



**UCL**

# Computation in Balanced Networks

David G.T. Barrett

B.A. (Mod) in Theoretical Physics, M.Sc.

Gatsby Computational Neuroscience Unit  
University College London

THESIS

Submitted for the degree of  
Doctor of Philosophy, University College London

2012

## Declaration

I, David Barrett, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

*David Barrett*

David Barrett

## Abstract

In the cortex, neural activity is noisy, irregular and asynchronous – a consequence of dynamically balancing excitatory and inhibitory input to neurons. Despite this noisy balancing, the brain is capable of performing a vast array of incredibly difficult computations. This is mysterious, because noise and irregularity are usually associated with poor performance.

We ask, how can the cortex compute in a noisy background? The observation of orientation tuning in the visual cortex suggests that structured connectivity is important. We propose a unifying model of cortical connectivity in which weak structured connectivity is embedded in strong random background connectivity. This connectivity can simultaneously produce orientation tuning and irregular, asynchronous dynamics. We find that structure can boost computational performance, by amplifying orientation tuning.

We then ask; why is cortical activity noisy? Surprisingly, we find that balanced network noise can also improve computational performance, by increasing the computational operating range of the cortex. The mechanism is simple; noise allows very large signals to become available for computation, despite the small operating range of individual neurons. However, this improvement comes at a price; for small signals, balanced network noise degrades performance. This exemplifies a performance-stability trade-off. As a corollary, we find that the contrast invariance of orientation tuned cells in the visual cortex is a consequence of this computational stability.

Finally, we ask; does noise co-variability impair computation? It is known that correlated variability can degrade the computational performance of a network, especially if many neurons are strongly co-variant. We find that correlations in balanced networks are weak, but not weak enough to be ignored in computation because they affect decoding.

Together, these results constitute an important link between neural computation and dynamics, opening the door to a reconciliation between conflicting theories of randomness and structure.

# Contents

Declaration . . . . .	2
Abstract . . . . .	3
Contents . . . . .	3
List of figures . . . . .	7
Acknowledgments . . . . .	9
<b>1 Introduction</b>	<b>12</b>
1.1 Computation in balanced networks: why do we care? . . . . .	12
1.2 The random versus structure debate . . . . .	13
1.3 Our approach . . . . .	14
<b>2 Literature Review</b>	<b>17</b>
2.1 Introduction . . . . .	17
2.2 Balanced networks . . . . .	18
2.2.1 Irregular, asynchronous spike trains in the cortex . . . . .	18
2.2.2 Where is the irregularity coming from? . . . . .	20
2.2.3 The synaptic background noise problem . . . . .	21
2.2.4 The standard model of cortical activity . . . . .	24
2.3 Orientation tuning . . . . .	28
2.3.1 Orientation tuning and contrast invariance . . . . .	28
2.3.2 Where does orientation tuning come from? . . . . .	30
2.3.3 Augmenting the standard model with structured connectivity . . . . .	32
2.4 Computation . . . . .	34
2.4.1 Orientation selectivity . . . . .	35
2.4.2 The role of correlations in computation . . . . .	37
2.4.3 The role of noise in computation . . . . .	39

<b>I</b>	<b>Dynamics</b>	<b>45</b>
<b>3</b>	<b>Orientation Tuning in Balanced Networks</b>	<b>47</b>
3.1	Introduction . . . . .	47
3.2	Balanced network model . . . . .	48
3.2.1	Single neuron dynamics . . . . .	50
3.2.2	Network connectivity . . . . .	52
3.2.3	Balance conditions . . . . .	54
3.3	Results . . . . .	55
3.3.1	Network simulations . . . . .	55
3.3.2	Balanced, irregular, asynchronous, spiking dynamics . . . . .	56
3.3.3	Orientation tuning . . . . .	62
3.3.4	Contrast invariance . . . . .	64
3.4	Discussion . . . . .	65
<b>4</b>	<b>Spike-train Correlations in Balanced Networks</b>	<b>69</b>
4.1	Introduction . . . . .	69
4.2	Correlation analysis . . . . .	70
4.2.1	Firing rate equations . . . . .	72
4.2.2	Correlation equations . . . . .	75
4.2.3	Solving the correlation equations . . . . .	79
4.3	Results . . . . .	81
4.3.1	The origin of correlations . . . . .	81
4.3.2	Simulations and predictions . . . . .	81
4.3.3	The contribution of structured connectivity and background connectivity to correlations . . . . .	84
4.3.4	Firing rate, preferred orientation and correlations . . . . .	88
4.3.5	Correlation scaling . . . . .	88
4.4	Discussion . . . . .	91
4.5	Supplementary Methods . . . . .	94
4.5.1	Master equation . . . . .	94
4.5.2	Firing rate equations . . . . .	95
4.5.3	Covariance derivation . . . . .	99

4.5.4	Correlation scaling . . . . .	104
<b>II</b>	<b>Computation</b>	<b>110</b>
<b>5</b>	<b>Information in Balanced Networks</b>	<b>112</b>
5.1	Introduction . . . . .	112
5.2	Orientation selectivity in balanced networks . . . . .	113
5.3	Information analysis . . . . .	115
5.3.1	Linear Fisher Information . . . . .	115
5.3.2	Tuning curve slopes . . . . .	117
5.3.3	Input information and output information . . . . .	118
5.4	Results . . . . .	119
5.4.1	The origin of information . . . . .	119
5.4.2	Simulations and predictions . . . . .	120
5.4.3	The contribution of structured connectivity to information . . . .	121
5.4.4	The contribution of background connectivity to information . . .	123
5.4.5	Correlations and information . . . . .	128
5.5	Discussion . . . . .	129
<b>6</b>	<b>Perspectives</b>	<b>134</b>
6.1	Open questions . . . . .	135
	<b>Bibliography</b>	<b>138</b>

# List of figures

2.1	Irregular spike trains . . . . .	19
2.2	Excitatory-Inhibitory balance . . . . .	27
2.3	Orientation tuning and contrast invariance . . . . .	29
2.4	Correlations and computation . . . . .	40
2.5	Stochastic resonance . . . . .	41
3.1	Network model . . . . .	49
3.2	Network input and structured connectivity . . . . .	52
3.3	Balance of excitation and inhibition . . . . .	57
3.4	Irregular, asynchronous dynamics . . . . .	58
3.5	Orientation tuning in balanced networks . . . . .	63
3.6	Contrast invariance . . . . .	64
4.1	Gaussian input . . . . .	74
4.2	Firing rate prediction, measurement and histogram . . . . .	82
4.3	Correlation distribution, prediction and measurement . . . . .	83
4.4	Correlation prediction accuracy . . . . .	84
4.5	Structured connectivity and correlations . . . . .	85
4.6	Background connectivity and correlations . . . . .	86
4.7	Firing rate, preferred orientation and correlations . . . . .	88
4.8	Correlation scaling . . . . .	90
4.9	Neuron model . . . . .	94
5.1	Orientation selectivity . . . . .	114
5.2	Predicted information . . . . .	121
5.3	Structure and information . . . . .	122

5.4	Contrast and information . . . . .	124
5.5	Background and information . . . . .	125
5.6	Correlations and information . . . . .	129

To my wife,  
Susie.

## Acknowledgements

Throughout my research, I have been extremely fortunate to have had the support of wise advisors, true friends and a wonderful family. My parents have been especially supportive. From the beginning, they encouraged me in my work, through the problems, to celebrating the successes.

This research would not have been possible without the insightful advice of my supervisor, Peter Latham. His insistence on the highest standards in neuroscience and his extensive mathematical experience enabled me to tackle some of the most important questions in theoretical neuroscience. I was also very lucky to have had Peter Dayan as a second-supervisor. His encyclopedic knowledge and sincere advice was invaluable.

It was a great privilege to work at the Gatsby Unit. I will miss the daily tea talks, the rigorous debate and free-flow of ideas and theories. In particular, many discussions with Jeff Beck, Alex Lerchner and Yasser Roudi had a great influence on my work.

I made many great friends during my PhD. In particular, it was a pleasure to share an office with Phillipp Hehrmann, whose research on pitch perception extended to regular music nights, with Ulrik and Gabi. Ross Williamson was a wonderful friend and post-conference road tripper, along with Loic, Lars, Marius, David and Ritwik. Many other friends provided welcome distraction throughout my work, including Amit, Andriy, Agnieszka, Biljana, Demis, Jakob, Charles, Lloyd, Jan, Arthur, Kai, Vinayak, Reza, Dilan, Rich, Shane and Hermann.

Funding for this work was provided predominantly by the Wellcome Trust, who were particularly generous and flexible. I am very grateful to David Attwell and the Wellcome Trust committee for their support and advice. The Gatsby charitable foundation also provided substantial financial support and Cell Press provided funding for travel to the Computational and Systems Neuroscience conference.

Above all, I want to thank Susie for her never-ending encouragement and support. A PhD can be a difficult and lonely journey, but thanks to her companionship, it has been overwhelmingly enjoyable and rewarding.



# Chapter 1

## Introduction

A professor cycling to work, and a cave man wandering through a forest have much in common; their survival depends on their ability to process the images around them. A speeding taxi or an encroaching pack of wolves must be detected so that appropriate evasive manoeuvres can be rapidly deployed.

Our natural ability at these tasks belies their awesome difficulty. We cannot yet build, or even imagine how to build computers capable of matching human computational power. This problem is compounded by the observation that neural activity in the cortex is noisy, irregular and asynchronous – a consequence of an ongoing balance of excitatory and inhibitory neural input. This is mysterious because noise and irregularity are usually associated with poor performance, not with astounding ability.

Our goal is to resolve this problem - to understand behaviour, in the form of computation, in terms of nervous system activity as characterised by balanced network dynamics.

### 1.1 Computation in balanced networks: why do we care?

In the last century, the theory of computation emerged as one of most important frameworks for describing animal behaviour and brain activity. *Computations* transform ugly incoming signals into desirable output. Efficient computation of sensory signals is important, lest we get devoured by wolves on the way to work. Indeed, there is mounting evidence that humans are capable of computing close to the theoretical optimum, suggesting that much of the cortex has actually evolved to compute optimally.

Large, interconnected networks of neurons provide the substrate for this behaviour. In recent years, *balanced network theory* has emerged as the standard model of cortical network dynamics. According to this theory, neurons in the cortex receive strong excitatory and inhibitory synaptic input which must balance to avoid pathological cortical states. However, the balancing is not exact, so neural input and output is noisy. This dynamic explains a broad spectrum of neural network activity and cortical phenomena, such as the irregular asynchronous spike trains observed in the cortex.

If we can understand computation in balanced networks, we will have achieved our goal of establishing a link between the extremes of behaviour on one hand and cellular activity on the other. Our task is difficult because the mathematical language used to describe optimal computation is very different from the language used to describe balanced network dynamics. Also, both theories impose different constraints on cortical input-output transformations. Most neural computation studies simply ignore balanced network dynamics. In this work, we go to great lengths to ensure that computation and balanced network dynamics are mutually consistent.

There is more to neural function than computation, and there is more to neural networks than dynamic balance. Indeed, both of these frameworks completely fail to describe a significant fraction of nervous system function and dynamics. For example, in some parts of the peripheral nervous system, neural activity is highly regular and temporally stereotypical, inconsistent with the irregularity predicted in balanced network theory. The peripheral nervous system is largely involved in information transmission from the sensory organs, and control of muscles. The language of communication theory and control theory are more appropriate for describing these functions.

## 1.2 The random versus structure debate

There is currently a major debate underway in theoretical neuroscience between proponents of two conflicting theories of balanced network computation. Traditionalists maintain that the brain is an exquisitely structured organ, precisely wired to perform carefully constructed computations. The opposing view holds that the brain is predominantly random, with random connections between neurons, performing random computations.

We propose a resolution to this conflict using a combination of weak structured

connectivity embedded in strong background. This connectivity has been used before to successfully explain memory in balanced networks. Using this model, we address some long-standing questions in theoretical neuroscience. First of all, we ask, can weak structured connectivity embedded in strong random background connectivity improve computational performance? If not, then perhaps structured connectivity is the same size as background connectivity. We also ask, why is the cortex noisy? Is it an evolutionary mistake or does cortical noise have some role in computation. Finally, we ask, how correlated is cortical noise, and do these correlations effect computation?

### 1.3 Our approach

All of these questions are important, so a huge body of experimental and theoretical knowledge already exists. In chapter 2, we review the most important literature, with a particular emphasis on the visual cortex, which has received most attention historically.

In chapter 3, we demonstrate that our balanced network model can produce orientation tuned, irregular, asynchronous spiking activity, just like the activity of simple cells in the visual cortex. We find that this orientation tuning is contrast invariant, because synaptic background noise effectively linearises single neuron dynamics.

In chapter 4, we derive mathematical expressions that relate firing rates and spike train correlations to network connectivity and input. This analysis is difficult because a neural network is an extremely complex dynamical system. However, using methods from statistical physics and dynamical systems theory we can understand how these are related. In particular, we find that the contribution of structured connectivity to correlations is about the same size as the contribution of background connectivity to correlations. We also find that correlations are very weak, in agreement with recent correlation measurements in the visual cortex.

Finally, in chapter 5, we return to our original computation questions. We study orientation selectivity, in which the orientation of an edge must be detected. This is an important computational building block for important survival-dependent computations such as object recognition. We quantify the ability of a balanced network to perform orientation selectivity by calculating linear Fisher information. When information is high, computational performance can be good, and when information is low, computational performance is poor.

Our most surprising result is that the noise produced in dynamic balancing can improve computational performance. It achieves this by maintaining the network in a highly informative state for a wide range of image contrasts. Furthermore, we find that the contribution of background connectivity does not preclude structured connectivity from contributing. In fact, structured connectivity can improve computational performance substantially. Finally, we find that although correlations are weak, their contribution to orientation selectivity cannot be ignored.

These results, though important, do not represent our primary contribution. Rather, the equations from which we have derived these results represent our primary result. These equations provide a well-defined mathematical relationship between computational performance, balanced network connectivity, network input, firing rates and spike train correlations. They provide insight into the origin of correlations, the origin of information and the precise relationship between these quantities. Furthermore, they provide a platform for future investigation into computation in balanced networks.



## Chapter 2

# Literature Review

### 2.1 Introduction

One of the central goals of neuroscience is to understand how animal behaviour emerges from nervous system activity (Dayan and Abbott, 2001). The field of *computational neuroscience* is largely dedicated to solving this problem (Sejnowski et al., 1988; Schwartz, 1990). In this section, we will review some of the most important results in this subject.

We begin by reviewing the dynamics of *balanced networks*. Experimental recordings indicate that most of the neural activity observed in the cortex is the result of a dynamic balance of excitatory and inhibitory input to neurons (Shu and Hasenstaub, 2003; Haider et al., 2006). This balancing is noisy so cortical activity is highly irregular (van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998). One problem with this is that cortical activity can also have structure, such as *orientation tuning*, in which cortical activity is tuned to the orientation of edge-like visual stimuli (Hubel and Wiesel, 1962). We will review the theories that attempt to resolve this problem.

Another problem is that many theories of neural computation are inconsistent with balanced network dynamics. One reason for this is that the irregularity produced by balanced networks appears to harm computation. However, this irregularity is not necessarily harmful (Destexhe and Contreras, 2006). Indeed, it may be the source of the brain's vast computational power. We will review computational theories that attempt to discover the computational role of balanced network dynamics, spike train irregularity and co-variability.

## 2.2 Balanced networks

Balanced network theory provides an explanation for a huge array of observed cortical phenomena, from the irregular asynchronous activity that occurs during awake behaviour to the up-down activity that occurs during sleep and anaesthesia (Lerchner and Latham, 2011). In recent years, this theory has emerged as the standard model of cortical network dynamics. In this section we review the principles of balanced network theory and the experiments that led to its development.

### 2.2.1 Irregular, asynchronous spike trains in the cortex

Spiking activity in the cortex is highly *irregular*, both within the spike trains of individual neurons and across neurons in a population (Burns and Webb, 1976; Softky and Koch, 1993; Bair et al., 1994). This irregularity can be characterized by counting the number of spikes emitted by a neuron during a fixed time window. Such *spike counts* have a broad range of values, across trials, with a distribution consistent with a truncated Gaussian (Gershon et al., 1998). Spike train irregularity can be quantified by calculating the *Fano factor*, given by the ratio of the spike count variance to the average spike count. Typically, the Fano factor is about 1 in the cortex (Softky and Koch, 1993), indicating a high degree of irregularity (Fig. 2.1).

Another measure of spike train irregularity is the *coefficient of variation*  $C_v$  of inter-spike intervals (Softky and Koch, 1993). The coefficient of variation is the ratio of the inter-spike interval standard deviation to the inter-spike interval mean. Typically,  $C_v \sim 1 - 1.5$  in the cortex (Shadlen and Newsome, 1998), again indicating a high degree of irregularity. Inter-spike intervals have a distribution consistent with a log-normal distribution (Gershon et al., 1998) (Fig. 2.1).

This irregular cortical activity suggests that spike trains in the brain result from some stochastic point process, such as a Poisson process with a brief refractory period (Bair et al., 1994). In other words, neurons seem to spike randomly with some temporally fluctuating firing rate. However, we must be careful to distinguish between noise, which is truly random and irregular fluctuations, which seem random but may carry some useful signal.

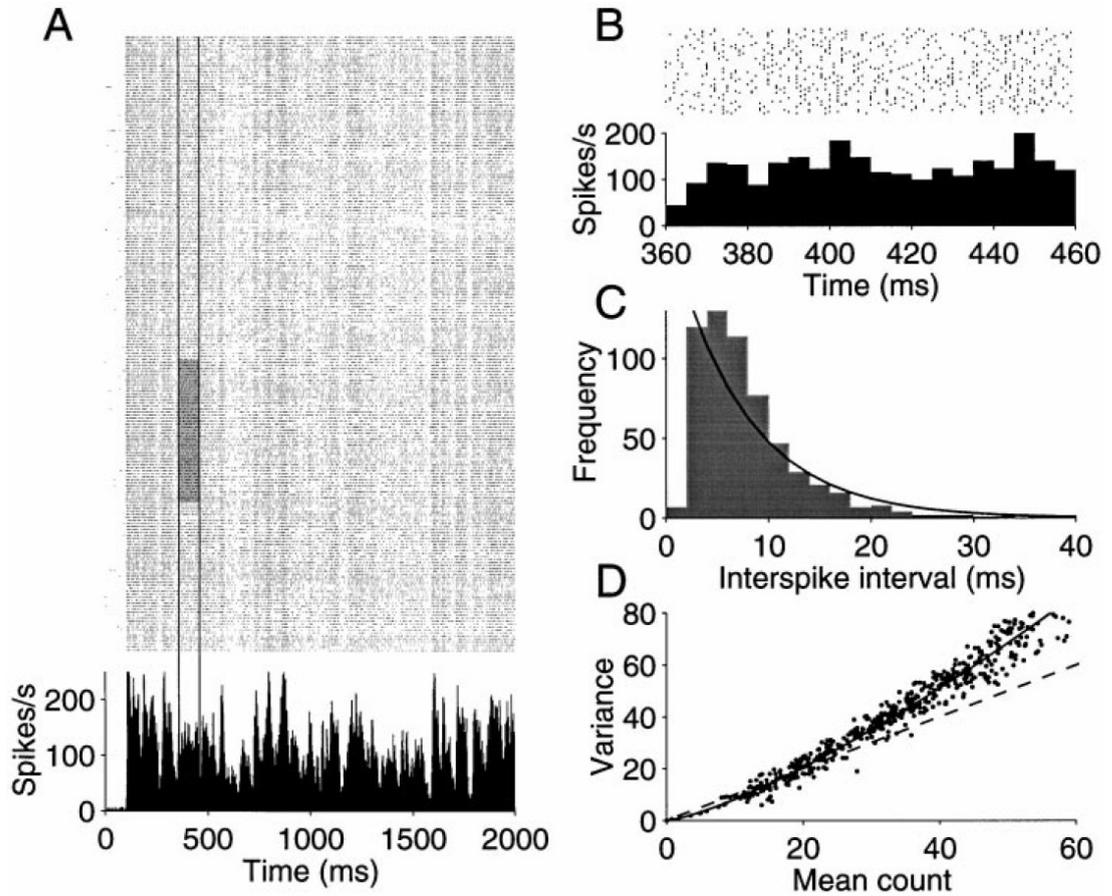


Figure 2.1: Irregular spike trains

The spike train of a cortical neuron recorded from an awake monkey is highly irregular. This irregularity is evident from the raster plot (A) of responses to repeated presentations of an identical random dot stimulus. Large fluctuations in the PSTH are predominantly stimulus dependent. However, a magnification of a relatively stationary response region (the shaded region) illustrates that spike train irregularity occurs even without stimulus related fluctuations (B). This irregularity is quantified by measuring inter-spike intervals (C) and spike counts (D). Reproduced from (Shadlen and Newsome, 1998). Permission to reproduce this figure has been granted by *The Journal of Neuroscience*

Spike train irregularity is not the only signature of cortical irregularity. Membrane potentials are also highly variable (Destexhe et al., 2003; Rudolph et al., 2005). For example, membrane potential recordings from the parietal cortex of an anaesthetised cat are approximately Gaussian distributed (Rudolph et al., 2005). This intracellular irregularity is an important factor in explaining the origin of spike train irregularity.

Spike trains are also *asynchronous*, as well as being irregular. The degree of synchronisation of neuron pairs can be quantified by measuring *Pearson's correlation coefficient* - a measure of the co-variability of a neuron pair. A correlation coefficient close

to 1 indicates that a pair of neurons are highly synchronised whereas a correlation coefficient close to zero indicates that neurons are asynchronous.

Correlations are notoriously difficult to measure, especially when they are small. However, carefully controlled recordings of neuron spike train pairs have recently been obtained from V1 of awake monkeys (Ecker et al., 2010). While the monkeys viewed orientated gratings, correlations were  $0.005 \pm 0.004$  (mean  $\pm$  standard deviation), with averages calculated across neuron pairs. Correlations were also recorded while the monkeys viewed natural images. In that case correlations were  $0.001 \pm 0.005$ . In another experiment, correlations from somatosensory and auditory cortices of urethane-anaesthetized rats were recorded during up-state activity (Renart et al., 2010). The median correlation was 0.0053 with an interquartile range of 0.0024 to 0.0094. These results indicate that spike trains are typically asynchronous in the cortex across animals and brain regions.

There have been some reports of strong spike train correlations in the cortex (Cohen and Kohn, 2011). Strong input fluctuations may be responsible for producing this large temporal co-variability (Hertz, 2010), and when inputs are stationary, cortical activity rapidly returns to an asynchronous spiking state.

The irregular, asynchronous spiking activity observed in cortical neurons stands in stark contrast to the highly regular spiking activity of neurons in peripheral regions of the nervous system, such as neurons in the glomerulus of the olfactory bulb (Laurent et al., 1996; Wehr and Laurent, 1996; MacLeod et al., 1998), or neurons in the cochlear nuclei of the auditory pathway (Joris et al., 1994; Joris et al., 1998). These neurons produce spikes at regular intervals, in response to specific preferred stimuli. Throughout this thesis, we focus on networks of cortical neurons.

### 2.2.2 Where is the irregularity coming from?

There are many possible sources of irregularity in the cortex (Faisal et al., 2008). The most obvious source is *intrinsic neuron noise*. This noise causes neurons to have different spike-train responses to identical stimuli across trials.

Intrinsic neuron noise can be categorised as either *synaptic* or *electrical*. Synaptic noise effects the transmission of spikes between neurons (Fatt and Katz, 1950), whereas electrical noise harms spike generation and propagation within a neuron. The largest source of electrical noise is the stochastic opening and closing of ion channels (White

and Rubinstein, 2000; Hodgkin and Huxley, 1952), with small neurons affected most (Faisal et al., 2005).

Surprisingly, intrinsic neural noise is not the source of cortical irregularity. This can be demonstrated by recording the responses of a cortical neurons to Gaussian input for a series of trials. If the same input is used for each trial, the neural response is almost identical (Mainen and Sejnowski, 1995). Therefore intrinsic noise does not harm spike propagation substantially.

The predominant source of irregularity in the cortex is fluctuating synaptic input. This fluctuating synaptic input is known as *synaptic background noise*, though, confusingly, it is not pure noise. Rather, it is the result of thousands of presynaptic inputs providing fluctuating input at the same time, or within the same membrane time constant (Fellous et al., 2003). Nonetheless, we will refer to this irregularity as noise, for historical consistency. The predominance of synaptic background noise can be measured by comparing cortical neuron responses to injected input in-vivo where synaptic input is intact and in-vitro where synaptic input is impaired. Spike train irregularity is much greater in-vivo, suggesting that synaptic input is the origin of most irregularity (Holt and Douglas, 1996; Rudolph et al., 2005).

### 2.2.3 The synaptic background noise problem

It is surprisingly difficult to build a network model in which spiking is irregular. For example, in the Hopfield network - the first network model of memory, the activity of the model neurons is regular (Hopfield, 1982), as are many other models of cortical activity and function.

This problem can be understood using the following toy network model in which the input to a neuron  $h$  is given by:

$$h = \sum_j^K w_j x_j, \quad (2.1)$$

where  $K$  is the number of synapses connecting to the neuron,  $w_j$  is the strength of the  $j^{\text{th}}$  synapse and  $x_j$  is the spiking state of the  $j^{\text{th}}$  presynaptic neuron. In this toy model,  $x_j = 1$  if the presynaptic neuron is spiking, and  $x_j = 0$  if the presynaptic neuron is silent.

In the cortex, neurons typically receive between 5,000 and 10,000 synaptic inputs

(Braitenber and Schuz, 1991; Binzegger et al., 2004), with each neuron receiving about 5 inputs from the same presynaptic neuron (Markram et al., 1997). Therefore,  $K$  is effectively between 1,000 and 2,000 (London et al., 2010), and consequently, we expect the mean synaptic input to be large:

$$\langle h \rangle \sim \mathcal{O}(K) , \quad (2.2)$$

where brackets  $\langle \dots \rangle$  denote an average over time. Here, we have estimated the size of the mean synaptic input in terms of the number of synapses  $K$ . We say that the mean synaptic input is order  $K$ , or that it *scales* with the number of synapses. This notation is often used in statistical physics, when studying a complex system. It allows us to make qualitative statements about a system, without having to specify all the system details, such as the values of the synaptic strengths.

We can estimate the size of synaptic fluctuations by calculating the standard deviation of the synaptic input. In a completely asynchronous (uncorrelated) network, the synaptic input variance is a sum of  $K$  terms, just as the mean input is a sum of  $K$  terms:

$$\text{var}(h) \simeq \sum_j^K w_j^2 \sigma_j^2 \sim \mathcal{O}(K) , \quad (2.3)$$

where  $\sigma_j^2$  is the spike-train variance of the  $j^{\text{th}}$  neuron. Therefore, the synaptic input standard deviation is:

$$\text{s.d.}(h) = \sqrt{\text{var}(h)} \sim \mathcal{O}(\sqrt{K}) . \quad (2.4)$$

This is the synaptic background noise problem - synaptic drive fluctuations are much smaller than the mean synaptic input, in a generic network model such as the model we have just discussed (Shadlen and Newsome, 1994; Softky, 1995). Specifically, the synaptic input standard deviation is  $\sqrt{K}$  times smaller than the mean synaptic input (Softky and Koch, 1992). This is inconsistent with intracellular recordings, which show that synaptic drive fluctuations are large (Rudolph et al., 2005). Also, this generic model will not produce irregular, asynchronous spike trains because synaptic background noise is the predominant source of irregularity in the cortex.

There are three categories of solution to this problem, following (Abbott, 2008). The first solution is to reduce the effective size of  $K$ . This can be achieved if  $K$  is treated as

the number of synapses on a dendrite, rather than the number of synapses on a neuron (Poirazi et al., 2003). Single neurons then behave as a 2-layer neural network, with synaptic inputs constituting the first layer and dendritic inputs constituting the second layer, with each layer providing a relatively small number of outputs. The effective size of  $K$  can also be reduced if a network uses sparse coding. In sparse coding, only a small number of neurons are active at any given time, so only a few synapses are active (Barlow, 1961; Barrett, 2007; Greene et al., 2009). Sparse coding also has interesting computational properties (Barlow, 2001). However, when neurons are active, sparse coding cannot explain why synaptic drive fluctuations are large.

Another solution is to assume that spike trains are synchronised. In that case, the synaptic input variance becomes a sum of  $K^2$  terms. Consequently, the standard deviation of the synaptic input is order  $K$  so fluctuations are about the same size as the mean synaptic input. However, this solution is inconsistent with cortical activity which is not synchronous as we have discussed (Ecker et al., 2010; Renart et al., 2010). This solution may be used in the early sensory pathways, where incoming sensory signals can be highly correlated (Joris et al., 1994; Laurent et al., 1996; Wehr and Laurent, 1996; MacLeod et al., 1998; Joris et al., 1998).

The third solution is to reduce the size of the total mean input down to the size of the synaptic drive fluctuations, or smaller. This occurs naturally in networks that are randomly connected with an average connection strength of zero. According to the central limit theorem, the sum of  $K$ , zero mean, uncorrelated random variables is order  $\sqrt{K}$ . Therefore, the mean input to a neuron in an asynchronous, randomly connected network will be order  $\sqrt{K}$ , the same order as the synaptic drive fluctuations. A problem with this solution is that it violates *Dale's law*, in that network connectivity allows neurons to be both excitatory and inhibitory simultaneously (Dale, 1935).

Randomly connected networks can be forced to obey Dale's law if all excitatory neurons are required to have positive connection strengths and all inhibitory neurons are required to have negative connection strengths. If there are  $K_E$  excitatory synapses of strength  $w^E$  and  $K_I$  inhibitory synapses of strength  $-w^I$ , where  $w^E$  and  $w^I$  are large, positive constants, then we can write the total mean input as:

$$\langle h \rangle = \sum_{j \in E}^{K_E} w^E \nu_j - \sum_{j \in I}^{K_I} w^I \nu_j, \quad (2.5)$$

where  $\nu_j \equiv \langle x_j \rangle$  is the trial averaged firing rate of the  $j^{\text{th}}$  neuron. Here we have used  $E$  to denote the set of excitatory presynaptic neurons and  $I$  to denote the set of inhibitory presynaptic neurons.

If  $w^E$  and  $w^I$  are chosen so that the background excitatory and inhibitory input to each neuron cancel, then  $\langle h \rangle \sim \mathcal{O}(\sqrt{K})$ . This can be accomplished by fine-tuning network connectivity to ensure that excitation and inhibition cancel (Vogels and Abbott, 2009; Rajan and Abbott, 2006). This is called *detailed balance*. A problem with this approach is that small perturbations to synaptic strengths can disrupt network balancing, causing spike trains to become regular. We will discuss a solution to this fine-tuning problem in the next section.

#### 2.2.4 The standard model of cortical activity

The most successful solution of the synaptic background noise problem was proposed by Carl van Vreeswijk and Haim Sompolinsky in a pair of seminal papers (van Vreeswijk and Sompolinsky, 1998; van Vreeswijk and Sompolinsky, 1996). They proposed that a *dynamic balance* of excitatory and inhibitory synaptic input is responsible for synaptic background noise (van Vreeswijk and Sompolinsky, 1998; van Vreeswijk and Sompolinsky, 1996; Brunel, 2000; Lerchner et al., 2004; Latham et al., 2000). This theory and its extensions has come to be regarded as the standard model of cortical activity (Lerchner and Latham, 2011).

Using mean field theory, van Vreeswijk and Sompolinsky demonstrated that a sparse, randomly connected network of excitatory and inhibitory neurons, can dynamically balance, producing a total synaptic input of about the same size as background synaptic noise. This dynamic balance is similar to detailed balance, except that fine tuning of network connectivity is not required. Instead, balancing is the result of neuron dynamics producing excitatory and inhibitory firing rates for which excitatory and inhibitory synaptic drive to all neurons cancel.

This can be understood by rewriting the mean synaptic input to the toy model that we discussed in the previous section (Eqn. 2.5) in terms of the mean excitatory population firing rate  $\nu^E$  and the mean inhibitory population firing rate  $\nu^I$ :

$$\langle h \rangle = K (w^E p^E \nu^E - w^I p^I \nu^I) + \mathcal{O}(\sqrt{K}), \quad (2.6)$$

where

$$\begin{aligned}\nu^E &\equiv \sum_{j \in E}^{K_E} \nu_j / K_E \\ \nu^I &\equiv \sum_{j \in I}^{K_I} \nu_j / K_I.\end{aligned}\tag{2.7}$$

Here,  $p^E = K^E/K$  is the proportion of excitatory synapses and  $p^I = K^I/K$  is the proportion of inhibitory synapses received by a neuron.

If the excitatory and inhibitory population firing rates evolve so that excitatory and inhibitory synaptic inputs cancel as follows;

$$w^E p^E \nu^E - w^I p^I \nu^I \sim \mathcal{O}\left(1/\sqrt{K}\right),\tag{2.8}$$

then the network is balanced and the total synaptic input and the synaptic background noise are about the same size.

Using methods from statistical physics, van Vreeswijk and Sompolinsky analysed network dynamics in the large network limit and found that dynamic balancing naturally occurs in networks of sparse, randomly connected excitatory and inhibitory neurons, as long as the synaptic weights obey some generous balance conditions (van Vreeswijk and Sompolinsky, 1998; van Vreeswijk and Sompolinsky, 1996). These conditions are often characterised as a requirement that inhibition dominates excitation. This solves the synaptic background noise problem. Changes to network parameters, such as synaptic weights and neuron time-scales do not effect network dynamics adversely. We will discuss the precise balance further in chapter 3.

A surprising consequence of dynamic balance is that excitatory and inhibitory neural populations track the activity of an input population linearly, despite the highly non-linear dynamics of the neurons in the network. This happens because dynamic balance effectively linearises the population dynamics of the network. This population tracking is very fast - much faster than the responses of individual neurons in the network, suggesting a possible functional role for balanced network dynamics.

The network model studied by van Vreeswijk and Sompolinsky consists of very simple binary point neurons, in which the neural states change at time intervals consistent with a Poisson distribution. Despite this, the balanced network results that they derive have important implications for networks of realistic neurons and real neurons. This is a

consequence of the fact that the spiking activity of a neuron in a balanced network is predominantly determined by network dynamics rather than neural dynamics. Indeed, van Vreeswijk and Sompolinsky took advantage of this fact in their choice of neuron model. Networks of simple, less realistic neuron models are analytically tractable to a greater extent than networks of more realistic neurons, such as *integrate and fire* models (Knight, 1972). Nonetheless, it is important to investigate the balanced state in more realistic models.

Irregular asynchronous spiking states that result from dynamically balancing synaptic input have been observed in simulations of current based integrate and fire neurons (Amit and Brunel, 1997b; Amit and Brunel, 1997a; Latham et al., 2000; Brunel, 2000). More recently, networks that consist of conductance based integrate and fire models have been simulated in the balanced state, or the *high conductance state* as it is known in such models (Latham et al., 2000; Lerchner et al., 2004; Lerchner et al., 2006; Hertz, 2010; Lerchner and Latham, 2011; Kumar et al., 2008). This high conductance state is characterised by large membrane potential fluctuations and a depolarised membrane potential (Destexhe et al., 2003).

A series of experiments have tested the predictions of balanced network theory. In-vitro intracellular recordings of ferret prefrontal and occipital cortex have revealed that excitatory and inhibitory conductances do balance (Shu and Hasenstaub, 2003). This balancing coincides with times in which spiking activity is irregular and asynchronous. However, there are many problems with in-vitro neural network experiments. In particular, many synaptic connections are broken in-vitro and it is not clear what effect this might have on the dynamics of the network. More recently, in-vivo intracellular recordings of ferret neocortex have been performed (Haider et al., 2006). During these recordings, ferrets were anaesthetised using ketamine–xylazine. Cortical activity under this anaesthetic fluctuates between up and down states, similar to activity fluctuations during sleep. During up-state activity, cortical spiking is irregular. Intracellular recordings during these up states demonstrate that excitatory and inhibitory conductances balance (Fig. 2.2).

