Department of Electrical & Electronic Engineering
The University of Melbourne

National ICT Australia (NICTA)

Mathematical Modeling of Brain Networks: From Synaptic Plasticity to Behavior

Robert R. Kerr

Submitted in total fulfillment of
the requirements of the degree of

Doctor of Philosophy

Produced on archival quality paper

June, 2014

Abstract

A fundamental goal of neuroscience is to understand how the brain encodes and processes information and how the networks and structures involved are formed. In this thesis, we use theoretical approaches to further our understanding of brain function. First, we investigate how experimentally-based learning rules lead to the formation of different network structures, through unsupervised learning. Second, we investigate how different experimentally-based neural models and network structures enable different types of information processing, such as goal-directed, top-down processing. Third, we consider how reinforcement learning arising from synaptic plasticity mechanisms can coexist with unsupervised learning during the operant conditioning of neural firing rates.

The unsupervised learning rule spiking-timing-dependent plasticity (STDP) has been shown to selectively potentiate feed-forward connections with specific axonal delays, enabling functions such as sound localization in the auditory brainstem of the barn owl. We demonstrate a similar selective potentiation for the recurrent connections in a network with axonal delays corresponding to the period of incoming oscillatory activity with frequencies in the range of 100-300Hz. For lower frequency oscillations, such as gamma (60Hz), we show that multiple, recurrently connected groups of neurons could encode not only the oscillation frequency but also a time lag between different sets of oscillations. These results have the potential to help explain missing fundamental pitch perception in the auditory brainstem and the formation of neuronal ensembles (or cell assemblies) in the cortex, respectively.

Neural systems are able to perform top-down processing of stimulus information and flexibly select behaviors appropriate to the environment and present goals. Based upon previous experimental and theoretical studies, we propose that information in higher-level areas of the cortex, such as the prefrontal cortex, is encoded in the amplitude and phase of neural oscillations, such as gamma oscillations, and that this activity is gated by two mechanisms: top-down feedback and coherence between these oscillations. By forming these units into circuits that can perform logic operations, we identify the different ways in which operations can be initiated and manipulated by top-down feedback. We demonstrate that more sophisticated and flexible top-down control is possible when the gain of units is modulated by two mechanisms. We explore how different network properties affect top-down control and make predictions about the likely connectivities between certain brain regions.

Typical and well-studied examples of behavioral learning are those in which the firing rates of individual cortical neurons in monkeys are increased using rewards. These results have been reproduced using reinforcement learning rules, such as a variant of STDP called reward-modulated spike-timing-dependent plasticity (RSTDP). However, these previous models have assumed that no unsupervised learning is present (i.e., no learning occurs without, or independent of, rewards). We show that these models cannot elicit firing rate reinforcement while exhibiting both reward learning and ongoing, stable unsupervised learning. To address this issue, we propose a new RSTDP model of synaptic plasticity, based upon the observed effects that dopamine has on long-term potentiation and depression, that is able to exhibit unsupervised learning and lead to firing rate reinforcement.

Acknowledgements

I would like to thank my supervisors Anthony Burkitt, David Grayden, and Doreen Thomas for their strong guidance, support and encouragement, and for all the time that they found to meet with me and give me feedback on my work.

I would like to thank my collaborator and unofficial supervisor Matthieu Gilson who gave me invaluable feedback and direction, almost entirely over email and via instant messaging and often from a sailing boat. I would also like to thank him for hosting my visit to Tokyo, where he provided me with a place to stay and introduced me to his colleagues and friends as well as karaoke and onsen.

I would like to thank my colleagues and friends in the NeuroEngineering Laboratory and in the CfNE building for discussion, comradeship, distractions, and generally making the lab a pleasant place to come into and work each day. These people include: Isabell Kiral-Kornek, Martin Spencer, Michael Eager, Dean Freestone, Andre Peterson, Kyle Slater, Farhad Goodarzy, Roland Diggelmann, Kerry Halupka, Elma O'Sullivan-Greene, Colin Hales, William Kentler, Nina Erfanian Saeedi, Paulius Stepanas, Emily O'Brien, Andrea Varsavsky, Craig Savage, Tatiana Kameneva, Bahman Tahayori, Kelvin Layton, Matias Maturana, Parvin Zarei, and Richard Balson.

I would also like to thank the members of my advisory panel Hamish Meffin and Leigh Johnston for their encouragement and support. And many thanks to Tracy Painter for all of her help with organizing my travel and for always being eager to help me with any administration problems I had.

My research was made possible by the financial support I received from my Australian Postgraduate Award, an Australian Research Council Discovery Project, and the Australian Federal and Victorian State Governments and the Australian Research Council through the ICT Centre of Excellence program, National ICT Australia (NICTA). I would like to thank the latter two for supporting my travel to a number of local and international conferences during my PhD. I would also like to thank the VLSCI for providing me with valuable computing resources, as well as for helpfully solving any issues I encountered with using these resources.

I would like to thank Geoff Goodhill who introduced me to, and inspired in me an interest in, computational neuroscience. I would also like to thank him for his continued support and guidance.

I would like to thank Gemma Tiernan for all of her support and patience. I would also like to thank her for convincing me to move down from Brisbane and out of my comfort zone in order to pursue a PhD.

Finally, I would like to thank my family for their continuing support and encouragement, as well as for keeping me grounded with frequent queries about when I was going to get a real job.

Publications

The chapters of this thesis have either been published, accepted, or are under review as journals papers:

- **Chapter 2:** Kerr RR, Burkitt AN, Thomas DA, Gilson M, and Grayden DB. (2013) "Delay Selection by Spike-Timing-Dependent Plasticity in Recurrent Networks of Spiking Neurons Receiving Oscillatory Inputs." PLoS Comput Biol 9(2): e1002897.
- **Chapter 3:** Kerr RR, Grayden DB, Thomas DA, Gilson M, and Burkitt AN. "Goal-directed control with cortical units that are gated by both top-down feedback and oscillatory coherence." Front Neural Circuits (under-review).
- **Chapter 4:** Kerr RR, Grayden DB, Thomas DA, Gilson M, and Burkitt AN. (2014) "Coexistence of Reward and Unsupervised Learning during the Operant Conditioning of Neural Firing Rates." PLoS ONE 9(1): e87123.

A number of refereed conference abstracts have also been published:

- **Chapter 2:** Kerr RR, Burkitt AN, Thomas DA, and Grayden DB. (2012) "STDP encodes oscillation frequencies in the connections of recurrent networks of spiking neurons." CNS*2012 Computational Neuroscience Meeting, BMC Neuroscience 2012, 13(Suppl 1), P130.
- **Chapter 4:** Kerr RR, Grayden DB, Thomas DA, Gilson M, and Burkitt AN. (2013) "Generalized reward-modulated STDP: a model of operant conditioning of cortical neurons." Cosyne Abstracts 2013, I-91, Salt Lake City USA.
- **Chapter 4:** Kerr RR, Grayden DB, Thomas DA, Gilson M, and Burkitt AN. (2013) "Requirements for the robust operant conditioning of neural firing rates." CNS*2013 Computational Neuroscience Meeting, BMC Neuroscience 2013, 14(Suppl 1), P48.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The field of neuroscience ambitiously aims to understand what is arguably the most complex and sophisticated system that we know of. Our brains are able to perform and learn a range of complex cognitive processes, including object identification, problem solving, and decision making. Even species with much smaller brains, for example honeybees, are able to exhibit many of the same abilities as us, such as navigation, short-term memory, and communication (Hammer and Menzel, 1995; Menzel and Giurfa, 2001).

During development, our brains learn how to detect and comprehend basic stimulus features, such as lines in our visual world or the pitches of sounds in our auditory world. These developments are mostly permanent and provide a foundation for further learning. From this, we are able to learn novel stimuli and recognize them again in the future. We learn how best to respond to these stimuli, based on the feedback we receive from the outside world regarding our behaviors. While this learning leads to long-term changes, we combine it with short-term or working memory in order to rapidly adjust this response depending on our recent experiences or on our current motivations or goals.

The biological mechanisms underlying these cognitive processes remain largely a mystery. An improved understanding of these mechanisms would lead to better techniques for preventing and treating neurological disorders, where some aspect of these mechanisms has been disrupted. Uncovering and understanding the biological mechanisms within the brain could also inspire technological advancements in many other areas, such as computing and robotics. In this thesis, we work toward an understanding of these mechanisms by proposing testable hypotheses for how certain biological processes lead to cognitive processes and, ultimately, to behavior.

1.1.1 Neurons

Neurons are thought to be the fundamental unit of computation in the brain. There are estimated to be on the order of 100 billion neurons in the human brain (Williams and Herrup, 1988), 100 000 times more than in a honeybee's brain (Menzel and Giurfa, 2001). They form networks through which they communicate via electrical bursts, known as action potentials or **spikes**. These spikes originate at the cell body, or **soma**, and travel along nerve fibers, called **axons**, which carry the spikes to other neurons. Axons connect to other neurons via **synapses**. Here the electrical activity in the axon elicits an electrical response in receiving fibers of the target neuron, known as **dendrites**, which in turn carry the electrical signals to the target neurons' somas. Neurons can either be **excitatory** or **inhibitory** depending on the types of synapses they make onto other neurons. Spikes from excitatory neurons elicit excitatory responses in their target neurons, where they increase the **membrane potential** (the difference in electrical potential of the inside of the cell to the outside) and, therefore, the chance of the neurons producing spikes. Inhibitory neurons, conversely, inhibit or suppress the firing of other neurons by decreasing the membrane potential. Inhibitory activity in the brain is required to prevent excitatory activity from building up and saturating. The change elicited in the membrane potential of a neuron due to an incoming spike from an excitatory or inhibitory neuron is referred to as the excitatory post-synaptic potential (EPSP) or inhibitory post-synaptic potential (IPSP), respectively.

1.1.2 Neural Activity and Codes

It is generally understood to be the firing patterns of neurons that carry the information being processed by the brain, from sensory information to information concerning decisions and actions to be performed. To first order, the firing of a single neuron is characterized by its firing rate. However, for a given firing rate, neurons can have significantly different firing statistics. At one extreme, they can fire regularly with little variance in the length of time between spikes (the inter-spike interval or ISI), while, at the other extreme, neurons can exhibit irregular firing where there is a large variance in their ISI and bursts of spikes are often exhibited. These two firing regimes are referred to as integration-driven and fluctuation-driven, respectively. Shu et al. (2003); Haider et al. (2006) observed that neurons in the cortex appear to operate in a fluctuation-driven regime, where the firing of neurons is driven by fluctuations in the balance of their excitatory and inhibitory inputs.

The firing between any two neurons may be independent or their firing may be correlated. These correlations have been observed throughout the brain and it is likely that, to some extent, they are functionally relevant to how information is encoded. If this is the case, then the activity of neurons cannot be decoded independently and instead the firing of many neurons together encodes information. Correlations between neurons can exist on a range of timescales, from correlations between the firing rates of neurons to correlations in the precise timing of spikes. Synchrony refers to a specific type of correlation where multiple neurons fire together. Oscillations refer to synchrony that

occurs in a rhythmic fashion with a particular frequency. Synchrony and oscillations are ubiquitous in the cortex. Internally generated oscillations with frequencies in a number of different bands have been observed throughout the cortex (Buzsáki and Draguhn, 2004; Wang, 2010). Gamma oscillations, in particular, have been shown to be important in higher brain functions (Bartos et al., 2007; Fries et al., 2007), such as (selective) attention (Womelsdorf and Fries, 2007) and top-down processing (Engel et al., 2001). In addition to these inherent brain oscillations, oscillations that are directly associated with a stimulus, such as those corresponding to the pitch of a sound, are observed in certain brain regions (e.g., auditory brainstem).

How information is encoded in neural activity in the brain is referred to as a neural code. Understanding this encoding requires consideration of how this information is becomes encoded, how it is decoded or utilized for behavior, and how it is processed in between. Different types of neural codes use different aspects of neural activity to carry information. There are two main types of neural code, the simpler of these is the rate code, where the frequency of spikes produced by different neurons, or their firing rates, carries all of the information, while the precise timing of the neuron's spikes carries no information. Contrary to this is the spike timing code, where the precise timing of spikes does carry information.

London et al. (2010) argued that the sensitivity of neural systems to perturbations of a single spike implies the use of rate codes in the cortex and that, unless they generate very large and fast depolarizing events, spike timing codes would not be possible. However, a number of such neural codes have been proposed, including synfire chains and polychronization. Synfire chains refer to multi-layer feed-forward networks (or diverging/converging chains of links), which allows synchronous propagation of action potentials (Abeles, 1991). For this synchronous propagation of activity to occur, these feed-forward networks must have equal (or almost equal) propagation delays on all connections between the two layers. Polychronization proposes an alternative to this where networks of neurons exhibit recurring firing patterns of time-locked but non-synchronous spikes (Izhikevich et al., 2004; Izhikevich, 2006). Each of these time-locked firing patterns is referred to as a polychronous group, and, similar to synfire chains, activity propagates through the coincident arrival of spikes at each neuron. Just as synfire chains require equal propagation delays between layers, polychronous networks require specific sets of connections and connection delays between the groups. Experimentally, it is difficult to find evidence supporting or disproving either synfire chains or polychronous groups because it is not clear which neurons may be part of the same synfire chain layer or polychronous group as the neurons may not be physically near each other. This difficulty is increased by the fact that neurons could belong to multiple synfire chains or polychronous groups.

1.1.3 Networks of Neurons

The ability for neural systems to encode, process, and decode information and to perform complex tasks is due to the specific connections between the neurons. Neurons in the

brain are connected together by synapses, forming neural networks. There are on average around 10 000 synapses from and onto each neuron in the human brain (Huttenlocher, 1990).

The cerebral cortex is a region of neural tissue unique to mammalian brains. It is made up of the neocortex, the hippocampus and the olfactory bulb. For humans, the largest of these is the neocortex, which by volume makes up approximately 80% of the entire human brain. It is the most recently evolved part of our brains with its volume representing the largest evolutionary expansion from our primate ancestors (Shepherd, 2004). The cortex plays a major role in high-level cognitive function like memory, thought, language, attention, awareness and consciousness. Anatomically, it is a sheet of cells only 2mm thick, which is considerably folded in larger mammals. This sheet is often divided into six different layers, each with a slightly different structure. There are approximately 50 000 neurons per cubic millimeter, 70% of which are pyramidal neurons (Shepherd, 2004). These pyramidal neurons form excitatory connections within the neocortex and to many other areas of the brain. Inhibitory neurons make up a large proportion of the remaining 30%.

Information travels in a particular direction through neuronal networks in the brain, from sensory inputs and, ultimately, to motor outputs. Connections that carry activity in this direction are generally referred to as **feedforward** connections, while those in the opposite direction are referred to as **feedback** connections. There are also lateral or **recurrent** connections, which are not in either direction but between neurons of a similar level or progression through the brain. In the cortex, a large fraction of the connections onto neurons come from other neurons in the cortex, making it highly recurrent in its connectivity (Martin, 2002; Douglas and Martin, 2007). A large proportion of these cortico-cortical connections are local and there is some evidence that, at least functionally, cortical neurons are organized into local groupings, called cortical columns, which extend from one side of the sheet to the other (Hubel and Wiesel, 1977; da Costa and Martin, 2010). In fact, anatomically, despite it's size, the cortex appear relatively homogeneous. This gives hope that there are common structures, circuits, and organization principles, which occur throughout different cortical areas, that can be uncovered.

1.1.4 Synaptic Plasticity

Learning is required in order to develop and adapt the cognitive abilities of the brain. It is predominately thought to occur through synaptic plasticity, the changing of the synaptic efficacies (or synaptic strengths), through which neural structures (networks of neurons) can be formed and altered. While not experimentally observed until much later (Bliss and Lomo, 1973), Hebb postulated that the efficacy of a synapse increased when the pre-synaptic neuron consistently caused the firing in the post-synaptic neuron (Hebb, 1949). Experimentally, this activity-dependent plasticity has become referred to as long-term potentiation (LTP), while the opposite phenomena, where synaptic efficacy decreases is referred to as long-term depression (LTD). Spike-timing-dependent plasticity (STDP) is

an experimentally observed learning rule that includes both LTP and LTD and that is based on the relative timing of pre- and post-synaptic spikes (Markram et al., 1997; Bi and Poo, 1998; Dan and Poo, 2004, 2006). STDP was first proposed theoretically by Gerstner et al. (1996) as an unsupervised Hebbian learning rule that could select feed-forward connections with specific axonal delays. STDP has subsequently been the focus of many theoretical and computation study of learning in networks of neurons.

In general, the learning due to synaptic plasticity can be separated into three types or components: **unsupervised**, **supervised**, and **reinforcement** learning (Barlow, 1989; Pfister et al., 2006; Frémaux et al., 2010). STDP is an example of an unsupervised learning rule, whereas supervised and reinforcement learning differ from unsupervised learning in that they depend upon an external supervisor and a reinforcement signal, respectively. Because of this, the latter types of learning can depend on the behavior of the system, agent, or animal. In this thesis, we focus on unsupervised learning and reinforcement learning.

1.2 Mathematical Models

Models allow us to formally and unambiguously describe our conception of how a system functions. Models that produces behaviors that can be compared to the actual system allow us to test whether we correctly understand how that system works.

1.2.1 Neuron Models

Neurons can be modeled in many different ways with many different degrees of complexity. Simple models use a single compartment, while more detailed models divide the cell into multiple compartments. For studies of networks of neurons, neurons are generally modeled as point neurons, which disregard any spatial aspect of the cell, and as spiking neurons, which output instantaneous impulses that represent spikes. Typically in these types of models there is a variable representing the membrane potential of the cell. Examples of such models are provided below.

Poisson neuron model

The Poisson neuron model (Kempter et al., 1999) is a stochastic model which outputs a spike train that is a realization of an inhomogeneous Poisson process with an intensity function. This intensity function is loosely analogous to the membrane potential of the neuron. It is made up of a spontaneous rate and the weighted sum of post-synaptic response kernels given by

$$\lambda_i(t) = \nu_0 + \sum_{j \neq i} K_{ij}(t) \sum_n \epsilon(t - t_{j,n} - d_{ij}), \tag{1.2.1}$$

where $\lambda_i(t)$ is the intensity function for the ith neuron at time t, ν_0 is the spontaneous rate, $K_{ij}(t)$ is the synaptic weight from neuron j to neuron i, $\epsilon(t)$ is the post-synaptic response kernel, $t_{j,n}$ is the time of the nth spike output by neuron j, and d_{ij} is the delay (comprising both dendritic and axonal delays) from neuron j to neuron i. Synapses here are modeled as current based. This means that synaptic input into the neuron is independent of the neuron's membrane potential (in this model, the intensity function).

Integrate-and-fire neuron model

Integrate-and-fire (IF) neurons, first introduced by Lapicque (1907), are used extensively as a simple spiking neuron model. Unlike the Poisson neuron, this is a deterministic model that integrates synaptic inputs into a membrane potential variable, resulting in the outputting of a spike each time it reaches a threshold value, followed by the resetting of the membrane potential variable. Typically, the membrane potential also decays to a equilibrium value and these models are referred to as leaky integrate-and-fire (LIF) neurons (Burkitt, 2006a,b). A typical differential equation describing the LIF model is

$$\frac{dV_i(t)}{dt} = \frac{1}{\tau_m}\Big[V_r - V_i(t) - \sum_{j\neq i} K_{ij}(t)\big(V_i(t) - E_{S,j}\big)\sum_n \epsilon(t - t_{j,n} - d_{ij})\Big], \tag{1.2.2}$$

where $V_i(t)$ is the membrane potential of the neuron, τ_m is the membrane time constant, V_r is the resting potential of the neuron, $E_{S,j}$ is the reversal potential for the synapse from neuron j, and the other parameters ($K_{ij}(t)$, $t_{j,n}$, d_{ij}, and $\epsilon(t)$) are the same as in the Poisson model. If $V_i(t)$ rises above the threshold potential, V_{th}, a spike is output and the potential set to the reset potential, V_{reset}. Additionally these models can have a refractory period, which is a set period of time after a spike is fired where the membrane potential is kept fixed. Synapses of LIF neurons are modeled as either current or conductance based. Current based models, as in the Poisson model, have a synaptic current into the neuron that is independent of the neuron's membrane potential. In conductance based models, however, the synaptic current is dependent on the neuron's membrane potential (see Equation 1.2.2).

Izhikevich neuron model

Izhikevich neurons extend LIF neurons by incorporating a refractory variable in addition to the membrane potential (Izhikevich, 2003). The state of each neuron is described by two differential equations, one for the membrane potential and the second one for the refractory variable. This adds complexity to the model but allows it to exhibit a wide variety of behaviors that have been experimentally observed. Like the LIF neuron, synapses can be either current or conductance based.

Hodgkin-Huxley neuron model

Unlike the previous models, the Hodgkin-Huxley model (Hodgkin and Huxley, 1952) does not model spikes as instantaneous pulses, but instead realistically models the fast depolarizing and subsequent repolarizing of the membrane potential during and after spiking, respectively. Unlike the other models we have described, the Hodgkin-Huxley model is suitable for multi-compartment modeling. The model typically uses four differential equations to explicitly describe not just the membrane potential but also the dynamics of synaptic gating variables. Hodgkin-Huxley neurons are used when a high level of accuracy is desired of the neuron's behavior; however, this requires careful calibration of the many parameters.

1.2.2 Synaptic Plasticity Models

Alongside models of neurons, exist models of how the synaptic connections between these neurons change. Models of activity-dependent synaptic plasticity can be divided into phenomenological and biophysical models. Phenomenological models of synaptic plasticity describe the changes that occur due to the neural activity but do so without considering the underlying biological mechanisms (Morrison et al., 2008). Because of this they are generally simpler and more analytically tractable, making them useful for analytical and simulation studies. Biophysical models on the other hand are based upon possible underlying mechanisms of the cells.

Spike-timing-dependent Plasticity (STDP)

Spike-timing-dependent plasticity (STDP) is a learning rule used to update synaptic weights that depends on the timing between pre- and post-synaptic spikes. A Hebbian learning rule dependent on this spike timing difference was first proposed as a mechanism for the precision achieved by barn owls in using inter-aural time differences to carry out azimuthal sound localization (Gerstner et al., 1996). Experimental observations confirming the existence of this mechanism were published soon after (Markram et al., 1997). Since then, STDP has been observed to govern the changes in synaptic weights in various regions of the brain (Bi and Poo, 1998; Dan and Poo, 2004, 2006) with a variety of learning windows observed (Caporale and Dan, 2008).

A general form of the STDP learning rule, as given by Gilson et al. (2010a), is

$$\Delta K = \eta \begin{cases} \omega_{in}, & \text{for each pre-synaptic spike} \\ \omega_{out}, & \text{for each post-synaptic spike} \\ W(\Delta t, K), & \text{for each pair of spikes (where } |\Delta t| \text{ is sufficiently small).} \end{cases} \tag{1.2.3}$$

Here, K is the synaptic weight, ΔK is the change in synaptic weight, η is the learning rate, and $\Delta t = t_{in} - t_{out}$ is the time of the pre-synaptic spike minus the time of the post-synaptic spike. ω_{in} and ω_{out} are homeostatic terms that change the synaptic weight every

time the pre- or post-synaptic neuron fires, respectively. These are not considered (i.e. they are set to zero) in many studies. The learning window, $W(t, K)$, that is typically used is of the form

$$W(t,K) = \begin{cases} f_+(K)e^{\frac{t}{\tau_+}}, & t < 0 \\ 0, & t = 0 \\ f_-(K)e^{\frac{-t}{\tau_-}}, & t > 0. \end{cases} \tag{1.2.4}$$

Early models of STDP assumed that the level of potentiation and depression induced was independent of the current strength of the synapse. This type of model is referred to as additive. For an additive STDP rule, f_+ and f_- are constants. The simplest model of weight dependent STDP is multiplicative STDP, where synaptic weights are changed by a fraction of their current strength instead of by a constant amount. Various other models of STDP exist, many of which model potentiation and depression differently from each other. A commonly used weight dependence model is the one proposed by Gütig et al. (2003), which allows interpolation between additive and multiplicative models resulting in a model that is a mix of the two. In this model, the weight-dependent amplitude of the potentiation and depression in the Gütig et al. (2003) model are, respectively,

$$\begin{aligned} f_+(K) &= (K_{max} - K)^\mu, \\ f_-(K) &= -\alpha K^\mu, \end{aligned} \tag{1.2.5}$$

where the weight, K, is allowed to take on values between 0 and K_{max}, and the parameter α controls the balance between potentiation and depression. The other parameter, μ, controls where the model lies between the additive and multiplicative models, with a value of 0 resulting in the additive model and a value of 1 resulting in the multiplicative model. Gilson and Fukai (2011) suggest another weight dependence model, where LTD has a logarithmic weight dependence. This produces the type of long-tailed synaptic weight distribution observed experimentally and provides a good balance between synaptic competition and stability.

Triplet STDP models

Since STDP was first proposed and observed experimentally, a number of other experimental observations have been made that STDP is unable to explain (Bi and Wang, 2002; Froemke and Dan, 2002; Wang et al., 2005; Froemke et al., 2006). In order to account for these observations, variants of STDP, such as triplet STDP, have been proposed (Sjöström et al., 2001; Froemke and Dan, 2002; Froemke et al., 2006; Appleby and Elliott, 2007). In the triplet STDP model, it is no longer sufficient to consider only pairs of spikes but instead the timing between spikes in triplets determines the synaptic changes that occur. Because of this, triplet STDP can capture a number of experimental results that classical STDP cannot. Triplet STDP can be thought of as modifying the learning window depending on the pre- and post-synaptic firing rates as well as capturing higher than second order correlations (Pfister and Gerstner, 2006).

Biophysical synaptic plasticity models

Graupner and Brunel (2012) proposed a synaptic plasticity model based on postsynaptic calcium concentrations of cells. This biophysically based model, which is based upon earlier models (Shouval et al., 2002, 2010), is able to exhibit the results of many plasticity experiments relating to different STDP windows, pairing with postsynaptic spikes and bursts, triplet and quadruplet STDP, firing rate effects, and the effects of dendritic location.

Reward-modulated STDP

Reward-modulated STDP is a variant of STDP in which synaptic changes depend not only on the activity of the pre- and post-synaptic neurons but also on a reward signal that is based on some behavior of the system (Izhikevich, 2007; Farries and Fairhall, 2007; Florian, 2007). This reward signal typically represents the extracellular concentration of a neuromodulator, such as dopamine, in that area of the brain that depends upon rewards or punishments received. In this model, the synaptic changes that STDP would elicit are not immediately applied but are instead stored or remembered in a synaptic or eligibility "trace", which decays over time. Changes are made to the weights of synapses by integrating the product of this synaptic trace with a reward signal, as given by

$$\Delta K(t) = \eta \int_{t}^{t+\Delta t} e(t')y(t')dt', \tag{1.2.6}$$

where η is the learning rate, $y(t)$ is the reward signal, and $e(t)$ is the eligibility trace. The eligibility trace is made up the synaptic changes that STDP would have made due to different pre- post-synaptic spike paris. These synaptic changes are only then implemented to the extent that the system is subsequently rewarded.

The RSTDP is based on the experimental evidence that neuromodulators affect the synaptic changes due to STDP. A number of experiments have considered the effect of dopamine, which is strongly linked to reinforcement learning in the brain (Schultz et al., 1997), on STDP (Bao et al., 2001; Reynolds et al., 2001; Reynolds and Wickens, 2002; Pawlak and Kerr, 2008; Zhang et al., 2009). In addition to dopamine, other neuromodulators have been observed to affect STDP, including: acetycholine (Sugisaki et al., 2011), octopamine (Cassenaer and Laurent, 2012), and norepinephrine (Salgado et al., 2012). In many of these experiments, the effect of the neuromodulator on STDP has been non-linear and, despite the way reward modulates learning in the RSTDP model, the presence of the neuromodulator is not necessarily required for LTP and LTD to occur.

R-max model

Other types of synaptic plasticity models that are used for reinforcement learning are those that have been derived theoretically to maximize the received reward (Seung, 2003; Xie

and Seung, 2004; Pfister et al., 2006; Florian, 2007), such as the R-max model (Frémaux et al., 2010). As discussed by Frémaux et al. (2010), the average change in synaptic weights due to reinforcement learning rules can be split into the unsupervised and reward learning components. The reward learning component depends on the covariance between neural activity and reward, while the unsupervised learning component is independent of this covariance, depending only the mean reward value. To maximize the reward that the system receives the unsupervised component needs to be as small as possible. Unlike RSTDP, the unsupervised component (or bias) is always zero in the R-max model. This is only possible because an assumption of the R-max model is that it has an unbiased estimator of the instantaneous firing rate of the post-synaptic neuron.

1.3 Theoretical Studies

Theoretical studies propose and develop models of how systems function in such a way that is consistent with existing experimental evidence, improves the understanding of a system, and makes testable predictions for future experimental work.

1.3.1 Unsupervised Learning Due to Synaptic Plasticity

Studies of networks of neurons subject to STDP learning rules began by analyzing the learning dynamics of a feed-forward network (Kempter et al., 1999). Analysis in this vein was extended to consider recurrent networks in a series of papers (Burkitt et al., 2007; Gilson et al., 2009a,b,c,d, 2010a,c,b). The general network structure used by these studies has a set of inputs that connect to neurons in a recurrent network. The inputs are considered to have constant intensity functions and have an input correlation structure that is either uncorrelated or delta-correlated. In both of these cases, the correlation structure either does not have a temporal dependence or the temporal dependence is trivial, in that there is only correlation for zero time-lag between the inputs. The inputs in these papers were often split into pools, where a pool could be assigned within-pool correlations (inputs within a pool were correlated) but there were no correlations between inputs of different pools. In this case, the correlation had a spatial dependence, in that the correlations were not identical for all pairs of inputs. A major finding from these papers was that a network with plastic recurrent connections, driven by plastic connections from two pools of correlated inputs, experiences symmetry breaking, where neurons specialize to just one pool of inputs and the network becomes split into two subnetworks.

Similar to these analytical studies, numerical studies by Song and Abbott (2001) looked at the effects of STDP when a network received inputs from a population of neurons with stimulus-response curves peaked about a preferred stimulus. STDP was shown, under certain conditions, to lead to both the formation and refinement of a map between the input neurons and the network neurons.

In the study by Gerstner et al. (1996), where STDP was first proposed, they con-

sidered the synaptic changes to the incoming connections of a neuron where correlations arose due to oscillations in the inputs. These correlations had an oscillatory temporal dependence and it was found that the STDP learning rule selected connections with specific axonal delays. These selected axonal delays differed from each other by multiples of the oscillation period and caused spikes to arrive at the post-synaptic neuron at the same time. The oscillations considered in this study had periods of the order of 500μs (frequency of the order of 2kHz), similar to the deviations in the range of axonal delays. This degree of temporal precision is suitable for the auditory system of barn owls, which use this mechanism to develop the circuits used to perform azimuthal sound localization.

Studies by Izhikevich and Edelman proposing polychronization showed through simulations that STDP selectively reinforced connections such that polychronous groups were created, as long as there was a sufficient spread of axonal conductance delays between recurrent connections in the network (Izhikevich et al., 2004; Izhikevich, 2006). Another simulation study showed that STDP can train a network that is presented with repetitions of an arbitrary Poisson-like input to respond with a single synchronous burst (Hosaka et al., 2008). These bursts can be thought of as the ignition of a synfire chain if the network is embedded in a larger feed-forward network. The timing of the synchrony was shown to depend on the network and on the input given. These studies highlight the ability for STDP to find and learn temporal patterns in the firing patterns of neurons.

1.3.2 Goal-directed Behavior and Neural Modulation

Goal-directed behavior is where the brain is able to perform fast switching between different "rules" that determine the appropriate response for a given stimulus. Wallis and Miller (2003) and Muhammad et al. (2006) performed such behavioral experiments, where a monkey was shown two successive visual stimuli and depending on a cue given to indicate the rule to apply, pull a lever if the stimuli matched or not. During this task, different neurons in the prefrontal (PRC), premotor (PMC), and inferior temporal (ITC) cortices and the striatum (STR) responded selectively to different parameters. This included the cue or task rule (the desired stimulus-response mapping), the behavioral response carried out, the visual stimulus being remembered, and whether or not the subsequent stimulus matched the remembered stimulus.

A theoretical study by Salinas examined how top-down gain modulation enabled networks of neurons to perform this type of fast stimulus-response remapping (Salinas, 2004). In this study, the behavior during goal-directed, visuomotor tasks was reproduced using different mappings between stimulus attributes, such as orientation and color, and responses, in the form of eye movements, to be performed and switched between. Their model contained two layers of neurons: a layer of gain modulated neurons and an output layer. Randomly chosen bottom-up feedforward and top-down feedback activities were assigned to each of the different stimuli and conditions (or rules), respectively, while the feedforward weights into the output neurons were determined such that they produced the intended responses. They showed that large changes in output response could be driven

by relatively small changes in the activities of the gain modulated neurons, which fit with experimental observations during attention tasks.

There are a number of different proposed mechanisms for top-down gain modulation, such as used by Salinas (2004). Amongst these is the hypothesis of Larkum (2013) that inputs to the apical dendrites, which tend to be feedback connections from higher-level cortical regions, modulate the gain of pyramidal neurons in the cortex to inputs to the soma and basal dendrites, which tend to be feedforward connections from lower-level regions. This is based on experiments by Larkum et al. (1999, 2001, 2004); Larkum (2013) where they observed that pyramidal neurons exhibited a much stronger response when they received inputs both to their soma (and basal dendrites) and to their apical dendrites than they did when they received only one of these types of inputs. They explained that this was due to a second spike initiation zone near the apical tuft of layer 5 pyramidal neurons, in addition to the one at the cell body that produces action potentials (sodium spikes) (Yuste et al., 1994; Schiller et al., 1997; Larkum and Zhu, 2002). This second initiation zone produces broad calcium spikes within the cell and its existence suggests that pyramidal neurons should be considered to have two functional compartments. In this way, the findings of Larkum et al. (1999, 2001, 2004); Larkum (2013) appear to invalidate single compartment neuron models, such as the LIF neuron model and suggest that at least two different sets of inputs need to be considered.

Another proposed mechanism is based on theoretical work that has shown that synchrony or coherence can act as a modulator of the gain of pyramidal neurons (Tiesinga et al., 2004; Börgers et al., 2005; Mishra et al., 2006; Tiesinga et al., 2008; Tiesinga and Sejnowski, 2009). This hypothesis is supported by experimental results and has been referred to as "communication-through-coherence" (Fries, 2005).

1.3.3 Reinforcement Learning Due to Synaptic Plasticity

A number of theoretical studies considered how reward-dependent synaptic plasticity models, such as RSTDP and R-max, can perform different reinforcement learning tasks and reproduce the results of behavioral experiments involving reinforcement learning (Izhikevich, 2007; Legenstein et al., 2008; Frémaux et al., 2010). Reward-modulated STDP provides a solution to the distal reward problem (also know as the credit assignment problem), which says that systems carrying out reinforcement learning often receive their reinforcement some time after the behavior deserving reinforcement, but somehow need to determine which of their recent behaviors should be reinforced (Izhikevich, 2007).

A behavioral experiment, which provides a simple example of reinforcement learning in the brain, showed that monkeys could learn to increase the firing of individual cortical neurons when given food or juice rewards for doing so (Fetz, 1969; Fetz and Baker, 1973). In these operant conditioning experiments, the monkey needed to be shown detailed visual feedback on the firing of the neuron. The monkey was also able to decrease the firing rate of a neuron (when its firing led to less rewards) and even increase the difference in firing rate between two nearby neurons. Legenstein et al. (2008) developed an analytical framework

for RSTDP and showed that it could reproduce the results of these operant conditioning experiments. In addition to this, they also investigated other learning tasks that RSTDP could perform, which were based on the timing of spikes.

Frémaux et al. (2010) compared the performance of the RSTDP and R-max models during reinforcement tasks. They showed that RSTDP was very sensitive to changes in the mean reward level (or dopamine concentration) as this would introduce a component of unsupervised learning, an "unsupervised bias". They showed that the R-max model was not susceptible to this as it cannot exhibit unsupervised learning, regardless of the mean reward value. For this reason, they claim that R-max but not RSTDP is suitable as a reinforcement learning rule in the brain because it requires a critic that is able to produce an unbiased estimator of the expected reward and subtract this from the actual reward received.

1.4 Overview of Thesis

This project aims to develop an understanding of the how the brain, and in particular the cortex, encodes and processes information and how the networks and structures involved are formed. This is carried out by proposing and analyzing mathematical models of neural systems, which are based on experimentally observed aspects of these systems. In order to develop understanding, sufficiently simple systems are considered or at least systems that can be analytically simplified and compared to simulations of the complete system.

In Chapter 2, we look at how unsupervised learning, due to STDP and within recurrently connected networks, can encode the frequency and phase of oscillations in neural activity. Specifically, we investigate how connections with certain propagation delays can be selectively potentiated in order to learn the frequency of oscillatory activity within the networks. We consider this selection for the synaptic connections within a recurrently connected group of neurons, where the propagation delays are entirely axonal and also where there is a range of axonal and dendritic propagation delays. We also investigate the axonal delay selection for connections between two groups of neurons with oscillatory activity, where the delay that is selected encodes both the oscillation frequency and the phase difference between the groups. In this chapter, we introduce the concepts of neural oscillations and activity-dependent synaptic plasticity, which are core to Chapters 3 and 4, respectively.

In Chapter 3, we build upon the ideas developed in Chapter 2 regarding the interactions between oscillating groups of neurons in the cortex. Rather than focussing on the synaptic plasticity of connections between these groups, we take a higher-level view and consider how top-down feedback can influence and control the way that these networks can process information and map stimuli to behaviors. We propose that sensory information in higher-levels of the cortex is encoded in the amplitude and phase of neural oscillations and that the propagation of this activity is gated by oscillatory, top-down feedback. We assume that the oscillation phase and propagation delays between cortical units cause

either reinforcement or interference, providing an additional gating (or gain modulation) mechanism to the top-down feedback. We show how this allows more sophisticated top-down control and manipulation of mappings in cortical networks and, therefore, more sophisticated goal-directed behavior.

In Chapter 4, we consider the case where, in addition to unsupervised learning, such as that considered in Chapter 2, there is reward or reinforcement learning present. We focus on the level of individual neurons and, in particular, a behavioral experiment in which a monkey, when presented with immediate feedback of the recorded firing of an individual cortical neuron, is able to increase its firing rate in order to receive food rewards. We investigate under what conditions a model of synaptic plasticity is able to exhibit both reward and unsupervised learning and also produce the experimentally observed reinforcement of neural firing rates. This is important to understand the ways that different types of learning can coexist in the brain and to understand how each of them are implemented at the level of individual synapses.

Chapter 2

Unsupervised Learning of Neural Oscillations

This chapter is a slightly modified version of the published article:

> *Kerr RR, Burkitt AN, Thomas DA, Gilson M, and Grayden DB. (2013) "Delay Selection by Spike-Timing-Dependent Plasticity in Recurrent Networks of Spiking Neurons Receiving Oscillatory Inputs." PLoS Comput Biol 9(2): e1002897.*

2.1 Abstract

Learning rules, such as spike-timing-dependent plasticity (STDP), change the structure of networks of neurons based on the firing activity. A network level understanding of these mechanisms can help infer how the brain learns patterns and processes information. Previous studies have shown that STDP selectively potentiates feed-forward connections that have specific axonal delays, and that this underlies behavioral functions such as sound localization in the auditory brainstem of the barn owl. In this chapter, we investigated how STDP leads to the selective potentiation of recurrent connections with different axonal and dendritic delays during oscillatory activity. We developed analytical models of learning with additive STDP in recurrent networks driven by oscillatory inputs, and supported the results using simulations with leaky integrate-and-fire neurons. Our results showed selective potentiation of connections with specific axonal delays, which depended on the input frequency. In addition, we demonstrated how this can lead to a network becoming selective in the amplitude of its oscillatory response to this frequency. We extended this model of axonal delay selection within a single recurrent network in two ways. First, we showed the selective potentiation of connections with a range of both axonal and dendritic delays. Second, we showed axonal delay selection between multiple groups receiving out-of-phase, oscillatory inputs. We discuss the application of these models to the formation

and activation of neuronal ensembles or cell assemblies in the cortex, and also to missing fundamental pitch perception in the auditory brainstem.

2.2 Introduction

Spike-timing-dependent plasticity (STDP) is an experimentally observed learning rule that changes synaptic strengths based on the relative timing of pre- and post-synaptic spikes (action potentials) (Markram et al., 1997; Bi and Poo, 1998; Dan and Poo, 2004, 2006). Gerstner et al. (1996) first proposed it as an unsupervised Hebbian learning rule that could select feed-forward connections with specific axonal delays. They showed that it could be used to achieve the high degree of temporal coherence that had been observed at frequencies of up to 8 kHz in the auditory brainstem of barn owls. This finding explained how a network could learn to perform sound localization using the time lag between the neural signals from the two ears. Their study also demonstrated that the precise timing of spikes could be captured by STDP and that this was sufficient to explain how neurons in the auditory pathway could learn to distinguish such fine temporal differences in an unsupervised fashion. In general, STDP has the ability to encode temporal correlations in neuronal activity, such as oscillations, into the functional structure of networks of neurons that have axonal and dendritic propagation delays.

The brain processes information through neuronal networks that contain specifically structured feed-forward and recurrent (lateral) connections. For example, only 5% of the input connections into cortical neurons are from the thalamus and, while these feed-forward connections tend to be strong, most of the remaining 95% are recurrent cortical connections (Martin, 2002; Douglas and Martin, 2007). For this reason, studies of neural learning that considered recurrent networks, rather than solely feed-forward networks, offered the possibility of providing new insight into how the brain processes and encodes information. While significant work has been carried out with learning in feed-forward networks (Kempter et al., 1999; Song and Abbott, 2001), it was only more recently that the same attention was paid to recurrent networks (Izhikevich, 2006; Burkitt et al., 2007; Morrison et al., 2007; Hosaka et al., 2008; Gilson et al., 2009a,b,c,d, 2010a; Kozloski and Cecchi, 2010).

Few analytical studies of spike-based learning in recurrent networks have been done, despite the ability for these studies to provide a more informative description of the mechanisms than studies that use simulations alone. A recent paper reviewed many of these studies (Gilson et al., 2010c). In one such analytical study, Gilson et al. (2009d) looked at the emergent structure that forms in recurrent networks due to STDP. They showed that spike correlations within two pools of inputs led to a form of symmetry breaking in the recurrent network receiving the inputs. Specifically, two sub-networks emerged with strong connections within the sub-networks but weak connections between them. In this way, the recurrent network encoded a spatial pattern of its inputs into its structure. The recurrent networks they considered contained only a narrow range of spike propagation delays. The inputs they considered contained instantaneous spike time correlations and

had firing rates that were constant in time.

Most inputs and activity in the brain are, however, not constant in time. Oscillations have been observed in many different regions of the brain, such as the cortex (Gray and Singer, 1989; Bragin et al., 1995; Buzsáki and Draguhn, 2004) and the auditory brainstem (Rose et al., 1967). In particular, gamma oscillations in the cortex have received considerable attention (Wang and Buzsáki, 1996; Bartos et al., 2007) and have been shown to play a role in attention, memory, and other cognitive functions (Gray and Singer, 1989; Jensen et al., 2007). For these reasons, it is important to consider the synaptic changes that occur due to these oscillations. Doing so may help elucidate the possible functions that oscillations play in cognitive processes.

A number of studies have explored the interaction between oscillatory activity and STDP using numerical simulations (Masquelier et al., 2009; Lee et al., 2009), but only few have performed analytical investigations. Pfister and Tass (2010) considered how STDP in recurrent networks can produce stable states of high and low synchrony (oscillations). They also examined how external stimulation can force the network out of a highly synchronous state into a state of lower synchrony. Muller et al. (2011) investigated how STDP can modify excitatory feed-forward connections into a single post-synaptic neuron such that it becomes phase-locked to oscillations in the inputs. Gilson et al. (2012) demonstrated a similar result for excitatory and inhibitory feed-forward connections with a range of dendritic delays. They further showed that the post-synaptic neuron became selective in its response to oscillatory inputs at the training frequency. These studies, however, did not consider networks that have a wide range of delays on the same timescale as the oscillation period, where the correlations due to the oscillations could drive delay selection. Gerstner et al. (1996) considered this situation for a specific neural system, but only for feed-forward connections and very high frequency oscillations. Though not specifically for oscillatory activity, further analysis has been performed for this concept of delay selection through STDP, although still only for feed-forward connections (Senn et al., 2002).

The broad question that motivated the work in this chapter was: what can be inferred about the ways that the brain learns patterns and processes information, given the role that STDP plays in determining network structure? We specifically aimed to address this for networks that have oscillatory firing patterns and a wide range of propagation delays, both axonal and dendritic. We investigated how additive STDP changes the strength of recurrent connections with a wide range of axonal delays and short dendritic delays when the network is driven by input spike trains that have oscillatory firing rates. We then looked at how these changes affect the oscillatory firing rate response of the network to inputs with different oscillation frequencies. We considered a range of oscillation frequencies from 100 to 300Hz. We discuss how this delay selection mechanism may suggest a possible explanation for how the auditory brainstem performs missing fundamental pitch perception.

We extended this simple situation and compared it to a network with a range of dendritic as well as axonal delays. We also extended the original model to one with multiple groups of neurons that were recurrently connected with connections that had a

range of axonal delays. In this case, the oscillatory inputs to each of the groups had the same frequency but were out of phase with each other. In both of these cases, we focused on frequencies in the gamma range (30-100Hz) (Jensen et al., 2007). We discuss how the second of these cases is relevant to the formation of oscillatory neuronal ensembles.

Throughout this chapter, we determined or estimated both the learning and dynamics of the networks analytically using the Poisson neuron model. We used numerical simulations with networks of leaky integrate-and-fire (LIF) neurons to support the results and conclusions. In the analysis and simulations, we considered only excitatory networks (i.e., without inhibition) to facilitate the mathematical analysis. We address the implications of this for the model in different contexts.

2.3 Methods

2.3.1 Poisson Neuron Model

Our analytical work used the Poisson neuron model (Kempter et al., 1999). This is a stochastic model which outputs a spike train that is a realization of an inhomogeneous Poisson process. The intensity function of this process is analogous to the membrane potential of the neuron. It is made up of a spontaneous rate and the weighted sum of post-synaptic response kernels given by

$$\lambda_i(t) = \nu_0 + \sum_{j \neq i} J_{ij}(t) \sum_n \epsilon(t - t_{j,n} - d_{ij}^{\text{ax}} - d_{ij}^{\text{den}}), \tag{2.3.1}$$

where $\lambda_i(t)$ is the intensity function for the ith neuron at time t, ν_0 is the spontaneous rate (assumed to be zero throughout this thesis), $J_{ij}(t)$ is the synaptic weight from neuron j to neuron i, $\epsilon(t)$ is the post-synaptic response kernel, or excitatory post-synaptic potential (EPSP) kernel, $t_{j,n}$ is the time of the nth spike output by neuron j, and d_{ij}^{ax} and d_{ij}^{den} are the axonal and dendritic delays, respectively, from neuron j to neuron i. Synapses here are modeled as current based. This means that synaptic input into the neuron is independent of the neuron's membrane potential (the intensity function in this model).

In this chapter, input spike trains are denoted $\hat{S}_k(t)$, neuron spike trains are $S_i(t)$, and both of these are represented as the sum of Dirac delta functions positioned at the times of spikes. These spike trains are realizations of the intensity functions, $\hat{\lambda}_k(t)$ and $\lambda_i(t)$, respectively, and have temporally averaged firing rates (or mean firing rates), $\hat{\nu}_k$ and ν_i, respectively.

All EPSP kernels used in this chapter are of the form given by

$$\epsilon(u) = \frac{1}{\tau_B - \tau_A}\Big(e^{-\frac{u}{\tau_B}} - e^{-\frac{u}{\tau_A}}\Big)h(u), \tag{2.3.2}$$

where $\tau_B > \tau_A$ and $h(u)$ is the Heaviside function such that for $u \geq 0$, $h(u) = 1$, and $h(u) = 0$ otherwise. There are three main EPSP kernels used in this chapter: 'slow',

'medium', and 'fast'. The values of the time constants for these EPSP kernels are shown in Table 2.3.1.

2.3.2 Spike-Timing-Dependent Plasticity (STDP)

In this chapter, learning refers to changes made to the network due to the additive STDP learning rule (Gilson et al., 2009d). The change in synaptic weight, J, due to this rule is

$$\Delta J = \eta \begin{cases} \omega_{\text{in}}, & \text{for each pre-synaptic spike} \\ \omega_{\text{out}}, & \text{for each post-synaptic spike} \\ W(\Delta t), & \text{for each pair of spikes (where } |\Delta t| \text{ is sufficiently small),} \end{cases} \tag{2.3.3}$$

where η is the learning rate, $\Delta t = t_{\text{in}} - t_{\text{out}} + d^{\text{ax}} - d^{\text{den}}$, t_{in} and t_{out} are the times of the spikes at the somas of the pre- and post-synaptic neurons, respectively, and d^{ax} and d^{den} are the axonal and dendritic delays of the synapse, respectively. This is illustrated in Figure 2.1A and B. Finally, ω_{in} and ω_{out} are rate-based parameters that change the synaptic weight for every pre- and post-synaptic spike, respectively. The learning window, $W(t)$, is of the form

$$W(t) = \begin{cases} c_{\text{p}} e^{\frac{t}{\tau_{\text{p}}}}, & t < 0 \\ 0, & t = 0 \\ -c_{\text{d}} e^{\frac{-t}{\tau_{\text{d}}}}, & t > 0. \end{cases} \tag{2.3.4}$$

where the values of the parameters used in this chapter are shown in Table 2.3.1. Figure 2.1C shows this learning window.

For a network with only axonal delays (i.e., the dendritic delays are sufficiently short to be neglected), the learning rule described in Equation (2.3.3) can be reformulated to give the rate of change of the weight J_{ij} as

$$\dot{J}_{ij}(t, d_{ij}^{\text{ax}}) = \eta \Big[\omega_{\text{in}} \nu_j(t) + \omega_{\text{out}} \nu_i(t) + \tilde{W} \nu_i(t) \nu_j(t) + C_{ij}^{W}(t, d_{ij}^{\text{ax}}) \Big], \tag{2.3.5}$$

where $\tilde{W} = \int_{-\infty}^{\infty} W(u) du$, and

$$C_{ij}^{W}(t, d_{ij}^{\text{ax}}) = \int_{-\infty}^{\infty} W(u) C_{ij}(t, u - d_{ij}^{\text{ax}}) du = \Big[W(-u) \underset{u}{*} C_{ij}(t, u) \Big](-d_{ij}^{\text{ax}}), \tag{2.3.6}$$

where $a(u) \underset{u}{*} b(u)$ denotes the convolution of functions $a(u)$ and $b(u)$ with respect to u. The axonal delay, d_{ij}^{ax}, can be seen to effectively shift the learning window in a positive direction. The correlation function for a pair of neurons in a recurrent network is defined by (Gilson et al., 2010c)

$$C_{ij}(t, u) = \frac{1}{T} \int_{t-T}^{t} \langle S_i(t') S_j(t' + u) \rangle dt' - \left[\frac{1}{T} \int_{t-T}^{t} \langle S_i(t') \rangle dt' \right] \left[\frac{1}{T} \int_{t-T}^{t} \langle S_j(t' + u) \rangle dt' \right]. \tag{2.3.7}$$

This notation generalizes that used previously (Gilson et al., 2009d), in which only constant input intensity functions were considered.

