. Sejnowski, and J. C. Doyle, "Diversity-enabled sweet spots in layered architectures and speed-accuracy trade-offs in sensorimotor control," *Proceedings of the National Academy of Sciences*, vol. 118, no. 22, 2021.

[38] M. E. Broucke, "Adaptive internal model theory of the oculomotor system and the cerebellum," *IEEE Transactions on Automatic Control*, vol. 66, no. 11, pp. 5444–5450, 2021.

[39] M. E. Broucke, "Adaptive internal models in neuroscience," *Foundations and Trends in Systems and Control*, vol. 9, no. 4, pp. 365–550, 2022.

[40] N. Sharma, N. Kirsch, N. A. Alibeji, and W. E. Dixon, "A non-linear control method to compensate for muscle fatigue during neuromuscular electrical stimulation," *Frontiers in Robotics and AI*, vol. 4, p. 68, 2017.

[41] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field," *Clinical Neurophysiology*, vol. 116, no. 10, pp. 2266–2301, 2005.

[42] S. G. Ferber, A. Weller, and H. Soreq, "Control systems theory revisited: new insights on the brain clocks of time-to-action," *Frontiers in Neuroscience*, vol. 17, p. 1171765, 2023.

[43] G. Acharya, S. F. Ruf, and E. Nozari, "Brain modeling for control: A review," *arXiv preprint arXiv:2210.15957*, 2022.

[44] E. Bullmore and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186–198, 2009.

[45] S. F. Muldoon, F. Pasqualetti, S. Gu, M. Cieslak, S. T. Grafton, J. M. Vettel, and D. S. Bassett, "Stimulation-based control of dynamic brain networks," *PLOS Computational Biology*, vol. 12, no. 9, p. e1005076, 2016.

[46] J. D. Medaglia, F. Pasqualetti, R. H. Hamilton, S. L. Thompson-Schill, and D. S. Bassett, "Brain and cognitive reserve: translation via network control theory," *Neuroscience & Biobehavioral Reviews*, vol. 75, pp. 53–64, 2017.

[47] Y. Yang, A. T. Connolly, and M. M. Shanechi, "A control-theoretic system identification framework and a real-time closed-loop clinical simulation testbed for electrical brain stimulation," *Journal of Neural Engineering*, vol. 15, no. 6, p. 066007, 2018.

[48] M. F. Singh, M. W. Cole, T. S. Braver, and S. Ching, "Developing control-theoretic objectives for large-scale brain dynamics and cognitive enhancement," *Annual Reviews in Control*, vol. 54, pp. 363–376, 2022.

[49] C. A. Coté, "A dynamic systems theory model of visual perception development," *Journal of Occupational Therapy, Schools, & Early Intervention*, vol. 8, no. 2, pp. 157–169, 2015.

[50] A. Kukona and W. Tabor, "Impulse processing: A dynamical systems model of incremental eye movements in the visual world paradigm," *Cognitive science*, vol. 35, no. 6, pp. 1009–1051, 2011.

[51] M. Kawato, K. Furukawa, and R. Suzuki, "A hierarchical neural-network model for control and learning of voluntary movement," *Biological Cybernetics*, vol. 57, pp. 169–185, 1987.

[52] K. V. Shenoy, M. T. Kaufman, M. Sahani, and M. M. Churchland, "A dynamical systems view of motor preparation: implications for neural prosthetic system design," *Progress in Brain Research*, vol. 192, pp. 33–58, 2011.

[53] C. I. Connolly, J. B. Burns, and M. S. Jog, "A dynamical-systems model for Parkinson's disease," *Biological Cybernetics*, vol. 83, pp. 47–59, 2000.

[54] E. J. Müller, S. J. van Albada, J. W. Kim, and P. A. Robinson, "Unified neural field theory of brain dynamics underlying oscillations in Parkinson's disease and generalized epilepsies," *Journal of Theoretical Biology*, vol. 428, pp. 132–146, 2017.

[55] F. L. D. Silva, W. Blanes, S. N. Kalitzin, J. Parra, P. Suffczynski, and D. N. Velis, "Epilepsies as dynamical diseases of brain systems: basic models of the transition between normal and epileptic activity," *Epilepsia*, vol. 44, pp. 72–83, 2003.