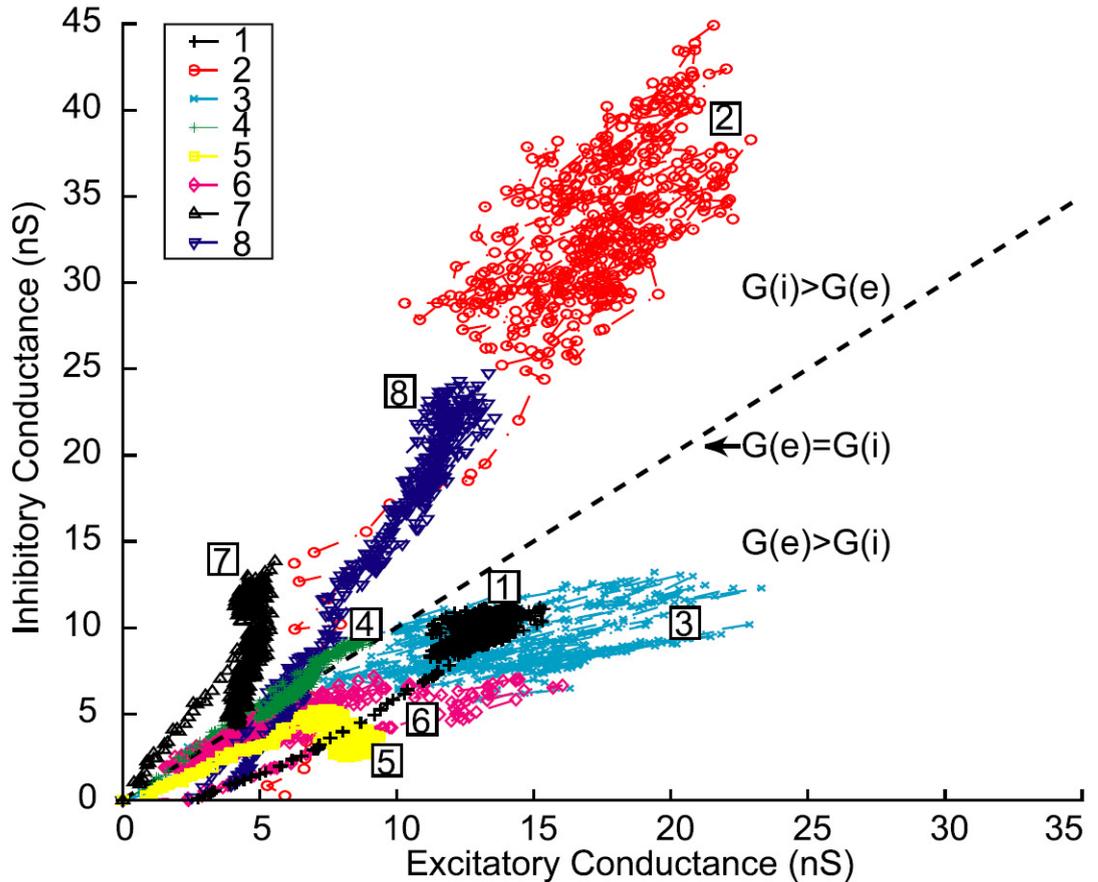


Figure 2.2: Excitatory-Inhibitory balance

Excitatory and inhibitory conductances are proportional to each other, indicating that the cortex operates in a balanced state. This proportionality is demonstrated for eight different neurons in a population (labelled 1 to 8). The conductances to all neurons are clustered around a conductance level that corresponds to equal excitation and inhibition (dashed line, labelled  $G(e) = G(i)$ ). Recordings are obtained during an up-state, starting at a low conductance level and continue for 500 ms. Adapted from (Haider et al., 2006). Permission to reproduce this figure has been granted by *The Journal of Neuroscience*

A number of experimental studies have reported cortical activity that seems inconsistent with balanced network theory, such as non-Gaussian synaptic input fluctuations (Okun and Lampl, 2008; DeWeese and Zador, 2006). However, recent work by Lerchner and Latham has shown that these observations are actually consistent with the theory of balanced networks. These results, along with a growing body of evidence from experiments and realistic network simulations (Destexhe et al., 2003) indicate that balanced network theory is a plausible theory of cortical network activity.

## 2.3 Orientation tuning

There is more to cortical activity than irregular, asynchronous spiking - it is known that the firing rates of spike trains in many cortical areas are tuned to specific features of sensory stimuli. In this *feature tuning*, the firing rate of a neuron is largest when a specific feature occurs in a stimulus. The firing rate decreases as that feature changes from the preferred feature value. For example, the firing rate of a place cell in the hippocampus of an animal is tuned to a specific location in the animals environment (O'Keefe and Dostrovsky, 1971). The firing rate of the place cell increases as the animal moves towards that location, and decreases as it moves away. The most famous example of feature tuning is in the visual cortex, where many neurons are tuned to the orientation of an edge in a visual stimulus (Hubel and Wiesel, 1962).

Firing rates of cortical neurons can also increase monotonically with changes to certain stimulus features. Again, the most famous example of this phenomenon occurs in the visual cortex, where the response of orientation tuned neurons increases monotonically as the contrast of a visual stimulus increases (Sclar and Freeman, 1982).

Since the discovery of orientation tuning, the visual cortex has become a veritable playground for experimentalists and theorists attempting to discover regularity in cortical activity (Olshausen and Field, 2005). In this section, we will discuss the key observations and the theories that attempt to explain visual cortex activity. The hope is that if we can understand the mechanisms and functions of this cortical region, the knowledge will transfer to the entire cortex.

### 2.3.1 Orientation tuning and contrast invariance

Hubel and Weisel discovered that neurons in layer 4 of the cat visual cortex are tuned to edges in visual stimuli (Hubel and Wiesel, 1959). They referred to these orientation tuned cells as *simple cells*. Specifically, these cells respond to gabor filter type stimuli, which consist of a bright line, flanked by dark lines on each side. These cells also respond to gabor filters of the opposite polarity, with dark lines, flanked by bright lines. Each simple cell has a preference for an edge at a particular orientation - firing vigorously when that orientation is presented, and responding less to edges at different orientations (Fig. 2.3). Edges that are perpendicular to the preferred orientation elicit the weakest response.

Orientation tuning has been observed in many animals, such as monkeys (Blasdel and Fitzpatrick, 1984) and ferrets (Chapman and Stryker, 1993). Typically, simple cells are found in layer 4 and layer 6 of the visual cortex, though some orientation tuned cells can be found in other cortical layers (Hubel and Wiesel, 1962; Kelly and Van Essen, 1974; Gilbert, 1977; Shatz and Stryker, 1978; Ferster and Miller, 2000). Neurons in layer 2/3 also respond to oriented edges. However, their response properties are slightly different in that the response is independent of the edge’s polarity. These cells are called *complex cells* (Hubel and Wiesel, 1962).

Orientation tuned cells also respond monotonically to the contrast of a visual stimulus (Sclar and Freeman, 1982; Sompolinsky and Shapley, 1997), as we mentioned earlier. As the image contrast increases, the firing rate increases almost linearly. This is called *contrast invariance* because the shape of the overall neural response and population response is approximately invariant to contrast changes (Fig. 2.3). Contrast invariance is often characterised by measuring tuning curve widths at different contrasts (Sclar and Freeman, 1982; Skottun et al., 1987). If the entire tuning curve is contrast invariant, the tuning curve widths will also be necessarily invariant to contrast.

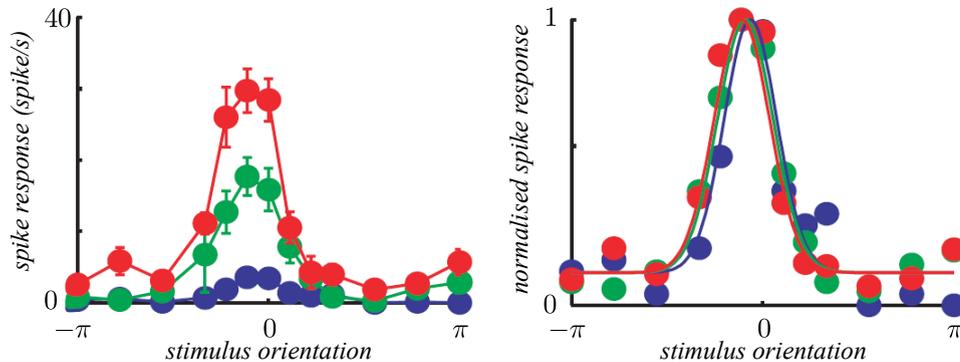


Figure 2.3: Orientation tuning and contrast invariance

Simple cells in the cat visual cortex are orientation tuned (left) and contrast invariant (right). The spiking response of this cell is largest for stimuli containing a grating orientated at  $0^\circ$ . The response becomes smaller as the stimulus orientation changes from the preferred orientation. Stimuli at low contrast (blue), medium contrast (green) and large contrast (red) are presented during this recording. When responses to all contrast are normalised (right), we can see that this cell is contrast invariant. From (Anderson et al., 2000). Reprinted with permission from AAAS.

### 2.3.2 Where does orientation tuning come from?

The anatomical and physiological origin of orientation tuning and contrast invariance have been the subject of intense investigation since their discovery (Ferster and Miller, 2000). From the outset, we know that network connectivity must play a central role in generating tuned responses, because an edge is a spatially extended feature and connectivity is the only mechanism that can integrate information across space. Therefore, feed forward connectivity or recurrent connectivity, or a combination of both are the obvious candidate sources of orientation tuning.

It is extremely difficult to measure connectivity directly (Reid and Alonso, 1995). However, the broad pattern of connectivity to simple cells is partially known (Ferster and Miller, 2000). Simple cells in layer 4 receive excitatory and inhibitory synaptic input from other cells within layer 4 and excitatory synaptic input from cells in the *lateral geniculate nucleus* (LGN). The LGN in turn receives input directly from the optic tract. LGN neurons are not tuned to orientation, but respond to *on-off* visual stimuli, which consist of a bright circle surrounded by a dark ring. LGN cells also respond to *off-on* stimuli, which are the polar opposite of on-off stimuli.

The *feedforward model* was the first model of orientation tuning (Hubel and Wiesel, 1962). In this model, connectivity from LGN neurons to layer 4 neurons produces orientation tuned responses, by integrating on-off cells along a straight line. The orientation of this line is the preferred orientation of the V1 cell receiving the feedforward input. Neurons in V1 only spike if synaptic input is sufficiently large because neurons are non-linear, and behave as threshold units in this model. Therefore, if a simple cell receives input from an edge with an orientation different to its preferred orientation, that cell will have a lower firing rate, because the synaptic input will be lower.

There is plenty of experimental evidence to support the feedforward model. Most compelling is the observation that the pattern of connectivity from the LGN to layer 4 in the cat resembles the connectivity pattern predicted by the feedforward model (Tanaka, 1983; Reid and Alonso, 1995). Another prediction is that synaptic input to layer 4 from the LGN is tuned to orientation. This has also been verified (Ferster et al., 1996; Chung and Ferster, 1998).

However, there are a number of problems with the feedforward model. The most serious difficulty is that it is not contrast invariant. As the contrast of a visual stimulus

increases, the width of tuning curves in the model also increases. This is called the *iceberg effect*, and is a consequence of the non-linear threshold behaviour of neurons. There is no single choice of spiking threshold that can produce realistic tuning curves for all contrast levels in this model (Ferster and Miller, 2000). Along with this fundamental theoretical problem, there is growing experimental evidence that the feedforward model is insufficient. In particular, an analysis of simple cell membrane potential changes elicited by visual stimuli has revealed that most membrane potential changes are not caused by LGN input (Ferster et al., 1996; Chung and Ferster, 1998). Clearly, this is a serious problem, because in the feedforward model, all visual stimulus related changes are caused by LGN input.

One solution to the iceberg effect is to combine inhibitory LGN input with excitatory LGN input in such a way that the total feed-forward input is close to threshold for all contrasts (Carandini and Heeger, 1994). However, synapses from LGN to layer 4 are excitatory, not inhibitory (Ferster and Lindstrom, 1983; Martin and Whitteridge, 1984). An alternative source of inhibitory input can be found in layer 4 itself. The *Push-Pull* model builds upon this idea, suggesting that layer 4 cells that are not contrast invariant themselves can provide the inhibitory input necessary to produce contrast invariant orientation tuning (Ferster and Miller, 2000).

A problem with all of these models is that they are inherently feedforward. Recurrent excitation is completely ignored and recurrent inhibition plays a minor modulatory role, if anything. Furthermore, these models are dynamically impoverished in their disregard for the contribution of recurrent connectivity.

In the *ring model* of orientation tuning, recurrent connectivity is the predominant source of orientation tuning (Ben-Yishai et al., 1995; Somers et al., 1995; Hansel and Sompolinsky, 1996). In this model, synaptic input from the LGN is weakly tuned to orientation. Recurrent connectivity sharpens input, by amplifying the response of neurons to their preferred orientations and suppressing the response to other orientations. The recurrent connectivity that performs this sharpening is *Mexican-Hat* shaped, meaning that cells with similar preferred orientations are more strongly connected than neurons with dissimilar preferred orientation.

Contrast invariance is a consequence of this mechanism - the same input features are sharpened by recurrent connectivity for all inputs, producing a similarly shaped

tuning curve at all contrast levels. However, if synaptic input to some neurons is sub-threshold, contrast invariance is disrupted (Sompolinsky and White, 2005). Apart from this theoretical problem, the ring model is broadly consistent with anatomy and observed tuning curves (Ferster, 1986; Douglas and Martin, 1991; Nelson et al., 1994).

### 2.3.3 Augmenting the standard model with structured connectivity

A serious problem with most orientation tuning models is that they produce regular spike trains. This is inconsistent with observed visual cortex spike trains, which are highly irregular and asynchronous (Ecker et al., 2010), as we discussed in section 2.2. This is a major failure, because it means these models cannot explain the bulk of simple cell activity.

Another problem with most orientation tuning models, such as the ring model and feed-forward models is that their tuning curves are typically homogeneous. This may not seem to be a serious problem, because most experimental evidence seems to suggest that tuning curves have a stereotypical shape (Ringach et al., 2002; Olshausen and Field, 2005). However, this belies a more complicated reality, in which tuning curves can have a bewildering array of possible shapes (Fregnac and Imbert, 1984; Maldonado, 1997; Ferster and Miller, 2000). Such tuning curves are often not reported in the literature, because they are difficult to interpret (Olshausen and Field, 2005). The reality is that in many animals, such as monkeys and ferrets, most cells in layer 4 do not have strong orientation tuning (Chapman and Stryker, 1993; Blasdel and Fitzpatrick, 1984), leading to estimates by some that we only understand between 10 – 20 % of V1 (Olshausen and Field, 2005).

To solve these problems, the theories that explain orientation tuning must somehow be reconciled with balanced network theory. This is not an easy task, because the connectivity underlying each phenomenon is very different. Balanced networks require random sparse background connectivity (van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998) and orientation tuning requires structured recurrent connectivity and structured feed-forward connectivity (Ferster and Miller, 2000; Sompolinsky and White, 2005).

Recently, a model has been developed that may solve these problems using a combination of random connectivity and structured connectivity (Lerchner et al., 2006; van

Vreeswijk and Sompolinsky, 2005). In this model, connectivity is sparse and random, as in balanced networks. However, the probability that a pair of neurons connect to each other depends on the preferred orientation of those neurons - neurons that have similar preferred orientations are more likely to connect to each other than neurons which have dis-similar preferred orientations. Although neurons connect randomly, the connectivity has some structure. The structured connectivity is similar to the Mexican Hat connectivity used in ring attractor models. The Mexican Hat description refers to the probability of a connection rather than the strength of a connection. All connections between the same neuron type are the same strength. Neurons in this model produce irregular spike trains and they are orientation tuned.

The background connectivity that produces irregular spike trains plays an active role in amplifying tuned input (Murphy and Miller, 2009). Balanced networks contain a hidden feedforward connectivity structure, which can be revealed by decomposing network connectivity along orthogonal directions in the dynamical space of the network. With appropriate structured connections, this hidden feedforward connectivity can amplify weakly tuned input. This *balanced amplification* can produce sharp orientation tuned responses.

This is a major theoretical breakthrough. It explains how neurons can produce irregular spike trains and have orientation tuning. Furthermore, it can also explain why orientation tuning is contrast invariant. Excitatory and inhibitory synaptic input is orientation tuned and balanced, so any non-linear changes induced by contrast are quickly counter-balanced by inhibition, producing linear, contrast invariant orientation tuning.

Unfortunately, the success of this balanced model comes at a price. Structured inhibitory input must balance structured excitatory input to produce irregular spiking. Learning this structure may be difficult, because synaptic weight changes must comply with these conditions. Changes to excitatory structure must be matched by changes to inhibitory structure.

Another potential problem is that population tuning curves are homogeneous in this model. This homogeneity is a consequence of the additional balance conditions, which guarantee that all tuning curves have similar, stereotypical shapes. This is consistent with the traditional view that orientation tuning is homogeneous, but is inconsistent

with the more recent realisation that population responses to different orientations can be quite diverse (Olshausen and Field, 2005).

There is a possibility that these problems may be resolved using an alternative form of structured connectivity. Recently, it has been shown that weak structured connectivity embedded in strong background connectivity can be used to store memories and generate realistic spiking activity (Roudi and Latham, 2007). However, because structured connectivity is weak, its form is less constrained. In the next chapter (Chapter 3), we will investigate orientation tuning in networks with this type of connectivity, where structured connectivity is much weaker than background random connectivity, rather than being the same size.

## 2.4 Computation

Computation is one of the primary functions of the brain (Dayan and Abbott, 2001). Computations transform sensory information from a complicated format into a useful format. For example, while going for a walk in a forest, we are performing object recognition - analysing complicated scenes of trees and leaves and detecting shadowy objects lurking in the dark, such as a hungry pack of wolves on the hunt for food. Our survival depends on our ability to compute.

Originally, single neurons were considered to be the fundamental units of computation. Neurons are extraordinarily complex cells, capable of supporting many simple transformations. However, they are fundamentally constrained by noise, and they can only transmit discrete all-or-nothing signals, in the form of action potentials (Hodgkin and Huxley, 1952).

During the last 20 years there has been a paradigm shift in how we think about neural activity and computation (Douglas and Martin, 1991). It is now widely believed that neural computation is mediated primarily by neural networks (Hertz et al., 1991). Neural network computations can be much more robust than single neuron computations. Also, the discrete nature of neural signals can be mitigated with a network of neurons operating in concert.

The computational capacity of neural networks is not known, mostly because a neural network is an extremely complicated dynamical system. Bridging the gap between complex computation and neural network dynamics promises to be one of the major

scientific challenges of the 21st century. Nonetheless, there has been much progress recently using methods from dynamical systems theory (Amit et al., 1985; Latham et al., 2000), stochastic calculus (Boerlin and Denève, 2011), information theory (Shannon, 1948) and control theory (Sutton and Barto, 1998; Todorov and Jordan, 2002). We review some of that progress in this section.

### 2.4.1 Orientation selectivity

*Orientation selectivity* is a simple computation in which the orientation of an edge in an image is computed. Naturally occurring images contain many orientated bars (Olshausen and Field, 1996; Bell and Sejnowski, 1997). Detecting the orientation of these bars is a critical building block for more sophisticated computations, such as object recognition. Orientation selectivity, like orientation tuning has become an important toy computation studied by experimentalists and theorists, who hope that general principles of computation may be discovered if we can understand this simple computation (Ferster and Miller, 2000).

The orientation of a bar is encoded in the activity of a network of recurrently connected neurons receiving visual input. This is called *population coding* (Pouget et al., 2000). Population codes are essential for computations such as orientation selectivity, where the stimulus is spatially extended. The standard model of population coding is given by (Zemel et al., 1998):

$$r_i = f_i(\theta) + \eta_i \tag{2.9}$$

where  $r_i$  is the firing rate of the  $i^{th}$  neuron in the population,  $\theta$  represents a stimulus feature such as the orientation of a bar and  $\eta_i$  is population noise (Zemel et al., 1998). The function  $f_i$  represents the tuning curve of the  $i^{th}$  neuron. This is typically a Gaussian function or a sigmoid-like non-linearity, depending on the type of population code. Usually, the noise is Gaussian, or Poisson.

This network model is a dramatically simplified version of the spiking models we have already discussed. The relationship between those spiking models and this population coding model is complicated. The tuning curves  $f_i$  and the noise statistics are related to network connectivity and neural dynamics in a convoluted, unspecified way. This may seem problematic, but this simplification has led to significant advances in our understanding of computation in neural populations, both in experiments that can be

easily fitted to this model (Paninski, 2004), and in theory, where analysis is tractable (Abbott and Dayan, 1999).

The computational performance of a population can be quantified by calculating the variance of an unbiased optimal orientation decoder. An *unbiased decoder* is a decoder that can decode a parameter correctly on average. An *optimal decoder* produces estimates of parameter  $\theta$  with the smallest possible variance. This variance is called the Cramér–Rao bound. If the optimal variance is large, then computational performance is low, because there is large uncertainty about orientation  $\theta$ . If the variance is very small then computational performance is high. Therefore, the inverse variance of an unbiased optimal decoder provides a good measure of the computational performance of a network. It is called the *Fisher Information*, and can be written as follows:

$$F(\theta) = \int (\partial_\theta \log p(\mathbf{r}|\theta))^2 p(\mathbf{r}|\theta) d\mathbf{r} \quad (2.10)$$

where we have written  $\mathbf{r} = (r_1, r_2, \dots, r_i, \dots, r_N)$  to denote the entire population response, with  $r_i$  representing the firing rate of the  $i^{\text{th}}$  neuron, given by equation 2.9.

Population codes with sharply peaked tuning curves have high Fisher Information (Paradiso, 1988; Seung and Sompolinsky, 1993). Therefore, we can conclude that orientation tuning is not simply phenomenological, but rather, it is functionally useful.

Orientation tuned neurons are also highly irregular. Unsurprisingly, population codes that have large noise levels have low Fisher Information indicating that cortical irregularity can harm orientation selectivity (Paradiso, 1988; Seung and Sompolinsky, 1993). The presence of cortical irregularity seems to be inconsistent with one of our most compelling theories for neural computation - that the brain performs computations optimally (Barlow, 1961). This mystery is confounded by the fact that cortical irregularity is not caused by intrinsic neural noise. Synaptic drive fluctuations are the predominant cause of spike train irregularity so it should be possible to eliminate most irregularity.

Fisher Information provides a theoretical bound on computational performance. However, it is important that a realistic network can implement a decoder that achieves this performance level in practice. A maximum-likelihood decoder or *ML decoder* is an example of a decoder that can be implemented in a realistic network. This decoding explicitly accounts for the noisy probabilistic nature of population coding by estimating

the orientation  $\theta$ , which maximises the probability of producing observed population activity:

$$\theta_{ML} = \max_{\theta} p(\mathbf{r}|\theta). \quad (2.11)$$

It turns out that it is actually possible for a large recurrently connected network to extract almost all the information available by performing ML decoding (Deneve et al., 2001). Indeed, the maximum likelihood decoder is asymptotically efficient (Amari and Nagaoka, 2000). Therefore, Fisher Information can be regarded as a good measure of computational performance.

There are other coding schemes that the brain might use, such as *temporal coding*, in which the precise timing of spikes is important. However, intrinsic neural noise is a serious problem for temporal coding. Rate based population codes can actually remove this noise by integrating activity across the population (Deneve et al., 2001). This is particularly important, because if noise increases in each layer of a network, noise will eventually dominate. A consequence of noise removal is that computations in populations can be much more accurate than single neuron computations (Paradiso, 1988). This noise removal becomes more effective in large networks, as long as the noise is not strongly correlated across neurons (Zohary et al., 1994). We will discuss the role of correlations in population coding in the next section.

Apart from noise removal and improved computational performance, population codes are useful because they can support complicated non-linear computations (Poggio, 1990). Population codes provide a basis set, which can be used to build computations such as object recognition (Poggio and Edelman, 1990) and object transformations (Salinas and Abbott, 1995; Pouget and Sejnowski, 1997). Population codes are also particularly robust to damage. If some neurons in the population die, or if synaptic connections fail, a population code will not be adversely effected, because the code does not rely on any single neuron. They can also be used for transmitting information robustly through time, or in other words, for storing memories (Hopfield, 1982).

#### **2.4.2 The role of correlations in computation**

Correlated variability can play an important, sometimes negative role in computation (Averbeck et al., 2006). For example, in a population code, if each neuron is completely correlated with every other neuron, it is not possible to remove noise by integrating the

population activity, just as it is impossible to remove noise from a single neuron. In this way, correlations can destroy the computational performance of a network.

Nonetheless, we can not simply dismiss correlations as universally harmful. Indeed, correlations may actually improve computational performance, depending on the structure of the correlations and the structure of the tuning curves in the population code (Averbeck et al., 2006; Ecker et al., 2010). Therefore, it is difficult to make definitive statements about the role of correlations in computation, if any.

Apart from being difficult to interpret, correlations are extremely difficult to measure (Ecker et al., 2010). Movement of the electrodes that record correlations can produce artificial correlations. Also, failure to adequately isolate individual cells in multi electrode recordings can produce artificial correlations. Anaesthetics that produce oscillatory activity, and internal changes in the dynamical state of the animal also cause problems for correlation measurements.

Nonetheless, carefully controlled correlation measurements have recently revealed that correlations can be very small in the cortex (Ecker et al., 2010; Renart et al., 2010). However, even very weak correlations can impair the computational performance of the cortex (Zohary et al., 1994; Sompolinsky et al., 2001; Wilke and Eurich, 2002; Averbeck and Lee, 2004). Whether or not weak correlations are significant in computation sensitively depends on the structure of the correlations and the population code (Abbott and Dayan, 1999).

This subtle relationship between correlations and computation is problematic for theorists and experimentalists attempting to understand computation in the cortex. If the contribution of correlations to computation were small enough to be ignored, we could make significant progress in our understanding of neural computation. In practice, most experimentalists and theorists simply ignore correlations, without justification, and hope that this does not negate their results.

The computational consequences of ignoring correlations can be quantified by calculating  $\Delta I_{diag}$ , defined as follows:

$$\Delta I_{diag} \equiv I - I_{diag}, \quad (2.12)$$

where  $I$  is a measure of computational performance with correlations, and  $I_{diag}$  is a measure of computational performance without correlations (?; Wu et al., 2001; Aver-

beck et al., 2006). For example,  $I$  could be the Fisher Information (Eqn. 2.10) of a correlated population code, and  $I_{diag}$  could be Fisher Information of the same population code but without correlated variability.

For a simple toy model containing just two neurons we can see that the consequences of ignoring correlations depend on the relationship between correlations and average neural responses (Fig. 2.4). If these interact so that there is no difference between a population decoder that ignores correlations and a decoder that does not ignore correlations, then  $\Delta I_{diag} = 0$  and correlations can be ignored without harming decoding. However, if correlations do effect decoding then  $\Delta I_{diag} > 0$  and ignoring correlations harms computation (Wu et al., 2001; Nirenberg and Latham, 2003).

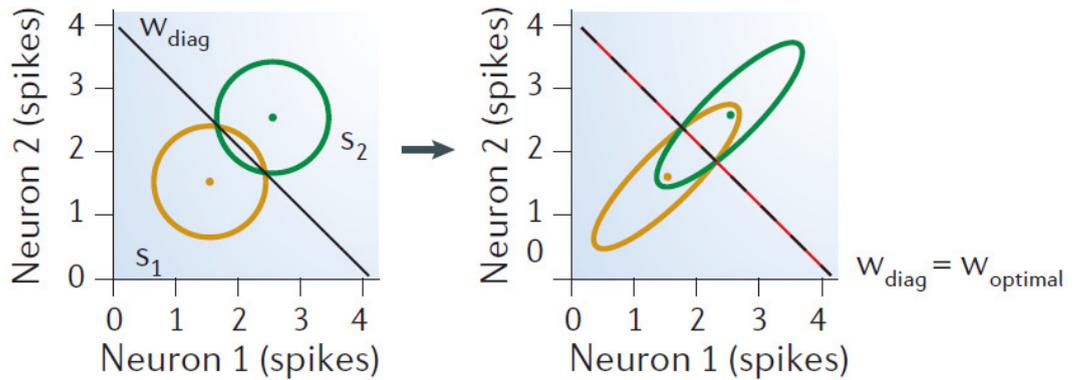
All the theoretical studies that we have discussed here are based on simple population coding models, where the correlation structure is chosen arbitrarily. In real neural populations, the correlation structure is determined by network connectivity and network input. To conclusively determine whether or not correlations can be ignored, a network model with realistic network correlations must be investigated. We will discuss this further in chapter 5.

### 2.4.3 The role of noise in computation

Most spike train irregularity in the cortex is caused by synaptic background noise (van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998). This is surprising, because this noise can be mostly eliminated using network connectivity that is completely structured (Destexhe and Contreras, 2006). So why are neural spike trains so irregular? Could this irregularity be useful in computation, or, is it an evolutionary mistake that fundamentally impairs cortical function?

There have been a number of intriguing theories of neural computation in which synaptic background noise is beneficial (Destexhe and Contreras, 2006). One proposal is that the cortex uses stochastic resonance (Longtin et al., 1991; Bulsara et al., 1991; Stemmler, 1996; Destexhe and Contreras, 2006; McDonnell and Abbott, 2009; Greenwood et al., 2000). *Stowe* imagine that structured connectivity produces orientation tuning while background connectivity provides dynamic balance. *chastic resonance* occurs when an increase in noise increases computational performance in a non-linear system (Fig. 2.5) (McDonnell and Abbott, 2009). Usually, this improvement is bound

**a**  $\Delta I_{\text{diag}} = 0$



**b**  $\Delta I_{\text{diag}} > 0$

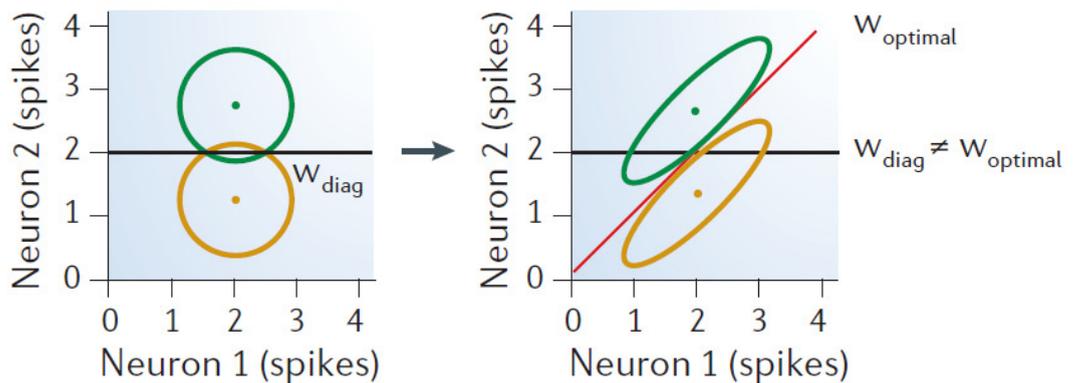


Figure 2.4: Correlations and computation

Ignoring correlations can harm computational performance dramatically, depending on the structure of the correlations. Here, the presence or absence of stimulus  $s_1$  (yellow) and stimulus  $s_2$  (green) is computed from the spiking activity of two neurons. The encoding of these stimuli is noisy, as illustrated using ellipses that represent 95% confidence intervals. The panels on the left show the neural encoding without correlations and the panels on the right show the encoding with correlations. (a) The computational performance of a decoder that ignores correlations (black line) can be just as good as a decoder that does not ignore correlations (red line), if the correlations have no effect on the decoder. The decoders, represented by these lines detect stimuli by reporting whether network activity is above or below the line. (b) If correlations do effect the decoder substantially, then computation can be harmed significantly by ignoring correlations. Adapted by permission from Macmillan Publishers Ltd: Nature Reviews Neuroscience (Averbeck et al., 2006), copyright 2006.

by a *stochastic resonance peak*, because too much noise eventually degrades incoming signals.

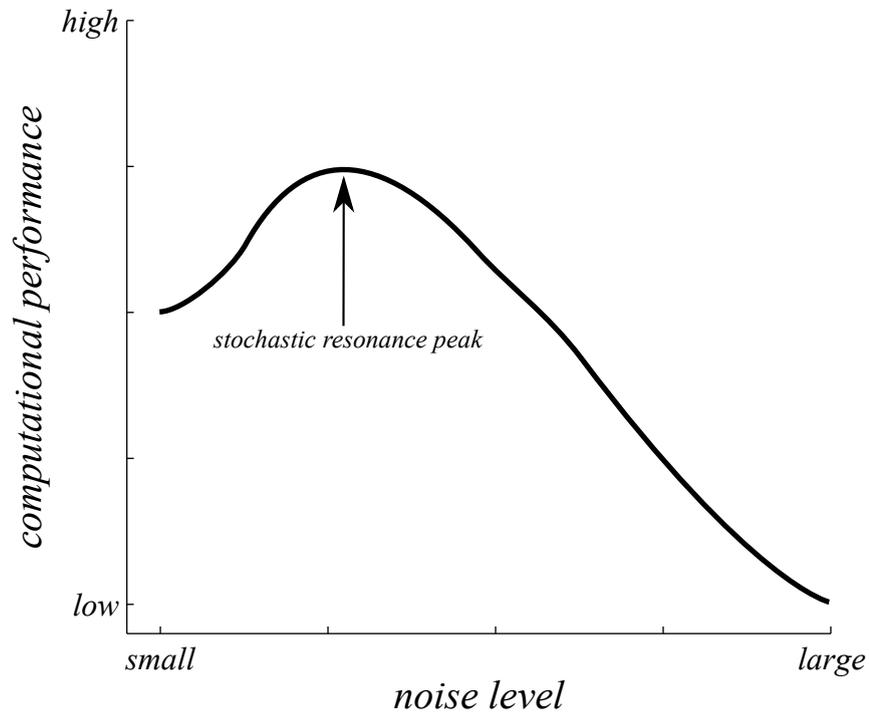


Figure 2.5: Stochastic resonance

The computational performance of a non-linear system can increase with noise, as in this cartoon. This is called *stochastic resonance*. Computational performance is maximised at a *stochastic resonance peak*. Excessive noise degrades the systems performance substantially, and insufficient noise also impairs performance.

A single neuron is an example of a non-linear dynamical system in which stochastic resonance might occur (Longtin et al., 1991; Bulsara et al., 1991; Stemmler, 1996). Typically, neurons behave as thresholding units, producing spikes when synaptic input is above some threshold and remaining silent otherwise. This non-linearity can impair the transmission of information through the neuron if the synaptic drive is not close to threshold. However, noise can effectively increase the range of synaptic signals that can be transmitted through a neuron (Lecar and Nossal, 1971), by allowing signals that are far from threshold to occasionally fluctuate across the spiking threshold. In this way, stochastic resonance occurs and computational performance is enhanced because of synaptic background noise.

Stochastic resonance has not yet been demonstrated in population codes. It is possible that the detrimental effects of synaptic background noise are so substantial

that there is no resonance peak in neural populations. We will discuss this further in chapter 5.

Another possible use of spike train irregularity is in *probabilistic population coding*, where a population of neurons represents an entire stimulus probability distribution, rather than a single stimulus parameter alone (Ma et al., 2006; Knill and Pouget, 2004). In probabilistic population coding, distributions are represented parametrically, using the statistics of neural activity. Probabilistic population codes are capable of supporting sophisticated computations such as Bayesian inference (Knill and Pouget, 2004).

Cortical irregularity might also play a role in memory. Network models that spike regularly can form spurious memory states that are irrelevant at best and harmful at worst (Hertz et al., 1991). The background connectivity responsible for irregular spiking can also act as a memory reservoir, from which relevant memories can be recalled and irrelevant memories can be ignored. Recently, this has been demonstrated in a balanced network with realistic spike train irregularity (Roudi and Latham, 2007).

Another approach to understanding the computational role of cortical irregularity is to treat irregularity as a complicated chaotic signal, rather than noise (Buonomano and Merzenich, 1995; Hopfield and Brody, 2001; Maass et al., 2002; Maass and Markram, 2004; Jaeger and Haas, 2004; Sussillo and Abbott, 2009). Chaotic network dynamics support computation by supplying a *computational reservoir*, from which a broad range of computations can be extracted using a simple linear decoder. The computational reservoir, as its name suggests, encodes a large number of computations in network activity. These computations are produced by network dynamics which repeatedly project network activity through non-linear neurons and network connectivity, producing a complicated superposition of computations.

A problem with this proposal is that it is particularly difficult to implement in a spiking network. However, it has recently been demonstrated that a linear operation can be used to robustly extract sophisticated Bayesian computations from a network of spiking neurons (Boerlin and Denève, 2011; Boerlin, 2011; Deneve, 2008a; Deneve, 2008b). In these models, neurons spike to signal a computational error. These spikes are then used for error correction. The membrane potential of each neuron is a measure of the error size. Computational robustness, dynamic balance and the consequent spike train irregularity are all features of this error reduction mechanism.

It is not clear which, if any, of these computational proposals explain the existence of cortical noise. It may be that the brain uses noise in several different ways and that more than one of these proposals is correct. At the other extreme, cortical irregularity may simply be a truly useless handicap - something that fundamentally limits the computational capacity of a brain. In any case, it is important that we identify as many unique theories of neural computation as possible, consistent with known physiology and anatomy.

In this thesis, we propose that cortical noise is a signature of stochastic resonance in a neural population (Chapter 5). We begin by describing a balanced network model that is capable of producing realistic spike train irregularity and feature tuning (Chapter 3). Before we can calculate the computational performance of this network, we must calculate the spike train correlations (Chapter 4). We can then understand how spike train irregularity, orientation tuning, computational performance and balanced network connectivity all interact to produce the rich, computational and dynamical diversity displayed by brains and their owners.



Part I

Dynamics



## Chapter 3

# Orientation Tuning in Balanced Networks

### 3.1 Introduction

Neurons in the visual cortex are *orientation tuned* (Hubel and Wiesel, 1962; Blasdel and Fitzpatrick, 1984; Chapman and Stryker, 1993), *contrast invariant* (Sclar and Freeman, 1982) and *dynamically balanced* (Haider et al., 2006; Ecker et al., 2010). Reconciling these diverse phenomena in a single cortical model is a major theoretical challenge, because the network connectivity underlying each observation is very different in the models proposed so far (Hubel and Wiesel, 1962; Ben-Yishai et al., 1995; Somers et al., 1995; Hansel and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998). Solving this problem is important because it will provide a long sought unifying explanation for many visual cortex observations, and more importantly, it may lead to a general theory of cortical activity and connectivity (Ferster and Miller, 2000).

Orientation tuned neurons respond vigorously to images containing edges at a preferred orientation (Hubel and Wiesel, 1962; Blasdel and Fitzpatrick, 1984; Chapman and Stryker, 1993). These responses are *contrast invariant*, with tuning curve shapes preserved as the image contrast changes (Sclar and Freeman, 1982). The spiking activity of these neurons is irregular and asynchronous (Ecker et al., 2010), a consequence of dynamically balancing excitatory and inhibitory synaptic input (van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998; Haider et al., 2006; Ecker

et al., 2010).