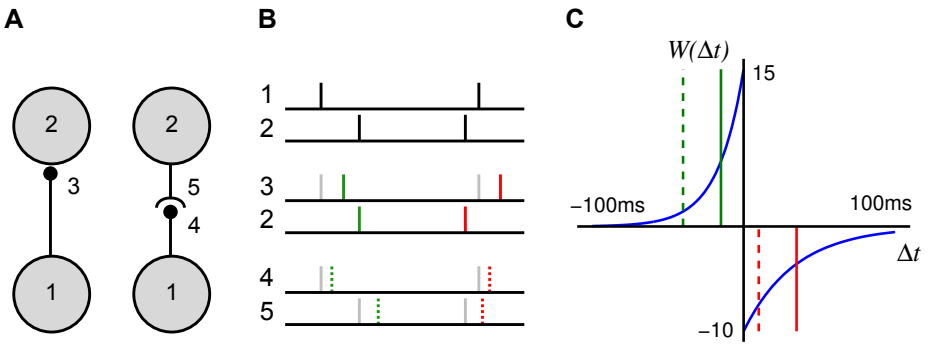


Figure 2.1: Additive STDP learning window. (A) Examples of two different synapses between the pre-synaptic neuron, 1, and the post-synaptic neuron, 2. The left one has a short dendritic delay and the right one has similar axonal and dendritic delays. (B) Examples of spike pairs. Top: Spike times are given at the somas of each of the neurons in each case in B. Middle: Pre- and post-synaptic spike times for the synapse with a short dendritic delay in B. Bottom: Pre- and post-synaptic spike times for the synapse with similar axonal and dendritic delays in B. (C) The learning window, W, used in this chapter that describes how the change in synaptic weight depends upon the difference in time between the pre- and post-synaptic spikes at the synapse. The form of this is described in Equation (2.3.4) with parameter values shown in Table 2.3.1. This window was used in an additive STDP learning rule along with two rate-based terms as described in Equation (2.3.3). The changes in synaptic strength due to the synaptic spike pairs (shown in B) for each of these two cases is shown by the red and green vertical lines. This shows that as the dendritic delay is increased, or the axonal delay decreased, the Δt for the spike pairs is shifted to the left on the learning window (the opposite occurs for increasing the axonal delay, or decreasing the dendritic delay).

2.3.3 Network Configuration

The network configuration that we considered, as illustrated in Figure 2.2A, consisted of a single network of N neurons. Each neuron received feed-forward connections from a set of M inputs and also recurrent connections from other neurons in the network. The inputs were spike trains, $\hat{S}_k(t)$, each a different realization of the same Poisson process with intensity function, $\hat{\lambda}(t)$. This intensity function was oscillatory in time; it can be thought of as the instantaneous firing rate of the inputs with mean (temporally averaged) firing rate, $\hat{\nu}$. Each neuron received N_K feed-forward connections, all with the same weight, $\bar{K}$, and axonal delay, $\hat{d}_0^{\text{ax}}$, and negligible dendritic delays. The neurons each produced spike trains, $S_i(t)$, according to the neuron model used. In this chapter, this was either the Poisson neuron model or the leaky integrate-and-fire (LIF) neuron model. There were N_J recurrent connections into each neuron. These were initially all the same weight but were modified by additive STDP. These connections each had different axonal delays, sampled uniformly from a range. Initially, we assumed these connections had negligible dendritic delays. This model is illustrated in Figure 2.2A, where $[K_{ik}]$ and $[J_{ij}]$ denote the matrices of feed-forward and recurrent connections just described.

In this chapter, we always used (unless otherwise stated) $N = M = 10\,000$, $N_K = N_J = 100$, and $\hat{d}_0^{\text{ax}}$ to 1ms. The axonal delay range (and later the dendritic delay range) used in this chapter was 1-10ms. This is consistent with the magnitude of axonal delays

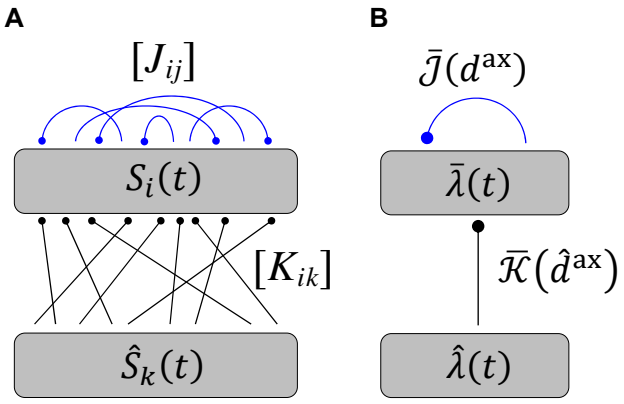


Figure 2.2: Diagram of single group network model. (A) Diagram of the full model used in simulations, which shows a network of N neurons with spike trains, $S_i(t)$, that receive inputs from M inputs, $\hat{S}_k(t)$, via fixed (black), feedforward connections denoted by $[K_{ik}]$, and from each other via plastic (blue), recurrent connections denoted by $[J_{ij}]$. (B) Diagram of the simplified, analytical model, which shows the same network represented by an ensemble averaged, instantaneous firing rate, $\bar{\lambda}(t)$, which is driven by inputs with instantaneous firing rate, $\hat{\lambda}(t)$. The (fixed; black) feedforward and (plastic; blue) recurrent connections are represented by the axonal delay profiles, $\bar{\mathcal{K}}(\hat{d}^{\text{ax}})$ and $\bar{\mathcal{J}}(d^{\text{ax}})$, respectively.

observed in the cortex (4-20ms) (González-Burgos et al., 2000), while perhaps less so for the auditory brainstem (0.4-1.4ms) (Beckius et al., 1999).

When the neurons were modeled using the Poisson neuron model, we simplified the full model analytically in two major ways. This simplification is illustrated in Figure 2.2B. First, instead of the full set of input and neuron spike trains, we considered only the ensemble averaged, instantaneous firing rates, $\bar{\hat{\lambda}}(t)$ and $\bar{\lambda}(t)$, for the input and network neurons, respectively (as the inputs have identical intensity functions, $\bar{\hat{\lambda}}(t) = \hat{\lambda}(t)$). Second, we represented the sets of feed-forward and recurrent connections as weighted axonal delay profiles (or simply axonal delay profiles or delay profiles), $\bar{\mathcal{K}}(\hat{d}^{\text{ax}})$ and $\bar{\mathcal{J}}(d^{\text{ax}})$, respectively. These delay profiles give the mean weight for connections with a specific axonal delay ($\hat{d}^{\text{ax}}$ or d^{ax}, respectively). When representing a set of recurrent connections that are uniformly sampled from a fixed range of axonal delays ($d^{\text{ax}}_{\text{min}}$ to $d^{\text{ax}}_{\text{max}}$), the integral of the recurrent axonal delay profile is

$$\int_{d^{\text{ax}}_{\text{min}}}^{d^{\text{ax}}_{\text{max}}} \bar{\mathcal{J}}(x)dx = (d^{\text{ax}}_{\text{max}} - d^{\text{ax}}_{\text{min}})\bar{J} = \Delta d^{\text{ax}}\bar{J}, \tag{2.3.8}$$

where $\bar{J}$ is the mean recurrent weight in the network and Δd^{ax} is the range of the axonal delays in the network. We relaxed our definition of the axonal delay profile representing the mean weight for a specific axonal delay when the range of the axonal delays, Δd^{ax}, was zero. This is the case for the input connections, as they all have the same axonal delay, $\hat{d}^{\text{ax}}_0$. The profile is instead given by $\bar{\mathcal{K}}(\hat{d}^{\text{ax}}) = \bar{K}\delta(\hat{d}^{\text{ax}} - \hat{d}^{\text{ax}}_0)$, where $\bar{K}$ is the mean feed-forward weight (and also the integral of the profile). Other feed-forward delay profiles (e.g. Gaussian) could have been considered but this was the simplest analytically and the

effect of other profiles would solely be to reduce the effective modulation amplitude of the input spike trains.

This chapter investigated the learning that occurs in the recurrent network through changes in the recurrent axonal delay profile. It also considered the amplitude of the oscillatory firing rate of the network (averaged over the neurons in the network) to different oscillatory inputs after this learning has occurred.

2.3.4 Learning with Axonal Delays

We investigated the learning of the recurrent connections in the network by considering the changes to the recurrent axonal delay profile due to Equation (2.3.5). We modified Equation (2.3.5) to apply to the recurrent axonal delay profile. The new learning equation is

$$\dot{\bar{\mathcal{J}}}(t, d^{\text{ax}}) = \eta\Big[\omega_{\text{in}}\bar{\nu}(t) + \omega_{\text{out}}\bar{\nu}(t) + \tilde{W}\bar{\nu}(t)^2 + \bar{C}^W(t, d^{\text{ax}})\Big], \tag{2.3.9}$$

where $\bar{\nu}(t)$ is the temporally averaged firing rate of the recurrent neurons and $\bar{C}^W(t, d^{\text{ax}})$ is the convolution of the learning window, $W(t)$, with the mean recurrent correlation function, $\bar{C}(t, u)$. The first three terms in this equation determine the evolution of the mean recurrent weight over all axonal delays. We were interested in the last term, which determines the average deviation from this mean for connections with different axonal delays. In this model, learning was assumed to happen on a longer timescale compared with that of the network activity and so we treated the recurrent correlation as quasi-stationary. For this reason, the t-dependence of the average recurrent correlation function is dropped and so is given by $\bar{C}(u)$ in the subsequent analysis of this chapter. Using the simplified model with a recurrent axonal delay profile, we found the (ordinary frequency) Fourier transform of $\bar{C}(u)$ (see Appendix A.1) to be approximated by

$$\mathcal{F}\bar{C}(f) = N_K^2\bar{K}^2|\mathcal{F}\epsilon(f)|^2\Gamma(f)\mathcal{F}\bar{\hat{C}}(f), \tag{2.3.10}$$

where

$$\Gamma(f) = \frac{1}{|1 - \mathcal{F}\epsilon(f)\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(f)|^2}, \tag{2.3.11}$$

$\mathcal{F}\bar{\hat{C}}(f)$ is the Fourier transform of the average input correlation function, $\bar{\hat{C}}(u)$, $\mathcal{F}\bar{\mathcal{J}}(f)$ is the Fourier transform of the axonal delay profile, $\bar{\mathcal{J}}(d^{\text{ax}})$, $\mathcal{F}\epsilon(f)$ is the Fourier transform of the EPSP kernel, $\epsilon(t)$, and $\tilde{N}_J = (\Delta d^{\text{ax}})^{-1}N_J$.

The input intensity function for a population of oscillatory inputs is defined as

$$\hat{\lambda}_k(t) = \langle\hat{S}_k(t)\rangle = \hat{\nu}_0 + a \cdot \cos\big[2\pi f_{\text{m}}(t + \hat{d}_k)\big], \tag{2.3.12}$$

where $\hat{\nu}_0$ is the mean input rate (in spikes/s), a is the amplitude in the oscillations (in spikes/s), f_{m} is the modulation frequency of the oscillations (in Hz), and $\hat{d}_k$ is the delay of the input (in seconds). In this model, all inputs are assumed to be in phase with each

other. As Appendix A.2 shows, the temporally averaged input firing rate is $\hat{\nu}_0$ and the correlation function for any pair of inputs is

$$\hat{C}_{kl}(u) = \frac{a^2}{2}\cos(2\pi f_{\mathrm{m}} u) = \bar{\hat{C}}(u) \tag{2.3.13}$$

and the Fourier transform of this is

$$\mathcal{F}\hat{C}_{kl}(f) = \frac{a^2}{4}\Big[\delta(f - f_{\mathrm{m}}) + \delta(f + f_{\mathrm{m}})\Big] = \mathcal{F}\bar{\hat{C}}(f). \tag{2.3.14}$$

It should be noted that no additional higher-order spike timing correlations were introduced in the input spike trains. The correlations described here are the rate correlations arising solely from the fact that all input neurons shared a common firing rate modulation.

With oscillatory inputs, the average recurrent correlation function becomes

$$\mathcal{F}\bar{C}(f) = \frac{a^2 N_K^2 \bar{K}^2}{4}|\mathcal{F}\epsilon(f)|^2\Gamma(f)\Big[\delta(f - f_{\mathrm{m}}) + \delta(f + f_{\mathrm{m}})\Big]. \tag{2.3.15}$$

It can be seen that $\Gamma(-f) = \Gamma(f)$ and $|\mathcal{F}\epsilon(-f)|^2 = |\mathcal{F}\epsilon(f)|^2$. Using this, the Fourier transform of the correlation function can be combined with the learning window (shifted by the axonal delay), as described by Equation (2.3.6), to give the contribution to learning from recurrent correlations for connections of axonal delay, d^{ax}, as

$$\begin{aligned}
\bar{C}^W(d^{\mathrm{ax}}) &= \Big[W(-u) \underset{u}{*} \bar{C}(u)\Big](-d^{\mathrm{ax}}) = \mathcal{F}^{-1}\Big[\mathcal{F}W(-f)\mathcal{F}\bar{C}(f)\Big](-d^{\mathrm{ax}}) \\
&= \frac{a^2 N_K^2 \bar{K}^2}{4}|\mathcal{F}\epsilon(f_{\mathrm{m}})|^2\Gamma(f_{\mathrm{m}})\mathcal{F}^{-1}\Big[\mathcal{F}W(-f_{\mathrm{m}})\delta(f - f_{\mathrm{m}}) + \mathcal{F}W(f_{\mathrm{m}})\delta(f + f_{\mathrm{m}})\Big](-d^{\mathrm{ax}}) \\
&= \frac{a^2 N_K^2 \bar{K}^2}{4}|\mathcal{F}\epsilon(f_{\mathrm{m}})|^2\Gamma(f_{\mathrm{m}})\Big[\mathcal{F}W(-f_{\mathrm{m}})e^{-2\pi i d^{\mathrm{ax}} f_{\mathrm{m}}} + \mathcal{F}W(f_{\mathrm{m}})e^{2\pi i d^{\mathrm{ax}} f_{\mathrm{m}}}\Big],
\end{aligned} \tag{2.3.16}$$

where $\mathcal{F}W(f)$ is the Fourier transform of the learning window, $W(u)$.

This was reformulated, by rewriting $\mathcal{F}\epsilon(f)$ as $r_\epsilon(f)e^{-i\phi_\epsilon(f)}$ and $\mathcal{F}W(f)$ as $r_W(f)e^{i\phi_W(f)}$, to be

$$\begin{aligned}
\bar{C}^W(d^{\mathrm{ax}}) &= \frac{a^2 N_K^2 \bar{K}^2 r_\epsilon^2(f_{\mathrm{m}}) r_W(f_{\mathrm{m}})\Gamma(f_{\mathrm{m}})}{4}\Big[e^{-i(2\pi d^{\mathrm{ax}} f_{\mathrm{m}} + \phi_W(f_{\mathrm{m}}))} + e^{i(2\pi d^{\mathrm{ax}} f_{\mathrm{m}} + \phi_W(f_{\mathrm{m}}))}\Big] \\
&= \frac{a^2 N_K^2 \bar{K}^2 r_\epsilon^2(f_{\mathrm{m}}) r_W(f_{\mathrm{m}})\Gamma(f_{\mathrm{m}})}{2}\cos\big[2\pi d^{\mathrm{ax}} f_{\mathrm{m}} + \phi_W(f_{\mathrm{m}})\big].
\end{aligned} \tag{2.3.17}$$

Expressions for functions $r_\epsilon(f)$, $\phi_\epsilon(f)$, $r_W(f)$, and $\phi_W(f)$ were derived, for the specific EPSPs and learning window used in this chapter, in Appendix A.6.

Assuming weak recurrent connections compared to the input connections, $\Gamma(f_{\mathrm{m}}) \approx 1$, we derived the approximation

$$\bar{C}^W(d^{\mathrm{ax}}) \approx \frac{a^2 N_K^2 \bar{K}^2 r_\epsilon^2(f_{\mathrm{m}}) r_W(f_{\mathrm{m}})}{2}\cos\big[2\pi d^{\mathrm{ax}} f_{\mathrm{m}} + \phi_W(f_{\mathrm{m}})\big]. \tag{2.3.18}$$

The deviation of the mean weight for a given delay, $\bar{\mathcal{J}}(d^{\text{ax}})$ from the mean weight over all delays, $\bar{J}$, is defined as $\Delta\bar{\mathcal{J}}(d^{\text{ax}})$. Although, the mean weight is driven towards a homeostatic equilibrium, $\bar{J}^*$, by the rate-based learning terms (see Appendix A.3), the evolution of the deviation of weights from this mean is described by

$$\Delta\dot{\bar{\mathcal{J}}}(d^{\text{ax}}) = \dot{\bar{\mathcal{J}}}(d^{\text{ax}}) - \dot{\bar{J}} \approx \eta\bar{C}^W(d^{\text{ax}}) \approx \frac{\eta a^2 N_K^2 \bar{K}^2 r_\epsilon^2(f_{\text{m}}) r_W(f_{\text{m}})}{2}\cos\big[2\pi d^{\text{ax}} f_{\text{m}} + \phi_W(f_{\text{m}})\big]. \tag{2.3.19}$$

2.3.5 Response to Oscillations after Axonal Delay Selection

To determine the response of a network, after learning, to oscillatory inputs, we first needed to consider the instantaneous firing rate of a single neuron, which is given by

$$\lambda_i(t) = \nu_0 + \sum J_{ij}(t)\int \epsilon(r-d_{ij})\lambda_j(t-r)dr + \sum K_{ik}(t)\int \epsilon(r-\hat{d}_{ik})\hat{\lambda}_k(t-r)dr, \tag{2.3.20}$$

where $d_{ij} = d_{ij}^{\text{ax}} + d_{ij}^{\text{den}}$ and $\hat{d}_{ij} = \hat{d}_{ij}^{\text{ax}} + \hat{d}_{ij}^{\text{den}}$ are the total recurrent and input delays, respectively, which are the sums of the axonal and dendritic delay components. This means that the network response only depends on the total delays, not directly on the axonal or dendritic components, so only total delays are referred to in the following derivation. In networks with short dendritic delays, the axonal delay is equivalent to the total delay.

We assumed that the input connections have equal total delay, $\hat{d}$, the inputs have identical rate functions, $\hat{\lambda}(t)$, and $\nu_0 = 0$. We also represented all the recurrent weights as a profile over total delay, $\bar{\mathcal{J}}(d)$. Therefore, the average response of the network is

$$\bar{\lambda}(t) = \tilde{N}_J \int_{d_{\min}}^{d_{\max}} \bar{\mathcal{J}}(x)\int \epsilon(r-x)\bar{\lambda}(t-r)drdx + N_K\bar{K}\int \epsilon(r-\hat{d})\hat{\lambda}(t-r)dr, \tag{2.3.21}$$

where $\bar{K}$ is the mean feedforward weight.

For oscillatory inputs, $\hat{\lambda}(t) = \hat{\nu}_0 + a\cos(2\pi f_{\text{m}} t)$, we showed that the expression for the response of the network becomes (see Appendix A.4)

$$\bar{\lambda}(t) = \bar{\nu} + aN_K\bar{K}r_\epsilon(f_{\text{m}})\sum_{j=0}^{\infty}\big[r_\epsilon(f_{\text{m}})\tilde{N}_J r_{\bar{\mathcal{J}}}(f_{\text{m}})\big]^j \cos\Big\{2\pi f_{\text{m}}(t-\hat{d}) - j\big[\phi_{\bar{\mathcal{J}}}(f_{\text{m}}) + \phi_\epsilon(f_{\text{m}})\big] - \phi_\epsilon(f_{\text{m}})\Big\}, \tag{2.3.22}$$

where $r_{\bar{\mathcal{J}}}(f)$ and $\phi_{\bar{\mathcal{J}}}(f)$ are defined by

$$\mathcal{F}\bar{\mathcal{J}}(f) = \int_{d_{\min}}^{d_{\max}} \bar{\mathcal{J}}(x)e^{-2\pi i f x}dx = r_{\bar{\mathcal{J}}}(f)e^{-i\phi_{\bar{\mathcal{J}}}(f)}, \tag{2.3.23}$$

and

$$\bar{\nu} = \frac{N_K\bar{K}\hat{\nu}_0}{1 - \tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(0)} = \frac{N_K\bar{K}\hat{\nu}_0}{1 - N_J\bar{J}}, \tag{2.3.24}$$

and, as before, $r_\epsilon(f)$ and $\phi_\epsilon(f)$ are given by Fourier transform, $\mathcal{F}\epsilon(f) = r_\epsilon(f)e^{-i\phi_\epsilon(f)}$.

To the second order, we approximated the network response as

$$\begin{aligned}\bar{\lambda}(t) \approx \bar{\nu} + aN_K\bar{K}r_\epsilon(f_\mathrm{m})\Big\{&\cos\big[2\pi f_\mathrm{m}(t-\hat{d}) - \phi_\epsilon(f_\mathrm{m})\big] \\ &+ r_\epsilon(f_\mathrm{m})\tilde{N}_J r_{\bar{\mathcal{J}}}(f_\mathrm{m})\cos\big[2\pi f_\mathrm{m}(t-\hat{d}) - \phi_{\bar{\mathcal{J}}}(f_\mathrm{m}) - 2\phi_\epsilon(f_\mathrm{m})\big]\Big\},\end{aligned} \tag{2.3.25}$$

and, since $r_\epsilon(f_\mathrm{m})r_J(f_\mathrm{m}) < 1$, we approximated the network response amplitude as

$$R(f_\mathrm{m}) \approx aN_K\bar{K}r_\epsilon(f_\mathrm{m})\sqrt{1 + 2r_\epsilon(f_\mathrm{m})\tilde{N}_J r_{\bar{\mathcal{J}}}(f_\mathrm{m})\cos\big[\phi_{\bar{\mathcal{J}}}(f_\mathrm{m}) + \phi_\epsilon(f_\mathrm{m})\big]}, \tag{2.3.26}$$

using the result from Appendix A.7.

We assumed a Gaussian delay profile with mean, $\bar{d}^\mathrm{ax}$, standard deviation, σ,

$$\bar{\mathcal{J}}(d) = \frac{\Delta d^\mathrm{ax}\bar{J}}{\sqrt{2\pi\sigma^2}}e^{-\frac{(d-\bar{d}^\mathrm{ax})^2}{2\sigma^2}}, \tag{2.3.27}$$

so we found that

$$\begin{aligned}\mathcal{F}\bar{\mathcal{J}}(f) &= \Delta d^\mathrm{ax}\bar{J}e^{-2(\pi\sigma f)^2 - 2\pi i\bar{d}^\mathrm{ax}f}, \\ r_{\bar{\mathcal{J}}}(f) &= \Delta d^\mathrm{ax}\bar{J}e^{-2(\pi\sigma f)^2}, \\ \phi_{\bar{\mathcal{J}}}(f) &= 2\pi\bar{d}^\mathrm{ax}f.\end{aligned} \tag{2.3.28}$$

We found the amplitude of the response function with this Gaussian delay profile by substituting Equation (2.3.28) into Equation (2.3.26).

2.3.6 Learning with Both Axonal and Dendritic Delays

As previously considered (Gilson et al., 2010a), when dendritic delays are included together with the axonal delays, the expression for the learning term, $C_{ij}^W(t)$, becomes

$$C_{ij}^W(t) = \int_{-\infty}^{\infty} W(u)C_{ij}(t, u - d_{ij}^\mathrm{ax} + d_{ij}^\mathrm{den})du = \Big[W(-u) \underset{u}{*} C_{ij}(t,u)\Big](d_{ij}^\mathrm{den} - d_{ij}^\mathrm{ax}). \tag{2.3.29}$$

We performed a similar derivation as for learning with only axonal delays. We found that, for oscillatory inputs, the learning term due to correlations is a function of both the axonal and dendritic delays,

$$\bar{C}^W(d^\mathrm{ax}, d^\mathrm{den}) \approx \frac{a^2N_K^2\bar{K}^2r_\epsilon^2(f_\mathrm{m})r_W(f_\mathrm{m})}{2}\cos\big[2\pi(d^\mathrm{ax} - d^\mathrm{den})f_\mathrm{m} + \phi_W(f_\mathrm{m})\big]. \tag{2.3.30}$$

Therefore, the deviation of the mean weight for a given axonal and dendritic delay evolves according to

$$\Delta\dot{\bar{\mathcal{J}}}(d^\mathrm{ax}, d^\mathrm{den}) = \beta(f_\mathrm{m})\cos\big[2\pi(d^\mathrm{ax} - d^\mathrm{den})f_\mathrm{m} + \phi_W(f_\mathrm{m})\big]. \tag{2.3.31}$$

2.3.7 Learning with Two Recurrently Connected Groups

The recurrent network was also considered to be made up of two groups of neurons with each group receiving inputs from a different group of oscillatory inputs, as shown in Figure 2.3. We once again considered networks with short dendritic delays.

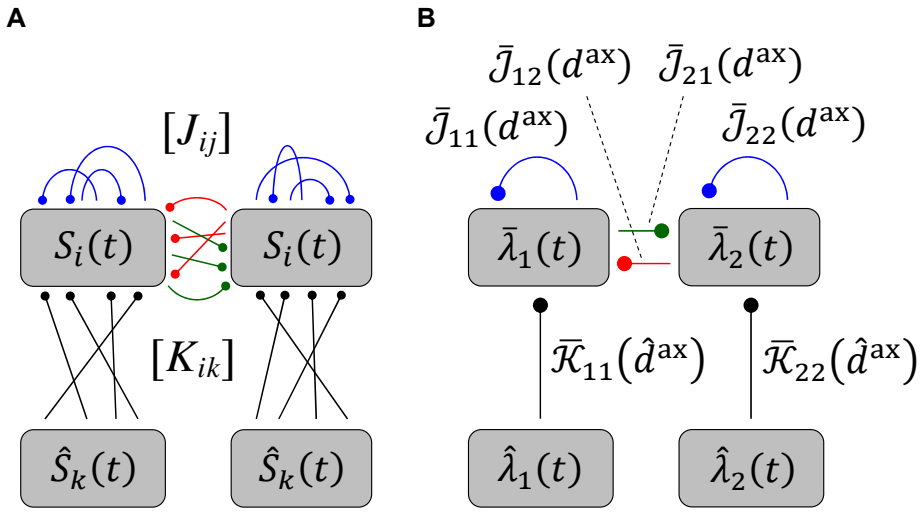


Figure 2.3: Diagram of two group network model. (A) Diagram of the full model used in simulations, which shows a network of N neurons with spike trains, $S_i(t)$, divided into two groups that each receive inputs from a different group of $\frac{M}{2}$ inputs, $\hat{S}_k(t)$, via fixed (black), feedforward connections, denoted by $[K_{ik}]$, and from each other via plastic (blue, red and green), recurrent connections, denoted by $[J_{ij}]$. (B) Diagram of the simplified, analytical model, which shows the same network represented by an ensemble averaged, instantaneous firing rate for each group, $\bar{\lambda}_1(t)$ and $\bar{\lambda}_2(t)$, respectively, that are driven by inputs with instantaneous firing rates, $\hat{\lambda}_1(t)$ and $\hat{\lambda}_2(t)$, respectively. The (fixed; black) feedforward and (plastic; blue, red, and green) recurrent connections are represented by the axonal delay profiles, $\bar{\mathcal{K}}_{ik}(\hat{d}^{\text{ax}})$ and $\bar{\mathcal{J}}_{ij}(d^{\text{ax}})$, respectively, where i denotes the group that the connections are to and k or j denote the group of inputs or neurons that the connections are from.

Here, $\bar{\mathcal{K}}_{ik}(\hat{d}^{\text{ax}})$ and $\bar{\mathcal{J}}_{ij}(d^{\text{ax}})$ are defined as the mean feed-forward and recurrent weights from group k or j to group i with delay $\hat{d}^{\text{ax}}$ or d^{ax}, respectively. We considered the case of two network groups, each with $\frac{N}{2}$ neurons, and two input groups, each with $\frac{M}{2}$ spike trains. The input connection matrix, $\bar{\mathcal{K}}(\hat{d}^{\text{ax}})$, is defined for the two input groups and two recurrent groups as

$$\bar{\mathcal{K}}(\hat{d}^{\text{ax}}) = \begin{bmatrix} \bar{K}\delta(\hat{d}^{\text{ax}} - \hat{d}_0^{\text{ax}}) & 0 \\ 0 & \bar{K}\delta(\hat{d}^{\text{ax}} - \hat{d}_0^{\text{ax}}) \end{bmatrix} = \bar{K}\delta(\hat{d}^{\text{ax}} - \hat{d}_0^{\text{ax}})I, \tag{2.3.32}$$

where, as before, $\bar{K}$ and $\hat{d}_0^{\text{ax}}$ are the mean feed-forward weight and the axonal delay of input connections, respectively. The spike trains in each input group were generated from the group's input intensity function. These are defined for each group of oscillatory inputs

as

$$\begin{aligned}\hat{\lambda}_1(t) &= \hat{\nu}_0 + a \cdot \cos\big[2\pi f_{\mathrm{m}}(t+\hat{d})\big],\\ \hat{\lambda}_2(t) &= \hat{\nu}_0 + a \cdot \cos\big[2\pi f_{\mathrm{m}}(t+\hat{d}+\hat{d}_{\mathrm{lag}})\big],\end{aligned} \tag{2.3.33}$$

where $\hat{\nu}_0$ is the mean input rate (in spikes/s), a is the amplitude in the oscillations (in spikes/s), f_{m} is the modulation frequency of the oscillations (in Hz), $\hat{d}$ is the delay of inputs in the first group (in seconds), and $\hat{d}_{\mathrm{lag}}$ is the time lag between the oscillations of the two input groups (in seconds). We determined that the average input correlation function matrix is (see Appendix A.2)

$$\bar{\hat{C}}(u) = \frac{a^2}{2}\begin{bmatrix} \cos\big[2\pi f_{\mathrm{m}}u\big] & \cos\big[2\pi f_{\mathrm{m}}(u+\hat{d}_{\mathrm{lag}})\big] \\ \cos\big[2\pi f_{\mathrm{m}}(u-\hat{d}_{\mathrm{lag}})\big] & \cos\big[2\pi f_{\mathrm{m}}u\big] \end{bmatrix}, \tag{2.3.34}$$

and the Fourier transform is

$$\mathcal{F}\bar{\hat{C}}(f) = \frac{a^2}{4}\Big[\delta(f-f_{\mathrm{m}})+\delta(f+f_{\mathrm{m}})\Big]\begin{bmatrix} 1 & e^{2\pi i\hat{d}_{\mathrm{lag}}f} \\ e^{-2\pi i\hat{d}_{\mathrm{lag}}f} & 1 \end{bmatrix}. \tag{2.3.35}$$

As with learning for a single group, we assumed weak recurrent connections. Therefore, we approximated the Fourier transform of the average recurrent correlation function as

$$\begin{aligned}\mathcal{F}\bar{C}(f) &= N_K^2\bar{K}^2|\mathcal{F}\epsilon(f)|^2\mathcal{F}\hat{C}(f)\\ &= \frac{a^2N_K^2\bar{K}^2}{4}|\mathcal{F}\epsilon(f)|^2\Big[\delta(f-f_{\mathrm{m}})+\delta(f+f_{\mathrm{m}})\Big]\begin{bmatrix} 1 & e^{2\pi i\hat{d}_{\mathrm{lag}}f} \\ e^{-2\pi i\hat{d}_{\mathrm{lag}}f} & 1 \end{bmatrix}.\end{aligned} \tag{2.3.36}$$

Therefore,

$$\begin{aligned}\bar{C}^W(d^{\mathrm{ax}}) &= \frac{a^2N_K^2\bar{K}^2|\mathcal{F}\epsilon(f_{\mathrm{m}})|^2}{4}\mathcal{F}^{-1}\Bigg\{\begin{bmatrix} 1 & e^{2\pi i\hat{d}_{\mathrm{lag}}f_{\mathrm{m}}} \\ e^{-2\pi i\hat{d}_{\mathrm{lag}}f_{\mathrm{m}}} & 1 \end{bmatrix}\mathcal{F}W(-f_{\mathrm{m}})\delta(f-f_{\mathrm{m}})\\ &\qquad + \begin{bmatrix} 1 & e^{-2\pi i\hat{d}_{\mathrm{lag}}f_{\mathrm{m}}} \\ e^{2\pi i\hat{d}_{\mathrm{lag}}f_{\mathrm{m}}} & 1 \end{bmatrix}\mathcal{F}W(f_{\mathrm{m}})\delta(f+f_{\mathrm{m}})\Bigg\}(-d^{\mathrm{ax}})\\ &= \frac{a^2N_K^2\bar{K}^2r_\epsilon^2(f_{\mathrm{m}})r_W(f_{\mathrm{m}})}{2}\\ &\qquad \times \begin{bmatrix} \cos\big[2\pi f_{\mathrm{m}}d^{\mathrm{ax}}+\phi_W(f_{\mathrm{m}})\big] & \cos\big[2\pi f_{\mathrm{m}}(d^{\mathrm{ax}}-\hat{d}_{\mathrm{lag}})+\phi_W(f_{\mathrm{m}})\big] \\ \cos\big[2\pi f_{\mathrm{m}}(d^{\mathrm{ax}}+\hat{d}_{\mathrm{lag}})+\phi_W(f_{\mathrm{m}})\big] & \cos\big[2\pi f_{\mathrm{m}}d^{\mathrm{ax}}+\phi_W(f_{\mathrm{m}})\big] \end{bmatrix}.\end{aligned} \tag{2.3.37}$$

2.3.8 Response of Two Groups after Axonal Delay Selection

For two recurrently connected groups, where the within group weights have been depressed and the inputs are as in Equation (2.3.33), each of the group responses is

$$\begin{aligned}\bar{\lambda}_1(t) &= \tilde{N}_J \int_{d_{\min}}^{d_{\max}} \bar{\mathcal{J}}_{12}(x) \int \epsilon(r-x)\bar{\lambda}_2(t-r)drdx + N_K\bar{K}\int \epsilon(r-\hat{d})\hat{\lambda}_1(t-r)dr,\\ \bar{\lambda}_2(t) &= \tilde{N}_J \int_{d_{\min}}^{d_{\max}} \bar{\mathcal{J}}_{21}(x) \int \epsilon(r-x)\bar{\lambda}_1(t-r)drdx + N_K\bar{K}\int \epsilon(r-\hat{d})\hat{\lambda}_2(t-r)dr.\end{aligned} \tag{2.3.38}$$

As derived in Appendix A.5, we approximated this as

$$\begin{aligned}\bar{\lambda}_1(t) \approx \frac{aN_K\bar{K}r_\epsilon(f_{\mathrm{m}})}{2}\Big\{&\cos\big[2\pi f_{\mathrm{m}}t - \phi_\epsilon(f_{\mathrm{m}}) - 2\pi\hat{d}f_{\mathrm{m}}\big]\\ &+ r_\epsilon(f_{\mathrm{m}})\tilde{N}_J r_{\bar{\mathcal{J}}_{12}}(f_{\mathrm{m}})\cos\big[2\pi f_{\mathrm{m}}t - 2\phi_\epsilon(f_{\mathrm{m}}) - \phi_{\bar{\mathcal{J}}_{12}}(f_{\mathrm{m}}) - 2\pi(\hat{d}_{\mathrm{lag}}+\hat{d})f_{\mathrm{m}}\big]\\ &+ r_\epsilon^2(f_{\mathrm{m}})\tilde{N}_J^2 r_{\bar{\mathcal{J}}_{12}}(f_{\mathrm{m}})r_{\bar{\mathcal{J}}_{21}}(f_{\mathrm{m}})\cos\big[2\pi f_{\mathrm{m}}t - 3\phi_\epsilon(f_{\mathrm{m}}) - \phi_{\bar{\mathcal{J}}_{12}}(f_{\mathrm{m}}) - \phi_{\bar{\mathcal{J}}_{21}}(f_{\mathrm{m}}) - 2\pi\hat{d}f_{\mathrm{m}}\big]\Big\},\\ \bar{\lambda}_2(t) \approx \frac{aN_K\bar{K}r_\epsilon(f_{\mathrm{m}})}{2}\Big\{&\cos\big[2\pi f_{\mathrm{m}}t - \phi_\epsilon(f_{\mathrm{m}}) - 2\pi(\hat{d}_{\mathrm{lag}}+\hat{d})f_{\mathrm{m}}\big]\\ &+ r_\epsilon(f_{\mathrm{m}})\tilde{N}_J r_{\bar{\mathcal{J}}_{21}}(f_{\mathrm{m}})\cos\big[2\pi f_{\mathrm{m}}t - 2\phi_\epsilon(f_{\mathrm{m}}) - \phi_{\bar{\mathcal{J}}_{21}}(f_{\mathrm{m}}) - 2\pi\hat{d}f_{\mathrm{m}}\big]\\ &+ r_\epsilon^2(f_{\mathrm{m}})\tilde{N}_J^2 r_{\bar{\mathcal{J}}_{21}}(f_{\mathrm{m}})r_{\bar{\mathcal{J}}_{12}}(f_{\mathrm{m}})\\ &\qquad\times\cos\big[2\pi f_{\mathrm{m}}t - 3\phi_\epsilon(f_{\mathrm{m}}) - \phi_{\bar{\mathcal{J}}_{21}}(f_{\mathrm{m}}) - \phi_{\bar{\mathcal{J}}_{12}}(f_{\mathrm{m}}) - 2\pi(\hat{d}_{\mathrm{lag}}+\hat{d})f_{\mathrm{m}}\big]\Big\},\end{aligned} \tag{2.3.39}$$

where $r_{\bar{\mathcal{J}}_{ij}}(f)$ and $\phi_{\bar{\mathcal{J}}_{ij}}(f)$ are the amplitude and negative phase of the Fourier transform of the axonal delay profile of connections from group j to group i, respectively. As we did for a single recurrent group, we assumed the between group delay profiles were Gaussian. Specifically, it was assumed that $r_{\bar{\mathcal{J}}_{ij}}(f) = \Delta d^{\mathrm{ax}}\bar{J}_{ij}e^{-2(\pi\sigma_{ij}f)^2}$ and $\phi_{\bar{\mathcal{J}}_{ij}}(f) = -2\pi\bar{d}_{ij}f$. The result from Appendix A.7 was used to approximate the amplitude of this response.

2.3.9 Numerical Simulations

Simulations were performed using the leaky integrate-and-fire (LIF) neuron model. A single state variable, $V_i(t)$, represents the membrane potential for each neuron i that evolves according to

$$\frac{dV_i(t)}{dt} = \frac{1}{\tau_{\mathrm{m}}}\left(V_{\mathrm{p}} - V_i(t) + \sum_{j\neq i}\Big\{J_{ij}(t)\big[E_{\mathrm{S},j} - V_i(t)\big]\sum_n \epsilon_{\mathrm{c}}(t - t_{j,n} - d_{ij}^{\mathrm{ax}} - d_{ij}^{\mathrm{den}})\Big\}\right), \tag{2.3.40}$$

where τ_{m} is the passive membrane time constant, V_{p} is the resting membrane potential, $E_{\mathrm{S},j}$ is the synaptic reversal potential of the (excitatory) synapses from neuron j, and $\epsilon_{\mathrm{c}}(t)$ represents the excitatory post-synaptic conductance (EPSC). This plays a similar role to the EPSP kernel, $\epsilon(t)$, in the Poisson neuron model and, because of this, we refer to both

$\epsilon(t)$ and $\epsilon_{\rm c}(t)$ as the EPSP or the EPSP kernel. $J_{ij}(t)$, $t_{j,n}$, $d_{ij}^{\rm ax}$ and $d_{ij}^{\rm den}$ are the same as for the Poisson neuron model. A spike was produced when the membrane potential reached a threshold value, $V_{\rm th}$, and it was reset to $V_{\rm r}$. An absolute refractory period was used, which prevented the membrane potential from changing during this time. The values of these parameters are shown in Table 2.3.1. It should be noted that different values for the membrane time constant were used for the three different EPSP kernels considered. Simulations were of a model with 10 000 LIF neurons that each received 100 randomly chosen input spike trains from a total of 10 000. These neurons also received 100 recurrent connections from other neurons in the network, which had no dendritic delay and axonal delays that were sampled uniformly from the range 1-10ms. The weights of the input (feed-forward) connections were fixed and chosen to be equal to each other and such that, without recurrent connections, the temporally averaged firing rate of the neurons was approximately equal to that of the inputs. The weights of the recurrent connections were updated by STDP during the simulation. They were initialized to be equal to each other and such that they significantly increased the firing rate of the neurons above the base rate caused by the inputs alone. Simulations were performed using an in-house neuron modeling software program, SpikeSim, used in previous studies (Gilson et al., 2009d, 2010a) and in Chapter 4.

The networks simulated and considered in this chapter contained only excitatory neurons and operated in super-threshold, mean-driven regimes. We address this and consider the limitations for this as a model of networks in the auditory brainstem or the cortex in the Discussion section.

Table 2.3.1: Model Parameters Parameters used in the model for the three different EPSPs ('slow', 'medium' and 'fast'). All parametersrs were used in simulations, but only EPSP and STDP parameters were used in the analytical model.

Type	**Parameter**	**Slow**	**Medium**	**Fast**
EPSP	Synaptic Rise Time, τ_A (ms)	1	0.5	0.1
EPSP	Synaptic Decay Time, τ_B (ms)	5	1	0.5
LIF	Membrane Time Constant, $\tau_{\rm m}$ (ms)	20	10	5
LIF	Threshold Potential, $V_{\rm th}$ (mV)	-50		
LIF	Resting Potential, $V_{\rm p}$ (mV)	-65		
LIF	Reset Potential, $V_{\rm r}$ (mV)	-65		
LIF	Synaptic Reversal Potentials, $E_{{\rm S},j}$ (mV)	0		
LIF	Refractory Period (ms)	1		
STDP	Potentiation Factor, $c_{\rm p}$	15		
STDP	Depression Factor, $c_{\rm d}$	10		
STDP	Potentiation Time Constant, $\tau_{\rm p}$ (ms)	17		
STDP	Depression Time Constant, $\tau_{\rm d}$ (ms)	34		

2.4 Results

In this chapter, we considered how STDP leads to delay selection in the recurrent connections of a network receiving oscillatory inputs. We used the Poisson neuron model to derive analytical results and the leaky integrate-and-fire (LIF) neuron model in simulations. The observed learning was due to additive STDP together with single-spike contributions, or rate-based terms. As in previous work (Gilson et al., 2009d), these rate-based terms, along with the learning window cause the mean recurrent weight to converge to a homeostatic equilibrium. However, in this chapter, we were concerned with the deviation of individual weights with specific delays from the mean weight.

2.4.1 Axonal Delay Selection within a Recurrent Network

We first considered how STDP changes the functional connectivity of a recurrent network receiving inputs from a single group of oscillatory inputs. The connections in the networks had a range of axonal delays (1-10ms) but very short dendritic delays. The modulation frequencies of the inputs were between 100 and 300Hz. This range is typical of the fundamental frequency of sounds encoded in modulation frequencies in the auditory brainstem. We modeled the recurrent connections as a weighted axonal delay profile $\bar{J}(d^{\text{ax}})$, which is the mean weight of connections with a given axonal delay. We analytically derived an expression for the changes made to this profile and showed that these predicted changes were supported by numerical simulations.

As detailed previous studies (Gilson et al., 2009d), the rate-based plasticity parameters ω_{in} and ω_{out}, together with the learning window bias $\tilde{W}$, caused the mean weight in the network to converge to a stable equilibrium value, $\bar{J}^*$. It is important to note that this equilibrium is for the mean weight only and it does not imply that the individual weights reach stable values. The mean weight of connections with a given axonal delay deviated from this homeostatic equilibrium as given by, $\Delta\bar{J}(d^{\text{ax}}) = \bar{J}(d^{\text{ax}}) - \bar{J}^*$. For inputs with a given modulation frequency, f_{m}, we predicted that this deviation would evolve according to (see Equation (2.3.19))

$$\Delta\dot{\bar{J}}(d^{\text{ax}}) = \beta(f_{\text{m}})\cos\big[2\pi d^{\text{ax}} f_{\text{m}} + \phi_W(f_{\text{m}})\big], \tag{2.4.1}$$

where $\beta(f_{\text{m}}) = \frac{1}{2}\eta a^2 N_K^2 \bar{K}^2 r_\epsilon^2(f_{\text{m}}) r_W(f_{\text{m}})$ is a positive factor that determines the rate of learning, a is the amplitude of the input modulation, $\bar{K}$ is the mean feed-forward weight. The functions $r_\epsilon(f)$ and $\phi_\epsilon(f)$ denote the amplitude and negative phase of the Fourier transform of $\epsilon(t)$ (i.e. $\mathcal{F}\epsilon(f) = r_\epsilon(f)e^{-i\phi_\epsilon(f)}$), respectively. The functions $r_W(f)$ and $\phi_W(f)$ denote the amplitude and phase of the Fourier transform of $W(t)$ (i.e. $\mathcal{F}W(f) = r_W(f)e^{i\phi_W(f)}$), respectively. A plot of $r_W(f)$ is shown in Figure 2.4. The functions $r_\epsilon(f)$ and $\phi_W(f)$ are considered in more detail in the next section.

Assuming upper and lower bounds on the synaptic weights, we determined the axonal delay profiles that resulted from this learning. An example of this is shown in Figure 2.5A for an input frequency of 120Hz. Eventually, a narrow range of delays was

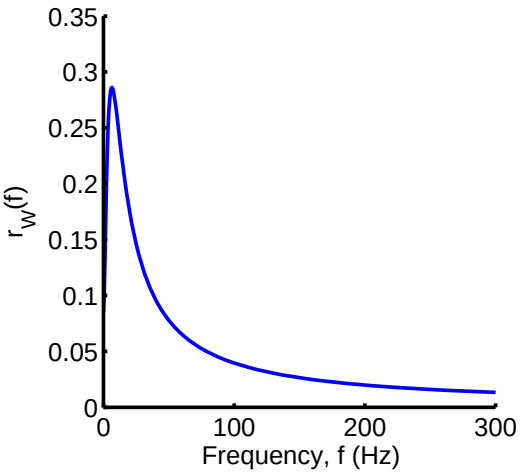


Figure 2.4: Amplitude of the Fourier transform of the learning window as a function of frequency. Given by Equations (31) in Appendix A.6. The standard window has parameters $c_P = 15$, $c_D = 10$, $\tau_P = 17$ms and $\tau_D = 34$ms.

uniquely selected because of the bounds on the weights. This narrow range was centered approximately on the delay that resonated with the input frequency. The shortest of these is given by

$$\bar{d}^{\text{ax}} = \frac{1}{f_{\text{m}}} - \frac{\phi_W(f_{\text{m}})}{2\pi f_{\text{m}}}. \tag{2.4.2}$$

If this delay is outside the range of axonal delays in the network, then the frequency cannot be encoded by the network. The minimum (maximum) delay in the recurrent connections of the network sets the limit on the maximum (minimum) frequency that can be learned.

Equation (2.4.1) shows that the synaptic rise and decay times only affect the learning rate, $\beta(f_{\text{m}})$, and not the delays that are selected. The learning rate is also dependent upon the square of the amplitude of oscillations and the square of the input strength. For the simulations with LIF neurons, the firing rate of neurons is no longer linear with the input strength so this learning rate dependence on input strength is different but it is still a non-decreasing dependence.

We compared the learning that occurs with 120Hz inputs, shown analytically in Figure 2.5A, to simulations with 10 000 LIF neurons. This is shown in Figure 2.5B. The shape of the delay profile learned was the same; it is only the rate of learning that differs. The simulations with the LIF neurons showed a significantly faster learning rate. Simulations with Poisson neurons, however, did not show this difference when compared to the analytical model (see Figure 2.6). The higher learning rate appears to be due to the differences between the Poisson and LIF neuron models. We saw that, after further learning occurred, the mean recurrent weight did not remain at the homeostatic equilibrium. Instead the mean recurrent weight increased up to a critical point (Figure 2.5C). The concept of a "critical point" and how it is relevant to the response of the network is

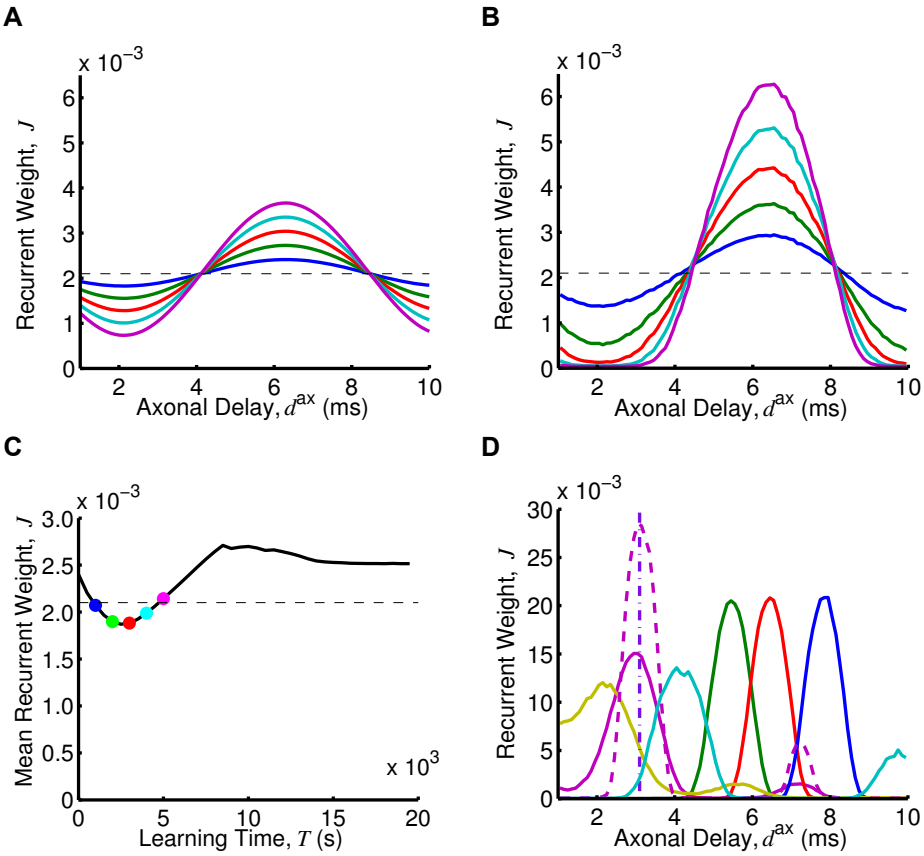


Figure 2.5: Learning through axonal delay selection with oscillatory inputs. (A) Axonal delay profiles predicted by the analytical model, Equation (2.4.1) (which uses the Poisson neuron model), with an input modulation frequency of 120Hz, after 1000s (blue), 2000s (green), 3000s (red), 4000s (cyan), and 5000s (magenta) of learning from the initial (dashed) profile. (B) As for A, but for a simulation using LIF neurons. (C) Mean recurrent weight, for connections of any axonal delay, over time for the simulation in B (black) and the stable mean recurrent weight assumed in the analytical model (dashed), with colored dots showing the mean recurrent weight in the simulation for the times corresponding to the profiles shown in B. (D) Axonal delay profiles after 20 000s of learning in simulations with LIF neurons for input modulation frequencies of 100Hz (blue), 120Hz (red), 140Hz (green), 180Hz (cyan), 240Hz (magenta), and 300Hz (yellow). These simulations used a 'medium' EPSP (solid), except for one that used a 'fast' EPSP (dashed). Also shown is a delta delay profile at 3.1ms (purple, dot-dashed). For both analytical and simulations, a modulation amplitude, a, of 5 spikes/s was used.

explained in more detail in the Discussion section. For learning within the simulation, it is sufficient to observe that the mean recurrent weight increased above the homeostatic equilibrium, providing another way that the simulation differed from the analytical model.