[56] C. M. Davidson, A. M. de Paor, and M. M. Lowery, "Insights from control theory into deep brain stimulation for relief from Parkinson's disease," in *2012 ELEKTRO*, pp. 2–7, IEEE, 2012.

[57] D. S. Bassett and P. Z. andJ. I. Gold, "On the nature and use of models in network neuroscience," *Nature Reviews Neuroscience*, vol. 19, no. 9, pp. 566–578, 2018.

[58] S. Denève, A. Alemi, and R. Bourdoukan, "The brain as an efficient and robust adaptive learner," *Neuron*, vol. 94, no. 5, pp. 969–977, 2017.

[59] F. Ratliff and H. K. Hartline, *Studies on Excitation and Inhibition in the Retina*. Rockefeller University Press, 1974.

[60] Z. Yi, L. Zhang, J. Yu, and K. K. Tan, "Permitted and forbidden sets in discrete-time linear threshold recurrent neural networks," *IEEE Transactions on Neural Networks*, vol. 20, no. 6, pp. 952–963, 2009.

[61] C. Curto, J. Geneson, and K. Morrison, "Stable fixed points of combinatorial threshold-linear networks," *Advances in Applied Mathematics*, vol. 154, p. 102652, 2024.

[62] T. Arakaki, S. Mahon, S. Charpier, A. Leblois, and D. Hansel, "The role of striatal feedforward inhibition in the maintenance of absence seizures," *Journal of Neuroscience*, vol. 36, no. 37, pp. 9618–9632, 2016.

[63] F. Celi, A. Allibhoy, F. Pasqualetti, and J. Cortés, "Linear-threshold dynamics for the study of epileptic events," *IEEE Control Systems Letters*, vol. 5, no. 4, pp. 1405–1410, 2021.

[64] E. Nozari and J. Cortés, "Hierarchical selective recruitment in linear-threshold brain networks. Part I: Intra-layer dynamics and selective inhibition," *IEEE Transactions on Automatic Control*, vol. 66, no. 3, pp. 949–964, 2021.

[65] E. Nozari and J. Cortés, "Hierarchical selective recruitment in linear-threshold brain networks. Part II: Inter-layer dynamics and top-down recruitment," *IEEE Transactions on Automatic Control*, vol. 66, no. 3, pp. 965–980, 2021.

[66] E. Nozari, R. Planas, and J. Cortés, "Structural characterization of oscillations in brain networks with rate dynamics," *Automatica*, vol. 146, p. 110653, 2022.

[67] X. Wang and J. Cortés, "Data-driven control of linear-threshold network dynamics," in *American Control Conference*, (Atlanta, Georgia), pp. 114–119, June 2022.

[68] P. Dayan and L. F. Abbott, *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Computational Neuroscience, Cambridge, MA: MIT Press, 2001.

[69] A. Destexhe and T. J. Sejnowski, "The Wilson-Cowan model, 36 years later," *Biological Cybernetics*, vol. 101, no. 1, pp. 1–2, 2009.

[70] K. P. Hadeler and D. Kuhn, "Stationary states of the Hartline-Ratliff model," *Biological Cybernetics*, vol. 56, no. 5-6, pp. 411–417, 1987.

[71] J. Feng and K. P. Hadeler, "Qualitative behaviour of some simple networks," *Journal of Physics A: Mathematical and General*, vol. 29, no. 16, pp. 5019–5033, 1996.

[72] Z. Yi and K. K. Tan, "Multistability of discrete-time recurrent neural networks with unsaturating piecewise linear activation functions," *IEEE Transactions on Neural Networks*, vol. 15, no. 2, pp. 329–336, 2004.

[73] R. H. R. Hahnloser, H. S. Seung, and J. J. Slotine, "Permitted and forbidden sets in symmetric threshold-linear networks," *Neural Computation*, vol. 15, no. 3, pp. 621–638, 2003.

[74] K. Morrison, A. Degeratu, V. Itskov, and C. Curto, "Diversity of emergent dynamics in competitive threshold-linear networks," *SIAM Journal on Applied Dynamical Systems*, vol. 23, no. 1, pp. 855–884, 2024.

[75] G. B. Ermentrout and D. H. Terman, *Mathematical Foundations of Neuroscience*, vol. 35. New York: Springer, 2010.

[76] E. M. Izhikevich, *Dynamical Systems in Neuroscience*. Cambridge, MA: MIT Press, 2007.