*Background connectivity*, which is strong, random and sparse, is responsible for dynamic balancing (van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998; Brunel, 2000). This connectivity is very different to the connectivity responsible for orientation tuning and contrast invariance, which is highly structured (Hubel and Wiesel, 1962; Ben-Yishai et al., 1995; Somers et al., 1995; Hansel and Sompolinsky, 1996; Ferster and Miller, 2000). If connectivity is structured and strong, it is possible to have orientation tuning and dynamic balance simultaneously (Lerchner et al., 2006). However, this structure is difficult to learn because the strong inhibitory structure must precisely balance the strong excitatory structure.

We propose that a combination of weak structured connectivity embedded in strong background connectivity is responsible for irregular, orientation tuned spiking activity. Connectivity of this form was recently used in a cortical model of memory, with structured connectivity responsible for memory storage and background connectivity responsible for irregular spiking (Roudi and Latham, 2007). Here, we imagine that structured connectivity produces orientation tuning while background connectivity provides dynamic balance. We expect this model to be as stable as unstructured balanced networks, because weak structure is unlikely to disrupt strong background connectivity.

A potential problem with this proposal is that the contribution of strong background connectivity might overwhelm the contribution of weak structured connectivity. We find that this does not happen, so neurons can be orientation tuned and balanced. However, the orientation tuning is noisy, because dynamic balancing is noisy. This is actually consistent with the much overlooked observation of noisy orientation tuning in experiments (Chapman and Stryker, 1993; Blasdel and Fitzpatrick, 1984).

We also find that orientation tuning is contrast invariant in our model. Large synaptic drive fluctuations from dynamic balancing effectively linearise single neuron dynamics, allowing tuning curve shapes to become invariant to image contrast. This linearisation mechanism has been observed experimentally (Anderson et al., 2000).

### 3.2 Balanced network model

Our model consists of three populations of neurons; an input population (X) that encodes orientation  $\theta$ , a population of excitatory neurons (E) and a population of in-

hibitory neurons (I) (Fig. 3.1). The orientation  $\theta$  of an edge in a visual stimulus is represented by a hill of spiking activity in the input population. The contrast of the stimulus,  $c$ , is represented by the height of the activity hill. We can loosely consider this to be a model of an orientation hyper-column, where the input population is layer 4 of V1 and the E-I population is layer 2/3 (Amit and Brunel, 1997a; Amit and Brunel, 1997b). For convenience, we refer to the excitatory and inhibitory population together as the E-I population (excitatory-inhibitory).

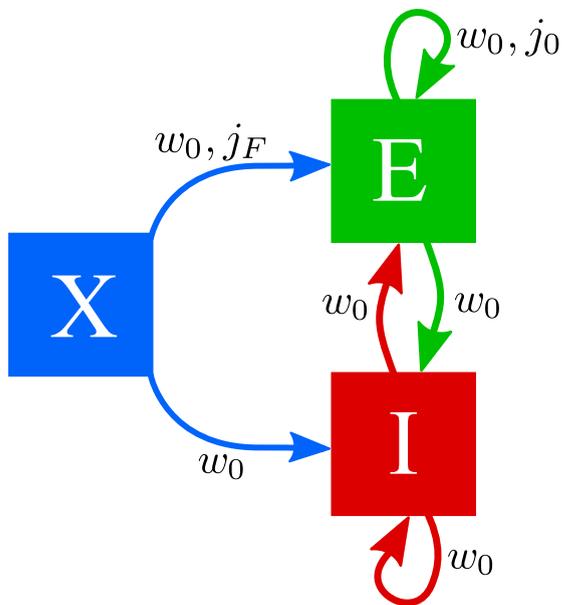


Figure 3.1: Network model

Our network model consists of three populations of neurons. The input population (X) contains excitatory neurons that project onto our recurrently connected excitatory (E) and inhibitory (I) populations. Background connectivity is parametrised by  $w_0$ . Structured connectivity between excitatory neurons is parametrised by  $j_F$  and  $j_0$ .

Network connectivity is sparse and random. Each neuron in the E-I population receives connections from  $K$  other neurons on average from the E-I population and  $K_X$  neurons on average from the input population. Neurons in the input population do not connect to each other. Synaptic strengths are not random, but are chosen so that E-I neurons can be orientation tuned and balanced.

The input population contains  $N_X$  excitatory neurons and the E-I population contains  $N$  neurons. The number of excitatory neurons in the E-I population is  $N_E$  and

the number of inhibitory neurons is  $N_I$ , where

$$N_E = p_E N \tag{3.1}$$

$$N_I = (1 - p_E) N \tag{3.2}$$

and  $p_E$  is the proportion of excitatory neurons. Here,  $N + N_X$  corresponds to the number of neurons in a cortical column, which is about 1 million neurons.

### 3.2.1 Single neuron dynamics

Neurons interact by transmitting action potentials, which are binary signals (Hodgkin and Huxley, 1952). Therefore, we use binary state neurons in our network model. This is a standard assumption in spiking network models (McCulloch and Pitts, 1943; Renart et al., 2010). Specifically, we use  $x_i(t)$  to denote the state of the  $i^{\text{th}}$  neuron at time  $t$ , where neurons may be in a spiking state denoted by  $x_i = 1$  or in a quiescent state, denoted by  $x_i = 0$ . We will often use the vector notation  $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_{N+N_X}(t))$  to denote the spiking state of the entire network.

Spiking is largely determined by synaptic input, through a non-linear, noisy transformation. We capture these dynamics using a simple spiking model in which the synaptic drive  $h_i$  determines the transition rates between the quiescent state and the spiking state of the  $i^{\text{th}}$  neuron at time  $t$ :

$$\begin{aligned} r(0 \rightarrow 1) &= f(h_i) / \tau_Q \\ r(1 \rightarrow 0) &= (1 - f(h_i)) / \tau_Q, \end{aligned} \tag{3.3}$$

where  $\tau_Q$  is the typical neuron response-time:

$$\tau_Q = \begin{cases} \tau_E & i \in E; \\ \tau_I & i \in I, \end{cases} \tag{3.4}$$

and  $f(h_i)$  is a sigmoidal non-linearity. Throughout much of the thesis, we use the Heaviside function for this non-linearity:

$$f(h_i) = \Theta(h_i), \tag{3.5}$$

where  $\Theta(h_i) = 1$  if  $h_i > 0$  and  $\Theta(h_i) = 0$  otherwise. Also, we use  $Q \in \{E, I\}$  to label a neural parameter as belonging to either the excitatory population  $E$  or the inhibitory population  $I$ . We will use this convention throughout the thesis when discussing properties specific to different neural populations.

The synaptic drive consists of a recurrent component and a feed-forward component:

$$h_i(t) = \sum_{j \in E, I}^N A_{ij} x_j(t) + \sum_{j \in X}^{N_X} F_{ij} x_j(t) - \theta^Q, \quad (3.6)$$

where  $A_{ij}$  is the strength of the synaptic connection from neuron  $j$  to neuron  $i$ ,  $F_{ij}$  is the strength of the synaptic connection from input neuron  $j$  to neuron  $i$ , and  $\theta^Q$  is the spiking threshold, where  $Q \in \{E, I\}$ .

Spiking dynamics of the input population are similar to the spiking dynamics of the E-I population. The transition rates between the quiescent state and the spiking state of the  $i^{\text{th}}$  input neuron at time  $t$  are:

$$\begin{aligned} r(0 \rightarrow 1) &= u_i(\theta) / \tau_X \\ r(1 \rightarrow 0) &= (1 - u_i(\theta)) / \tau_X, \end{aligned} \quad (3.7)$$

where  $\tau_X$  is the typical input neuron response-time and  $u_i(\theta)$  determines how the input population encodes orientation  $\theta$ .

Each input neuron represents a different orientation by spiking most when its preferred orientation is present (Fig. 3.2). Specifically, the  $i^{\text{th}}$  neuron in the input population represents orientation  $\theta$  as follows:

$$u_i(\theta) = c \exp\left(-\frac{\sin^2(\theta_i - \theta)}{2\kappa^2}\right), \quad (3.8)$$

where  $c \in [0, 1]$  is the input contrast,  $\kappa$  represents the width of the spiking hill of activity relative to background spiking activity, and

$$\theta_i = \pi i / N_X, \quad (3.9)$$

is the preferred orientation of the  $i^{\text{th}}$  neuron. The spiking activity of the population, therefore, is peaked at the neuron that responds most strongly to  $\theta$ .

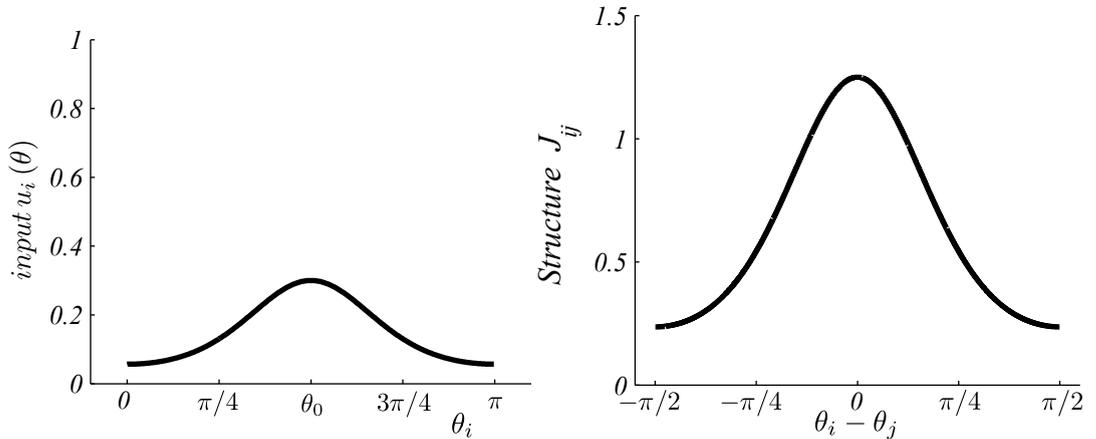


Figure 3.2: Network input and structured connectivity

The input population equilibrium firing rate is a hill of activity with a peak at  $\theta_o$  (left). Recurrent structured connectivity is Mexican-Hat shaped, with stronger connections between neurons with similar preferred orientations and weaker connections between neurons with dissimilar preferred orientations (right). Structured connectivity like this can amplify orientation tuned population activity, such as the input population here.

The neurons in our model capture some of the most important features of real neurons such as non-linear dynamics and spiking activity. They have often been used before in neuroscience, because they can be understood quantitatively (Hertz et al., 1991). Indeed, they were originally developed as analytically tractable models of magnetism (Glauber, 1963). However, as with all models, there are many neuron features that are not captured, such as complicated synaptic dynamics, dendritic dynamics and action potential generation. This is not a serious problem when studying balanced network dynamics, because balanced networks can produce realistic cortex-like spiking activity even with artificial neurons, as we shall see (Douglas and Martin, 1991). Furthermore, excessively complex models can be intractable and uninterpretable, so simple neuron models are often better.

### 3.2.2 Network connectivity

We propose that a combination of weak structured connectivity embedded in strong background connectivity can produce orientation-tuned, balanced network activity. The connection strength from neuron  $j$  to neuron  $i$  is chosen randomly according to the

following equation:

$$A_{ij} = \begin{cases} \frac{w_0}{\sqrt{K}} W_{ij} + \frac{j_0}{K} J_{ij} & \text{prob } K/N; \\ 0 & \text{prob } 1 - K/N, \end{cases} \quad (3.10)$$

Here,  $\mathbf{W}$  is the *background connectivity* required for balanced cortical dynamics and  $\mathbf{J}$  is the *structured connectivity* that may produce orientation tuned responses. The relative contributions are determined by the parameters  $w_0$  and  $j_0$ .

Structured connectivity is mexican hat shaped between excitatory neurons (Fig. 3.2, right). We have chosen this connectivity because it matches the structure of our input. It is stronger between neurons that have similar orientation preferences than between those that have a preference for very different orientations. Therefore, it amplifies the response of neurons that are firing vigorously in response to a particular orientation. Specifically, structure is given by:

$$J_{ij} = p_E^{-1} \begin{cases} \exp \left[ -\frac{\sin^2(\theta_i - \theta_j)}{2\sigma_J^2} \right] & i, j \text{ excitatory neurons;} \\ 0 & \text{otherwise,} \end{cases} \quad (3.11)$$

where

$$\theta_j = \pi j / N_E, \quad (3.12)$$

and  $\sigma_J$  is the full width of this structured connectivity.

Background connectivity is responsible for the dynamic balance between excitatory and inhibitory synaptic drive. We let the background connection strengths depend only on neuron type, so we write:

$$W_{ij} = \begin{cases} W^{QE} & \text{neuron } j \text{ is excitatory;} \\ -W^{QI} & \text{neuron } j \text{ is inhibitory,} \end{cases} \quad (3.13)$$

where  $W^{QE}$  and  $W^{QI}$  are positive constants with  $Q = E$  if neuron  $i$  is excitatory and  $Q = I$  if it is inhibitory.

Feed-forward connectivity from the input population X is similar to the E-I connectivity that we have just described, except that it is excitatory (Ferster and Lind-

strom, 1983; Martin and Whitteridge, 1984). It consists of weak structured connectivity embedded in strong background connectivity:

$$F_{ij} = \begin{cases} \frac{w_o}{\sqrt{K}} W_{ij}^F + \frac{j_F}{K} J_{ij}^F & \text{prob } K_x/N_x; \\ 0 & \text{prob } 1 - K_x/N_x, \end{cases} \quad (3.14)$$

Here,  $\mathbf{W}^F$  is the *feed-forward background connectivity* and  $\mathbf{J}^F$  is *feed-forward structured connectivity*. The relative contributions are determined by  $w_o$  and  $j_F$ . Input neurons do not connect to each other.

As before, we use mexican hat shaped connectivity as feed-forward structure. This amplifies the activity of neurons in the input population whose preferred orientation is close to orientation  $\theta$ :

$$J_{ij}^F = \begin{cases} \exp\left[-\frac{\sin^2(\theta_i - \theta_j^X)}{2\sigma_F^2}\right] & i \in E \text{ and } j \in X; \\ 0 & \text{otherwise,} \end{cases} \quad (3.15)$$

where,

$$\theta_j^X = \pi j/N_X \quad \text{where } i \in [1, N_X],$$

and  $\sigma_F$  determines the full width of the feed-forward structure.

Background connectivity from the input population is determined by neuron type in the E-I population:

$$W_{ij}^F = \begin{cases} W^{EX} & \text{neuron } i \text{ is excitatory;} \\ W^{IX} & \text{neuron } i \text{ is inhibitory,} \end{cases} \quad (3.16)$$

where  $W^{EX}$  and  $W^{IX}$  are positive.

### 3.2.3 Balance conditions

In strongly connected networks there is a danger that connectivity may produce pathological dynamical states in which neurons spike incessantly (or remain completely silent). This can happen because excitatory and inhibitory connections are strong enough to overwhelm neurons with synaptic input.

To prevent this, background connectivity must satisfy the following balance conditions (*van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998*):

$$\begin{aligned} W^{EX}/W^{IX} &> W^{EI}/W^{II} > W^{EE}/W^{IE} \\ W^{EI}(1 - p_E) &> W^{EE}p_E. \end{aligned} \tag{3.17}$$

If these conditions are satisfied, strong excitatory synaptic drive and strong inhibitory synaptic drive will dynamically balance producing a total synaptic drive that fluctuates around the spiking threshold. The resulting network state is called the balanced state.

### 3.3 Results

We have described a network model which has the potential to produce orientation tuned balanced network activity, similar to the spiking activity of V1, using weak structured connectivity embedded in strong background connectivity. There are a number of potential problems with our proposal which must be investigated. For example, strong background connectivity might completely overwhelm the contribution of weak structure. Also, orientation tuned cells might not be contrast invariant.

#### 3.3.1 Network simulations

We investigate the capacity of this network to produce orientation tuned responses and balanced network dynamics by measuring firing rates and spike trains of the neurons in a simulation of the network.

We simulate the network using the Gillespie Algorithm (Gillespie, 1977). This algorithm was originally developed for simulating stochastic discrete-state systems and was recently used to simulate weakly connected spiking neural networks (Benayoun et al., 2010). The algorithm is exact in that it provides spiking dynamics that are perfectly consistent with our network dynamics (Eqns. 3.3, 3.7).

The parameters that we use in our simulations are compiled in table 3.1. We do not need to fine-tune these parameters because the balanced state is very robust to parameter choice. Therefore, we use the same parameters throughout our work, unless stated otherwise.

$N$	500		$N_X$	400
$K$	100		$K_X$	160
$W^{EE}$	0.312		$W^{EX}$	0.65
$W^{IE}$	0.312		$W^{IX}$	0.56
$W^{II}$	3.37		$j_F$	5
$W^{EI}$	3.75		$c$	0.3
$p_E$	0.8		$\kappa$	0.775
$j_0$	1		$\theta_0$	$\pi/2$
$w_o$	1		$\sigma_F$	0.775
$\sigma_J$	0.775			

Table 3.1: Network parameters

We use these network parameters in our numerical calculations and simulations, unless otherwise stated. Connectivity parameters were chosen so that different input components, such as average excitatory input, average inhibitory input and average structured input are approximately the same size. Also, parameters are chosen so that the mean excitatory firing rate is given by  $\nu^E \simeq 0.2$  and the mean inhibitory firing rate is given by  $\nu^I \simeq 0.2$ , in the large network size limit. The relative sizes of the E-I population  $N$ , the input population  $N_x$ , the average number of synapses per neuron  $K$  and  $K_X$  are held constant throughout this work.

### 3.3.2 Balanced, irregular, asynchronous, spiking dynamics

We begin by demonstrating that our network is balanced and that resulting spike trains are irregular and asynchronous. A problem for some models of orientation tuning is that structured connectivity might disrupt the balance of excitation and inhibition. Indeed, this is a major difficulty for models that use strong structured connectivity (Lerchner et al., 2006; van Vreeswijk and Sompolinsky, 2005). We don't expect this to be a problem for our network because structured connectivity is weak.

We measure the excitatory and inhibitory synaptic drive to a randomly chosen neuron from our network (Fig. 3.3). These dynamically balance producing a total input that is close to the spiking threshold. This dynamic balancing produces irregular asynchronous spiking activity, as can be seen from a raster plot of network activity (Fig. 3.4). We will quantify this asynchronicity explicitly in chapter 4 by calculating spike train correlations numerically and analytically.

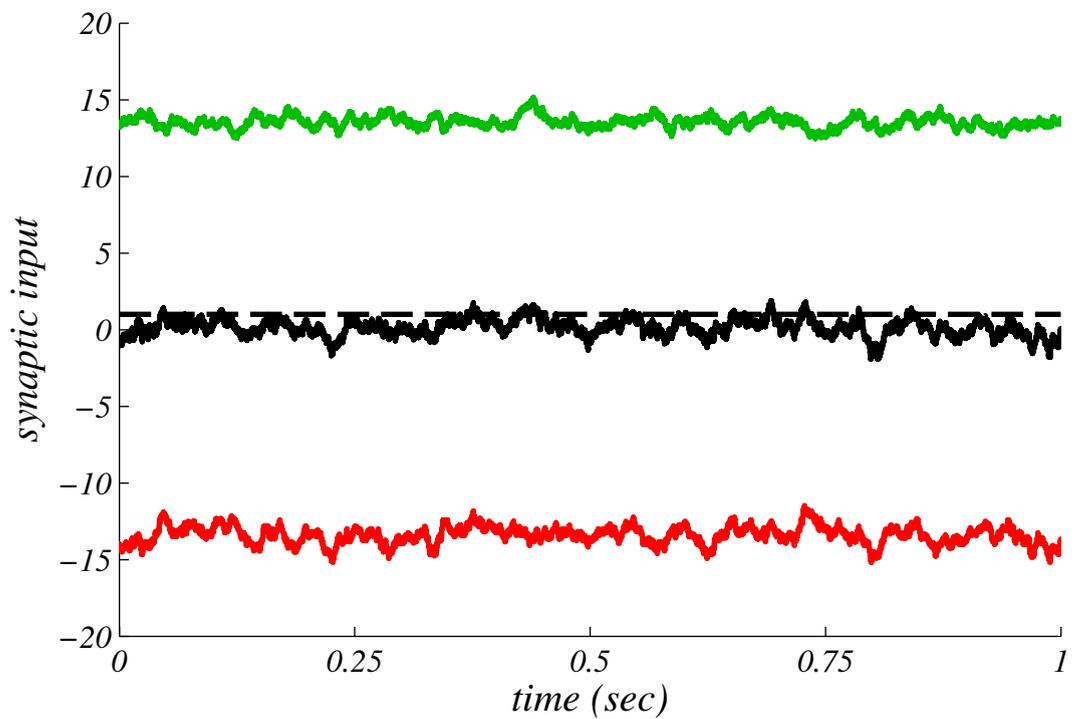


Figure 3.3: Balance of excitation and inhibition

In a balanced network, the excitatory input current (green) and inhibitory input current (red) are large in magnitude relative to the neural threshold (dashed line). However, these inputs balance dynamically so that the total input current (black) is close to the threshold. This dynamic balancing is noisy, so neurons spike irregularly, as in the cortex. The input and spike train is given for an excitatory neuron in a network with  $N = 10000$   $N_X = 8000$   $K = 2000$   $K_X = 3200$ .

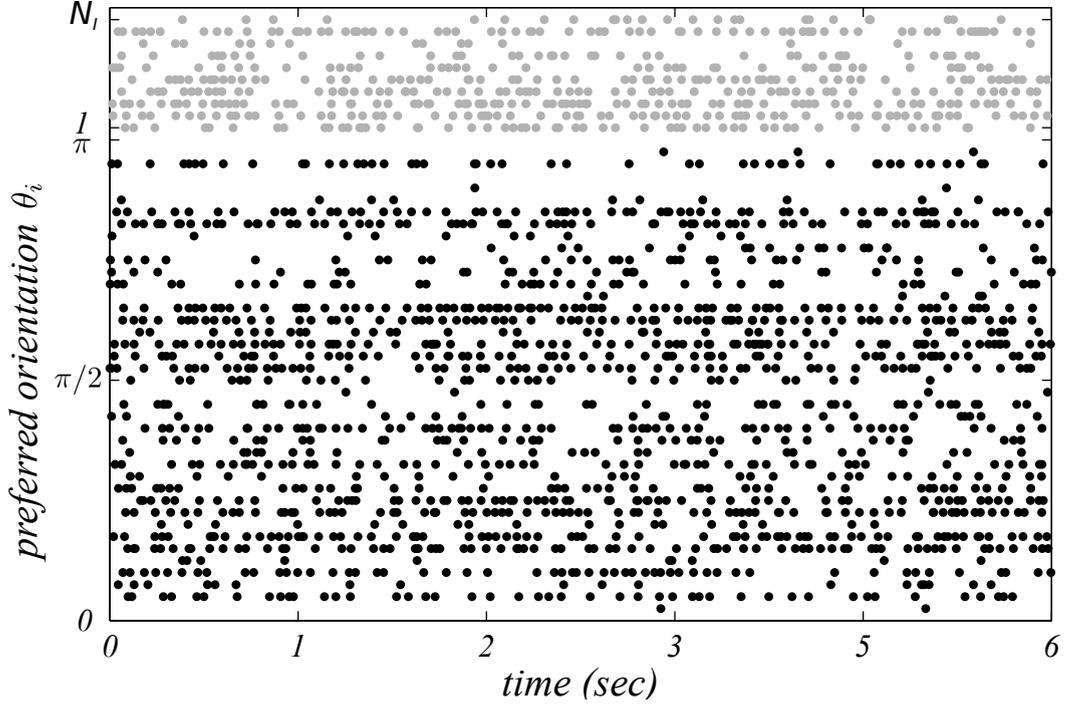


Figure 3.4: Irregular, asynchronous dynamics

In a balanced network, neurons are asynchronous and irregular. This raster plot shows spike trains from 50 example neurons. Excitatory neurons (black) have an bump of activity around  $\theta = \pi/2$  whereas inhibitory neurons (grey) do not have an activity bump. These bumps are partially obscured because of temporal irregularity and irregularity across the population. A spike here represents a transition from state 0 to state 1.

We can understand why neural activity is balanced and irregular by analysing the synaptic drive to the neurons in our network. We begin by writing the synaptic drive  $h_i(t)$  (Eqn. 3.6) in terms of the average synaptic drive  $m_i(t)$  and fluctuations about this average:

$$h_i(t) = m_i(t) + \sigma_i \xi_i, \quad (3.18)$$

where  $\xi_i$  represents  $\mathcal{O}(1)$  fluctuations, and

$$m_i(t) \equiv \langle h_i(t) \rangle \quad (3.19)$$

$$\sigma_i \equiv \sqrt{\langle (\delta h_i)^2 \rangle}, \quad (3.20)$$

with brackets  $\langle \dots \rangle$  denoting an average over different realisations of spiking network activity. This average is equivalent to a trial-average, but in a dream experiment, where an infinite number of trials are performed and network statistics are identical for each

trial.

We can calculate the mean synaptic drive, by taking an average of equation 3.6, leading to:

$$m_i(t) = \sum_{j \in E, I}^N A_{ij} \nu_j(t) + \sum_{j \in X}^{N_X} F_{ij} \nu_j(t) - \theta^Q \quad (3.21)$$

where

$$\nu_j(t) = \langle x_j(t) \rangle, \quad (3.22)$$

is the instantaneous firing rate of the  $j^{\text{th}}$  neuron.

Then, substituting our equations for network connectivity (Eqns. 3.10 and 3.14) into the mean synaptic drive equation (Eqn. 3.21), we obtain equations for the mean synaptic drive to an excitatory neuron  $m_i^E(t)$  and the mean synaptic drive to an inhibitory neuron  $m_i^I(t)$ :

$$\begin{aligned} m_i^E(t) &= \sqrt{K} (w_o W^{EE} p^E \nu^E(t) - w_o W^{EI} p^I \nu^I(t) + w_0 W^{EX} p^X \nu^X(t)) \\ &\quad + j_o \sum_{j \in E} J_{ij} \nu_j(t) p^E / N_E + j_F \sum_{j \in X} J_{ij}^F \nu_j(t) p^X / N_X \\ &\quad + \sum_{j \in E, I} \delta A_{ij} \nu_j(t) + \sum_{j \in X} \delta F_{ij} \nu_j(t) - \theta^E \end{aligned} \quad (3.23)$$

$$\begin{aligned} m_i^I(t) &= \sqrt{K} (w_o W^{IE} p^E \nu^E(t) - w_o W^{II} p^I \nu^I(t) + w_0 W^{IX} p^X \nu^X(t)) \\ &\quad + \sum_{j \in E, I} \delta A_{ij} \nu_j(t) + \sum_{j \in X} \delta F_{ij} \nu_j(t) - \theta^I, \end{aligned} \quad (3.24)$$

where

$$\nu^E(t) \equiv \sum_{j \in E} \nu_j(t) / N_E \quad (3.25)$$

$$\nu^I(t) \equiv \sum_{j \in I} \nu_j(t) / N_I \quad (3.26)$$

$$\nu^X(t) \equiv \sum_{j \in X} \nu_j(t) / N_X, \quad (3.27)$$

representing the mean excitatory, inhibitory and external population firing rates, re-

spectively, and

$$\begin{aligned}\delta A_{ij} &= A_{ij} - w_o W^{QR} p^R (-1)^{\delta_{RI}} \sqrt{K}/N_R - j_o J_{ij} p^E \delta_{QE} \delta_{RE}/N_E \\ \delta F_{ij} &= F_{ij} - w_o W^{QX} p^X \sqrt{K}/N_X - j_o J_{ij}^F p^X \delta_{QE}/N_X,\end{aligned}\quad (3.28)$$

represent connectivity noise, resulting from the random component of network connectivity (Eqns. 3.14, 3.10), where  $Q = E$  if  $i \in E$  and  $Q = I$  if  $i \in I$ , and similarly  $R = E$  if  $i \in E$  and  $R = I$  if  $i \in I$ . Also,  $p^I \equiv 1 - p^E$  and  $p^X = K_X/K = N_X/N$ . This connectivity noise is analogous to random differences in connectivity between different animals from the same species.

The contribution of the mean background network connectivity to the mean synaptic drive seems to be much larger than the contribution of connectivity noise and the contribution of structured connectivity. We can see this by calculating the approximate size of each term in equations 3.23 and 3.24:

$$\begin{aligned}j_o \sum_{j \in E} J_{ij} \nu_j(t) p^E / N_E &\sim \mathcal{O}(N_E/N_E) \\ j_F \sum_{j \in X} J_{ij}^F \nu_j(t) p^X / N_X &\sim \mathcal{O}(N_X/N_X).\end{aligned}\quad (3.29)$$

Both of these terms are  $\mathcal{O}(1)$ , much smaller than the contribution of the mean background network connectivity terms which are  $\mathcal{O}(\sqrt{K})$  (Eqns. 3.23, 3.24).

Similarly, the contribution of connectivity noise is  $\mathcal{O}(1)$ :

$$\begin{aligned}\sum_{j \in E, I} \delta A_{ij} \nu_j(t) &\sim \mathcal{O}(\sqrt{N}/\sqrt{N}) \\ \sum_{j \in X} \delta F_{ij} \nu_j(t) &\sim \mathcal{O}(\sqrt{N_X}/\sqrt{N_X}),\end{aligned}\quad (3.30)$$

because  $\delta A_{ij} \sim \mathcal{O}(1/\sqrt{N})$ ,  $\delta F_{ij} \sim \mathcal{O}(1/\sqrt{N_X})$  and according to the central limit theorem, the sum of  $N$  zero-mean uncorrelated random variables is  $\mathcal{O}(\sqrt{N})$ .

These equations seem to indicate that the contribution of background connectivity is  $\sqrt{K}$  times larger than the spiking thresholds  $\theta^E, \theta^I \sim \mathcal{O}(1)$ . In the cortex,  $K \sim 2000$  (Braitenberg and Schuz, 1991; Binzegger et al., 2004) so the contribution of background may be excessively large, producing pathological dynamical states in which neurons are

completely silent or spiking incessantly. In balanced networks, this does not happen because strong excitatory and inhibitory input dynamically balance:

$$w_o W^{EE} p^E \nu^E(t) - w_o W^{EI} p^I \nu^I(t) + w_o W^{EX} p^X \nu^X(t) \sim \mathcal{O}\left(1/\sqrt{K}\right) \quad (3.31)$$

$$w_o W^{IE} p^E \nu^E(t) - w_o W^{II} p^I \nu^I(t) + w_o W^{IX} p^X \nu^X(t) \sim \mathcal{O}\left(1/\sqrt{K}\right). \quad (3.32)$$

This dynamic balancing is possible because  $W^{EE}$ ,  $W^{EI}$ ,  $W^{IE}$ ,  $W^{II}$ ,  $W^{EX}$  and  $W^{IX}$  satisfy the balance conditions (Eqn. 3.17) (van Vreeswijk and Sompolinsky, 1998).

Therefore, by inserting equation 3.32, 3.31, 3.30 and 3.29 into equation 3.23 we see that the approximate size of the mean synaptic drive:

$$m_i(t) \sim \mathcal{O}\left(\sqrt{K}/\sqrt{K}\right) \sim \mathcal{O}(1). \quad (3.33)$$

This is about the same size as the spiking thresholds  $\theta^E, \theta^I \sim \mathcal{O}(1)$ . Therefore, pathological dynamical states can be avoided. Structured connectivity does not disrupt this balance. Indeed, the balance equations for a network with background connectivity (Eqns. 3.31) are the same as the balance equations for a network without structured connectivity.

Next, we demonstrate that an irregular, asynchronous spiking state is consistent with our dynamical equations. We do this by calculating the size of synaptic drive fluctuations, represented by the synaptic drive standard deviation  $\sigma_i$  (Eqn. 3.20). If these fluctuations are the same size as the mean synaptic drive then the synaptic drive will often cross the spiking threshold stochastically and spiking dynamics will be irregular.

We begin by writing the synaptic drive standard deviation in terms of the variance of neural spike trains:

$$\sigma_i^2 \simeq \sum_{j \in E}^{N_E} A_{ij}^2 \rho_{jj} + \sum_{j \in I}^{N_I} A_{ij}^2 \rho_{jj} + \sum_{j \in X}^{N_X} F_{ij}^2 \rho_{jj}, \quad (3.34)$$

where

$$\rho_{ii} \equiv \left\langle (\delta x_i)^2 \right\rangle, \quad (3.35)$$

is the variance of the  $i^{th}$  neuron. Here we have neglected the contribution of synaptic drive correlations, because in balanced networks, synaptic drive correlations balance dynamically in the same way that the mean synaptic drive balances dynamically (Renart

et al., 2010) (Chapter 4).

Now, substituting our equations for network connectivity (Eqns. 3.14, 3.10) into equation 3.34 we can write:

$$\begin{aligned} \sigma_i^2 = & (w_o W^{EE})^2 p^E \sum_{j \in E} \rho_{jj}/N_E + (w_o W^{EI})^2 p^I \sum_{j \in I} \rho_{jj}/N_I \\ & + (w_o W^{EX})^2 p^X \sum_{j \in X} \rho_{jj}/N_X + \mathcal{O}\left(1/\sqrt{K}\right), \end{aligned} \quad (3.36)$$

where the  $\mathcal{O}(1/\sqrt{K})$  term represents the contribution of connectivity noise and structured connectivity. Spiking dynamics are irregular, so  $\rho_{ii} \sim \mathcal{O}(1)$  for all neurons  $i \in E, I, X$ . Therefore synaptic drive fluctuations are about the same size as the mean synaptic drive (Eqn. 3.33):

$$\sigma_i \sim \mathcal{O}(1). \quad (3.37)$$

This analysis explains how spike trains in a balanced network with weak structured connectivity can be irregular and asynchronous. Synaptic drive fluctuations are about the same size as the mean synaptic drive, so the total synaptic drive will often cross the spiking threshold stochastically, producing irregular spike trains. If the network only had structured connectivity, then the synaptic drive fluctuations would be very small ( $\sigma_i \sim \mathcal{O}(1/K)$ ) and spike trains would be highly regular for most network inputs.

### 3.3.3 Orientation tuning

We have proposed that weak structured connectivity embedded in strong background connectivity can produce orientation tuned, balanced network activity. We demonstrate this by simulating our network and measuring the firing rates of all neurons. For computational reasons, we assume ergodicity and estimate firing rates across time. All statistical estimates converge towards the true statistics when simulations and measurements become infinitely long.

We find that the neurons in our network do have orientation tuning (Fig. 3.5). The orientation representation is noisy across the population. Similar population noise is also observed in experiments (Fregnac and Imbert, 1984; Blasdel and Fitzpatrick, 1984; Chapman and Stryker, 1993; Maldonado, 1997; Ferster and Miller, 2000; Olshausen and Field, 2005). The inhibitory population has no orientation preference (Fig. 3.5).

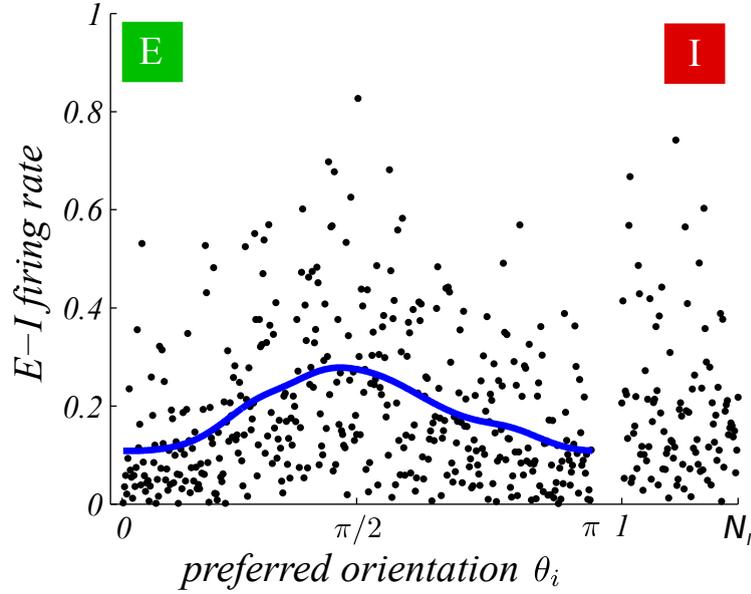


Figure 3.5: Orientation tuning in balanced networks

The equilibrium firing rates of neurons in the excitatory population (E) and inhibitory population (I) are shown. The population response is noisy so we show a smoothed version of the firing rate (blue). Smoothing is performed by convolving the population firing rate with a Gaussian function with standard deviation  $\pi/16$ . Firing rates are estimated during an 11 second simulation for a network with parameters given by table 3.1.

We can understand why background connectivity does not overwhelm the contribution of structured connectivity by inserting our dynamic balance equations (Eqns. 3.31 and 3.32) into our equation for mean synaptic drive (Eqn. 3.23):

$$m_i(t) = w_0 \mathcal{O}(1) + j_o \sum_{j \in E} J_{ij} \nu_j(t) p^E / N_E + j_F \sum_{j \in X} J_{ij}^F \nu_j(t) p^X / N_X - \theta^E . \quad (3.38)$$

The contribution of structured connectivity and background connectivity to the mean synaptic drive are about the same size. Therefore, structured connectivity that matches the structure of the network input can produce orientation tuned responses.