We observed that different delays were selected for different input frequencies. This is shown in the delay profiles in Figure 2.5D. These are the result of 20 000s of learning in simulations with 10 000 LIF neurons. It can be seen that for the higher frequencies used (180, 240 and 300Hz) there was a second smaller peak at a longer delay. Equation

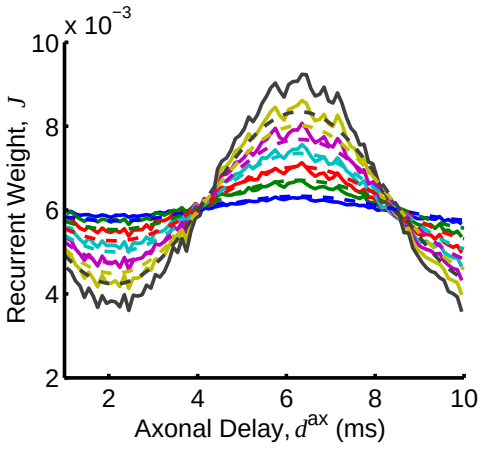


Figure 2.6: Comparison of learning predicted analytically (dashed) and from a simulation with Poisson neurons (solid). The axonal delay profile of a network at homeostatic equilibrium with 120Hz oscillatory inputs after 1000s (blue), 2000s (green), 3000s (red), 4000s (cyan), 5000s (magenta), 6000s (yellow), and 7000s (black). A 'medium' EPSP was used here.

(2.4.1) predicts that this second delay (and others that are within the allowed range for axonal delays) should be equally selected for. However, the simulations showed that, while each of these delays was initially potentiated, eventually the shortest of these delays was selected over the longer ones. We used a 'medium' EPSP kernel (0.5ms rise time, 1ms decay time) in all previously mentioned simulations. The learning for 240Hz with a 'fast' EPSP kernel (0.1ms rise time, 0.5ms decay time) is shown in Figure 2.5D.

2.4.2 Frequency Selective Response after Delay Selection

Next, we considered how this learning changed the way the network responds to different input frequencies. Being driven by oscillatory inputs, the network always had an oscillatory response at the same modulation frequency. We derived an approximation to the amplitude of this oscillatory response, $R(f_{\mathrm{m}})$, as a function of the modulation frequency of the inputs (see Equation (2.3.26))

$$R(f_{\mathrm{m}}) \approx aN_K\bar{K}r_\epsilon(f_{\mathrm{m}})\sqrt{1+2r_\epsilon(f_{\mathrm{m}})N_J\bar{J}e^{-2(\pi\sigma f_{\mathrm{m}})^2}\cos\left[2\pi f_{\mathrm{m}}\bar{d}^{\mathrm{ax}}+\phi_\epsilon(f_{\mathrm{m}})\right]}, \qquad (2.4.3)$$

where the network has a Gaussian axonal delay profile centered about $\bar{d}^{\mathrm{ax}}$ with a standard deviation of σ (and short dendritic delays, $d^{\mathrm{den}} \approx 0$). Additionally, $r_\epsilon(f)$ and $\phi_\epsilon(f)$ denote the amplitude and negative phase, respectively, of the Fourier transform of the EPSP kernel, $\epsilon(t)$ (i.e. $\mathcal{F}\epsilon(f) = r_\epsilon(f)e^{-i\phi_\epsilon(f)}$) and $\bar{J}$ is the mean recurrent weight.

The shape of this response function, $R(f_{\mathrm{m}})$, is highly dependent on the amplitude of the Fourier transform of the EPSP, $r_\epsilon(f_{\mathrm{m}})$, being used (see Figure 2.7A). This depends on the decay time and, to a lesser extent, on the rise time of the EPSP in the model (see

Appendix A.6). Figure 2.7C shows that, when the 'slow' EPSP (1ms rise time, 5ms decay time) was used, the oscillatory response of the network was very small in the frequency range considered regardless of how the delays in the network were tuned. The 'medium' EPSP (0.5ms rise time, 1ms decay time); however, gave rise to an oscillatory response with an amplitude peaked at a particular frequency. This frequency depended on the axonal delay, $\bar{d}^{\text{ax}}$, about which the profile was centered. Using the 'fast' EPSP (0.1ms rise time, 0.5ms decay time), this selective response was even more pronounced. This difference was larger when the axonal delay profile was such that the peak response amplitude was at higher frequencies.

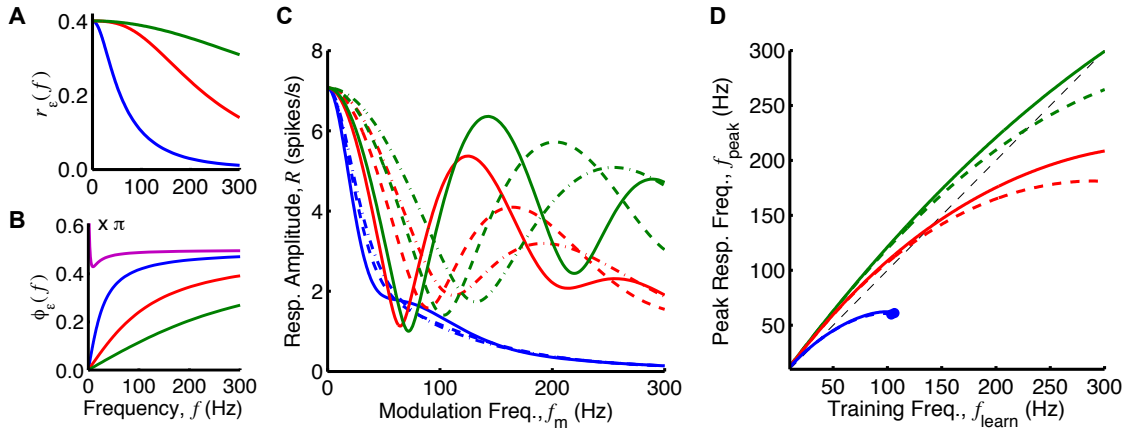


Figure 2.7: Analytical investigation of network response after delay selection. Responses considered for 'slow' (blue), 'medium' (red), and 'fast' (green) EPSPs. (A) Amplitude of the Fourier transform of the three EPSPs as functions of frequency, as given by Equation (33) in Appendix A.6. (B) Negative phase of the Fourier transform of the three EPSPs compared to the phase of the Fourier transform of the learning window (magenta) as functions of frequency, as given by Equations (31) and (33) in Appendix A.6. (C) Analytically determined response of networks with delay profiles centered about the delay selected due to learning (Equation (2.4.2)) with input frequencies of 120Hz (solid), 180Hz (dashed), and 240Hz (dot-dashed) inputs, with a profile width, σ, of 0.5ms. These curves are given by Equation (2.4.3). (D) Peak response frequency, f_{peak}, as a function of the training frequency, f_{learn}, of the network, for delay profiles with a width, σ, of 0.5ms (solid) and 1ms (dashed). The peak response frequency was numerically determined from the analytical formula in Equation (2.4.3). The dashed line represents $f_{\text{peak}} = f_{\text{learn}}$. Note that the dot ending some of the lines represents that, for higher training frequencies, there was no peak in the response amplitude with frequency. For plots C and D, a recurrent strength, $N_J\bar{J}$, of 0.5 and a modulation amplitude, a, of 5 spikes/s were used.

The frequency of the peak in the response amplitude function (excluding the peak at 0Hz) is denoted by f_{peak}. This frequency does not necessarily correspond to the frequency present during learning, f_{learn}. Equation (2.4.2) shows how f_{learn} determines the selected axonal delay, $\bar{d}^{\text{ax}}$. The correspondence between these two frequencies depends on the difference between $\phi_W(f)$ and $\phi_\epsilon(f)$. This is shown in Figure 2.7B for 'slow', 'medium', and 'fast' EPSPs. This shows that $\phi_W(f)$ was larger than $\phi_\epsilon(f)$ across the frequency range, for any of the EPSPs considered. This tended to cause f_{peak} to be higher than f_{learn}. However, there is a second factor affecting the correspondence between f_{learn} and

f_{peak}. This is the decay with frequency that was evident in both $r_\epsilon(f)$ (see Figure 2.7A) and the $e^{-2(\pi\sigma f_{\text{m}})^2}$ term. These decays tended to make f_{peak} lower than f_{learn}. Figure 2.7D shows the correspondence between f_{peak} and f_{learn}, for 'slow', 'medium', and 'fast' EPSPs for narrow and wide delay profiles. We generated this plot by first considering the response amplitude function, $R(f_{\text{m}})$, that resulted from assuming the selected axonal delay produced by the training frequency, f_{learn}. We then numerically found f_{peak} as the location of the first peak (after 0Hz) in this function. A similar plot for different learning window parameters is shown in Figure 2.8. This demonstrates the robustness of the mechanism to the learning window used. It is important to note that these plots do not take into account the membrane time constant (not present in the Poisson neuron model). This was present in simulations using the LIF neuron model and worked to effectively increase $\phi_\epsilon(f)$, bringing it closer to $\phi_W(f)$. Later in this section, results of simulations with LIF neurons show how this affected the frequency of the peak response.

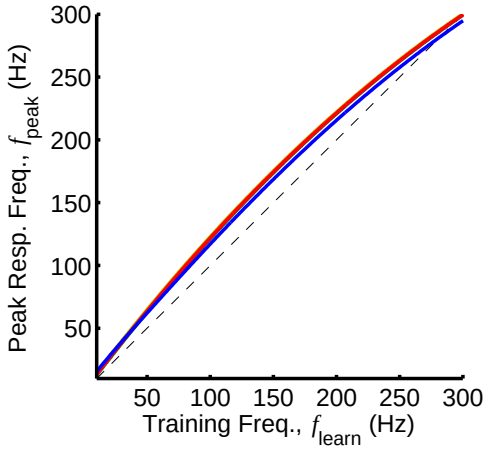


Figure 2.8: Analytical comparison of frequency correspondence between learning and response for different learning windows. Plot of the training frequency and corresponding peak response frequency for different learning windows for a 'fast' EPSP, a delay profile with width, σ, of 0.5ms, strength, J_0, of 0.5, and with a modulation amplitude, a, of 5 spikes/s. The different learning windows shown are: the standard window with $c_P = 15$, $c_D = 10$, $\tau_P = 17$ms and $\tau_D = 34$ms (green), the standard window with τ_P and τ_D multiplied by 0.1 (blue), the standard window with τ_P and τ_D multiplied by 10 (red), a balanced window with $c_P = c_D = 10$ and $\tau_P = \tau_D = 20$ (magenta), and a window biased in the reverse way to the standard with $c_P = 10$, $c_D = 15$, $\tau_P = 34$ms and $\tau_D = 17$ms (yellow). The dashed line represents $f_{\text{peak}} = f_{\text{learn}}$.

We compared the analytical expression for the network response amplitude for various input frequencies, $R(f_{\text{m}})$, to simulations using the Poisson neuron model. We carried out simulations with networks of 10 000 neurons. These simulations were done before any learning had occurred in the network (all weights were equal) and then after 20 000s of learning with 120Hz oscillatory inputs. The axonal delay profiles of this network before and after learning, along with a Gaussian profile fit to the after-learning profile, are shown in Figure 2.9A. Simulations were run multiple times, each for 10s of modeled time and

with a different input modulation frequency, for frequencies ranging from 30Hz to 300Hz in steps of 10Hz. The amplitude of the responses for these two networks as a function of the input modulation frequency is compared to an analytical approximation in Figure 2.9B. We determined this analytical approximation using Equation (2.4.3) and assuming the Gaussian delay profile that was fitted to the after-learning network. The analytical approximation closely matched the responses observed in the simulations.

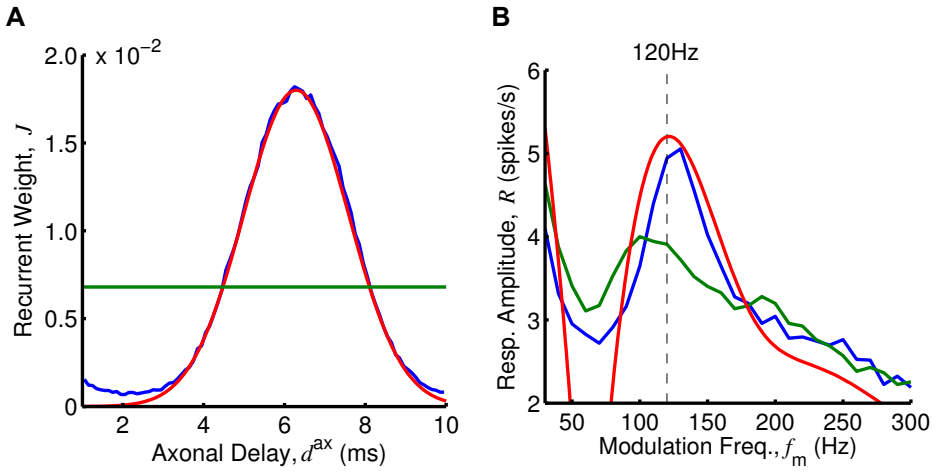


Figure 2.9: Comparison of analytical expression for network response with simulations using the Poisson neuron model. (A) The axonal delay profile of a network before (green) and after (blue) 40 000s of learning, with STDP and 120Hz inputs and a Gaussian delay profile (red), which closely approximates the profile after learning. (B) The amplitude of the network response to different input modulation frequencies, from simulations with networks of the same color in A (green and blue), or analytically determined from the Gaussian delay profile of the same color in A (red), using Equation (2.4.3). A 'medium' EPSP was used here.

Similar to the above simulations with the Poisson neuron model, we carried out simulations using networks of 10 000 LIF neurons (and 'medium' EPSPs). These networks had each learnt over 20 000s while receiving inputs with different modulation frequencies ('training frequencies'). The resulting axonal delay profiles of these networks are shown in Figure 2.5D. We ran simulations multiple times with all of these networks, each for 10s of modeled time and with input modulation frequencies ('testing frequencies') ranging from 50Hz to 300Hz in steps of 5Hz. The peak instantaneous firing rates of the periodic responses (averaged over neurons), or the peak response, observed are shown in Figure 2.10A. The peak instantaneous firing rate is presented instead of the amplitude because the response, while still periodic with the same frequency as the input, was no longer a cosine function. This was due to the non-linear nature of the LIF model. For networks trained with 100Hz, 120Hz, 140Hz, and 180Hz, these response curves showed a clear selectivity toward the input modulation frequency at which they were trained.

While LIF networks were able to encode higher training frequencies (240Hz and 300Hz) in their selected axonal delays (Figure 2.5D), they did not respond selectively to

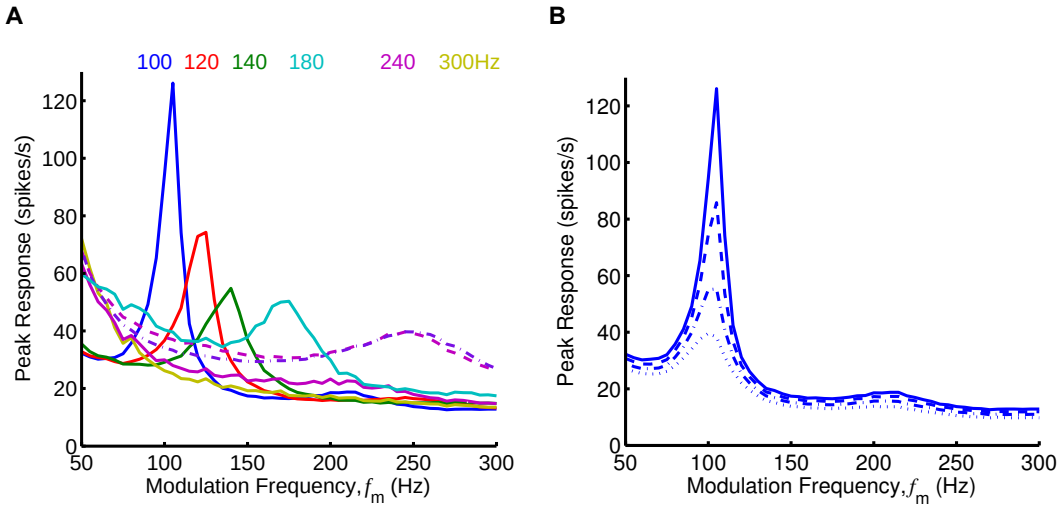


Figure 2.10: Simulations of peak network response after delay selection. Response plots showing the peak in the periodic response of networks of LIF neurons plotted as a function of the modulation frequency, f_m ('medium' EPSPs used unless otherwise specified). (A) Response plot for networks after 20 000s of learning with input modulation frequencies of 100Hz (blue), 120Hz (red), 140Hz (green), 180Hz (cyan), 240Hz (magenta), and 300Hz (yellow), as shown in Figure 2.5D. The simulations used 'medium' EPSPs (solid), except for two which used the fast EPSP (dashed and dot-dashed). The weights in the networks trained with 240Hz and 300Hz inputs were scaled down slightly (to about 0.99 of their original value) so that the networks were below criticality. (B) Response plot for the network trained with 100Hz inputs in A, with the weights all scaled by 0.90 (dotted), 0.95 (dot-dashed), 0.98 (dashed), 1.00 (solid).

this frequency after learning. This was largely due to the fact that the network was not able to respond with these higher frequencies regardless of the delays in the network. We hypothesized that networks with faster synapses and neurons would be able to show a stronger response at these higher frequencies. We considered this situation by running simulations using a network with faster synapses and neurons that was trained with an input frequency of 240Hz. This is described in the Methods section and the learned network is shown in Figure 2.5D. Its response is shown in Figure 2.10A. We observed that the network showed selectivity to an input frequency of 250Hz. This was very close to the trained frequency. The response of the network with all of the axonal delays set to 3.1ms (also shown in Figure 2.5D) showed a response with only slightly improved selectivity. Another point to notice is that the response of the networks trained with higher frequencies (180, 240, and 300Hz) to frequencies in the lower range (50-100Hz) was higher than networks trained with 100, 120, or 140Hz. This was likely due to the fact that the potentiated delays in these networks were relatively short. It may be that these short delays were providing recurrent feedback within the same oscillation peak, which for lower frequencies like 50Hz was relatively wide.

The recurrent connections in a network of excitatory neurons provided positive feedback to the network. For weak recurrent connections, this positive feedback did not greatly affect the firing of the neurons in the network. As this feedback increased, these connec-

tions caused higher firing rates. This continued up to a critical point where the feedback caused the firing in the network to continue increasing in an unstable manner. A network with its mean recurrent weight at this point can be said to be critical, or to have reached criticality. The trained networks we considered ended up just below criticality after learning. Figure 2.10B shows the change to the response of the network caused by scaling down the recurrent weights in the network trained with 100Hz. This shows a decreasing frequency selectively as the network moves away from criticality. In Figure 2.10A, it was necessary for us to scale down all of the recurrent weights in the networks trained with 240 and 300Hz by a slight amount (down to 0.99 of their original value) so that they were below criticality (for all frequencies).

2.4.3 Axonal and Dendritic Delay Selection

We extended the learning of oscillatory inputs by axonal delay selection to consider the networks with connections that had a range of dendritic as well as axonal delays. To do this, we needed to consider the recurrent connections as a weighted delay profile over both axonal and dendritic delays, $\bar{\mathcal{J}}(d^{\mathrm{ax}}, d^{\mathrm{den}})$. We derived an expression for how this evolves due to STDP (see the Methods section). As previously, we predicted the evolution of the deviation of this profile from the homeostatic equilibrium

$$\Delta\dot{\bar{\mathcal{J}}}(d^{\mathrm{ax}}, d^{\mathrm{den}}) = \beta(f_{\mathrm{m}})\cos\big[2\pi(d^{\mathrm{ax}} - d^{\mathrm{den}})f_{\mathrm{m}} + \phi_W(f_{\mathrm{m}})\big], \tag{2.4.4}$$

where, as before, $\beta(f_{\mathrm{m}}) = \frac{1}{2}\eta a^2 N_K^2 \bar{K}^2 r_\epsilon^2(f_{\mathrm{m}}) r_W(f_{\mathrm{m}})$.

This analytic result can be visualized using a heat map of the deviation of weights from the homeostatic equilibrium for connections of different axonal and dendritic delays. Figure 2.11A shows the resulting heat map after learning with 'medium' EPSPs and inputs with modulation frequencies of 120Hz. This same result is shown in Figure 2.11B for gamma frequency inputs (60Hz) and 'slow' EPSPs (typical of pyramidal neurons) to model how this mechanism may work in the cortex. In both of these cases, the two-dimensional delay profiles that were learned showed a bias towards connections that have a linear relationship between their axonal and dendritic delays (with a slope of 1.0). We compared these analytic results (Figure 2.11A and B) to simulations of networks of 10 000 LIF neurons. As shown in Figure 2.11C and D, these results supported the analytic model.

In order for the network to show a selective response, it is the sum of the axonal and dendritic delays (not the difference between them) that is required to be tuned to a particular value. The diagonal black lines in Figure 2.11A and B show the connections that have the specific axonal and dendritic delays required for the network to have its largest response amplitude at the training frequency. It can be seen that these lines did not match at all with the delays selected during learning. The implications of this are addressed in the Discussion section.

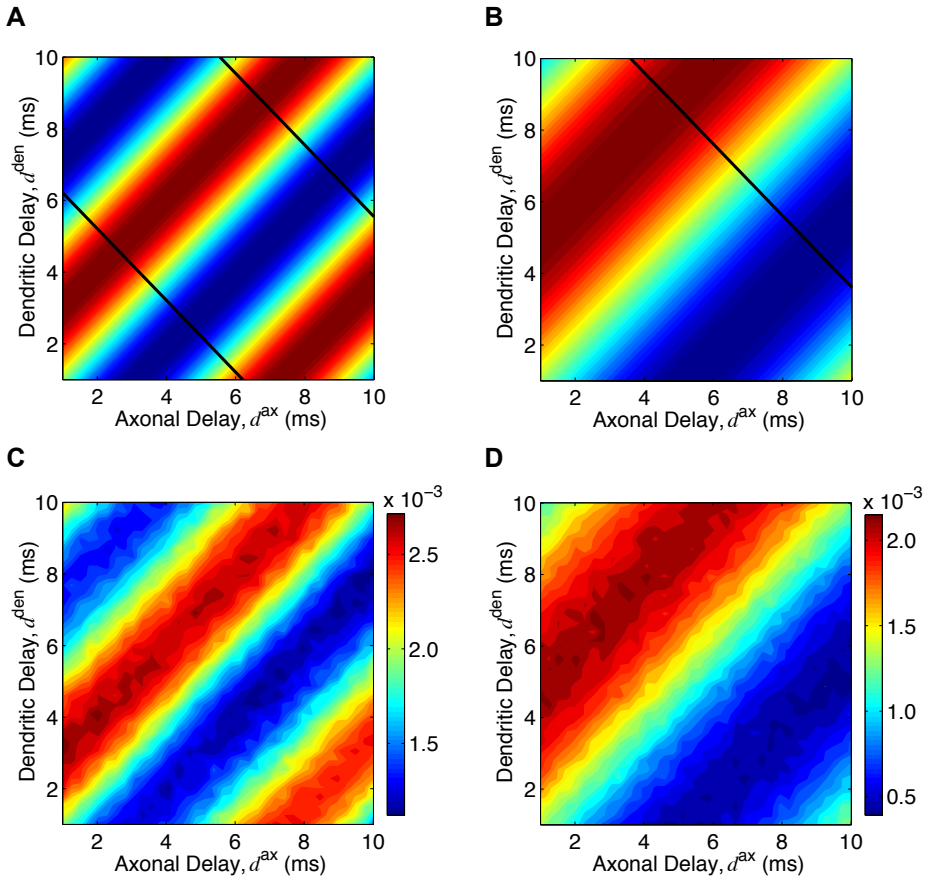


Figure 2.11: Axonal and dendritic delay selection. (A) Analytically determined heat map of potentiation/depression of connections with different axonal and dendritic delay with 120Hz inputs and 'medium' EPSP, as given by Equation (2.4.4). Regions of red correspond to potentiation and regions of blue correspond to depression. Black lines correspond to the delays that maximize the response at 120Hz. (B): Same as A but with 60Hz inputs and 'slow' EPSPs (black lines correspond to the delays that maximize the response at 60Hz). (C) Resulting heat map of mean connection strengths for different axonal and dendritic delays after simulating 500s of learning with 10 000 LIF neurons, 'medium' EPSPs, and 120Hz inputs. (D) Same as C but with 'slow' EPSPs, 60Hz inputs, and learning for only 50s. Note that no color bars are shown in A and B as the value of the weights is arbitrary; the mean depended on the homeostatic equilibrium and the amplitude on the learning duration.

2.4.4 Delay Selection with Multiple, Out-of-phase, Oscillatory Groups

Gamma oscillations (40-60Hz) are the highest modulation frequencies typically observed in the cortex. Axons within a single group of neurons would need to have delays of approximately 10-20ms to be selected by the STDP delay selection mechanism considered thus far. This may not be realistic for most axons in the cortex. We showed that for multiple groups of neurons receiving out-of-phase, oscillatory inputs it was possible for multiple, shorter delays (e.g., 1-10ms) to encode these lower frequencies (e.g., in the gamma

range). More interestingly, these delays could simultaneously encode the time lag between the groups. These different groups of neurons can be thought of as being in different cortical regions.

We extended the learning of oscillatory inputs through axonal delay selection to learning within and between multiple groups of neurons. These groups each received oscillatory inputs of the same frequency but with different phases to each other. This extension of the model to two groups of neurons (and inputs) is described in the Methods section. In this case, we defined $\bar{\mathcal{J}}_{ij}(d^{\mathrm{ax}})$ as the recurrent, weighted axonal delay profile for connections from group j to group i. These sets of connections are shown in Figure 2.12A using different colors. This matrix of axonal delay profiles (specifically the deviation of these profiles from the homeostatic equilibrium) was predicted to evolve, due to STDP, according to

$$\Delta\dot{\bar{\mathcal{J}}}(d^{\mathrm{ax}}) = \beta(f_{\mathrm{m}}) \begin{bmatrix} \cos\left[2\pi f_{\mathrm{m}} d^{\mathrm{ax}} + \phi_W(f_{\mathrm{m}})\right] & \cos\left[2\pi f_{\mathrm{m}}(d^{\mathrm{ax}} - \hat{d}_{\mathrm{lag}}) + \phi_W(f_{\mathrm{m}})\right] \\ \cos\left[2\pi f_{\mathrm{m}}(d^{\mathrm{ax}} + \hat{d}_{\mathrm{lag}}) + \phi_W(f_{\mathrm{m}})\right] & \cos\left[2\pi f_{\mathrm{m}} d^{\mathrm{ax}} + \phi_W(f_{\mathrm{m}})\right] \end{bmatrix}, \tag{2.4.5}$$

where, as before, $\beta(f_{\mathrm{m}}) = \frac{1}{2}\eta a^2 N_K^2 \bar{K}^2 r_\epsilon^2(f_{\mathrm{m}}) r_W(f_{\mathrm{m}})$.

An example of this delay selection at 60Hz with a phase difference of 0.78π (6.5ms) and 'slow' EPSPs (typical of pyramidal synapses in the cortex) is shown in Figure 2.12B. This shows the analytical prediction for the resulting delay profiles between the groups (red and green) after 25 000s of learning. It also shows the supporting simulations, which used two groups each of 5000 LIF neurons. In both the analytical and simulation results, the two within-group axonal delay profiles are shown in blue but are not easily seen. This is because these connections were almost completely depressed and their plots lie close to the horizontal axis. We investigated analytically how the axonal delays that were selected between the groups depended on the time lag in Figure 2.13. For 60Hz inputs, time lags of about 5ms to 12ms made it possible for the between-group connections to encode the frequency and time lag (for axonal delay ranges of 1-10ms).

The response of each of the two groups, after delay selection, depended upon both the frequency of the inputs, f_{m}, and the time lag between them, $\hat{d}_{\mathrm{lag}}$ (their relative phases). As with a single group, the response of the groups was oscillatory with the same frequency as the inputs. We considered only the case where the within-group connections were completely depressed and played no role in the response. Given this, the amplitudes of the responses of groups 1 and 2, respectively, (derived in the Methods section) are

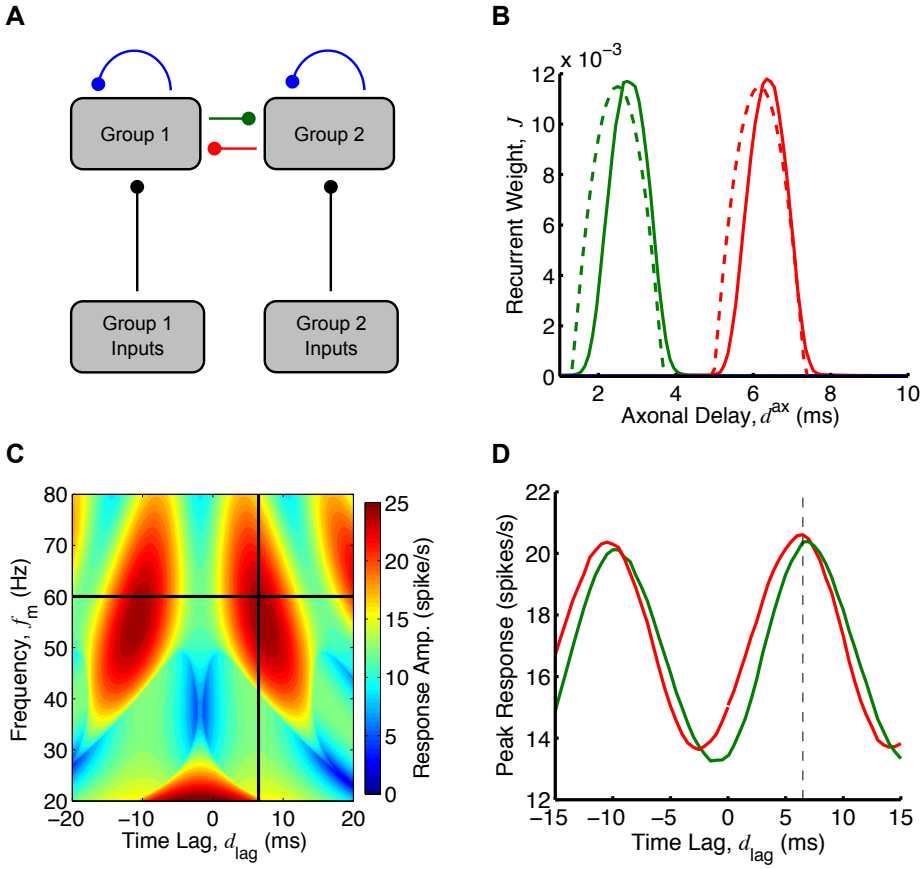


Figure 2.12: Axonal delay selection between two recurrently connected groups. (A) Diagram of two group model simplified from Figure 2.3. (B) Comparison between analytical (dashed) and simulation (solid) for axonal delay profiles between two groups each with oscillatory inputs with a modulation frequency of 60Hz and where the inputs into group 2 are 6.5ms behind the inputs into group 1. Analytical result was for 40 000s of learning and as given by Equation (2.4.5) and used a 'slow' EPSP. Simulation result was for 20 000s of learning with two groups of 5000 LIF neurons each. Shown are the delay profiles for the connections from group 2 to group 1 (red), from group 1 to group 2 (green), and within groups (blue) for which the mean weight for all delays was zero. (C) Analytically determined heat map (Equation (2.4.6)) of the mean of the group response amplitudes for different input frequencies and time lags. An EPSP with 1ms rise time and 3ms decay time was used instead of the 'slow' EPSP (Figure 2.14C), and recurrent strengths, $N_J\bar{J}_{12}$ and $N_J\bar{J}_{21}$, of 0.9. Black lines represent the training frequency (60Hz) and time lag (6.5ms). (D) Peak responses from simulations of group 1 (red) and group 2 (green) for different input time lags at the training frequency (60Hz), for the network after learning, shown in B. Dashed vertical line represents training time tag. Both B, C and D use a modulation amplitude, a, of 5 spikes/s, and the analytical plots in B and C used $N_K\bar{K} = 4$, to match the network response to the network response during the simulation with the nonlinear LIF neurons.

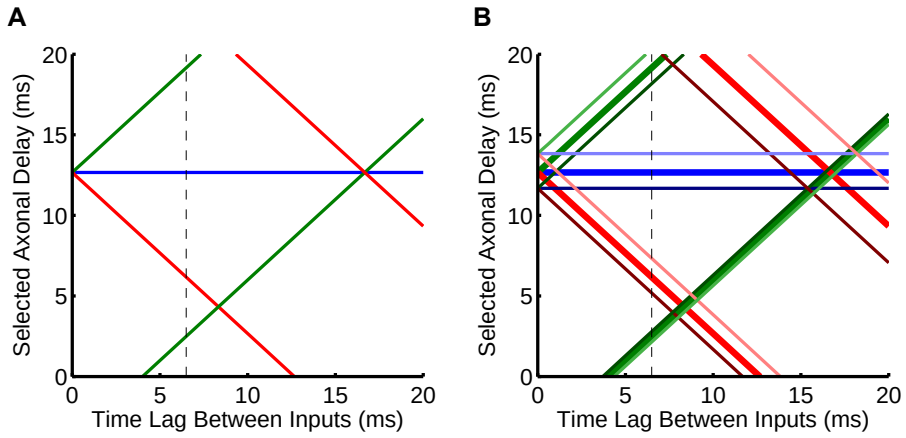


Figure 2.13: Selected axonal delays for different time lags between the inputs into two groups. (A) Analytical plot of the axonal delays selected by STDP for connections within each of two groups (blue), from group 1 to group 2 (green), and from group 2 to group 1 (red), with the time lag between the 60Hz oscillatory inputs into each group. The dashed line represents the 6.5ms time lag considered in more detail. (B) Same as A (thick lines) with additional lines for 55Hz (paler lines) and 65Hz (darker lines). Note that the three green lines (pale, thick and dark) in the bottom right of B are very close together.

approximated by

$$
\begin{aligned}
R_1(f_\mathrm{m},\hat{d}_\mathrm{lag}) \approx aN_K\bar{K}r_\epsilon(f_\mathrm{m})\Big\{1 &+ r_\epsilon^2(f_\mathrm{m})N_J^2\bar{J}_{12}^2 e^{-4(\pi\sigma_{12}f_\mathrm{m})^2} \\
&+ 2r_\epsilon(f_\mathrm{m})N_J\bar{J}_{12}e^{-2(\pi\sigma_{12}f_\mathrm{m})^2}\cos\big[\phi_\epsilon(f_\mathrm{m}) + 2\pi f_\mathrm{m}(\bar{d}_{12}+\hat{d}_\mathrm{lag})\big] \\
&+ 2r_\epsilon^2(f_\mathrm{m})N_J^2\bar{J}_{12}\bar{J}_{21}e^{-2(\pi f_\mathrm{m})^2(\sigma_{12}^2+\sigma_{21}^2)}\cos\big[2\phi_\epsilon(f_\mathrm{m}) + 2\pi f_\mathrm{m}(\bar{d}_{12}+\bar{d}_{21})\big]\Big\}^{\frac{1}{2}}, \\
R_2(f_\mathrm{m},\hat{d}_\mathrm{lag}) \approx aN_K\bar{K}r_\epsilon(f_\mathrm{m})\Big\{1 &+ r_\epsilon^2(f_\mathrm{m})N_J^2\bar{J}_{21}^2 e^{-4(\pi\sigma_{21}f_\mathrm{m})^2} \\
&+ 2r_\epsilon(f_\mathrm{m})N_J\bar{J}_{21}e^{-2(\pi\sigma_{21}f_\mathrm{m})^2}\cos\big[\phi_\epsilon(f_\mathrm{m}) + 2\pi f_\mathrm{m}(\bar{d}_{21}-\hat{d}_\mathrm{lag})\big] \\
&+ 2r_\epsilon^2(f_\mathrm{m})N_J^2\bar{J}_{21}\bar{J}_{12}e^{-2(\pi f_\mathrm{m})^2(\sigma_{21}^2+\sigma_{12}^2)}\cos\big[2\phi_\epsilon(f_\mathrm{m}) + 2\pi f_\mathrm{m}(\bar{d}_{21}+\bar{d}_{12})\big]\Big\}^{\frac{1}{2}},
\end{aligned}
\tag{2.4.6}
$$

where $\bar{J}_{ij}$, $\bar{d}_{ij}$, and σ_{ij} are the mean recurrent weight, and the mean and standard deviation (width) of the axonal delay profile, respectively, from group j to group i.

Figure 2.12D shows the peak response of each group in the network from Figure 2.12B observed in simulations with various time lags between the inputs. Figure 2.12C shows the mean of the analytically determined group response amplitudes, $R_1(f_\mathrm{m},\hat{d}_\mathrm{lag})$ and $R_2(f_\mathrm{m},\hat{d}_\mathrm{lag})$, respectively, against both the input frequency, f_m, and the time lag between the inputs, $\hat{d}_\mathrm{lag}$. The individual group response amplitudes of this average are shown in Figure 2.14A and B. Both the analytical results and the simulations showed that the network response was selective to the trained time lag. For these results, an EPSP with a rise time of 1ms and a decay time of 3ms was used. This was in the parameter

range between the 'slow' and 'medium' EPSPs, the results of which are shown in Figure 2.14C and D, respectively. The frequency that the network was predicted to be selective to differs from the training frequency, and these plots show that this difference strongly depended on the EPSP used. The strength of this selectivity, as with the response of a single group, depended on the mean strength of the recurrent connections. This is shown in Figure 2.14E, where the mean group response amplitudes are plotted for a network with weaker between group connections.

2.5 Discussion

In this chapter, we examined how STDP can lead to the selection of axonal and dendritic propagation delays within a recurrently connected network receiving oscillatory synaptic input. We found that the recurrent delays selected depend on the frequency of the oscillations and that this delay selection influences the response of the network to different input frequencies. We identified the conditions under which the resulting network was selective to the frequency of oscillation that it received during learning.

For learning with only axonal delays (assuming short dendritic delays), the range of frequencies that can be learned is limited by the range of axonal delays in the network. This can be seen from Equation (2.4.2). Here, the maximum delay (10ms in this chapter) sets the limit on the minimum frequency (76Hz) and the minimum delay (1ms) sets the limit on the maximum frequency (750Hz).

After delay selection, the frequencies that a network can possibly respond selectively to differ from the frequencies that it can learn. While the minimum frequency is the same, the maximum frequency is limited by how reliably the neurons and synapses in the network respond to these higher frequencies. This is illustrated in Figure 2.10A and B. We also observed in Figure 2.10A that networks trained with higher frequencies (e.g. 240Hz) have a higher response for low frequencies (e.g. 50Hz). This is likely due to the width of the oscillation peaks at these low frequencies being much larger than potentiated axonal delays in the network. In this case, the oscillation peaks become reinforced by recurrent feedback multiple times during a single oscillation, increasing the response at lower frequencies.

We found that short dendritic delays are necessary in order for a network to become selective in its response to a particular frequency. This can be seen with a range of axonal and dendritic delays, where the connections that are selected during learning (red stripes in Figure 2.11) have a large range of total propagation delays ($d^{\text{ax}} + d^{\text{den}}$). In other words, the learning is no longer a one-to-one mapping between frequency and delay. It has been suggested that, by considering unreliable synaptic transmission, STDP can differentiate between connections with the same difference in axonal and dendritic delays ($d^{\text{ax}} - d^{\text{den}}$) (Senn et al., 2002). In this chapter, however, the effect of unreliable synaptic transmission is minimal because the post-synaptic activity arises predominately from the inputs rather than the recurrent connections. Without this one-to-one mapping between frequency and delay, the learned network is not selective to any frequency (let alone the training

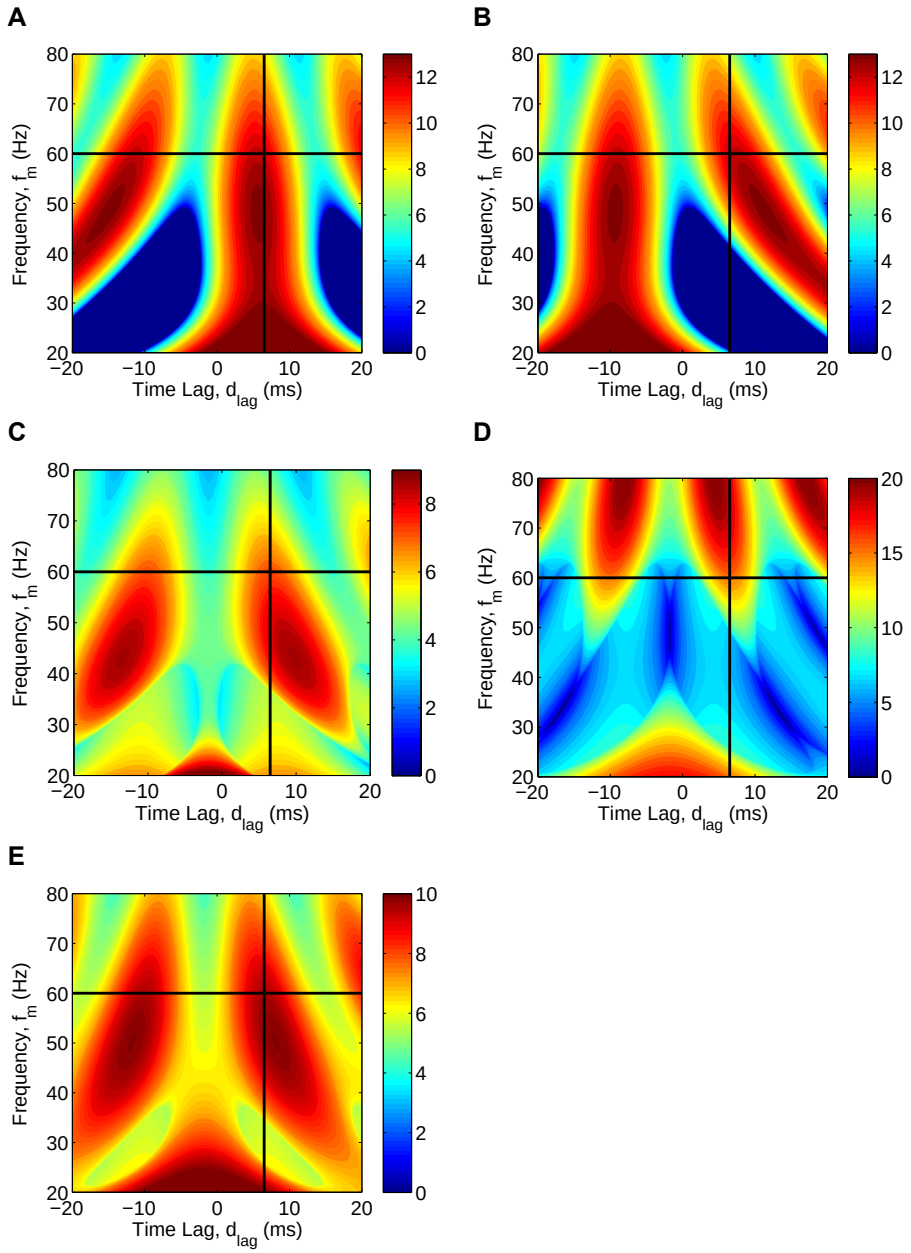


Figure 2.14: Analytical estimations of two group response amplitude to different inputs. (A) Response amplitude of group 1 for inputs with different frequencies and relative time lags, using Equation (2.4.6) with an EPSP with a rise time of 1ms and a decay time of 3ms, a modulation amplitude, a, of 5 spikes/s, feedforward strengths, $N_K\bar{K}$, of 1.0, and recurrent strengths, $N_J\bar{J}_{12}$ and $N_J\bar{J}_{21}$, of 0.9. (B) Same as A but for group 2. (C) Plot of the average between the response amplitudes of groups 1 and 2 as plotted for A and B but with a 'slow' EPSP. (D) Same as C but with a 'medium' EPSP. (E) Same as C but with the EPSP used in A and B and weaker recurrent strengths of 0.5.

frequency). If dendritic delays are not short but are sharply tuned, the network would become selective to a frequency but it would not be the training frequency. In a network with a range of dendritic delays and short axonal delays, the network would also become frequency selective but again not to the training frequency. While there are unlikely to be regions in the brain that contain only connections with short dendritic delays, the sub-networks made up of only these short dendritic connections would exist. Depending on the relative strength and correlation of the inputs from connections with longer dendritic delays, the learning and response of the networks would be as we have considered here.

The STDP learning window used here has the typical bimodality (depression for post-pre spike pairs and potentiation for pre-post pairs) observed with excitatory synapses. The axonal delays of the recurrent connections temporally shift this learning window, and are potentiated if they do so by the right amount and capture a peak in the neurons' periodic correlogram. It is important to note that, because the oscillatory activity of the neurons is due to the inputs, coincidence detection by means of recurrent axonal delays is not necessary for learning. However, this recurrent coincidence detection is relevant when considering the amplitude of the network's response for different input frequencies.

As discussed in the Results section, recurrent excitatory networks have a critical point where the positive feedback of the potentiated recurrent connections causes run-away increases in firing. The mean strength of the recurrent connections converges to a homeostatic equilibrium that depends on the homeostatic learning parameters, the learning window, and the level of correlation in the inputs. This may be above or below this critical point. For networks where the mean recurrent weights tends to increase, Lubenov and Siapas (2008) showed that STDP keeps this mean weight just below the critical point, provided the connections have axonal delays that exceed their dendritic delays. In this situation, synchronous firing is caused by the super-critical feedback. This synchronous firing is propagated along the recurrent axons arriving at neurons shortly after post-synaptic spikes, leading to synaptic depression. We observed this to be the case for the learning that was performed in this chapter (see Figure 2.5C). The networks taken after 20 000s of learning (see Figure 2.5D) all had mean recurrent weights just less than this critical value. We showed that the frequency selectively of the network increases dramatically as it approaches criticality.

We found our results to be robust with respect to the shape of the learning window used, provided it has the bimodality typical of excitatory synapses (see Figure 2.8). We also explored multiple neuron models (Poisson and LIF) and a range of synaptic time constants ('slow', 'medium' and 'fast' EPSPs). The delay selection learning within a single group requires a significantly large number of neurons in the network (see Appendix A.1 and Figure 2.15A). In spiking networks, a particular synaptic input may be the input that induces an output spike, which is referred to as a spike triggered correlation (Kempter et al., 1999; Gilson et al., 2009d). In the Poisson neuron model, this is captured by the autocorrelation of the external inputs and the autocorrelation of the neurons in the network. This effect is small if many inputs are required for a neuron to generate a spike. If the network is too small, these correlations dominate the shape of the delay

profile learned and prevent the encoding of the input frequency. For axonal delay selection between groups, where the spike triggering correlations are not present, smaller numbers of neurons (e.g. 1000) can demonstrate axonal delay selection than is possible within a single group (see Figure 2.15B).

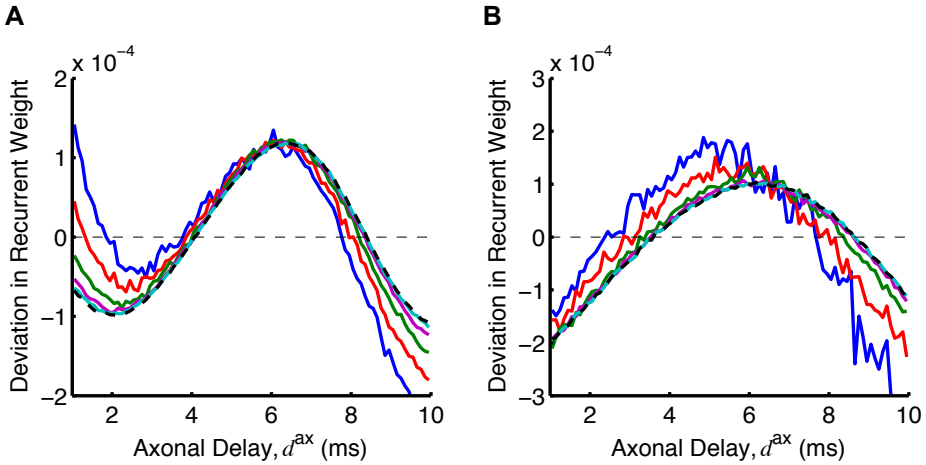


Figure 2.15: Axonal delay profile learned in networks of different sizes. (A) Deviation of axonal delay profile (from mean network weight) after 250s of learning for a single group of 500 (blue), 1000 (red), 2000 (green), 5000 (magenta), 10 000 (cyan), 20 000 (black, dashed) LIF neurons receiving 120Hz inputs and with 'medium' EPSPs. (B) Deviation of axonal delay profile (from mean network weight), for connections from group 2 to group 1, after 100s of learning for two groups, each with 250 (blue), 500 (red), 1000 (green), 2000 (magenta), 5000 (cyan), and 10 000 (black, dashed) LIF neurons, receiving different, out-of-phase (6.5ms), 60Hz inputs and with 'slow' EPSPs.