[77] M. Breakspear, "Dynamic models of large-scale brain activity," *Nature Neuroscience*, vol. 20, no. 3, pp. 340–352, 2017.

[78] X. Li, D. Coyle, L. Maguire, T. M. McGinnity, and H. Benali, "A model selection method for nonlinear system identification based fMRI effective connectivity analysis," *IEEE Transactions on Medical Imaging*, vol. 30, no. 7, pp. 1365–1380, 2011.

[79] K. E. Stephan, L. Kasper, L. M. Harrison, J. Daunizeau, H. E. M. den Ouden, M. Breakspear, and K. J. Friston, "Nonlinear dynamic causal models for fMRI," *Neuroimage*, vol. 42, no. 2, pp. 649–662, 2008.

[80] G. T. Einevoll, C. Kayser, N. K. Logothetis, and S. Panzeri, "Modelling and analysis of local field potentials for studying the function of cortical circuits," *Nature Reviews Neuroscience*, vol. 14, no. 11, pp. 770–785, 2013.

[81] S. Keeley, Á. Byrne, A. Fenton, and J. Rinzel, "Firing rate models for gamma oscillations," *Journal of Neurophysiology*, vol. 121, no. 6, pp. 2181–2190, 2019.

[82] T. Schwalger and A. V. Chizhov, "Mind the last spike—firing rate models for mesoscopic populations of spiking neurons," *Current Opinion in Neurobiology*, vol. 58, pp. 155–166, 2019.

[83] A. N. Khambhati, A. Sizemore, R. Betzel, and D. Bassett, "Modeling and interpreting mesoscale network dynamics," *NeuroImage*, vol. 180, pp. 337–349, 2018.

[84] D. A. Henze, Z. Borhegyi, J. Csicsvari, M. A. Mamiya, K. D. Harris, and G. Buzsaki, "Intracellular features predicted by extracellular recordings in the hippocampus in vivo," *Journal of Neurophysiology*, vol. 84, no. 1, pp. 390–400, 2000.

[85] D. A. Henze, K. D. Harris, Z. Borhegyi, J. Csicsvari, A. Mamiya, H. Hirase, A. Sirota, and G. Buzsaki, "Simultaneous intracellular and extracellular recordings from hippocampus region ca1 of anesthetized rats." CRCNS.org, 2009.

[86] B. Wang, W. Ke, J. Guang, G. Chen, L. Yin, S. Deng, Q. He, Y. Liu, T. He, R. Zheng, Y. Jiang, X. Zhang, T. Li, G. Luan, H. D. Lu, M. Zhang, X. Zhang, and Y. Shu, "Firing frequency maxima of fast-spiking neurons in human, monkey, and mouse neocortex," *Frontiers in Cellular Neuroscience*, vol. 10, p. 239, 2016.

[87] C. Curto, A. Degeratu, and V. Itskov, "Encoding binary neural codes in networks of threshold-linear neurons," *Neural Computation*, vol. 25, no. 11, pp. 2858–2903, 2013.

[88] Q. Liu, A. Ulloa, and B. Horwitz, "Using a large-scale neural model of cortical object processing to investigate the neural substrate for managing multiple items in short-term memory," *Journal of Cognitive Neuroscience*, vol. 29, no. 11, pp. 1860–1876, 2017.

[89] M. Kajiwara, R. Nomura, F. Goetze, M. Kawabata, Y. Isomura, T. Akutsu, and M. Shimono, "Inhibitory neurons exhibit high controlling ability in the cortical microconnectome," *PLOS Computational Biology*, vol. 17, no. 4, p. e1008846, 2021.

[90] J. Chen, U. Hasson, and C. Honey, "Processing timescales as an organizing principle for primate cortex," *Neuron*, vol. 88, no. 2, pp. 244–246, 2015.

[91] M. Bear, B. Connors, and M. A. Paradiso, *Neuroscience: Exploring the Brain*. Jones & Bartlett Learning, 2020.

[92] S. M. Sherman and R. W. Guillery, *Exploring the Thalamus and Its Role in Cortical Function*. MIT press, 2006.

[93] K. George and J. M. Das, "Neuroanatomy, thalamocortical radiations," *StatPearls [Internet]*, 2020.

[94] S. M. Sherman, "Thalamocortical interactions," *Current Opinion in Neurobiology*, vol. 22, no. 4, pp. 575–579, 2012.