However, dynamic balance is not exact. By pure chance, some neurons receive more synaptic input than others. This is called *quenched noise*, because it is frozen into the network and does not disappear over time like dynamical noise. Quenched noise manifests itself as tuning curve inhomogeneity. It explains why the orientation tuning in our model is noisy and may explain some of the inhomogeneity observed in visual cortex tuning curves (Fregnac and Imbert, 1984; Maldonado, 1997; Ferster and Miller,

2000; Olshausen and Field, 2005).

Inhibitory neurons do not have orientation tuning because they do not have structured connectivity in our model. In theory, there is no reason why inhibitory neurons cannot have weak structured connectivity so this is something that can be included in extensions of our model.

### 3.3.4 Contrast invariance

In the visual cortex, when the contrast of the stimulus is increased, the response of orientation tuned cells increases linearly (Sclar and Freeman, 1982). This is called contrast invariance because the shape of the overall population response is invariant to contrast changes. Any plausible model of visual cortex responses must have contrast invariance, along with orientation tuning.

We investigate the contrast dependence of our model by measuring firing rates in a series of network simulations, each one receiving inputs at different contrast levels. We find that networks with background connectivity are contrast invariant, whereas networks without background connectivity are not (Fig. 3.6).

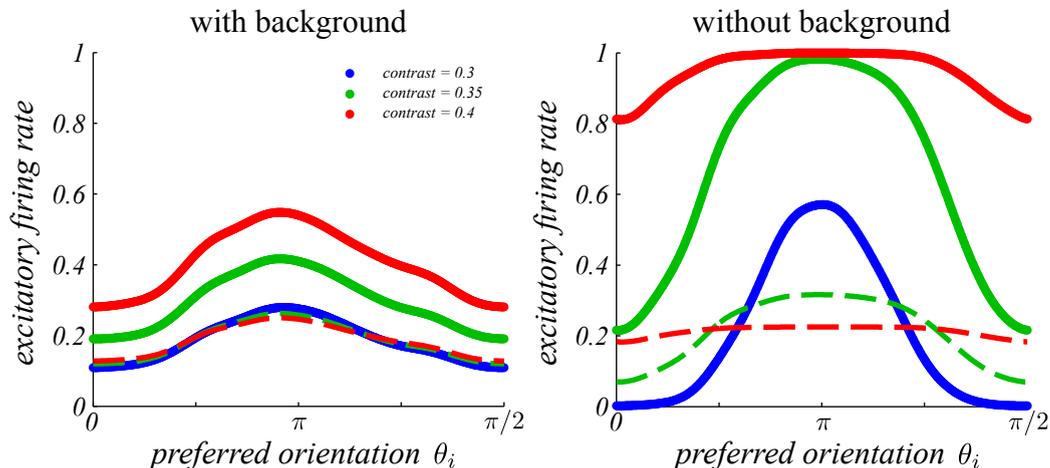


Figure 3.6: Contrast invariance

Networks with background connectivity (left,  $w_0 = 1$ ) are contrast invariant, whereas networks without background connectivity (right,  $w_o = 1/\sqrt{100 \times K}$ ) are not contrast invariant. The equilibrium firing rates of the excitatory population receiving input at three different contrasts (solid lines) are shown. Rescaled firing rates are calculated to test contrast invariance (dashed lines). The population response is noisy so we show a smoothed version of the population firing rate. Smoothing is performed by convolving the population firing rate with a Gaussian function with standard deviation  $\pi/16$ . Firing rates are estimated during an 10 second simulation.

Networks without background connectivity are not contrast invariant, because the relationship between neural firing rates and input contrast is highly non-linear.

In balanced networks, background connectivity approximately linearises the relationship between the mean input and output. This linearisation is exact for the mean population firing rates in large networks, as we can see by solving the dynamic balance equations (Eqns. 3.31):

$$\nu^E(t) = c^E \nu^X(t) + \mathcal{O}\left(1/\sqrt{K}\right) \quad (3.39)$$

$$\nu^I(t) = c^I \nu^X(t) + \mathcal{O}\left(1/\sqrt{K}\right), \quad (3.40)$$

where

$$c^E = (W^{EI}W^{IX} - W^{II}W^{EX}) / (W^{EE}W^{II} - W^{EI}W^{IE}) \times p^X/p^E \quad (3.41)$$

$$c^I = (W^{EE}W^{IX} - W^{IE}W^{EX}) / (W^{EE}W^{II} - W^{EI}W^{IE}) \times p^X/p^I. \quad (3.42)$$

These equations show that the E-I population tracks the input population firing rate with great precision. The input firing rates are proportional to contrast (Eqn. 3.8). Therefore, the mean E-I population firing rate is proportional to contrast. Incidentally, this tracking does not depend on structured connectivity.

### 3.4 Discussion

Cortical activity can be irregular and uniform simultaneously - from the uniformity of orientation tuning and contrast invariance (Hubel and Wiesel, 1962; Sclar and Freeman, 1982) to the irregularity of Poisson-like spike trains and synaptic background noise (Burns and Webb, 1976; Softky and Koch, 1993; Bair et al., 1994). The coexistence of these opposites prompts us to ask: how can network connectivity produce such wildly disparate phenomena?

We find that weak structured connectivity embedded in a strong background connectivity can solve this problem. Structured connectivity produces orientation tuning and background connectivity is responsible for irregular spiking. This orientation tuning is contrast invariant, because structured connectivity provides contrast invariant synaptic input, and background connectivity allows firing rates to adopt this contrast

invariance, by linearising single neuron dynamics. Background connectivity also causes orientation tuning inhomogeneity, by introducing quenched noise.

An alternative to our proposal is that strong structured connectivity produces contrast invariant orientation tuning (Lerchner et al., 2006). One problem with that proposal is that learning may be difficult, because the balance conditions require inhibitory synaptic plasticity to match excitatory synaptic plasticity.

Testing theories of cortical connectivity is difficult, because measuring cortical connectivity directly is difficult (DeFelipe, 2010). This is especially true when structure is weak, because the strong background component of connectivity will mask the structured connectivity that we want to measure. Nonetheless, recent observation of promiscuous inhibitory connectivity in the mouse cortex suggests that structure may be weak (Fino and Yuste, 2011; Hofer et al., 2011), because if structure is strong, then inhibitory connectivity must be structured so that it can balance the strong structured excitation. Also, the historical failure to discover precisely wired cortical circuits is in itself consistent with our proposal, because it suggests that structured connectivity is hidden. In future work, it may be possible to infer the presence of weak structured connectivity by systematically perturbing connection strengths by small amounts according to preferred orientation and measuring the effect on orientation tuning.

There is plenty of indirect evidence for our proposed connectivity model, apart from the observation of orientation tuning and contrast invariance. First of all, the facilitation of contrast invariance using synaptic background noise has been observed in anaesthetised cats (Anderson et al., 2000). This synaptic background noise was also found to have a very weak dependence on orientation, just like the synaptic background noise in our model (Eqn. 3.36). This is inconsistent with the synaptic background noise in a network with strong structure, which is strongly dependent on orientation (Lerchner et al., 2006). Further indirect support for our proposal is provided by the widely ignored observations of *‘poorly tuned’* cells in the visual cortex (Blasdel and Fitzpatrick, 1984; Chapman and Stryker, 1993; Olshausen and Field, 2005), which is consistent with the quenched noise caused by strong background connectivity. An explicit comparison of observed orientation tuning inhomogeneity with orientation tuning inhomogeneity in our model would provide another interesting indirect test of our model.

Apart from orientation tuning, contrast invariance and spike train irregularity, spike

trains in V1 are also asynchronous (Ecker et al., 2010; Renart et al., 2010). Therefore, spike trains in our network model must also be asynchronous. In the next chapter (Chapter 4), we calculate the size of correlations in our network. We will study the computational performance of this network in chapter 5.

We have presented our connectivity model as an orientation hyper-column model. However, there are many differences between this model and a real orientation hyper-column. For example, the probability that two excitatory neurons connect depends on the preferred orientation (Ko et al., 2011), whereas neurons in our model connect promiscuously. It is possible that this observed connectivity may be a form of weak structure, different to the type of weak structure that we consider, but broadly consistent with our proposal. In future work it would be interesting to study networks with weak structured connectivity of this form.

In truth, the comparison to an orientation hyper-column does not reflect our primary interest. Our goal is to understand how an arbitrary piece of cortex can be tuned to an arbitrary input feature and have irregular asynchronous spiking activity. Our simple model allowed us to distill the essential properties of feature tuning, whereas an excessively complex model may have been intractable. A combination of weak structure embedded in strong background connectivity may unify a wide variety of cortical phenomena across a wide range of cortical regions.



## Chapter 4

# Spike-train Correlations in Balanced Networks

### 4.1 Introduction

Spike train correlations play an important role in cortical computation and dynamics (Averbeck et al., 2006). For example, the correlation between pre-synaptic and post-synaptic spiking activity is important in synaptic plasticity and learning (Hebb, 1949). Also, the capacity of a network to encode and decode information depends on correlations (Wu et al., 2001; Nirenberg and Latham, 2003), and, the measurement of correlations might be useful as a practical experimental tool for inferring network connectivity, if, the relationship between correlation and connectivity can be understood (Abeles, 1991). In this chapter, we study spike-train correlations in a balanced network.

Despite their importance for many cortical functions, correlations are poorly understood (Abbott and Dayan, 1999). The main reason for this is that correlations are very difficult to measure (Ecker et al., 2010). Small movements of recording electrodes, errors in single cell isolation from multi-electrode recordings and oscillations induced by anaesthetics all create spurious correlation measurements. When accurate measurements are available, the quantity of correlation pairs is not large enough to quantify the relationship between computation and correlation.

Rather than attempting to resolve these experimental problems, we study correlations analytically. We calculate the correlation between the spike-trains of all pairs of neurons in a balanced network. These calculations are complicated because neurons in

balanced networks interact strongly. By performing a series of biologically reasonable approximations we obtain mathematical expressions that relate correlations directly to network connectivity and input. This is the first full prediction of correlations in balanced networks and as such, it is an important step forward in our understanding of cortical network dynamics.

We can use our mathematical expressions to understand the relationship between correlations and structured connectivity. If structured connectivity does not produce substantial correlations, Hebbian learning may be impossible and attempts to infer structured connectivity using correlation measurements might also fail. We find that structured connectivity does increase correlations substantially. However, the contribution of background connectivity to correlations is on the same order of magnitude as the contribution of structured connectivity. This result is not encouraging for experimentalists who are attempting to use correlations as a proxy for connectivity, because it will be extremely difficult to infer structured connectivity separately from random connectivity using correlations. However, there is hope for Hebbian plasticity, which might still work, but in a background of quenched noise.

We also calculate the typical size of correlations in balanced networks with structure. The size of correlations is important because it can determine whether or not correlations play a role in computation. We find that the typical size of correlations is 0.02. This is very weak, though it is stronger than the typical size of correlations in networks without structure (Renart et al., 2010). These predictions match recent measurements from the visual cortex of awake monkeys viewing oriented gratings, in which correlations were about  $0.005 \pm 0.004$  (Ecker et al., 2010).

## 4.2 Correlation analysis

Our goal is to derive a mathematical expression for the correlation  $C_{ij}$  between a pair of neurons,  $i$  and  $j$  in a balanced network. We will relate the correlation coefficient  $C_{ij}$  directly to recurrent network connectivity  $\mathbf{A}$  (Eqn. 3.10), feedforward network connectivity  $\mathbf{F}$  (Eqn. 3.14) and the mean firing rate of the input population  $\mathbf{u}(\theta)$  (Eqn. 3.8):

$$C_{ij} = f_{ij}(\mathbf{A}, \mathbf{F}, \mathbf{u}(\theta)) . \quad (4.1)$$

We calculate correlations in the same network model that we introduced in chapter 3. Network connectivity consists of weak structured connectivity, parametrised by  $j_o$  and  $j_F$  embedded in strong background connectivity, parametrised by  $w_o$  (Eqns. 3.10, 3.14). Neurons in the network model are not very realistic. However, because much of the spiking activity in the cortex is the result of strong connectivity, we can still have realistic irregular asynchronous spike trains and orientation tuning, similar to neurons in V1 (Hubel and Wiesel, 1959; Softky and Koch, 1993; Ecker et al., 2010).

Calculating correlations in a strongly interacting network is difficult, and requires knowledge of dynamical systems theory, differential equations and linear algebra. Readers may prefer to skip our calculation and jump to the result, which appears in equation 4.46. However, we do not recommend this, as much insight into neural network dynamics may be acquired by following our calculation, which is, in itself an important result. With this in mind, we have relegated many of the less interesting technical components of our calculation to the methods section, where interested readers can find all the details.

Our calculation is in three parts:

1. We calculate the firing rates of neurons, by exploiting the fact that in balanced networks the synaptic drive is approximately Gaussian (Rudolph et al., 2005).
2. We derive a dynamical equation for the covariance between spike trains, by exploiting the fact that in balanced networks spike trains are irregular (Burns and Webb, 1976; Softky and Koch, 1993; Bair et al., 1994) and asynchronous (Ecker et al., 2010; Renart et al., 2010).
3. Finally, we solve the firing rate dynamical equations and covariance dynamical equations at equilibrium, and use the solutions to obtain an expression for balanced network correlations  $C_{ij}$ .

We will test our predictions by comparing them to simulation measurements, before answering questions about the origin of correlations.

### 4.2.1 Firing rate equations

We begin our analysis by deriving a dynamical equation for the firing rates of neurons in our network. As before, the firing rate that we calculate is defined as

$$\nu_i \equiv \langle x_i \rangle , \quad (4.2)$$

where brackets denote an average over all neuron states at time  $t$ :

$$\langle f(\mathbf{x}) \rangle \equiv \sum_{\mathbf{x}} f(\mathbf{x}) P(\mathbf{x}, t) . \quad (4.3)$$

where  $P(\mathbf{x}, t)$  denotes the probability that the network is in state  $\mathbf{x}$  at time  $t$  and  $f(\mathbf{x})$  is an arbitrary function of  $\mathbf{x}$ , where  $\mathbf{x} = (x_1, x_2, \dots, x_i, \dots, x_{N+N_X})$ . This firing rate is equivalent to a trial-averaged firing rate, but in an ideal experiment, where an infinite number of trials are performed and experiment conditions do not change, apart from random fluctuations.

We can derive a dynamical equation for firing rates using the equations for the dynamics of individual neurons (Eqns. 3.3, 3.7). Individual neurons are noisy, and so, the equation that describes the spiking state of the network is probabilistic. This equation, known as the Master Equation is given by ( Supplementary Methods 4.5.1) (Glauber, 1963; Ginzburg and Sompolinsky, 1994):

$$\frac{dP(\mathbf{x}, t)}{dt} = - \sum_i r(x_i \rightarrow 1 - x_i) P(\mathbf{x}, t) + \sum_i r(1 - x_i \rightarrow x_i) P(\mathbf{x}^{(i)}, t) . \quad (4.4)$$

Here,  $P(\mathbf{x}^{(i)}, t)$  denotes the probability that the network is in state  $\mathbf{x}^{(i)}$  at time  $t$ , where  $\mathbf{x}^{(i)} = (x_1, x_2, \dots, 1 - x_i, \dots, x_{N+N_X})$ . If neuron  $i$  is in the E-I population, the transition rate  $r(x_i \rightarrow 1 - x_i)$  is given by equation 3.3. If neuron  $i$  is in the input population, the transition rate is given by equation 3.7.

Combining the Master Equation (Eqn. 4.4) with our firing rate definition (Eqn. 4.2) we can easily derive dynamical equations for the firing rates (Supplementary Methods

4.5.2) (Ginzburg and Sompolinsky, 1994):

$$\tau_0 \frac{d}{dt} \nu_i = -\nu_i + \langle f(h_i) \rangle \quad i \in E, I; \quad (4.5)$$

$$\tau_0 \frac{d}{dt} \nu_i = -\nu_i + u_i \quad i \in X, \quad (4.6)$$

where  $f(h_i)$  is a sigmoidal function of  $h_i$ , the synaptic drive to the  $i^{\text{th}}$  neuron (Eqn. 3.6) and  $u_i$  determines how the input population encodes orientation (Eqn. 3.8).

The firing rate equations for neurons in the input population X can be solved easily, because they are uncoupled linear differential equations, whereas the equations for neurons in the E-I population are more complicated, because of the term  $\langle f(h_i) \rangle$ . This term is a complicated average over all neuron states  $\mathbf{x}$  (Eqn. 3.6, 4.3). However, we can simplify this if we rewrite it as an average over total synaptic input  $h_i$ :

$$\langle f(h_i) \rangle = \sum_{\mathbf{x}} f(h_i) P(\mathbf{x}, t) = \int f(h_i) p(h_i, t) dh_i. \quad (4.7)$$

In a sense, this step moves the calculation from the axon to the soma.

To calculate this new average, we must identify the synaptic drive probability distribution  $p(h_i, t)$ . In a balanced network,  $h_i$  is approximately Gaussian (Fig. 4.1), similar to cortical input (Destexhe et al., 2003; Rudolph et al., 2005). Mathematically, this is a consequence of the fact that  $h_i$  is the sum of a large number of weakly correlated random variables and the central limit theorem states that the sum of a large number of independent random variables is approximately Gaussian. Therefore we can write

$$p(h_i) = \frac{1}{(2\pi\sigma_i^2)^{1/2}} \exp \left[ -\frac{(h_i - m_i)^2}{2\sigma_i^2} \right], \quad (4.8)$$

where  $m_i \equiv \langle h_i \rangle$  is the mean synaptic drive (Eqn. 3.21) and  $\sigma_i^2 \equiv \langle (\delta h_i)^2 \rangle$  is the synaptic drive variance (Eqn. 3.34).

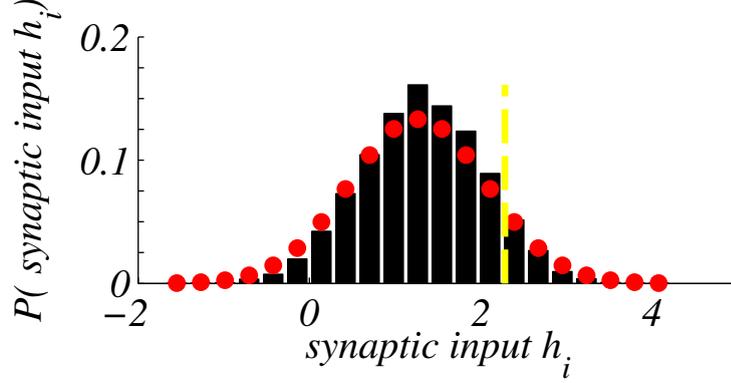


Figure 4.1: Gaussian input

The synaptic input current to a neuron in a balanced network is Gaussian distributed close to the neural threshold (blue line). The predicted input distribution (red) matches the measured distribution (black). The input current is measured for an excitatory neuron during a 10 second simulation.

We can relate the mean synaptic drive to the firing rate by taking an average of equation 3.6:

$$m_i = \sum_{j \in E, I}^N A_{ij} \nu_j + \sum_{j \in X}^{N_X} F_{ij} \nu_j - \theta^Q, \quad (4.9)$$

where  $A_{ij}$  is the strength of the synaptic connection from neuron  $j$  to neuron  $i$ ,  $F_{ij}$  is the strength of the synaptic connection from input neuron  $j$  to neuron  $i$ , and  $\theta^Q$  is the spiking threshold, where  $Q \in \{E, I\}$ .

A similar calculation gives us the variance (Supplementary Methods 4.5.2):

$$\sigma_i^2 \simeq \sum_{j \in E, I} A_{ij} \nu_j (1 - \nu_j) A_{ji}^T + \sum_{j \in X} F_{ij} \nu_j (1 - \nu_j) F_{ji}^T, \quad (4.10)$$

where  $\mathbf{A}^T$  denotes the transpose of  $\mathbf{A}$ .

We can now solve the integral in equation 4.7. For the particular case where  $f(x) = \Theta(x)$ , the Heaviside function, this integral yields the following non-linear differential equations for the firing rates:

$$\tau_0 \frac{d\nu_i}{dt} = -\nu_i + \Phi\left(\frac{m_i}{\sigma_i}\right) \quad i \in E, I, \quad (4.11)$$

where  $\Phi$  is the cumulative normal function (Supplementary Methods 4.5.2),

$$\Phi(x) \equiv \int_{-\infty}^x dz \frac{\exp[-z^2/2]}{(2\pi)^{1/2}}. \quad (4.12)$$

Our goal is to calculate correlations at equilibrium, and so, we must calculate firing rates at equilibrium:

$$\nu_i = \Phi\left(\frac{m_i}{\sigma_i}\right) \quad i \in E, I \quad (4.13)$$

$$\nu_i = u_i \quad i \in X. \quad (4.14)$$

These equations are non-linear but they can be solved numerically using the Newton–Raphson method (Supplementary Methods 4.5.2). We will discuss these solutions later in this chapter.

## 4.2.2 Correlation equations

We can now begin our correlation analysis. Our goal is to derive an expression for the correlation between a pair of neurons

$$C_{ij} = \rho_{ij} / \sqrt{\rho_{ii}\rho_{jj}}, \quad (4.15)$$

where

$$\rho_{ij} = \langle \delta x_i \delta x_j \rangle. \quad (4.16)$$

This analysis is more complicated than the firing rate derivation, though some of the steps are similar.

First, we calculate the diagonal elements of  $\rho$ . Using the fact that  $x_i^2 = x_i$ , we find:

$$\rho_{ii} \equiv \langle (\delta x_i)^2 \rangle = \langle x_i^2 \rangle - \langle \nu_i^2 \rangle = \nu_i - \nu_i^2. \quad (4.17)$$

Next we calculate the off-diagonal elements of  $\rho$ . This calculation is more complicated. Similar to the firing rate derivation, we can combine the Master Equation for our network dynamics with our covariance definition. This results in the following dynamical equation for the equal-time covariance between neurons (Supplementary Methods 4.5.3) (Ginzburg and Sompolinsky, 1994):

$$\tau_0 \frac{d\rho_{ij}}{dt} = -2\rho_{ij} + \langle \delta x_i \delta f_j \rangle + \langle \delta f_i \delta x_j \rangle, \quad (4.18)$$

where  $i, j \in E, I$  and  $i \neq j$ . Here we have used the notation  $\delta f_i = f(h_i) - \langle f(h_i) \rangle$ .

This expression is complicated because of the  $\langle \delta x_i \delta f_j \rangle$  term and the  $\langle \delta f_i \delta x_j \rangle$  term. We had to deal with a similar term in our firing rate calculation (Eqn. 4.7). In that calculation, we replaced an average over  $\mathbf{x}$  with an average over  $\mathbf{h}$ . Here, we would like to do something similar. However,  $\langle \delta x_i \delta f_j \rangle$  and  $\langle \delta f_i \delta x_j \rangle$  contain terms that depend on both  $\mathbf{x}$  and  $\mathbf{h}$ , so simply replacing the average over  $\mathbf{x}$  with an average over  $\mathbf{h}$  is not useful.

We must transform  $\langle \delta x_i \delta f_j \rangle$  into a more tractable form before we take an average over  $\mathbf{h}$ . We achieve this by multiplying  $\delta \mathbf{x}$  by a *pseudo connectivity matrix*  $\mathbf{H}$  to produce a *pseudo synaptic-drive*  $\delta h'_i = \sum_j H_{ij} \delta x_j$ , where  $\mathbf{H}$  is chosen so that  $(\delta h'_i, \delta h_j)$  can be treated as a weakly correlated zero mean two-dimensional Gaussian random variable. A matrix of this form will exist if the network is balanced because spike trains in balanced networks are asynchronous (Ecker et al., 2010) and the sum of a large number of uncorrelated random variables can be treated as a Gaussian random variable, by the central limit theorem. We also require  $\mathbf{H}$  to be invertible, for reasons that shall become clear shortly.

Now, using  $\mathbf{H}$ , we can transform the problematic  $\langle \delta x_i \delta f_j \rangle$  terms into the following form:

$$\langle \delta h'_i \delta f_j \rangle = \sum_k H_{ik} \langle \delta x_k \delta f_j \rangle . \quad (4.19)$$

If we replace the average over  $\mathbf{x}$  with an average over  $h'_i$  and  $h_j$ , this becomes tractable:

$$\langle \delta h'_i \delta f_j \rangle = \sum_{\mathbf{x}} \delta h'_i \delta f(h_j) P(\mathbf{x}, t) = \int \delta h'_i \delta f(h_j) p(h'_i, h_j, t) dh'_i dh_j . \quad (4.20)$$

Given our definition of  $\delta h'_i$ , we can treat the synaptic drive distribution  $p(h'_i, h_j, t)$  as a two-dimensional Gaussian with mean  $(m'_i, m_j)$  and covariance  $\mathbf{\Gamma}$ :

$$\mathbf{\Gamma} \equiv \begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} \\ \Sigma_{ji} & \Sigma_{jj} \end{pmatrix} \quad (4.21)$$

where

$$\Sigma_{ij} \equiv \langle \delta h'_i \delta h_j \rangle . \quad (4.22)$$

This integral is still quite complicated. However, it can be greatly simplified if we perform a Taylor series expansion of  $p(h'_i, h_j, t)$ , assuming  $\Sigma_{ij} \ll \Sigma_{ii} \Sigma_{jj}$  (Supplementary

Methods 4.5.2):

$$p(h'_i, h_j, t) = \left( 1 + \delta h'_i \delta h_j \left( \frac{\Sigma_{ij}}{\Sigma_{ii} \Sigma_{jj}} \right) + \mathcal{O} \left( \left( \frac{\Sigma_{ij}}{\Sigma_{ii} \Sigma_{jj}} \right)^2 \right) \right) p(h'_i, t) p(h_j, t) . \quad (4.23)$$

We can ignore higher order terms in this series, because in balanced networks, synaptic drive fluctuations  $\Sigma_{ii}$  are large (Rudolph et al., 2005) and we can make the ansatz that synaptic drive fluctuations  $\Sigma_{ij}$  are small, because the spike trains that they produce are weakly correlated (Ecker et al., 2010). We will check that this ansatz produces self-consistent correlation equations later.

To first order, we can now solve the integral in equation 4.20 (Supplementary Methods 4.5.2):

$$\begin{aligned} & \int \delta h'_i \delta f(h_j) p(h'_i, h_j, t) dh'_i dh_j \simeq \\ & \simeq \int \delta h'_i \delta f(h_j) \left( 1 + \frac{\delta h'_i \delta h_j}{\Sigma_{ii} \Sigma_{jj}} \Sigma_{ij} \right) p(h'_i, t) p(h_j, t) dh'_i dh_j = \Sigma_{ij} g_{jj} , \end{aligned} \quad (4.24)$$

where

$$g_{ij} = \delta_{ij} \frac{\partial \nu_j}{\partial m_j} = \delta_{ij} \frac{\exp[-m_j^2/2\sigma_j^2]}{(2\pi\sigma_j^2)^{1/2}} . \quad (4.25)$$

We call  $g_{ij}$  the *gain*. It represents the sensitivity of a neuron's firing rate to changes in the average synaptic drive to that neuron. The neuron gain shall appear again in various calculations throughout the remainder of this thesis.

Only a few simple steps now remain in our derivation. First of all we must relate  $\Sigma$  to  $\rho$  (Eqn. 4.26). This is straightforward

$$\begin{aligned} \Sigma &= \langle \delta \mathbf{h}' \otimes \delta \mathbf{h} \rangle \\ &= \sum_{\mathbf{x}} (\mathbf{H} \cdot \delta \mathbf{x}) \otimes (\mathbf{A} \cdot \delta \mathbf{x} + \mathbf{F} \cdot \delta \mathbf{x}) p(\mathbf{x}, t) \\ &= \mathbf{H} \rho \mathbf{A}^T + \mathbf{H} \mathbf{r}^T \mathbf{F}^T , \end{aligned} \quad (4.26)$$

where  $\mathbf{r}$  represents the covariance between the spiking activity of neurons in the E-I population and neurons in the input population X:

$$r_{ij} \equiv \langle \delta x_i \delta x_j \rangle \quad i \in X \quad j \in E, I . \quad (4.27)$$

Now, inserting this (Eqn. 4.26) into equation 4.24 we can obtain a simple expression for our troublesome term:

$$\langle \delta \mathbf{x} \otimes \delta \mathbf{f} \rangle = \mathbf{H}^{-1} (\mathbf{H} \boldsymbol{\rho} \mathbf{A}^T + \mathbf{H} \mathbf{r}^T \mathbf{F}^T) \mathbf{g} = \boldsymbol{\rho} \mathbf{A}^T \mathbf{g} + \mathbf{r}^T \mathbf{F}^T \mathbf{g}. \quad (4.28)$$

We also have to transform  $\langle \delta f_i \delta x_j \rangle$  into a tractable form. Similar to the  $\langle \delta x_i \delta f_j \rangle$  calculation, we can multiply  $\langle \delta f_i \delta x_j \rangle$  by a pseudo connectivity matrix, but on the right hand side this time. The resulting expression can be simplified using a Taylor series expansion, as before, leading to:

$$\langle \delta \mathbf{f} \otimes \delta \mathbf{x} \rangle = \mathbf{g} \mathbf{A} \boldsymbol{\rho} + \mathbf{g} \mathbf{F} \mathbf{r}. \quad (4.29)$$

Finally, we substitute equation 4.28 and equation 4.29 into equation 4.18. This gives us a first order linear differential equation for the covariance between two neurons in our E-I population:

$$\tau_0 \frac{d}{dt} \rho_{ij} = -2\rho_{ij} + [\mathbf{g} \mathbf{A} \boldsymbol{\rho} + \boldsymbol{\rho} \mathbf{A}^T \mathbf{g}]_{ij} + [\mathbf{g} \mathbf{F} \mathbf{r} + \mathbf{r}^T \mathbf{F}^T \mathbf{g}]_{ij} \quad ; i \neq j \quad i, j \in E, I \quad . \quad (4.30)$$

This equation describes the time evolution of the spike train covariance  $\boldsymbol{\rho}$  of recurrently connected excitatory and inhibitory neurons in a balanced network.

The covariance between neurons in the E-I population depends on the covariance  $\mathbf{r}$  between neurons in external population X and neurons in the E-I population. We can derive a linear differential equation for  $\mathbf{r}$  using similar methods to our derivation of  $\boldsymbol{\rho}$  (Supplementary Methods 4.5.2):

$$\tau_0 \frac{d\mathbf{r}}{dt} = -2\mathbf{r} + \mathbf{r} \mathbf{A}^T \mathbf{g} + \mathbf{n} \mathbf{F}^T \mathbf{g}, \quad (4.31)$$

where  $\mathbf{n}$  is the covariance of our input population X:

$$n_{ij} \equiv \langle \delta x_i \delta x_j \rangle = \delta_{ij} (\nu_i - \nu_i^2) \quad ; i, j \in X \quad . \quad (4.32)$$

Neurons in the input population are unconnected, so their spike trains are uncorrelated. This is captured by the fact that  $\mathbf{n}$  is a diagonal matrix.

Correlations in balanced networks were recently studied by Renart et al. (Renart

et al., 2010). Their work is similar to ours in that they derive the same equilibrium covariance equation that we have derived (Eqn. 4.30). However, there are a number of important differences. The balanced network that we study contains structured connectivity, along with background connectivity, whereas Renart et al. study networks that only contain background connectivity. Also, our derivation is quite different and may be more easily adapted to the analysis of correlations in networks of more realistic neurons. In the next section, we go further than Renart et al. and solve the covariance equations at equilibrium.

### 4.2.3 Solving the correlation equations

In this section we solve the covariance equations and the firing rate equations that we have derived to finally obtain an expression for the spike train correlations  $C_{ij}$  in a balanced network. We begin by writing down the balanced network covariance equations 4.30, 4.31 and 4.32 at equilibrium:

$$2\boldsymbol{\rho} = [\mathbf{g}\mathbf{A}\boldsymbol{\rho} + \boldsymbol{\rho}\mathbf{A}^T\mathbf{g}]^{od} + [\mathbf{g}\mathbf{F}\mathbf{r} + \mathbf{r}^T\mathbf{F}^T\mathbf{g}]^{od} + 2\boldsymbol{\rho}^d, \quad (4.33)$$

$$2\mathbf{r} = \mathbf{r}\mathbf{A}^T\mathbf{g} + \mathbf{n}\mathbf{F}^T\mathbf{g}, \quad (4.34)$$

$$n_{ij} = \nu_i(1 - \nu_i)\delta_{ij}, i, j \in X \quad (4.35)$$

where we have used the following notation

$$M_{ij}^{od} = M_{ij}(1 - \delta_{ij}) \quad (4.36)$$

$$M_{ij}^d = M_{ij}\delta_{ij}. \quad (4.37)$$

These are linear equations so we can use linear algebra to solve them. This is particularly straightforward for the covariance between E-I neurons and neurons in our external population (Eqn. 4.34):

$$\mathbf{r} = \mathbf{n}\mathbf{F}^T\mathbf{g}(2\mathbf{I} - \mathbf{A}^T\mathbf{g})^{-1}. \quad (4.38)$$

However, the solution of the E-I covariance equation is more complicated. This is because the mathematical structures of the covariance between two different neurons and the covariance of a neuron with itself (its variance) are very different (Eqn. 4.33).

We deal with this by rewriting equation 4.33 in the following form:

$$2\rho^{od} = \mathbf{gA}\rho^{od} + \rho^{od}\mathbf{A}^T\mathbf{g} + \boldsymbol{\chi} + \boldsymbol{\chi}^T - \boldsymbol{\xi}, \quad (4.39)$$

$$\boldsymbol{\chi} = \mathbf{gA}\rho^d + \mathbf{gFr}, \quad (4.40)$$

$$\boldsymbol{\xi} \equiv \left[ \rho^{od}\mathbf{A}^T\mathbf{g} + \mathbf{gA}\rho^{od} \right]^d + \left[ \mathbf{r}^T\mathbf{F}^T\mathbf{g} + \mathbf{gFr} \right]^d. \quad (4.41)$$

Now, multiplying both sides of equation 4.39 by the conjugate eigenvalues of  $\mathbf{gA}$  allows us to write:

$$2\mathbf{e}_\mu^\dagger \rho^{od} \mathbf{e}_\nu^\dagger = \lambda_\mu \mathbf{e}_\mu^\dagger \rho^{od} \mathbf{e}_\nu^\dagger + \lambda_\nu \mathbf{e}_\nu^\dagger \rho^{od} \mathbf{e}_\mu^\dagger + \mathbf{e}_\mu^\dagger \boldsymbol{\chi} \mathbf{e}_\nu^\dagger + \mathbf{e}_\mu^\dagger \boldsymbol{\chi}^T \mathbf{e}_\nu^\dagger - \mathbf{e}_\mu^\dagger \boldsymbol{\xi} \mathbf{e}_\nu^\dagger, \quad (4.42)$$

where

$$\mathbf{e}_\mu^\dagger \mathbf{gA} = \lambda_\mu \mathbf{e}_\mu^\dagger \quad (4.43)$$

$$\mathbf{gA} \mathbf{e}_\mu = \lambda_\mu \mathbf{e}_\mu. \quad (4.44)$$

The term  $\mathbf{e}_\mu^\dagger \boldsymbol{\xi} \mathbf{e}_\nu^\dagger$  is much smaller than the other terms so its contribution to equation 4.42 can be ignored (Supplementary Methods 4.5.3). Multiplying both sides of this equation by the eigenvectors of  $\mathbf{gA}$  and solving for  $\rho^{od}$  gives:

$$\rho^{od} = \sum_{\nu, \mu} \frac{\mathbf{e}_\mu \mathbf{e}_\mu^\dagger (\boldsymbol{\chi} + \boldsymbol{\chi}^T) \mathbf{e}_\nu^\dagger \mathbf{e}_\nu}{2 - (\lambda_\mu + \lambda_\nu)}. \quad (4.45)$$

Finally, inserting equation 4.45 and equation 4.17 into equation 4.1 we can write down a single equation for the correlation between two different neurons,  $i$  and  $j$  in a balanced network:

$$C_{ij} = (\nu_i - \nu_i^2)^{-1/2} \left[ \sum_{\nu, \mu} \frac{\mathbf{e}_\mu \mathbf{e}_\mu^\dagger (\boldsymbol{\chi} + \boldsymbol{\chi}^T) \mathbf{e}_\nu^\dagger \mathbf{e}_\nu}{2 - (\lambda_\mu + \lambda_\nu)} \right]_{ij} (\nu_j - \nu_j^2)^{-1/2} \quad (4.46)$$

$$\boldsymbol{\chi} = \mathbf{gA}\rho^d + \mathbf{gFnF}^T\mathbf{g} (2\mathbf{I} - \mathbf{A}^T\mathbf{g})^{-1}. \quad (4.47)$$

This equation provides a prediction for the spiking covariance between all pairs of neurons in our network, given network connectivity and input statistics. It is the first full prediction for the correlations of a strongly connected balanced network.

Correlations in networks without strong background connectivity have been calcu-

lated analytically before (Ginzburg and Sompolinsky, 1994). The correlation equations for weakly correlated networks are similar to the equations for strongly connected networks except that the neuron gain is  $g_{ij} = \delta_{ij} \partial f(m_i) / \partial m_i$ , compared to  $g_{ij} = \delta_{ij} \partial \langle f(h_i) \rangle / \partial m_i$  in strongly connected networks. This difference is a consequence of synaptic background noise, generated by background connectivity, which effectively smooths the non-linear dynamics of individual neurons.