This chapter focused on additive STDP and more general STDP models remain to be fully explored (Morrison et al., 2008). Although not investigated here, we would expect a weight dependent (non-additive) learning rule to produce qualitatively similar results as long as the weight dependency leads to sufficient competition between the recurrent weights (van Rossum et al., 2000; Gütig et al., 2003; Morrison et al., 2007; Gilson et al., 2010a). Another variant of STDP that was not considered here is triplet STDP (Sjöström et al., 2001; Froemke and Dan, 2002; Froemke et al., 2006; Appleby and Elliott, 2007). Triplet STDP effectively modifies the learning window for different firing rates and captures correlations beyond the second order (Pfister and Gerstner, 2006). The correlations in our model arise solely from oscillations and there are no higher-than-second-order correlations (see Appendix A.8). Also, our results have been shown to be reasonably insensitive to the precise shape of the typical excitatory STDP learning window for the frequency range considered here (see Figure 2.8). Therefore, we would expect qualitatively similar results with a triplet STDP model, provided that the mean firing rates are in the range such that there is both LTP and LTD (approximately 5-30Hz (Pfister and Gerstner, 2006)).

In this chapter, we looked only at the learning of oscillatory (sinusoidal) patterns. However, this delay selection learning could potentially encode many other temporal pat-

terns. This chapter suggests the ranges of the frequency spectra of the types of signals that could be learned. A slightly different pattern to consider would be a periodic, phase-locked firing pattern. Here, the firing rate would be made up of a small base level with narrow, high intensity peaks that periodically occur. In this situation, we would expect delay selection to occur in a manner similar to that described in this chapter, or possibly faster. Trained networks would be expected to respond less to low frequencies (much lower than the training frequency). This is because the narrow peaks within the inputs would not allow the behavior observed in this chapter, where recurrent connections reinforced activity within the same oscillation peak (see Figure 2.10A).

Delay selection can also encode the oscillation frequency of signals in the feed-forward connection delays (Gerstner et al., 1996). This requires selecting two (or more) different delays, the difference of which would need to be a multiple of the oscillation period. Depending on the frequency, this may require an unrealistic range of delays in the connections, especially if these are connections within a local network.

First proposed by Hebb (1949), there has been growing experimental evidence for spatially distributed neuronal ensembles or cell assemblies (Harris et al., 2003; Buzsáki, 2004; Carrillo-Reid et al., 2008; Canolty et al., 2010) in which neurons far apart in the brain are synchronously active. This type of coordinated activity between distant regions (or cortical 'columns' in these regions) may be mediated by long range synaptic connections. The learning considered in this chapter, where delay selection is observed between multiple groups that are driven by out-of-phase, gamma oscillations, provides a possible mechanism by which this type of behavior could arise. Gamma (and other cortical) oscillations are known to be stronger during certain cognitive tasks (Jensen et al., 2007). This strengthening may correspond to the "activation" of particular ensembles between certain regions/columns, as illustrated in Figure 2.16A. These "activations", as described here, would be triggered by specific sets of phase differences as shown in Figure 2.16B. This would be similar to proposed gating (Vogels and Abbott, 2009) and routing mechanisms (Akam and Kullmann, 2010). If the learning we have considered here does lead to the formation of these types of neuronal ensembles, then it can be seen that it would be possible for these ensembles to merge with one another, split into multiple ensembles, grow in size, shrink, temporally bind with one another, or trigger, promote, stop, or block one another. Such ensemble activity is thus related to studies on neural syntax (Buzsáki, 2010), neural Darwinism (Edelman, 1993), synfire chains (Abeles, 1991), and polychronous groups (Izhikevich, 2006) (although at the larger scale of regions/columns than polychronous groups).

This chapter focused on the oscillatory activity in the network being due to oscillations in the inputs. These inputs may represent activity coming from sensory areas and representing a stimulus (e.g., sound), but they may also be involved in the coding of information coming from other brain regions. In this case, the oscillations would not necessarily represent the oscillatory nature of a stimulus, but instead have some functional role (such as the above mentioned cognitive role postulated for gamma oscillations). The oscillations in the networks may even be intrinsically generated and not due to the inputs at all. In a

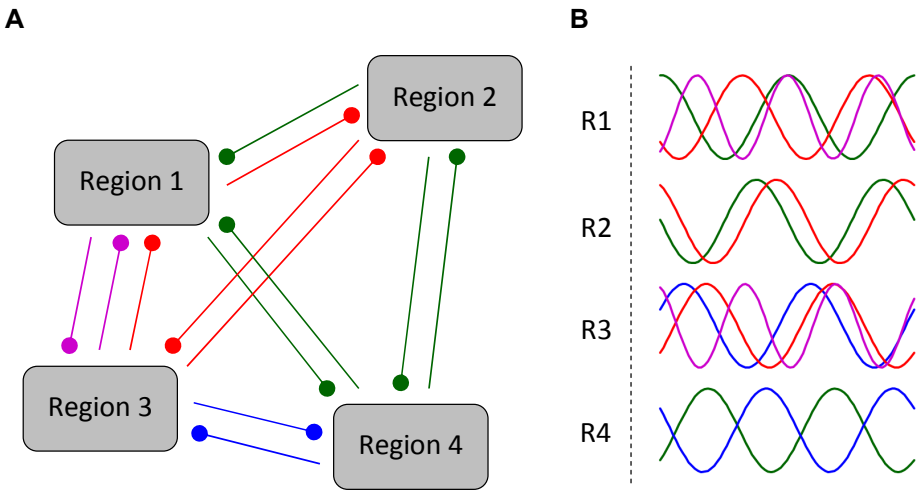


Figure 2.16: Illustration of different neuronal ensembles formed by delay selection. (A) Multiple groups/regions of recurrently connected neurons where connections between them have undergone axonal delay selection appropriate to the different out-of-phase, oscillatory activities present during the activation of each neuronal ensemble (blue, red, green and magenta). (B) Different out-of-phase, oscillatory activity of each group/region for the different neuronal ensembles (blue, red, green and magenta). In this network, only one neuronal ensemble can ever be 'active' at once. For each neuronal ensemble, the activities of the participating regions have the specific frequencies and relative phases that drove the delay selection and, after learning, resonate with the selected axonal delays of the recurrent connections.

related study, Câteau et al. (2008) looked at the oscillations that arose in a network due to STDP and the specific properties of the neuron model used. They observed functional clustering of neurons into groups with only weak within group connections. Since there was only a narrow range of axonal delays in the study, STDP was not able to perform delay selection. The study demonstrated that STDP instead selected the number of out-of-phase groups of neurons such that the connections between the groups all "shared" the oscillation delay.

Pyramidal-interneuron gamma (PING) is an established network configuration that generates intrinsic oscillations via the interactions between excitatory and inhibitory populations (Whittington et al., 2000; Brunel and Wang, 2003). If it is assumed that inhibitory connections in the cortex only act locally, local populations producing these oscillations would be connected to other local populations by only excitatory connections. These interconnected groups are similar to those considered in this chapter. The main difference is that their oscillations are internally generated rather than due to inputs. Though it remains to be considered, we would expect the synaptic changes to the excitatory connections between these groups to occur in a qualitatively similar manner to what was observed between the multiple groups considered in this chapter. Furthermore, when oscillations are internally generated, the phase of their oscillations is not fixed and can be perturbed. In this situation, it is possible that the delay selection between multiple groups could lead to the formation of stable patterns of relative phases or oscillatory states or modes. Al-

though the phase of the oscillations is not fixed, the frequency of the oscillations would be fixed at a frequency determined by the excitatory-inhibitory interactions (assuming that the within group connections remain fixed). Because of this, the frequency selectivity of the network response due to the delay selection would not be of interest. However, the phase selectivity of the network response would be relevant. We would expect a stronger response when the activity is in one of the stable oscillatory modes. Inputs to the network (not necessarily oscillatory) could possibly steer the network in and out of these different, self-sustaining modes. This formation of neuronal ensembles from intrinsic network oscillations may even be possible with multiple, distinct oscillation frequencies present (e.g., gamma and beta frequencies) (Kopell et al., 2010, 2011). This suggests another extension for the learning between multiple groups, which only considered a single frequency. Investigating how these networks could encode two (or more) different frequencies, instead of a single frequency and a time lag, is left as an interesting challenge.

Another example of where the delay selection mechanism might be employed in the brain is in the auditory brainstem. Here, sound information is encoded both spatially, with different frequency components being carried by spike trains of different nerve fibers, and temporally, with the precise timing of these spikes. The delay selection mechanism considered in this chapter may provide a way to extract the temporal information. Specifically, it could help explain how the brain can perceive the pitch of a sound with a missing fundamental, such as in telephone speech. The frequency range we have considered (100-300Hz) is typical of the pitches present in speech. The neuronal and synaptic time constants that we have used ('fast' EPSP) are consistent with those observed in the auditory brainstem. We demonstrated how oscillations in this range can be encoded into the axonal delays of a network, which becomes selectively responsive to this trained frequency. It remains to be explored whether networks could be trained instead with a complex sound containing a fundamental frequency (e.g., 120Hz) as well as its harmonics (e.g., 240, 360, and 480Hz). It would then be of interest whether the network became selective not only to this signal but also to the corresponding signal without the fundamental frequency. If this were shown with a sufficiently detailed simulation with realistic inputs, this mechanism would be a candidate for describing missing fundamental pitch perception. However, other possible mechanisms are likely to exist and these would each need to be tested against experimental data.

Throughout this chapter, we only considered networks of excitatory neurons. This is an important point to note because, as mentioned, studies have shown that interactions between excitatory and inhibitory populations can lead to the generation of intrinsic oscillations (e.g., PING) (Whittington et al., 2000; Brunel and Wang, 2003). Furthermore, the networks considered all operated in super-threshold, mean-driven regimes. This was done to facilitate the mathematical analysis and reduce the number of parameters to consider. The present analytic framework has not been extended to incorporate inhibitory post-synaptic potentials and this could be an area for future investigation. The current model provides a suitable description of the feed-forward behavior of neural processing in the auditory brainstem (Fay and Popper, 1992). Conversely, the cortex is generally considered to be a balanced network, operating in a fluctuation-driven regime (Tsodyks

and Sejnowski, 1995; van Vreeswijk and Sompolinsky, 1996; Burkitt, 2001; Burkitt et al., 2003; Shu et al., 2003; Meffin et al., 2004; Haider et al., 2006). Because of this, it is not clear how applicable the results in this chapter are to such networks. However, as we discussed, our results for multiple, out-of-phase groups would be expected to extend to the situation where local populations of excitatory and inhibitory neurons, internally generating (gamma) oscillations (through a mechanism such as PING) are connected to each other through long-range, excitatory connections. This situation may provide a more suitable model of the cortex (Roerig and Chen, 2002; Stepanyants et al., 2009).

STDP in a network regularly driven by oscillatory activity introduces a bias towards strong connections with a specific linear relationship between their axonal and dendritic delays, as shown in Figure 2.11. If a structural plasticity mechanism exists in the brain that physically removes weak connections not being used, then this predicts that a bias should be observed in the axonal and dendritic delays of connections in regions of the brain known to exhibit oscillatory activity. For example, the relationship between these delays shown in Figure 2.11B may be what is observed in regions of the cortex that show gamma oscillations. This prediction assumes that STDP works in the same manner for connections with long dendritic delays. While it is usually modeled in this way (Morrison et al., 2008), it is not clear whether this is consistent with experimental work (Sjöström and Häusser, 2006; Froemke et al., 2010).

Chapter 3

Cognitive Control with Cortical Units

This chapter is a slightly modified version of the article submitted to Frontiers in Neural Circuits:

> *Kerr RR, Grayden DB, Thomas DA, Gilson M, and Burkitt AN. "Goal-directed control with cortical units that are gated by both top-down feedback and oscillatory coherence."*

3.1 Abstract

The brain is able to flexibly select behaviors that adapt to both its environment and its present goals. This cognitive control is understood to occur within the hierarchy of the cortex and relies strongly on the prefrontal and premotor cortices, which sit at the top of this hierarchy. Pyramidal neurons, the principal neurons in the cortex, have been observed to exhibit much stronger responses when they receive inputs at their soma/basal dendrites that are coincident with inputs at their apical dendrites. This corresponds to inputs from both lower-order regions (feedforward) and higher-order regions (feedback), respectively. In addition to this, coherence between oscillations, such as gamma oscillations, in different neuronal groups has been proposed to modulate and route communication in the brain. In this chapter, we develop a simple, but novel, neural mass model in which cortical units (or ensembles) exhibit gamma oscillations when they receive coherent oscillatory inputs from both feedforward and feedback connections. By forming these units into circuits that can perform logic operations, we identify the different ways in which operations can be initiated and manipulated by top-down feedback. We demonstrate that more sophisticated and flexible top-down control is possible when the gain of units is modulated by not only top-down feedback but by coherence between the activities of the oscillating units. With

these types of units, it is possible to not only add units to, or remove units from, a higher-level unit's logic operation using top-down feedback, but also to modify the type of role that a unit plays in the operation. Finally, we explore how different network properties affect top-down control and processing in large networks. Based on this, we make predictions about the likely connectivities between certain brain regions that have been experimentally observed to be involved in goal-directed behavior and top-down attention.

3.2 Introduction

Our perception of the world around us and the way in which we respond to it depend on more than just the sensory information that is sent to our brains. It also depends on our recent and past experiences and on our current motivations and goals. While plasticity can make changes based upon past experiences, top-down processing allows numerous, faster changes (or switches) between stimulus-response mapping that can depend on recent events and current goals, as well as a more efficient way to allow interactions between concurrent stimuli.

The brain, in particular the cortex, exhibits a hierarchy both anatomically and functionally. Within this hierarchy, sensory information progresses "forward" through a series of regions. For example, in the visual system, stimuli cause neural activity that begins in the retina, propagates through the lateral geniculate nucleus (LGN) to the visual cortex, where it progresses through levels V1 and V2 before splitting into the dorsal (the "where" or "how" pathway) and ventral (the "what" pathway) streams and continuing further "upstream" (Goodale and Milner, 1992). In addition to this "forward" flow of information, there is much evidence that information also flows "backward" through this hierarchy. Buffalo et al. (2010) observed attentional effects that propagated from higher-order visual areas back to lower-order visual areas (i.e., V4 to V2 to V1).

This "backward" propagation of information, or top-down feedback, explains the observations by Womelsdorf et al. (2006, 2008) of context-dependent changes in the receptive field of neurons in visual cortical area MT. These changes included shifts of the centers of the receptive fields toward the focus of attention and narrowings of the receptive fields. Similar to this, Cohen and Newsome (2008) observed context-dependent changes in the noise correlations of MT neurons. Such top-down effects are also evident in goal-directed behavior, where the brain is able to perform fast switching between different "rules" that determine the appropriate response for a given stimulus. Wallis and Miller (2003) and Muhammad et al. (2006) showed how, during such a behavioral task, different neurons in the prefrontal (PRC), premotor (PMC), and inferior temporal (ITC) cortices and the striatum (STR) responded selectively to either the task rule (desired stimulus-response mapping), the behavioral response carried out, the visual stimulus being remembered, or whether or not the subsequent stimulus matched this remembered stimulus.

In order to perform tasks such as top-down attention and goal-directed behavior, the functional connectivity of cortical networks must be rapidly and flexibly modifiable.

Haider and McCormick (2009) reviewed the evidence for neural activity producing this rapid modulation in the functional connectivity. In this study, we focus on two different mechanisms for rapidly rearranging the functional connectivity of cortical networks: **gain modulation** and **communication-through-coherence**.

Gain modulation is where one type of input modulates the gain or sensitivity of a neuron to another type of input (Salinas and Sejnowski, 2001). Top-down gain modulation of neuronal responses that is dependent on contextual information or a different type of stimulus has been observed; however, the neuronal mechanisms underlying it have not been well understood. Larkum and colleagues found that pyramidal neurons exhibit a much stronger response when they receive inputs from both feedforward and feedback connections (Larkum et al., 1999; Larkum, 2013), which tend to be targeted to the cell's soma and basal dendrites and to the cell's apical dendrites, respectively (Felleman and Van Essen, 1991). This nonlinearity is due to interactions between the sodium and calcium spike initiation zones of pyramidal neurons, which are located at the soma and apical branch, respectively. This suggests that feedback connections to pyramidal neurons from higher-order regions can be thought of as modulating the gain of the neurons they target. While gain modulation provides a means for top-down processing or control, this has not been fully explored and there are limitations to the influence that is possible.

Synchronization and oscillations are ubiquitous in the cortex. Gamma oscillations, in particular, have been shown to be important in higher brain functions (Bartos et al., 2007; Fries et al., 2007), such as (selective) attention (Womelsdorf and Fries, 2007) and top-down processing (Engel et al., 2001). Communication-through-coherence (CTC) proposes that coherence between the oscillations of different neuronal groups modulates and routes communication through the brain (Fries, 2005). Supporting this hypothesis, synchronization and phase relations have been observed to govern interactions between neuronal groups (Womelsdorf et al., 2007). Gregoriou et al. (2009) showed that the prefrontal cortex and V4 exhibited long-range coupling of activity at gamma frequencies, initiated in the prefrontal cortex. There is much evidence suggesting that gamma (and beta) oscillations are involved in top-down and bottom-up interactions between the prefrontal and visual cortices (Benchenane et al., 2011).

Theoretical work has also shown how synchrony or coherence can act as a modulator of the gain of pyramidal neurons (Tiesinga et al., 2004; Börgers et al., 2005; Mishra et al., 2006; Tiesinga et al., 2008; Tiesinga and Sejnowski, 2009) and has also examined how top-down gain modulation can enable networks of neurons to perform fast stimulus-response remappings (Salinas, 2004). However, this situation has not been explored theoretically with neurons whose gain is simultaneously modulated by two different mechanisms: top-down (apical-targeted) feedback and oscillatory coherence. Furthermore, there has not been sufficient attention paid to understanding how gain modulation behaves and is controlled in hierarchical networks with several levels/layers.

In this chapter, we develop a simple neural mass model in which units exhibit gamma oscillations when they receive coherent oscillatory inputs to both the apical dendrites (feedback) and the soma/basal dendrites (feedforward). In this way, activity is modulated by

two different mechanisms: apical-targeted, top-down feedback and oscillatory coherence. We explore how these units can be formed into circuits to perform the same types of logic operations (e.g., "AND", "OR", and "NOT") considered by Vogels and Abbott (2005). Similar to previous studies involving gain modulated units (Salinas, 2004), we consider how these logic operations can be initiated and controlled (i.e., altered) by top-down feedback. However, unlike previous studies, we identify the different ways in which this top-down control can be implemented in hierarchical networks. As top-down gain modulation can strengthen or weaken the activity of neurons, we show that units can be added to or removed from a higher-level unit's logic operation by altering the feedback activity given. Furthermore, by modeling units as oscillating with a particular phase, we show that it is possible for feedback to modify the type of role that a unit has in the operation. This is not possible with top-down gain modulation alone and requires the additional coherence modulation. We explore how different network properties affect top-down control and processing in the networks, and make predictions about the likely connectivities between the different brain regions that have been experimentally observed to be involved in goal-directed behavior and top-down attention.

3.3 Methods

3.3.1 Cortical Unit Model

We model the cortex as being composed of a network of small units of pyramidal neurons and inhibitory interneurons (Figure 3.1A). These units are modeled as neural masses and the individual neurons are not explicitly modeled. The units receive two types of inputs: feedforward inputs to the soma and basal dendrites (blue) and feedback inputs to the apical dendrites (red). As proposed by Larkum (2013) for individual pyramidal neurons, we hypothesize that these units are associative and generate much stronger output when they are activated simultaneously by both of these types of inputs. We further hypothesize that these units exhibit gamma oscillations (at a particular frequency) when they are activated. Importantly, it is an assumption of the model that the activity is oscillatory in this way - the model does not generate these oscillations. While the model does not generate them, these oscillations represent fluctuations of the instantaneous spiking rate of neural populations that arise in the active units due to the reciprocal excitation and inhibition within the population. In addition to receiving both feedforward and feedback input, activation of units requires that these inputs are in phase, or coherent (Figure 3.1B). The requirement for units to receive both feedforward and feedback activity in order to become active can be thought of as binary gain modulation or a gating of the unit's activity (see Figure 3.1C).

Our model has a coarse time-step equal to half the period of a typical gamma oscillation (about 7-10ms). At a given time-step, the state of each unit, $s_{i,t}$, takes on one of three possible values:

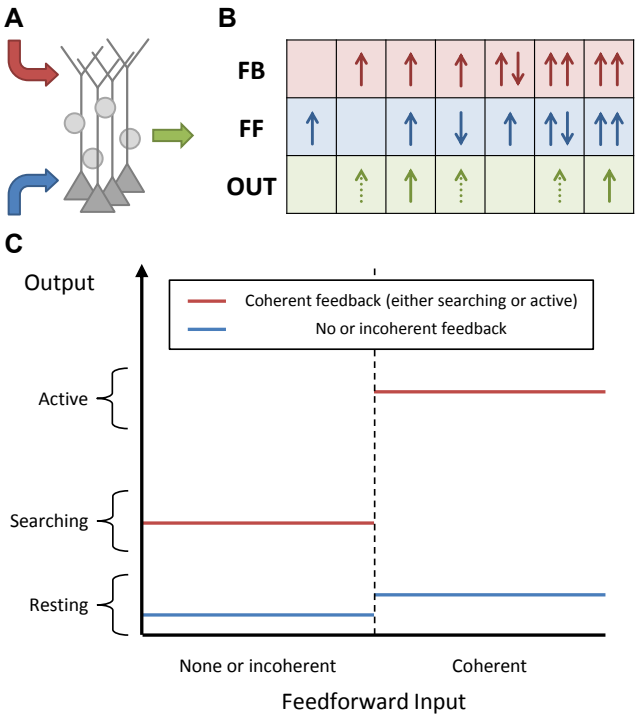


Figure 3.1: Diagram of Model. **A:** A cortical unit, composed of pyramidal neurons and inhibitory interneurons, exhibits activity (green) based on the feedforward, basal/soma-targeted (blue) and feedback, apical-target (red) inputs it receives. **B:** Table describing how the unit activity depends on these inputs, as described by Equations (3.3.2) and (3.3.3). The inputs and outputs are shown by solid and dashed arrows, which correspond to active and searching inputs/outputs, respectively. The direction of each arrow indicates the phase of gamma oscillations (active) or the timing of sporadic, feedback-propagating bursts (searching). The different rows correspond to feedforward and feedback inputs, and unit output, respectively. Multiple feedforward or feedback arrows indicate multiple inputs of these types. Note that the same effects are achieved with sporadic, bursting feedback inputs (but not so for feedforward inputs). **C:** Modulating effect of feedback on a unit's responsiveness to feedforward input, as described by Equation (3.3.2). Without feedback, the unit will remain in the resting state, regardless of the feedforward input. Coherent feedforward input must be coherent within itself but also with any feedback activity.

- **Resting:** units exhibit insufficient activity to affect other units;
- **Searching:** units exhibit strong but sporadic bursts of activity that can propagate and affect other units via feedback connections;
- **Active:** units exhibit strong, gamma-frequency activity that affect other units via feedforward or feedback connections.

The activity of units is confined to gamma oscillations and, consequently, units can only be active every second time-step and also must have one of two possible phases. While, the resting and active states correspond to the on and off states of binary models, the searching state represents a novel type of state, where units have not been fully activated

but are still able to pass down the feedback they receive to lower-levels. We refer to this as the searching state as it can be thought of as searching lower-level units that the unit sends feedback connections to. It is then able to ignite activity in units that are receiving feedforward activity. In this way, top-down feedback allows feedforward activity to propagate to higher-levels, which it would be unable to do otherwise.

While individual neurons are able to fire in response to only feedforward or feedback input, we are hypothesizing that, in higher-level areas of the cortex, groups/ensembles of neurons (units) are generally only able to be activated to a sufficient degree when the neurons in these groups receive both feedforward and feedback activity. Active units in our model exhibit strong gamma oscillations and can significantly affect the firing of other units that they are connected to. However, within a resting unit, the neurons are still assumed to be firing (perhaps even as gamma oscillations), although we have assumed that the firing is at a lower rate and is not sufficient to significantly affect the activity of other units that they are connected to. In this way, we are still modeling feedback as modulating the activity of groups of neurons (as illustrated by Figure 3.1C). However, because we only consider three different levels of activity (resting, searching, and active), the feedback modulation effectively becomes a gating of unit activity.

The state of each unit is determined by the inputs that it received from other units in the previous and current time-steps. These inputs come from other units via connections with short (∼0ms, negligible) or long time lags (∼7-10ms, one time-step). We denote the sets of short connections into unit i as $\hat{F}_i$ and $\hat{B}_i$ (feedforward and feedback, respectively) and the sets of long connections into unit i as $\bar{F}_i$ and $\bar{B}_i$ (feedforward and feedback, respectively). The presence of feedforward and feedback inputs into each unit are summarized by the Boolean expressions

$$\begin{aligned} f_{i,t} &= \left\{ \bigcup_{j \in \hat{F}_i} \left[s_{j,t} \text{ is active} \right] \right\} \cup \left\{ \bigcup_{j \in \bar{F}_i} \left[s_{j,t-1} \text{ is active} \right] \right\}, \\ b_{i,t} &= \left\{ \bigcup_{j \in \hat{B}_i} \left[s_{j,t} \text{ is active or searching} \right] \right\} \cup \left\{ \bigcup_{j \in \bar{B}_i} \left[s_{j,t-1} \text{ is active or searching} \right] \right\}, \end{aligned} \tag{3.3.1}$$

respectively. As mentioned above, the activity of units can be thought of as having a phase. In this view, activity arriving at a target unit will have the same phase as the source unit for connections with short time delays or the opposite phase to the source unit for connections with long time delays.

The state of unit i is given by

$$s_{i,t} = \begin{cases} \text{resting} & \text{if } \neg b^*_{i,t} \\ \text{searching} & \text{if } b^*_{i,t} \cap \neg f^*_{i,t} \\ \text{active} & \text{if } b^*_{i,t} \cap f^*_{i,t}, \end{cases} \tag{3.3.2}$$

where $f^*_{i,t}$ and $b^*_{i,t}$ are Boolean expressions for whether the basal/soma and apical compartments, respectively, of the pyramidal neurons in unit i receive coherent inputs and

become activated in time-step t. This is illustrated in Figure 3.1C and shows that units are only activated if both of these compartments are activated in its pyramidal neurons. The unit is in the searching state if the apical compartment is activated but the soma/basal compartment is not. The unit is in the resting state if the apical compartment is not activated.

Gamma oscillations are generated in the cortex through activation of either networks of inhibitory neurons, via the interneuron gamma (ING) mechanism, or reciprocally connected networks of excitatory and inhibitory neurons, via the pyramidal-interneuron gamma (PING) mechanism (Whittington et al., 2000; Brunel and Wang, 2003; Tiesinga and Sejnowski, 2009). Given the role that inhibitory neurons therefore have in producing gamma oscillations, we have assumed that the (basal/soma or apical) compartments of a unit are shut down by the inhibitory neurons if they receive non-coherent inputs. Put another way, the incoherent inputs interfere with the rhythmic interaction between the excitatory and inhibitory neurons in the unit. This is described by

$$\begin{aligned} f^*_{i,t} &= f_{i,t} \cap \neg f_{i,t-1}, \\ b^*_{i,t} &= b_{i,t} \cap \neg b_{i,t-1}. \end{aligned} \tag{3.3.3}$$

The compartments cannot be activated in consecutive time-steps as the inhibitory population constrains the activity to gamma oscillations. In this way, there is a phase given to the activity of units. By considering the model in terms of phases, the input/output relations in Figure 3.1B present another perspective. Provided there is coherent feedback inputs, there is at least the sporadic, searching signal (green dotted arrows) that can propagate down to lower levels. If, additionally, there are coherent feedforward inputs that are also in phase with the feedback, then the unit becomes active and exhibits strong gamma oscillations (green solid arrows).

3.3.2 Cortical Networks

In this chapter, we consider that these cortical units are organized into architectures similar to that presented in Figure 3.2. Here, the system receives sensory inputs (left) and produces motor outputs (right). Units in the system represent abstract concepts, such as percepts and actions, that depend on the sensory inputs and determine the behavior, respectively. In Figure 3.2, we divided the architecture into levels (using vertical black lines). These levels embody a hierarchy in the processing of information. Feedforward connections are made from units in lower levels to units in higher levels while the reverse is true for feedback connections. Here, the number of levels depicted is arbitrary and for illustrative purposes; the actual number of levels is most likely much greater. Similarly, the multiple vertical lines between the sensory and the percepts, and between the actions and motor, are only intended to indicate that there would be a number of levels of processing (e.g., for the visual pathway: those in the retina, LGN, V1, etc.) in between. The levels aim to convey the idealized version of the functional architecture that we consider in this paper.

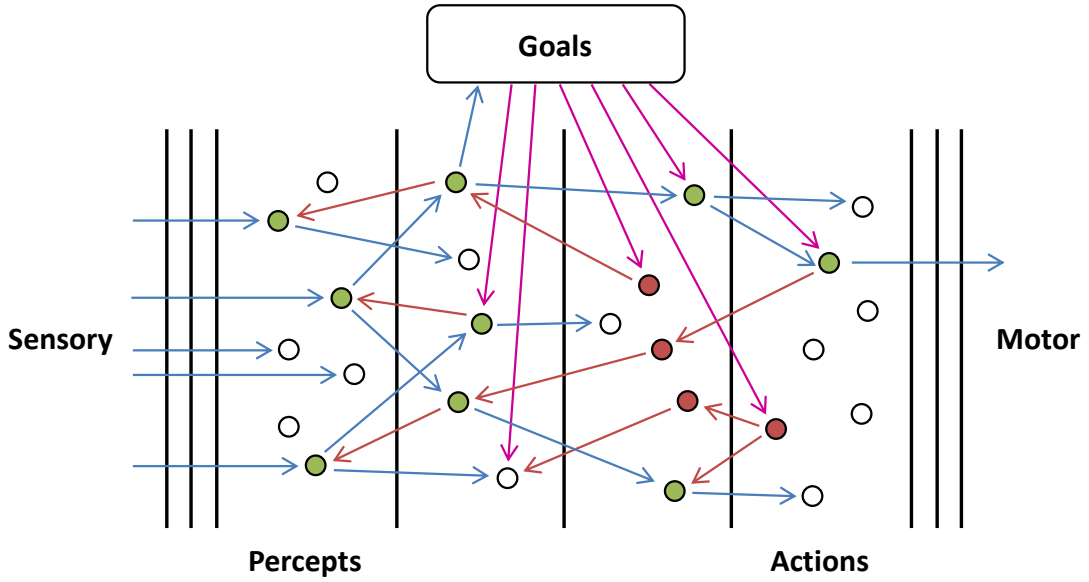


Figure 3.2: Goal-directed Network. An illustration of the proposed cortical architecture. Sensory, feedforward input (left) is mapped to percepts, actions, and finally motor responses (right), and this mapping is controlled by goal-dependent feedback (top). In the diagram, blue, red, and magenta arrows correspond to feedforward, internal feedback and external feedback (feedback corresponding to the goals of the system) connections, respectively. It should be noted that only the connections from active or searching units have been shown and they would exist other connections which have not been shown. White, green, and red units correspond to resting, active, and searching units, respectively.

Units in a network require feedback in order to become activated. For units in the networks that we are considering, this feedback must arrive from an external source, otherwise no units can become activated regardless of the sensory, feedforward inputs that they receive. We assume this external feedback arrives from higher-level networks or areas of the brain. We rely on the assumption that there exists at least one high-level region that provides this feedback to the rest of the brain without receiving feedback itself. This feedback would be dependent on the goals, motivation, and state of the system (working memory), and would control the way in which the network causes percepts to lead to actions. While these goals must be reasonably persistent, feedforward activity of certain percepts would assumedly have the ability to affect these goals; however, we are not going to consider how these goals persist or change in this paper.

In order for arbitrary mappings from percepts to actions to be made, units receive feedback activity and perform logic operations on their inputs. In a sense, they are "asking" questions or testing hypotheses regarding the state of these inputs. Higher-level units will in turn use the outputs of these lower-level units as inputs. As illustrated in Figure 3.2, lower-level units, or groups of units, represent different percepts formed about the sensory information received while higher-level units, or groups of units, begin to more appropriately resemble different courses of action for the system to perform. The logic

operations that each unit performs can be thought of as hypotheses about the state of the external world and hypotheses about what (if any) actions should be carried out.

We consider networks of units to be composed of smaller subnetworks. For the subnetwork being considered, we refer to the lowest-level units as being the inputs and the highest-level units as being the output units. These units are identical in terms of how they are modeled but play different roles in the larger network. We assume that only input units receive feedforward activity from lower-level units outside the subnetwork and only output units send feedforward activity to higher-level units outside the subnetwork (it is only their activity that matters to higher-level areas).

Just as the units in our model require feedback to become active, networks (and subnetworks) of these units require external feedback in order for any of their units to become active. For a given subnetwork, external feedback can not only be to output units, it can also be to input units and intermediate units (units that are neither input nor output units). We refer to external feedback that arrives at output units as **initiating** feedback (as it initiates the output units) and external feedback that arrives at lower-level units in the subnetwork as **orchestrating** feedback (as it orchestrates or manipulates the operations performed by the output units). These two types of feedback are not different in the way that they affect units but are distinguished by the different functional roles they play. Importantly, their roles are specified relevant to the subnetwork being considered. For example, orchestrating feedback for one subnetwork may be initiating feedback for another subnetwork. In the network in Figure 3.2, external feedback is represented by the dotted red arrows that are dependent on the current goals.

3.3.3 Logical Operations

We consider examples of simple subnetworks, or motifs, in order to illustrate the functional roles of different types of connections. Shown in Figure 3.3, these motifs perform simple logic operations when initiated by external feedback. The output units, Y1, Y2, and Y3, send feedback activity to input units, X1, X2, X3, X4, X5, and X6, and become activated if they receive feedforward activity in return. Y1 ("X1 or X2") becomes activated if either X1 or X2 receives coherent, feedforward input, which, because of the long time lag of the connections, must be out-of-phase with the activity of Y1 so that they can provide returning feedforward activity that is coherent with Y1's activity. Y2 ("X3 and not X4") becomes activated only if X3 receives coherent, feedforward input and X4 (which makes a short feedforward connection onto Y2) does not, as activity from this unit would arrive out-of-phase with Y2's activity. In the last motif, the unit in the intermediate layer performs the same operation ("AND NOT") on its inputs as Y2. Y3 ("X5 and X6"), in turn, also performs the same ("AND NOT") operation as Y2 except that the intermediate unit is initiated in phase with Y3 and so the time lags of the connections between them are reversed.

Generally, we denote whether or not the output unit i is activated by y_i, which can either be true or false. This is determined by the operation that the unit performs, which

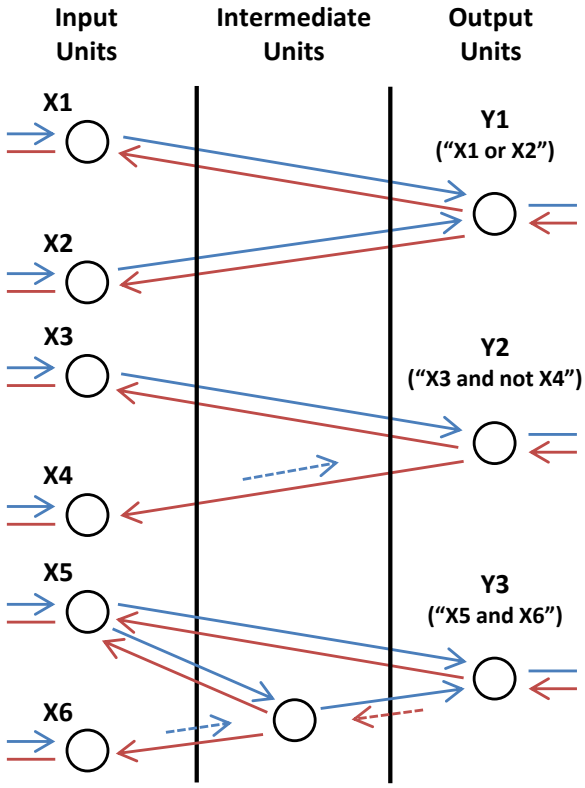


Figure 3.3: Basic Logic Motifs. From top to bottom, motifs in which the output units compute the operations "X1 or X2", "X3 and not X4", and "X5 and X6", respectively. Arrows correspond to feedforward (blue) and feedback (red) connections with short (dashed and shorter length) and long (solid and longer length) time lags. Arrows are also shown connecting the input units to lower-level units (not shown) and connecting the output units to higher-level units (not shown) as these motifs function as circuits in a larger network.

is given by the binary function

$$y_i = g_i(\mathbf{x}; \mathcal{B}; \hat{\mathcal{B}}), \tag{3.3.4}$$

where the binary vector $\mathbf{x}$ denotes whether or not each of the input units are receiving (external) feedforward activity of the appropriate phase. In this way, $\mathbf{x}$ depends of the state of the units that are external to the subnetwork and that make feedforward connections onto each of the inputs. More specifically, for a desired phase, $\mathbf{x}_i$ indicates whether $f^*_{i,t}$ is true on the appropriate time steps. The other variables $\mathcal{B}$ and $\hat{\mathcal{B}}$ denote the set of external feedback (and phase of the feedback) that each output unit receives (initiating feedback) and each lower-level unit receives (orchestrating feedback), respectively. For output units, $\mathcal{B}$ determines the values of $b^*_{i,t}$. For the other units in the subnetwork, $\hat{\mathcal{B}}$, in combination with the states of other units in the subnetwork, which make feedback connections onto these units, determines the values of $b^*_{i,t}$. We refer to $\mathcal{B}^\phi$ and $\hat{\mathcal{B}}^\phi$ as the empty feedback sets, where there is no initiating and orchestrating feedback, respectively, and to $\mathcal{B}^i$ as the

set in which only output unit i receives initiating feedback. Unless $\mathcal{B}$ includes feedback to unit i, $g_i(\mathbf{x}; \mathcal{B}; \hat{\mathcal{B}}) = 0$ for all inputs. As we have motifs that can perform the "OR" and "NOT" operations, we can compose these together to form networks to perform arbitrary logic operations.

3.4 Results

The brain is able to perform arbitrary mappings from different stimuli to different behavioral responses and to rapidly modify its mappings based on current motivations and goals. In order to understand how this occurs, we have proposed a model of groups or ensembles of neurons that when connected into networks can produce arbitrary mappings. We explore the different ways in which they can be controlled by top-down feedback from higher brain regions and how this depends on the connectivity of these networks.

3.4.1 Top-down Processing

Goals influence the operations of units through providing external feedback to the network. For a given set of output units, this feedback can be divided into initiating feedback, which targets the output units, and orchestrating feedback, which targets lower-level units. Where the goals send initiating feedback to multiple output units, the operations performed by these units may be either **non-interacting** or **interacting**. Where the goals send orchestrating feedback to units in the network, the operations may be **orchestrated**. Each of these types of operations shall now be described in turn.

Non-interacting operations

Operations performed by output units are non-interacting if the units perform the same operations when they are initiated together as they did when initiated separately. This means that the operations can be performed in parallel without affecting each other. Using functional notation, we say that the operation of unit j does not interact with the operation of unit i for orchestrating feedback to the network $\hat{\mathcal{B}}$ if, for all sets of inputs $\mathbf{x}$,

$$g_i(\mathbf{x}; \mathcal{B}^i \cup \mathcal{B}^j; \hat{\mathcal{B}}) = g_i(\mathbf{x}; \mathcal{B}^i; \hat{\mathcal{B}}), \tag{3.4.1}$$

where $\mathcal{B}^i$ denotes the set with initiating feedback to only output unit i and $\mathcal{B}^i \cup \mathcal{B}^j$ denotes the union of the sets $\mathcal{B}^i$ and $\mathcal{B}^j$ with initiating feedback to only output units i and j.

Interacting operations

Interactions occur when a unit's operation is modified by other units being initiated alongside it (i.e, when output units are initiated together). The number of operations that can be performed in parallel is limited by the number of interactions that occur. In the most

extreme case, there is only one information channel due to the dependencies between the units and collectively the output units only perform a single, more complex operation. Interactions allow top-down processing, as feedback into one unit can affect the operations performed by other units in the network. Using functional notation as for non-interacting operations, we say that the operation of unit j interacts with the operation of unit i for orchestrating feedback to the network $\hat{\mathcal{B}}$ if there exists a set of inputs $\mathbf{x}$ for which

$$g_i(\mathbf{x}; \mathcal{B}^i \cup \mathcal{B}^j; \hat{\mathcal{B}}) \neq g_i(\mathbf{x}; \mathcal{B}^i; \hat{\mathcal{B}}). \tag{3.4.2}$$

Figure 3.4 illustrates the contrast between interacting and non-interacting operations. It also demonstrates how small changes to a network with non-interacting operations cause it's operations to interact. Figure 3.4A shows a network where two operations have overlapping inputs but no interactions occur. Figure 3.4B and C provide examples of networks with overlapping motifs where the same operations are performed when the outputs are initiated separately but interactions occur when they are initiated together. These interactions are evident in the table in Figure 3.4D, where there exist inputs for which a different output is produced depending on whether the two hypotheses are initiated separately or together. The last row in Figure 3.4D contains the number of inputs that are involved in the operations performed. Interactions can cause this to either increase or decrease.

Orchestrated operations

Orchestrating feedback (from an external source) can alter the operations that an output unit in a subnetwork performs. In this case, the feedback manipulates or controls the operations that are performed. We found dependencies between interactions that occur between units and the level and type of control that is possible by orchestrating feedback. We say that the operation of unit i has been orchestrated if the orchestrating feedback $\hat{\mathcal{B}}$ causes there to exist a set of inputs $\mathbf{x}$ for which

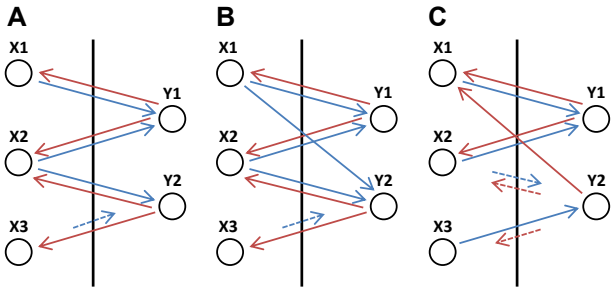
$$g_i(\mathbf{x}; \mathcal{B}^i; \hat{\mathcal{B}}^\phi) \neq g_i(\mathbf{x}; \mathcal{B}^i; \hat{\mathcal{B}}), \tag{3.4.3}$$

where $\hat{\mathcal{B}}^\phi$ denotes the empty set where there is no orchestrating feedback to the network.

Figure 3.5 shows the mechanisms by which external feedback **adds** or **removes** units from an operation. Without any external feedback, the operation performed is "X2 or X3". By adding the external feedback, X2 is removed from the operation and X1 is added, making the operation "X1 or X3". This could be orchestrating feedback from outside the network (as we have shown) but it could also be feedback from another output unit that is initiated with Y1.

3.4.2 Goal-directed Behavior

Figure 3.6 shows examples of networks that can perform possible stimulus-response experiments, where switching between different rules or goals, is required. Each of the

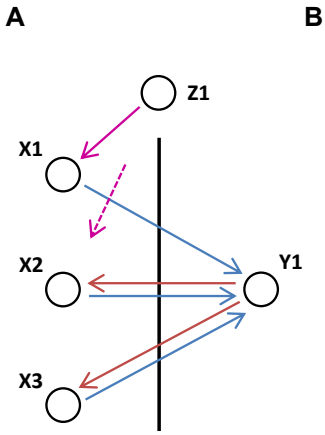


D

			A/B/C				A		B		C	
Inputs			Y1 Alone		Y2 Alone		Y1 + Y2		Y1 + Y2		Y1 + Y2	
X1	X2	X3	Y1	Y2	Y1	Y2	Y1	Y2	Y1	Y2	Y1	Y2
0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	1	0	0	0	1	0	1	1	0	0
0	1	0	1	0	0	1	1	1	1	1	1	1
0	0	1	0	0	0	0	0	0	0	0	0	0
1	1	0	1	0	0	1	1	1	1	1	1	1
1	0	1	1	0	0	0	1	0	1	0	0	0
0	1	1	1	0	0	0	1	0	1	0	1	0
1	1	1	1	0	0	0	1	0	1	0	1	0
# of inputs			2	0	0	2	2	2	2	3	1	2

Figure 3.4: Non-interacting and Interacting Operations. A: Two output units, Y1 and Y2, which individually perform operations "X1 or X2" and "X2 and not X3", respectively. Feedforward connections to input units from lower-level units and feedforward connections from output units to higher-level units (as shown in Figure 3.3) have been omitted. **B:** Same as A but with an additional feedforward connection which does not change the individual operations but introduces interactions when they are initiated together. **C:** Same as A but Y2 instead needs to be initiated in phase with the input units for it to perform the same operation. There is also an additional feedback connection that, similar to the additional connection in B, does not change the individual operations but introduces interactions. **D:** Input-output table for the networks in A, B, and C. The input units (or cues), X1, X2, and X3, either receive feedforward input (1) or not (0), and the output units, Y1 and Y2, are either activated (1) or not (0) for each of the networks initiated with external feedback to only Y1, only Y2, or to both Y1 and Y2. Green (red) outputs are ones that are activated (not activated) when the units are initiated together but were not activated (activated) when the units were initiated separately. The final row indicates the number of inputs that the output unit's operation depends upon (relevant inputs), where green (red) indicates that the number has increased (decreased) from being initiated separately to being initiated together.

networks in Figure 3.6A, B, and C, has two different percepts (stimulus cues) as inputs, two different actions (levers to pull) as outputs, and a number of different goals, rules, or stimulus-response mappings that direct how these inputs lead to different outputs. The tables correspond to the binary functions, $g_i(\mathbf{x}; \mathcal{B}; \hat{\mathcal{B}})$, where i corresponds to the two levers, or output units, $\mathbf{x}$ denotes the cues that are present, and $\mathcal{B}$ and $\hat{\mathcal{B}}$ denote the initiating



B

Cues			Outputs	
X1	X2	X3	Y1 (-Z1)	Y1 (+Z1)
0	0	0	0	0
1	0	0	0	1
0	1	0	1	0
0	0	1	1	1
1	1	0	1	1
1	0	1	1	1
0	1	1	1	1
1	1	1	1	1
# of inputs			2	2

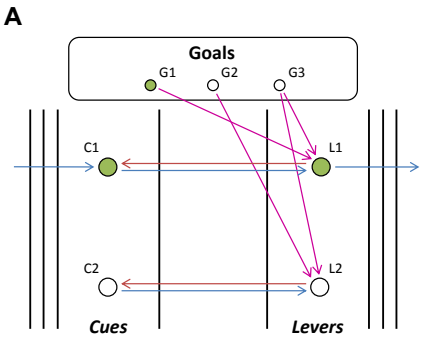
Figure 3.5: Orchestrated Operations. A: Motif performing an "OR" operation over the two input units X2 and X3 has the input unit X2 removed and X1 added by orchestrating feedback (magenta arrows) from unit Z1 which, like Y1, is initiated out-of-phase with the inputs. **B:** Input-output table for the network in A. The input units (or cues), X1, X2, and X3, either receive feedforward input (1) or not (0), and the output unit Y1 is either activated (1) or not (0), in the cases where there is feedback or not from Z1. The final row indicates the number of inputs that the output unit's operation depends on (relevant inputs).

and orchestrating feedback, respectively, from the goal units.

In Figure 3.6A and C, each of the goals only sends initiating feedback to the network. In Figure 3.6A, there are three different goals: G1, to pull lever L1 when cue C1 is presented (and ignore C2); G2: to pull lever L2 when cue C2 is presented (and ignore C1); and G3, to pull lever L1 when cue C1 is presented and pull lever L2 when cue C2 is presented (perform both goals in parallel). These goals control the network through feedback to the two units corresponding to the two lever actions. The inputs of the output units do not overlap, so trivially their operations (performed separately by G1 and G2) are non-interacting and so can be performed in parallel (by G3). Figure 3.6B shows the input-output table for this.

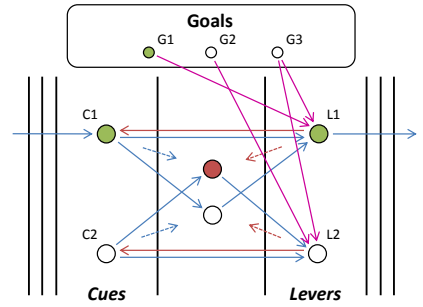
In Figure 3.6C, there are also three different goals, the first two of which are the same as the first two in the network in Figure 3.6A. The third goal is to pull levers L1 and L2 when both cues C1 and C2 are presented (and ignore both C1 and C2 presented alone). The same feedback from the three goals as in Figure 3.6A is used to control the network, but different behavior (Figure 3.6D) arises due to differences in the networks. In Figure 3.6B, we see that the conditions for L1 and L2 remain the same regardless of whether they are included in the task or not (i.e., the logic operations can be performed in parallel). However, we see that this is not the case in Figure 3.6D, where different logic operations are performed when both L1 and L2 are included compared to when they are considered alone. In this case, the two operations interact.