[95] L. Gabernet, S. P. Jadhav, D. E. Feldman, M. Carandini, and M. Scanziani, "Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition," *Neuron*, vol. 48, no. 2, pp. 315–327, 2005.

[96] S. Cruikshank, T. J. Lewis, and B. Connors, "Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex," *Nature Neuroscience*, vol. 10, no. 4, pp. 462–468, 2007.

[97] K. Delevich, J. Tucciarone, Z. J. Huang, and B. Li, "The mediodorsal thalamus drives feedforward inhibition in the anterior cingulate cortex via parvalbumin interneurons," *Journal of Neuroscience*, vol. 35, no. 14, pp. 5743–5753, 2015.

[98] J. T. Porter, C. K. Johnson, and A. Agmon, "Diverse types of interneurons generate thalamus-evoked feedforward inhibition in the mouse barrel cortex," *Journal of Neuroscience*, vol. 21, no. 8, pp. 2699–2710, 2001.

[99] H. A. Swadlow, "Thalamocortical control of feed-forward inhibition in awake somatosensory 'barrel'cortex," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 357, no. 1428, pp. 1717–1727, 2002.

[100] E. K. Kosmidis and J. Vibert, "Feed-forward inhibition in the visual thalamus," *Neurocomputing*, vol. 44, pp. 479–487, 2002.

[101] J. S. Isaacson and M. Scanziani, "How inhibition shapes cortical activity," *Neuron*, vol. 72, no. 2, pp. 231–243, 2011.

[102] H. Khalil, *Nonlinear Systems, 3rd ed.* Englewood Cliffs, NJ: Prentice Hall, 2002.

[103] L. M. Ward, "Synchronous neural oscillations and cognitive processes," *Trends in Cognitive Sciences*, vol. 7, no. 12, pp. 553–559, 2003.

[104] A. Engel and P. Fries, "Neuronal oscillations, coherence, and consciousness," in *The Neurology of Conciousness*, pp. 49–60, Elsevier, 2016.

[105] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proceedings of the National Academy of Sciences*, vol. 79, no. 8, pp. 2554–2558, 1982.

[106] G. E. Hinton and T. J. Sejnowski, "Learning and relearning in Boltzmann machines," in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, p. 2, 1986.

[107] O. M. Velarde, H. A. Makse, and L. C. Parra, "Architecture of the brain's visual system enhances network stability and performance through layers, delays, and feedback," *PLOS Computational Biology*, vol. 19, no. 11, p. e1011078, 2023.

[108] A. Allibhoy, F. Celi, F. Pasqualetti, and J. Cortés, "Optimal network interventions to control the spreading of oscillations," *IEEE Open Journal of Control Systems*, vol. 1, pp. 141–151, 2022.

[109] A. Bel, R. Cobiaga, W. Reartes, and H. G. Rotstein, "Periodic solutions in threshold-linear networks and their entrainment," *SIAM Journal on Applied Dynamical Systems*, vol. 20, no. 3, pp. 1177–1208, 2021.

[110] C. Curto, J. Geneson, and K. Morrison, "Fixed points of competitive threshold-linear networks," *Neural Computation*, vol. 31, no. 1, pp. 94–155, 2019.

[111] C. Parmelee, J. Alvarez, C. Curto, and K. Morrison, "Sequential attractors in combinatorial threshold-linear networks," *SIAM Journal on Applied Dynamical Systems*, vol. 21, no. 2, pp. 1597–1630, 2022.

[112] C. Parmelee, S. Moore, K. Morrison, and C. Curto, "Core motifs predict dynamic attractors in combinatorial threshold-linear networks," *PloS One*, vol. 17, no. 3, p. e0264456, 2022.

[113] K. Morrison and C. Curto, "Predicting neural network dynamics via graphical analysis," in *Algebraic and Combinatorial Computational Biology*, Elsevier, 2018.

[114] A. M. C. Kelly, L. Q. Uddin, B. B. Biswal, F. X. Castellanos, and M. P. Milham, "Competition between functional brain networks mediates behavioral variability," *Neuroimage*, vol. 39, no. 1, pp. 527–537, 2008.

[115] D. Dupret, B. Pleydell-Bouverie, and J. Csicsvari, "Inhibitory interneurons and network oscillations," *Proceedings of the National Academy of Sciences*, vol. 105, no. 47, pp. 18079–18080, 2008.