## 4.3 Results

### 4.3.1 The origin of correlations

We can now explain the origin of correlations in the cortex. Our correlation equations (Eqns. 4.46 and 4.47) reveal that there are two sources of spike correlations: recurrent input from cortical populations (the first term in  $\chi$ ); and correlated input from external populations (the second term in  $\chi$ ). These sources are transformed by the connectivity and gain of individual neurons by projecting the *effective input covariance*  $\chi + \chi^T$  onto eigenvectors of the 'effective connectivity matrix'  $\mathbf{gA}$ . Projections onto eigenvectors whose eigenvalues are closest to 1 contribute most to network correlations.

Our description of the origin of correlations is technical. This reflects the exact nature of our prediction. Such exact descriptions are necessary for studying many properties of balanced networks, such as the contribution of structured connectivity to correlations and the contribution of correlations to information. However, many properties of correlations can be understood without resorting to exact predictions, such as the dependence of correlations on the number of connections per neuron. We will investigate balanced network correlations qualitatively and quantitatively.

### 4.3.2 Simulations and predictions

In deriving our predictions, we made a number of approximations, which though reasonable, must be checked in network simulations. We do this by measuring the firing rates and correlations of all the neurons in a simulation of our network.

As before, we simulate our balanced network using the Gillespie Algorithm (Gillespie, 1977). During each simulation, we measure firing rates and correlations. For computa-

tional reasons, we assume ergodicity and estimate statistics across time. All statistical estimates converge towards the true statistics when simulations (and measurements) become infinitely long. This is also the limit for which our analytic predictions apply - when the network is at equilibrium.

For each simulation and prediction, we use the network parameters given in table 3.1, unless otherwise stated. We do not need to fine-tune these parameters because the balanced state is very robust to parameter choice.

We can accurately predict the firing rate of every neuron in our network (Fig. 4.2, inset). The neuron firing rate distribution is skewed, with a peak close to zero (Fig. 4.2). This shape is consistent with a range of firing rate recording experiments (Hromádka et al., 2008; O'Connor et al., 2010).

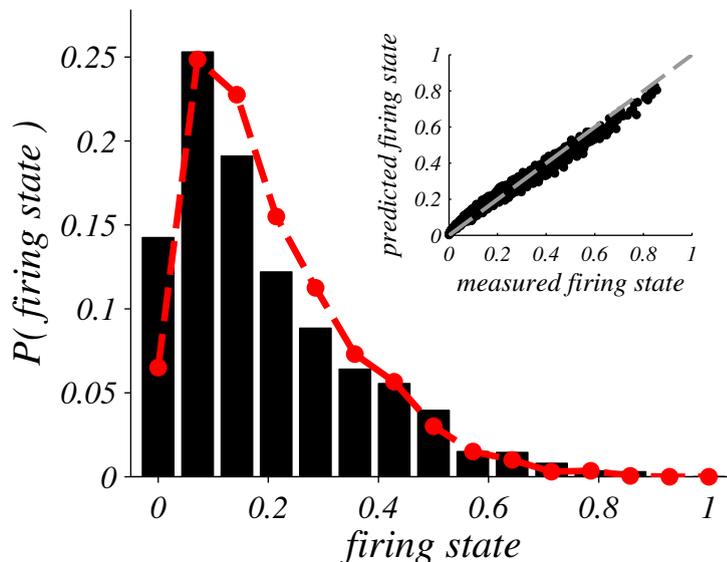


Figure 4.2: Firing rate prediction, measurement and histogram

The firing rate distribution in a balanced network is skewed towards zero with a low firing rate peak (black bars). We can accurately predict the firing rate distribution (red dots) and the individual firing rates  $\nu_i$  of neurons in our network (inset) for a network containing  $N = 2000$  neurons. The mean prediction error is 0.023.

We find that our correlation predictions closely match measured correlations (Fig. 4.3, inset). This is the first analytic prediction of correlations in a balanced network. Correlations are weak on average but are widely distributed relative to this average. The distribution is sharply peaked and skewed towards zero (Fig. 4.3). This is consistent with recent correlation recordings (Ecker et al., 2010).

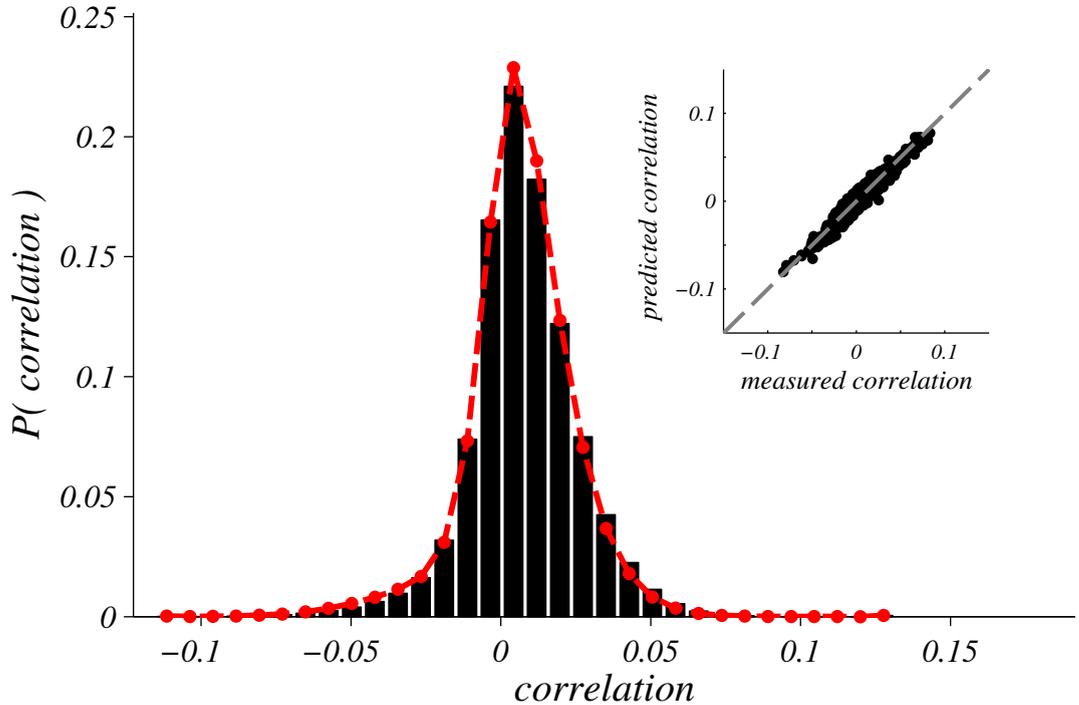


Figure 4.3: Correlation distribution, prediction and measurement

Correlations in balanced networks are weak on average and widely distributed relative to this average (black bars). The correlation distribution is skewed towards zero. We can accurately predict this distribution (red dots, left) and the correlations between pairs of neurons in our network (inset). Each point represents the correlation  $C_{ij}$  between a pair of randomly selected neurons  $i$  and  $j$ . The dashed line is the 45° line. The correlation error  $\sum_{ij}^N |C_{ij} - C_{ij}^m| / N^2$  is 0.004, where  $C_{ij}^m$  is the measured correlation between neuron  $i$  and  $j$ .

Although our predictions are accurate, they are not perfect. Indeed, we don't expect them to be perfect, because some of the approximations in our analysis are only applicable to large networks. However, this is not a serious problem, because a typical cortical column is extremely large, containing millions of neurons. Nonetheless, we must demonstrate that our predictions do improve as we increase the size of our networks. We quantify the accuracy of our correlation predictions by calculating  $\sum_{ij}^N |C_{ij} - C_{ij}^m| / N^2$ , where  $C_{ij}^m$  is the measured correlation between neuron  $i$  and  $j$ . We find that this error decreases with network size (Fig. 4.4).

It is also important that our predictions are accurate for networks with different levels of background connectivity and structured connectivity. Our correlation predictions consists of  $(N^2 - N) / 2$  distinct numbers. To compare predicted correlations to measured correlations for a range of networks we use  $\bar{\mathbf{C}} + s.d.(\mathbf{C})$  as a summary stat-

istic, where  $\overline{\mathbf{C}}$  is the mean correlation between spike trains in our network and  $s.d.(\mathbf{C})$  is the standard deviation of correlation coefficients. This quantity captures the typical size of correlations in our network. We find that measured correlations closely match predicted correlations (Fig. 4.4) for a range of networks with different amounts of background connectivity and structured connectivity. Therefore, we can rest assured that our predictions are accurate for a range of network parameters and network sizes.

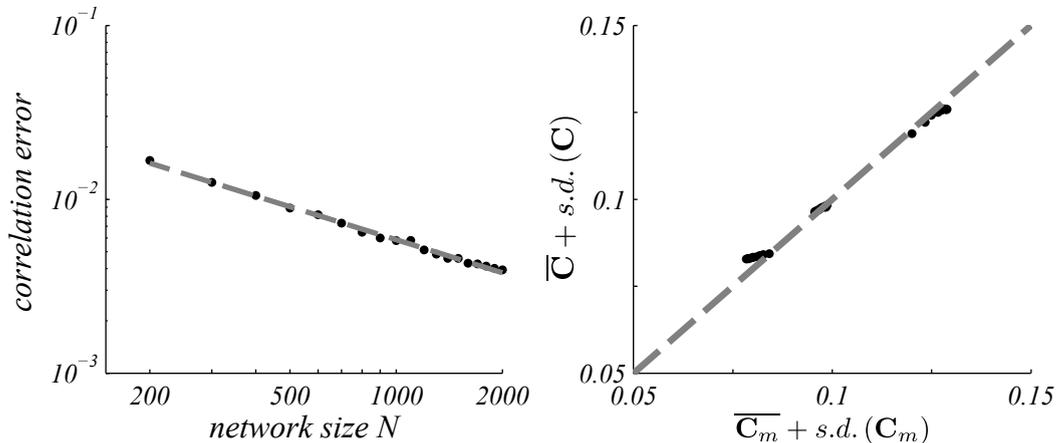


Figure 4.4: Correlation prediction accuracy

The error in our correlation predictions decreases as the network size increases (left). A straight line fit to the error indicates that the error decreases with  $N^{-0.63}$ . For each prediction, the same network parameters are used (Table 3.1), apart from  $N$ ,  $N_X$ ,  $K$  and  $K_X$  which all increase by the same proportion. The average firing rate is held constant, by adjusting the spiking thresholds. Also, we increase the measurement time  $T$  according to  $T = N \times 50 \text{ sec}$ , because correlations become smaller with  $1/\sqrt{N}$  (Fig. 4.8) and this requires more accurate correlation estimates. This accuracy increases with  $\sqrt{T}$ , so measurement time must increase linearly with network size to provide a fair comparison between networks. We also find that our correlation predictions represented by  $\overline{\mathbf{C}} + s.d.(\mathbf{C})$  closely match predicted correlations, represented by  $\overline{\mathbf{C}}_m + s.d.(\mathbf{C}_m)$  for a wide range of parameters ( $w_0 \in [0.01, 2]$ ,  $j_0 \in [0.01, 2] \times w_0$ ), where the network size is  $N = 500$  and the measurement time is  $T = 250 \text{ sec}$ .

### 4.3.3 The contribution of structured connectivity and background connectivity to correlations

We now investigate the relationship between correlations and connectivity. It is important to understand this relationship because the interaction of connectivity and correlations is central to the theory of cortical memory formation (Hebb, 1949). If the contribution of structured connectivity to correlations is overwhelmed by the contribution of background connectivity, Hebbian learning may be impossible. Also, attempts

to infer structured connectivity using correlation measurements may be hopeless.

We find that structured connectivity parametrised by  $j_0$  increases correlations substantially (Fig. 4.5). The average correlation becomes more positive and the correlation standard deviation increases substantially.

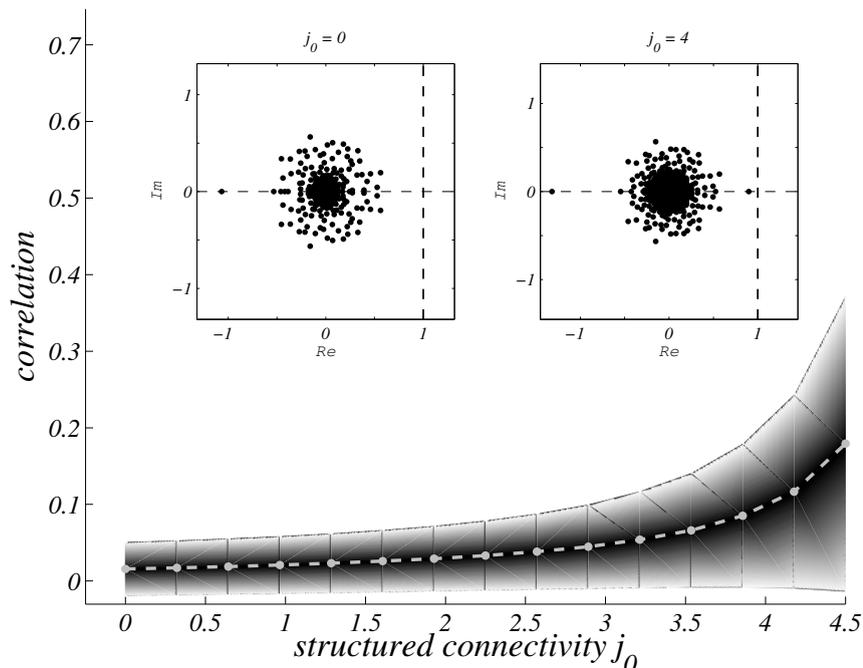


Figure 4.5: Structured connectivity and correlations

Structured connectivity increases correlations. Here the predicted mean correlation (grey dashed line), and the mean  $\pm$  standard deviation (grey shading) are shown. For each prediction, the average excitatory firing rate is held constant, by adjusting the spiking thresholds. The eigenvalue spectrum of our 'effective connectivity matrix'  $\mathbf{gA}$  is shown without structure (left inset) and with structure (right inset). Structured connectivity can increase correlations if some eigenvalues  $j_0$  are close to 1.

It is not surprising that the increase in the magnitude of average correlations is positive, rather than negative. In our model, structured connections only occur between excitatory neurons. Consequently, the contribution of structure to total synaptic drive correlations is positive on average, which, in turn, causes positive increases in average spike correlations.

It is surprising however that this increase is so substantial, because structured connectivity is much weaker than background connectivity. We can understand this mathematically, in terms of the eigenvalues  $\lambda_n$  and eigenvectors  $\mathbf{e}_n$  of the effective connectivity matrix  $\mathbf{gA}$  where  $\mathbf{gA}\mathbf{e}_n = \lambda_n\mathbf{e}_n$ . Eigenvectors whose eigenvalues are closest to 1 make

the largest contribution to correlations (Eqn. 4.46). Although structured connectivity is weak it has some  $\mathcal{O}(1)$  eigenvalues. As  $j_o$  increases, these eigenvalues increase and consequently, the correlations increase.

Next, we study the contribution of background connectivity to correlations. We find that background connectivity, parametrised by  $w_0$ , decreases average correlations, but increases the correlation standard deviation (Fig. 4.6). However, this contribution eventually saturates.

Background connectivity is much stronger than structured connectivity so it is not surprising that it contributes to correlations. The decrease in mean correlation strengths is a result of synaptic drive correlations between excitatory and inhibitory synaptic inputs balancing the positive correlations from structured connectivity.

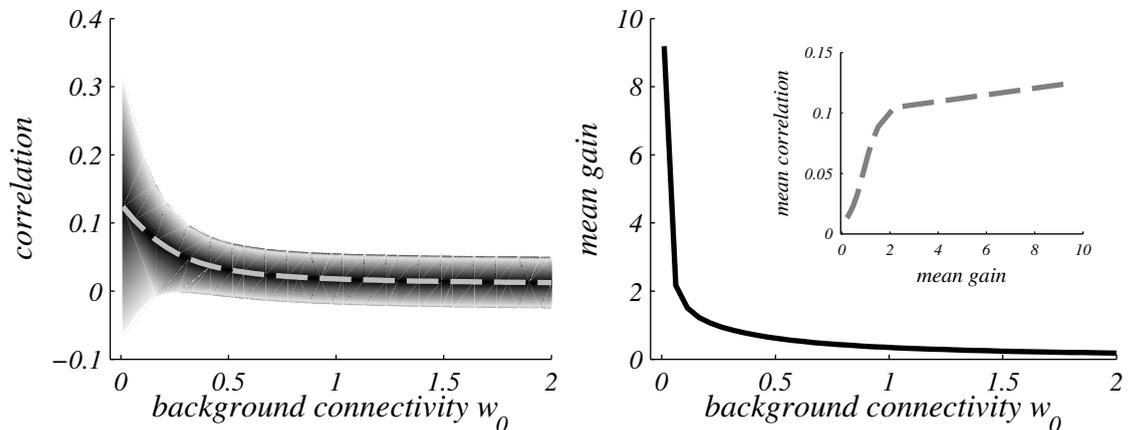


Figure 4.6: Background connectivity and correlations

Background connectivity decreases mean correlations (grey dashed line, left) and increases the standard deviation of correlations (grey shading, left). These correlation changes eventually saturate. Correlations decrease with background connectivity because correlations are largely determined by the gain (inset, right), and background connectivity decreases the mean gain (right). For each prediction here, the average firing rate is held constant, by adjusting the spiking thresholds.

The decrease in correlations before saturation can be understood in terms of the mean gain  $\bar{g} \equiv \sum_i g_{ii}/N$  (Eqn. 4.25), because the gain plays an important role in determining the size of correlations (Eqns. 4.44, 4.46, 4.47). If the gain is small, then correlations will be small, regardless of network connectivity (Fig. 4.6).

We can understand this relationship mathematically by estimating the gain of a

neuron in terms of background connectivity, parametrised by  $w_o$ :

$$g_{ii} = \frac{\exp[-(m_i/\sigma_i)^2/2]}{(2\pi\sigma_i^2)^{1/2}} \quad (4.48)$$

We can simplify this by observing that background connectivity is predominantly responsible for synaptic background noise, so  $\sigma_i \sim w_0$  (Eqn. 4.10). Also, the ratio  $m_i/\sigma_i$  is constant because we have required firing rates to be constant (Eqn. 4.13):

$$m_i/\sigma_i = \phi^{-1}(\nu_i) . \quad (4.49)$$

Therefore, for typical amounts of background connectivity ( $w_0 \sim \mathcal{O}(1)$ ) we have:

$$g_{ii} \sim w_o^{-1} . \quad (4.50)$$

The gain decreases as background increases, at a given firing rate, so correlations decrease as background connectivity increases.

The saturating contribution of background connectivity can also be understood mathematically by looking at the eigenvalue spectrum of the effective connectivity matrix  $\mathbf{gA}$ . These eigenvalues determine the contribution of different connectivity components to correlations, as we have already discussed (Eqn. 4.46). If these eigenvalues saturate with background connectivity, then the contribution of background connectivity to correlations will also saturate.

When  $w_0$  becomes large, the typical size of these eigenvalues in terms of background connectivity, parametrised by  $w_o$ , is given by:

$$\begin{aligned} \lambda_\mu &\sim g_{ii} \times A_{ij} \\ &\longrightarrow \mathcal{O}(w_o^{-1} \times w_o) . \end{aligned} \quad (4.51)$$

Therefore, these eigenvalues become independent of background connectivity and consequently, the contribution of background connectivity to correlations saturates eventually.

### 4.3.4 Firing rate, preferred orientation and correlations

The relationship between correlations and firing rates is important for computation. If correlations are higher between neurons with similar preferred orientations, then the ability of the network to encode and decode information may be harmed (Averbeck et al., 2006) (Fig. 2.4). Here, we study this relationship in balanced networks.

We find that if a pair of neurons in a balanced network have high firing rates, they are more likely to be correlated (Fig. 4.7). Consequently, neurons that encode stimuli of similar orientation have higher correlations on average. This relationship is not very strong, so it is not clear that it would harm computations significantly. This is something that we will address in the next chapter.

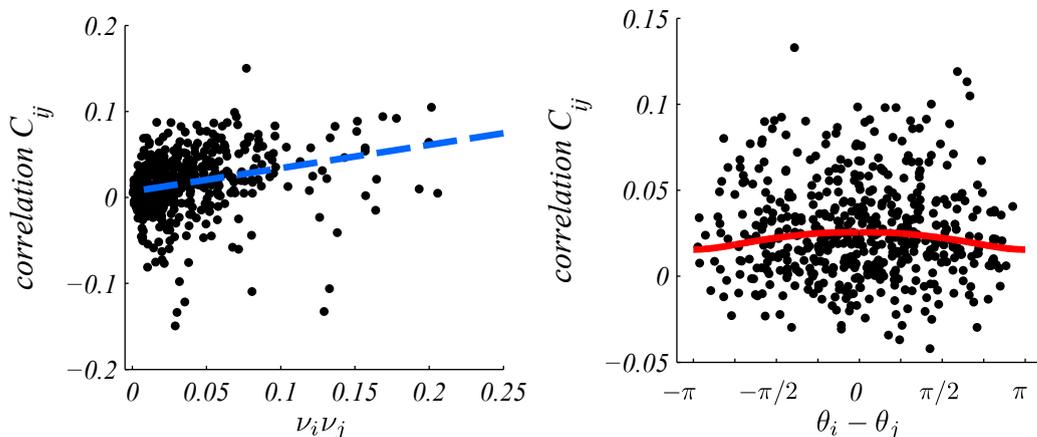


Figure 4.7: Firing rate, preferred orientation and correlations

Pairs of neurons with high firing rates are more likely to be correlated than neurons with low firing rates. A least squares fit to the full set of correlations has a slope of 0.27 (dashed blue line, left). Consequently, neurons that receive input tuned to similar orientations are more correlated (right). This can be seen clearly from the average correlation (red line, right). Here, averages are given by the average correlation at a particular orientation difference, and then smoothed by convolving with a Gaussian with standard deviation of  $\pi/8$ . As before, averaging and regression are performed for the entire population, whereas only a randomly selected subset of correlations are plotted, for clarity.

### 4.3.5 Correlation scaling

In this section we estimate the size of correlations in balanced networks by calculating how correlations scale with  $K$ , the average number of synapses per neuron. The size of

correlations is important because it can determine whether or not correlations play a role in computation (Abbott and Dayan, 1999). The scaling of correlations with  $K$  is a good indicator of correlation size because  $K$  is very large in the cortex (Braitenber and Schuz, 1991; Binzegger et al., 2004) and therefore plays an important role in determining spike train statistics, as we have seen in the previous chapters.

We quantify correlation size by calculating the mean correlation  $\bar{\mathbf{C}}$  of the E-I population and the correlation standard deviation  $s.d.(\mathbf{C})$  (Supplementary Methods 4.5.4). Our approach is to calculate the relative size of each term in our covariance equation (Eqn. 4.39). Some terms will be much smaller than others because of their dependence on  $K$ . The smallest terms can be ignored and the resulting equation can easily be solved. We find that correlations get smaller as the number of synapses per neuron increases:

$$\begin{aligned}\bar{\mathbf{C}} &\sim 1/K \\ s.d.(\mathbf{C}) &\sim 1/\sqrt{K}.\end{aligned}\tag{4.52}$$

We made a number of approximations in this analysis which can be corroborated by solving our covariance equation numerically (Eqn. 4.46) for a series of networks of increasing size. Again, we find that  $\bar{\mathbf{C}} \sim 1/K$  and  $s.d.(\mathbf{C}) \sim 1/\sqrt{K}$  (Fig. 4.8). This is in agreement with our analysis.

These scaling results allow us to predict the typical size of correlations in the cortex. Neurons in the cortex typically have about 1000 – 2000 connections (Braitenber and Schuz, 1991; Binzegger et al., 2004). Consequently, correlations in balanced networks are about 0.02. This is consistent with the most recent measurements of correlations in cortical networks (Ecker et al., 2010; Renart et al., 2010).

We also find that structured connectivity and background connectivity both make a similar sized contribution to  $\bar{\mathbf{C}}$  and  $s.d.(\mathbf{C})$  (Supplementary Methods 4.5.4). This is consistent with our previous observations that structured connectivity can increase correlations substantially (Figs. 4.5 and 4.7).

Our scaling results are qualitatively similar to the results of Renart et al. (Renart et al., 2010) who also found that  $\bar{\mathbf{C}} \sim 1/K$  in their balanced network model. However, the result is quantitatively different because our networks contain structured connectivity and we have shown that the contribution of structured connectivity is just as large

as the contribution of background connectivity.

An important contribution of (Renart et al., 2010) was their investigation into the size of synaptic drive correlations. They observed that the average correlation of excitatory synaptic drives,  $c_{EE}$  is large and positive;  $c_{EE} \sim \mathcal{O}(1)$ . Similarly, the average correlation of inhibitory synaptic drives,  $c_{II}$  is large and positive;  $c_{II} \sim \mathcal{O}(1)$ . The average correlation between inhibitory and excitatory synaptic drives  $c_{EI}$  is large and negative;  $c_{EI} \sim \mathcal{O}(-1)$ . However, the total synaptic drive correlation  $c_{tot}$  is weak  $c_{tot} \sim \mathcal{O}(1/\sqrt{K})$ . They proposed that this is a consequence of correlation components balancing each other, similar to the balancing of excitatory and inhibitory mean synaptic drive in balanced networks (Renart et al., 2010).

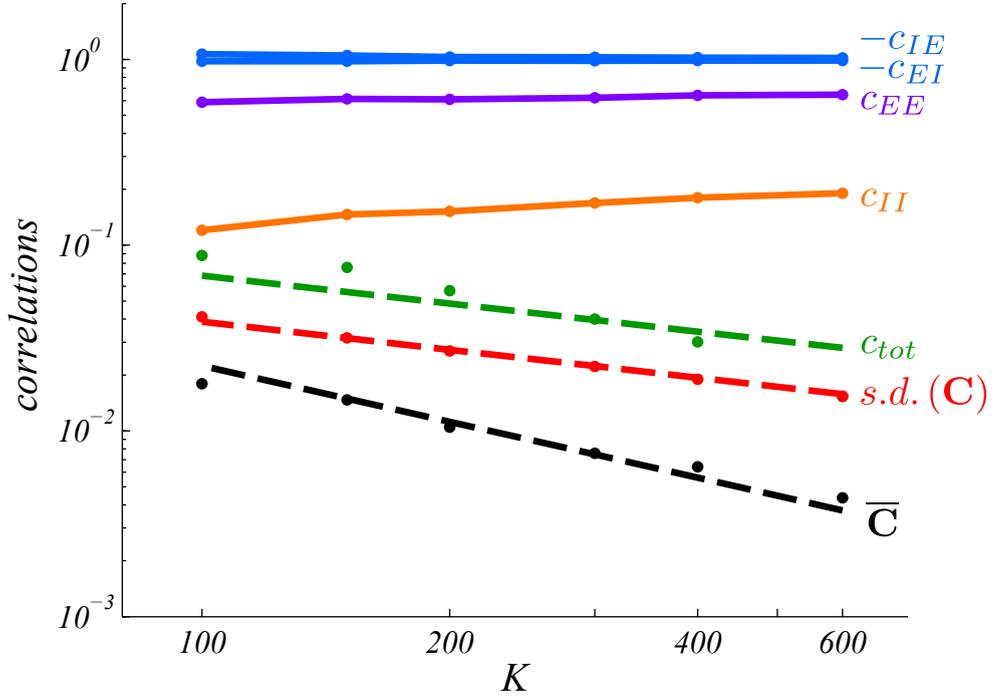


Figure 4.8: Correlation scaling

Correlation scaling shows that correlations are weak in a balanced network. Here, the average spike correlation  $\overline{C}$  between neurons in a balanced network are shown in black and the correlation standard deviation is shown in red  $s.d.(\mathbf{C})$ . Dashed lines of order  $1/K$  (black) and  $1/\sqrt{K}$  (red) show the scaling of spike train correlations. We also show the average correlations between excitatory and inhibitory synaptic drives  $c_{EE}$  (purple),  $c_{II}$  (orange),  $-c_{IE}$  and  $-c_{EI}$  (blue). These correlations cancel so that the total synaptic drive correlation  $c_{tot}$  (green) is small and is of order  $1/\sqrt{K}$  (dashed green line).

We find that synaptic drive correlations also balance in networks that have structured connectivity and background connectivity (Fig. 4.8). The background connectiv-

ity correlations balance the structured connectivity correlations producing synaptic drive correlations that are weak (Fig. 4.6). This is important because in our analysis, we assumed that the contribution of correlations to synaptic drive variance was small enough to be ignored (Eqn. 4.10). It is also important for computation, because potentially harmful correlations induced by structured connectivity are reduced in balanced networks.

## 4.4 Discussion

Since the seminal work of Adrian and Zotterman (Adrian and Zotterman, 1926), the firing rates of neurons have become the subject of intense investigation. Spike train correlations have only recently received the same attention, mostly because they are difficult to measure (Ecker et al., 2010; Cohen and Kohn, 2011). Nonetheless, correlations can play just as important a role in cortical function (Averbeck et al., 2006).

We study correlations analytically using techniques from statistical physics and dynamical systems theory. We expect this analysis to be difficult, if not intractable, because single neuron dynamics are highly non-linear. However, synaptic background noise is large in balanced networks (Destexhe et al., 2003; Rudolph et al., 2005) and spike trains are asynchronous (Ecker et al., 2010; Cohen and Kohn, 2011), allowing us to linearise the complicated non-linear network dynamics and derive a tractable mathematical relationship between correlations and connectivity.

Our analysis reveals that correlations in the cortex are weak, about 0.02, and highly heterogeneous across a population. Specifically, the correlation standard deviation is of order  $1/\sqrt{K}$  and the mean correlation is of order  $1/K$ . This is consistent with recent correlation measurements of  $0.005 \pm 0.004$  (Ecker et al., 2010; Renart et al., 2010). Although this is weak, it is much larger than correlations in networks that do not have background connectivity, about  $\sqrt{K}$  times larger (Ginzburg and Sompolinsky, 1994; Hertz, 2010). We investigate the implications of correlation size for computation in the next chapter and find that balanced network correlations are too small to play a critical role in computation, though they are not small enough to be ignored. Recent work by Renart et al. also found that correlations are weak (Renart et al., 2010), but in a network model that does not contain structured connectivity.

We find that the contribution of structured connectivity to correlations can be just

as large as the contribution of background connectivity, despite the fact that structured connectivity is much weaker than background connectivity in our model. Therefore, we predict that it will be difficult, if not impossible to unravel the contributions of background connectivity and structured connectivity in experiments. Consequently, experimentalists hoping to infer structure using correlation measurements will face analysis and measurement difficulties. We also anticipate difficulties for Hebbian learning in the cortex, because background connectivity produces spurious correlations which impairs the ability of a network to learn. Some form of population averaging may be necessary to stabilise Hebbian learning in balanced networks.

We find that the distribution of correlations in a balanced network can be sharply peaked close to zero. This is reminiscent of firing rate distributions which are also skewed and sharply peaked (Hromádka et al., 2008; O’Connor et al., 2010). Recently it was proposed that such firing rate distributions are produced by neural tuning curves stretching synaptic drives that are Gaussian distributed (Roxin et al., 2011). A similar mechanism may explain the skewed correlation distributions that we have calculated.

One problem with our analysis is that the neurons in our network model are not very realistic. This is not a serious problem because spiking activity in our network model can still be realistic - background connectivity produces synaptic background noise which in turn produces irregular asynchronous spiking activity. Nonetheless, in future work it would be interesting to investigate correlations in a network of realistic neurons. This may be difficult, because typically, there is a trade-off between model complexity and analytic tractability. Nonetheless, there has been some progress using simulations of conductance based integrate and fire neurons (Kumar et al., 2008; Hertz, 2010; Lerchner and Latham, 2011). These studies indicate that correlations are weak in the balanced state but can be much larger in the unbalanced state. This may explain some of the surprisingly large correlation measurements that are often reported (Hertz, 2010; Cohen and Kohn, 2011). However, a problem with all simulation based results is that they depend on simulation parameters in some unknown way. Even if a phenomenon is observed for some broad range of parameter values, there may be some undetected singularities where network activity is different. Future analytic calculations of correlations between realistic neurons may reveal as yet unknown subtleties in balanced network spike train correlations that may be important in learning and

computation.

In this chapter we calculated correlations in a balanced network at equilibrium. However, correlations can fluctuate in time, just as firing rates fluctuate in time. These fluctuations may play an important role in computations, especially for computations that are performed rapidly. Also, correlation fluctuations may be important for short time-scale learning. It is difficult to measure the time course of correlations, because averaging across a single trial is not possible for such measurements. Analysis is also difficult, because time-dependent correlation equations are more complicated than equilibrium equations. In this chapter, we have already derived time-dependent correlation equations for a balanced network, but we have not solved them. In future work, the role of correlation fluctuations in learning and computation could be investigated by solving these time dependent equations.

The most important prediction arising from this work is that structured connectivity increases correlations by amplifying salient inputs. Already, there is some evidence for this from experiments where monkeys view orientated bars, which are particularly salient, producing correlations in V1 of  $0.005 \pm 0.004$ , compared to natural image viewing which produce correlations of  $0.001 \pm 0.005$  (Ecker et al., 2010). Furthermore, we also predict that similarly tuned cells are slightly more correlated than cells that have dissimilar tuning. There is also some evidence for this from experiments where correlations are  $0.023 \pm 0.005$  for similarly tuned cells compared to  $0.008 \pm 0.002$  for cells with dissimilar tuning (Ecker et al., 2010). However, a more systematic investigation of our prediction is necessary, preferably across animals and sensory modalities.

We have made substantial progress in the analysis of cortical correlations. The correlation equations that we derive can be used to calculate the typical correlation strength in the cortex and the relative contribution of structured connectivity and background connectivity. However, our analysis has revealed a complicated relationship between correlations and structured connectivity and between correlations and network input. This complexity is not surprising, given that the cortex consists of many interconnected non-linear dynamical spiking neurons. In learning and computation, the details matter, so it is important that we understand them. In the next chapter, we will quantify the role of correlations in computation.

## 4.5 Supplementary Methods

In this section, we provide derivations for all the equations in this chapter. Although all of these derivations are important mathematically, they provide little additional neuroscience insight and they are quite technical. Therefore, we have not included them in the main text. For the sake of completeness, we have reproduced some derivations from previous work, as noted.

### 4.5.1 Master equation

Here we derive the Master equation for our network dynamics (Eqn. 4.4). This Master equation has been derived before (Ginzburg and Sompolinsky, 1994).

We begin by considering the dynamics of the neurons in our network (Fig. 4.9). If the network is in state  $\mathbf{x}$  at time  $t + \Delta t$ , then for sufficiently small  $\Delta t$ , the network was either in the same state  $\mathbf{x}$ , or in a different state  $\mathbf{x}^{(i)}$  at time  $t$ .

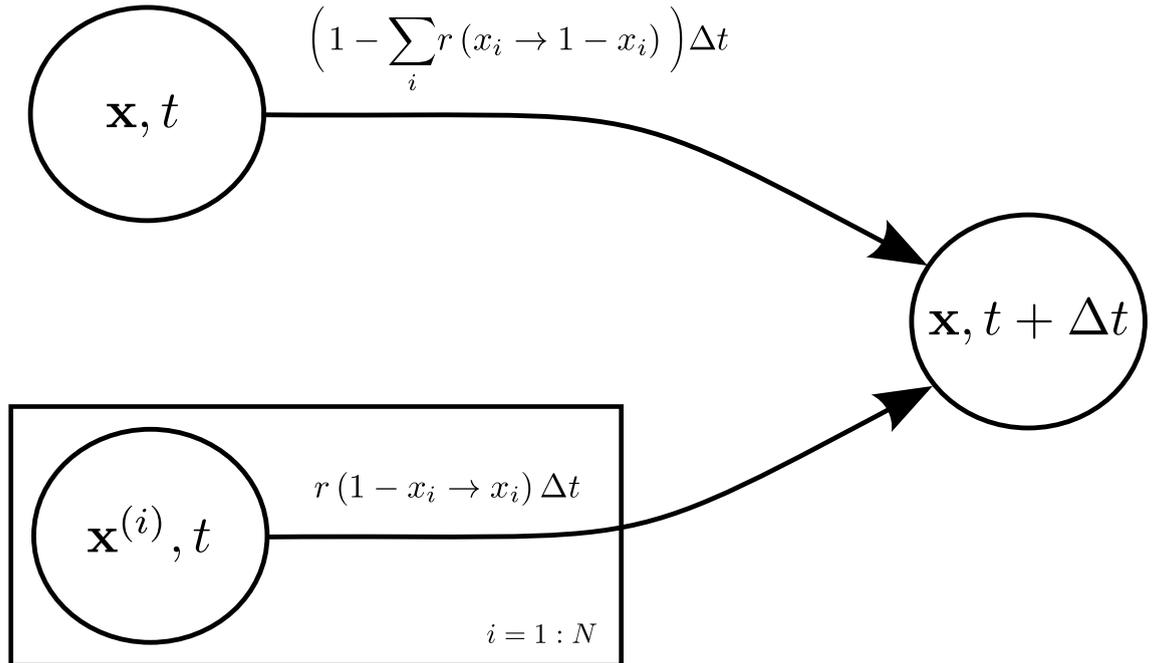


Figure 4.9: Neuron model

This graphical model represents the stochastic dynamics of our network. The network can evolve to the state  $\mathbf{x}$  at time  $t + \Delta t$  from the same state  $\mathbf{x}$  at time  $t$ , or from a different state  $\mathbf{x}^{(i)}$  at time  $t$ . The network transition probabilities are notated beside the directed edge of the corresponding network state transition. We have used plate notation to denote all the possible transitions from past network states.