In Figure 3.6E, unlike in Figure 3.6A and C, the goals (except G*, which is always active) send orchestrating feedback, which targets intermediate units in the network. This



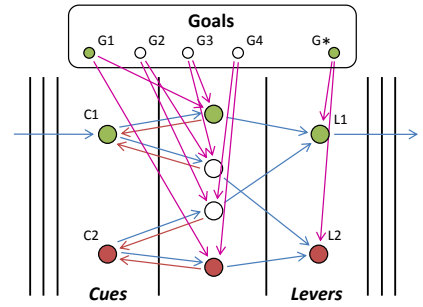
B

Cues	G1 (C1-L1)		G2 (C2-L2)		G3 (parallel)	
C1	C2	L1	L2	L1	L2	L1	L2
0	0	0	0	0	0	0	0
1	0	1	0	0	0	1	0
0	1	0	0	0	1	0	1
1	1	1	0	0	1	1	1
inputs		1	0	0	1	1	1



D

Cues	G1 (C1-L1)		G2 (C2-L2)		G3 (both)	
C1	C2	L1	L2	L1	L2	L1	L2
0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0
0	1	0	0	0	1	0	0
1	1	1	0	0	1	1	1
inputs		1	0	0	1	2	2



F

Cues	G1 (standard)		G2 (reversed)		G3 (L1-only)		G4 (L2-only)	
C1	C2	L1	L2	L1	L2	L1	L2	L1	L2
0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	1	1	0	0	1
0	1	0	1	1	0	1	0	0	1
1	1	1	1	1	1	1	0	0	1
inputs		1	1	1	1	2	0	0	2

Figure 3.6: Stimulus-response Tasks. **A:** An example of how the cortical architecture would be utilized for a stimulus-response task where a subject pulls one of two levers when presented with one of two sensory cues. The task switches between one of three "goals": lever L1 should be pulled for cue C1 and cue C2 ignored (G1), L2 for C2 and C1 ignored (G2), and L1 for C1 and L2 for C2 (G3). Network activity is shown for when G1 is active. Similarly, the arrows to the cue units from the left and those leaving the lever units depict the inputs and outputs of the network (i.e., only "active" connections) for a particular set of inputs. **B:** Input-output tables for the network shown in A for the three different goals. The final row indicates the number of inputs that the output unit's operation depends on (the relevant inputs). **C:** Same as A but the third task (G3) now involves pulling both levers if and only if both cues occur together. Note that the feedback from the three goals is the same as in A but there is an extra layer in the network. **D:** Same as B but for the network in C. **E:** Similar to A and C but with four different goals: L1 should be pulled for C1 and L2 for C2 (G1), L2 for C1 and L1 for C2 (G2), L1 for either C1 or C2 (G3), and L2 either C1 or C2 (G4). G* is not actually one of the four goals but instead always provides feedback (each of the goals could instead provide this feedback). The feedback from the goals is no longer only to output units. **F:** Same as B and D but for the network in E.

allows Figure 3.6E to demonstrate a more complex stimulus-response situation with four different goals (tasks) that the system needs to switch between. These are: G1, to pull lever L1 when cue C1 is presented and L2 when C2 is; G2, to pull L2 for C1 and L1 for C2; G3, to pull L1 for C1 and L1 for C2; and G4, to pull L2 for C1 and L2 for C2. It is due to the orchestrating feedback, which modifies the way the network maps its sensory inputs to its behavioral outputs, that the network is able to perform each of these tasks (Figure 3.6F).

3.4.3 Phase-dependent Operations

We found that the relative phase between the output units and the units to which they send feedback can also affect the computations performed. If a network has feedback connections with a mix of short and long time lags and all the inputs are in phase, then some of the units receiving feedback will be activated and others blocked. Figure 3.7A shows this situation, where switching the phase of the feedback switches which of the input units are activated and which are blocked. It, therefore, modifies the set of inputs that are used in the operation. For example, if Y1 is initiated in phase with the inputs, then only X2 and X5 will be involved in the operation ("X2 and not X5"). But, if Y1 is initiated out-of-phase with the inputs, then only X3 and X4 will be included ("X3 and not X4"). Inputs, such as X1, are persistent and involved in the operation regardless of the phase if they do not receive the feedback from the output units but instead feedback from another source (Z1 in this case).

Unlike gain modulation models where units typically do not have phase, we discovered that, rather than simply adding a unit to an operation (increasing its gain), units can be added with different phases and play a different role in the network. This is shown in Figure 3.7B, where Y2 and Y3 each send feedback of a different phase to an intermediate unit, X5, adding it to the operation performed by Y1. While both send feedback that adds X5 to the operation of Y1, the different phases of the feedback cause X5 to play a different role in the operation of Y1. With feedback from Y2, X5 is initiated in phase with Y1 and so, due to its long feedforward connection, the operation of Y1 becomes "X4 and not X5". With feedback from Y3, X5 is initiated out-of-phase with Y1 and the operation of Y1 instead becomes "X4 or X5". In turn, the different phased feedback causes X5 to perform different operations on its own inputs (X2 and X3). In this situation, we are only concerned with the operation of Y1. However, Y2 and Y3, in addition to modifying the operation of Y1 may also be performing their own operations with their own sets of inputs but these are not shown.

3.4.4 The Role of Network Properties

We investigated how the properties of the feedforward and feedback connections in the network determine the extent to which interactions occur and the operations can be orchestrated. For example, the network in Figure 3.4B is the same as the one in Figure

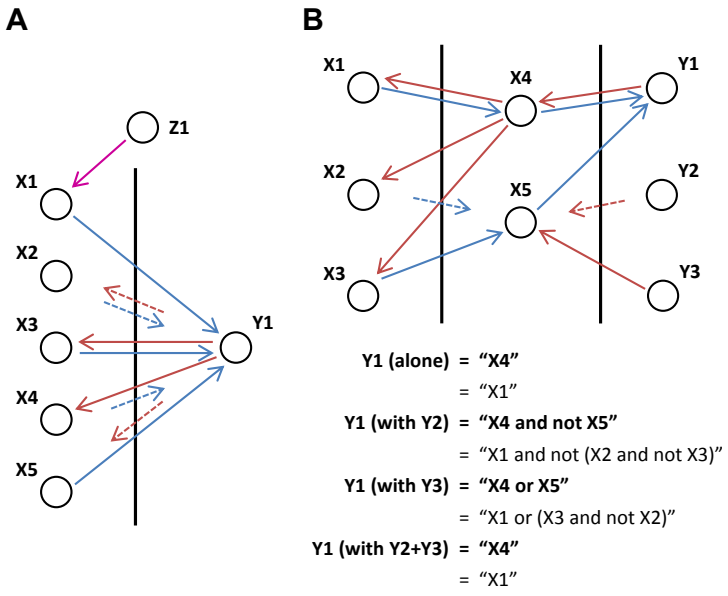


Figure 3.7: Phase-dependent Operations. A: An output unit connected to the five input units, X1-X5, which receive orchestrating feedback from unit Z1 that is out-of-phase with the inputs, performs the operation "(X1 or X3) and not X4" when it is out-of-phase with the inputs and "X3 and not X1 and not X5" when it is in phase with the inputs. **B:** The operation performed by output unit Y1 is changed depending on which other output (or external) units are initiated with it. Intermediate unit X5 is added by either Y2 or Y3 but in different ways, causing it to play a different role in the operation of Y1, and to perform different operations on its own inputs (X2 and X3).

3.4A but with an additional feedforward connection. While this does not affect the operations when they are initiated separately, this additional feedforward connection changes the operations when they are performed together: the input X1 is added to the operation performed by Y2. Figure 3.4C also performs the same operations as Figure 3.4A and B when they are performed separately, provided that the feedback to Y2 is in phase with the inputs. However, due to a feedback connection from Y2 to X1, the operations interact when they are initiated together: the input X1 is removed from Y1's operation. Similar to feedback from initiating another output unit, orchestrating feedback from Z1 in Figure 3.5A modifies the operation of Y1 by adding X1 and removing X2. Adding and removing units is analogous to strengthening or weakening inputs using gain modulation. However, in our model, as shown in Figure 3.7B, a unit can be added with feedback of a different phase causing it to play a different role in the operation that it is added to. We quantitatively explored how these interaction effects depend on different network connections probabilities in large networks.

Quantifying top-down effects

We quantified the effect of top-down influences by considering the number of inputs that feedback adds or removes from operations of each of the two possible phases. The feedback may be either from other outputs that are initiated or it may be external, orchestrating feedback. No inputs will be added or removed if, and only if, no other operations interact with the operations and orchestrating feedback does not modify the operation. We considered a two-layer network with N_I input units. We defined the following two-component vectors:

- $\mathbf{N_{R0}}$ whose components are the number of input units of each phase that are involved in the computation performed by an output unit (the number of **relevant inputs** of each phase in an operation) when it is initiated without any other outputs initiated or any external feedback.
- $\mathbf{N_{R+}}$ whose components are the number of input units of each phase that are **added** to the set of relevant inputs in an operation when another output unit is initiated or orchestrating feedback from an external unit is present.
- $\mathbf{N_{R-}}$ whose components are the number of input units of each phase that are **removed** from the set of relevant inputs in an operation when another output unit is initiated or orchestrating feedback from an external unit is present.

In order to understand how network properties affect these metrics, we considered a two-layer network with only feedforward and feedback connections (no lateral connections). The connection probabilities for feedforward and feedback connections is $p_{\text{ff}} = p_{\text{ff-only}} + p_{\text{ff+fb}}$ and $p_{\text{fb}} = p_{\text{fb-only}} + p_{\text{ff+fb}}$, respectively, where $p_{\text{ff-only}}$, $p_{\text{fb-only}}$, and $p_{\text{ff+fb}}$ are the probabilities that pairs of units in each layer are connected with only feedforward, only feedback, and both feedforward and feedback connections, respectively. The probability of a feedforward or feedback connection having a long (short) time lag is given by $p_{\bar{F}}$ ($p_{\hat{F}} = 1 - p_{\bar{F}}$) and $p_{\bar{B}}$ ($p_{\hat{B}} = 1 - p_{\bar{B}}$), respectively. The connection probabilities for the feedback from other output units or from external units is given by p^*_{fb} and the probability of them being long (short) is $p^*_{\bar{B}}$ ($p^*_{\hat{B}} = 1 - p^*_{\bar{B}}$). For this situation, we determined the expressions,

$$\begin{aligned}
\frac{\mathbf{N_{R0}}}{N_I} &= p_{\text{ff+fb}}\Big[\alpha, \quad (1-\alpha)\Big], \\
\frac{\mathbf{N_{R+}}}{N_I} &= p^*_{\text{fb}} p_{\text{ff-only}}\Big[\alpha^*, \quad (1-\alpha^*)\Big], \\
\frac{\mathbf{N_{R-}}}{N_I} &= p^*_{\text{fb}} p_{\text{ff+fb}}\Big[\alpha(1-\alpha^*), \quad (1-\alpha)\alpha^*\Big],
\end{aligned} \tag{3.4.4}$$

where

$$\begin{aligned}
\alpha &= 1 - p_{\bar{B}}, \\
\alpha^* &= \beta(1 - p^*_{\bar{B}}) + (1-\beta)p^*_{\bar{B}},
\end{aligned} \tag{3.4.5}$$

and β is the probability that the unit providing the additional feedback is in phase with the output unit being considered. The values α and α^* are the probabilities that feedback

from the output unit or the other source of feedback, respectively, arrives with the same phase as the main output unit. It is only the time lag of the feedback connection that affects α; however, α^* also depends on the likelihood of the other output source being in or out of phase with the main output unit (i.e., β).

When the output unit of interest is initiated alone ($\mathbf{N_{R0}}$), only reciprocally connected units where the feedback is in phase with the inputs will be involved in the operation. Only input units which are not reciprocally connected but make a feedforward connection to the output unit can be added ($\mathbf{N_{R+}}$) and they are added by receiving feedback that is in phase. This is shown in Figure 3.8A, where we plot the total number of units (of either phase) originally in the operation and the total number added and removed as functions of the ratio $p_{\text{ff-only}}/p_{\text{ff}}$. As expected, when there are only reciprocal connections (i.e., $p_{\text{ff-only}}/p_{\text{ff}} = 0$), no units can be added; when there are no reciprocal connections (i.e., $p_{\text{ff-only}}/p_{\text{ff}} = 1$), no units are originally in the operation (and so none can be removed either).

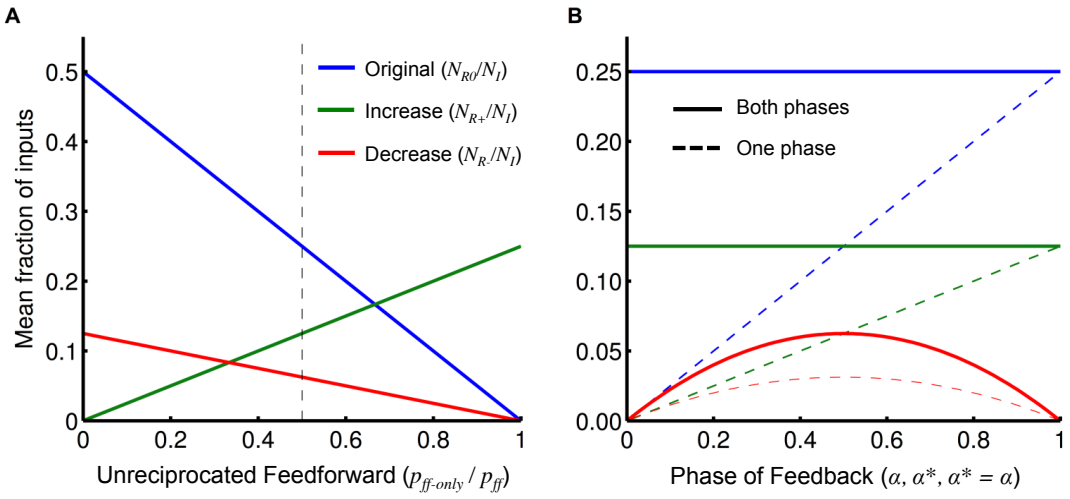


Figure 3.8: Interaction Effects with Network Parameters. A: The mean fraction of relevant inputs (either phase) for an operation initiated alone, $\mathbf{N_{R0}}/N_I$ (blue), and the mean increase and decrease in the fraction of relevant inputs (either phase) when feedback from a second operation or external unit is also present, $\mathbf{N_{R+}}/N_I$ (green) and $\mathbf{N_{R-}}/N_I$ (red), respectively, plotted as functions of the fraction $p_{\text{ff-only}}/p_{\text{ff}}$ (the fraction of unreciprocated feedforward connections) as given by Equation (3.4.4). The values of other network parameters used were: $p_{\text{ff}} = 0.5$, $p_{\text{fb}} = p^*_{\text{fb}} = 0.5$, and $\alpha = \alpha^* = 0.5$. The dashed vertical line shows the fraction of $p_{\text{ff-only}}/p_{\text{ff}}$ used in B. **B:** Same as A but varying the probability of the phase of the different types of feedback: α (phase probability of initiating feedback), α^* (phase probability of orchestrating feedback), and $\alpha = \alpha^*$ (phase probability of any external feedback), for $\mathbf{N_{R0}}/N_I$, $\mathbf{N_{R+}}/N_I$, and $\mathbf{N_{R-}}/N_I$, respectively. Also shown is the fraction of relevant inputs of a particular phase (dashed) that, compared to the fraction of relevant inputs of either phase (solid), illustrates the split between the two phases.

For a unit to be removed ($\mathbf{N_{R-}}$), it must originally be in the operation and then receive new feedback that is out-of-phase. This is shown in Figure 3.8B, where we plot

the original number of units in the operation and the number of units that are added and removed for different feedback phase probabilities. For the original number of inputs and the number of inputs added, the total number of inputs is fixed but the split between the phases changes linearly with the value of α and α^*, respectively. The total number of input units removed is zero when both $\alpha = \alpha^* = 0$ or 1. This is because both types of feedback (from the output unit or the other unit) always arrives with the same phase and so input cannot receive incoherent feedback. Similarly, Equation (3.4.4) shows that if $\alpha = 0$ and $\alpha^* = 1$, or $\alpha = 1$ and $\alpha^* = 0$, then feedback from the two sources will always be out of phase. In this case, input units of only one phase would be involved in the operation originally and those that receive additional feedback will always be removed.

From Figure 3.8, we see that there are separate network properties controlling the number of units that are added ($p_{\text{ff-only}}/p_{\text{ff}}$) and the network property controlling the number of units that are removed (α and α^*). However, in a random network, it would not be possible to have a mix of phases in the original input units or in the input units that were added and also avoid having units removed from operations.

Interactions and shared inputs

In the random networks we consider, there will be some overlap between the inputs that comprise the operations of different output units but this will not depend on network properties except the likelihood of reciprocal connections and the number of input units. Whether shared or non-shared inputs are added or removed from the operations due to interactions depends on the types of connections involved. Shared inputs are added by interactions due to feedback from reciprocal connections, whereas feedback without a reciprocal feedforward connection adds non-shared inputs. However, input units removed by interactions will always be non-shared as the second output unit must make a interfering feedback connection to the unit.

Interactions in orchestrated networks

Orchestrated networks provide much flexibility for networks to be modified to perform arbitrary operations and this control through high-level, external feedback is a commonly envisioned architecture for gain modulated networks. There are two possible extremes for these types of networks. The first extreme is networks in which many non-interacting operations are performed in parallel. The second are networks in which a larger number of output unit combinations interact to perform a single but potentially more complicated operation. In these two cases, the orchestrating feedback controls and modifies the operations or single operation, respectively.

Considering the first type of network, we investigated the constraints on orchestrating the network if there are to be no interactions or if interactions are to be restricted in some way. An interesting result, shown by Figure 3.8B, is that $p_{\text{ff-only}} > 0$ is required in order to allow external feedback to add additional units to operations but, as long as

$p_{\text{fb}} > 0$, this means that similar interactions will occur between the operations (see Equation (3.4.4)). This suggests the first of the following three possible network conditions that can control the interactions:

- **No internal feedback:** In this case ($p_{\text{fb}} = 0$), operations would require external feedback in order to exist (feedforward activity would not be able to propagate without external feedback) but no interactions would be possible due to the fact that output units could not influence lower-level units at all.
- **Homogenous delays for internal feedback:** In this case, all of the internal feedback would have 'short' or 'long' delays but not a mix of the two ($\alpha = 0$ or 1). This means that interactions could cause units to be added to the operations of other units but not removed. Provided that units performed operations that contained at least one unit when they are initiated separately, this network condition would also ensure that if there were no interactions without orchestrating feedback then the addition of arbitrary orchestrating feedback would not change this.
- **Homogenous delays for internal feedback and no non-reciprocal feedforward connections:** In this case ($\alpha = 0$ or 1, and $p_{\text{ff-only}} = 0$), no interactions between output units would be possible and orchestrating feedback would only be able to remove units from operations.

3.5 Discussion

3.5.1 Relation to Cognitive Phenomena

The networks of cortical units that we have proposed and investigated provides a high-level model of various cognitive phenomena, including goal-directed behavior and top-down attention. We described a general architecture for goal-directed behavior in Figure 3.2C and demonstrated simple examples in Figure 3.6. We considered it out of the scope of this chapter to explore how these goals are generated, maintained, or changed; however, we considered how feedback from goals could quickly switch and modify stimulus-response mappings. This is crucial in behavioral settings, where goals or information held in working memory need to influence the way that stimuli are responded to. In this chapter, we have identified the different ways that this influence can be implemented. We proposed a model with oscillatory, gain-modulated units that allows feedback to more flexibly manipulate stimulus-response mappings than models with only gain modulation.

Top-down attention naturally arises in this situation because units and subnetworks of units are only activated if they receive feedback corresponding to attention. Stimuli that are not relevant to a particular task will be ignored and activity they elicit will not propagate to higher brain regions. Therefore, bottom-up attention must work via a different means to those described in this chapter, so that salient stimuli can interrupt top-down tasks and perhaps alter these tasks or goals.

3.5.2 Relation to Experimental Findings

Context-dependent changes to neural responses

Womelsdorf et al. (2006, 2008) observed context-dependent changes to the receptive field of neurons in the middle temporal area (MT). The changes included shifts of the centers of the receptive fields toward the focus of attention and narrowings of the receptive fields. Cohen and Newsome (2008) similarly observed that noise correlations of MT neurons depended on the current behavioral task being performed. In both of these experiments, the stimuli were not being changed and, according to our model, these context dependent changes are due to changes in the top-down feedback (either initiating or orchestrating) to these neurons.

In low-level areas of the auditory cortex, Zion Golumbic et al. (2013) observed that attention boosted the activity corresponding to "attended speech", but that "ignored speech" remained represented. However, in higher-order regions, attention becomes more "selective" and activity representing ignored speech was not present. Similarly, in the visual system, Hupé et al. (1998) showed that feedback connections serve to amplify and focus activity of neurons in lower-order areas and that they were important in discriminating between a figure and the background. Schroeder et al. (2010) refer to this interaction between sensory and attentional, top-down signals as "active sensing". This is consistent with the model we are proposing where attention, determined by the goals of the system, "selects" the relevant sensory stimuli, while ignoring irrelevant stimuli.

Abstract rules and operations

Wallis and Miller (2003), Muhammad et al. (2006), and Buschman et al. (2012) considered abstract rules that could be applied in a very similar manner to many different stimuli-response mappings. The ability of the brain to create such abstract mappings suggests that it reuses the same circuitries or networks for multiple analogous purposes. This is consistent with the way networks in our model can be composed together and embedded into larger networks. In this case, it is the role of orchestrating feedback to make sure the reused network receives the appropriate inputs and that its outputs are used correctly. Badre (2008) reviewed the evidence for hierarchies within goals and rules used for cognitive control in the PFC where there were increasing levels of abstraction for higher-level goals. This hierarchy of goals suggests the existence of different levels of goal-dependent feedback, each orchestrating different parts of the stimulus-response mapping required for the overarching goal.

Buschman et al. (2012) showed that during a stimulus-response task there was a dominant rule (based on the orientation of a visual stimulus), which alpha oscillations appeared to suppress in order for a different rule (based on the color of a visual stimulus) to be employed. In our model, this type of behavior may be exhibited by having orchestrating feedback that would modify the original, dominant operation or mapping performed by

the network to a secondary mapping. However, our model does not suggest an explanation as to why alpha rhythms would be involved in this top-down remapping.

Fast signal propagation and neural coding

In our model, activity takes at most half the oscillation period (about 7-10ms for gamma oscillations) to propagate from one unit to the next. The target unit does not need to integrate its inputs but can very quickly pass along the "signal" provided that it receives coherent feedforward and feedback inputs. In other words, units in the model are assumed to exist in a fluctuation-driven regime, where, unlike models in which units need to integrate their inputs, activity can be more rapidly altered. This is consistent with the range of reaction times (about 300-400ms) observed by Wallis and Miller (2003) in their rule-based behavioral experiments. In our model, both the phases and the levels of activation (absolute firing rate) are important for performing computations. Our model does not predict that absolute spike rates are irrelevant but it does make the assumption that they are only relevant in concert with the appropriate phases.

Pyramidal neurons

Larkum et al. (1999, 2001, 2004) and Larkum (2013) observed that pyramidal neurons exhibited a much stronger response when they received inputs both to their soma (and basal dendrites) and to their apical dendrites than they did when they received only one of these types of inputs. In addition to the spike initiation zone at the cell body for action potentials (sodium spikes), there is a second initiation zone near the apical tuft of layer 5 pyramidal neurons (Yuste et al., 1994; Schiller et al., 1997; Larkum and Zhu, 2002). This second initiation zone produces broad calcium spikes within the cell and its existence suggests that pyramidal neurons should be considered to have two functional compartments. Larkum et al. (1999, 2001, 2004) and Larkum (2013) discuss how interactions between these two initiation zones, where spikes from either one lower the firing threshold of the other, provide the associative mechanism whereby a stronger response occurs when both somatic and apical inputs are present.

We proposed our analogous model for interconnected groups of pyramidal neurons based on this experimentally-based description of how pyramidal neurons respond to different types of inputs. In our model, groups of neurons behave similarly to individual pyramidal neurons in that they produce a much stronger response when receiving somatic feedforward activity as well as apical feedback. However, our model differs in that the groups of neurons also contain inhibitory interneurons and because of this they exhibit oscillatory activity in the gamma frequency range.

3.5.3 Experimental Predictions

Requirements of neural activation

In our model, the activation of groups/ensembles of neurons requires strong coherent feedforward and feedback activity. We are predicting that, at least during goal-directed tasks, neuronal ensembles in high-level areas of the cortex are only activated if they receive feedback from higher regions, or if there has been recent feedback (regions involved in working memory may sustain activity without feedback). Similarly, without receiving activity from lower-level units, our model predicts that high-level units would at most be able to exhibit sporadic, searching activity and not strong oscillatory (e.g. gamma frequency) activity.

This prediction does not necessarily extend to lower-level areas of the cortex, such as V1, in which sensory input alone may be sufficient to activate groups of neurons (Hupé et al., 1998). There may also be top-down feedback present during non-goal-driven behavior, or during resting states, that provides a "default" set of operations for the network. Similar to this, the presence of a neuromodulator may remove (or introduce) the need for top-down feedback, allowing feedforward activity alone to activate units and propagate into higher-level regions. For example, there is evidence that cholinergic neurons increase the amount that attention modulates the activity of cortical neurons (Herrero et al., 2008; Thiele, 2009; Goard and Dan, 2009; Herrero et al., 2013). In this situation, acetylcholine may actually decrease the excitability of the neurons, pushing them into a more goal-driven mode, where they are forced to rely on both feedforward and feedback activity to become active.

In addition to this, we hypothesize that there must exist coherence between neurons within an active ensemble and between neurons in different active ensembles that are strongly connected. This prediction is most relevant to the activation of neuronal ensembles during attentional and behavioral tasks. This type of experimental result has been observed for alpha and beta frequencies by Buschman et al. (2012), where there was coherence between neurons in the PFC during behavioral tasks that involved switching between different abstract rules.

Different cortical regions

The results from Equation (3.4.4) and Figure 3.8 for large networks suggest a trade-off between the ability to perform many operations in parallel and the ability to control these operations in a top-down manner with feedback. Given that different regions of the cortex would have different priorities in this regard, this makes experimental predictions for the connectivity within and between different regions of the cortex. For instance, within regions where units would correspond to percepts, interactions between hypotheses and goal-directed manipulation of hypotheses would be expected to be low as our perceptions are relatively stable with respect to our goals. In this case, we would expect that $\alpha^* \approx$

1, $p_{\text{ff-only}} \approx 0$, and $p^*_{\text{fb}} \approx 0$. In other words, feedback connections from units in one hierarchical level to another would be expected to have quite similar propagation delays, feedforward and feedback connections between units would mostly occur together (i.e., reciprocally), and there would be few feedback connections coming directly from regions involved in motivations/goals/behavior. For higher-level regions where units correspond to actions and more abstract concepts, goal-directed orchestration of operations, and perhaps also interactions between different operations, would be desired. For the regions involved in motivation and goals to orchestrate the operations and control the network, there needs to be feedback connections from these regions to the units (i.e., $p^*_{\text{fb}} > 0$).

Wallis and Miller (2003) and Muhammad et al. (2006) recorded from neurons in the prefrontal (PRC), premotor (PMC), and inferior temporal (ITC) cortices and the striatum (STR) during a stimulus-matching task. During the task, two visual stimuli were presented and, depending on the rule (which was indicated via a visual cue presented with the first stimulus), the subject was required to either continue holding a lever or release the lever. They observed different neurons that responded selectively to the rule (desired stimulus-response mapping), the behavioral response carried out, the visual stimulus being remembered, or whether the subsequent stimulus matched this remembered stimulus. We constructed a possible network to carry out this task (Figure 3.9A and B). In addition to external feedback that depends on the rule to be employed, this network receives external feedback based on the stimulus being remembered (held in working memory). Based on the selectivity that was observed of neurons in different cortical regions, we divided this network into these different regions (Figure 3.9C). While this is not necessarily the exact network used for this task or the correct allocation of units to cortical regions, this demonstrates how our model may be useful in understanding the role of neurons in different regions.

Searching feedback and neuronal avalanches

The sporadic, bursting feedback activity that we proposed to be exhibited by units during the searching state is based on the observations by Larkum (2013) of the activity of pyramidal neurons that receive only strong input to their apical dendrites. We propose that this mechanism exists for ensembles of pyramidal neurons and that it is used to pass internal predictions/expectations from higher-level ensembles down to lower-level ensembles. This relies on sharp, sporadic, bursts of activity being able to propagate along feedback connections down to lower-levels but not along feedforward connections up to higher-levels. Neuronal avalanches observed *in vitro* (Beggs and Plenz, 2003) may correspond to spontaneous examples of these searching signals within networks that are not receiving any sensory inputs. Of interest would be the calcium activity, due to the activation of the calcium spike initiation zone on the apical branch, of pyramidal neurons during neuronal avalanches.

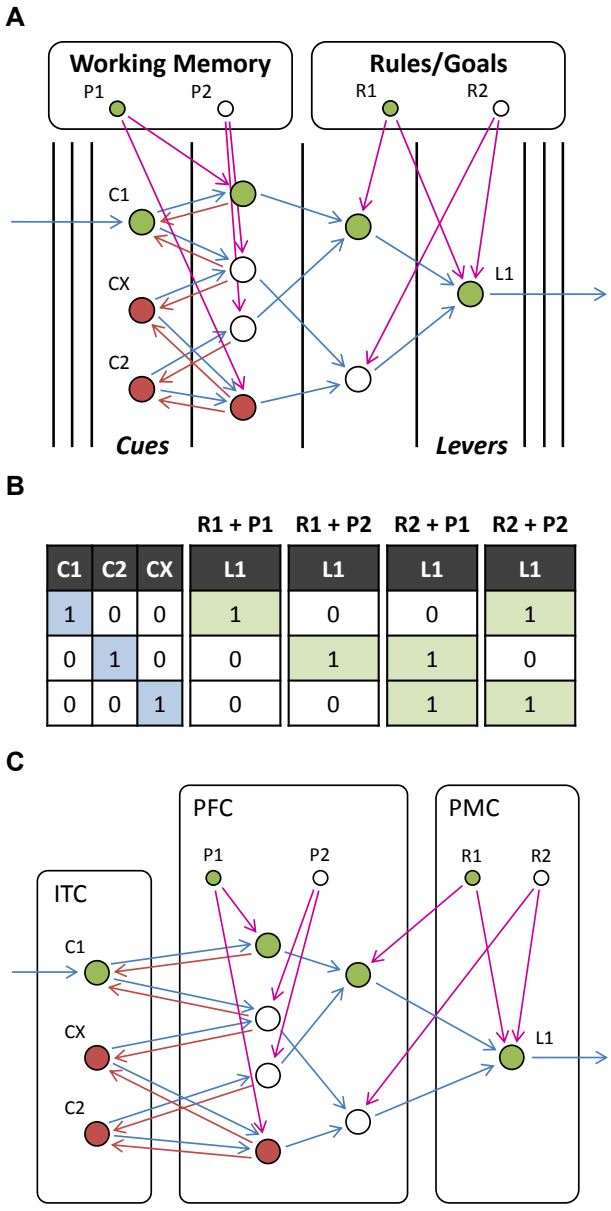


Figure 3.9: Stimulus-matching Experiment. A: Network for responding to a stimulus (C1, C2 or any other stimulus CX), which is dependent on the current rule that determines whether to respond to a match or non-match with the previous stimulus. In contrast to previous tasks, working memory is required to remember the previous stimulus (P1 and P2 which correspond to the same stimulus as C1 and C2, respectively) in the same way that the rules (R1 and R2) are remembered. **B:** Input-output table for the network presented with a stimulus for combinations of rules and remembered stimuli. **C:** Same network as in A but with units assigned to anatomical regions of the cortex.

Apical-targeted feedforward connections

We have assumed that higher level areas connect to lower level areas via feedback (apical-targeted) connections and that, in return, the lower level areas connect to the higher level areas via feedforward (soma/basal-targeted) connections. While this seems to be the prevalent pattern in the cortex, exceptions are likely to exist. Apical-targeted feedforward connections, for example, would allow salient, lower-level features/hypotheses to prompt/pose higher level hypotheses. The opposite (soma/basal-targeted feedback connections) may also exist but it is not clear the functional role that these connections could play.

3.5.4 Neuronal Ensembles and Lateral Connections

Units in our model are not necessarily local or spatially distinct groups of neurons, but are instead defined by their functional connectivities. In other words, units are better thought of as strongly interconnected ensembles of neurons. We have assumed, in this chapter, that units consist of distinct, non-overlapping groups of neurons. It may be, however, that there is a large overlap between the neurons that make up each unit.

We considered feedforward and feedback connections between units but not lateral connections. Assuming that units in the same layer consist of overlapping groups of neurons, it may be better to think of lateral connections between units as interactions between these overlapping ensembles, in which the ensembles either inhibit each other (both cannot be active: winner-take-all), reinforce each other (performing an "OR" operation), or require co-activation (performing an "AND" operation). In the situation where units laterally inhibit each other, the hypotheses they represent are incompatible. This is similar to multistable perceptual phenomena, such as binocular and monocular rivalry, where there is competition between two incompatible perceptions. Leopold and Logothetis (1999) showed that, during binocular rivalry experiments, a greater number of neurons in higher-level areas are correlated with the perception than in lower-level areas, suggestive of top-down processing.

3.5.5 Relation to Gain Modulation

Previous studies consider gain modulation, caused by feedback signals, as a means of performing top-down processing and cognitive control. In the study by Salinas (2004), fixed feedback dependent on the current rule was used to modulate the gain of neurons with feedforward connections on output neurons. This is functionally similar to the last two layers of the network we considered in Figure 3.6E, where there are only feedforward connections between the layers and external feedback orchestrates/modulates the inputs to be considered. In contrast to such studies, our model exhibits an exaggerated and simplified example of non-linear gain modulation, where feedback modulates feedforward signals above a threshold, which cannot be otherwise achieved and which permits further

propagation of the signal (see Figure 3.1C). It remains to be explored how our findings, regarding interacting, orchestrated, and phase-dependent operations, extend to the case where activities and connection strengths are continuous. Instead, our model has only binary inputs, states (besides the searching state), and connections. Because of this, our simplified model is not well suited to integrating many individual inputs and determining whether all are present or a threshold has been reached. A more detailed model with continuous activation levels and connection strengths would be more suited to these types of computation. However, we suggest that this type of processing is potentially more prevalent in lower-level regions of the cortex and that, in higher-level, associative areas, it may play a smaller role.

With only gain modulation, feedback can either increase or decrease the gain of a neuron and, therefore, how strongly it is involved in an operation, but it cannot change the role that it plays in the various operations in which it is involved. Here, however, the top-down feedback can have different phases that can initiate units with different phases giving them different roles in operations.

3.5.6 Future Extensions

Synaptic plasticity

When investigating how network properties affected top-down processing, we only considered randomly connected networks, whereas when we considered a specific task, the networks we used had a very specific structure. This specific structure would need to emerge due to some form of activity-dependent synaptic plasticity. Spike-timing-dependent plasticity, for instance, would be expected to reinforce connections between active, coherent units, as in Chapter 2 (Kerr et al., 2013). It may also be that, in the case where a set of connections only ever inhibits the activity of the target unit, synaptic plasticity only maintains the connections onto the inhibitory neurons that cause this inhibition. This would mean that this set of connections becomes only able to inhibit and not activate the target unit. In addition to this, it remains to be investigated how robust certain networks or motifs are to the introduction (removal) of units to (from) operations through synaptic plasticity. We speculate that networks with fewer interactions may be more robust in this regard but this remains to be explored.

Analysis of networks with three or more levels

Our exploration of network properties only considered the case of a two-layer network of units. We would expect that interactions between different operations would be amplified in a network with more layers because more units and connections, which can cause the interactions, are involved. However, connections do not need to be restricted to being between adjacent layers. For instance, this is not the case in Figure 3.6 and Figure 3.7B. An analysis of how different network properties affect top-down processing and interactions

becomes more complicated in this case.

Other phases and frequencies

We have assumed that active units in our model oscillate at one of only two different phases. This simplified model could be extended to include a continuous range of phases, better capturing the complexity of networks in the brain. More possible phases would increase the number of different operations in which units/ensembles could be used. In addition, neurons within an ensemble may actually exhibit a range of phases rather than just a single phase. This would further complicate the ways that units affect each other through lateral "connections", where neurons are shared between units, or actual feedforward and feedback connections, where the phase of the individual source and target neurons differs within the units.

Our model would similarly apply to other inhibitory-based rhythms (Whittington et al., 2000), such as beta frequency oscillations. In fact, there is experimental evidence to suggest that beta frequency oscillations, either alone or interacting with gamma oscillations, may represent a better candidate for top-down modulations (Engel and Fries, 2010; Benchenane et al., 2011). For example, Buschman et al. (2012) showed that neurons in the PFC synchronized to beta frequencies during a rule-based behavioral task.

A number of other roles for gamma oscillations have been proposed based on experimental observations. Schroeder and Lakatos (2009) argued that the amplitude of gamma oscillations is often coupled to the phase of, or "enslaved" to, lower frequency oscillations (e.g. delta or theta) and propose that non-enslaved gamma oscillations are only exhibited during a "vigilance" mode when there is no task relevant rhythm. Arnal et al. (2011) proposed that gamma oscillations represent bottom-up prediction errors, indicating when sensory signals misalign with top-down predictions represented by beta oscillations.

This study focuses on a single spectral band associated with sensory processing and motor pattern generation. However, multiple frequencies are likely present at the same time and future work exploring this situation would be very interesting. For example, there is evidence to suggest that feedback activity would likely be at lower frequencies (e.g., beta) while feedforward activity would be at higher frequencies (e.g., gamma) (Arnal et al., 2011; Bastos et al., 2012). The role of these different frequency oscillations, and how they may interact in situations such as this, while out of the scope of the current study, promises a rich area for exploration. Unless the lower frequencies are subharmonics of the higher frequencies, how different frequency oscillations would interact poses a problem that needs to be investigated. Alternatively, computations with different frequencies could potentially operate in parallel to each other.

Detailed models of neural activation

In this study, we have considered only three discrete activation levels for units of neurons (resting, searching, and active). An area for future work is to consider more detailed mod-

els of neural activation in which units have continuous levels of activation. In this case, feedback could be modeled as smoothly (although most likely nonlinearly) modulating the sensitivity, or gain, of units to feedforward input, rather than simply gating the activity of units. The conditions under which top-down feedback would play a major role in activating and modifying neural ensembles and the computations that they preform remains to be explored. Another interesting area of investigation would be modeling the effects of neuromodulators, such as acetylcholine, on neural activation and exploring how such neuromodulators could switch a network between bottom-up, forward-driven and top-down, feedback-driven modes of operation.

Phase-locking of top-down oscillations

Feedforward and feedback phases propagate forwards and backwards, respectively, according to connections that have been established through plasticity mechanisms, which remain to be explored in this context. However, there would ultimately be the outer-most feedforward phases and the inner-most feedback phases and, in order for them to synchronize and activate along the appropriate network pathways, some sort of matching between these phases would be necessary. In order to accomplish this, high-level areas of the brain would need to perform "phase-locking" between their top-down signals and the bottom-up signals in lower-level areas. This type of synchronization must be ubiquitous in the brain and at high frequencies, such as gamma and beta, it should be possible to perform this "phase-locking" quickly. Together with an investigation of a model in which there are continuous phase ranges, future work lies in investigating how this type of "phase-locking" could be carried out between top-down and bottom-up oscillations.

More specific connections

While we separated inputs into two types (apical and basal), each of these could be further split up into individual dendritic branches that locally integrate their own inputs (Larkum et al., 2009) and can be targeted specifically by certain inhibitory inputs (Palmer et al., 2012). Targeted inhibition to specific branches and, therefore, inputs would allow neurons/units to perform much more complicated computations and, in particular, would be useful in allowing neurons to be re-used for different operations. In addition to this, excitatory connections to a unit could specifically target either the excitatory or inhibitory neurons in the unit.

Chapter 4

Coexistence of Reward and Unsupervised Learning

This chapter is a slightly modified version of the published article:

> *Kerr RR, Grayden DB, Thomas DA, Gilson M, and Burkitt AN. (2014) "Coexistence of Reward and Unsupervised Learning during the Operant Conditioning of Neural Firing Rates." PLoS ONE 9(1): e87123.*

4.1 Abstract

A fundamental goal of neuroscience is to understand how cognitive processes, such as operant conditioning, are performed by the brain. Typical and well studied examples of operant conditioning, in which the firing rates of individual cortical neurons in monkeys are increased using rewards, provide an opportunity for insight into this. Studies of reward-modulated spike-timing-dependent plasticity (RSTDP), and of other models such as R-max, have reproduced this learning behavior, but they have assumed that no unsupervised learning is present (i.e., no learning occurs without, or independent of, rewards). We show that these models cannot elicit firing rate reinforcement while exhibiting both reward learning and ongoing, stable unsupervised learning. To fix this issue, we propose a new RSTDP model of synaptic plasticity based upon the observed effects that dopamine has on long-term potentiation and depression (LTP and LTD). We show, both analytically and through simulations, that our new model can exhibit unsupervised learning and lead to firing rate reinforcement. This requires that the strengthening of LTP by the reward signal is greater than the strengthening of LTD and that the reinforced neuron exhibits irregular firing. We show the robustness of our findings to spike-timing correlations, to the synaptic weight dependence that is assumed, and to changes in the mean reward. We also consider our model in the differential reinforcement of two nearby neurons. Our model

aligns more strongly with experimental studies than previous models and makes testable predictions for future experiments.

4.2 Introduction

Operant conditioning refers to an individual modifying its behavior based on some consequence of that behavior. Understanding how this process arises from neural mechanisms in the brain will provide a promising step toward linking neural mechanisms with behavior and learning and discovering how the brain gives rise to cognitive functions in general. It is also applicable to brain-computer interfaces, where operant conditioning can be used to develop control of external prostheses rather than tailoring them to existing neuronal circuitry (Moritz et al., 2008).

Operant conditioning experiments have shown that the firing rate of individual neurons in the precentral motor cortex and prefrontal cortex of monkeys could be significantly increased by giving positive reinforcement, provided that the monkeys were also given immediate feedback on the neuron's firing (Fetz, 1969; Fetz and Baker, 1973; Kobayashi et al., 2010). A visual display presented the monkeys with a time-decaying signal that was incremented for each action potential that an implanted electrode measured. Upon reaching a threshold value, the signal was reset and the monkey received a food reward. Negative punishment (i.e., the removal of reward in order to decrease a particular behavior) was performed with a similar setup, where measured spikes decremented the signal (and artificially generated spikes incremented the signal) (Fetz and Baker, 1973). In this case, low firing rates were elicited. Through a combination of positive reinforcement and negative punishment, they also showed that a differential between the firing rates of two neurons could be elicited.

Current theories hold that learning at the behavioral level is ultimately due to changes at the synaptic level. Reinforcement learning models of synaptic plasticity depend on neuronal activity and also on a reward signal (Pawlak et al., 2010) that, due to the evidence linking dopamine to reward learning in the brain (Schultz et al., 1997), typically represents the amount of extracellular dopamine present. Similar to Frémaux et al. (2010), we identify two main types of existing models. First, there are models that have been derived theoretically to maximize the received reward (Seung, 2003; Xie and Seung, 2004; Pfister et al., 2006; Florian, 2007), such as the R-max model (Frémaux et al., 2010). Secondly, there is reward-modulated spike-timing-dependent plasticity (STDP) (Izhikevich, 2007; Farries and Fairhall, 2007; Florian, 2007), or RSTDP, where the amplitudes of synaptic changes that would have been made by STDP (Markram et al., 1997; Bi and Poo, 1998) are modulated by subsequent rewards.

A reinforcement learning model of synaptic plasticity exhibits unsupervised learning (i.e., learning that occurs independently of any rewards) if there is long-term potentiation (LTP) or long-term depression (LTD) at the mean reward level. Additionally, for models where LTP and LTD do not depend on the current synaptic weight (additive models),

unsupervised learning is only present if the LTP and LTD do not cancel with each other. Studies with existing models find that there should be no unsupervised learning in order to perform reinforcement learning tasks, such as the operant conditioning of neuronal firing rates (Legenstein et al., 2008; Frémaux et al., 2010). However, even after development, the brain receives large amounts of novel sensory information without any associated rewards or punishments (Barlow, 1989). Any learning based on this information is necessarily unsupervised, suggesting an ongoing role for unsupervised learning after development. This likely depends on the brain region. In synapses onto GABAergic spiny neurons in the rat striatum, Pawlak and Kerr (2008) showed that no LTP or LTD occurred when D1-receptors (dopamine receptors) were blocked. In synapses onto pyramidal neurons in the rat hippocampus, however, Zhang et al. (2009) observed classical STDP learning windows without any dopamine present. When extracellular dopamine was added, Zhang et al. (2009) observed increased LTP for pre-post spike pairs and that LTD had switched to LTP for post-pre spike pairs. Based on this, it seems unlikely that there would be no LTP or LTD at the base level of dopamine, which suggests that unsupervised learning can coexist with reward learning.

Here, we consider the case where unsupervised learning does occur (unlike in the situation considered in previous studies (Legenstein et al., 2008; Frémaux et al., 2010)) and so, even without reinforcement learning, a balance of LTP and LTD produces stable firing rates. Under this assumption, we demonstrate that existing RSTDP models are unable to elicit increased firing rates in neurons that are rewarded for firing. We propose a new RSTDP model that can elicit reinforcement learning, in which LTP and LTD are modulated separately by the reward signal. This is more consistent with the experimental observations that dopamine affects LTP and LTD differently, even causing LTD to switch to LTP for high concentrations (Zhang et al., 2009). We show that these findings are robust to the introduction of spike-timing correlations, the synaptic weight dependence that is assumed, and the reward signal used. We demonstrate that our model is also able to reproduce the differential reinforcement of two neurons observed by Fetz and Baker (1973). Finally, we compare the learning induced by the operant conditioning of firing rates using our model with the R-max model to highlight the impact of including unsupervised learning with reward learning.

4.3 Results

4.3.1 RSTDP Model

To better incorporate the effects that neuromodulators have been observed to have on synaptic plasticity (Figure 4.1A), we propose a new RSTDP model in which LTP and LTD can be modulated differently by a neuromodulator (e.g., dopamine). In this model, there are a pair of modulated parameters for each of LTP and LTD. Each pair describes the linear effect that a neuromodulator has on the amplitude of LTP and LTD. The modulation offsets, q_+ and q_-, give the amplitudes of LTP and LTD, respectively, when

the reward signal is zero. The modulation slopes, p_+ and p_-, give the rates of change of the amplitudes of LTP and LTD, respectively. By setting both modulation offsets to zero (i.e., $q_\pm = 0$), the classical RSTDP model is recovered (dashed blue line in Figure 4.1A). In this chapter, we focus on a particular set of modulation parameters (solid blue line in Figure 4.1A) that leads to the effect that Zhang et al. (2009) observed dopamine to have on STDP (blue circles in Figure 4.1A). We refer to this parameterization as dopamine RSTDP. Figure 4.1B illustrates the effective learning windows corresponding to changes in the reward signal, as compared to classical RSTDP shown in Figure 4.1C.

Our new RSTDP model introduces two qualitatively new features. The first is that there can be LTD and LTP when the reward is zero (provided that $q_\pm \neq 0$). This differs from previous studies in which firing rate reinforcement was demonstrated (Legenstein et al., 2008; Frémaux et al., 2010), where the base reward signal was zero and, at this level, there was no LTP or LTD. This difference is illustrated in Figure 4.1A and Figure 4.2A. However, we consider the case where the base reward level is positive and so, for both our RSTDP model and classical RSTDP, there is LTD and LTP present at the base reward level and, therefore, there is unsupervised learning. The second new feature, introduced by our new RSTDP model, is that LTD and LTP are modulated separately by the reward signal. This means that it is possible for a balance of LTP and LTD to be disrupted by an increase (or decrease) in reward. It also means it is possible for the LTP (LTD) caused by pre-post (post-pre) spike pairs to be differentially switched to LTD (LTP) for high reward signal values. The latter of these, where LTD transitions to LTP, is demonstrated with dopamine RSTDP (Figure 4.1B) and matches observed effects of dopamine of STDP (Zhang et al., 2009). In classical RSTDP, the only point at which both LTP and LTD switch is when the rewards become negative (or below baseline in previous studies (Legenstein et al., 2008; Frémaux et al., 2010)).

The model is able to exhibit differential modulation of LTP and LTD because it stores the effects of the pre-post and post-pre spike pairs in two separate eligibility traces, $e_{ik}^{\pm}(t)$. This is in contrast to classical RSTDP, which combines these effects into a single eligibility trace. Figure 4.1D shows the two eligibility traces for an individual synapse, as well as the reward signal, $y(t)$ (determined by the post-synaptic spike train, $S_i(t)$), and the changes elicited in the synaptic weight, $K_{ik}(t)$.

4.3.2 Analytical Predictions

To apply this model to operant conditioning experiments, we considered the feed-forward network shown in Figure 4.1E, containing three different types of post-synaptic neurons:

- **Reinforced:** The firing of the reinforced neuron is recorded and determines the amount of reward delivered. In operant conditioning experiments, the firing rate of this neuron was observed to increase.
- **Surround:** The surround neuron is located near the reinforced neuron but its firing does not affect the reward delivered.

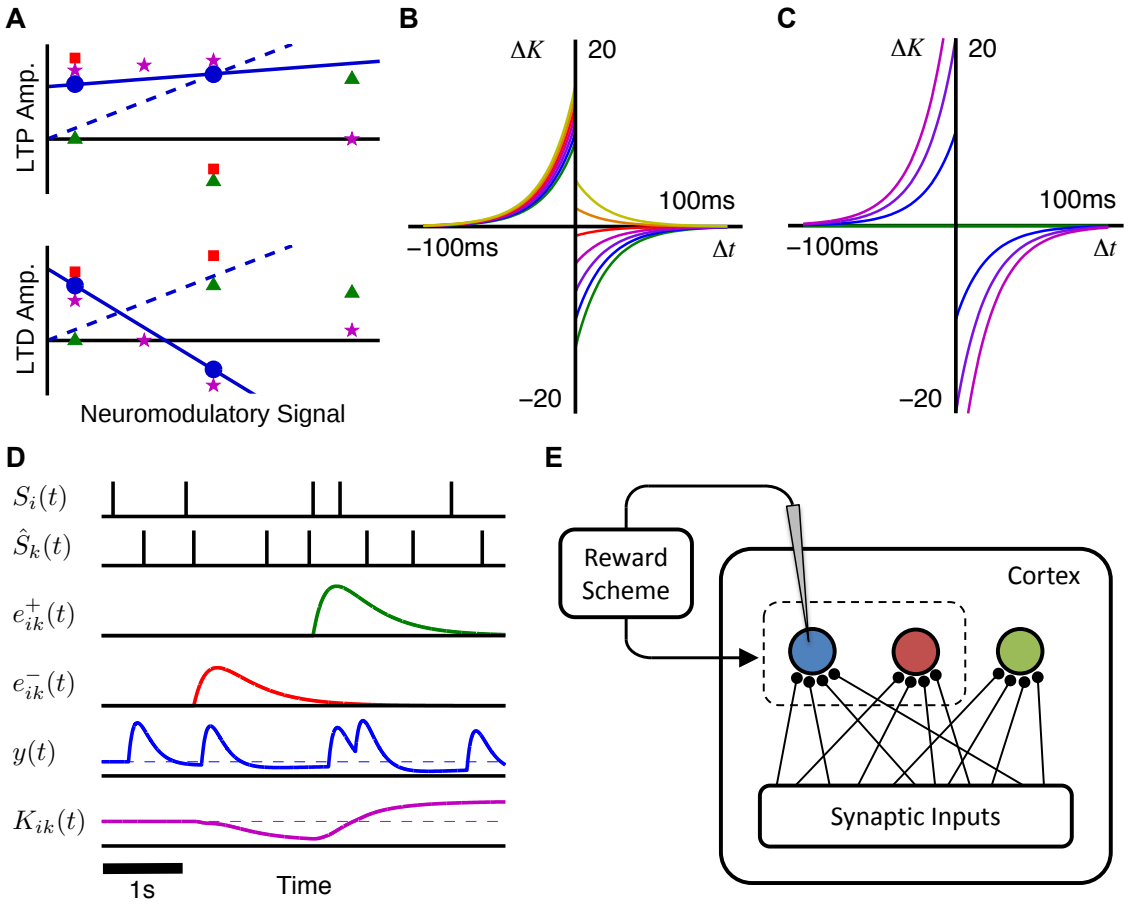


Figure 4.1: Modulation of STDP. A: Qualitative summary of the observed modulation of LTP and LTD amplitudes with increasing concentrations of dopamine (blue circles), octopamine (red squares), norepinephrine (green triangles), and acetycholine (magenta stars). These are based on observations by Zhang et al. (2009), Cassenaer and Laurent (2012), Salgado et al. (2012), and Sugisaki et al. (2011), respectively. The markers show qualitative effects only and the scales between the different modulators are not necessarily comparable. An example of our new RSTDP model parameterized to exhibit the same effect on STDP as dopamine (solid blue line). This is compared to an example of classical RSTDP model (dashed blue line). **B:** Effective learning windows for dopamine RSTDP for reward levels of 0 (green), 1 (blue), 2 (purple), 3 (magenta), 4 (red), 5 (orange), and 6 (yellow). The modulation factors are $p_+ = 1$, $p_- = -3$, $q_+ = 9$, and $q_- = 13$. **C:** Effective learning windows for classical RSTDP. Same axes and lines (not all are shown) as in B. The modulation parameters are $p_+ = 10$, $p_- = 10$, $q_+ = 0$, and $q_- = 0$. **D:** Conceptual plot of RSTDP variables during an operant conditioning experiment. Variables are (from the top down): post- and pre-synaptic spike trains, LTP and LTD eligibility traces, reward signal (dashed line shows the mean value), and synaptic weight (dashed line shows the initial value). **E:** Feedforward network where reinforced neuron (blue) is recorded from, determining the reward, which in turn influences changes made to the synapses into the reinforced and surround (red) neurons. The control neuron (green) represents either neuron before the operant conditioning experiment was preformed.