[116] C. Curto, A. Degeratu, and V. Itskov, "Flexible memory networks," *Bulletin of Mathematical Biology*, vol. 74, no. 3, pp. 590–614, 2012.

[117] T. Donoghue, M. Haller, E. J. Peterson, P. Varma, P. Sebastian, R. Gao, T. Noto, A. H. Lara, J. D. Wallis, R. T. Knight, A. Shestyuk, and B. Voytek, "Parameterizing neural power spectra into periodic and aperiodic components," *Nature Neuroscience*, vol. 23, no. 12, pp. 1655–1665, 2020.

[118] L. Perko, *Differential Equations and Dynamical Systems*, vol. 7 of *Texts in Applied Mathematics*. New York: Springer, 3rd ed., 2000.

[119] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.

[120] M. Fiedler and V. Ptak, "On matrices with non-positive off-diagonal elements and positive principal minors," *Czechoslovak Mathematical Journal*, vol. 12, no. 3, pp. 382–400, 1962.

[121] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[122] B. G. Windham, B. Deere, M. E. Griswold, W. Wang, D. C. Bezerra, D. Shibata, K. Butler, D. Knopman, R. F. Gottesman, G. Heiss, and T. H. M. Jr, "Small brain lesions and incident stroke and mortality: a cohort study," *Annals of Internal Medicine*, vol. 163, no. 1, pp. 22–31, 2015.

[123] N. Tinbergen, "The hierarchical organization of nervous mechanisms underlying instinctive behaviour," in *Symposium for the Society for Experimental Biology*, vol. 4, pp. 305–312, 1950.

[124] A. R. Luria, "The functional organization of the brain," *Scientific American*, vol. 222, no. 3, pp. 66–79, 1970.

[125] S. J. Kiebel, J. Daunizeau, and K. J. Friston, "A hierarchy of time-scales and the brain," *PLOS Computational Biology*, vol. 4, no. 11, p. e1000209, 2008.

[126] U. Hasson, J. Chen, and C. J. Honey, "Hierarchical process memory: memory as an integral component of information processing," *Trends in Cognitive Sciences*, vol. 19, no. 6, pp. 304–313, 2015.

[127] M. I. Rabinovich, I. Tristan, and P. Varona, "Hierarchical nonlinear dynamics of human attention," *Neuroscience & Biobehavioral Reviews*, vol. 55, pp. 18–35, 2015.

[128] K. Hwang, M. A. Bertolero, W. B. Liu, and M. D'Esposito, "The human thalamus is an integrative hub for functional brain networks," *Journal of Neuroscience*, vol. 37, no. 23, pp. 5594–5607, 2017.

[129] R. D. D'Souza and A. Burkhalter, "A laminar organization for selective cortico-cortical communication," *Frontiers in Neuroanatomy*, vol. 11, p. 71, 2017.

[130] J. M. Alonso and H. A. Swadlow, "Thalamus controls recurrent cortical dynamics," *Nature Neuroscience*, vol. 18, no. 12, pp. 1703–1704, 2015.

[131] J. A. Harris, S. Mihalas, K. E. Hirokawa, J. D. Whitesell, H. Choi, A. Bernard, P. Bohn, S. Caldejon, L. Casal, A. Cho, A. Feiner, D. Feng, N. Gaudreault, C. R. Gerfen, N. Graddis, P. A. Groblewski, A. M. Henry, A. Ho, R. Howard, J. E. Knox, L. Kuan, X. Kuang, J. Lecoq, P. Lesnar, Y. Li, J. Luviano, S. McConoughey, M. T. Mortrud, M. Naeemi, L. Ng, S. W. Oh, B. Oullette, E. Shen, S. A. Sorenson, W. Wakeman, Q. Wang, Y. Wang, A. Williford, J. W. Phillips, A. R. Jones, C. Koch, and H. Zeng, "Hierarchical organization of cortical and thalamic connectivity," *Nature*, vol. 575, no. 7781, pp. 195–202, 2019.

[132] M. M. Halassa and L. Acsády, "Thalamic inhibition: diverse sources, diverse scales," *Trends in Neurosciences*, vol. 39, no. 10, pp. 680–693, 2016.

[133] D. Liberzon, *Switching in Systems and Control*. Systems & Control: Foundations & Applications, Birkhäuser, 2003.