Therefore, the probability that the network is in state  $\mathbf{x}$  at time  $t + \Delta t$  is given by:

$$\begin{aligned}
P(\mathbf{x}, t + \Delta t) = & P(\mathbf{x}, t) \left( 1 - \sum_i r(x_i \rightarrow 1 - x_i) \right) \Delta t \\
& + \sum_i r(1 - x_i \rightarrow x_i) P(\mathbf{x}^{(i)}, t) \Delta t, \tag{4.53}
\end{aligned}$$

where  $P(\mathbf{x}, t)$  denotes the probability that the network is in state  $\mathbf{x}$  at time  $t$  and  $P(\mathbf{x}^{(i)}, t)$  denotes the probability that the network is in state  $\mathbf{x}^{(i)}$  at time  $t$ .

By dividing both sides of equation 4.53 by  $\Delta t$  and taking the small  $\Delta t$  limit we obtain the Master equation for our network dynamics:

$$\begin{aligned}
\frac{d}{dt} P(\mathbf{x}, t) & \equiv \lim_{\Delta t \rightarrow 0} \frac{P(\mathbf{x}, t + \Delta t) - P(\mathbf{x}, t)}{\Delta t} \\
& = - \sum_i r(x_i \rightarrow 1 - x_i) P(\mathbf{x}, t) + \sum_i r(1 - x_i \rightarrow x_i) P(\mathbf{x}^{(i)}, t). \tag{4.54}
\end{aligned}$$

Q.E.D.

## 4.5.2 Firing rate equations

Here we derive the firing rate equations for our network (Eqn. 4.5). These equations were also derived previously (Ginzburg and Sompolinsky, 1994). We begin by combining our Master Equation (Eqn. 4.4) and our firing rate definition (Eqn. 4.2) as follows:

$$\begin{aligned}
\frac{d}{dt} v_i & = \sum_{\mathbf{x}} x_i \frac{d}{dt} P(\mathbf{x}, t) \\
& = - \sum_{\mathbf{x}} \sum_j x_i r(x_j) P(\mathbf{x}, t) + \sum_{\mathbf{x}} \sum_j x_i r(1 - x_j) P(\mathbf{x}^{(j)}, t) \\
& = \sum_{\mathbf{x}} \left[ - \sum_j x_i r(x_j) + \sum_{j \neq i} x_i r(x_j) + (1 - x_i) r(x_i) \right] P(\mathbf{x}, t) \\
& = \sum_{\mathbf{x}} (1 - 2x_i) r(x_i) P(\mathbf{x}, t) \\
& = \langle (1 - 2x_i) r(x_i) \rangle. \tag{4.55}
\end{aligned}$$

Here, we have abused notation, using  $r(x_j)$  to denote  $r(x_j \rightarrow 1 - x_j)$  and we have used brackets  $\langle \dots \rangle$  to denote an average over all neuron states at time  $t$ .

We can derive dynamical equations for our firing rates by inserting our transition rate expressions (Eqns. 4.56 and 4.57) into equation 4.55. For neurons in our external

population the transition rate can be written as

$$r(x_i) = \frac{1}{2\tau_X} (1 + (2x_i - 1)(2u_i - 1)), \quad (4.56)$$

and for neurons in our recurrently connected population the transition rate can be written as:

$$r(x_i) = \frac{1}{2\tau_Q} (1 + (2x_i - 1)(2f(h_i) - 1)), \quad (4.57)$$

with time constants given by

$$\tau_Q = \begin{cases} \tau_E & i \in E \\ \tau_I & i \in I. \end{cases} \quad (4.58)$$

In this chapter we assume that  $\tau_E = \tau_0$  and  $\tau_I = \tau_0$ .

Now, inserting 4.57 into 4.55 we find:

$$\begin{aligned} \frac{d}{dt}\nu_i &= \frac{1}{2\tau_0} \langle (1 - 2x_i)(1 + (2x_i - 1)(2f(h_i) - 1)) \rangle \\ &= \frac{1}{2\tau_0} \langle (1 - 2x_i) + (2f(h_i) - 1) \rangle \\ &= -\frac{1}{\tau_0} \langle x_i - f(h_i) \rangle. \end{aligned} \quad (4.59)$$

Here, we have used the fact that  $(1 - 2x_i)^2 = 1$ . A similar calculation can also be performed for neurons in our input population. This produces the firing rate equations for our network (Eqns. 4.5):

$$\begin{aligned} \tau_0 \frac{d}{dt}\nu_i &= -\nu_i + \langle f(h_i) \rangle & i \in E, I \\ \tau_X \frac{d}{dt}\nu_i &= -\nu_i + u_i & i \in X. \end{aligned} \quad (4.60)$$

Q.E.D.

### Synaptic drive variance

Next, we calculate the variance of the synaptic drive to a neuron in our network (Eqn. 4.10):

$$\begin{aligned}
\sigma_i^2 &= \left\langle [\mathbf{A} \cdot \delta \mathbf{x} + \mathbf{F} \cdot \delta \mathbf{x}]_i^2 \right\rangle \\
&= \left[ (\mathbf{A} + \mathbf{F}) \langle \delta \mathbf{x} \delta \mathbf{x}^T \rangle (\mathbf{A} + \mathbf{F})^T \right]_{ii} \\
&= [\mathbf{A} \boldsymbol{\rho} \mathbf{A}^T + \mathbf{F} \mathbf{n} \mathbf{F}^T + 2\mathbf{F} \mathbf{r} \mathbf{A}^T]_{ii},
\end{aligned} \tag{4.61}$$

where  $\mathbf{n}$  is the covariance of the input population X:

$$n_{ij} = \delta_{ij} (\langle x_i^2 \rangle - \nu_i^2) = \delta_{ij} (\nu_i - \nu_i^2) \quad i, j \in X, \tag{4.62}$$

and  $\mathbf{r}$  is the covariance between the input population X and the E-I population:

$$r_{ij} = \langle \delta x_i \delta x_j \rangle \quad i \in X, j \in E, I. \tag{4.63}$$

This synaptic drive variance equation (Eqn. 4.61) is complicated because of the contribution of correlations. However, in a balanced network the synaptic drive correlations balance, just as the mean synaptic input balances (Renart et al., 2010). Therefore, we can approximate the synaptic drive variance as follows:

$$\sigma_i^2 \simeq \sum_{j \in E, I} A_{ij} \nu_j (1 - \nu_j) A_{ji}^T + \sum_{j \in X} F_{ij} \nu_j (1 - \nu_j) F_{ji}^T. \tag{4.64}$$

Q.E.D.

## Equilibrium firing rate

Next we derive an equation for the equilibrium firing rate of a neuron in a balanced network (Eqn. 4.13):

$$\begin{aligned}
\nu_i &= \int \Theta(h_i) p(h_i, t) dh_i \\
&= \int \Theta(h_i) \exp[-(h_i - m_i)^2 / 2\sigma_i^2] (2\pi\sigma_i^2)^{-1/2} dh_i \\
&= \int \Theta(m_i + \delta h_i) \exp[-(\delta h_i)^2 / 2\sigma_i^2] (2\pi\sigma_i^2)^{-1/2} d\delta h_i \quad (4.65) \\
&= \int_{-m_i/\sigma_i}^{\infty} \exp[-z^2/2] (2\pi)^{-1/2} dz \\
&= \int_{-\infty}^{m_i/\sigma_i} \exp[-z^2/2] (2\pi)^{-1/2} dz \\
&= \Phi(m_i/\sigma_i), \quad (4.66)
\end{aligned}$$

where we have rewritten equation 4.65 using  $z \equiv \delta h_i / \sigma_i$ .

Q.E.D.

## Newton–Raphson method

We use the Newton-Raphson algorithm to solve our equilibrium firing rate equations (Algorithm 4.1). This algorithm converges towards the zeros of the following function:

$$z_i(\boldsymbol{\nu}) \equiv \nu_i - \Phi(m_i/\sigma_i), \quad (4.67)$$

using the Jacobian

$$\begin{aligned}
J_{ij} &= \frac{\partial}{\partial \nu_j} z_i \\
&= \frac{\partial \nu_i}{\partial \nu_j} - \frac{\partial}{\partial \nu_j} \Phi(m_i/\sigma_i) \\
&= \delta_{ij} - \sum_k g_{ii} \tilde{A}_{ik} \frac{\partial \nu_k}{\partial \nu_j} + \sum_k g_{ii} \tilde{F}_{ik} \frac{\partial u_k}{\partial \nu_j} \\
&= \delta_{ij} - g_{ii} \tilde{A}_{ij}, \quad (4.68)
\end{aligned}$$

where

$$g_{ii} = \frac{\exp[-m_i^2/2\sigma_i^2]}{(2\pi\sigma_i^2)^{1/2}}, \quad (4.69)$$

$$\tilde{A}_{ij} = A_{ij} \left( 1 - \frac{m_j}{2\sigma_i^2} A_{ij} (1 - 2\nu_j) \right), \quad (4.70)$$

$$\tilde{F}_{ij} = F_{ij} \left( 1 - \frac{m_j}{2\sigma_i^2} F_{ij} (1 - 2u_j) \right). \quad (4.71)$$

---

**Algorithm 4.1** Newton–Raphson method

---

Initialise:	Randomly choose firing rates $\boldsymbol{\nu}$
	Solve $\mathbf{J}\Delta\boldsymbol{\nu} = -\mathbf{z}(\boldsymbol{\nu})$ for $\Delta\boldsymbol{\nu}$
Iterate:	Update $\boldsymbol{\nu} \rightarrow \Delta\boldsymbol{\nu} + \boldsymbol{\nu}$
	Stop when $\max_i  z_i  < z_{threshold}$

---

### 4.5.3 Covariance derivation

Here we derive the equal-time covariance equations for our network (Eqn. 4.18). This derivation is reproduced from (Ginzburg and Sompolinsky, 1994). We begin by combining the Master Equation (Eqn. 4.4) with the covariance definition (Eqn. 4.16):

$$\begin{aligned}
\frac{d}{dt} \langle x_i x_j \rangle &= \sum_{\mathbf{x}} x_i x_j \frac{d}{dt} P(\mathbf{x}, t) \\
&= - \sum_{\mathbf{x}} \sum_l x_i x_j r(x_l) P(\mathbf{x}, t) + \sum_{\mathbf{x}} \sum_l x_i x_j r(1 - x_l) P(\mathbf{x}^{(l)}, t) \\
&= \sum_{\mathbf{x}} \left[ - \sum_l x_i x_j r(x_l) + \sum_{l \neq i, j} x_i x_j r(x_l) \right. \\
&\quad \left. + (1 - x_i) x_j r(x_i) + x_i (1 - x_j) r(x_j) \right] P(\mathbf{x}, t) \\
&= \langle x_j (1 - 2x_i) r(x_i) \rangle + \langle x_i (1 - 2x_j) r(x_j) \rangle \\
&= - \langle x_j (x_i - f(h_i)) \rangle / \tau_0 - \langle x_i (x_j - f(h_j)) \rangle / \tau_0.
\end{aligned} \quad (4.72)$$

This can be rewritten as follows:

$$\tau_0 \frac{d}{dt} \langle x_i x_j \rangle = -2 \langle x_i x_j \rangle + \langle x_i f(h_j) \rangle + \langle x_j f(h_i) \rangle. \quad (4.73)$$

Now, inserting equation 4.73 and equation 4.13 into our equal-time covariance defin-

ition (Eqn. 4.16) we find the following:

$$\begin{aligned}
\tau_o \frac{d}{dt} \rho_{ij} &= \tau_o \frac{d}{dt} \langle x_i x_j \rangle - \tau_o \langle x_j \rangle \frac{d}{dt} \langle x_i \rangle - \tau_o \langle x_i \rangle \frac{d}{dt} \langle x_j \rangle \\
&= -2 (\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle) + (\langle x_i f_j \rangle - \langle x_i \rangle \langle f_j \rangle) + (\langle f_i x_j \rangle - \langle f_i \rangle \langle x_j \rangle) \\
&= -2\rho_{ij} + \langle \delta x_i \delta f_j \rangle + \langle \delta f_i \delta x_j \rangle ,
\end{aligned} \tag{4.74}$$

where we have abused notation to write  $f_j = f(h_j)$  and  $\delta f_i \equiv f(h_i) - \langle f(h_i) \rangle$ .

Incidentally, if  $i \in X$  and  $j \in E, I$ , and  $\tau_X = \tau_o$  we obtain an equation for the covariance between neurons in the input population and neurons in the E-I population:

$$\begin{aligned}
\tau_o \frac{d}{dt} r_{ij} &= -2 (\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle) + (\langle x_i u_j \rangle - \langle x_i \rangle \langle u_j \rangle) + (\langle x_i f_j \rangle - \langle x_i \rangle \langle f_j \rangle) \\
&= -2r_{ij} + \langle \delta x_i \delta f_j \rangle .
\end{aligned} \tag{4.75}$$

Q.E.D.

### Taylor series expansion

Here we will show that the synaptic drive probability distribution  $p(h'_i, h_j, t)$  can be simplified using a Taylor series expansion (Eqn. 4.23). This distribution is a two-dimensional Gaussian with mean  $(m'_i, m_j)$  and covariance:

$$\mathbf{\Gamma} \equiv \begin{pmatrix} \Sigma_{ii} & \Sigma_{ij} \\ \Sigma_{ji} & \Sigma_{jj} \end{pmatrix} = \begin{pmatrix} \sigma_i^2 & \rho \\ \rho & \sigma_j^2 \end{pmatrix} . \tag{4.76}$$

Here, we have relabelled the elements of  $\mathbf{\Gamma}$  with  $\sigma_i$ ,  $\sigma_j$  and  $\rho$  for notational convenience.

First we need to calculate the inverse of  $\mathbf{\Gamma}$ :

$$\mathbf{\Gamma}^{-1} = \frac{1}{\sigma_i^2 \sigma_j^2 - \rho^2} \begin{pmatrix} \sigma_j^2 & -\rho \\ -\rho & \sigma_i^2 \end{pmatrix} . \tag{4.77}$$

We can then write the synaptic drive probability distribution  $p(h'_i, h_j, t)$  as follows:

$$\begin{aligned}
p(h'_i, h_j, t) &= \exp \left[ -\frac{1}{2} (\delta h'_i, \delta h_j) \cdot \Gamma^{-1} \cdot \begin{pmatrix} \delta h'_i \\ \delta h_j \end{pmatrix} \right] |2\pi\Gamma|^{-1/2} \\
&= \exp \left[ \left( \rho \frac{\delta h'_i \delta h_j}{\sigma_i^2 \sigma_j^2} - \frac{1}{2} \left( \frac{\delta h'_i}{\sigma_i} \right)^2 - \frac{1}{2} \left( \frac{\delta h_j}{\sigma_j} \right)^2 \right) / \left( 1 - \left( \frac{\rho}{\sigma_i \sigma_j} \right)^2 \right) \right] \\
&\quad \times 2\pi (\sigma_i \sigma_j)^{-1} \left( 1 - \rho^2 / (\sigma_i \sigma_j)^2 \right)^{-1/2} \tag{4.78}
\end{aligned}$$

Now, if we make the ansatz that  $\rho / (\sigma_i \sigma_j)^2$  is small, we can simplify  $p(h'_i, h_j, t)$  with a Taylor series expansion:

$$\begin{aligned}
p(h'_i, h_j, t) &= \left( 1 + \rho \frac{\delta h'_i \delta h_j}{\sigma_i^2 \sigma_j^2} + \mathcal{O}(\rho^2) \right) \frac{1}{2\pi \sigma_i \sigma_j} \exp \left[ -\frac{1}{2} \left( \frac{\delta h'_i}{\sigma_i} \right)^2 - \frac{1}{2} \left( \frac{\delta h_j}{\sigma_j} \right)^2 \right] \\
&= \left( 1 + \delta h'_i \delta h_j \left( \frac{\Sigma_{ij}}{\Sigma_{ii} \Sigma_{jj}} \right) + \mathcal{O} \left( \left( \frac{\Sigma_{ij}}{\Sigma_{ii} \Sigma_{jj}} \right)^2 \right) \right) p(h'_i, t) p(h_j, t) . \tag{4.79}
\end{aligned}$$

Q.E.D

### Covariance Integral

Here, we solve the integral in equation 4.24. We begin by writing:

$$\langle \delta h'_i \delta f_j \rangle = \langle h'_i f(h_j) \rangle - \langle h'_i \rangle \langle f(h_j) \rangle . \tag{4.80}$$

The first term can be integrated as follows:

$$\begin{aligned}
\langle h'_i f(h_j) \rangle &= \int h'_i f(h_j) p(\delta h'_i, \delta h_j, t) d\delta h'_i d\delta h_j \\
&\simeq \int h'_i f(h_j) \left( 1 + \delta h'_i \delta h_j \frac{\Sigma_{ij}}{\Sigma_{ii} \Sigma_{jj}} \right) p(\delta h'_i, t) p(\delta h_j, t) d\delta h'_i d\delta h_j \\
&= \langle h'_i \rangle \langle f(h_j) \rangle + \frac{\Sigma_{ij}}{\Sigma_{ii} \Sigma_{jj}} \int h'_i f(h_j) \delta h'_i \delta h_j p(\delta h'_i, t) p(\delta h_j, t) d\delta h'_i d\delta h_j \\
&= \langle h'_i \rangle \langle f(h_j) \rangle + I_i \Sigma_{ij} g_{jj} , \tag{4.81}
\end{aligned}$$

where

$$I_i = \Sigma_{ii}^{-1} \int h'_i \delta h'_i p(\delta h'_i, t) d\delta h'_i = \Sigma_{ii}^{-1} \left( m'_i \langle \delta h'_i \rangle + \langle (\delta h'_i)^2 \rangle \right) = 1 , \tag{4.82}$$

and

$$\begin{aligned}
g_{jj} &= \Sigma_{jj}^{-1} \int f(h_j) \delta h_j p(\delta h_j, t) d\delta h_j \\
&= -f(m_j + \delta h_j) p(\delta h_j, t) \Big|_{-\infty}^{\infty} + \int f'(m_j + \delta h_j) p(\delta h_j, t) d\delta h_j \\
&= \langle f'(m_j + \delta h_j) \rangle \\
&= \frac{\partial \nu_j}{\partial m_j}.
\end{aligned} \tag{4.83}$$

Together, these give

$$\langle h'_i f(h_j) \rangle = \langle h'_i \rangle \langle f(h_j) \rangle + \Sigma_{ij} g_{jj}, \tag{4.84}$$

and finally

$$\langle \delta h'_i \delta f_j \rangle = \Sigma_{ij} g_{jj}. \tag{4.85}$$

Q.E.D.

### Covariance between input population $\mathbf{X}$ and the E-I population

Here, we derive an expression for the covariance  $\mathbf{r}$  between the input population  $\mathbf{X}$  and the E-I population (Eqn. 4.31). We have already derived a dynamical equation for this covariance (Eqn. 4.75):

$$\tau_o \frac{d}{dt} r_{ij} = -2r_{ij} + \langle \delta x_i \delta f_j \rangle \quad ; i \in X \quad j \in E, I. \tag{4.86}$$

We can simplify the term  $\langle \delta x_i \delta f_j \rangle$ , just as we simplified the E-I population covariance equation (Eqn. 4.74). We begin by making the following transformation:

$$\langle \delta \mathbf{h}'_r \otimes \delta \mathbf{f} \rangle = \mathbf{H}_r \cdot \langle \delta \mathbf{x} \otimes \delta \mathbf{f} \rangle, \tag{4.87}$$

where we have defined a 'pseudo synaptic-drive':  $\delta \mathbf{h}'_r = \mathbf{H}_r \cdot \delta \mathbf{x}$ , where  $\mathbf{H}_r$  is an invertible matrix that is chosen so that  $(\delta h'_{ri}, \delta h_j)$  can be treated as a weakly correlated, zero mean two dimensional Gaussian random variable. Now, replacing averages over spikes with averages over synaptic drives, and performing a Taylor expansion, we can write:

$$\langle \delta \mathbf{h}'_r \otimes \delta \mathbf{f} \rangle_{ij} = \Sigma_{ij}^r g_{jj}. \tag{4.88}$$

Now all that remains is to write  $\Sigma^r$  in terms of  $\mathbf{r}$  and  $\mathbf{n}$ :

$$\begin{aligned}
\Sigma^r &= \langle \delta \mathbf{h}'_r \otimes \delta \mathbf{h} \rangle \\
&= \sum_{\mathbf{x}} (\mathbf{H}_r \cdot \delta \mathbf{x}) \otimes (\mathbf{A} \cdot \delta \mathbf{x} + \mathbf{F} \cdot \delta \mathbf{x}) p(\mathbf{x}, t) d\mathbf{x} \\
&= \mathbf{H}_r \mathbf{r} \mathbf{A}^T + \mathbf{H}_r \mathbf{n} \mathbf{F}^T.
\end{aligned} \tag{4.89}$$

Now, substituting equation 4.89 and equation 4.88 into equation 4.87, we can write

$$\langle \delta \mathbf{x} \otimes \delta \mathbf{f} \rangle = \mathbf{H}_r^{-1} (\mathbf{H}_r \mathbf{r} \mathbf{A}^T + \mathbf{H}_r \mathbf{n} \mathbf{F}^T) \mathbf{g} = \mathbf{r} \mathbf{A}^T \mathbf{g} + \mathbf{n} \mathbf{F}^T \mathbf{g}. \tag{4.90}$$

Finally, substituting equation 4.90 into equation 4.86 we find:

$$\tau_0 \frac{d\mathbf{r}}{dt} = -2\mathbf{r} + \mathbf{r} \cdot \mathbf{A}^T \cdot \mathbf{g} + \mathbf{n} \cdot \mathbf{F}^T \cdot \mathbf{g}. \tag{4.91}$$

Q.E.D.

### Equilibrium solution approximation

Here, we show that the term  $\mathbf{e}_\mu^\dagger \cdot \xi \cdot \mathbf{e}_\nu^\dagger$  is much smaller than the other terms  $\mathbf{e}_\mu^\dagger \cdot \boldsymbol{\rho}^{od} \cdot \mathbf{e}_\nu^\dagger$  and  $\mathbf{e}_\mu^\dagger \cdot \boldsymbol{\chi} \cdot \mathbf{e}_\nu^\dagger$  in equation 4.42. Our approach is to calculate the size of each term in equation 4.42 in terms of the E-I population size  $N$  and the average number of synapses per neuron  $K$ .

We begin by writing the spike train covariance in terms of the mean covariance  $\bar{\boldsymbol{\rho}}$  and the covariance standard deviation  $s.d.(\boldsymbol{\rho})$ :

$$\rho_{ij} \equiv \bar{\rho} + s.d.(\boldsymbol{\rho}) \eta_{ij}, \tag{4.92}$$

where averages are population averages and  $\eta_{ij}$  represents quenched fluctuations in the

population covariance. We can now write:

$$\begin{aligned}
\mathbf{e}_\mu^\dagger \cdot \boldsymbol{\rho}^{od} \cdot \mathbf{e}_\nu^\dagger &= \sum_{i \neq j}^N e_{\mu,i}^\dagger \rho_{ij} e_{\nu,j}^\dagger \\
&= \sum_{i \neq j}^N e_{\mu,i}^\dagger (\bar{\rho} + s.d.(\boldsymbol{\rho}) \eta_{ij}) e_{\nu,j}^\dagger \\
&\sim \mathcal{O}(N^2 K^{-1}) .
\end{aligned} \tag{4.93}$$

Here, we have used the fact that  $\bar{\rho} \sim 1/K$  and  $s.d.(\boldsymbol{\rho}) \sim 1/\sqrt{K}$  (Supplementary Methods 4.5.4).

Next we calculate the size of  $\mathbf{e}_\mu^\dagger \cdot \boldsymbol{\chi} \cdot \mathbf{e}_\nu^\dagger$ . In a strongly connected network we have  $\chi_{ij} \sim \mathcal{O}(1/\sqrt{K})$  (Supplementary Methods 4.5.4). Therefore

$$\mathbf{e}_\mu^\dagger \cdot \boldsymbol{\chi} \cdot \mathbf{e}_\nu^\dagger \sim \sum_{i \neq j}^N e_{\mu,i}^\dagger \chi_{ij} e_{\nu,j}^\dagger \sim \mathcal{O}(NK/\sqrt{K}) . \tag{4.94}$$

Finally, we can calculate the size of  $\mathbf{e}_\mu^\dagger \cdot \boldsymbol{\xi} \cdot \mathbf{e}_\nu^\dagger$  by observing that  $\xi_{ii} \sim 1/\sqrt{K}$ . Therefore

$$\mathbf{e}_\mu^\dagger \cdot \boldsymbol{\xi} \cdot \mathbf{e}_\nu^\dagger \sim \sum_i^N e_{\mu,i}^\dagger \xi_{ii} e_{\nu,i}^\dagger \sim \mathcal{O}(NK^{-1/2}) . \tag{4.95}$$

This term is much smaller than the other terms in equation 4.42, as required.

#### 4.5.4 Correlation scaling

Here, we calculate the scaling of the mean correlation  $\bar{\mathbf{C}}$  and the correlation standard deviation  $s.d.(\mathbf{C})$ . Our approach is to calculate the relative size of each term in our covariance equation. We must begin by calculating the size of the covariance  $\mathbf{r}$  between the input population X and the E-I population (Eqn. 4.33). We then calculate the size of the E-I covariance  $\boldsymbol{\rho}$  using equation 4.34 before finally calculating the size of  $\bar{\mathbf{C}}$  and  $s.d.(\mathbf{C})$ .

##### Scaling of covariance $\mathbf{r}$

Here, we calculate the scaling of  $\mathbf{r}$  with  $K$  by quantifying the size of the mean covariances  $r^E$  and  $r^I$  and the covariance standard deviations,  $\sigma_r^E$  and  $\sigma_r^I$ , defined as

follows:

$$r^E \equiv \sum_{i \in X, j \in E} r_{ij} / (N_X N_E) \quad (4.96)$$

$$r^I \equiv \sum_{i \in X, j \in I} r_{ij} / (N_X N_I) \quad (4.97)$$

$$(\sigma_r^E)^2 \equiv r_E^{(2)} - (r^E)^2 \quad (4.98)$$

$$(\sigma_r^I)^2 \equiv r_I^{(2)} - (r^I)^2. \quad (4.99)$$

where

$$r_E^{(2)} \equiv \sum_{i \in X, j \in E} r_{ij}^2 / (N_X N_E) \quad (4.100)$$

$$r_I^{(2)} \equiv \sum_{i \in X, j \in I} r_{ij}^2 / (N_X N_I) \quad (4.101)$$

We begin by calculating the mean covariance  $r^E$  and  $r^I$  by averaging our equilibrium equation for  $\mathbf{r}$  (Eqn. 4.34) using average connectivity (Eqns. 3.10, 3.14):

$$2r^E \simeq w_o \sqrt{K} \left( (g^E W^{EE} p^E) r^E - (g^E W^{EI} p^I) r^I \right) + w_0 \left( g^E W^{EX} p^X / \sqrt{K} \right) \bar{n} \\ + j_o \sum_{i, k \in E, j \in X} g_i J_{ik} r_{kj}^T p^E / (N_E^2 N_X) + j_F \sum_{i, e \in E, j \in X} g_i J_{ij}^F n_{jj} p^x / (N_X^2 N_E) \quad (4.102)$$

$$2r^I \simeq w_o \sqrt{K} \left( (g^I W^{IE} p^E) r^E - (g^I W^{II} p^I) r^I \right) + w_0 \left( g^I W^{IX} p^X / \sqrt{K} \right) \bar{n}, \quad (4.103)$$

where

$$g^Q \equiv \sum_{i \in Q} g_{ii} / N_Q$$

$$\bar{n} \equiv \sum_{i \in X} n_{ii} / N_X.$$

There are no non-zero solutions to these equations if  $r^E, r^I \sim \mathcal{O}(1)$  or if  $r^E, r^I \sim \mathcal{O}(1/\sqrt{K})$ . However, if  $r^E, r^I \sim \mathcal{O}(1/K)$  there is a solution, following Renart et. al. (Renart et al., 2010). This solution can be obtained for large  $K$  by solving the following

equations for  $r^E$  and  $r^I$ :

$$\begin{aligned}
0 &= w_o (g^E W^{EE} p^E) r^E - w_o (g^E W^{EI} p^I) r^I + w_o (g^E W^{EX} p^X / K) \bar{n} \\
&\quad + j_o \sum_{i,k \in E, j \in X} g_i J_{ik} r_{kj}^T p^E / (N_E^2 N_X \sqrt{K}) \\
0 &= w_o (g^I W^{IE} p^E) r^E - w_o (g^I W^{II} p^I) r^I + w_o (g^I W^{IX} p^X / K) \bar{n}
\end{aligned} \tag{4.104}$$

Therefore,

$$r^E, r^I \sim \mathcal{O}(1/K). \tag{4.105}$$

Also, we see that the contribution of structured connectivity to the covariance  $r^E$  and  $r^I$  is the same order of magnitude as the contribution of background connectivity. This is consistent with our previous results (Fig. 4.5).

Similarly we calculate the the covariance standard deviation  $\sigma_r^E$  and  $\sigma_r^I$ . We begin by calculating the typical size of  $r_E^{(2)}$  and  $r_I^{(2)}$ :

$$4r_{ij}^2 = [\mathbf{gAr}^T + \mathbf{gFn}]_{ij}^2$$

There are no non-zero solutions to these equations if  $r_E^{(2)}, r_I^{(2)} \sim \mathcal{O}(1)$  or if  $r_E^{(2)}, r_I^{(2)} \sim \mathcal{O}(1/\sqrt{K})$ . However, if  $r_E^{(2)}, r_I^{(2)} \sim \mathcal{O}(1/K)$  there is a solution and we can write:

$$\begin{aligned}
\mathcal{O}(1/K) &= (w_o W^{EE})^2 p^E r_E^{(2)} + (w_o W^{EI})^2 p^I r_I^{(2)} + (w_o W^{EX})^2 p^X n^{(2)} / K \\
\mathcal{O}(1/K) &= (w_o W^{IE})^2 p^E r_E^{(2)} + (w_o W^{II})^2 p^I r_I^{(2)} + (w_o W^{IX})^2 p^X n^{(2)} / K
\end{aligned}$$

where

$$n^{(2)} \equiv \sum_{i \in X} n_{ii} / N_X.$$

Therefore, using the fact  $r^E, r^I \sim \mathcal{O}(1/K)$ , we can see from equations 4.98 and 4.99 that:

$$\begin{aligned}
\sigma_r^E &\sim \mathcal{O}(1/\sqrt{K}) \\
\sigma_r^I &\sim \mathcal{O}(1/\sqrt{K})
\end{aligned}$$

### Scaling of covariance $\rho$

Next, we calculate the size of the covariance  $\rho$  by calculating the mean covariances  $\rho^{QP}$  and the covariance standard deviation  $\sigma_\rho^{QP}$ , defined as follows:

$$\rho^{QP} \equiv \sum_{i \in Q, j \in P} \rho_{ij} / (N_Q N_P) \quad (4.106)$$

$$(\sigma_\rho^{QP})^2 \equiv \sum_{i \in Q, j \in P} \rho_{ij}^2 / (N_Q N_P) - (\rho^{QP})^2 \quad (4.107)$$

This calculation is similar to the previous calculation, except that there are more terms:

$$\begin{aligned} \rho^{EE} = & \sqrt{K} [w_o (g^E W^{EE} p^E) \rho^{EE} - w_o (g^E W^{EI} p^I) \rho^{IE} + w_o (g^E W^{EX} p^X) r^E] \\ & + w_o (W^{EE} p^E / \sqrt{K}) \left( \sum_{i \in E} g_i \rho_{ii} / N_E \right) + j_o \sum_{i, k, j \in E} g_i J_{ik} \rho_{kj} p^E / N_E^3 \\ & + j_F \sum_{i, j \in E, k \in X} g_i J_{ik}^F r_{kj} p^E / (N_E^2 N_X) \end{aligned} \quad (4.108)$$

$$\begin{aligned} 2\rho^{EI} = & \sqrt{K} [w_o (g^E W^{EE} p^E) \rho^{EI} - w_o (g^E W^{EI} p^I) \rho^{II} + w_o (g^E W^{EX} p^X) r^I \\ & w_o \rho^{EE} (p^E W^{EI} g^I) + w_o \rho^{EI} (p^I W^{II} g^I) + w_o r^E (p^X W^{EX} g^E)] \\ & - w_o (W^{EI} p^I / \sqrt{K}) \left( \sum_{i \in I} g_i \rho_{ii} / N_I \right) + w_o (W^{EI} p^E / \sqrt{K}) \left( \sum_{i \in E} g_i \rho_{ii} / N_E \right) \\ & + j_o \sum_{i, k \in E, j \in I} g_i J_{ik} \rho_{kj} p^I / (N_E^2 N_I) + j_F \sum_{i \in E, k \in X, j \in I} g_i J_{ik}^F r_{kj} p^I / (N_E N_I N_X) \end{aligned} \quad (4.109)$$

$$\begin{aligned} \rho^{II} = & \sqrt{K} [w_o (g^I W^{IE} p^E) \rho^{EI} - w_o (g^I W^{II} p^I) \rho^{II} + w_o (g^I W^{IX} p^X) r^I] \\ & - w_o (W^{II} p^I / \sqrt{K}) \left( \sum_{i \in I} g_i \rho_{ii} / N_I \right). \end{aligned} \quad (4.110)$$

Again, there are no non-zero solutions to these equations if  $\rho^{EE}, \rho^{EI}, \rho^{IE}, \rho^{II} \sim \mathcal{O}(1)$  or if  $\rho^{EE}, \rho^{EI}, \rho^{IE}, \rho^{II} \sim \mathcal{O}(1/\sqrt{K})$ . However, if  $\rho^{EE}, \rho^{EI}, \rho^{IE}, \rho^{II} \sim \mathcal{O}(1/K)$ , there is a solution. This solution can be obtained by solving the following equations for

$\rho^{EE}, \rho^{EI}, \rho^{IE}$  and  $\rho^{II}$ :

$$\begin{aligned}
0 = & w_o (g^E W^{EE} p^E) \rho^{EE} - w_o (g^E W^{EI} p^I) \rho^{IE} + w_0 (g^E W^{EX} p^X) r^E \\
& + w_o (W^{EE} p^E / K) \left( \sum_{i \in E} g_i \rho_{ii} / N_E \right) + j_o \sum_{i, k, j \in E} g_i J_{ik} \rho_{kj} p^E / (N_E^3 \sqrt{K}) \\
& + j_F \sum_{i, j \in E, k \in X} g_i J_{ik}^F r_{kj} p^E / (N_x N_E^2 \sqrt{K})
\end{aligned} \tag{4.111}$$

$$\begin{aligned}
0 = & w_o (g^E W^{EE} p^E) \rho^{EI} - w_o (g^E W^{EI} p^I) \rho^{II} + w_0 (g^E W^{EX} p^X) r^I \\
& + w_o \rho^{EE} (p^E W^{EI} g^I) + w_o \rho^{EI} (p^I W^{II} g^I) + w_0 r^E (p^X W^{EX} g^E) \\
& - w_o (W^{EI} p^I / K) \left( \sum_{i \in I} g_i \rho_{ii} / N_I \right) + w_o (W^{EI} p^E / K) \left( \sum_{i \in E} g_i \rho_{ii} / N_E \right) \\
& + j_o \sum_{i, k \in E, j \in I} g_i J_{ik} \rho_{kj} p^I / (N_E^2 N_I \sqrt{K}) + j_F \sum_{i \in E, k \in X, j \in I} g_i J_{ik}^F r_{kj} p^I / (N_E N_I N_X \sqrt{K})
\end{aligned} \tag{4.112}$$

$$\begin{aligned}
0 = & w_o (g^I W^{IE} p^E) \rho^{EI} - w_o (g^I W^{II} p^I) \rho^{II} + w_0 (g^I W^{IX} p^X) r^I \\
& - w_o (W^{II} p^I / K) \left( \sum_{i \in I} g_i \rho_{ii} / N_I \right).
\end{aligned} \tag{4.113}$$

Therefore,

$$\rho^{EE}, \rho^{EI}, \rho^{IE}, \rho^{II} \sim \mathcal{O}(1/K). \tag{4.114}$$

Also, we see that the contribution of structured connectivity to the covariance  $\rho^{EE}, \rho^{EI}, \rho^{IE}$  and  $\rho^{II}$  is the same order of magnitude as the contribution of background connectivity. This is consistent with our previous results (Fig. 4.5).

We calculate the size of covariance standard deviation  $\sigma_\rho^{QP}$ , just as we calculated  $\sigma_r^E$  and  $\sigma_r^I$ . We find that:

$$\sigma_\rho^{QP} \sim \mathcal{O}(1/\sqrt{K}) \tag{4.115}$$

Finally, we can calculate  $\bar{\mathbf{C}}$  and the correlation standard deviation *s.d.* ( $\mathbf{C}$ ). This is straightforward, because  $\rho_{ii} \sim \mathcal{O}(1)$ . Therefore, we can conclude from equation 4.114 and equation 4.115 that

$$\begin{aligned}
\bar{\mathbf{C}} & \sim 1/K \\
s.d.(\mathbf{C}) & \sim 1/\sqrt{K}.
\end{aligned} \tag{4.116}$$

Q.E.D.