- **Control:** The control neuron represents either the reinforced or surround neuron before the operant conditioning experiment was performed.

Each spike from the reinforced neuron produced a perturbation of the reward signal, referred to as the reward kernel. The reward kernel has a mass, m, between 0 and 1. We initially focussed on the case where $m = 0$ and hence the mean of the reward signal, $\bar{y}$, is equal to the base level, y_0. This is the case in Figure 4.1D, where the kernel has a negative tail. The kernel is scaled by a reward strength, γ, which is positive to reinforce a high firing rate and negative to reinforce a low firing rate.

Analytically, we found that, for there to be reinforcement and unsupervised learning, rewards must produce a large increase in LTP then LTD and the reinforced neuron's firing must be irregular. We determined this by considering the changes to the mean feed-forward weight into neuron i, which is given by $\bar{K}_i = \frac{1}{N_k}\sum_{k=1}^{N_k} K_{ik}$, where there are N_K inputs and K_{ik} is the weight from input k to neuron i. Focussing on the case where the inputs to the neurons are uncorrelated, this mean weight evolves according to (see Appendix B.1 for derivation)

$$\dot{\bar{K}}_i \approx \eta\nu_i\Big\{(p_+\tilde{y}_i + q_+)f_+(\bar{K}_i)(\tilde{W}_+\hat{\nu} + \theta c_i) + (p_-\tilde{y}_i + q_-)f_-(\bar{K}_i)\tilde{W}_-\hat{\nu}\Big\}, \tag{4.3.1}$$

where $+$ and $-$ refer to the LTP and LTD parts of the learning window, respectively, η is the learning rate, ν_i is the firing rate of neuron i, c_i is the normalized, mean strength of the cross-covariances between neuron i and its inputs, θ describes the effect of these cross-covariances on learning, $\hat{\nu}$ is the input firing rate, $f_\pm(K_i)$ and $\tilde{W}_\pm$ are the weight dependence function and mass of the learning window parts, respectively, and $\tilde{y}_i$ gives the mean effective reward following the spikes of neuron i. For weights into the control and surround neurons, $\tilde{y}_i = y_0$ and $\tilde{y}_i = \bar{y}$, respectively. For weights into the reinforced neuron, $\tilde{y}_i = \bar{y} + \gamma a\eta_r$, where η_r describes the interaction between the reward kernel and the eligibility kernel, and γ and a are the reward strength and the net area of the auto-covariance function of the reinforced neuron, respectively. The statistic a provides a measure of irregularity in the firing of a neuron. In this way, reinforcement of a neuron occurs based on the average value of the reward signal following spike pairs.

We consider the case where the mean firing rates of the inputs are equal and only small spike correlations exist. In this case, the firing rate of a neuron is dependent on the mean excitatory synaptic weight of its inputs (assuming no, or fixed, inhibitory inputs). Therefore, for the reinforced neuron to increase its firing rate for a given set of inputs, the mean weight into it must increase compared to the mean weight into the control neuron (i.e., $\dot{\bar{K}}_{\text{reinforced}} > \dot{\bar{K}}_{\text{control}}$). From Equation (4.3.1), this requires that

$$\frac{p_+(\bar{y} + \gamma a\eta_r) + q_+}{p_+y_0 + q_+} > \frac{p_-(\bar{y} + \gamma a\eta_r) + q_-}{p_-y_0 + q_-}. \tag{4.3.2}$$

Assuming that $\gamma a\eta_r > 0$ and that $\bar{y} = y_0 = 1$, the requirement for reinforcement given by Equation (4.3.2) can be further simplified as

$$\frac{p_+}{p_+ + q_+} > \frac{p_-}{p_- + q_-}. \tag{4.3.3}$$

In classical RSTDP, where $q_+ = q_- = 0$, this requirement cannot be satisfied and neither an increase nor a decrease in the reinforced firing rate will occur. This is because, in classical RSTDP, LTP and LTD must both be zero at the same reward level and so, for there to be linear modulation of LTP and LTD that produces a balance of LTP and LTD at the base reward level, LTP and LTD necessarily match/balance at any reward level (dashed lines in Figure 4.2A). In the study by Legenstein et al. (2008), the reward signal was shifted so that there was zero LTP and LTD at the base level (dot-dashed lines in Figure 4.2A) and so, except at this point, no balancing of the amounts of LTP and LTD were necessary. In that case, reward above the base level produced Hebbian STDP while reward below the base level produced anti-Hebbian STDP. Therefore, provided that correlations between the inputs and the neurons caused there to be a greater amount of LTP than LTD while the reward was above the base level, RSTDP would lead to a stable increase in the synaptic weights and the firing rate of the reinforced neuron. However, in this situation, no unsupervised learning was present, as there was no LTP and LTD at the average reward level. If, in the study by Legenstein et al. (2008), the reward signal had not been shifted and there was LTP and LTD at the base reward level, unsupervised learning would be present but there would not be a balance of LTP and LTD at the base reward level. In this situation, the synaptic weights would either grow or decay unstably even without any rewards being given to the system.

In our RSTDP model, LTP and LTD are not necessarily both zero at the same reward level and so, to balance each other at the base reward level, they are not required to balance for all reward levels (solid lines in Figure 4.2A). In this case, it depends on the particular parameters as to whether reinforcement occurs or whether the 'rewards' actually behave as punishments and lead to a decrease in the firing rate of the neuron. For the dopamine inspired modulation parameters that we focus on, this requirement is met and reinforcement occurs. The inequality in Equation (4.3.3) and the illustration in Figure 4.2B show that, relative to the amounts of LTP and LTD at the base reward level, the increase in the amount of LTP with reward must be greater than the increase in the amount of LTD in order for the firing rate to increase (be reinforced). If the increase in LTP is the same as (less than) the increase in LTD, then the firing rate remains the same (decreases). Therefore, the parameters we consider here, which correspond to the results of Zhang et al. (2009), are just one of many possible sets of modulation parameters that we predict would lead to firing rate reinforcement.

Figure 4.1B shows that, for high values of dopamine, there is only LTP (post-pre spike pairs lead to LTP, instead of LTD). Because of this, if $\tilde{y}_i$, the mean effective reward following the spikes of neuron i, is sufficiently large then on average post-pre spike pairs with neuron i would lead to LTP and weights into neuron i would grow in an unstable manner. However, we found that there is a broad range of modulation parameters for which a stable fixed point for the mean input weight exists.

In addition to the modulation parameters, Equations (4.3.1) and (4.3.2) predict that the amount of reinforcement that occurs depends on the value of a, which we show depends on how irregular the firing of the reinforced neuron is.

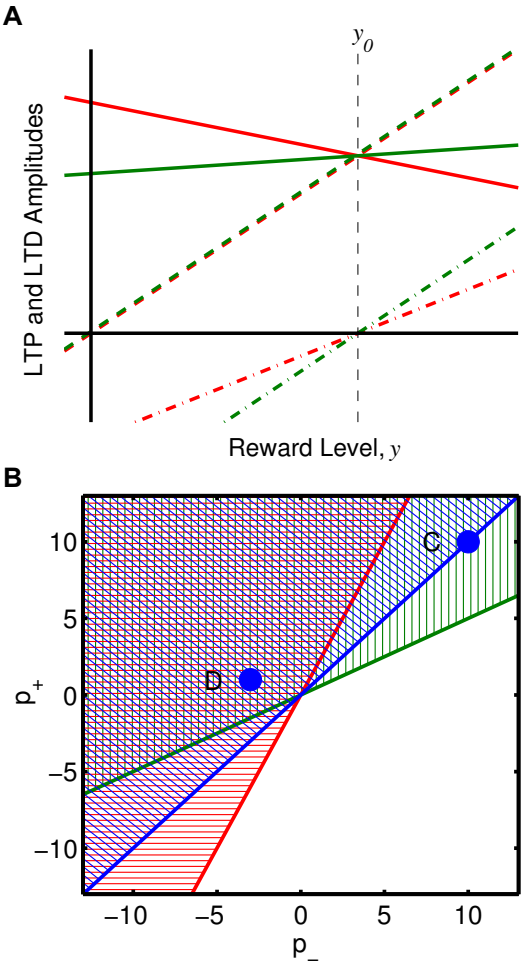


Figure 4.2: Comparison of our RSTDP model and classical RSTDP models. A: Amounts of LTP (green) and LTD (red) vs. reward level, with our RSTDP model (solid) and with classical RSTDP with and without unsupervised learning (dashed and dot-dashed, respectively) at the equilibrium synaptic weight. For classical RSTDP without unsupervised learning the reward signal has been shifted such that there is no LTP and LTD at the base reward level, y_0 (vertical, black, dashed line) instead of at zero reward, $y = 0$. **B:** An increase (decrease) in firing rate is predicted to occur in the hatched (unhatched) regions for LTP:LTD ratios at the base reward level (i.e., $(p_+ + q_+)$:$(p_- + q_-)$) of 2:1 (red), 1:2 (green), and 1:1 (blue). On the lines that divide these regions no increase or decrease is predicted. The points marked as C and D correspond to the classical and dopamine parameter sets used in this chapter (see Figure 4.1) with a base level ratio of 1:1.

4.3.3 Operant Conditioning Simulations

To support our analytical predictions, we simulated the learning during the operant conditioning of a neuron's firing rate using leaky integrate-and-fire (LIF) neurons in two different cases. In the first, the neurons received 10 000 excitatory inputs (E), while in the second, they received 8000 excitatory and 2000 inhibitory inputs (E+I). In the E+I case, only the

excitatory input weights changed due to RSTDP (i.e., the inhibitory inputs' weights were fixed). While we assume no covariance between the inputs, the correlations arising due to the influence of individual input spikes on the firing of the output neuron (spike triggered correlations) are significant and need to be taken into account. Figures 4.3A and 4.3B show numerically determined values for the strengths of these correlations (normalized by the firing rate) varying with mean input weight for the two cases. While the correlation strength increases with the mean input weight, it does so in a weaker fashion than the firing rate and so the normalized correlation strength decreases with mean input weight. The auto-covariance functions of the LIF neurons had a negative region for short time lags (Figure 4.3C). Negative regions represent spike time differences that are less likely to occur. In the integrator regime (E), the negative region is due to a minimum inter-spike-interval exhibited by the neuron. This minimum inter-spike-interval was smaller in the E+I case than the E case because the neuron exhibited more irregular firing. The net area of the auto-covariance function, a, is affected by the irregularity in firing: lower values occur for more regular firing and higher values for more irregular firing. Figure 4.3D shows how the value of a (the firing irregularity) changes as the balance between excitation and inhibition is varied.

We compared the analytical predictions to simulations with LIF neurons (see Appendix B.2 for derivation of the predicted weights/rates). While our analytical predictions hold for any weight dependence, for simulations we chose logLTD weight dependence (and also additive STDP). These results are shown in Figures 4.4A, 4.4B, and 4.4C. As predicted, classical RSTDP did not lead to an increase in the firing rate of the reinforced neuron in either E or E+I case. With dopamine RSTDP, this increase is seen but it is much smaller in the E case than in the E+I case. This has a number of causes, the most significant of which is that the negative region in the auto-covariance function, caused by the regular firing of the neuron in this case, almost completely cancels out the delta function at zero time lag (see Figure 4.3D), resulting in a small value for a. This has the effect of decorrelating the output spike train from itself and, therefore, the reward signal. This appears clearly in the average reward signal following spikes from the reinforced neuron (see Figure 4.4D). With low values of a (regular firing), the inter-spike-intervals of the reinforced neuron are large and this causes the spikes to occur less during times of high reward. This is the reason that less reinforcement occurs in the E case.

Other reasons for the smaller amount of reinforcement observed in the E case (compared with the E+I case) are that the correlation strength decreases faster with mean weight and that a larger increase in the mean input weight is required for the same increase in the firing rate (Figures 4.3A and 4.3B). The latter of these influences is somewhat made up for by the larger value of α used in the E case. Figures 4.4A, 4.4B, and 4.4C include analytical predictions that assume $a = 1$ and others that take the correct value of a into account ($a \approx 0.15$ for E and $a \approx 0.92$ for E+I). This shows the contribution that the value of a, the irregularity of the firing, has on the reduced reinforcement in E compared with the other factors.

Cases E and E+I typify mean- and fluctuation-driven regimes, respectively, for the

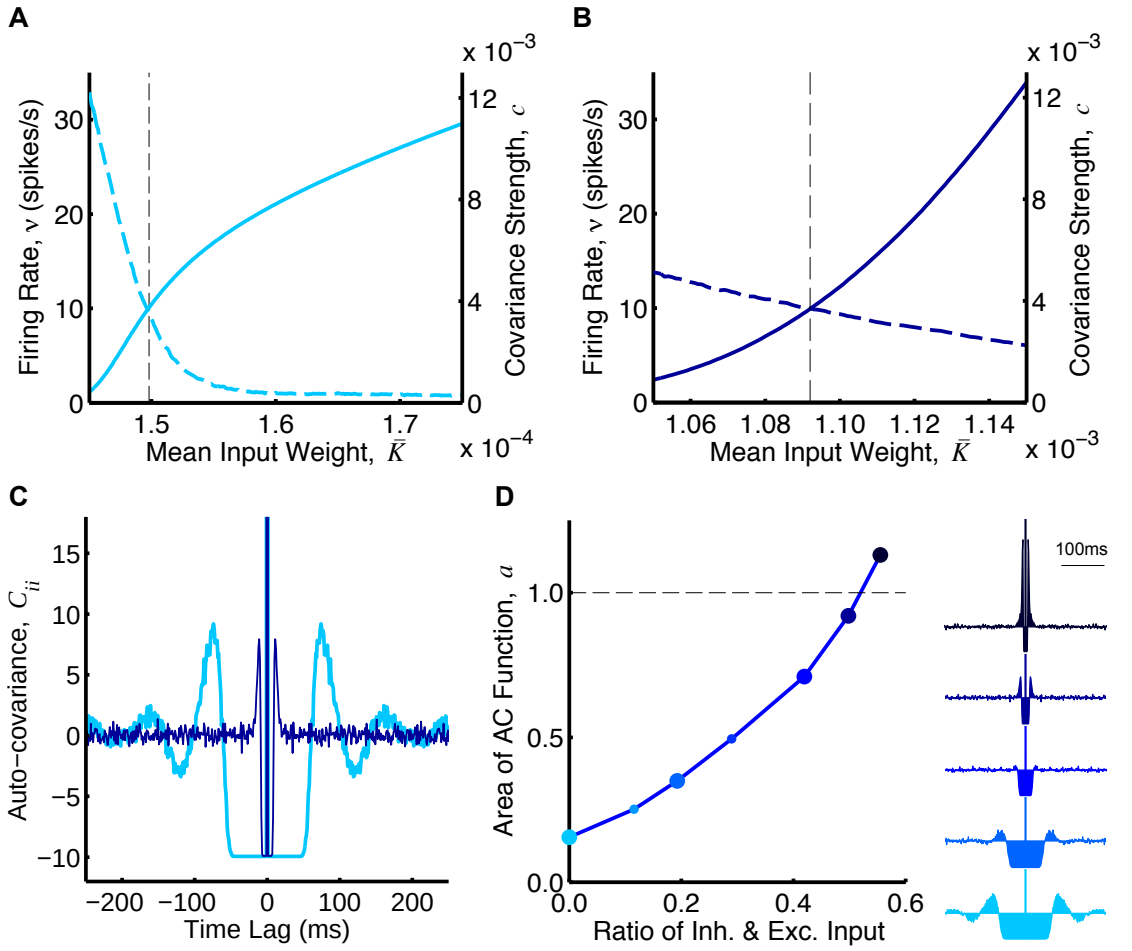


Figure 4.3: Numerically determined spiking statistics using the LIF neuron model. A: Mean output firing rate (ν, solid) and mean cross-covariance strength (covariance normalized by the firing rate) between the input and output spike trains (c, dashed) for different mean input weights, $\bar{K}$, for a LIF neuron with 10 000 excitatory inputs. **B:** Same as A but for a LIF neuron with 8000 excitatory inputs and 2000 inhibitory inputs. **C:** The auto-covariance function of the output spike trains, C_{ii}, of the LIF neurons in A (light blue) and B (dark blue) with mean input weights of 1.498×10^{-4} and 1.092×10^{-3}, respectively (dashed vertical lines in A and B). **D:** The net area of the auto-covariance (AC) functions, a, of LIF neurons (with input and output rates of 10 spikes/s) with 8000 excitatory inputs and 2000 inhibitory inputs for different ratios of the inhibitory and excitatory input currents. The auto-covariance functions for the first, third, fifth, sixth, and seventh points are shown to the right from bottom to top. The first point is the case in A and C (light blue), except with only 8000 excitatory inputs, and the fifth point is the case in B and C (dark blue). Table 4.5.1 shows the parameters used in the LIF neuron model.

neurons. We observed that varying the relative amount of inhibitory input controls a smooth transition between these two regimes (Figure 4.3D). The correlation between the firing of the reinforced neuron and the reward signal and, therefore, the amount of reinforcement, perfectly follows this transition (Figure 4.4D).

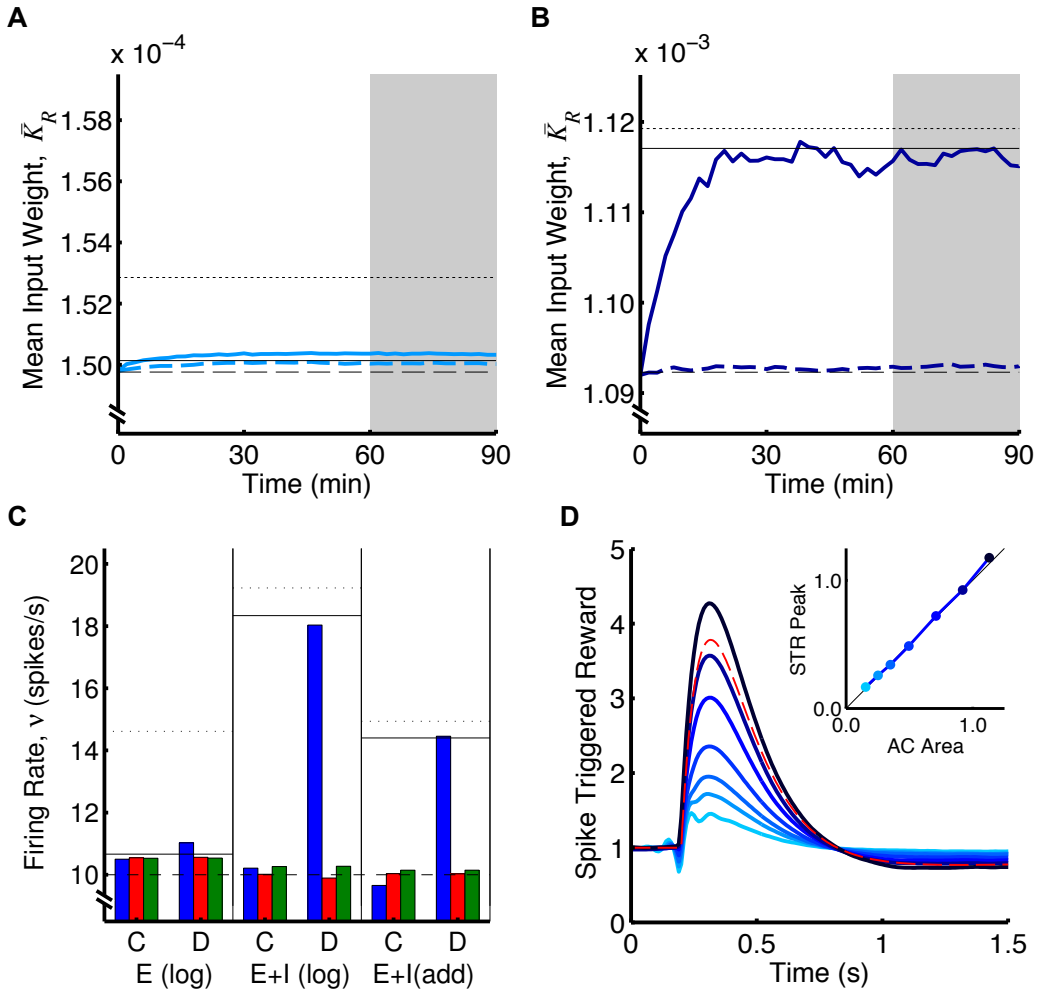


Figure 4.4: Operant conditioning experiment with LIF neurons. **A:** Mean weight into the reinforced neuron ($\gamma = 0.06$) over time for LIF neurons receiving 10 000 excitatory inputs where the weights are updated using the classical (dashed) and dopamine (solid) RSTDP models. Horizontal lines represent analytical predictions for classical RSTDP (dashed), dopamine RSTDP where $a = 1$ (dotted), and dopamine RSTDP where the correct value of a is assumed (solid). **B:** Same as A with 8000 excitatory inputs and 2000 inhibitory inputs (inhibitory synaptic strengths were fixed at 0.01). **C:** The mean firing rates of the reinforced (blue), surround (red), and control (green) neurons for the last 30 minutes of the simulations (shaded areas in A and B) with classical (C) and dopamine (D) RSTDP in A (E (log)), B (E+I (log)), and as in B but with additive weight dependence (E+I (add)), as described by Equation (4.5.12). Horizontal lines represent analytical predictions as in A and B. **D:** The average reward signal after the reinforced neuron's spikes (spike triggered reward) for neurons with different ratios between the excitatory and inhibitory input currents. The different ratios shown increase from no inhibitory inputs (lightest blue) up to the strongest inhibitory inputs (darkest blue), and correspond to the points in Figure 4.3D. The first line corresponds to the E case in A and C while the sixth line corresponds to the E+I case in B and C. The inset shows the relationship between the net area of the auto-covariance (AC) function and the peak of the spike triggered reward (STR) curve normalized by the peak of the reward kernel (red dashed line).

Figure 4.4C also shows an example of this reinforcement learning with an additive weight dependence (E+I case only). This weight dependence includes rate-based learning terms, as used by Gilson et al. (2009b), and used slightly different modulation parameters to achieve stable equilibria (see Table 4.5.3). These simulations show similar results as for the logLTD weight dependence.

4.3.4 Correlated Inputs

We simulated the learning during the operant conditioning experiment where the inputs (excitatory and inhibitory) contained pairwise spike correlations and found that reinforcement still occurs and that the firing rate of the surround neuron also increased. We used two different methods of generating input correlations: the single and multiple interaction process (SIP and MIP, respectively) models (Kuhn et al., 2003). Introducing correlations to the inputs leads to a higher firing rate even without providing the system with rewards. As shown in the inset of Figure 4.5A, we used smaller values of the modulation offset, q_+, with dopamine RSTDP so that the stable firing rate of the control neurons remained at 10 spikes/s. For classical RSTDP, equal reductions were made to p_+ to achieve the same outcome. Figure 4.5A shows the resulting firing rates of the reinforced and surround neurons from simulations with different input correlations with dopamine RSTDP. Using either method, we observed a lower firing rate after learning for the reinforced neuron than for the uncorrelated case but reduction was larger with SIP correlations. We also observed an increase in the firing rate of the surround neurons above baseline (10 spikes/s) using either method. While this reduction may not have completely saturated with a covariance strength of $\hat{c} = 0.02$, the trend appears to be sufficiently captured. Also, as the increase in the firing rate of surround neuron is due to its firing becoming correlated with the reinforced neuron's, our model does not predict that the surround neuron would ever increase its firing rate more than the reinforced neuron. Figure 4.5B shows the firing rates for only $\hat{c} = 0.01$ with both classical and dopamine RSTDP and compares them to the case with uncorrelated inputs. There is no apparent reinforcement of the firing rates of either neuron for classical RSTDP with input correlations.

4.3.5 Non-zero Reward Kernel Mass

We found a similar result to adding correlated inputs, when we considered the case where the mass of the reward kernel, m, is no longer zero. In this case, the mean of the reward signal, $\bar{y}$, is not fixed at the base level, y_0. Instead, it is given by

$$\bar{y} = y_0 + \gamma m \bar{\nu}_{\mathrm{R}}, \tag{4.3.4}$$

where γ is the reward strength and $\bar{\nu}_{\mathrm{R}}$ is the firing rate of the reinforced neuron. Figure 4.6A shows the analytical predictions for the mean firing rates of the neurons after learning for different reward strengths for $m = 0.00$ and $m = 0.05$. These results are supported by simulations, as shown in Figures 4.6A and 4.6C. For dopamine RSTDP, we observed that

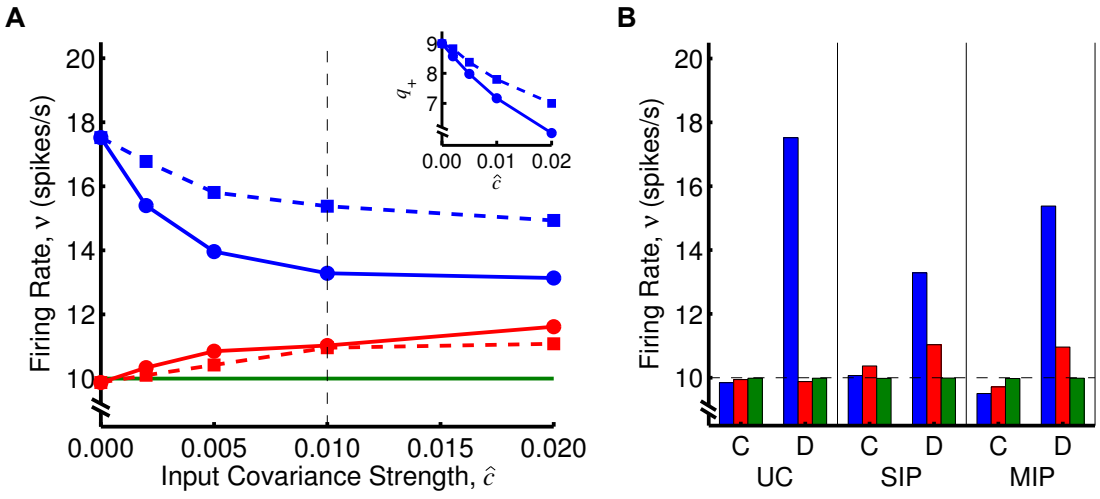


Figure 4.5: Operant conditioning experiment with correlations between inputs. A: Firing rates of reinforced (blue), surround (red), and control (green) neurons after learning in simulations with dopamine RSTDP for different input correlations ($\hat{c} = 0.000$, 0.002, 0.005, 0.010, and 0.020) introduced using two different methods. The first method (SIP, solid with circles) leads to common spikes across many spike trains, while the second (MIP, dashed with squares) does not. *Inset:* Smaller values of the modulation offset, q_+, were used so that the stable firing rate of the control neurons remained at 10 spikes/s. **B:** Firing rates of the three neurons after learning with classical (C) and dopamine (D) RSTDP for uncorrelated inputs ($\hat{c} = 0.000$) and with input correlation ($\hat{c} = 0.010$, dashed vertical line in A) introduced using the two different methods.

the firing rate of the surround neuron (as well as the reinforced neuron) increased above that of the control neuron when using a non-zero mass reward kernel. This was because the reward signal mean was no longer fixed but increased according to Equation (4.3.4). Because of this, we observed that the reinforced firing rate was unstable if the reward strength and kernel mass were too large. For classical RSTDP, neither the reinforced nor the surround firing rates increased.

4.3.6 Differential Reinforcement

We also considered the case where there are two differentially reinforced neurons (i.e., the neurons have positive and negative reward strength, respectively). In this case, the mean reward is given by

$$\bar{y} = y_0 + \gamma_{\mathrm{HR}} m \bar{\nu}_{\mathrm{HR}} - \gamma_{\mathrm{LR}} m \bar{\nu}_{\mathrm{LR}}, \tag{4.3.5}$$

where γ_{HR} and γ_{LR} and $\bar{\nu}_{\mathrm{HR}}$ and $\bar{\nu}_{\mathrm{LR}}$ are the reward strengths and firing rates of the neurons reinforced for high and low firing rates, respectively. Figure 4.6B shows the analytical predictions for the mean firing rates of the four neurons (two differentially reinforced neurons and surround and control neurons) after learning for a positive reward strength of 0.035 and different negative reward strengths for $m = 0.00$ and $m = 0.05$. These results are supported by simulations, as shown in Figures 4.6B and 4.6C. As was

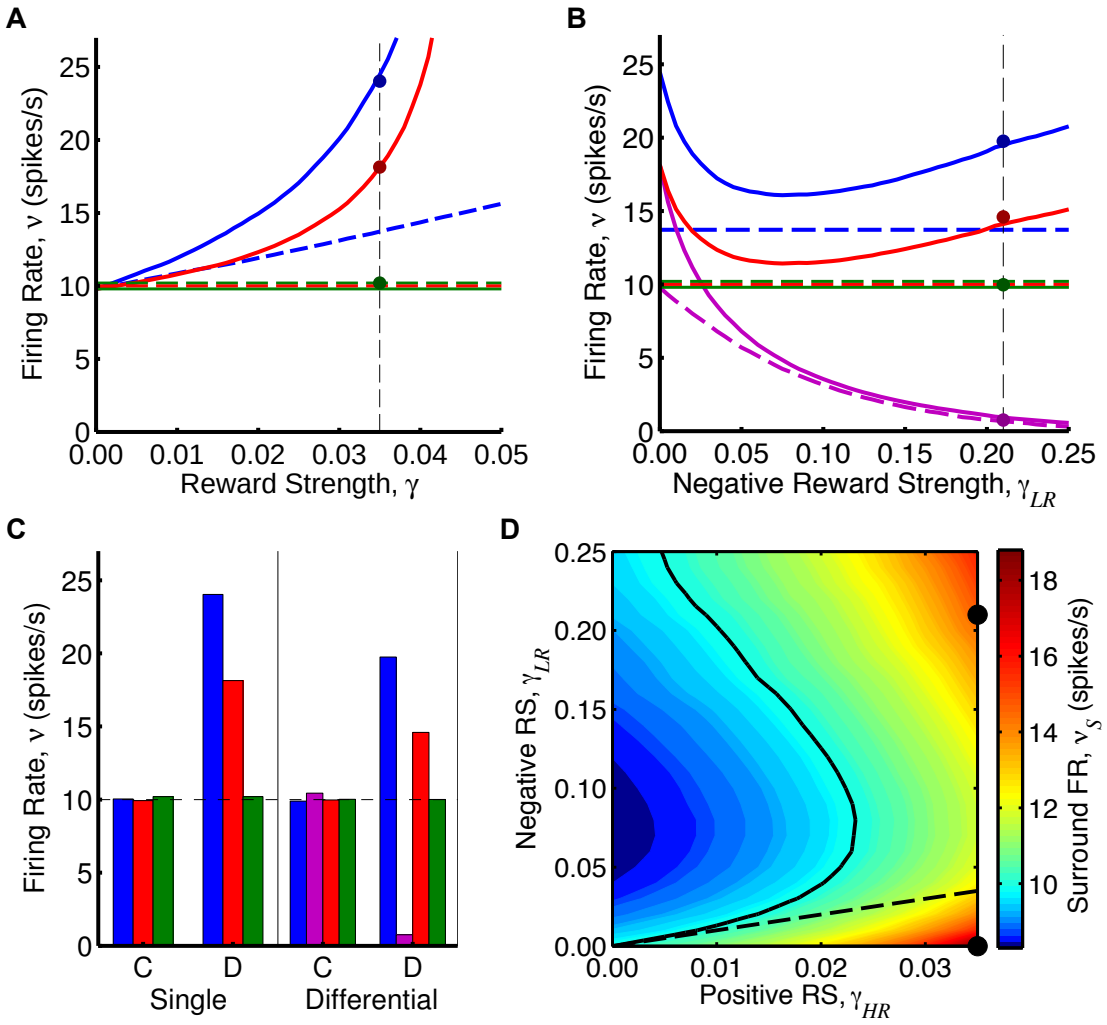


Figure 4.6: Operant conditioning experiment with non-zero-mass reward kernels. A: Firing rates of reinforced (blue), surround (red), and control (green) neurons after learning with dopamine RSTDP using reward kernels with masses of 0.05 (solid) and 0.00 (dashed) with reward strength. The green solid line and the red and green dashed lines are shown separate but are actually in line. Dots show the firing rates after learning from a simulation using the 0.05 mass reward kernel. **B:** Same as A but with an additional neuron (magenta) that is reinforced for a low firing rate. The high-rate reinforced neuron has fixed reward strength of 0.035, while the negative reward strength is varied. **C:** Firing rates of the three neurons (same colors as in A and B) after learning with classical (C) and dopamine (D) RSTDP for the single reinforced neuron, $\gamma = 0.035$, and differentially reinforced neurons, $\gamma = 0.035$ and -0.21 (vertical dashed lines and dots in A and B). **D:** Heat map of the firing rate (FR) of the surround neuron as the reward strengths (RSs) of the two neurons are varied. The solid line shows where the firing rate is unchanged from the base level (10 spikes/s) and the dashed line shows where the positive and negative reward strengths are equal in magnitude.

the case with only one reinforced neuron, classical RSTDP did not lead to changes in the firing rates of any of the neurons. For dopamine RSTDP, we observed a decrease in the firing rate of the low-rate reinforced neuron, either for all values of γ_{LR} (with $m = 0.00$) or for values of γ_{LR} above a certain threshold (with $m = 0.05$), in addition to the increase in the firing rate of the high-rate reinforced neuron. Interestingly, as the negative reward signal increased, there was an initial decrease in the stable firing rate of the high-rate reinforced and surround neurons followed by a slow increase. This increase is due to the decreasing stable firing rate of the low-rate reinforced neuron having less of an effect on the mean of the reward signal. Figure 4.6D shows how the stable firing rate of the surround neuron depends on the two reward strengths. Depending on the two reward strengths, the stable firing rate of the surround neuron is above or below the firing rate of the control neuron.

4.3.7 Comparison with R-max Model

As discussed by Frémaux et al. (2010), the average change in synaptic weights due to reinforcement learning rules can be split into the unsupervised and reward learning components. The reward learning component depends on the covariance between neural activity and reward, while the unsupervised learning component is independent of this covariance, depending only the mean reward value. This separation of components is given by

$$\dot{K}_{ik} \propto \left\langle \mathbb{C}\Big[p_+ e^+_{ik}(t) + p_- e^-_{ik}(t), y(t)\Big] + \mathbb{E}\Big[e^+_{ik}(t)\Big]\Big[p_+\bar{y} + q_+\Big] + \mathbb{E}\Big[e^-_{ik}(t)\Big]\Big[p_-\bar{y} + q_-\Big]\right\rangle_T, \tag{4.3.6}$$

where $\mathbb{C}[A, B]$ denotes the covariance between A and B, $\mathbb{E}[A]$ denotes the expected value of A and $\langle x\rangle_T$ denotes the temporal average of signal x. The first term in the equation is the reward learning component and the second and third terms combine to give the unsupervised learning component. For R-max and classical RSTDP, this simplifies to

$$\dot{K}_{ik} \propto \left\langle \mathbb{C}\Big[e_{ik}(t), y(t)\Big] + \mathbb{E}\Big[e_{ik}(t)\Big]\bar{y}\right\rangle_T, \tag{4.3.7}$$

where $e_{ik}(t) = e^+_{ik}(t) + e^-_{ik}(t)$. To maximize the reward that the system receives the unsupervised component needs to be as small as possible. The major difference between R-max and RSTDP is that, in the R-max model, the unsupervised component (or bias) is always zero (i.e., $\mathbb{E}[e_{ik}(t)] = 0$). This is only possible because an assumption of the R-max model is that it has an unbiased estimator of the instantaneous firing rate of the post-synaptic neuron. In contrast, RSTDP is only able to have zero unsupervised bias if, in the classical case, the mean value of the reward signal is zero (or can be removed), or if, in our model, the mean value of the reward signal is such that $p_+\bar{y} + q_+ = 0$ and $p_-\bar{y} + q_- = 0$. However, we are interested in when this is not the case and there is an unsupervised learning component. The unsupervised learning component without any reward learning leads to a stable base firing rate, and the introduction of the reward learning component, during operant conditioning, should result in a shift of this stable point. As we have

shown, classical RSTDP is not able to both exhibit an ongoing unsupervised learning component that produces such a stable point and also elicit a shift in this stable point due to reinforcement learning.

In order to demonstrate how the operant conditioning experiment is different with and without an unsupervised learning component present, we used the Spike Response Model (Gerstner and Kistler, 2002) to compare our dopamine RSTDP model (with logLTD) to the R-max model (Frémaux et al., 2010). This is shown in Figure 4.7. Both models are able to elicit an increased firing rate in the reinforced neuron. For the same learning rate, the R-max model leads to much faster firing rate reinforcement so for comparison we have set the learning rate for the R-max model to be 60 times smaller than for the dopamine RSTDP model. Aside from the differences in learning rate and the size of the firing rate increase, there are two important differences between the models. They are both due to the fact that there is an unsupervised component (or bias) to the changes elicited by the dopamine RSTDP model but not with the R-max model. The first difference is that, using dopamine RSTDP, the firing rate returned to the base level during extinction, as observed in operant conditioning experiments (Fetz, 1969; Fetz and Baker, 1973; Kobayashi et al., 2010), while in the R-max model it did not. The second difference is that the firing rate saturated in the dopamine RSTDP model, also as observed experimentally, while in the R-max model it did not. With our RSTDP model, there is a transient drop in the firing rate of the surround neuron at the beginning of the extinction period. This is due to a transient decrease in the mean value of the reward signal due to rewards no longer being delivered and the negative tail of the reward kernel. A transient increase in this firing rate similarly occurs at the beginning of the reinforcement period.

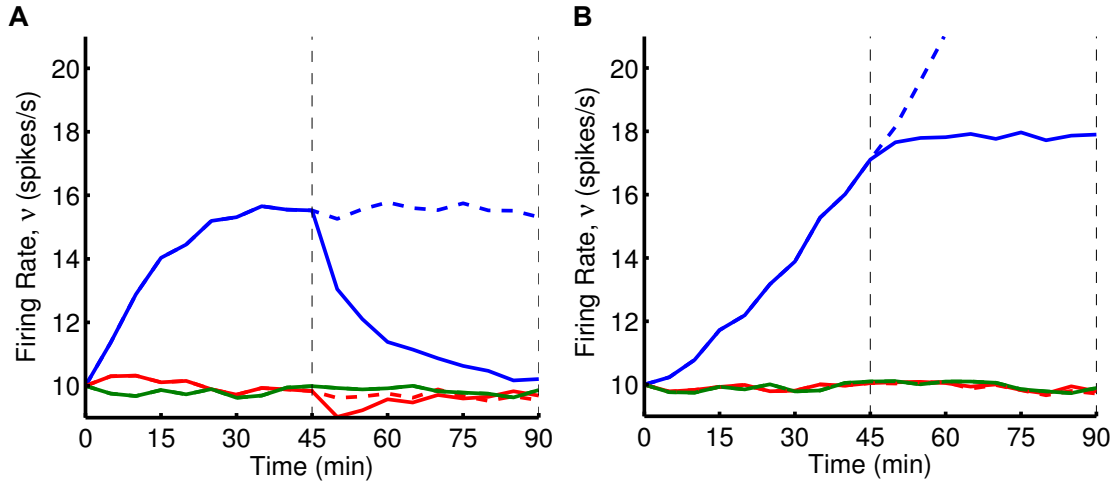


Figure 4.7: Comparison between dopamine RSTDP and R-max. A: Change in the firing rate over time for the reinforced (blue), surround (red), and control (green) neurons using the dopamine RSTDP model ($\eta = 2.00 \times 10^{-8}$), SRM neurons, and 10 000 excitatory inputs. A reward strength of 0.2 is used during the first 45mins and this is either maintained for the second 45mins (dashed) or reduced to 0.0 (solid). **B:** Same as A but using the R-max model ($\eta = 3.33 \times 10^{-10}$).

4.3.8 Summary of Results

We considered RSTDP in the case where LTP and LTD exist, both without any rewards and also at the mean reward level, which means that unsupervised learning is present. We showed that, in this situation, classical RSTDP is not able to elicit the neuronal firing rate reinforcement that is observed in experiments and in models which assume that there is no unsupervised learning. We proposed a new RSTDP model, which better captures the experimentally observed modulation of STDP by dopamine, and showed that it is able to elicit firing rate reinforcement. Without any rewards, the unsupervised learning component led to a stable base firing rate (this was demonstrated with the control neuron) and, during an operant conditioning experiment, a reward learning component was introduced and, coexisting with the unsupervised learning component, led to a shift in the firing rate of the reinforced neuron. We identified that this reinforcement is much stronger when the neurons are in a fluctuation-driven regime (as opposed to a mean-driven regime), such as when they receive a balance of excitatory and inhibitory inputs. We demonstrated that our findings are robust to the weight dependency used, the input correlations, and whether the mean of the reward signal is fixed or dependent on the reinforced firing rate.

4.4 Discussion

4.4.1 Related Models of Operant Conditioning

Previous reinforcement learning models, such as classical RSTDP and R-max, are able to perform operant conditioning tasks only when they do not have an unsupervised component (or bias) to the synaptic changes they elicit (Legenstein et al., 2008; Frémaux et al., 2010). For R-max, this is the case regardless of the reward signal statistics, but, for classical RSTDP, this is only true when there is no LTP and LTD at the average reward value. However, there is much experimental evidence suggesting that unsupervised learning occurs in the brain. This includes all experiments in which STDP is observed to occur and especially the findings of Zhang et al. (2009), which show that LTP and LTD are always present regardless of the dopamine concentration. An unsupervised learning component is also evident in the operant conditioning experiments when the reinforced firing rate returns to its original level during extinction (Fetz, 1969; Fetz and Baker, 1973). Figure 4.7 shows that our dopamine RSTDP model, with its unsupervised bias, can capture this behavior, unlike a model without an unsupervised component, such as R-max. A further aspect to the R-max model is that it requires an unbiased estimator of the instantaneous firing rate of the post-synaptic neuron in order to ensure there is never an unsupervised bias.

While a learning rule with an unsupervised learning component cannot always maximize the rewards received, it is not clear that learning rules employed by the brain are able to either. For example, in certain learning tasks, such as where perceptual roving is involved, R-max has been shown to out-perform the human brain (Herzog et al., 2012).

This was our reason for considering the operant conditioning learning task in this chapter. This simple situation can be compared directly with experiments and it is important to understand cases such as this before considering more general and complex learning situations. While out of the scope of this chapter, we would expect our model to perform similarly in more complex reinforcement learning tasks. As in this simple task, the unsupervised learning component would work against the reward-based changes but given sufficiently strong reinforcement learning components there is no reason why these learning tasks could not be performed.

Frémaux et al. (2010) argued that RSTDP is not an appropriate model of reward learning because it is sensitive to changes in the mean of the reward signal and will only work if the mean reward can be estimated without bias and subtracted from the current reward. However, in the simple operant conditioning protocol corresponding to published experiments (Fetz, 1969; Fetz and Baker, 1973; Kobayashi et al., 2010), we show that reward learning can coexist with unsupervised learning provided that certain conditions are imposed on how the STDP learning window changes with the value of the reward signal. Also, while Frémaux et al. (2010) considered a system in which rewards with positive mass (net area) were given and the mean reward over multiple trials had to be estimated and removed, we considered a model of dopamine dynamics in which this was unnecessary. Similar to Legenstein et al. (2008), we assumed that rewards (bursts of dopamine) that the system received had zero mass, with dopamine dropping below baseline after an initial burst. Although this remains to be explored, it results in a mean reward value that is fixed and, therefore, the presence of a critic to accurately estimate this mean (as discussed by Frémaux et al. (2010)) would be unnecessary.

4.4.2 Reward Prediction

In the actual operant conditioning experiments, rewards are not given for each of the output spikes. However, visual feedback is presented to the monkey at the level of individual spikes and, through classical conditioning, we assume that the dopamine response comes to be elicited by the more frequent and earlier feedback of the spikes (conditioned stimuli) as this is predictive of the less frequent and delayed rewards (unconditioned stimuli). For this reason, we believe the reward signal we have used, in which kernels for each of the output spikes are summed, is consistent with the evidence that dopamine encodes reward prediction error (RPE) (Schultz et al., 1997). While dopamine ceases to be released for the actual rewards, no further predictor of the reinforced spikes exists and we expect that dopamine continues being released as these spikes occur.

We made the same type of assumptions for the case where a differential firing rate was being reinforced. As in the simple case, the visual feedback of the spikes is completely predictive of the rewards received. The only difference is that spikes from the neuron that is negatively punished for firing (the low-rate neuron) predict less (or later) rewards and so we assumed that these spikes should lead to a drop in the dopamine concentration.

4.4.3 Firing Regimes

Neurons can operate as integrators that accumulate inputs over time to reach a threshold or as coincidence detectors that are sensitive to inputs arriving at the same time. These two different modes are referred to as mean-driven and fluctuation-driven regimes, respectively. In simple network models that only include excitatory synapses, neurons can only operate in a mean-driven regime, where firing is regular. However, when neurons receive a balance of excitatory and inhibitory inputs, they operate in a fluctuation-driven regime with high firing variability (Tsodyks and Sejnowski, 1995; van Vreeswijk and Sompolinsky, 1996; Burkitt, 2001). Experimental studies suggest that this is how cortical neurons operate (Shu et al., 2003; Haider et al., 2006).

In this chapter, we found that firing rate reinforcement is stronger for irregular firing neurons. This is consistent with previous reinforcement learning studies (Seung, 2003; Xie and Seung, 2004; Christodoulou and Cleanthous, 2011), which found that firing variability is important for ensuring correlation between the reward signal and the neural firing to be reinforced. Here, we controlled the firing variability of LIF neurons by varying the relative amounts of excitatory and inhibitory inputs to the neurons.

In all the simulations in this chapter, the input firing rates (and the control firing rate) were 10 spikes/s. This was based on the observed firing rates in the corresponding experimental studies (Fetz, 1969; Fetz and Baker, 1973; Kobayashi et al., 2010). For lower firing rates, Equation (4.3.1) predicts a lower learning rate and a stronger influence of the cross-covariances between neurons and inputs, but it still predicts qualitatively similar outcomes for the firing rate changes.

4.4.4 Experimental Predictions

We suggest three different types of possible experiments in which our model makes testable predictions. The first relates to the firing regime of the reinforced neuron. We predict that the effectiveness of the reinforcement learning is dependent on the firing regime of the neuron being reinforced. Fetz and Baker (1973) describe the reinforced neuron in their experiments as firing in bursts. This type of firing regime would have an auto-covariance function with a net area greater than 1. This fits with our chapter, which predicts that this type of firing is beneficial to the reinforcement of firing rates (Figures 4.3D and 4.4D). To further test this prediction, operant conditioning experiments could be performed on neurons with different firing regimes, in particular, differently shaped auto-covariance functions. These could be different neurons, potentially in different brain regions, which are observed to naturally produce different firing behaviors. Alternatively, it may be possible to experimentally modify the firing statistics in a single neuron.

The second type of experiment relates to directly controlling a particular neuromodulator, such as dopamine, in the manner described in this chapter and observing the firing rate changes. This would allow the RSTDP mechanism to be investigated more explicitly, without assuming the dopamine signal based on the reward scheme. As mentioned in

the Introduction, other neuromodulators have been observed to affect STDP (see Figure 4.1A). It would be of particular interest to carry out this experiment with one of these modulators. This chapter predicts that neurons could either be reinforced or punished with the same reward signal depending on the neuromodulator and concentrations used. For example, a burst of octopamine could be injected into an area of the mushroom body of a locust for each spike from an arbitrarily chosen neuron such that it resembles the reward signal considered in this chapter. A similar experiment to this was performed by Nargeot et al. (1999), where an analogue of the operant conditioning of Aplysia was performed by stimulating the esophageal nerve, releasing dopamine.

The third type of experiment relates to the behavior of a nearby neuron, especially during the differential reinforcement experiment. During operant conditioning experiments, where a high firing rate was being reinforced, the firing rates of nearby neurons, which were not being reinforced, were also observed to significantly increase (Fetz and Baker, 1973). This increase was much more variable and in some cases was larger than the increase in the reinforced neuron. In our chapter, while the increase would never be more for the surround neuron than the reinforced neuron, this is consistent with there being correlated inputs (and, therefore, correlations between the neurons) or with a reward kernel with positive mass (and, therefore, an increase in the mean of the reward signal), or with both of these. Fetz and Baker (1973) qualitatively observed correlations between the neurons but did not carry out more quantitative measurements or analysis. During the operant conditioning of the firing rate of a neuron, correlations between the reinforced neuron and a nearby neuron could be measured and compared with the increases of the firing rate of the two neurons. Alternatively, the firing of a nearby neuron could be controlled and made to fire independently of its inputs and, more importantly, independently of the reinforced neuron. After the firing rate of the reinforced neuron has increased, the control of the nearby neuron could be released and the firing rate that it exhibits immediately afterwards due to its inputs could be observed. Our model predicts that the firing rate of a nearby neuron will increase less if it is not correlated with the reinforced neuron. If there was still a firing rate increase, this would assumedly be due to an increase in the mean reward value. In this case, another experiment could be performed, observing the change in firing rate of a nearby neuron during the differential firing rate reinforcement of two neurons. Figure 4.6D shows that whether the firing rate of the surround neuron increased or decreased depended on the relative reward strengths of the two differentially reinforced neurons.