[134] P. V. Kokotović and H. K. Khalil, eds., *Singular Perturbation Methods in Control: Analysis and Design*. SIAM, 1999.

[135] V. Veliov, "A generalization of the Tikhonov theorem for singularly perturbed differential inclusions," *Journal of Dynamical & Control Systems*, vol. 3, no. 3, pp. 291–319, 1997.

[136] R. Chaudhuri, K. Knoblauch, M. Gariel, H. Kennedy, and X. Wang, "A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex," *Neuron*, vol. 88, no. 2, pp. 419–431, 2015.

[137] S. M. Sherman and R. W. Guillery, "Distinct functions for direct and transthalamic corticocortical connections," *Journal of Neurophysiology*, vol. 106, no. 3, pp. 1068–1077, 2011.

[138] M. Wolff and S. D. Vann, "The cognitive thalamus as a gateway to mental representations," *Journal of Neuroscience*, vol. 39, no. 1, pp. 3–14, 2019.

[139] H. Gudden, "Klinische und anatomische beiträge zur kenntniss der multiplen alkoholneuritis nebst bemerkungen über die regenerationsvorgänge im peripheren nervensystem," *Archiv für Psychiatrie und Nervenkrankheiten*, vol. 28, no. 3, pp. 643–741, 1896.

[140] E. Ahissar and T. Oram, "Thalamic relay or cortico-thalamic processing? Old question, new answers," *Cerebral Cortex*, vol. 25, no. 4, pp. 845–848, 2015.

[141] Y. B. Saalmann and S. Kastner, "Cognitive and perceptual functions of the visual thalamus," *Neuron*, vol. 71, no. 2, pp. 209–223, 2011.

[142] S. M. Sherman and R. W. Guillery, "Functional organization of thalamocortical relays," *Journal of Neurophysiology*, vol. 76, no. 3, pp. 1367–1395, 1996.

[143] J. A. Gallego, M. G. Perich, R. H. Chowdhury, S. A. Solla, and L. E. Miller, "Long-term stability of cortical population dynamics underlying consistent behavior," *Nature Neuroscience*, vol. 23, no. 2, pp. 260–270, 2020.

[144] K. E. Chu, "Generalization of the Bauer-Fike theorem," *Numerische Mathematik*, vol. 49, no. 6, pp. 685–691, 1986.

[145] G. K. Kostopoulos, "Involvement of the thalamocortical system in epileptic loss of consciousness," *Epilepsia*, vol. 42, pp. 13–19, 2001.

[146] V. Vuksanović and P. Hövel, "Functional connectivity of distant cortical regions: role of remote synchronization and symmetry in interactions," *NeuroImage*, vol. 97, pp. 1–8, 2014.

[147] Y. Qin, Y. Kawano, and M. Cao, "Stability of remote synchronization in star networks of Kuramoto oscillators," in *IEEE Conf. on Decision and Control*, (Miami Beach, USA), pp. 5209–5214, 2018.

[148] Z. Ye, X. Yu, C. M. Houston, Z. Aboukhalil, N. P. Franks, W. Wisden, and S. G. Brickley, "Fast and slow inhibition in the visual thalamus is influenced by allocating $GABA_A$ receptors with different γ subunits," *Frontiers in Cellular Neuroscience*, vol. 11, p. 95, 2017.

[149] S. Lee and H. Ahn, "Robust stability of slowly varying nonlinear systems having a continuum of equilibria," in *IEEE Conf. on Decision and Control*, (Osaka, Japan), IEEE, Dec. 2015.

[150] M. Marconi, N. D. C. Blanco, C. Zimmer, and A. Guyon, "Eye movements in response to different cognitive activities measured by eyetracking: a prospective study on some of the neurolinguistics programming theories," *Journal of Eye Movement Research*, vol. 16, no. 2, 2023.

[151] A. P. Vaz, J. H. Wittig Jr., S. K. Inati, and K. A. Zaghloul, "Replay of cortical spiking sequences during human memory retrieval," *Science*, vol. 367, no. 6482, pp. 1131–1134, 2020.

[152] G. Ariani, J. A. Pruszynski, and J. Diedrichsen, "Motor planning brings human primary somatosensory cortex into action-specific preparatory states," *eLife*, vol. 11, p. e69517, 2022.