Part II

Computation



## Chapter 5

# Information in Balanced Networks

### 5.1 Introduction

Naturally occurring images typically contain many orientated bars (Olshausen and Field, 1996; Bell and Sejnowski, 1997). In *orientation selectivity*, the orientation of these bars must be detected. It is important for the cortex to perform this computation because it is an essential building block for many other spatial computations such as object recognition (Ferster and Miller, 2000). In this chapter we investigate how orientation selectivity can be implemented in balanced networks.

Information quantifies the ability of a network to perform orientation selectivity (Averbeck et al., 2006). If information is high, computational performance can be good. If information is low, it is difficult to detect orientation, and computational performance is poor.

We quantify the ability of a network to perform orientation selectivity by calculating information analytically. Information cannot be measured experimentally because it requires measurements of correlations between all neuron pairs (Ecker et al., 2010). Using the correlation equations that we derived in chapter 4, we can derive a mathematical relationship between information and network connectivity. This allows us to answer questions about the role of connectivity, correlations and spike train irregularity in spatial computation.

We begin by considering the role of structured connectivity. We find that structured connectivity increases information dramatically, by tuning neural responses to orientation. Therefore, orientation tuning is more than phenomenological - it is important for

cortical function. This is interesting because the structured connectivity in our balanced network model is much weaker than background connectivity. The orientation tuning produced by structured connectivity is similar to the orientation tuning recorded in many visual cortex neurons (Hubel and Wiesel, 1959).

Next, we investigate the computational role of background connectivity. Dynamically, background connectivity is responsible for synaptic background noise and spike train irregularity, so we expect it to harm computation (van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998). As expected, we find that background connectivity reduces information. However, this computational damage only occurs over a narrow range of contrasts. At most contrast levels, information is higher in networks with background connectivity. This exemplifies a performance-stability trade-off: the cost of high performance is a lower operating range; the trade-off arises because while background connectivity produces noise, it also maintains the network in a highly informative state. It thus provides computational stability, just as it provides dynamic stability. Furthermore, we find that the contrast invariance of orientation tuned cells is a signature of this computational stability.

Finally, we investigate the role of correlations in computation. This is an important question because correlations may increase or decrease information dramatically, depending on which neurons are correlated (Wu et al., 2001; Nirenberg and Latham, 2003). There has been much debate about this because information calculations, both in theory and experiment would be much easier if the contribution of correlations could be ignored (Averbeck et al., 2006). We find that correlations cannot be ignored, despite the fact that correlations are weak in balanced networks. Furthermore, we find that information increases with correlations. However, this relationship is not causal. The connectivity that produces correlations is the same connectivity that produces orientation tuning. The complicated relationship between correlations and tuning curves exemplifies the difficulty in prescribing computational functions to particular spike train statistics.

## 5.2 Orientation selectivity in balanced networks

We quantify orientation selectivity in a balanced network of recurrently connected excitatory and inhibitory neurons. We use the same network model that we have analysed

throughout this thesis (Fig. 5.1). The neurons in the model produce irregular, asynchronous spike trains and have orientation tuning (Chapter 3), similar to neurons in V1 (Hubel and Wiesel, 1959; Softky and Koch, 1993; Ecker et al., 2010).

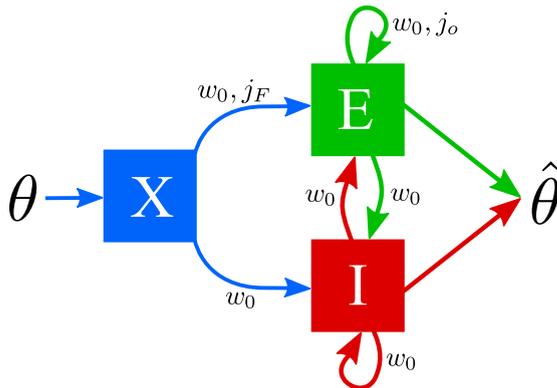


Figure 5.1: Orientation selectivity

Our network model consists of three populations of neurons; an input population (X), a population of excitatory neurons (E) and a population of inhibitory neurons (I). The input population contains neurons that encode the orientation  $\theta$  of a stimulus. These neurons project onto our excitatory and inhibitory populations which provide an estimate  $\hat{\theta}$  of the input orientation. As before, background connectivity is parameterized by  $w_o$ . Structured connectivity between excitatory neurons is parameterised by  $j_F$  and  $j_o$ .

As before, network connectivity consists of weak structured connectivity embedded in strong background connectivity (Eqns. 3.10, 3.14). The structured connectivity, parameterised by  $j_o$  and  $j_F$  is responsible for producing orientation tuned responses. The background connectivity, parameterised by  $w_o$  is responsible for the irregular asynchronous spiking activity.

The balanced network spiking activity encodes the orientation  $\theta$  of an edge from a visual stimulus. In orientation selectivity, this spiking activity must be decoded to provide an estimate,  $\hat{\theta}$  of the edge's orientation. We will quantify the ability of a balanced network to support this computation by calculating information.

This model may be loosely considered to be a V1 orientation hypercolumn model. However, we can easily treat it as a model for any feature tuned cortical region. Indeed, most of our results and analysis can be generalised to the entire cortex. However, this modeling flexibility comes at a price. There are many V1 phenomena and simple cell phenomena that our model does not currently capture, such as color tuning and adaptation. These phenomena can be easily accommodated without adversely effecting

our analysis. In truth, however, we are more interested in discovering general principles of cortical computation than idiosyncratic properties of any particular region of cortex.

### 5.3 Information analysis

We quantify orientation selectivity performance by calculating the information that a balanced network contains about the angular variable,  $\theta$ , as a function of recurrent connectivity  $\mathbf{A}$ , feedforward connectivity  $\mathbf{F}$  and input population firing rate  $\mathbf{u}(\theta)$ :

$$I(\theta) = f(\mathbf{A}, \mathbf{F}, \mathbf{u}(\theta)) . \quad (5.1)$$

Ideally, we would like to calculate Fisher information, which provides a lower bound on the variance of the optimal unbiased estimator. However, this requires the calculation of the full probability distribution of activity  $p(\mathbf{x}|\theta)$ , which is intractable for most network models.

Instead we calculate a related quantity called *linear Fisher information*. This calculation is surprisingly simple, as it only requires knowledge of the spike correlations and firing rates, which we have calculated already.

#### 5.3.1 Linear Fisher Information

*Linear Fisher information* is defined as the inverse variance of a locally unbiased optimal linear estimator (Deneve et al., 1999; Sompolinsky et al., 2001). It provides a good measure of a network’s ability to perform orientation selectivity, and it is analytically tractable, as we shall see (Seriès et al., 2004).

To calculate linear Fisher information, we must first derive an expression for an unbiased optimal linear estimator. When we say *unbiased*, we mean that the estimator is correct on average, and when we say *optimal*, we mean that the variance of the estimates is as small as possible. A linear estimator,  $\mathbf{w}$ , is an estimator that can be used to estimate  $\theta$  from population activity in the following way:

$$\hat{\theta} = \mathbf{w} \cdot (\mathbf{x}(\theta) - \boldsymbol{\nu}(\theta_0)) , \quad (5.2)$$

where

$$\boldsymbol{\nu}(\theta_0) = \langle \mathbf{x}(\theta_0) \rangle, \quad (5.3)$$

is the firing rate of the population in response to orientation  $\theta_0$ .

We begin by requiring our estimate to be correct on average, or more specifically, we require it to be locally unbiased around  $\theta_o$ :

$$\left. \frac{\partial}{\partial \theta} \langle \hat{\theta} \rangle \right|_{\theta_o} = 1. \quad (5.4)$$

This is satisfied if we constrain  $\mathbf{w}$  to have the following form:

$$\mathbf{w} = (\tilde{\mathbf{w}} \cdot \boldsymbol{\nu}'(\theta_o))^{-1} \tilde{\mathbf{w}}, \quad (5.5)$$

where  $\tilde{\mathbf{w}}$  is another linear decoder and we have used a prime to denote a derivative with respect to  $\theta$ .

Next, we require the variance of our estimates to be as small as possible. The variance of  $\hat{\theta}$  can be written in terms of  $\tilde{\mathbf{w}}$ :

$$var(\hat{\theta}) = \tilde{\mathbf{w}} \cdot \boldsymbol{\rho} \cdot \tilde{\mathbf{w}} / (\tilde{\mathbf{w}} \cdot \boldsymbol{\nu}'(\theta_o))^2. \quad (5.6)$$

The decoder with the smallest variance can be derived by minimising this variance analytically:

$$\left. \frac{\partial}{\partial \tilde{\mathbf{w}}} var(\hat{\theta}) \right|_{\tilde{\mathbf{w}}_{opt}} = \left[ 2\boldsymbol{\rho} \cdot \tilde{\mathbf{w}} / (\tilde{\mathbf{w}} \cdot \boldsymbol{\nu}'(\theta_o))^2 - 2\boldsymbol{\nu}'(\theta_o) (\tilde{\mathbf{w}} \cdot \boldsymbol{\rho} \cdot \tilde{\mathbf{w}}) / (\tilde{\mathbf{w}} \cdot \boldsymbol{\nu}'(\theta_o))^3 \right] \Big|_{\tilde{\mathbf{w}}_{opt}} = 0. \quad (5.7)$$

Solving this, we find that

$$\tilde{\mathbf{w}}_{opt} \propto \boldsymbol{\rho}^{-1} \cdot \boldsymbol{\nu}'(\theta_o). \quad (5.8)$$

We obtain an exact expression for the locally unbiased optimal linear estimator  $\mathbf{w}_{opt}$  by inserting equation 5.8 into equation 5.5:

$$\mathbf{w}_{opt} = \boldsymbol{\rho}^{-1} \cdot \boldsymbol{\nu}'(\theta_o) / (\boldsymbol{\nu}'(\theta_o) \cdot \boldsymbol{\rho}^{-1} \cdot \boldsymbol{\nu}'(\theta_o)). \quad (5.9)$$

This optimal estimator can decode the orientation encoded in the spiking activity of a balanced network. The performance of this decoder is best in a locally linear region

around  $\theta_0$ . We will not actually use this decoder in this work, because we do not need it to calculate information. Nonetheless, we have verified that it can perform orientation selectivity optimally (data not shown).

It is now straightforward to calculate linear Fisher information. By inserting equation 5.8 into equation 5.6 we can write:

$$I(\theta_0) = 1/\text{var}(\hat{\theta}) = \boldsymbol{\nu}'(\theta_0) \cdot \boldsymbol{\rho}^{-1} \cdot \boldsymbol{\nu}'(\theta_0) . \quad (5.10)$$

Linear Fisher information is a good measure of the ability of our network to perform orientation selectivity because accurate orientation selectivity corresponds to high information. Also, linear Fisher information is tractable because it only depends on correlations, which we have already calculated, and tuning curve slopes  $\boldsymbol{\nu}'(\theta_0)$ , which we will calculate in the next section.

### 5.3.2 Tuning curve slopes

The sensitivity of firing rates to changes in  $\theta$  is represented by  $\boldsymbol{\nu}'(\theta_0)$ . We can write down an equation for  $\nu'_i$  by differentiating our firing rate equation (Eqn. 4.13):

$$\nu'_i = (2\pi)^{-1/2} \exp[-m_i^2/2\sigma_i^2] \partial_\theta (m_i/\sigma_i) \quad (5.11)$$

$$= g_{ii} \partial_\theta m_i - g_{ii} (m_i / (2\sigma_i^2)) \partial_\theta \sigma_i^2 , \quad (5.12)$$

where  $g_{ii}$  is the *gain* of neuron  $i$ , the same gain function that appeared in our correlation calculation (Eqn. 4.25).

Using the fact that correlations cancel in balanced networks, we calculate the derivative of the synaptic drive variance (Eqn. 4.10):

$$\partial_\theta \sigma_i^2 \simeq \sum_j A_{ij}^2 \partial_\theta \rho_{jj} + \sum_j F_{ij}^2 \partial_\theta n_{jj} \quad (5.13)$$

$$= \sum_j A_{ij}^2 (1 - 2\nu_j) \partial_\theta \nu_j + \sum_j F_{ij}^2 (1 - 2u_j) \partial_\theta u_j . \quad (5.14)$$

We can now write down a linear equation for  $\boldsymbol{\nu}'$ , the tuning curve slopes of neurons in our network:

$$\boldsymbol{\nu}' = \mathbf{g}\tilde{\mathbf{A}} \cdot \boldsymbol{\nu}' + \mathbf{g}\tilde{\mathbf{F}} \cdot \mathbf{u}' , \quad (5.15)$$

where

$$\tilde{A}_{ij} = A_{ij} - \frac{m_j}{2\sigma_i^2} A_{ij}^2 (1 - 2\nu_j) , \quad (5.16)$$

and

$$\tilde{F}_{ij} = F_{ij} - \frac{m_j}{2\sigma_i^2} F_{ij}^2 (1 - 2u_j) . \quad (5.17)$$

This equation shows that the sensitivity of firing rates to  $\theta$  is equal to the sensitivity of firing rates to the mean synaptic drive  $g_{ii}$  multiplied by the sensitivity of the effective mean synaptic drive to  $\theta$ . Together, these terms determine the sensitivity of a balanced network. Solving this gives:

$$\boldsymbol{\nu}' = \left( \mathbf{1} - \mathbf{g}\tilde{\mathbf{A}} \right)^{-1} \mathbf{g}\tilde{\mathbf{F}} \cdot \mathbf{u}' . \quad (5.18)$$

This expression characterises how the population firing rate depends on the orientation  $\theta$ . Here,  $\mathbf{u}'$  tells us how changes to  $\theta$  effect the input population firing rate.

### 5.3.3 Input information and output information

We can now write down an expression for output linear Fisher information. Substituting our expression for  $\boldsymbol{\nu}'$  (Eq. 5.18) into our linear Fisher information expression (Eqn. 5.10) we find:

$$I_{out}(\theta_o) = \mathbf{u}' \cdot \tilde{\mathbf{F}}^T \mathbf{g} \left( \left( \mathbf{1} - \mathbf{g}\tilde{\mathbf{A}} \right) \boldsymbol{\rho} \left( \mathbf{1} - \tilde{\mathbf{A}}^T \mathbf{g} \right) \right)^{-1} \mathbf{g}\tilde{\mathbf{F}} \cdot \mathbf{u}' , \quad (5.19)$$

where  $\boldsymbol{\rho}$  is the matrix of covariances between neuron pairs that we calculated in the last chapter (Eqn. 4.45).

To evaluate the computational performance of our network, we must compare this output information to the input information. Input information is considerably easier to derive than output information. Inserting our expression for the input covariance  $\mathbf{n}$  (Eqn. 4.35) into our linear Fisher information expression (Eqn. 5.10) we find:

$$I_{in}(\theta_o) = \mathbf{u}' \cdot \mathbf{n}^{-1} \cdot \mathbf{u}' = \sum_{i \in X}^{N_X} (u'_i)^2 / (u_i - u_i^2) . \quad (5.20)$$

Incidentally, this is also the input Fisher information.

We quantify the ability of a network to perform orientation selectivity by calculating the ratio of output information to input information. Throughout the remainder of this work we will simply refer to this quantity as the information:

$$F(\theta_0) = I_{out}(\theta_o) / I_{in}(\theta_o) . \quad (5.21)$$

This is a complicated function of recurrent connectivity  $\mathbf{A}$ , feedforward connectivity  $\mathbf{F}$  and the firing rate of the input population  $\mathbf{u}(\theta)$ .

## 5.4 Results

Although we now have an equation that relates information directly to connectivity and input (Eqn. 5.21), it is difficult to interpret this equation. The reason for this is that information is naturally expressed as a function of correlations  $\rho$  and tuning curve slopes  $\nu'$ . However, these statistics are inextricably linked to each other. They are both determined by the same connectivity and input. Here we will attempt to unravel our information equation so that we can understand how correlations and connectivity contribute to information.

### 5.4.1 The origin of information

Immediately, we can see that there are two sources of information in our network. The first source is the external population. Feedforward connectivity  $\mathbf{F}$  transmits information from the external population to the E-I population. This connectivity must match the structure of the input if it is to successfully transmit information. The second source of information is past network activity. This information is transmitted from the past to the present through recurrent network connectivity  $\mathbf{A}$ . Again, this connectivity must match the structure of the balanced network spiking activity if it is to successfully transmit information from the past.

The contribution of both sources is modulated by the neural gain function  $\mathbf{g}$ . The neural gain captures the sensitivity of individual neurons to changes in the average synaptic drive. If the gain is small, information will be low, regardless of the network connectivity.

Finally, information is strongly dependent on the amount of noise,  $\rho$ , in our network.

As we know, balanced networks are highly irregular (Softky and Koch, 1993). Therefore, we expect intrinsic neuronal variability represented by  $\rho^d$  to reduce information. Balanced networks are also asynchronous, so we do not expect covariance  $\rho^{od}$  to have an enormous impact on information, though it may contribute to some extent (Shadlen and Newsome, 1994). We will discuss the contribution of all these factors in further detail.

### 5.4.2 Simulations and predictions

Before using our information expression (Eqn. 5.21) to study orientation selectivity, we must check that the approximations we make in our analysis are acceptable. We do this by comparing our predicted information to information estimated from network simulations.

We would like to measure linear Fisher information directly. However, linear Fisher information characterises the sensitivity of an unbiased decoder to infinitesimal changes in parameter  $\theta$  and we cannot measure infinitesimally small changes. Instead, we measure firing rates and correlations in two simulations of a network; one receiving input encoding orientation  $\theta^+ = \theta + \Delta\theta/2$  and the other receiving input encoding  $\theta^- = \theta - \Delta\theta/2$ . We can then use these measurements to calculate the *output Fisher Criteria* and the *input Fisher Criteria*:

$$I_{out}^m(\theta) = \left( \frac{\boldsymbol{\nu}(\theta^+) - \boldsymbol{\nu}(\theta^-)}{\Delta\theta} \right) \cdot \left( \frac{\boldsymbol{\rho}(\theta^+) + \boldsymbol{\rho}(\theta^-)}{2} \right)^{-1} \cdot \left( \frac{\boldsymbol{\nu}(\theta^+) - \boldsymbol{\nu}(\theta^-)}{\Delta\theta} \right), \quad (5.22)$$

$$I_{in}^m(\theta) = \sum_{j \in X}^{N_X} \left( \frac{\nu_j(\theta^+) - \nu_j(\theta^-)}{\Delta\theta} \right)^2 \left( \frac{2}{\rho_{jj}(\theta^+) + \rho_{jj}(\theta^-)} \right). \quad (5.23)$$

The ratio of these gives a good measure of *Information*, if  $\Delta\theta$  is small:

$$F^m = \frac{I_{out}^m(\theta)}{I_{in}^m(\theta)}. \quad (5.24)$$

It is much more difficult to estimate information than firing rates or correlations because the Fisher Criteria depends on the difference  $\nu_i(\theta^+) - \nu_i(\theta^-)$  between two very similar firing rates. We must estimate these firing rates to a precision that is much higher than the difference. This requires very long simulations.

As before, we simulate our network using the Gillespie Algorithm (Gillespie, 1977).

We find that we can accurately predict the information in balanced networks using equation 5.21, for a wide range of network parameters (Fig. 5.2). We can now proceed to study the role of correlations, structured connectivity, background connectivity and spike train irregularity in orientation selectivity.

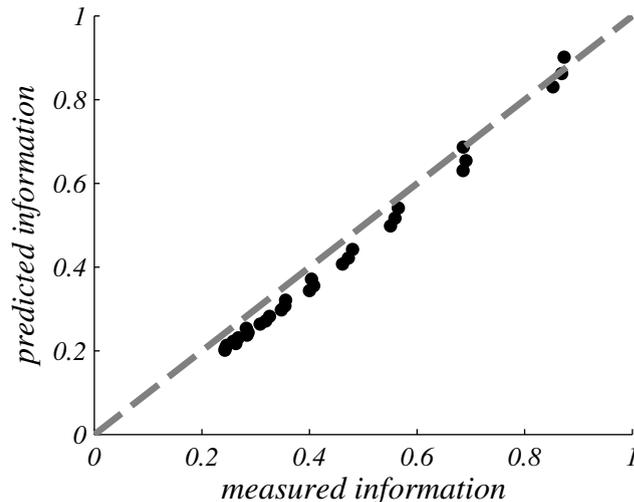


Figure 5.2: Predicted information

Predicted information matches measured information for a wide range of networks with different levels of structure and background connectivity ( $w_o \in [0.5, 2]$ ,  $j_o \in w_o \times [0, 2]$ ). The average percentage error is  $\sum_i^{N_s} |(F_i - F_i^m) / F_i^m| \times 100 / N_s = 10.7\%$ . Each measurement is obtained in a 1000 second network simulation using  $\Delta\theta = 10^\circ$ .

### 5.4.3 The contribution of structured connectivity to information

We begin by investigating the contribution of structured connectivity to information. We have already demonstrated that weak structured connectivity can produce orientation tuned responses in balanced networks (Fig. 3.5). However, this does not imply that structured connectivity increases information. It may be that orientation tuning is an epiphenomenon, and that correlations and noise completely obstruct their contribution to computation.

We calculate information (Eqn. 5.21) for a series of networks containing increasing amounts of structure (parameterised by  $j_0$ ). For each calculation we adjust the spiking thresholds  $\theta^E$  and  $\theta^I$  so that the mean firing rate of the network is constant. This allows us to assess the contribution of structure, independent of changes to the mean firing rate. We find that structured connectivity increases information dramatically (Fig. 5.3). This is a clear demonstration that weak structured connectivity embedded

in random connectivity can be useful for computations such as orientation selectivity and that orientation tuning is useful rather than being an epiphenomenon.

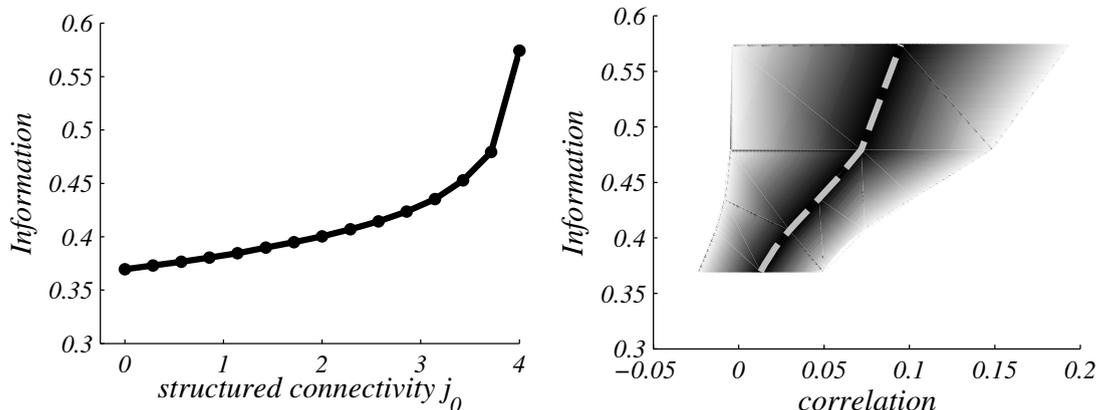


Figure 5.3: Structure and information

Weak structured connectivity embedded in strong background connectivity increases information (left). Correlations that result from structured connectivity increase as information increases (right). The shaded area corresponds to the correlation mean (dashed line)  $\pm$  correlation standard deviation. Spiking thresholds are chosen so that the mean excitatory firing rate is  $\nu^E = 0.2$  for all  $j_0$ .

We also find that correlations originating from structured connectivity increase with information (Fig. 5.3). However, this does not mean that correlations cause information to increase. Structure also increases orientation tuning, which is known to increase information (Sompolinsky and Shapley, 1997). Correlations originating from structure are more phenomenological than functional.

We can understand these results mathematically, by calculating the contribution of structured connectivity to the network response. We begin by writing the eigenvalues  $\kappa_\mu$  and eigenvectors  $\mathbf{d}_\mu$  of structured connectivity  $\mathbf{J}/K$  in the following form:

$$K^{-1}\mathbf{J}\mathbf{d}_\mu = \kappa_\mu\mathbf{d}_\mu \quad (5.25)$$

We have chosen the structured connectivity so that its eigenvectors  $\mathbf{d}_\mu$  are similar to the mean network input. When network input is projected through structured connectivity, the components that encode  $\theta$  are amplified:

$$\mathbf{A}\mathbf{d}_\mu = \left( \frac{w_o}{\sqrt{K}}\mathbf{W} + \frac{j_o}{K}\mathbf{J} \right) \mathbf{d}_\mu = \mathbf{b}_\mu + j_o\kappa_\mu\mathbf{d}_\mu, . \quad (5.26)$$

Here,  $\mathbf{b}_\mu$  is the noisy contribution of random background connectivity. The size of

the amplification is determined by the size of the connectivity eigenvalues  $\kappa_\mu$ , and this structure size  $j_o$ . This explains how structured connectivity, parameterised by  $j_o$  can be used to amplify particular inputs that match the structure, thereby increasing computational performance.

#### 5.4.4 The contribution of background connectivity to information

A growing body of experimental evidence indicates that background connectivity is responsible for a wide range of dynamical phenomena in the cortex (Shu and Hasenstaub, 2003; Haider et al., 2006). However, the functional role of this connectivity is not understood and is currently the subject of great debate and investigation in theoretical neuroscience (Destexhe and Contreras, 2006).

We find that background connectivity contributes to orientation selectivity by providing computational stability. To demonstrate this, we calculate information in a series of networks receiving inputs of increasing contrast (Fig. 5.4, top). Information is high for a wide range of contrasts in networks with background connectivity, whereas networks without background connectivity have high information for a narrow range of contrast levels only. This represents a performance-stability trade off. Networks without background connectivity can have higher performance levels but pay the price of computational instability, whereas the cost of computational stability is reduced performance.

This computational stability is a consequence of the fact that firing rates are not too small or too large for a wide range of contrasts, so information can be transmitted from the synaptic drives into spike trains (Fig. 5.4, middle). Contrast invariance is a signature of this computational stability. Networks with background connectivity have contrast invariant tuning curves, whereas networks without background connectivity are not contrast invariant, as we saw in chapter 3 (Fig. 3.6).

The optimal amount of background connectivity depends on contrast (Fig. 5.4, bottom). At most contrast levels, some background connectivity increases information. However, if there is too much background connectivity, information decreases.

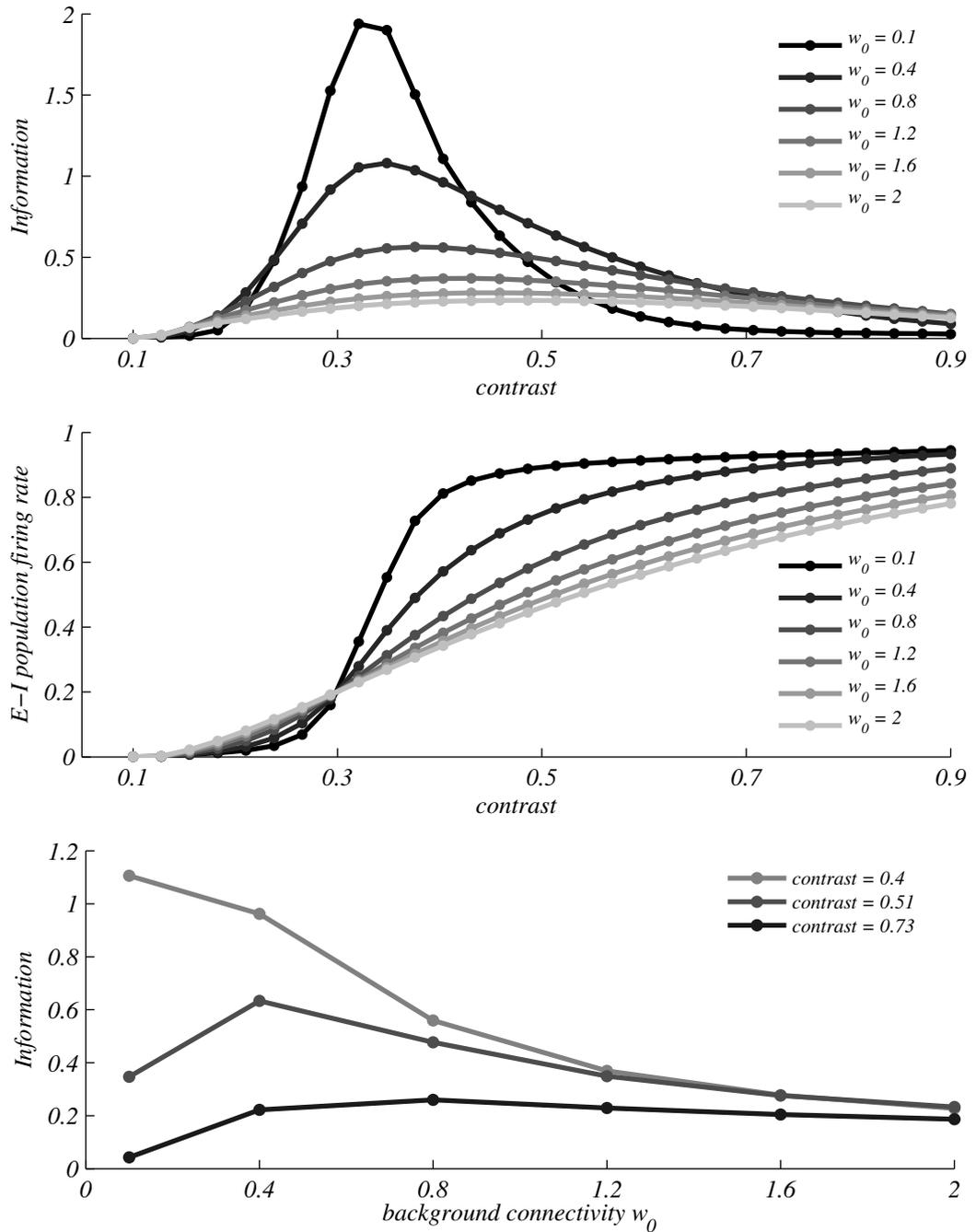


Figure 5.4: Contrast and information

Background connectivity provides computational stability - information is relatively high for a wide range of contrasts in networks with background connectivity, whereas networks without background connectivity have low information for most contrast levels (top). This is a consequence of the fact that firing rates are either very low or close to maximum for most contrast levels in networks without background connectivity (middle). The optimal amount of background connectivity depends on contrast (bottom). The average information for  $w_0 = 0.1$  is  $I_{avg} = 0.43$ , for  $w_0 = 0.4$  it is  $I_{avg} = 0.45$  and for  $w_0 = 0.8$  it is  $I_{avg} = 0.33$ . Thresholds are chosen so that the mean excitatory firing rate is  $\overline{\nu^E} = 0.2$  at contrast  $c = 0.3$ , for all networks.

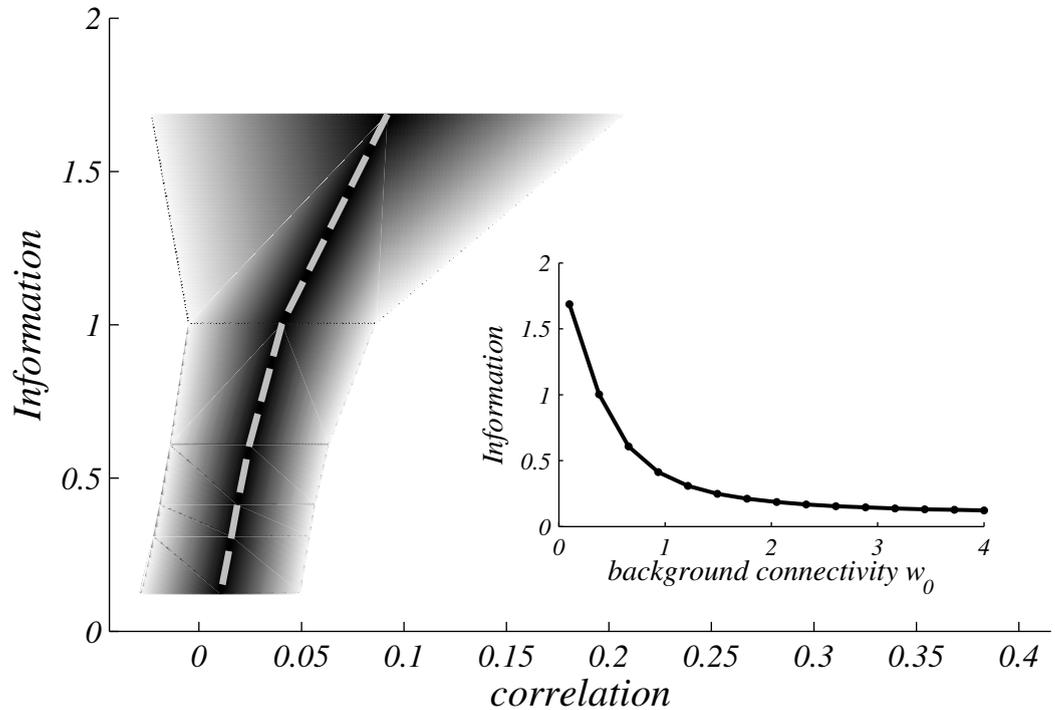


Figure 5.5: Background and information

Correlations that result from background connectivity decrease as information decreases. The shaded area corresponds to the correlation mean (dashed line)  $\pm$  correlation standard deviation. This relationship between correlations and information is particular to a choice of contrast and mean firing rate, where background connectivity, parameterised by  $w_0$  decreases information (inset).

Correlations originating from background connectivity increase with information at the contrast for which information is largest (Fig. 5.5). Again, we cannot conclude from this that correlations cause information to increase. Indeed, for a different choice of contrast, the relationship between information and correlation can be more complicated.

We can understand these results mathematically, by looking at the information equation (Eqn. 5.21). In this equation, we see that the contribution of connectivity and spike train correlations to information are modulated by the gain,  $\mathbf{g}$  (Eqn. 4.25). This is a diagonal matrix that represents the sensitivity of firing rates to the mean synaptic drive. When the elements of the gain are small, information is small, regardless of what connectivity we use. If we can understand how background connectivity, parameterised by  $w_0$  affects the gain, we can begin to understand how background connectivity affects information. In particular, we can understand why the optimal amount of background connectivity is zero for some contrasts and non-zero for other contrasts.

The amount of background connectivity that maximizes the gain of neuron  $i$  can be calculated by solving the following equation:

$$\partial_{w_o} g_{ii} = -\frac{g_{ii}}{\sigma_i} \left( m_i \partial_{w_o} \left( \frac{m_i}{\sigma_i} \right) + \partial_{w_o} \sigma_i \right) = 0. \quad (5.27)$$

This equation is very general so it is difficult to solve directly, and the exact solution is difficult to interpret. However, there are particular cases that are analytically tractable. For example, at a fixed firing rate we can use equation 4.13 to show that:

$$\partial_{w_o} \left( \frac{m_i}{\sigma_i} \right) = \partial_{w_o} (\phi^{-1}(\nu_i)) = 0. \quad (5.28)$$

Now, using this result, we can calculate amount of background connectivity that maximizes the gain of neuron  $i$ :

$$\partial_{w_o} g_{ii} = -\frac{g_{ii}}{\sigma_i} \partial_{w_o} \sigma_i = -\frac{g_{ii}}{2\sigma_i^2} \partial_{w_o} \sigma_i^2. \quad (5.29)$$

We can conclude that at a fixed firing rate, the gain is maximized when  $w_o = 0$ , because the synaptic drive variance  $\sigma_i^2$  of a neuron is always smaller in a network without background connectivity (Eqn. 4.10) than in a network with background connectivity. This fixed firing rate solution corresponds to the information peak in figure 5.4 (top). Information is largest when there is no background connectivity ( $w_0 = 0$ ) and decreases as background connectivity increases.

It should be noted that this is not the complete story. The gain is a diagonal matrix, so its shape across the population also contributes to information. Also, correlations depend on background connectivity and contribute to information. However, all of these factors are modulated by the elements of the gain matrix, so if the gain is small, the contribution of these terms will be small.

As we have discussed, networks without background connectivity have higher information across a narrow range of contrasts. However, at most contrast levels, information is higher in networks with background connectivity, because the firing rate is neither too high nor too low (Fig. 5.4, top). We can understand this result mathematically by examining the gain, again, but this time allowing the firing rate to vary as we vary

background connectivity (Eqn. 5.27):

$$\partial_{w_o} g_{ii} = -\frac{g_{ii}}{\sigma_i} \left( \frac{m_i}{\sigma_i} \partial_{w_o} m_i - \left( \frac{m_i}{\sigma_i} \right)^2 \partial_{w_o} \sigma_i + \partial_{w_o} \sigma_i \right) = 0. \quad (5.30)$$

As we have noted, this equation is very general so it is difficult to solve directly. However, when  $\partial_{w_o} \sigma_i \neq 0$  we can write

$$\partial_{w_o} g_{ii} = \frac{g_{ii}}{\sigma_i} \left( \left( \frac{m_i}{\sigma_i} \right)^2 - \left( \frac{m_i}{\sigma_i} \right) \left( \frac{\partial_{w_o} m_i}{\partial_{w_o} \sigma_i} \right) - 1 \right) \partial_{w_o} \sigma_i = 0. \quad (5.31)$$

This is now a quadratic equation for  $m_i/\sigma_i$ , which can be easily solved:

$$\frac{m_i}{\sigma_i} = \frac{1}{2} \left( \frac{\partial_{w_o} m_i}{\partial_{w_o} \sigma_i} \right) \pm \sqrt{\frac{1}{4} \left( \frac{\partial_{w_o} m_i}{\partial_{w_o} \sigma_i} \right)^2 + 1}. \quad (5.32)$$

This is a complicated solution. However, when  $\partial_{w_o} m_i = 0$  we see that

$$\frac{|m_i|}{\sigma_i} = 1. \quad (5.33)$$

This tells us that when  $\partial_{w_o} m_i = 0$  the gain is maximized when the mean synaptic drive and synaptic background noise are the same size. More generally, we can see from equations 5.32, 4.9 and 4.10 that the gain is maximized when

$$\frac{|m_i|}{\sigma_i} \sim \mathcal{O}(1). \quad (5.34)$$

This is an intriguing result. It explains how background connectivity can facilitate an increase in information. In our our literature review (Chapter 2) we saw that networks without random background connectivity have  $m_i/\sigma_i \sim \mathcal{O}(\sqrt{K})$ , which leads to regular spike trains, inconsistent with observed cortical dynamics. However, networks that do have background connectivity have relatively large synaptic background noise, or  $m_i/\sigma_i \sim \mathcal{O}(1)$ . Now, we have shown (Eqn. 5.34) that when  $|m_i|/\sigma_i \sim \mathcal{O}(1)$ , the gain is maximized, facilitating high performance at orientation selectivity. Therefore, background connectivity is important for cortical function as well as cortical dynamics.