4.4.5 Other Plasticity Models

We focussed on two specific weight dependencies (logLTD and additive), but Equation (4.3.1) holds for any pair of weight functions. Because the mechanism for the firing rate reinforcement is in the differential modulation of LTP and LTD, we would expect similar findings regardless of the weight dependence. It remains to be seen how more detailed models such as triplet STDP (Froemke and Dan, 2002; Pfister and Gerstner, 2006) and voltage-based STDP (Clopath et al., 2010) could be incorporated into RSTDP and how

this would affect the results of this chapter.

Building upon earlier models (Shouval et al., 2002, 2010), Graupner and Brunel (2012) proposed a synaptic plasticity model based on postsynaptic calcium concentrations of cells. This biophysically based model is able to exhibit the results of many plasticity experiments relating to different STDP windows, pairing with postsynaptic spikes and bursts, triplet and quadruplet STDP, firing rate effects, and the effects of dendritic location. While our RSTDP model allows the change in the STDP learning window that Zhang et al. (2009) observed to occur with the addition of dopamine, this same dopamine dependence could be more simply incorporated by the modulation of just one of the parameters in the calcium-based plasticity model.

4.5 Methods

4.5.1 Neuron Models

We considered three neuron models: the Poisson neuron model, the leaky integrate-and-fire (LIF) neuron model, and the Spike Response Model (SRM) (Gerstner and Kistler, 2002). The Poisson neuron model was used in the analytical derivations, together with numerically determined functions for the firing rate and auto- and cross-correlations for the spike trains with mean input weight for the LIF neuron model. This aided the comparison between our analytical results and simulations with the LIF neuron model. The SRM is only used when comparing our RSTDP model to the R-max model.

The Poisson neuron model is a stochastic model that outputs a spike train that is a realization of an inhomogeneous Poisson process (Kempter et al., 1999). The intensity function of this process is analogous to the membrane potential of the neuron. It is made up of a spontaneous rate and the weighted sum of post-synaptic response kernels given by

$$\lambda_i(t) = \nu_0 + \sum_k K_{ik}(t) \sum_n \epsilon(t - t_{k,n} - \hat{d}_{ik}), \tag{4.5.1}$$

where $\lambda_i(t)$ is the intensity function for the ith neuron at time t, ν_0 is the spontaneous rate (assumed to be zero in this chapter), $K_{ik}(t)$ is the synaptic weight from input k to neuron i, $\epsilon(t)$ is the excitatory post-synaptic potential (EPSP) kernel, $t_{k,n}$ is the time of the nth spike output by neuron k, and $\hat{d}_{ik}$ is the axonal delay from neuron k to neuron i. Synapses here are modeled as current based. This means that synaptic input into the neuron is independent of the neuron's membrane potential (the intensity function in this model). The EPSP kernel used in this chapter has the form

$$\epsilon(u) = \frac{1}{\tau_B - \tau_A}\left(e^{-\frac{u}{\tau_B}} - e^{-\frac{u}{\tau_A}}\right)h(u), \tag{4.5.2}$$

where $\tau_B > \tau_A$ and $h(u)$ is the Heaviside function (i.e., $h(u) = 1$ for $u \geq 0$ and $h(u) = 0$ otherwise).

The leaky integrate-and-fire neuron is modeled using a single variable, $V_i(t)$. This represents the membrane potential for each neuron, i, and evolves according to

$$\frac{dV_i(t)}{dt} = \frac{1}{\tau_{\mathrm{m}}}\left(V_{\mathrm{p}} - V_i(t) + \sum_k \left\{K_{ik}(t)\big[E_{\mathrm{S},k} - V_i(t)\big]\sum_n \epsilon_{\mathrm{c}}(t - t_{k,n} - \hat{d}_{ik})\right\}\right), \qquad (4.5.3)$$

where τ_{m} is the passive membrane time constant, V_{p} is the resting membrane potential, $E_{\mathrm{S},k}$ is the synaptic reversal potential of the (excitatory) synapses from neuron k, and $\epsilon_{\mathrm{c}}(t)$ is the excitatory post-synaptic conductance (EPSC). The EPSC plays a similar role to the EPSP kernel, $\epsilon(t)$, in the Poisson neuron model and, because of this, we refer to both $\epsilon(t)$ and $\epsilon_{\mathrm{c}}(t)$ as EPSPs or EPSP kernels. $K_{ik}(t)$, $t_{k,n}$, and $\hat{d}_{ik}$ are the same as for the Poisson neuron model. A spike is produced when the membrane potential reaches a threshold value, V_{th}, and it is reset to V_{r}. An absolute refractory period is used, which prevents the membrane potential from changing during this time. The values of these parameters are given in Table 4.5.1. Similarly, the parameters for the Spike Response Model (the same as those used by Frémaux et al. (2010)) are shown in Table 4.5.2. Simulations with the LIF neuron model and the SRM were performed using an in-house neuron modeling software program, SpikeSim, used in previous studies (Gilson et al., 2009b,d, 2010a) and in Chapter 2 (Kerr et al., 2013).

Table 4.5.1: LIF Neuron Parameters

Parameter	**Value**
Synaptic Rise and Decay Times: τ_A, τ_B (ms)	1, 5
Membrane Time Constant: τ_{m} (ms)	20
Threshold, Resting and Reset Potentials: V_{th}, V_{p}, V_{r} (mV)	-50, -65, -65
Excitatory/Inhibitory Reversal Potentials: $E_{\mathrm{S},k}$ (mV)	0, -70
Refractory Period (ms)	1

Table 4.5.2: SRM Neuron Parameters

Parameter	**Value**
Synaptic Rise Time: τ_{s} (ms)	5
Membrane Time Constant: τ_{m} (ms)	20
Firing Rate at Threshold: ρ_0 (spikes/s)	60
Threshold and Reset Potentials: θ, u_{reset} (mV)	16, -5
Escape Noise Control: Δu (mV)	1

We considered the feed-forward network shown in Figure 4.1E, which has three different post-synaptic neurons: the reinforced, surround, and control neurons. Unless

otherwise stated, we have considered the case where there is a single reinforced neuron and an arbitrary number of surround and control neurons (the number does not affect the results). Each neuron outputs a spike train, S_i, with a mean firing rate, $\bar{\nu}_i$. They receive synaptic inputs from 10 000 input spike trains, $\hat{S}_k$, with strength, K_{ik}, and equal axonal delay, $\hat{d}$ (dendritic delays are assumed to be negligible). The input spike trains are assumed to be uncorrelated and have the same mean firing rate, $\hat{\nu}$. The mean feed-forward weights and mean firing rates of the reinforced, surround, and control neurons are denoted $\bar{K}_R$ and $\bar{\nu}_R$, $\bar{K}_S$ and $\bar{\nu}_S$, and $\bar{K}_C$ and $\bar{\nu}_C$, respectively. In simulations, the weights are initially the same and set to be approximately equal to $\bar{K}_C$.

4.5.2 Reward Signal

As in previous studies (Legenstein et al., 2008), we assumed that rewards given to the monkey affect the concentration of dopamine in the neural network. This is based upon the evidence linking dopamine to reward learning in the brain (Schultz et al., 1997). Dopamine is delivered to different brain regions by the axons of neurons located in the ventral tegmental area (VTA), whose activity is dependent not only on rewards received but also on predicted or expected rewards.

In the operant experiments by Fetz (1969); Fetz and Baker (1973); Kobayashi et al. (2010), monkeys were presented with a screen showing a signal that decayed with time but was incremented for each action potential measured from an electrode implanted in their precentral motor cortex or prefrontal cortex. If the signal reached a threshold value, a reward was given and the signal returned to a reset value. With this setup, the experiments showed that high firing rates were elicited. Negative punishment (i.e., the removal of reward in order to decrease a particular behavior) was performed with a similar setup, where measured spikes decremented the signal (and artificially generated spikes incremented the signal). In this case, low firing rates were elicited. Through a combination of positive reinforcement and negative punishment, they also showed that a differential between the firing rates of two neurons could be elicited.

In our model, the reward signal, which is related to the dopamine concentration, is driven by the firing of the reinforced neuron(s) and is given by

$$y(t) = y_0 + \sum_i \gamma_i \int_0^\infty S_i(t - d_r - r) g_r(r) dr, \tag{4.5.4}$$

where y_0 is the base level of the reward signal, $S_i(t)$ is the spike train of reinforced neuron i, d_r is the reward delay, and γ_i is the reward strength for neuron i (this can be either positive or negative for neurons whose firing affects the signal, or zero for neurons whose firing does not). Reward strengths correspond to the heights of the voltage pulses delivered to the feedback signal for each spike of reinforced neurons in the operant conditioning

experiments (Fetz, 1969; Fetz and Baker, 1973). The reward kernel, $g_r(t)$, is given by

$$g_r(t) = \left[\left(\frac{1}{\tau_{r,B} - \tau_{r,A}}\right)\left(e^{\frac{-t}{\tau_{r,B}}} - e^{\frac{-t}{\tau_{r,A}}}\right) - (1-m)\left(\frac{1}{\tau_{r,C} - \tau_{r,B}}\right)\left(e^{\frac{-t}{\tau_{r,C}}} - e^{\frac{-t}{\tau_{r,B}}}\right)\right] h(t), \tag{4.5.5}$$

where $\tau_{r,A}$, $\tau_{r,B}$, and $\tau_{r,C}$ are the rise, decay, and recovery time constants, respectively, and m is the normalized kernel mass. As in the chapter by Legenstein et al. (2008), we initially focussed on the case where the reward kernel has zero mass (i.e., $m = 0$ and so $\int_0^\infty g_r(r)dr = 0$). If this is the case, the mean of the dopamine signal is fixed ($\bar{y} = y_0$). This dopamine signal affects the synapses to the reinforced and surround neurons but not the control neurons. The dopamine signal that affects the control neuron(s) is one that remains fixed at the base level, y_0. The reward kernel parameters used in this chapter are shown in Table 4.5.3. Figure 4.1D shows an example of a reward signal, $y(t)$, dependent on the spike train of neuron i, $S_i(t)$.

4.5.3 RSTDP Model

Based upon the experimental results of Zhang et al. (2009), Figure 4.1B shows the observed effect that the concentration of dopamine has on the amplitudes of LTP and LTD (blue circles). These experimental observations suggest that LTD and LTP are non-zero when there is no dopamine, that as the concentration of dopamine increases, LTD and LTP change in different ways, and that for high dopamine concentrations, LTD switches to LTP. In addition to dopamine, other neuromodulators have been observed to affect STDP. These neuromodulators include acetycholine (Sugisaki et al., 2011) in the hippocampus of rats, octopamine in the mushroom body of locusts (Cassenaer and Laurent, 2012), and norepinephrine in the visual cortex of mice (Salgado et al., 2012). Their effects on LTP and LTD are illustrated with the markers in Figure 4.1A.

In the existing RSTDP model, "classical RSTDP", both LTP and LTD are modulated equally by the reward signal (i.e., the dopamine concentration) such that no synaptic changes can occur when there is no reward. This is illustrated in Figure 4.1A (dashed blue line). Figure 4.1C shows this as different learning windows (relationships between the timing difference of spike pairs and the change in synaptic weight) for different dopamine concentrations. This chapter introduces a new RSTDP model that can better capture experimental findings (Zhang et al., 2009; Sugisaki et al., 2011; Cassenaer and Laurent, 2012; Salgado et al., 2012). In our RSTDP model, the potentiation (LTP) and depression (LTD) parts of the STDP learning window ($\Delta t < 0$ and $\Delta t > 0$, respectively) are modulated separately by the reward signal. This new model is shown in Figure 4.1A (solid blue line) and with different learning windows in Figure 4.1B.

In our RSTDP model, changes to the feed-forward weights are given by

$$\Delta K_{ik}(t) = \eta \int_t^{t+\Delta t} \Big\{ e_{ik}^+(t')\big[p_+ y(t') + q_+\big] + e_{ik}^-(t')\big[p_- y(t') + q_-\big] \Big\} dt', \tag{4.5.6}$$

and so the time and ensemble averaged rate of change of these feed-forward weights is given by

$$\dot{K}_{ik}(t) = \eta\left\langle \mathbb{E}\Big\{ e^{+}_{ik}(t)\big[p_{+}y(t) + q_{+}\big] + e^{-}_{ik}(t)\big[p_{-}y(t) + q_{-}\big] \Big\} \right\rangle_{T}, \tag{4.5.7}$$

where η is the learning rate, $\mathbb{E}\{X\}$ is the expected value of a random variable X, and $\langle x(t)\rangle_T = \frac{1}{T}\int_t^{t+T} x(t')dt'$ is the temporal average of the signal, $x(t)$, over a timescale, T, that is slower than both the neuronal and reward signal dynamics. The eligibility traces for LTP and LTD are given by

$$e^{\pm}_{ik}(t) = f_{\pm}(K_{ik}(t)) \int_0^{\infty} g_c(r) \int_{-\infty}^{\infty} W_{\pm}(u) S_i(t-r) \hat{S}_k(t-r+u) du dr, \tag{4.5.8}$$

where $W_{\pm}(u)$ and $f_{\pm}(K)$ are the learning windows and weight dependence functions for LTP $(+)$ and LTD $(-)$, respectively. The modulation offsets, $q_{\pm}$, give the amplitude of LTP and LTD for zero reward, while the modulation slopes, $p_{\pm}$, describe how the reward signal affects the amplitudes of LTP and LTD, respectively. The eligibility kernel, $g_c(t)$, is given by

$$g_c(t) = \left(\frac{1}{\tau_{c,B} - \tau_{c,A}}\right)\left(e^{\frac{-t}{\tau_{c,B}}} - e^{\frac{-t}{\tau_{c,A}}}\right)h(t). \tag{4.5.9}$$

This learning process is described in Figure 4.1D.

The learning window, which is divided into the LTP and LTD windows, is given by

$$W_{+}(t) = e^{\frac{t}{\tau_{+}}}h(-t), \qquad W_{-}(t) = -e^{\frac{-t}{\tau_{-}}}h(t), \tag{4.5.10}$$

where τ_{+} and τ_{-} are the time constants for LTP and LTD, respectively. As the relative amplitudes of LTP and LTD are determined by the modulation parameters, the amplitudes of the learning windows were both set to to 1 to avoid redundancy in the parameters. For the same reason, the base value of the reward signal (which for zero-mass reward kernels is equal to the signal mean) is set to 1.

The type of weight dependence, $f_{\pm}(K)$, that we focussed on in this chapter was one with additive LTP and logarithmically dependent LTD. This was inspired by the weight dependence considered by Gilson and Fukai (2011). This weight dependence is referred to as "logLTD". The functions for logLTD are given by

$$f_{+}(K) = 1, \qquad f_{-}(K) = \frac{\log(1 + \alpha\frac{K}{K_0})}{\log(1+\alpha)}, \tag{4.5.11}$$

where α and K_0 are parameters defining the shape of the LTD weight dependence. This weight dependence was chosen because it provides an intermediate between additive and multiplicative weight dependencies. Additive STDP leads to strong competition between the synapses and a bimodal weight distribution. Multiplicative STDP leads to a unimodal weight distribution but only weak competition (Gilson and Fukai, 2011). LogLTD elicits

strong competition between the synapses, while producing a stable, unimodal weight distribution. We also considered additive weight dependence, where the functions are given by

$$f_+(K) = 1, \qquad f_-(K) = 1. \tag{4.5.12}$$

Additive weight dependence was considered with rate-based learning terms (Kempter et al., 1999), which are not modulated by the reward signal. These are given by ω_{in} and ω_{out}, which either increase or decrease the synaptic weight for each pre- or post-synaptic spike, respectively. When using an additive weight dependence, these rate-based terms are necessary to achieve a stable mean weight.

The parameters values for the eligibility kernel, learning window, and weight dependence functions are shown in Table 4.5.3 (the parameters of the weight dependence functions were chosen to produce the desired stable firing rate for the control neuron and to exhibit sufficient sensitivity to being reinforced). Equation (4.3.1) was derived from Equations (4.5.4), (4.5.7) and (4.5.8) using results from Bohrnstedt and Goldberger (1969) (see Appendix B.1). The analytical predictions for the resulting mean input weights, for the two different weight dependencies in Equations (4.5.11) and (4.5.12), are based on Equation (4.3.1) (see Appendix B.2).

Table 4.5.3: RSTDP Parameters

Parameter	**Value**
Reward Rise, Decay & Recovery Times: $\tau_{r,A}$, $\tau_{r,B}$, $\tau_{r,C}$ (s)	0.10, 0.15, 3.00
Reward Delay: d_r (s)	0.20
Base Reward Level: y_0	1
Eligibility Rise and Decay Times: $\tau_{c,A}$, $\tau_{c,B}$ (s)	2.0, 5.0
LTP/LTD Window Time Constants: τ_+, τ_- (ms)	20, 20
LogLTD Parameters (E): K_0, α	1.4541×10^{-4}, 5.0
LogLTD Parameters (E+I): K_0, α	1.0692×10^{-3}, 1.5
LogLTD Parameters (SRM): K_0, α	1.4550×10^{-4}, 15.0
Additive Input/Output Rate Parameters: ω_{in}, ω_{out}	0.1, 0.0
Dopamine Modulation Parameters (log): p_+, p_-, q_+, q_-	1, −3, 9, 13
Classical Modulation Parameters (log): p_+, p_-, q_+, q_-	10, 10, 0, 0
Dopamine Modulation Parameters (add): p_+, p_-, q_+, q_-	1, −3, 9, 13.64
Classical Modulation Parameters (add): p_+, p_-, q_+, q_-	10, 10.64, 0, 0

4.5.4 Covariances in the Network

We have focussed on the case where the inputs are uncorrelated and the neurons receive separate (non-overlapping) sets of input spike trains. While the inputs are uncorrelated,

correlations between the neurons and inputs arise due to the influence of individual input spikes on the firing of the output neuron. These are referred to as "spike triggered correlations". Therefore, for neurons i and j and one of the inputs, k, into neuron i, we have mean neuron-input cross-covariances, $\bar{F}_{ik}(u)$ and $\bar{F}_{jk}(u)$, and mean neuron-neuron auto- and cross-covariances, $\bar{C}_{ii}(u)$ and $\bar{C}_{ij}(u)$, given by

$$\begin{aligned} \bar{F}_{ik}(u) &= c_i \nu_i \epsilon(-u + \hat{d}), \\ \bar{F}_{jk}(u) &= 0, \\ \bar{C}_{ij}(u) &= 0, \\ \bar{C}_{ii}(u) &\approx a \nu_i \delta(u), \end{aligned} \tag{4.5.13}$$

where c_i is the magnitude of the spike triggering effect, $\epsilon(u)$ is the EPSP kernel, and a is net area of the auto-covariance function of neuron i for short time lags. For Poisson neurons, $c_i = N_K^{-1}$, where N_K is the number of input spike trains into each neuron. However, for LIF neurons, c_i is not constant but depends on the strength of the inputs into neuron i. Figures 4.3A and 4.3B show numerically determined values for c_i when there are only excitatory inputs and when there is a balance of excitatory and inhibitory inputs, respectively. For Poisson neurons, $a = 1$ (as the auto-covariance function is a Dirac delta function), while for LIF neurons, this is not necessarily the case. This discrepancy is often due to the minimum inter-spike interval that LIF neurons exhibit. While we approximated $\bar{C}_{ii}(u)$ as a delta function, Figure 4.3C shows that this is not the case on short time scales. Figure 4.3D shows how a and the shape of the auto-covariance function change with the ratio of inhibitory to excitatory input currents. These curves agree with analytical studies that considered the statistics of LIF neuron outputs (Moreno-Bote and Parga, 2006; de la Rocha et al., 2007).

For correlated inputs, $\bar{F}_{jk}(u)$ and $\bar{C}_{ij}(u)$ would no longer be zero and new curves for the output firing rate and the neuron-input and neuron-neuron covariance strengths with mean input weight would need to be determined. While this would be more complex, the analytical framework presented is able to incorporate these differences and make predictions for reinforcement learning with input correlations. However, in this chapter, we considered operant conditioning experiments with correlated inputs through simulations only, and did not analytically derive expressions for this case. In these simulations, we considered two methods for generating inputs with constant firing rates and pairwise covariances. The first, referred to as the single interaction process (SIP) model, introduces the pairwise covariances between inputs through common spike events, in which many inputs participate (Kuhn et al., 2003; Gütig et al., 2003; Meffin et al., 2006). The second, referred to as the multiple interaction process (MIP) model, introduces pairwise covariances without these common spike events (Kuhn et al., 2003). We considered input correlations of up to 0.02, consistent with the range of correlations typically observed in the cortex (Jermakowicz et al., 2009).

Chapter 5

Conclusion

The brain is able to perform and learn many difficult and complicated cognitive tasks. These include learning and recognizing sensory stimuli, navigation, attention, producing and comprehending language, reasoning, decision making, and the production of behavior and motor actions. However, these different tasks are rarely distinct or separable from each other and, in fact, most tasks involve a number of different cognitive processes working together. Furthermore, learning how to perform a task is intrinsically linked to the performance of the task. For this reason, a particular cognitive process should not be considered in complete isolation and how the processes interact needs to be considered.

The brain is an extremely complex system and can be investigated on a number of different scales, from individual ion channels and receptors to neurons, networks, and whole brain regions. In order to understand a system of this complexity, it is necessary to consider only small parts at a time. However, while the brain as a whole appears to be composed of smaller modules, these modules could never function separately from the rest of the brain. Therefore, because of its irreducibility, it is important to continually consider the way in which individually investigated modules fit into the larger system. In other words, as is the case for different cognitive tasks, it is important to not focus on only one scale or level of complexity, but instead consider the system on multiple levels simultaneously.

5.1 Summary

5.1.1 Unsupervised Learning of Neural Oscillations

In this project, we have considered three different cognitive tasks and investigated these with theoretical models on three different scales or levels. Each of these tasks related to the others, as did each of the different models. In Chapter 2, we explored how the unsupervised learning that results from synaptic plasticity (STDP) can produce structure in recurrently-connected local networks of neurons by learning temporal firing patterns in the inputs to

this network. The temporal patterns that we considered were simple oscillations, as well as, multiple, phase-offset oscillations. These patterns could represent stimulus in auditory areas or abstract neural codes using, for example, gamma oscillations. We showed how the structure that was formed encoded the stimulus or pattern and changed the neural responses to be selective to the encoded pattern. This analytical exploration of how delays in a recurrent network can encode temporal firing patterns provided insight into the role of unsupervised learning in the brain. Its generality allows for it to be applied or extended to many different types of firing patterns and initial network structures.

5.1.2 Cognitive Control with Cortical Units

In Chapter 3, we zoomed out to a higher-level and considered the connections between groups of neurons, where these groups were modeled as neural masses. Instead of considering a learning task, we instead considered how networks, which would have been formed by learning, could be rapidly remapped and controlled by top-down feedback. We proposed a model in which the oscillatory activity of the neural masses, or cortical units, was gated by top-down, feedback, which targeted the distal dendrites of the pyramidal neurons, and also the coherence between the oscillations of the feedforward and feedback inputs. We demonstrated how these units could be formed into circuits to perform logic operations and identified the different ways in which operations could be controlled by top-down feedback. We showed that more sophisticated and flexible top-down control was possible when the units were gated by the two mechanisms rather than only top-down feedback. Our proposed model and exploration of top-down control in cortical networks provide a novel way of considering how cortical networks can be remapped depending on different contexts or goals. This type of high-level cognitive control is central to understanding high-level cognition in general and our findings provide not only a framework for considering this but also suggest how previously proposed mechanisms may work together.

5.1.3 Coexistence of Reward and Unsupervised Learning

In Chapter 4, we went down to the level of individual neurons and considered how synaptic plasticity rules could produce a simple reinforcement learning phenomenon that has been observed experimentally. Here, we did not consider networks of neurons, but only the synapses onto single neurons and how rewards based on the firing of individual neurons could affect the changes to these synapses. We showed that previous models cannot elicit firing rate reinforcement while exhibiting both reward learning and ongoing, stable unsupervised learning. We proposed a new experimentally-based RSTDP model that is able to exhibit reward and unsupervised learning, while also leading to firing rate reinforcement similar to that observed experimentally. Both unsupervised and reward learning have been shown to play an important role in the brain and our findings provide an explanation for how they can coexist. Therefore, we provide an understanding of operant conditioning, a fundamental cognitive process, while presenting a model that is consistent with other learning models and can be applied to unsupervised learning tasks.

5.2 Future Areas of Investigation

Theoretical research should not only propose models that explain experimental findings and make experimentally testable predictions, it should also inspire and suggest further research and investigation.

5.2.1 Unsupervised Learning of Neural Oscillations

As discussed in Chapter 2, there are a number of different avenues of further research that this work provides. Firstly, it sets up a framework that could be easily extended to investigate delay selection with different temporal firing patterns, such as phase-locked, oscillatory firing patterns rather than sinusoidal oscillations. It would be expected that this would lead to a response that was more selective to the train oscillation as the narrower grouping of spikes should make discrimination of these inputs simpler. Delay selection could also be performed with more complex, oscillatory firing patterns, which corresponded to activity elicited by sounds containing a fundamental frequency as well as harmonic frequencies. This would be a useful model for explaining missing fundamental pitch perception; however, a more detailed model of the auditory neural system would need to be used in order to make this case.

Another large opportunity for future work lies in considering balanced networks with both excitatory and inhibitory neurons, where activity is in a fluctuation-driven regime and where oscillations, such as gamma frequency oscillations, can be generated internally and intrinsically by the recurrently connected groups of neurons. Demonstrating recurrent delay selection in this situation would provide a much stronger experimental link with the cortex. Furthermore, it would be of great interest to explore how delay selection would work between multiple, intrinsically oscillating groups of neurons. While the oscillation frequency of the groups would be fixed by the internal mechanisms generating them, the oscillation phases would not be and it may be possible for multiple sets of stable phase differences would arise between the groups. Inputs into these groups would potentially be able to shift the system between these different stable points.

5.2.2 Cognitive Control with Cortical Units

Chapter 3, which also considers interconnected, oscillating groups of neurons, opens up many areas for further investigation. First and most importantly, there remains the task of investigating how synaptic plasticity, both unsupervised and reward learning, could lead to the development of the types of cortical networks considered and how with new inputs it could adjust these networks. Further to this, the robustness of different networks to the addition of new units and to the removal of units due to synaptic plasticity could be explored. Introducing plasticity to this model would also strengthen the link between this work and that of Chapter 2.

Secondly, it could be considered how multiple frequencies (e.g., gamma and beta

frequencies) might coexist in these networks. This would relate our model more strongly with experimental work by Buschman et al. (2012), and would also likely result in new types of top-down control. Next, something that the model is currently missing is lateral connections between cortical units that are distinct from both feedforward and feedback connections. Units may actually be composed of overlapping sets of neurons and as such lateral connections may not correspond as directly to sets of synaptic connections as for feedforward and feedback connections. However, this means that lateral connections may be able to produce a range of different types of interactions between units that could be considered.

Finally, there is the potential for investigation of more detailed models that produce similar behaviors. This may simply be models that contain continuous activation levels and/or continuous oscillation phases. It would be of interest to observe whether any new types of interactions or behaviors arise due to these changes. Models in which individual neurons and synaptic connections are considered could also be investigated. In this case, it would be interesting to see under what conditions such models produced similar behaviors. More detailed models could also include connections that targeted only excitatory or inhibitory neurons, or even connections that targeted specific dendritic branches. For example, inhibitory synapses have been experimentally observed to target particular parts of a pyramidal neuron's dendritic tree and block only inputs into this area (Palmer et al., 2012).

5.2.3 Coexistence of Reward and Unsupervised Learning

An obvious extension to the work of Chapter 4 is to investigate the ability of the new RSTDP learning rule to perform more complicated learning tasks, such as learning different stimulus-response mappings. The stimuli and expected responses may be encoded using different firing rates between neurons or different correlation structures, and it would be interesting to observe how the learning performance differed with each of these encodings.

Further research could also extend existing calcium-based plasticity models, such as the one proposed by Graupner and Brunel (2012), to incorporate a dopamine dependence similar to that observed experimentally by Zhang et al. (2009). Such a model should be able to similarly reproduce the experimentally observed reinforcement of neural firing rates that we considered (Fetz, 1969; Fetz and Baker, 1973; Kobayashi et al., 2010). The model proposed by Graupner and Brunel (2012) also contained a bistable, unsupervised learning component in which synapses could permanently rest in either a potentiated or depressed state. It would be interesting to explore how reward learning could coexist this bistability and if this would result in reinforcement that may be either temporary or permanent. This bistability may also allow multiple stimulus-response behaviors to be learnt and maintained within the same network.

5.3 Final Remarks

This thesis has examined a variety of neuronal and synaptic processes and, with consideration of previous work, proposed roles and functions for them within higher-level systems and contexts, such as networks, brain regions, and at a behavioral level. We proposed experiments in which our theoretical models could be investigated. In doing so we hope to have helped bridge the gap that exists in neuroscience between theoretical and experimental neuroscience. It is crucial that theoretical neuroscience always remains always aware of the empirical evidence that has been acquired and experimental neuroscience can benefit hugely from the direction and insight that theoretical work can provide. Specifically, theoretical neuroscience has the potential to, as we have done here, link low-level biological processes with high-level cognitive processes and behaviors. Hopefully our work will inspire and provide a structure for future experimental investigations.

Bibliography

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. Cambridge University Press.

Akam, T. and Kullmann, D. M. (2010). Oscillations and filtering networks support flexible routing of information. *Neuron*, 67(2):308–320.

Appleby, P. A. and Elliott, T. (2007). Multispike interactions in a stochastic model of spike-timing-dependent plasticity. *Neural Comput*, 19(5):1362–1399.

Arnal, L. H., Wyart, V., and Giraud, A.-L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nat Neurosci*, 14(6):797–801.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro-caudal organization of the frontal lobes. *Trends Cogn Sci*, 12(5):193–200.

Bao, S., Chan, V. T., and Merzenich, M. M. (2001). Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature*, 412(6842):79–83.

Barlow, H. B. (1989). Unsupervised learning. *Neural Comput*, 1:295–311.

Bartos, M., Vida, I., and Jonas, P. (2007). Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat Rev Neurosci*, 8(1):45–56.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, 76(4):695–711.

Beckius, G. E., Batra, R., and Oliver, D. L. (1999). Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *J Neurosci*, 19(8):3146–3161.

Beggs, J. and Plenz, D. (2003). Neuronal avalanches in neocortical circuits. *J Neurosci*, 23(35):11167–11177.

Benchenane, K., Tiesinga, P. H., and Battaglia, F. P. (2011). Oscillations in the prefrontal cortex: a gateway to memory and attention. *Curr Opin Neurobiol*, 21(3):475–485.

Bi, G. Q. and Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *J Neurosci*, 18(24):10464–10472.

Bi, G. Q. and Wang, H. X. (2002). Temporal asymmetry in spike timing-dependent synaptic plasticity. *Physiol Behav*, 77(4-5):551–555.

Bliss, T. V. and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol*, 232(2):331–356.

Bohrnstedt, G. W. and Goldberger, A. S. (1969). On the exact covariance of products of random variables. *J Am Stat Assoc*, 64(328):1439–1442.

Börgers, C., Epstein, S., and Kopell, N. J. (2005). Background gamma rhythmicity and attention in cortical local circuits: a computational study. *Proc Natl Acad Sci USA*, 102(19):7002–7007.

Bragin, A., Jandó, G., Nádasdy, Z., Hetke, J., Wise, K., and Buzsáki, G. (1995). Gamma (40-100 Hz) oscillation in the hippocampus of the behaving rat. *J Neurosci*, 15(1 Pt 1):47–60.

Brunel, N. and Wang, X.-J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J Neurophysiol*, 90(1):415–430.

Buffalo, E. A., Fries, P., Landman, R., Liang, H., and Desimone, R. (2010). A backward progression of attentional effects in the ventral stream. *Proc Natl Acad Sci USA*, 107(1):361–365.

Burkitt, A. N. (2001). Balanced neurons: analysis of leaky integrate-and-fire neurons with reversal potentials. *Biol Cybern*, 85(4):247–255.

Burkitt, A. N. (2006a). A review of the integrate-and-fire neuron model: I. Homogeneous synaptic input. *Biol Cybern*, 95(1):1–19.

Burkitt, A. N. (2006b). A review of the integrate-and-fire neuron model: II. Inhomogeneous synaptic input and network properties. *Biol Cybern*, 95(2):97–112.

Burkitt, A. N., Gilson, M., and van Hemmen, J. L. (2007). Spike-timing-dependent plasticity for neurons with recurrent connections. *Biol Cybern*, 96(5):533–546.

Burkitt, A. N., Meffin, H., and Grayden, D. B. (2003). Study of neuronal gain in a conductance-based leaky integrate-and-fire neuron model with balanced excitatory and inhibitory synaptic input. *Biol Cybern*, 89(2):119–125.

Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D., and Miller, E. K. (2012). Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron*, 76(4):838–846.

Buzsáki, G. (2004). Large-scale recording of neuronal ensembles. *Nat Neurosci*, 7(5):446–451.

Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron*, 68(3):362–385.

Buzsáki, G. and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science*, 304(5679):1926–1929.

Canolty, R. T., Ganguly, K., Kennerley, S. W., Cadieu, C. F., Koepsell, K., Wallis, J. D., and Carmena, J. M. (2010). Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proc Natl Acad Sci USA*, 107(40):17356–17361.

Caporale, N. and Dan, Y. (2008). Spike timing-dependent plasticity: a Hebbian learning rule. *Annu Rev Neurosci*, 31:25–46.

Carrillo-Reid, L., Tecuapetla, F., Tapia, D., Hernández-Cruz, A., Galarraga, E., Drucker-Colin, R., and Bargas, J. (2008). Encoding network states by striatal cell assemblies. *J Neurophysiol*, 99(3):1435–1450.

Cassenaer, S. and Laurent, G. (2012). Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature*, 482(7383):47–52.

Câteau, H., Kitano, K., and Fukai, T. (2008). Interplay between a phase response curve and spike-timing-dependent plasticity leading to wireless clustering. *Phys Rev E*, 77(5 Pt 1):051909.

Christodoulou, C. and Cleanthous, A. (2011). Does high firing irregularity enhance learning? *Neural Comput*, 23(3):656–663.

Clopath, C., Büsing, L., Vasilaki, E., and Gerstner, W. (2010). Connectivity reflects coding: a model of voltage-based stdp with homeostasis. *Nat Neurosci*, 13(3):344–352.

Cohen, M. R. and Newsome, W. T. (2008). Context-dependent changes in functional circuitry in visual area MT. *Neuron*, 60(1):162–173.

da Costa, N. M. and Martin, K. A. C. (2010). Whose cortical column would that be? *Front Neuroanat*, 4:16.

Dan, Y. and Poo, M.-M. (2004). Spike timing-dependent plasticity of neural circuits. *Neuron*, 44(1):23–30.

Dan, Y. and Poo, M.-M. (2006). Spike timing-dependent plasticity: From synapse to perception. *Physiol Rev*, 86(3):1033–1048.

de la Rocha, J., Doiron, B., Shea-Brown, E., Josic, K., and Reyes, A. (2007). Correlation between neural spike trains increases with firing rate. *Nature*, 448(7155):802–806.

Douglas, R. J. and Martin, K. A. C. (2007). Recurrent neuronal circuits in the neocortex. *Curr Biol*, 17(13):R496–R500.

Edelman, G. M. (1993). Neural Darwinism: selection and reentrant signaling in higher brain function. *Neuron*, 10(2):115–125.

Engel, A. K. and Fries, P. (2010). Beta-band oscillations–signalling the status quo? *Curr Opin Neurobiol*, 20(2):156–165.

Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci*, 2(10):704–716.

Farries, M. A. and Fairhall, A. L. (2007). Reinforcement learning with modulated spike timing dependent synaptic plasticity. *J Neurophysiol*, 98(6):3648–3665.

Fay, R. R. and Popper, A. N., editors (1992). *The Mammalian Auditory Pathway: Neurophysiology*. Springer-Verlag.

Felleman, D. J. and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1(1):1–47.

Fetz, E. E. (1969). Operant conditioning of cortical unit activity. *Science*, 163(870):955–958.

Fetz, E. E. and Baker, M. A. (1973). Operantly conditioned patterns on precentral unit activity and correlated responses in adjacent cells and contralateral muscles. *J Neurophysiol*, 36(2):179–204.

Florian, R. V. (2007). Reinforcement learning through modulation of spike-timing-dependent synaptic plasticity. *Neural Comput*, 19(6):1468–1502.

Frémaux, N., Sprekeler, H., and Gerstner, W. (2010). Functional requirements for reward-modulated spike-timing-dependent plasticity. *J Neurosci*, 30(40):13326–13337.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci*, 9(10):474–480.

Fries, P., Nikolić, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci*, 30(7):309–316.

Froemke, R. C. and Dan, Y. (2002). Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature*, 416(6879):433–438.

Froemke, R. C., Letzkus, J. J., Kampa, B. M., Hang, G. B., and Stuart, G. J. (2010). Dendritic synapse location and neocortical spike-timing-dependent plasticity. *Front Synaptic Neurosci*, 2:29.

Froemke, R. C., Tsay, I. A., Raad, M., Long, J. D., and Dan, Y. (2006). Contribution of individual spikes in burst-induced long-term synaptic modification. *J Neurophysiol*, 95(3):1620–1629.

Gerstner, W., Kempter, R., van Hemmen, J. L., and Wagner, H. (1996). A neuronal learning rule for sub-millisecond temporal coding. *Nature*, 383(6595):76–81.

Gerstner, W. and Kistler, W. M. (2002). Mathematical formulations of hebbian learning. *Biol Cybern*, 87(5-6):404–415.

Gilson, M., Bürck, M., Burkitt, A. N., and van Hemmen, J. L. (2012). Frequency selectivity emerging from spike-timing-dependent plasticity. *Neural Comput*, 24(9):2251–2279.

Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009a). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks I: Input selectivity–strengthening correlated input pathways. *Biol Cybern*, 101(2):81–102.

Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009b). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks II: Input selectivity–symmetry breaking. *Biol Cybern*, 101(2):103–114.

Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009c). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks III: Partially connected neurons driven by spontaneous activity. *Biol Cybern*, 101(5-6):411–426.

Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2009d). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks IV: Structuring synaptic pathways among recurrent connections. *Biol Cybern*, 101(5-6):427–444.

Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2010a). Emergence of network structure due to spike-timing-dependent plasticity in recurrent neuronal networks V: Self-organization schemes and weight dependence. *Biol Cybern*, 103(5):365–386.

Gilson, M., Burkitt, A. N., Grayden, D. B., Thomas, D. A., and van Hemmen, J. L. (2010b). Representation of input structure in synaptic weights by spike-timing-dependent plasticity. *Phys Rev E*, 82(2 Pt 1):021912.

Gilson, M., Burkitt, A. N., and van Hemmen, J. L. (2010c). STDP in recurrent neuronal networks. *Front Comput Neurosci*, 4(July):23.

Gilson, M. and Fukai, T. (2011). Stability versus neuronal specialization for STDP: long-tail weight distributions solve the dilemma. *PLoS One*, 6(10):e25339.

Goard, M. and Dan, Y. (2009). Basal forebrain activation enhances cortical coding of natural scenes. *Nat Neurosci*, 12(11):1444–1449.

González-Burgos, G., Barrionuevo, G., and Lewis, D. A. (2000). Horizontal synaptic connections in monkey prefrontal cortex: an *in vitro* electrophysiological study. *Cereb Cortex*, 10(1):82–92.

Goodale, M. A. and Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends Neurosci*, 15(1):20–25.

Graupner, M. and Brunel, N. (2012). Calcium-based plasticity model explains sensitivity of synaptic changes to spike pattern, rate, and dendritic location. *Proc Natl Acad Sci USA*, 109(10):3991–3996.

Gray, C. M. and Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci USA*, 86(5):1698–1702.

Gregoriou, G. G., Gotts, S. J., Zhou, H., and Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science*, 324(5931):1207–1210.

Gütig, R., Aharonov, R., Rotter, S., and Sompolinsky, H. (2003). Learning input correlations through nonlinear temporally asymmetric Hebbian plasticity. *J Neurosci*, 23(9):3697–3714.

Haider, B., Duque, A., Hasenstaub, A. R., and McCormick, D. A. (2006). Neocortical network activity *in vivo* is generated through a dynamic balance of excitation and inhibition. *J Neurosci*, 26(17):4535–4545.

Haider, B. and McCormick, D. A. (2009). Rapid neocortical dynamics: cellular and network mechanisms. *Neuron*, 62(2):171–189.

Hammer, M. and Menzel, R. (1995). Learning and memory in the honeybee. *J Neurosci*, 15(3 Pt 1):1617–1630.

Harris, K. D., Csicsvari, J., Hirase, H., Dragoi, G., and Buzsáki, G. (2003). Organization of cell assemblies in the hippocampus. *Nature*, 424(6948):552–556.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Herrero, J. L., Gieselmann, M. A., Sanayei, M., and Thiele, A. (2013). Attention-induced variance and noise correlation reduction in macaque v1 is mediated by nmda receptors. *Neuron*, 78(4):729–739.

Herrero, J. L., Roberts, M. J., Delicato, L. S., Gieselmann, M. A., Dayan, P., and Thiele, A. (2008). Acetylcholine contributes through muscarinic receptors to attentional modulation in v1. *Nature*, 454(7208):1110–1114.

Herzog, M. H., Aberg, K. C., Frémaux, N., Gerstner, W., and Sprekeler, H. (2012). Perceptual learning, roving and the unsupervised bias. *Vision Res*, 61:95–99.

Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117(4):500–544.

Hosaka, R., Araki, O., and Ikeguchi, T. (2008). STDP provides the substrate for igniting synfire chains by spatiotemporal input patterns. *Neural Comput*, 20(2):415–435.

Hubel, D. H. and Wiesel, T. N. (1977). The functional architecture of the macaque visual cortex. *Proc R Soc Lond*, 198(1130):1–59.

Hupé, J. M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by v1, v2 and v3 neurons. *Nature*, 394(6695):784–787.

Huttenlocher, P. R. (1990). Morphometric study of human cerebral cortex development. *Neuropsychologia*, 28(6):517–527.

Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Trans Neural Netw*, 14(6):1569–1572.

Izhikevich, E. M. (2006). Polychronization: computation with spikes. *Neural Comput*, 18(2):245–282.

Izhikevich, E. M. (2007). Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cereb Cortex*, 17(10):2443–2452.

Izhikevich, E. M., Gally, J. A., and Edelman, G. M. (2004). Spike-timing dynamics of neuronal groups. *Cereb Cortex*, 14(8):933–944.

Jensen, O., Kaiser, J., and Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci*, 30(7):317–324.

Jermakowicz, W. J., Chen, X., Khaytin, I., Bonds, A. B., and Casagrande, V. A. (2009). Relationship between spontaneous and evoked spike-time correlations in primate visual cortex. *J Neurophysiol*, 101(5):2279–2289.

Kempter, R., Gerstner, W., and van Hemmen, J. L. (1999). Hebbian learning and spiking neurons. *Phys Rev E*, 59:4498–4514.

Kerr, R. R., Burkitt, A. N., Thomas, D. A., Gilson, M., and Grayden, D. B. (2013). Delay selection by spike-timing-dependent plasticity in recurrent networks of spiking neurons receiving oscillatory inputs. *PLoS Comput Biol*, 9(2):e1002897.

Kobayashi, S., Schultz, W., and Sakagami, M. (2010). Operant conditioning of primate prefrontal neurons. *J Neurophysiol*, 103(4):1843–1855.

Kopell, N., Kramer, M. A., Malerba, P., and Whittington, M. A. (2010). Are different rhythms good for different functions? *Front Hum Neurosci*, 4:187.

Kopell, N., Whittington, M. A., and Kramer, M. A. (2011). Neuronal assembly dynamics in the beta1 frequency range permits short-term memory. *Proc Natl Acad Sci USA*, 108(9):3779–3784.

Kozloski, J. and Cecchi, G. A. (2010). A theory of loop formation and elimination by spike timing-dependent plasticity. *Front Neural Circuits*, 4:7.

Kuhn, A., Aertsen, A., and Rotter, S. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Comput*, 15(1):67–101.

Lapicque, L. (1907). Recherches quantitatives sur l'excitation electrique des nerfs traitee comme une polarization. *Journal de Physiologie et Pathologie General*, 9:620–635.

Larkum, M. (2013). A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci*, 36(3):141–151.

Larkum, M. E., Nevian, T., Sandler, M., Polsky, A., and Schiller, J. (2009). Synaptic integration in tuft dendrites of layer 5 pyramidal neurons: a new unifying principle. *Science*, 325(5941):756–760.

Larkum, M. E., Senn, W., and Lüscher, H.-R. (2004). Top-down dendritic input increases the gain of layer 5 pyramidal neurons. *Cereb Cortex*, 14(10):1059–1070.

Larkum, M. E. and Zhu, J. J. (2002). Signaling of layer 1 and whisker-evoked Ca2+ and Na+ action potentials in distal and terminal dendrites of rat neocortical pyramidal neurons in vitro and in vivo. *J Neurosci*, 22(16):6991–7005.

Larkum, M. E., Zhu, J. J., and Sakmann, B. (1999). A new cellular mechanism for coupling inputs arriving at different cortical layers. *Nature*, 398(6725):338–341.

Larkum, M. E., Zhu, J. J., and Sakmann, B. (2001). Dendritic mechanisms underlying the coupling of the dendritic with the axonal action potential initiation zone of adult rat layer 5 pyramidal neurons. *J Physiol*, 533(Pt 2):447–466.

Lee, S., Sen, K., and Kopell, N. (2009). Cortical gamma rhythms modulate nmdar-mediated spike timing dependent plasticity in a biophysical model. *PLoS Comput Biol*, 5(12):e1000602.

Legenstein, R., Pecevski, D., and Maass, W. (2008). A learning theory for reward-modulated spike-timing-dependent plasticity with application to biofeedback. *PLoS Comput Biol*, 4(10):e1000180.

Leopold and Logothetis (1999). Multistable phenomena: changing views in perception. *Trends Cogn Sci*, 3(7):254–264.

London, M., Roth, A., Beeren, L., Husser, M., and Latham, P. E. (2010). Sensitivity to perturbations *in vivo* implies high noise and suggests rate coding in cortex. *Nature*, 466(7302):123–127.

Lubenov, E. V. and Siapas, A. G. (2008). Decoupling through synchrony in neuronal circuits with propagation delays. *Neuron*, 58(1):118–131.

Markram, H., Lübke, J., Frotscher, M., and Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, 275(5297):213–215.

Martin, K. A. C. (2002). Microcircuits in visual cortex. *Curr Opin Neurobiol*, 12(4):418–425.

Masquelier, T., Hugues, E., Deco, G., and Thorpe, S. J. (2009). Oscillations, phase-of-firing coding, and spike timing-dependent plasticity: An efficient learning scheme. *J Neurosci*, 29(43):13484–13493.

Meffin, H., Besson, J., Burkitt, A., and Grayden, D. (2006). Learning the structure of correlated synaptic subgroups using stable and competitive spike-timing-dependent plasticity. *Physical Review E*, 73(4):384–388.

Meffin, H., Burkitt, A. N., and Grayden, D. B. (2004). An analytical model for the 'large, fluctuating synaptic conductance state' typical of neocortical neurons *in vivo*. *J Comput Neurosci*, 16(2):159–175.

Menzel, R. and Giurfa, M. (2001). Cognitive architecture of a mini-brain: the honeybee. *Trends Cogn Sci*, 5(2):62–71.

Mishra, J., Fellous, J.-M., and Sejnowski, T. J. (2006). Selective attention through phase relationship of excitatory and inhibitory input synchrony in a model cortical neuron. *Neural Netw*, 19(9):1329–1346.

Moreno-Bote, R. and Parga, N. (2006). Auto- and crosscorrelograms for the spike response of leaky integrate-and-fire neurons with slow synapses. *Phys Rev Lett*, 96(2):028101.

Moritz, C. T., Perlmutter, S. I., and Fetz, E. E. (2008). Direct control of paralysed muscles by cortical neurons. *Nature*, 456(7222):639–642.

Morrison, A., Aertsen, A., and Diesmann, M. (2007). Spike-timing-dependent plasticity in balanced random networks. *Neural Comput*, 19(6):1437–1467.

Morrison, A., Diesmann, M., and Gerstner, W. (2008). Phenomenological models of synaptic plasticity based on spike timing. *Biol Cybern*, 98(6):459–478.

Muhammad, R., Wallis, J. D., and Miller, E. K. (2006). A comparison of abstract rules in the prefrontal cortex, premotor cortex, inferior temporal cortex, and striatum. *J Cogn Neurosci*, 18(6):974–989.

Muller, L., Brette, R., and Gutkin, B. (2011). Spike-timing dependent plasticity and feed-forward input oscillations produce precise and invariant spike phase-locking. *Front Comput Neurosci*, 5:45.

Nargeot, R., Baxter, D. A., Patterson, G. W., and Byrne, J. H. (1999). Dopaminergic synapses mediate neuronal changes in an analogue of operant conditioning. *J Neurophysiol*, 81(4):1983–1987.

Palmer, L., Murayama, M., and Larkum, M. (2012). Inhibitory regulation of dendritic activity *in vivo*. *Front Neural Circuits*, 6:26.

Pawlak, V. and Kerr, J. N. D. (2008). Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *J Neurosci*, 28(10):2435–2446.

Pawlak, V., Wickens, J. R., Kirkwood, A., and Kerr, J. N. D. (2010). Timing is not everything: Neuromodulation opens the STDP gate. *Front Synaptic Neurosci*, 2:146.

Pfister, J.-P. and Gerstner, W. (2006). Triplets of spikes in a model of spike timing-dependent plasticity. *J Neurosci*, 26(38):9673–9682.

Pfister, J.-P. and Tass, P. A. (2010). STDP in oscillatory recurrent networks: Theoretical conditions for desynchronization and applications to deep brain stimulation. *Front Comput Neurosci*, 4:22.

Pfister, J.-P., Toyoizumi, T., Barber, D., and Gerstner, W. (2006). Optimal spike-timing-dependent plasticity for precise action potential firing in supervised learning. *Neural Comput*, 18(6):1318–1348.

Reynolds, J. N., Hyland, B. I., and Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, 413(6851):67–70.

Reynolds, J. N. J. and Wickens, J. R. (2002). Dopamine-dependent plasticity of corticostriatal synapses. *Neural Netw*, 15(4-6):507–521.

Roerig, B. and Chen, B. (2002). Relationships of local inhibitory and excitatory circuits to orientation preference maps in ferret visual cortex. *Cereb Cortex*, 12(2):187–198.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol*, 30(4):769–793.

Salgado, H., Köhr, G., and Treviño, M. (2012). Noradrenergic 'tone' determines dichotomous control of cortical spike-timing-dependent plasticity. *Sci Rep*, 2:417.

Salinas, E. (2004). Fast remapping of sensory stimuli onto motor actions on the basis of contextual modulation. *J Neurosci*, 24(5):1113–1118.

Salinas, E. and Sejnowski, T. J. (2001). Gain modulation in the central nervous system: where behavior, neurophysiology, and computation meet. *Neuroscientist*, 7(5):430–440.

Schiller, J., Schiller, Y., Stuart, G., and Sakmann, B. (1997). Calcium action potentials restricted to distal apical dendrites of rat neocortical pyramidal neurons. *J Physiol*, 505 (Pt 3):605–616.

Schroeder, C. E. and Lakatos, P. (2009). The gamma oscillation: master or slave? *Brain Topogr*, 22(1):24–26.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Curr Opin Neurobiol*, 20(2):172–176.

Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, 275(5306):1593–1599.

Senn, W., Schneider, M., and Ruf, B. (2002). Activity-dependent development of axonal and dendritic delays, or, why synaptic transmission should be unreliable. *Neural Comput*, 14(3):583–619.

Seung, H. S. (2003). Learning in spiking neural networks by reinforcement of stochastic synaptic transmission. *Neuron*, 40(6):1063–1073.

Shepherd, G. (2004). *The synaptic organization of the brain*. Oxford University Press.

Shouval, H. Z., Bear, M. F., and Cooper, L. N. (2002). A unified model of NMDA receptor-dependent bidirectional synaptic plasticity. *Proc Natl Acad Sci USA*, 99(16):10831–10836.

Shouval, H. Z., Wang, S. S.-H., and Wittenberg, G. M. (2010). Spike timing dependent plasticity: a consequence of more fundamental learning rules. *Front Comput Neurosci*, 4.

Shu, Y., Hasenstaub, A., and McCormick, D. A. (2003). Turning on and off recurrent balanced cortical activity. *Nature*, 423(6937):288–293.

Sjöström, P. J. and Häusser, M. (2006). A cooperative switch determines the sign of synaptic plasticity in distal dendrites of neocortical pyramidal neurons. *Neuron*, 51(2):227–238.

Sjöström, P. J., Turrigiano, G. G., and Nelson, S. B. (2001). Rate, timing, and cooperativity jointly determine cortical synaptic plasticity. *Neuron*, 32(6):1149–1164.

Song, S. and Abbott, L. F. (2001). Cortical development and remapping through spike timing-dependent plasticity. *Neuron*, 32(2):339–350.

Stepanyants, A., Martinez, L. M., Ferecsk, A. S., and Kisvrday, Z. F. (2009). The fractions of short- and long-range connections in the visual cortex. *Proc Natl Acad Sci USA*, 106(9):3555–3560.

Sugisaki, E., Fukushima, Y., Tsukada, M., and Aihara, T. (2011). Cholinergic modulation on spike timing-dependent plasticity in hippocampal ca1 network. *Neuroscience*, 192:91–101.

Thiele, A. (2009). Optimizing brain processing. *Nat Neurosci*, 12(11):1359–1360.

Tiesinga, P., Fellous, J.-M., and Sejnowski, T. J. (2008). Regulation of spike timing in visual cortical circuits. *Nat Rev Neurosci*, 9(2):97–107.

Tiesinga, P. and Sejnowski, T. J. (2009). Cortical enlightenment: are attentional gamma oscillations driven by ING or PING? *Neuron*, 63(6):727–732.

Tiesinga, P. H., Fellous, J.-M., Salinas, E., Jos, J. V., and Sejnowski, T. J. (2004). Inhibitory synchrony as a mechanism for attentional gain modulation. *J Physiol Paris*, 98(4-6):296–314.

Tsodyks, M. and Sejnowski, T. (1995). Rapid state switching in balanced cortical network models. *Network: Computation in Neural Systems*, 6:111–124.

van Rossum, M., Bi, G., and Turrigiano, G. (2000). Stable Hebbian learning from spike timing-dependent plasticity. *J Neurosci*, 20(23):8812–8821.

van Vreeswijk, C. and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science*, 274(5293):1724–1726.

Vogels, T. P. and Abbott, L. F. (2005). Signal propagation and logic gating in networks of integrate-and-fire neurons. *J Neurosci*, 25(46):10786–10795.

Vogels, T. P. and Abbott, L. F. (2009). Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. *Nat Neurosci*, 12(4):483–491.

Wallis, J. D. and Miller, E. K. (2003). From rule to response: neuronal processes in the premotor and prefrontal cortex. *J Neurophysiol*, 90(3):1790–1806.

Wang, H.-X., Gerkin, R. C., Nauen, D. W., and Bi, G.-Q. (2005). Coactivation and timing-dependent integration of synaptic potentiation and depression. *Nat Neurosci*, 8(2):187–193.

Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol Rev*, 90(3):1195–1268.

Wang, X. J. and Buzsáki, G. (1996). Gamma oscillation by synaptic inhibition in a hippocampal interneuronal network model. *J Neurosci*, 16(20):6402–6413.

Whittington, M. A., Traub, R. D., Kopell, N., Ermentrout, B., and Buhl, E. H. (2000). Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int J Psychophysiol*, 38(3):315–336.

Williams, R. W. and Herrup, K. (1988). The control of neuron number. *Annu Rev Neurosci*, 11:423–453.

Womelsdorf, T., Anton-Erxleben, K., Pieper, F., and Treue, S. (2006). Dynamic shifts of visual receptive fields in cortical area MT by spatial attention. *Nat Neurosci*, 9(9):1156–1160.

Womelsdorf, T., Anton-Erxleben, K., and Treue, S. (2008). Receptive field shift and shrinkage in macaque middle temporal area through attentional gain modulation. *J Neurosci*, 28(36):8934–8944.

Womelsdorf, T. and Fries, P. (2007). The role of neuronal synchronization in selective attention. *Curr Opin Neurobiol*, 17(2):154–160.

Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., and Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science*, 316(5831):1609–1612.

Xie, X. and Seung, H. S. (2004). Learning in neural networks by reinforcement of irregular spiking. *Phys Rev E Stat Nonlin Soft Matter Phys*, 69(4 Pt 1):041909.

Yuste, R., Gutnick, M. J., Saar, D., Delaney, K. R., and Tank, D. W. (1994). Ca2+ accumulations in dendrites of neocortical pyramidal neurons: an apical band and evidence for two functional compartments. *Neuron*, 13(1):23–43.

Zhang, J.-C., Lau, P.-M., and Bi, G.-Q. (2009). Gain in sensitivity and loss in temporal contrast of STDP by dopaminergic modulation at hippocampal synapses. *Proc Natl Acad Sci USA*, 106(31):13028–13033.

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., Goodman, R. R., Emerson, R., Mehta, A. D., Simon, J. Z., Poeppel, D., and Schroeder, C. E. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron*, 77(5):980–991.

Appendix A

Supporting Material for Chapter 2

A.1 Recurrent Correlation

As given by Equation (58) of Gilson et al. (2009d), the (ordinary frequency) Fourier transform, $\mathcal{F}g(f) = \int_{-\infty}^{\infty} g(x)e^{-2\pi ixf}dx$, of the recurrent correlation function for a network with only axonal delays is

$$\mathcal{F}C(f) = Q(f)\Big\{P(f)\big[\mathcal{F}\hat{C}(f) + \mathrm{diag}(\hat{\nu})\big]P^T(-f) + \mathrm{diag}(\nu)\Big\}Q^T(-f) - \mathrm{diag}(\nu), \tag{A.1.1}$$

where

$$\begin{aligned} Q_{jk}(f) &= [I - J_{jk}e^{2\pi i d^{\mathrm{ax}}_{jk} f}\mathcal{F}\epsilon(-f)]^{-1}, \\ P_{jk}(f) &= K_{jk}e^{2\pi i \hat{d}^{\mathrm{ax}}_{jk} f}\mathcal{F}\epsilon(-f). \end{aligned} \tag{A.1.2}$$

It can be considered be to make up of three components

$$\begin{aligned} C(u) &= C_1(u) + C_2(u) + C_3(u) \\ \mathcal{F}C(f) &= \mathcal{F}C_1(f) + \mathcal{F}C_2(f) + \mathcal{F}C_3(f), \end{aligned} \tag{A.1.3}$$

where

$$\begin{aligned} \mathcal{F}C_1(f) &= Q(f)P(f)\mathcal{F}\hat{C}(f)P^T(-f)Q^T(-f), \\ \mathcal{F}C_2(f) &= Q(f)P(f)\mathrm{diag}(\hat{\nu})P^T(-f)Q^T(-f), \\ \mathcal{F}C_3(f) &= Q(f)\mathrm{diag}(\nu)Q^T(-f) - \mathrm{diag}(\nu). \end{aligned} \tag{A.1.4}$$

These components are due to correlations in the inputs, spike triggering effects from the inputs, and recurrent spike triggering effects, respectively. The last two of these are assumed to be negligible to the learning for large numbers of inputs, M, and large numbers of neurons, N, respectively. This is the same assumption made in Gilson et al. (2009d). Because of this only the first correlation component was considered (i.e., $\mathcal{F}\bar{C}(f) \approx \mathcal{F}\bar{C}_1(f)$).

To determine how large a network was sufficient for the spike triggering components to be negligible, simulations with LIF neurons were run to observe the shape of the learned axonal delay distribution after 250s of learning. This is shown in Figure 2.15. For simulations it was decided that the network size would always be the same as the number of inputs (i.e. $N = M$). It can be seen that as the number of neurons (and inputs) increases, the resulting delay distribution becomes a perfect cosine function. We decided that 10 000 neurons (and inputs) was sufficient for simulations in this study.

A.2 Oscillatory Inputs

Input intensity functions are defined for oscillatory inputs as

$$\hat{\lambda}_k(t) = \langle \hat{S}_k(t) \rangle = \hat{\nu}_0 + a\cos\big[2\pi f_m(t + \hat{d}_k)\big], \tag{A.2.1}$$

where $\hat{\nu}_0$ is the mean input rate (in Hz), a is the magnitude of the oscillations (in Hz), f_m is the modulation frequency of the oscillations (in Hz), and $\hat{d}_k$ is the delay of the input (in seconds). Inputs within the same group have the same delay, meaning that they are in phase.

The mean input firing rate of neuron k is

$$\begin{aligned} \hat{\nu}_k &= \frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t') \rangle dt' = \frac{1}{T}\int_{t-T}^{t} \Big\{ \hat{\nu}_0 + a\cos\big[2\pi f_m(t + \hat{d}_k)\big] \Big\} dt' \\ &= \hat{\nu}_0 + \frac{a}{T}\int_{t-T}^{t} \cos\big[2\pi f_m(t + \hat{d}_k)\big] dt' = \hat{\nu}_0. \end{aligned} \tag{A.2.2}$$

The correlation function for a pair of inputs (k and l) is

$$\begin{aligned} \hat{C}_{kl}(t,u) &= \frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\hat{S}_l(t'+u) \rangle dt' - \left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t') \rangle dt'\right)\left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_l(t'+u) \rangle dt'\right) \\ &= \frac{1}{T}\int_{t-T}^{t} \Big\{ \hat{\nu}_0 + a\cos\big[2\pi f_m(t' + \hat{d}_k)\big] \Big\} \Big\{ \hat{\nu}_0 + a\cos\big[2\pi f_m(t' + u + \hat{d}_l)\big] \Big\} dt' - \hat{\nu}_0^2 \\ &= \hat{\nu}_0^2 + \frac{\hat{\nu}_0 a}{T}\int_{t-T}^{t} \Big\{ \cos\big[2\pi f_m t'\big] + \cos\big[2\pi f_m(t' + u + \hat{d}_{\text{lag}})\big] \Big\} dt' \\ &\quad + \frac{a^2}{T}\int_{t-T}^{t} \cos\big[2\pi f_m t'\big]\cos\big[2\pi f_m(t' + u + \hat{d}_{\text{lag}})\big] dt' - \hat{\nu}_0^2 \\ &= \frac{a^2}{2T}\int_{t-T}^{t} \Big\{ \cos\big[2\pi f_m(u + \hat{d}_{\text{lag}})\big] + \cos\big[2\pi f_m(2t' + u + \hat{d}_{\text{lag}})\big] \Big\} dt' \\ &= \frac{a^2}{2}\cos\big[2\pi f_m(u + \hat{d}_{\text{lag}})\big], \end{aligned} \tag{A.2.3}$$

where $\hat{d}_{\text{lag}} = \hat{d}_l - \hat{d}_k$, and the Fourier transform of this is

$$\mathcal{F}\hat{C}_{kl}(f) = \frac{a^2}{4}\Big[\delta(f - f_m) + \delta(f + f_m)\Big] e^{2\pi i \hat{d}_{\text{lag}} f}. \tag{A.2.4}$$

If the inputs are from the same group, then $\hat{d}_{\text{lag}} = 0$, and so

$$\begin{aligned} \hat{C}_{kl}(u) &= \frac{a^2}{2}\cos\big(2\pi f_m u\big), \\ \mathcal{F}\hat{C}_{kl}(f) &= \frac{a^2}{4}\Big[\delta(f - f_m) + \delta(f + f_m)\Big]. \end{aligned} \tag{A.2.5}$$

A.3 Homeostatic Equilibrium in a Recurrent Network

The rate of change of the recurrent axonal delay distribution is

$$\dot{\bar{\mathcal{J}}}(t, d^{\text{ax}}) = \eta\Big[\omega_{\text{in}}\bar{\nu}(t) + \omega_{\text{out}}\bar{\nu}(t) + \tilde{W}\bar{\nu}(t)^2 + \bar{C}^W(t, d^{\text{ax}})\Big], \tag{A.3.1}$$

where $\bar{\nu}(t)$ is the mean firing rate of the recurrent group given by

$$\begin{aligned} \bar{\nu} &= \frac{\nu_0 + N_K\bar{K}\hat{\nu}_0}{1 - \mathcal{F}\epsilon(0)\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(0)} \\ &= \frac{\nu_0 + N_K\bar{K}\hat{\nu}_0}{1 - \tilde{N}_J\int_{d_{\min}}^{d_{\max}}\bar{\mathcal{J}}(x)dx} \\ &= \frac{\nu_0 + N_K\bar{K}\hat{\nu}_0}{1 - N_J\bar{J}}, \end{aligned} \tag{A.3.2}$$

where ν_0 is the spontaneous firing rate of the neurons, $\hat{\nu}_0$ is the mean firing rate of the inputs, and $\bar{J}$ is the mean recurrent weight averaged over all axonal delays.

The stable mean firing rate, $\bar{\nu}^*$, and stable mean weight, $\bar{J}^*$, are found from

$$\begin{aligned} \dot{\bar{J}} &\propto (\omega_{\text{in}} + \omega_{\text{out}})\bar{\nu} + \tilde{W}\bar{\nu}^2 + \int_{d_{\min}}^{d_{\max}}\bar{C}^W(x)dx \\ 0 &= (\omega_{\text{in}} + \omega_{\text{out}})\bar{\nu}^* + \tilde{W}(\bar{\nu}^*)^2 + \bar{C}^W. \end{aligned} \tag{A.3.3}$$

Assuming $\bar{C}^W$ is small and that $\nu_0 = 0$, the solution to this is

$$\bar{\nu}^* = \frac{-(\omega_{\text{in}} + \omega_{\text{out}})}{\tilde{W}}, \tag{A.3.4}$$

and by substituting in Equation (A.3.2) we have that

$$\begin{aligned} 1 - N_J\bar{J}^* = \frac{N_K\bar{K}\hat{\nu}_0}{\bar{\nu}^*} &= \frac{N_K\bar{K}\hat{\nu}_0\tilde{W}}{-(\omega_{\text{in}} + \omega_{\text{out}})} \\ \bar{J}^* &= \frac{1}{N_J}\left(1 + \frac{N_K\bar{K}\hat{\nu}_0\tilde{W}}{\omega_{\text{in}} + \omega_{\text{out}}}\right). \end{aligned} \tag{A.3.5}$$

A.4 Network Response for a Single Group

Given the average response is

$$\bar{\lambda}(t) = \tilde{N}_J \int_{d_{\min}}^{d_{\max}} \bar{\mathcal{J}}(x) \int \epsilon(r-x)\bar{\lambda}(t-r)drdx + N_K \bar{K} \int \epsilon(r-\hat{d})\hat{\lambda}(t-r)dr, \quad \text{(A.4.1)}$$

where $\hat{d}$ is the delay of the inputs. The Fourier transform of this is

$$\mathcal{F}\bar{\lambda}(f) = \tilde{N}_J \mathcal{F}\bar{\mathcal{J}}(f)\mathcal{F}\epsilon(f)\mathcal{F}\bar{\lambda}(f) + N_K \bar{K} e^{-2\pi i \hat{d} f}\mathcal{F}\epsilon(f)\mathcal{F}\hat{\lambda}(f), \quad \text{(A.4.2)}$$

and by rearranging this we get

$$\mathcal{F}\bar{\lambda}(f) = \frac{N_K \bar{K} e^{-2\pi i \hat{d} f}\mathcal{F}\epsilon(f)\mathcal{F}\hat{\lambda}(f)}{1 - \tilde{N}_J \mathcal{F}\bar{\mathcal{J}}(f)\mathcal{F}\epsilon(f)}. \quad \text{(A.4.3)}$$

For oscillatory inputs where $\hat{\lambda}(t) = \hat{\nu}_0 + a\cos(2\pi f_m t)$ and $\mathcal{F}\hat{\lambda}(f) = \hat{\nu}_0\delta(f) + \frac{a}{2}\big[\delta(f - f_m) + \delta(f + f_m)\big]$, the expression for the response of the network becomes

$$\begin{aligned}
\mathcal{F}\bar{\lambda}(f) &= \frac{N_K \bar{K}\mathcal{F}\epsilon(0)\hat{\nu}_0\delta(f)}{1 - \tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(0)\mathcal{F}\epsilon(0)} + \frac{aN_K\bar{K}e^{-2\pi i\hat{d}f}\mathcal{F}\epsilon(f)\big[\delta(f-f_m)+\delta(f+f_m)\big]}{2\big[1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(f)\mathcal{F}\epsilon(f)\big]} \\
&= \frac{N_K\bar{K}\hat{\nu}_0\delta(f)}{1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(0)} + \frac{aN_K\bar{K}e^{-2\pi i\hat{d}f_m}\mathcal{F}\epsilon(f_m)\delta(f-f_m)}{2\big[1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(f_m)\mathcal{F}\epsilon(f_m)\big]} + \frac{aN_K\bar{K}e^{2\pi i\hat{d}f_m}\mathcal{F}\epsilon(-f_m)\delta(f+f_m)}{2\big[1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(-f_m)\mathcal{F}\epsilon(-f_m)\big]} \\
\bar{\lambda}(t) &= \frac{N_K\bar{K}\hat{\nu}_0}{1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(0)} + \frac{aN_K\bar{K}e^{-2\pi i\hat{d}f_m}\mathcal{F}\epsilon(f_m)e^{2\pi i f_m t}}{2\big[1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(f_m)\mathcal{F}\epsilon(f_m)\big]} + \frac{aN_K\bar{K}e^{2\pi i\hat{d}f_m}\mathcal{F}\epsilon(-f_m)e^{-2\pi i f_m t}}{2\big[1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(-f_m)\mathcal{F}\epsilon(-f_m)\big]} \\
&= \bar{\nu} + aN_K\bar{K}\mathrm{Re}\left[\frac{\mathcal{F}\epsilon(f_m)e^{2\pi i f_m(t+\hat{d})}}{1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(f_m)\mathcal{F}\epsilon(f_m)}\right] \\
&= \bar{\nu} + aN_K\bar{K}r_\epsilon(f_m)\mathrm{Re}\left\{\frac{e^{i[2\pi f_m(t-\hat{d})-\phi_\epsilon(f_m)]}}{1-r_\epsilon(f_m)e^{-i\phi_\epsilon(f_m)}\tilde{N}_J r_{\bar{\mathcal{J}}}(f_m)e^{-i\phi_{\bar{\mathcal{J}}}(f_m)}}\right\},
\end{aligned} \quad \text{(A.4.4)}$$

where $\mathcal{F}\epsilon(f) = r_\epsilon(f)e^{-i\phi_\epsilon(f)}$, $\mathcal{F}\mathcal{J}(f) = \int_{d_{\min}}^{d_{\max}}\mathcal{J}(x)e^{-2\pi i f x}dx = r_{\bar{\mathcal{J}}}(f)e^{-i\phi_{\bar{\mathcal{J}}}(f)}$, and $\bar{\nu} = \frac{N_K\bar{K}\hat{\nu}_0}{1-\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}(0)} = \frac{N_K\bar{K}\hat{\nu}_0}{1-N_J J}$. This gives Equation (2.3.22).

A.5 Network Response for Two Groups

For two recurrently connected groups where the within group weights have been depressed each of the group responses are given in Equation (2.3.38). The Fourier transforms of these is

$$\begin{aligned}
\mathcal{F}\bar{\lambda}_1(f) &= \tilde{N}_J\mathcal{F}\bar{\mathcal{J}}_{12}(f)\mathcal{F}\epsilon(f)\mathcal{F}\bar{\lambda}_2(f) + N_K\bar{K}e^{-2\pi i\hat{d}f}\mathcal{F}\epsilon(f)\mathcal{F}\hat{\lambda}_1(f), \\
\mathcal{F}\bar{\lambda}_2(f) &= \tilde{N}_J\mathcal{F}\bar{\mathcal{J}}_{21}(f)\mathcal{F}\epsilon(f)\mathcal{F}\bar{\lambda}_1(f) + N_K\bar{K}e^{-2\pi i\hat{d}f}\mathcal{F}\epsilon(f)\mathcal{F}\hat{\lambda}_2(f),
\end{aligned} \quad \text{(A.5.1)}$$

and by rearranging these we get

$$\begin{aligned}\mathcal{F}\bar{\lambda}_1(f) &= \frac{N_K\bar{K}e^{-2\pi i\hat{d}f}\mathcal{F}\epsilon(f)\Big[\mathcal{F}\epsilon(f)\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}_{12}(f)\mathcal{F}\hat{\lambda}_2(f)+\mathcal{F}\hat{\lambda}_1(f)\Big]}{1-\mathcal{F}\epsilon^2(f)\tilde{N}_J^2\mathcal{F}\bar{\mathcal{J}}_{12}(f)\mathcal{F}\bar{\mathcal{J}}_{21}(f)},\\ \mathcal{F}\bar{\lambda}_2(f) &= \frac{N_K\bar{K}e^{-2\pi i\hat{d}f}\mathcal{F}\epsilon(f)\Big[\mathcal{F}\epsilon(f)\tilde{N}_J\mathcal{F}\bar{\mathcal{J}}_{21}(f)\mathcal{F}\hat{\lambda}_1(f)+\mathcal{F}\hat{\lambda}_2(f)\Big]}{1-\mathcal{F}\epsilon^2(f)\tilde{N}_J^2\mathcal{F}\bar{\mathcal{J}}_{21}(f)\mathcal{F}\bar{\mathcal{J}}_{12}(f)},\end{aligned} \tag{A.5.2}$$

which can be approximated as

$$\begin{aligned}\mathcal{F}\bar{\lambda}_1(f) \approx\ & aN_K\bar{K}e^{-2\pi i\hat{d}f}r_\epsilon(f)e^{-i\phi_\epsilon(f)}\Big[1+r_\epsilon(f)e^{-i\phi_\epsilon(f)}\tilde{N}_Jr_{\bar{\mathcal{J}}_{12}}(f)e^{-i\phi_{\bar{\mathcal{J}}_{12}}(f)}e^{-2\pi i\hat{d}_{\text{lag}}f}\\ &+r_\epsilon^2(f)e^{-2i\phi_\epsilon(f)}\tilde{N}_Jr_{\bar{\mathcal{J}}_{12}}(f)e^{-i\phi_{\bar{\mathcal{J}}_{12}}(f)}\tilde{N}_Jr_{\bar{\mathcal{J}}_{21}}(f)e^{-i\phi_{\bar{\mathcal{J}}_{21}}(f)}\Big]\Big[\delta(f-f_{\text{m}})+\delta(f+f_{\text{m}})\Big],\\ \mathcal{F}\bar{\lambda}_2(f) \approx\ & aN_K\bar{K}e^{-2\pi i(\hat{d}+\hat{d}_{\text{lag}})f}r_\epsilon(f)e^{-i\phi_\epsilon(f)}\Big[1+r_\epsilon(f)e^{-i\phi_\epsilon(f)}\tilde{N}_Jr_{\bar{\mathcal{J}}_{21}}(f)e^{-i\phi_{\bar{\mathcal{J}}_{21}}(f)}e^{2\pi i\hat{d}_{\text{lag}}f}\\ &+r_\epsilon^2(f)e^{-2i\phi_\epsilon(f)}\tilde{N}_Jr_{\bar{\mathcal{J}}_{21}}(f)e^{-i\phi_{\bar{\mathcal{J}}_{21}}(f)}\tilde{N}_Jr_{\bar{\mathcal{J}}_{12}}(f)e^{-i\phi_{\bar{\mathcal{J}}_{12}}(f)}\Big]\Big[\delta(f-f_{\text{m}})+\delta(f+f_{\text{m}})\Big].\end{aligned} \tag{A.5.3}$$

This is then used to give Equation (2.3.39).

A.6 Learning Window and EPSP Kernel

It is assumed that $W(u)$ and $\epsilon(u)$ are given by

$$W(u) = -c_de^{-\frac{u}{\tau_d}}h(u)+c_pe^{\frac{u}{\tau_p}}h(-u), \tag{A.6.1}$$

and

$$\epsilon(u) = \frac{1}{\tau_B-\tau_A}\Big(e^{-\frac{u}{\tau_B}}-e^{-\frac{u}{\tau_A}}\Big)h(u), \tag{A.6.2}$$

where $\tau_B > \tau_A$.

From this, it can be seen that

$$\begin{aligned}\mathcal{F}W(f) &= \frac{c_p\tau_p}{1-2\pi i\tau_pf}-\frac{c_d\tau_d}{1+2\pi i\tau_df}\\ &= \frac{c_p\tau_p}{1+4\pi^2\tau_p^2f^2}-\frac{c_d\tau_d}{1+4\pi^2\tau_d^2f^2}+2\pi if\left(\frac{c_p\tau_p^2}{1+4\pi^2\tau_p^2f^2}+\frac{c_d\tau_d^2}{1+4\pi^2\tau_d^2f^2}\right)\\ &= \frac{c_p\tau_p-c_d\tau_d+4\pi^2f^2\tau_p\tau_d(c_p\tau_d-c_d\tau_p)+2\pi if\big[c_p\tau_p^2+c_d\tau_d^2+4\pi^2\omega^2\tau_p^2\tau_d^2(c_p+c_d)\big]}{(1+4\pi^2\tau_p^2f^2)(1+4\pi^2\tau_d^2f^2)},\end{aligned} \tag{A.6.3}$$

and

$$\begin{aligned}\mathcal{F}\epsilon(f) &= \frac{1}{\tau_B-\tau_A}\left(\frac{\tau_B}{1+2\pi i\tau_B f}-\frac{\tau_A}{1+2\pi i\tau_A f}\right)\\ &= \frac{1}{(\tau_B-\tau_A)}\left[\frac{\tau_B}{1+4\pi^2\tau_B^2f^2}-\frac{\tau_A}{1+4\pi^2\tau_A^2f^2}-2\pi if\Big(\frac{\tau_B^2}{1+4\pi^2\tau_B^2f^2}-\frac{\tau_A^2}{1+4\pi^2\tau_A^2f^2}\Big)\right]\\ &= \frac{(\tau_B-\tau_A)+4\pi^2f^2\tau_B\tau_A(\tau_B^2-\tau_A^2)-2\pi if(\tau_B^2-\tau_A^2)}{(\tau_B-\tau_A)(1+4\pi^2\tau_B^2f^2)(1+4\pi^2\tau_A^2f^2)}.\end{aligned} \tag{A.6.4}$$

From this, it can be seen that $\mathcal{F}W(-f) = (\mathcal{F}W(f))^*$ and $\mathcal{F}\epsilon(-f) = (\mathcal{F}\epsilon(f))^*$.

Writing $\mathcal{F}W(f)$ in polar form gives

$$\mathcal{F}W(f) = r_W(f)e^{i\phi_W(f)}, \tag{A.6.5}$$

where

$$\begin{aligned}r_W(f) &= \frac{\sqrt{\big[c_p\tau_p-c_d\tau_d+4\pi^2f^2\tau_p\tau_d(c_p\tau_d-c_d\tau_p)\big]^2+4\pi^2f^2\big[c_p\tau_p^2+c_d\tau_d^2+4\pi^2f^2\tau_p^2\tau_d^2(c_p+c_d)\big]^2}}{(1+4\pi^2\tau_p^2f^2)(1+4\pi^2\tau_d^2f^2)},\\ \phi_W(f) &= \begin{cases}\arctan(\frac{x}{y}) & \text{for } y>0\\ \frac{\pi}{2} & \text{for } y=0\\ \arctan(\frac{x}{y})+\pi & \text{for } y<0\end{cases},\end{aligned} \tag{A.6.6}$$

where $x = 2\pi f[c_p\tau_p^2+c_d\tau_d^2+4\pi^2f^2\tau_p^2\tau_d^2(c_p+c_d)]$ and $y = c_p\tau_p-c_d\tau_d+4\pi^2f^2\tau_p\tau_d(c_p\tau_d-c_d\tau_p)$. Plots of $r_W(f)$ and $\phi_W(f)$ are shown in Figures 2.4 and 5B, respectively.

Writing $\mathcal{F}\epsilon(f)$ in polar form gives

$$\mathcal{F}\epsilon(f) = r_\epsilon(f)e^{-i\phi_\epsilon(f)}, \tag{A.6.7}$$

where

$$\begin{aligned}r_\epsilon(f) &= \frac{\sqrt{\big[(\tau_B-\tau_A)+4\pi^2f^2\tau_B\tau_A(\tau_B^2-\tau_A^2)\big]^2+4\pi^2f^2(\tau_B^2-\tau_A^2)^2}}{(\tau_B-\tau_A)(1+4\pi^2\tau_B^2f^2)(1+4\pi^2\tau_A^2f^2)}\\ &= \frac{\sqrt{1+8\pi^2f^2\tau_B\tau_A(\tau_B+\tau_A)+16\pi^4f^4\tau_B^2\tau_A^2(\tau_B+\tau_A)^2+4\pi^2f^2(\tau_B+\tau_A)^2}}{(1+4\pi^2\tau_B^2f^2)(1+4\pi^2\tau_A^2f^2)}\\ &= \frac{\sqrt{1+4\pi^2f^2(\tau_B+\tau_A)\big[2\tau_B\tau_A+4\pi^2f^2\tau_B^2\tau_A^2(\tau_B+\tau_A)+\tau_B+\tau_A\big]}}{(1+4\pi^2\tau_B^2f^2)(1+4\pi^2\tau_A^2f^2)},\\ \phi_\epsilon(f) &= \arctan\left[\frac{2\pi f(\tau_B^2-\tau_A^2)}{(\tau_B-\tau_A)+4\pi^2f^2\tau_B\tau_A(\tau_B^2-\tau_A^2)}\right]\\ &= \arctan\left[\frac{2\pi f(\tau_B+\tau_A)}{1+4\pi^2f^2\tau_B\tau_A(\tau_B+\tau_A)}\right].\end{aligned} \tag{A.6.8}$$

It can be seen from this that $\mathcal{F}W(0) = \tilde{W}$ and $\mathcal{F}\epsilon(0) = 1$. Plots of $r_\epsilon(f)$ and $\phi_\epsilon(f)$ are shown in Figure 5A and B, respectively.

A.7 Estimating the Amplitude of a Sum of Cosines

The amplitude of

$$S(x) = \cos(x + a) + \sum_i B_i \cos(x + b_i), \tag{A.7.1}$$

is unchanged under a shift in the x axis. So

$$S(x - a) = \cos(x) + \sum_i B_i \cos(x + b'_i), \tag{A.7.2}$$

where $b'_i = b_i - a$, will have the same amplitude. This can be written as

$$\begin{aligned} S(x-a) &= \cos(x) + \sum_i \left[B_i \cos(b'_i)\cos(x) - B_i \sin(b'_i)\sin(x) \right] \\ &= \left[1 + \sum_i B_i \cos(b'_i) \right] \cos(x) - \left[\sum_i B_i \sin(b'_i) \right] \sin(x) \\ &= P\cos(x) + Q\sin(x), \end{aligned} \tag{A.7.3}$$

where $P = 1 + \sum_i B_i \cos(b'_i)$ and $Q = -\sum_i B_i \sin(b'_i)$. This can be written in the form

$$S(x - a) = W \cos(x + \gamma), \tag{A.7.4}$$

where the amplitude, W, is given by

$$\begin{aligned} W^2 = P^2 + Q^2 &= \left[1 + \sum_i B_i \cos(b'_i) \right]^2 + \left[\sum_i B_i \sin(b'_i) \right]^2 \\ &= 1 + 2\sum_i B_i \cos(b'_i) + 2 \sum_{i,j \neq i} B_i B_j \cos(b'_i)\cos(b'_j) + \sum_i B_i^2 \cos^2(b'_i) \\ &\quad + \sum_i B_i^2 \sin^2(b'_i) + 2 \sum_{i,j \neq i} B_i B_j \sin(b'_i)\sin(b'_j) \\ &= 1 + 2\sum_i B_i \cos(b'_i) + \sum_i B_i^2 + \sum_{i,j \neq i} B_i B_j \left[\cos(b'_i - b'_j) + \cos(b'_i + b'_j) \right] \\ &\quad + \sum_{i,j \neq i} B_i B_j \left[\cos(b'_i - b'_j) - \cos(b'_i + b'_j) \right] \\ &= 1 + 2\sum_i B_i \cos(b'_i) + \sum_i B_i^2 + 2 \sum_{i,j \neq i} B_i B_j \cos(b'_i - b'_j) \\ &= 1 + \sum_i B_i \left[2\cos(b_i - a) + B_i + 2 \sum_{j \neq i} B_j \cos(b_i - b_j) \right] \\ W &= \sqrt{1 + \sum_i B_i \left[2\cos(b_i - a) + B_i + 2 \sum_{j \neq i} B_j \cos(b_i - b_j) \right]}. \end{aligned} \tag{A.7.5}$$

For the case where we have $B_i \propto X^i$, $X < 1$, and it is an infinite sum of cosines, we can estimate the square of the amplitude to the $(k+1)$th order with

$$W^2 = 1 + 2\sum_{i}^{k} B_i \cos(b_i - a) + \sum_{i}^{\lfloor k/2 \rfloor} B_i^2 + 2\sum_{i}^{k}\sum_{j\neq i}^{k-i} B_i B_j \cos(b_i - b_j), \tag{A.7.6}$$

where $\lfloor x \rfloor$ is the floor of x.

A.8 Third-Order Covariance of Oscillatory Inputs

Similar to the second-order input covariance,

$$\hat{C}_{kl}(t,u) = \frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\hat{S}_l(t'+u)\rangle dt' - \left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\rangle dt'\right)\left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_l(t'+u)\rangle dt'\right), \tag{A.8.1}$$

we defined the third-order input covariance as

$$\begin{aligned}
\hat{C}_{klm}(t,u,r) = {} & \frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\hat{S}_l(t'+u)\hat{S}_m(t'+u+r)\rangle dt' \\
& - \left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\rangle dt'\right)\left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_l(t'+u)\hat{S}_m(t'+u+r)\rangle dt'\right) \\
& - \left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_l(t'+u)\rangle dt'\right)\left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\hat{S}_m(t'+u+r)\rangle dt'\right) \\
& - \left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_m(t'+u+r)\rangle dt'\right)\left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\hat{S}_l(t'+u)\rangle dt'\right) \\
& - \left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\rangle dt'\right)\left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_l(t'+u)\rangle dt'\right)\left(\frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_m(t'+u+r)\rangle dt'\right) \\
= {} & \frac{1}{T}\int_{t-T}^{t} \langle \hat{S}_k(t')\hat{S}_l(t'+u)\hat{S}_m(t'+u+r)\rangle dt' \\
& - \hat{\nu}_k \hat{C}_{lm}(t+u,r) - \hat{\nu}_l \hat{C}_{km}(t,u+r) - \hat{\nu}_m \hat{C}_{kl}(t,u) - \hat{\nu}_k\hat{\nu}_l\hat{\nu}_m.
\end{aligned} \tag{A.8.2}$$

So for inputs which are simple realizations identical, sinusoidal intensity functions

given by $\hat{\nu}_0 + a\cos(2\pi f_m t)$, this is

$$\begin{aligned}
\hat{C}_{klm}(t,u,r) &= \frac{1}{T}\int_{t-T}^{t} \big\{\hat{\nu}_0 + a\cos[2\pi f_m t']\big\}\big\{\hat{\nu}_0 + a\cos[2\pi f_m(t'+u)]\big\}\big\{\hat{\nu}_0 + a\cos[2\pi f_m(t'+u+r)]\big\}dt' \\
&\quad - \frac{a^2\hat{\nu}_0}{2}\Big\{\cos[2\pi f_m r] + \cos[2\pi f_m(u+r)] + \cos[2\pi f_m u]\Big\} - \hat{\nu}_0^3 \\
&= \frac{a^3}{T}\int_{t-T}^{t} \cos[2\pi f_m t']\cos[2\pi f_m(t'+u)]\cos[2\pi f_m(t'+u+r)]dt' \\
&= \frac{a^3}{2T}\int_{t-T}^{t} \Big\{\cos[2\pi f_m u] + \cos[2\pi f_m(2t'+u)]\Big\}\cos[2\pi f_m(t'+u+r)]dt' \\
&= \frac{a^3}{4T}\int_{t-T}^{t} \Big\{\cos[2\pi f_m(t'+r)] + \cos[2\pi f_m(t'+2u+r)] + \cos[2\pi f_m(t'-r)] \\
&\quad + \cos[2\pi f_m(3t'+2u+r)]\Big\}dt' \\
&= 0.
\end{aligned} \tag{A.8.3}$$

Appendix B

Supporting Material for Chapter 4

B.1 Derivation of the Learning Equation

Given the general RSTDP model (Equations (4.5.7) and (4.5.8)), a description of how the reward signal depends on the network activity (Equation (4.5.4)), and the details of the neuron models, we derived the specific learning equations that govern how a network evolves in time. We considered only the case where both the firing rates and the correlations between neurons are quasi-stationary (change very slowly over time) and are negligible for large time lags (of the order of the reward delay, d_r). Different definitions for the learning window, and correlations and covariances were used by Legenstein et al. (2008); Gilson et al. (2009b); the time lags of the functions are reversed. We used the definitions from the latter.

We first considered the learning due to an arbitrary reward signal with fixed mean. This is given by

$$\begin{aligned}\dot{K}_{ik} = \eta\Bigg\{ \int_0^\infty g_c(s)\Big[p_+ f_+(K_{ik}) Y_{ik}^{W_+}(t,s) + p_- f_-(K_{ik}) Y_{ik}^{W_-}(t,s)\Big] ds \\ + (p_+\bar{y} + q_+) f_+(K_{ik}) D_{ik}^{W_+}(t) + (p_-\bar{y} + q_-) f_-(K_{ik}) D_{ik}^{W_-}(t)\Bigg\},\end{aligned} \tag{B.1.1}$$

where $\hat{d}$ is the axonal delay from the inputs, $D_{ik}^{\psi}(t) = \int_{-\infty}^{\infty} \psi(u) D_{ik}(t, u-\hat{d}) du$, $D_{ik}(t,u)$ is the neuron-input cross-correlation (i.e., $D_{ik}(t,u) = F_{ik}(u) + \nu_i \hat{\nu}_k$), $Y_{ik}^{\psi}(t,s) = \int_{-\infty}^{\infty} \psi(u) Y_{ik}(t,s,u-\hat{d}) du$, and $Y_{ik}(t,s,u)$ is given by

$$Y_{ik}(t,s,u) = \Big\langle \mathbb{E}\big[\Delta y(t) S_i(t-s) \hat{S}_k(t-s+u)\big] \Big\rangle_T, \tag{B.1.2}$$

where $\Delta y(t) = y(t) - \bar{y}$.

For the operant conditioning experiment, we substituted Equation (4.5.4) into Equa-

tion (B.1.2) to give

$$Y_{ik}(t,s,u) = \sum_j \gamma_j \int_0^\infty g_r(r)\mathbb{E}\Big[S_j(t-d_r-r)S_i(t-s)\hat{S}_k(t-s+u)\Big]dr. \tag{B.1.3}$$

Using the results of Bohrnstedt and Goldberger (1969), we found that

$$\begin{aligned}\mathbb{E}\big[S_j(t_1)S_i(t_2)\hat{S}_k(t_3)\big] &= \mathbb{E}\big[S_j(t_1)\big]\mathbb{E}\big[S_i(t_2)\big]\mathbb{E}\big[\hat{S}_k(t_3)\big] + \mathbb{E}\big[S_j(t_1)\big]\mathbb{C}\big[S_i(t_2),\hat{S}_k(t_3)\big] \\ &\quad + \mathbb{E}\big[S_i(t_2)\big]\mathbb{C}\big[S_j(t_1),\hat{S}_k(t_3)\big] + \mathbb{E}\big[\hat{S}_k(t_3)\Big]\mathbb{C}\big[S_j(t_1),S_i(t_2)\big] \\ &\quad + \mathbb{C}\big[S_j(t_1),S_i(t_2),\hat{S}_k(t_3)\big],\end{aligned} \tag{B.1.4}$$

where $\mathbb{C}[A,B]$ is the joint cumulant of random variables A and B (the covariance), and $\mathbb{C}[A,B,C]$ is the joint cumulant of random variables A, B, and C.

We have assumed that only pairwise correlations exist (i.e., for $i \neq j$, $\mathbb{C}\big[S_j(t_1),S_i(t_2),\hat{S}_k(t_3)\big] = 0$). However, when $i = j$, we found that

$$\mathbb{C}\big[S_i(t_1),S_i(t_2),\hat{S}_k(t_3)\big] = \nu_i^{-1}\bar{C}_{ii}(t_2-t_1)\bar{F}_{ik}(t_3-t_1). \tag{B.1.5}$$

Therefore, using the covariances given in Equation (4.5.13), we derived that

$$\begin{aligned}\mathbb{E}\big[S_i(t_1)S_i(t_2)\hat{S}_k(t_3)\big] &= \nu_i\nu_i\hat{\nu} + \nu_i\bar{F}_{ik}(t_3-t_2) + \nu_i\bar{F}_{ik}(t_3-t_1) \\ &\quad + \hat{\nu}\bar{C}_{ii}(t_2-t_1) + \nu_i^{-1}\bar{C}_{ii}(t_2-t_1)\bar{F}_{ik}(t_3-t_1) \\ &= \nu_i\nu_i\hat{\nu} + c_i\nu_i\nu_i\epsilon(t_2-t_3+\hat{d}) + c_i\nu_i\nu_i\epsilon(t_1-t_3+\hat{d}) \\ &\quad + a\hat{\nu}\nu_i\delta(t_2-t_1) + ac_i\nu_i\delta(t_2-t_1)\epsilon(t_1-t_3+\hat{d}),\end{aligned} \tag{B.1.6}$$

and

$$\begin{aligned}\mathbb{E}\big[S_j(t_1)S_i(t_2)\hat{S}_k(t_3)\big] &= \nu_j\nu_i\hat{\nu} + \nu_j\bar{F}_{ik}(t_3-t_2) + \nu_i\bar{F}_{jk}(t_3-t_1) + \hat{\nu}\bar{C}_{ij}(t_2-t_1) \\ &= \nu_j\nu_i\hat{\nu} + c_i\nu_j\nu_i\epsilon(t_2-t_3+\hat{d}).\end{aligned} \tag{B.1.7}$$

Substituting Equations (B.1.6) and (B.1.7) into Equation (B.1.3), we derived an expression for $Y_{ik}(t,s,u)$ as

$$Y_{ik}(t,s,u) = \gamma_i\nu_i\Big\{ag_r(s-d_r)\big[\hat{\nu} + c_i\epsilon(-u+\hat{d})\big] + c_i\nu_i g_r(s-d_r-u+\hat{d})\Big\}, \tag{B.1.8}$$

where $\epsilon(t)$ is the excitatory post-synaptic potential (EPSP) and c_i is the mean correlation strength between neuron i and its inputs. This also includes a correction factor, $a = \int_{-\mathcal{U}}^{\mathcal{U}} \bar{C}_{ii}(t,u)du$, where $\bar{C}_{ii}(t,u)$ is the mean auto-covariance function of neuron i at time t, and $\mathcal{U}$ is a period of time longer than the time scale of the learning window but shorter than the time scale of the reward and eligibility kernels. For Poisson neurons with constant inputs, the auto-covariance is a simple delta-function and so $a = 1$. However, for LIF neurons, this is not necessarily the case.

We substituted Equation (B.1.8) into Equation (B.1.1) and obtained the rates of change of the mean weights into the reinforced, surround, and control neurons as, respectively,

$$\begin{aligned}
\dot{\bar{K}}_R &= \eta\bar{\nu}_R\Big\{\big[p_+(\bar{y}+a\gamma\eta_r)+q_+\big]f_+(\bar{K}_R)\big[\tilde{W}_+\hat{\nu}+\bar{c}_R\theta\big]+\big[p_-(\bar{y}+a\gamma\eta_r)+q_-\big]f_-(\bar{K}_R)\tilde{W}_-\hat{\nu}\\
&\qquad+\gamma\eta_r\bar{\nu}_R\bar{c}_R\big[p_+f_+(\bar{K}_R)\tilde{W}_++p_-f_-(\bar{K}_R)\tilde{W}_-\big]\Big\},\\
\dot{\bar{K}}_S &= \eta\bar{\nu}_S\Big\{\big[p_+\bar{y}+q_+\big]f_+(\bar{K}_S)\big[\tilde{W}_+\hat{\nu}+\bar{c}_S\theta\big]+\big[p_-\bar{y}+q_-\big]f_-(\bar{K}_S)\tilde{W}_-\hat{\nu}\Big\},\\
\dot{\bar{K}}_C &= \eta\bar{\nu}_C\Big\{\big[p_+y_0+q_+\big]f_+(\bar{K}_C)\big[\tilde{W}_+\hat{\nu}+\bar{c}_C\theta\big]+\big[p_-y_0+q_-\big]f_-(\bar{K}_C)\tilde{W}_-\hat{\nu}\Big\},
\end{aligned} \tag{B.1.9}$$

where $\bar{y}$ is the mean value of the reward signal, $\tilde{W}_+$ and $\tilde{W}_-$ are the integrals over the LTP and LTD parts of the learning window, respectively, $\eta_r = \int_0^\infty g_c(s)g_r(s-d_r)ds$, $\theta = (W_+ * \epsilon)(0)$, and $\bar{\nu}_R$ and $\bar{c}_R$, $\bar{\nu}_S$ and $\bar{c}_S$, and $\bar{\nu}_C$ and $\bar{c}_C$ are the mean firing rates and mean spike triggered correlations of the reinforced, surround, and control neurons, respectively. For the specific functions and kernels used in this study, $\eta_r = 0.76$ and $\theta = 0.76$. For the small covariances due to the spike triggering effect, the third term for the evolution of $\bar{K}_R$ in Equation (B.1.9) can be neglected and this gives Equation (4.3.1).

B.2 Resulting Mean Input Weights

Using logLTD weight dependence (Equation (4.5.11)) and uncorrelated input spike trains, the equations describing the stable equilibria of the mean synaptic weights into the reinforced, surround, and control neurons, $\bar{K}_R$, $\bar{K}_S$, and $\bar{K}_C$, respectively, are

$$\begin{aligned}
\frac{\log(1+\alpha\bar{K}_R^*/K_0)}{\log(1+\alpha)} &= \frac{\gamma\eta_r\bar{\nu}_R^*\bar{c}_R^*p_+\tilde{W}_+ + \big[p_+(\bar{y}+a\gamma\eta_r)+q_+\big]\big[\tilde{W}_+\hat{\nu}+\bar{c}_R^*\theta\big]}{\gamma\eta_r\bar{\nu}_R^*\bar{c}_R^*p_-\tilde{W}_- + \big[p_-(\bar{y}+a\gamma\eta_r)+q_-\big]\tilde{W}_-\hat{\nu}},\\
\frac{\log(1+\alpha\bar{K}_S^*/K_0)}{\log(1+\alpha)} &= \frac{\big[p_+\bar{y}+q_+\big]\big[\tilde{W}_+\hat{\nu}+\bar{c}_S^*\theta\big]}{\big[p_-\bar{y}+q_-\big]\tilde{W}_-\hat{\nu}},\\
\frac{\log(1+\alpha\bar{K}_C^*/K_0)}{\log(1+\alpha)} &= \frac{\big[p_+y_0+q_+\big]\big[\tilde{W}_+\hat{\nu}+\bar{c}_C^*\theta\big]}{\big[p_-y_0+q_-\big]\tilde{W}_-\hat{\nu}},
\end{aligned} \tag{B.2.1}$$

where K_0 and α are the parameters of the weight dependence function.

Using additive weight dependence (Equation (4.5.12)) with rate-based learning terms, we have the equilibria

$$\begin{aligned}
\frac{\bar{\nu}_R^*}{\hat{\nu}} &= \frac{-\omega_{\text{in}}}{\big[p_+(\bar{y}+a\gamma\eta_r)+q_+\big]\big[\tilde{W}_+\hat{\nu}+\bar{c}_R\theta\big]+\big[p_-(\bar{y}+a\gamma\eta_r)+q_-\big]\hat{\nu}\tilde{W}_-+\omega_{\text{out}}},\\
\frac{\bar{\nu}_S^*}{\hat{\nu}} &= \frac{-\omega_{\text{in}}}{[p_+\bar{y}+q_+][\hat{\nu}\tilde{W}_++\bar{c}_S\theta]+[p_-\bar{y}+q_-]\hat{\nu}\tilde{W}_-+\omega_{\text{out}}},\\
\frac{\bar{\nu}_C^*}{\hat{\nu}} &= \frac{-\omega_{\text{in}}}{[p_+y_0+q_+][\hat{\nu}\tilde{W}_++\bar{c}_C\theta]+[p_-y_0+q_-]\hat{\nu}\tilde{W}_-+\omega_{\text{out}}},
\end{aligned} \tag{B.2.2}$$

where the rate-based learning terms, ω_{in} and ω_{out}, give the changes to the synaptic strength for pre- and post-synaptic spikes, respectively.