[153] A. N. Burkitt, "A review of the integrate-and-fire neuron model: I. homogeneous synaptic input," *Biological Cybernetics*, vol. 95, pp. 1–19, 2006.

[154] M. Khosla, G. H. Ngo, K. Jamison, A. Kuceyeski, and M. R. Sabuncu, "Cortical response to naturalistic stimuli is largely predictable with deep neural networks," *Science Advances*, vol. 7, no. 22, p. eabe7547, 2021.

[155] F. Shao and Z. Shen, "How can artificial neural networks approximate the brain?," *Frontiers in Psychology*, vol. 13, p. 970214, 2023.

[156] N. Kriegeskorte, "Deep neural networks: a new framework for modeling biological vision and brain information processing," *Annual Review of Vision Science*, vol. 1, pp. 417–446, 2015.

[157] Y. Li, H. Yang, and S. Gu, "Upgrading voxel-wise encoding model via integrated integration over features and brain networks," *BioRxiv*, 2022.

[158] G. Baggio, D. S. Bassett, and F. Pasqualetti, "Data-driven control of complex networks," *Nature Communications*, vol. 12, no. 1, pp. 1–13, 2021.

[159] Y. Qin, T. Menara, S. Oymak, S. Ching, and F. Pasqualetti, "Representation learning for context-dependent decision-making," in *American Control Conference*, (Atlanta, GA), pp. 2130–2135, 2022.

[160] V. Narayanan, J. T. Ritt, J. Li, and S. Ching, "A learning framework for controlling spiking neural networks," in *American Control Conference*, pp. 211–216, IEEE, 2019.

[161] J. Z. Kim and D. S. Bassett, "A neural machine code and programming framework for the reservoir computer," *Nature Machine Intelligence*, vol. 5, no. 6, pp. 622–630, 2023.

[162] L. M. Smith, J. Z. Kim, Z. Lu, and D. S. Bassett, "Learning continuous chaotic attractors with a reservoir computer," *Chaos*, vol. 32, no. 1, 2022.

[163] M. Lukoševičius, H. Jaeger, and B. Schrauwen, "Reservoir computing trends," *KI-Künstliche Intelligenz*, vol. 26, pp. 365–371, 2012.

[164] W. Maass, T. Natschläger, and H. Markram, "Real-time computing without stable states: A new framework for neural computation based on perturbations," *Neural Computation*, vol. 14, no. 11, pp. 2531–2560, 2002.

[165] F. Damicelli, C. C. Hilgetag, and A. Goulas, "Brain connectivity meets reservoir computing," *PLoS Computational Biology*, vol. 18, no. 11, p. e1010639, 2022.

[166] P. Enel, E. Procyk, R. Quilodran, and P. F. Dominey, "Reservoir computing properties of neural dynamics in prefrontal cortex," *PLoS Computational Biology*, vol. 12, no. 6, p. e1004967, 2016.

[167] C. Merkel, Q. Saleh, C. Donahue, and D. Kudithipudi, "Memristive reservoir computing architecture for epileptic seizure detection," *Procedia Computer Science*, vol. 41, pp. 249–254, 2014.

[168] D. Canaday, A. Pomerance, and D. J. Gauthier, "Model-free control of dynamical systems with deep reservoir computing," *Journal of Physics: Complexity*, vol. 2, no. 3, p. 035025, 2021.

[169] D. J. Gauthier, E. Bollt, A. Griffith, and W. A. S. Barbosa, "Next generation reservoir computing," *Nature Communications*, vol. 12, no. 1, p. 5564, 2021.

[170] R. Kent, W. A. S. Barbosa, and D. J. Gauthier, "Controlling chaotic maps using next-generation reservoir computing," *Chaos*, vol. 34, no. 2, 2024.

[171] M. Lukoševičius and H. Jaeger, "Reservoir computing approaches to recurrent neural network training," *Computer Science Review*, vol. 3, no. 3, pp. 127–149, 2009.

[172] G. Tanaka, T. Yamane, J. B. Héroux, R. Nakane, N. Kanazawa, S. Takeda, H. Numata, D. Nakano, and A. Hirose, "Recent advances in physical reservoir computing: A review," *Neural Networks*, vol. 115, pp. 100–123, 2019.

[173] S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, "Highly nonrandom features of synaptic connectivity in local cortical circuits," *PLOS Biology*, vol. 3, no. 3, p. e68, 2005.