### 5.4.5 Correlations and information

Finally, we ask whether correlations increase or decrease information in balanced networks. Naively, we might expect correlations to always decrease information, though in theory, we know that it is possible for correlations to increase or decrease information (Averbeck et al., 2006).

We have found that correlations in balanced networks can increase with information (Figs. 5.3 and 5.5). However, this relationship is not causal. The connectivity that produces correlated spiking is the same connectivity that produces orientation tuning and contrast invariance. Therefore, it is impossible to attribute information increases to correlation increases.

The complicated relationship between correlations and tuning curves exemplifies the difficulty in prescribing computational functions to particular spike train statistics. Rather than looking for direct relationships between correlations and information, we quantify the effect of ignoring correlations in orientation selectivity. Theorists and experimentalists would like to ignore the contribution of correlations because calculating correlations and measuring correlations is difficult (Ecker et al., 2010; Averbeck et al., 2006). To quantify the effect of ignoring correlations, we calculate

$$\Delta I_{diag}/I = (I - I_{diag})/I, \quad (5.35)$$

where  $I_{diag}$  is the inverse variance of an unbiased decoder that ignores correlations but is optimal otherwise (?; Wu et al., 2001; Averbeck et al., 2006).

We calculate  $I_{diag}$  by writing down an expression for the unbiased sub-optimal decoder that ignores correlations:

$$\mathbf{w}_{diag} = \boldsymbol{\rho}_d^{-1} \boldsymbol{\nu}' / (\boldsymbol{\nu}' \boldsymbol{\rho}_d^{-1} \boldsymbol{\nu}'). \quad (5.36)$$

We can then use this to calculate  $I_{diag}$ :

$$I_{diag} = \text{var}(\theta_{diag})^{-1} = (\mathbf{w}_{diag} \cdot \boldsymbol{\rho} \cdot \mathbf{w}_{diag})^{-1} = \frac{(\boldsymbol{\nu}' \cdot \boldsymbol{\rho}_d^{-1} \cdot \boldsymbol{\nu}')^2}{\boldsymbol{\nu}' \cdot \boldsymbol{\rho}_d^{-1} \boldsymbol{\rho} \boldsymbol{\rho}_d^{-1} \cdot \boldsymbol{\nu}'}. \quad (5.37)$$

Using this expression, we calculate  $\Delta I_{diag}/I$  for a series of networks with different

levels of structured connectivity and background connectivity (Fig. 5.6). We find that correlations cannot be ignored, despite the fact that correlations are weak. It is particularly bad to ignore correlations in networks with lots of structure ( $j_0 = 3$ , for example). However, correlations cancel in networks with background connectivity, so ignoring correlations is not disastrous in balanced networks, although up to 40% of information is lost in the example shown here.

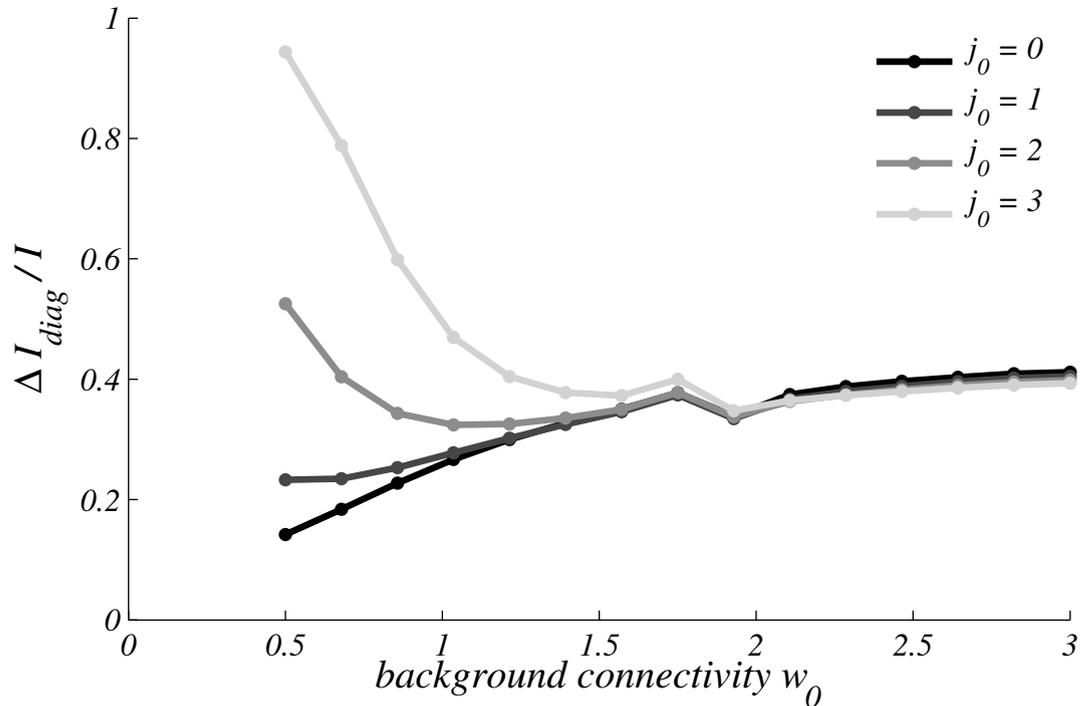


Figure 5.6: Correlations and information

The effect of ignoring correlations on information for a series of networks with different levels of background connectivity and structured connectivity. When  $\Delta I_{diag}/I = 0$  correlations can be ignored. When  $\Delta I_{diag}/I = 1$ , all information is lost if correlations are ignored.

## 5.5 Discussion

The computational role of spike train correlations and spike train irregularity are currently the subject of intense debate in theoretical neuroscience (Destexhe and Contreras, 2006; Averbeck et al., 2006). Given that the bulk of cortical activity is irregular and asynchronous, it is especially disturbing that we do not yet know its functional role. Spike train irregularity is a particularly mysterious phenomenon because the synaptic background noise that causes irregularity can harm computation.

Previous attempts to discover the role of irregularity and correlation in computation have been impaired by the absence of a plausible theory of cortical activity (Abbott and Dayan, 1999; Seriès et al., 2004; Beck et al., 2011). None of these models have background connectivity, so synaptic background noise is not realistic. However, in recent years, a credible explanation of cortical dynamics has been provided by balanced network theory (van Vreeswijk and Sompolinsky, 1996; Amit and Brunel, 1997a; van Vreeswijk and Sompolinsky, 1998; Brunel, 2000; Lerchner et al., 2004; Lerchner et al., 2006; Hertz, 2010; Lerchner and Latham, 2011). In this thesis, we extend this theory to explain orientation tuning, using weak structured connectivity (Chapter 3). We also calculate spike train correlations in a balanced network (Chapter 4). That preliminary work has allowed us to calculate linear Fisher information, which quantifies the ability of a balanced network to perform orientation selectivity. We obtain mathematical expressions which relate information to network connectivity and input, enabling us to address the role of irregularity, correlations and orientation tuning in computation.

We find that orientation tuning improves orientation selectivity. This is not entirely inevitable, because the structured connectivity that produces orientation tuning is much weaker than the background connectivity in our model.

This observation suggests that a balanced network model with weak structured connectivity embedded in strong background connectivity is a plausible model for cortical connectivity. Recently, similar connectivity was used in a cortical memory model (Roudi and Latham, 2007). In this memory model, background connectivity acts as a memory reservoir, effectively enveloping irrelevant memories. In orientation selectivity, the contribution of background connectivity is very different.

Surprisingly, we find that background connectivity improves computational performance across a wide range of stimulus contrasts. This computational stability comes at a cost. For a narrow range of stimulus contrasts, computational performance is reduced. The mechanism is simple - synaptic background noise increases the operating range of the network, providing computational stability, in the same way that it provides dynamical stability. This is an example of stochastic resonance (McDonnell and Abbott, 2009). Stochastic resonance has been demonstrated before in single neurons (Longtin et al., 1991; Bulsara et al., 1991; Stemmler, 1996; Destexhe and Contreras, 2006; McDonnell and Abbott, 2009; Greenwood et al., 2000) and has been proposed as an explanation

for the presence of synaptic background noise in V1 (Anderson et al., 2000). However, this is the first demonstration of stochastic resonance in a cortical network model with realistic spiking activity. The contrast invariance of orientation tuned cells is a signature of this stochastic resonance.

The discovery of a simple computational role for synaptic background noise suggests that the brain is optimal, in some sense (Barlow, 1961). The original observation that background connectivity is responsible for synaptic noise (van Vreeswijk and Sompolinsky, 1996; van Vreeswijk and Sompolinsky, 1998) had raised the prospect that the cortex was sub-optimal, because synaptic background noise is avoidable noise. Our work indicates that, rather than being an evolutionary mistake, networks with background connectivity are optimal.

It is difficult to unravel the contribution of correlations to computation. Unsurprisingly, correlations are larger in networks with structured connectivity, because connectivity produces correlations. Structure also produces orientation tuning, which we know increases information. Therefore, we cannot conclude that correlations cause information to increase. The relationship between correlations and information cannot be disentangled in a network model with realistic spike train irregularity such as our balanced network model.

Nonetheless, we can calculate the effect of ignoring correlations. This is an important practical issue for experimentalists and theorists, because ignoring correlations would simplify both theoretical analysis and data analysis (Averbeck et al., 2006). We find that although correlations are weak, they are not weak enough to be ignored completely. Neither are they strong enough to cause catastrophic computational damage if ignored.

The equilibrium state of a balanced network and computation in this state have been the focus of our investigation, so far. This is an important computational state. However, in future work it would be interesting to calculate information contained in a network out of equilibrium. This should be possible using the non-equilibrium firing rate equations and covariance equations that we derived for balanced networks (Chapter 4).

It would also be interesting to implement a more complex computation than orientation selectivity, such as object recognition for example. Orientation selectivity is a particularly simple computation, especially in a network that receives orientation tuned

input. It has served to unveil some interesting computational properties. A more involved computation, such as object recognition, or some Bayesian inference, should provide additional insight into the computational roles of neural irregularity, asynchronicity and connectivity.

A number of experimental predictions arise from our work. Behavioural studies should reveal impaired orientation selectivity at most contrast levels in animals with reduced synaptic background noise, with improved performance for some narrow range of contrasts. We can also predict that small targeted perturbations to synaptic weights should be enough to destroy orientation tuning, if structured connectivity is weak.

The work that we have presented provides an important link between cortical function and activity. Our calculations have enabled us to answer some long standing computational neuroscience questions. However, the equations we derive also illustrate that the relationship between computational performance and connectivity is complicated, underlying the fallacy of assigning simple functions to specific network statistics such as correlations.



## Chapter 6

# Perspectives

The air surrounding us contains an enormous number of atoms, violently smashing against each other, following extremely complicated paths. Measuring the movement of all these atoms and predicting their trajectories is practically impossible. Nonetheless, easily measurable quantities such as temperature, pressure and volume provide an excellent description of a gaseous state and can be easily related to each other using *Boyle's Law* and *Charles' Law*. These *macroscopic* quantities emerge from the *microscopic* movements of individual atoms in a gas and can be related to each other using statistical mechanics. For example, using statistical mechanics, the temperature of a gas can be linearly related to the average kinetic energy of individual atoms.

In neuroscience, we are faced with a similar problem, with hundreds of millions of neurons, all interacting with each other according to complicated neural dynamics to produce irregular, asynchronous spike trains. Our goal is to derive a mathematical relationship between the macroscopic and microscopic properties of a balanced network. We use techniques from statistical mechanics and dynamical systems theory to relate macroscopic quantities such as information, correlations and firing rates to microscopic quantities such as network connectivity and spike transition rates.

Just as the relationship between the macroscopic and microscopic provided important insights in thermodynamics, the equations that we derive provide new insight into correlations and computational performance in the cortex. In particular, we find:

1. A unifying model of network connectivity with weak structured connectivity embedded in strong background connectivity can produce orientation tuned, contrast invariant, irregular, asynchronous neural responses, consistent with a large body

of cortical experiments.

2. Spike train correlations in a balanced network are weak, but not weak enough to be ignored in computation.
3. Synaptic background noise can improve computational performance by maintaining a network in a highly informative state for a broad range of inputs.

This last result is perhaps the most important, as it provides a simple functional explanation for the existence of synaptic background noise in the cortex. Without a functional explanation, such as this, we would have to conclude that the brain is severely sub-optimal in computation, contradicting one of the most successful theories of cortical function.

## 6.1 Open questions

Building upon the results and analysis of this thesis, there are many new dynamical and computational questions that we can now ask. We have already developed some of these new directions, in unpublished work. We finish by discussing some of these possibilities.

To begin, we ask if higher order spike train statistics play an important role in computation? We have already calculated first order statistics and second order statistics of balanced network spike trains and quantified their role in orientation selectivity. One of the most important steps in this analysis was treating synaptic background noise as a one, or two-dimensional Gaussian input. Perhaps, by treating synaptic background noise as high dimensional Gaussian input, we might calculate higher order spike train statistics and quantify their computational role, if any.

Another interesting direction for generalisation is increased computational complexity. The brain is capable of performing computations that are much more complex than orientation selectivity. For example, object recognition is an important, but difficult, survival dependent computation that animals perform. The insight and analysis that we have developed investigating orientation selectivity may provide clues for the implementation of such difficult computations.

Temporal differentiation is another example of a difficult computation that the brain must perform. The brain must compute derivatives when predicting the trajectory of

temporally varying quantities, such as the position of a moving object. In recent, unpublished work, we have found that balanced networks can naturally compute derivatives. There are many other temporal computations whose neural implementation has not yet been discovered.

Learning in balanced networks is another interesting direction for generalisation of our work. We have demonstrated that much of the spiking activity observed in the cortex can be explained using a connectivity model with weak structured connectivity embedded in strong background connectivity. It would be interesting to understand how this weak structure can be learned.

One possible learning mechanism is a Hebbian learning rule, in which synapses between correlated neurons increase in strength and synapses between uncorrelated neurons decrease in strength. A problem with this mechanism is that the contribution of background connectivity to correlations is about the same size as the contribution of structure, and this may interfere with learning. However, the noisy contribution of background connectivity correlations might actually help, by preventing the formation of unstable memory states.

The extent to which network connectivity is random or structured is unknown. A number of ambitious experimental programmes are currently addressing this question by directly measuring cortical connectivity. Our proposed model for network connectivity provides a new and promising target for these experiments.

How does the brain compute? This is the central question in computational neuroscience. We have made some progress towards an answer, describing the role of synaptic background noise, structured connectivity and correlations in orientation selectivity. In the end, however, we have raised as many new questions as we have answered. This is a reflection of the depth and breadth of this subject. Indeed, disentangling computation in balanced networks promises to be one of the great challenges for 21st century neuroscience.



# Bibliography

- Abbott, L. (2008). Theoretical neuroscience rising. *Neuron*, pages 489–495.
- Abbott, L. F. and Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural computation*, 11(1):91–101.
- Abeles, M. (1991). *Corticonics: Neural Circuits of the Cerebral Cortex*. Cambridge University Press.
- Adrian, E. and Zotterman, Y. (1926). The impulses produced by sensory nerve endings. *The Journal of physiology*, 61(1):49.
- Amari, S.-i. and Nagaoka, H. (2000). *Methods of Information Geometry*. Oxford University Press.
- Amit, D. and Brunel, N. (1997a). Dynamics of a recurrent network of spiking neurons before and following learning. *Network: Computation in Neural Systems*, 8(4):373–404.
- Amit, D., Gutfreund, H., and Sompolinsky, H. (1985). Storing infinite numbers of patterns in a spin-glass model of neural networks. *Physical Review Letters*, 55(14):1530–1533.
- Amit, D. J. and Brunel, N. (1997b). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cerebral cortex (New York, N.Y. : 1991)*, 7(3):237–52.
- Anderson, J. S., Lampl, I., Gillespie, D. C., and Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science (New York, N.Y.)*, 290(5498):1968–72.

- Averbeck, B. B., Latham, P. E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7(5):358–366.
- Averbeck, B. B. and Lee, D. (2004). Coding and transmission of information by neural ensembles. *Trends in neurosciences*, 27(4):225–30.
- Bair, W., Koch, C., Newsome, W., and Britten, K. (1994). Power spectrum analysis of bursting cells in area MT in the behaving monkey. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 14(5 Pt 1):2870–92.
- Barlow, H. (1961). Possible principles underlying the transformation of sensory messages. *Sensory Communication* ed WA Rosenblith.
- Barlow, H. (2001). Redundancy reduction revisited. *Network: Computation in Neural Systems*, 12(3):241–253.
- Barrett, D. G. T. (2007). *Sparse Coding: from natural sounds to music*. M. sc., Trinity College Dublin.
- Beck, J., Bejjanki, V. R., and Pouget, A. (2011). Insights from a simple expression for linear fisher information in a recurrently connected population of spiking neurons. *Neural computation*, 23(6):1484–502.
- Bell, A. and Sejnowski, T. (1997). The independent components of natural scenes are edge filters. *Vision research*, 37(23):3327.
- Ben-Yishai, R., Bar-Or, R., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences*, 92(9):3844.
- Benayoun, M., Cowan, J. D., van Drongelen, W., and Wallace, E. (2010). Avalanches in a stochastic model of spiking neurons. *PLoS computational biology*, 6(7):e1000846.
- Binzegger, T., Douglas, R. J., and Martin, K. a. C. (2004). A quantitative map of the circuit of cat primary visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(39):8441–53.
- Blasdel, G. and Fitzpatrick, D. (1984). Physiological organization of layer 4 in macaque striate cortex. *The Journal of neuroscience*, 4(3):880–895.

- Boerlin, M. (2011). *Probabilistic Inference and Working Memory in Networks of Spiking Neurons*. PhD thesis.
- Boerlin, M. and Denève, S. (2011). Spike-Based Population Coding and Working Memory. *PLoS Computational Biology*, 7(2):e1001080.
- Braitenber, V. and Schuz, A. (1991). *Anatomy of the cortex. Statistics and Geometry*. Springer.
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of computational neuroscience*, 8(3):183–208.
- Bulsara, a., Jacobs, E. W., Zhou, T., Moss, F., and Kiss, L. (1991). Stochastic resonance in a single neuron model: theory and analog simulation. *Journal of theoretical biology*, 152(4):531–55.
- Buonomano, D. and Merzenich, M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science*, 267(5200):1028.
- Burns, B. D. and Webb, a. C. (1976). The spontaneous activity of neurones in the cat’s cerebral cortex. *Proceedings of the Royal Society of London. Series B, Containing papers of a Biological character. Royal Society (Great Britain)*, 194(1115):211–23.
- Carandini, M. and Heeger, D. (1994). Summation and division by neurons in primate visual cortex. *Science*, 264(5163):1333.
- Chapman, B. and Stryker, M. (1993). Development of orientation selectivity in ferret visual cortex and effects of deprivation. *The Journal of neuroscience*, 3(December).
- Chung, S. and Ferster, D. (1998). Strength and orientation tuning of the thalamic input to simple cells revealed by electrically evoked cortical suppression. *Neuron*, 20(6):1177–89.
- Cohen, M. R. and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nature neuroscience*, 14(7):811–9.
- Dale, H. (1935). Pharmacology and nerve endings. *The Journal of Nervous and Mental Disease*, 82(4):457.

- Dayan, P. and Abbott, L. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. The MIT Press.
- DeFelipe, J. (2010). From the connectome to the synaptome: an epic love story. *Science (New York, N.Y.)*, 330(6008):1198–201.
- Deneve, S. (2008a). Bayesian spiking neurons I: inference. *Neural computation*, 20(1):91–117.
- Deneve, S. (2008b). Bayesian spiking neurons II: learning. *Neural computation*, 20(1):118–45.
- Deneve, S., Latham, P., and Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. *nature neuroscience*, 4(8):826–831.
- Deneve, S., Latham, P. E., and Pouget, A. (1999). Reading population codes: a neural implementation of ideal observers. *Nature neuroscience*, 2(8):740–5.
- Destexhe, A. and Contreras, D. (2006). Neuronal computations with stochastic network states. *Science (New York, N.Y.)*, 314(5796):85–90.
- Destexhe, A., Rudolph, M., and Paré, D. (2003). The high-conductance state of neocortical neurons in vivo. *Nature reviews. Neuroscience*, 4(9):739–51.
- DeWeese, M. R. and Zador, A. M. (2006). Non-Gaussian membrane potential dynamics imply sparse, synchronous activity in auditory cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(47):12206–18.
- Douglas, R. J. and Martin, K. a. (1991). Opening the grey box. *Trends in neurosciences*, 14(7):286–93.
- Ecker, A. S., Berens, P., Keliris, G. a., Bethge, M., Logothetis, N. K., and Tolias, A. S. (2010). Decorrelated neuronal firing in cortical microcircuits. *Science (New York, N.Y.)*, 327(5965):584–7.
- Faisal, a. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nature Reviews Neuroscience*, 9(4):292–303.
- Faisal, a. A., White, J. a., and Laughlin, S. B. (2005). Ion-channel noise places limits on the miniaturization of the brain’s wiring. *Current biology : CB*, 15(12):1143–9.

- Fatt, P. and Katz, B. (1950). Some observations on biological noise. *Nature*, 166(4223):597.
- Fellous, J., Rudolph, M., Destexhe, A., and Sejnowski, T. (2003). Synaptic background noise controls the input/output characteristics of single cells in an in vitro model of in vivo activity. *Neuroscience*, 122(3):811–829.
- Ferster, D. (1986). Orientation selectivity of synaptic potentials in neurons of cat primary visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 6(5):1284–301.
- Ferster, D., Chung, S., Wheat, H., and Others (1996). Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature*, 380(6571):249–252.
- Ferster, D. and Lindstrom, S. (1983). An intracellular analysis of geniculate-cortical connectivity in area 17 of the cat. *The Journal of physiology*, (1983):181–215.
- Ferster, D. and Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annual review of neuroscience*, 23:441–71.
- Fino, E. and Yuste, R. (2011). Dense inhibitory connectivity in neocortex. *Neuron*, 69(6):1188–203.
- Fregnac, Y. and Imbert, M. (1984). Development of neuronal selectivity in primary visual cortex of cat. *Physiological reviews*, 64(1).
- Gershon, E., Wiener, M., Latham, P. E., and Richmond, B. (1998). Coding Strategies in Monkey V1 and Inferior Temporal Cortices. *Journal of Neurophysiology*, 79(3):1135.
- Gilbert, C. (1977). Laminar differences in receptive field properties of cells in cat primary visual cortex. *The Journal of physiology*, 268(2):391.
- Gillespie, D. (1977). Exact stochastic simulation of coupled chemical reactions. *The journal of physical chemistry*, 93555(1):2340–2361.
- Ginzburg, I. and Sompolinsky, H. (1994). Theory of correlations in stochastic neural networks. *Physical Review E*, 50(4):3171–3191.
- Glauber, R. J. (1963). Time-Dependent Statistics of the Ising Model. *Journal of Mathematical Physics*, 4(2):294.

- Greene, G., Barrett, D. G. T., Sen, K., and Houghton, C. (2009). Sparse coding of birdsong and receptive field structure in songbirds. *Network: Computation in neural systems*, 20(3):162–77.
- Greenwood, P. E., Ward, L. M., Russell, D. F., Neiman, a., and Moss, F. (2000). Stochastic resonance enhances the electrosensory information available to paddlefish for prey capture. *Physical review letters*, 84(20):4773–6.
- 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. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(17):4535–45.
- Hansel, D. and Sompolinsky, H. (1996). Chaos and synchrony in a model of a hypercolumn in visual cortex. *Journal of computational neuroscience*, 3(1):7–34.
- Hebb, D. (1949). *The organization of behavior*. New York: Wiley and Sons.
- Hertz, J. (2010). Cross-correlations in high-conductance states of a model cortical network. *Neural computation*, 22(2):427–47.
- Hertz, J., Krogh, A., and Palmer, R. (1991). *Introduction to the theory of neural computation*. Westview Press.
- Hodgkin, A. L. and Huxley, A. F. (1952). Currents carried by sodium and potassium ions through the membrane of the giant axon of Loligo. *Journal of Physiology*, 116:449–472.
- Hofer, S. B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., Lein, E., Lesica, N. a., and Mrsic-Flogel, T. D. (2011). Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature Neuroscience*, 14(8).
- Holt, G. R. and Douglas, J. (1996). Comparison of Discharge Variability Visual Cortex Neurons In Vitro and In Vivo in Cat Visual Cortex Neurons. *Neurophysiology*, 75(5):1806–1814.
- Hopfield, J. J. (1982). Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558.

- Hopfield, J. J. and Brody, C. D. (2001). What is a moment? Transient synchrony as a collective mechanism for spatiotemporal integration. *Proceedings of the National Academy of Sciences of the United States of America*, 98(3):1282–7.
- Hromádka, T., DeWeese, M., and Zador, A. (2008). Sparse representation of sounds in the unanesthetized auditory cortex. *PLoS biology*, 6(1):e16.
- Hubel, D. and Wiesel, T. (1959). Receptive fields of single neurones in the cat’s striate cortex. *The Journal of physiology*, 148(3):574–591.
- Hubel, D. and Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex. *The Journal of physiology*, 160(1):106.
- Jaeger, H. and Haas, H. (2004). Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science (New York, N. Y.)*, 304(5667):78–80.
- Joris, P. X., Smith, P. H., and Yin, T. C. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. II. Responses in the tuning curve tail. *Journal of neurophysiology*, 71(3):1037–51.
- Joris, P. X., Smith, P. H., and Yin, T. C. T. (1998). Coincidence Detection in the Auditory System: 50 Years after Jeffress. *Neuron*, 21:1235–1238.
- Kelly, J. P. and Van Essen, D. C. (1974). Cell structure and function in the visual cortex of the cat. *The Journal of physiology*, 238(3):515–47.
- Knight, B. W. (1972). Dynamics of encoding in a population of neurons. *The Journal of general physiology*, 59(6):734–66.
- Knill, D. and Pouget, A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *TRENDS in Neurosciences*, 27(12).
- Ko, H., Hofer, S. B., Pichler, B., Buchanan, K. a., Sjöström, P. J., and Mrsic-Flogel, T. D. (2011). Functional specificity of local synaptic connections in neocortical networks. *Nature*, 473(7345):87–91.
- Kumar, A., Schrader, S., Aertsen, A., and Rotter, S. (2008). The high-conductance state of cortical networks. *Neural computation*, 20(1):1–43.

- Latham, P., Richmond, B., Nelson, P., and Nirenberg, S. (2000). Intrinsic dynamics in neuronal networks. I. Theory. *Journal of Neurophysiology*, 83(2):808.
- Laurent, G., Wehr, M., and Davidowitz, H. (1996). Temporal representations of odors in an olfactory network. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 16(12):3837–47.
- Lecar, H. and Nossal, R. (1971). Theory of threshold fluctuations in nerves. II. Analysis of various sources of membrane noise. *Biophysical journal*, 11(12):1068–84.
- Lerchner, a., Ahmadi, M., and Hertz, J. (2004). High-conductance states in a mean-field cortical network model. *Neurocomputing*, 58-60:935–940.
- Lerchner, A. and Latham, P. E. (2011). A unifying framework for state-dependent dynamics in cortex.
- Lerchner, A., Sterner, G., Hertz, J., and Ahmadi, M. (2006). Mean field theory for a balanced hypercolumn model of orientation selectivity in primary visual cortex. *Network (Bristol, England)*, 17(2):131–50.
- London, M., Roth, A., Beeren, L., Häusser, M., and Latham, P. E. (2010). Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature*, 466(7302):123–7.
- Longtin, A., Bulsara, A., and Moss, F. (1991). Time-interval sequences in bistable systems and the noise-induced transmission of information by sensory neurons. *Physical Review Letters*, 67(5):656–659.
- Ma, W. J., Beck, J. M., Latham, P. E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature neuroscience*, 9(11):1432–8.
- Maass, W. and Markram, H. (2004). On the computational power of circuits of spiking neurons. *Journal of computer and system sciences*, 69(4):593–616.
- Maass, W., Natschläger, T., and Markram, H. (2002). Real-time computing without stable states: a new framework for neural computation based on perturbations. *Neural computation*, 14(11):2531–60.

- MacLeod, K., Bäcker, A., and Laurent, G. (1998). Who reads temporal information contained across synchronized and oscillatory spike trains? *Nature*, 395(6703):693–698.
- Mainen, Z. F. and Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science (New York, N.Y.)*, 268(5216):1503–6.
- Maldonado, P. E. (1997). Orientation Selectivity in Pinwheel Centers in Cat Striate Cortex. *Science*, 276(5318):1551–1555.
- Markram, H., Lübke, J., Frotscher, M., Roth, a., and Sakmann, B. (1997). Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. *The Journal of physiology*, 500 ( Pt 2:409–40.
- Martin, K. and Whitteridge, D. (1984). Form, function and intracortical projections of spiny neurones in the striate visual cortex of the cat. *The Journal of physiology*, 26(3):72–72.
- McCulloch, W. and Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5:115–133.
- McDonnell, M. D. and Abbott, D. (2009). What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. *PLoS computational biology*, 5(5):e1000348.
- Murphy, B. K. and Miller, K. D. (2009). Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron*, 61(4):635–48.
- Nelson, S., Toth, L., Sheth, B., and Sur, M. (1994). Orientation selectivity of cortical neurons during intracellular blockade of inhibition. *Science*, 265(5173):774.
- Nirenberg, S. and Latham, P. E. (2003). Decoding neuronal spike trains: how important are correlations? *Proceedings of the National Academy of Sciences of the United States of America*, 100(12):7348–53.
- O’Connor, D. H., Peron, S. P., Huber, D., and Svoboda, K. (2010). Neural activity in barrel cortex underlying vibrissa-based object localization in mice. *Neuron*, 67(6):1048–61.

- O'Keefe, J. and Dostrovsky, J. (1971). The hippocampus as a spatial map: Preliminary evidence from unit activity in the freely-moving rat. *Brain research*, 34:171–175.
- Okun, M. and Lampl, I. (2008). Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. *Nature neuroscience*, 11(5):535–7.
- Olshausen, B. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609.
- Olshausen, B. a. and Field, D. J. (2005). How close are we to understanding v1? *Neural computation*, 17(8):1665–99.
- Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. *Network: Computation in Neural Systems*, 15(4):243–262.
- Paradiso, M. (1988). A theory for the use of visual orientation information which exploits the columnar structure of striate cortex. *Biological Cybernetics*, 58(1):35–49.
- Poggio, T. (1990). A theory of how the brain might work.
- Poggio, T. and Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, 343,:263 – 266.
- Poirazi, P., Brannon, T., and Mel, B. W. (2003). Pyramidal neuron as two-layer neural network. *Neuron*, 37(6):989–99.
- Pouget, a., Dayan, P., and Zemel, R. (2000). Information processing with population codes. *Nature reviews. Neuroscience*, 1(2):125–32.
- Pouget, A. and Sejnowski, T. J. (1997). Spatial Transformations in the Parietal Cortex Using Basis Functions. *Journal of Cognitive Neuroscience*, 9(2):222–237.
- Rajan, K. and Abbott, L. (2006). Eigenvalue spectra of random matrices for neural networks. *Physical review letters*, 97(18):188104.
- Reid, R. and Alonso, J. (1995). Specificity of monosynaptic connections from thalamus to visual cortex. *Nature*, 378(6554):281–283.
- Renart, A., Rocha, J. D. L., Bartho, P., Hollender, L., and N (2010). The asynchronous state in cortical circuits. *science*, 587(January).

- Ringach, D. L., Shapley, R. M., and Hawken, M. J. (2002). Orientation selectivity in macaque V1: diversity and laminar dependence. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(13):5639–51.
- Roudi, Y. and Latham, P. E. (2007). A balanced memory network. *PLoS Computational Biology*, 3(9):1679–1700.
- Roxin, A., Brunel, N., Hansel, D., Mongillo, G., and van Vreeswijk, C. (2011). On the Distribution of Firing Rates in Networks of Cortical Neurons. *The Journal of Neuroscience*, 31(45):16217–16226.
- Rudolph, M., Pelletier, J. G., Paré, D., and Destexhe, A. (2005). Characterization of synaptic conductances and integrative properties during electrically induced EEG-activated states in neocortical neurons in vivo. *Journal of neurophysiology*, 94(4):2805–21.
- Salinas, E. and Abbott, L. F. (1995). Transfer of coded information from sensory to motor networks. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 15(10):6461–74.
- Schwartz, E. (1990). *Computational neuroscience*. MIT Press.
- Sclar, G. and Freeman, R. D. (1982). Orientation selectivity in the cat’s striate cortex is invariant with stimulus contrast. *Experimental brain research. Experimentelle Hirnforschung. Expérimentation cérébrale*, 46(3):457–61.
- Sejnowski, T. J., Koch, C., and Churchland, P. S. (1988). Computational Neuroscience. *Science*, 241:1299 – 1306.
- Seriès, P., Latham, P., and Pouget, A. (2004). Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nature Neuroscience*, 7:1129–1135.
- Seung, H. and Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences*, 90(1):1–5.
- Shadlen, M. N. and Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current opinion in neurobiology*, 4(4):569–79.

- Shadlen, M. N. and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 18(10):3870–96.
- Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical journal*, 27:379–423.
- Shatz, C. and Stryker, M. (1978). Ocular dominance in layer IV of the cat’s visual cortex and the effects of monocular deprivation. *The Journal of Physiology*, 281(1):267.
- Shu, Y. and Hasenstaub, A. (2003). Turning on and off recurrent balanced cortical activity. *Nature*, 423(May).
- Skottun, B. C., Bradley, a., Sclar, G., Ohzawa, I., and Freeman, R. D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: a comparison of single cells and behavior. *Journal of neurophysiology*, 57(3):773–86.
- Softky, W. (1995). Simple codes versus efficient codes. *Current Opinion in Neurobiology*.
- Softky, W. R. and Koch, C. (1992). Cortical Cells Should Fire Regularly, But Do Not. *Neural Computation*, 4(5):643–646.
- Softky, W. R. and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 13(1):334–50.
- Somers, D., Nelson, S., and Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *The Journal of neuroscience*, 15(8):5448–5465.
- Sompolinsky, H. and Shapley, R. (1997). New perspectives on the mechanisms for orientation selectivity. *Current opinion in neurobiology*, 7(4):514–522.
- Sompolinsky, H. and White, O. (2005). THEORY OF LARGE RECURRENT NETWORKS: FROM SPIKES TO BEHAVIOR. In *Methods and models*.
- Sompolinsky, H., Yoon, H., Kang, K., and Shamir, M. (2001). Population coding in neuronal systems with correlated noise. *Physical Review E*, 64(5):51904.
- Stemmler, M. (1996). A single spike suffices: the simplest form of stochastic resonance in model neurons. *Network: Computation in Neural Systems*, 7(4):687–716.

- Sussillo, D. and Abbott, L. F. (2009). Generating coherent patterns of activity from chaotic neural networks. *Neuron*, 63(4):544–57.
- Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press.
- Tanaka, K. (1983). Cross-correlation analysis of geniculostriate neuronal relationships in cats. *Journal of neurophysiology*, 49(6):1303–18.
- Todorov, E. and Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature neuroscience*, 5(11):1226–35.
- van Vreeswijk, C. and Sompolinsky, H. (1996). Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science (New York, N.Y.)*, 274(5293):1724–1726.
- van Vreeswijk, C. and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural computation*, 10(6):1321–1371.
- van Vreeswijk, C. and Sompolinsky, H. (2005). Irregular activity in large networks of neurons. *Methods and models in neurophysics*.
- Vogels, T. P. and Abbott, L. F. (2009). Gating multiple signals through detailed balance of excitation and inhibition in spiking networks. *Nature neuroscience*, 12(4):483–91.
- Wehr, M. and Laurent, G. (1996). Odour encoding by temporal sequences of firing in oscillating neural assemblies. *Nature*, 384(6605):162–166.
- White, J. and Rubinstein, J. (2000). Channel noise in neurons. *J. Neurosci*, 23(3):131–7.
- Wilke, S. D. and Eurich, C. W. (2002). Representational accuracy of stochastic neural populations. *Neural computation*, 14(1):155–89.
- Wu, S., Nakahara, H., and Amari, S. (2001). Population coding with correlation and an unfaithful model. *Neural computation*, 13(4):775–97.
- Zemel, R. S., Dayan, P., and Pouget, a. (1998). Probabilistic interpretation of population codes. *Neural computation*, 10(2):403–30.
- Zohary, E., Shadlen, M., and Newsome, W. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance.