The impact of synaptic depression on network activity and implications for neural coding

Lawrence Christopher York

Doctor of Philosophy
Institute for Adaptive and Neural Computation
School of Informatics
University of Edinburgh
2011

Abstract

Short - term synaptic depression is the phenomena where repeated stimulation leads to a decreased transmission efficacy. In this thesis, the impact of synaptic depression on the responses and dynamics of network models of visual processing is investigated, and the coding implications are examined. I find that synaptic depression can fundamentally change the operation of previously well - understood networks, and explain temporal nonlinearities present in neural responses to multiple stimuli. Furthermore, I show, more generally, how nonlinear interactions can be beneficial with respect to neural coding.

I begin chapter 1 with a short introduction. In chapter 2 of this thesis, the behaviour of a ring attractor network is examined when its recurrent connections are subject to short term synaptic depression. I find that, in the presence of a uniform background current, the activity of the network settles to one of three states: a stationary attractor state, a uniform state or a rotating attractor state. I show that the rotation speed can be adjusted over a large range by changing the background current, opening the possibility to use the network as a variable frequency oscillator or pattern generator, and use mathematical analysis to determine an approximate maximum rotation speed. Using simulations, I then extend the network into two - dimensions, and find a rich range of possible behaviours.

Processing in the visual cortex can be non - linear: the response to two objects or other visual stimuli presented simultaneously is often less than the sum of the responses to the individual objects. A maximum function has in some cases been proposed to describe these competitive interactions. More recent data has emphasised that such interactions have temporal aspects as well, namely that the response to an initially presented stimulus can suppress the response to a stimulus presented subsequently, especially if the first stimulus is presented at high contrast. Chapter 3 of this thesis will present a simple neuronal network featuring synaptic depression which can account for much of the temporal aspects of this behaviour, whilst remaining consistent with older data and models. Furthermore, it will show how this model leads to several strong predictions regarding the processing of low contrast stimuli sequences, as well suggesting a link between response latency and suppression strength. The response of the model to a structured sequence of input stimuli also appears to anticipate future stimuli, and we predict that the magnitude of this stimulus anticipation will decrease as contrast is decreased.

Following on from investigating the temporal aspects of responses to stimuli pairs, in

chapter 4 this thesis examines an abstract model of how coding is impacted by non - linear interactions, for both structured and unstructured stimuli spaces. I find that non-linear methods of responding to pairs of stimuli presented simultaneously can have a beneficial effect on coding capacity, with linearly combined responses generally leading to the highest decoding errors rates. This thesis goes on to examine the interplay between this models noise assumptions and the decoding performance, and finds that many of the assumptions made can be weakened without changing, qualitatively, these findings.

In chapter 5, this thesis examines layered networks of noisy spiking neurons with recurrent connectivity and featuring depressing synapses. The contrast dependent latency and spike count statistics of the model are analysed and are found to be strongly dependent on the parameters of the noise. The tuning of parameters for models containing noisy IF neurons is discussed, and an information theoretic approach to tuning is outlined which successfully reproduces earlier work in which noise was tuned to linearise the response of a spiking network. The approach is applied to maximise the ability of the network to filter rapid noise transients at low contrast. I finish the thesis with a short concluding chapter.

Acknowledgements

Many thanks to my supervisors, Mark van Rossum (Edinburgh) and Mike Oram (St Andrews). Thanks to Jesus Cortes and Dmitri Bibitchkov for discussion of the paper forming the basis of chapter 2 (published in the Journal of Computational Neuroscience, http://www.springerlink.com/content/b4513l30232g6082/) - I also thank the reviewers for their comments. Thanks to the EPSRC for funding me through the University of Edinburgh's Neuroinformatics Doctoral Training Centre. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF). (http://www.ecdf.ed.ac.uk/). The ECDF is partially supported by the eDIKT initiative (http://www.edikt.org.uk).

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Lawrence Christopher York*)

Dedicated to my parents and sister

Contents

Chapter 1

Introduction

In this introductory section, we outline the contents of this thesis, describing what we investigate in each chapter and some of the existing research which is most relevant. Firstly, we note that much of the work presented in this thesis involves short - term synaptic depression, and its impact on network activity. Synaptic depression is the name given to the phenomenon by which the strength of a synapse is temporarily weakened by repetitive stimulation, a property of synapses thought to be found throughout the cortex (Markram and Tsodyks, 1996; Wang et al., 2006; Thomson and Lamy, 2007). The fatiguing of synapses during short term depression is due to multiple biophysical mechanisms, both pre - and post - synaptic. As a result, detailed models of it can be highly complex (e.g. see Hennig et al., 2008), however, simplified models are sufficient for the applications we consider here. Synaptic depression has previously been incorporated in models seeking to explain some of the response properties of visual neurons, such as the suppression of responses to visual stimuli by mask stimuli or the strong dependence on contrast of response latency in deep layers of the visual system (Carandini et al., 2002; van Rossum et al., 2008). In the work we present here, we use a frequently employed approach to model synaptic depression in which the efficacy of the synapse is reduced (in a multiplicative fashion) in response to a spike, as well as a related firing rate model (Abbott et al., 1997; Tsodyks and Markram, 1997; Tsodyks et al., 1998).