[174] U. S. Bhalla, "How to record a million synaptic weights in a hippocampal slice," *PLoS Computational Biology*, vol. 4, no. 6, p. e1000098, 2008.

[175] I. B. Yildiz, H. Jaeger, and S. J. Kiebel, "Re-visiting the echo state property," *Neural Networks*, vol. 35, pp. 1–9, 2012.

[176] E. Bollt, "On explaining the surprising success of reservoir computing forecaster of chaos? The universal machine learning dynamical system with contrast to VAR and DMD," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 31, no. 1, 2021.

[177] S. Haeusler and W. Maass, "A statistical analysis of information-processing properties of lamina-specific cortical microcircuit models," *Cerebral Cortex*, vol. 17, no. 1, pp. 149–162, 2007.

[178] Y. Xue, L. Yang, and S. Haykin, "Decoupled echo state networks with lateral inhibition," *Neural Networks*, vol. 20, no. 3, pp. 365–376, 2007.

[179] J. T. Serences and S. Kastner, "A multi-level account of selective attention," *The Oxford Handbook of Attention*, p. 76, 2014.

[180] J. W. Sander and S. D. Shorvon, "Epidemiology of the epilepsies," *Journal of Neurology, Neurosurgery, and Psychiatry*, vol. 61, no. 5, p. 433, 1996.

[181] M. Abouelleil, N. Deshpande, and R. Ali, "Emerging trends in neuromodulation for treatment of drug-resistant epilepsy," *Frontiers in Pain Research*, vol. 3, p. 839463, 2022.

[182] P. Buteneers, D. Verstraeten, P. van Mierlo, T. Wyckhuys, D. Stroobandt, R. Raedt, H. Hallez, and B. Schrauwen, "Automatic detection of epileptic seizures on the intracranial electroencephalogram of rats using reservoir computing," *Artificial Intelligence in Medicine*, vol. 53, no. 3, pp. 215–223, 2011.

[183] A. H. Shoeb, *Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment*. PhD thesis, Massachusetts Institute of Technology, 2009.

[184] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, and H. E. Stanley, "Physionet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[185] P. Detti, G. Z. M. de Lara, R. Bruni, M. Pranzo, F. Sarnari, and G. Vatti, "A patient-specific approach for short-term epileptic seizures prediction through the analysis of EEG synchronization," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 6, pp. 1494–1504, 2018.

[186] M. Vinck, R. Oostenveld, M. van Wingerden, F. Battaglia, and C. M. A. Pennartz, "An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias," *Neuroimage*, vol. 55, no. 4, pp. 1548–1565, 2011.

[187] R. Cherian and E. G. Kanaga, "Theoretical and methodological analysis of EEG based seizure detection and prediction: An exhaustive review," *Journal of Neuroscience Methods*, vol. 369, p. 109483, 2022.

[188] A. Zygmund, *Trigonometric Series*, vol. 1. Cambridge University Press, 2002.

[189] K. K. Majumdar and P. Vardhan, "Automatic seizure detection in ECoG by differential operator and windowed variance," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 4, pp. 356–365, 2011.

[190] P. Meritam Larsen, S. Wüstenhagen, D. Terney, E. Gardella, H. Aurlien, and S. Beniczky, "Duration of epileptic seizure types: A data-driven approach," *Epilepsia*, vol. 64, no. 2, pp. 469–478, 2023.

[191] C. Zrenner, D. Galevska, J. O. Nieminen, D. Baur, M. Stefanou, and U. Ziemann, "The shaky ground truth of real-time phase estimation," *Neuroimage*, vol. 214, p. 116761, 2020.

[192] M. Rosenblum, A. Pikovsky, A. A. Kühn, and J. L. Busch, "Real-time estimation of phase and amplitude with application to neural data," *Scientific Reports*, vol. 11, no. 1, p. 18037, 2021.

[193] J. R. McIntosh and P. Sajda, "Estimation of phase in EEG rhythms for real-time applications," *Journal of Neural Engineering*, vol. 17, no. 3, p. 034002, 2020.

[194] A. Shakeel, T. Tanaka, and K. Kitajo, "Time-series prediction of the oscillatory phase of EEG signals using the least mean square algorithm-based AR model," *Applied Sciences*, vol. 10, no. 10, p. 3616, 2020.