In chapter 2 of this thesis, we investigate the effect of synaptic depression on the ring attractor network. The ring attractor network is a standard model in computational neuroscience. In such a network, neurons are arranged on a ring with their position corresponding to an orientation angle. These neurons are connected recurrently to each other in a rotationally invariant, centre - surround fashion. The ring attractor model

was originally introduced to model the contrast invariant tuning of orientation selective neurons in V1 (Ben-Yishai et al., 1995), but networks with similar organisational principles (such as a periodic structure, and centre surround connectivity) have been used to model many phenomena such as prefrontal working memory and the response properties of head direction cells (Compte et al., 2000; Zhang, 1996; McNaughton et al., 2006; van der Meer et al., 2007) - this range of modelling applications make the ring attractor network a good choice for examining the impact of synaptic depression on network activity.

The mathematical formulation of ring attractor networks has allowed an in depth analysis of its behaviour. In particular, it has been found that if the strength of this local, recurrent connectivity is strong enough, a weakly tuned input can lead to a sharply tuned, stationary, profile of network activity. In the absence of input, and provided the recurrent connections are strong enough, such a 'bump state' can persist as a stable attractor state in which a group of neighbouring neurons maintain their activity due to reinforcement by local excitatory connections (Ben-Yishai et al., 1995; Hansel and Sompolinsky, 1998). Whilst in the original ring attractor model outlined in Ben-Yishai et al. (1995) this bump state is stationary, the incorporation of either strong local inhibition or spike frequency adaptation into the model can lead to such an attractor state becoming unstable - in such situations there can instead exist rotating attractor states, with network activity settling to a travelling bump of activity which propagates with constant speed (Ben-Yishai et al., 1997; Hansel and Sompolinsky, 1998).

We incorporate short term synaptic depression into the recurrent connections of the ring attractor model, investigating how this changes affects network activity in the presence of a constant, homogeneous, background input. By keeping the mathematical formulation as simple as possible, we analyse how synaptic depression interacts with the stability of different attractor states as background input is altered. We show that if the magnitude of the synaptic depression is strong enough, the stationary bump states observed in the original model are not stable. Instead, we find rotating attractor states in a similar fashion as is observed in Hansel and Sompolinsky (1998), with network activity settling to a travelling bump of activity, rotating around the network. However, rather than rotation speeds remaining fixed over a range of input strengths, we instead find that the rotation speed of a propagating bump of activity varies over a large range, as the amplitude of the background input is changed. Possible applications for such a network include a neural oscillator or pattern generator with a variable frequency. Finally, we examine a network in which neurons are arranged in two dimensional struc-

tures, outlining some of the different modes of behaviour such networks can operate in.

In chapter 3, we continue to investigate networks featuring short term synaptic depression, motivated by experimental results relating to the temporal properties of responses to multiple visual stimuli. As a first order approximation, the responses of some neurons in the early stages of the visual system, such as simple cells in V1, can be modelled by the rectification of a linear receptive field applied to set of inputs - however nonlinearities in the responses to multiple stimuli exist which such an approach is unable to capture. Theoretical studies into object recognition have suggested that neurons might exist which implement the maximum operator - in which the neural response to a group of inputs is equal to the maximum of those inputs (Riesenhuber and Poggio, 1999, 2002). Neurophysiological evidence relating to this maximum - like computation has been obtained by recording the response of a neuron to two stimuli presented simultaneously, and finding that this is approximately equal to the largest of the responses of the same neuron to those stimuli presented singly (Sato, 1989; Lampl et al., 2004; Gawne and Martin, 2002 - although other data suggest the average, rather than the maximum, as the appropriate operator - Miller et al., 1993; Zoccolan et al., 2005 and again Sato, 1989).

Most of the experimental studies into the responses to multiple stimuli have focused on multiple stimuli presented simultaneously, however, a recent study examined the temporal properties of how V4 neurons respond to pairs of stimuli presented asynchronously (Gawne, 2008). This data indicates that the time course of the response to a pair of stimuli most closely resembled that of the response with the earliest onset latency. The effect is strongest if the 'first' stimulus is presented a high contrast - in such a situation, the second stimulus can appear to provide no change in the response. Similar data has been obtained for neurons in the superior temporal sulcus (STS, a part of the brain known to contain face and object sensitive cells) in which the response of a neuron to a stimulus is reduced by a preceding presentation of another, different, stimulus (Perrett et al., 2009).

These investigations into the temporal dynamics of the responses to multiple stimuli motivate the model we present in chapter 3. As in chapter 2, short term synaptic depression plays a key role. The network receives two inputs, and during the course of responding to these inputs activate a number of shared synaptic connections. We propose it is the depression of these shared connections following an early, high contrast stimulus which leads to such a strongly reduced response when the second stimulus is

presented. In line with the experimental data, we show how this effect is diminished when the first stimulus is presented at low contrast. We will show that the time - course of the recovery of this suppression is a reasonable match to that observed experimentally, and also how such a network could contribute to stimulus anticipation. We also use our model to outline some experimental predictions.

Chapter 4 marks a temporary departure from short term depression to consider, using an abstract model, how the responses of neural populations to multiple stimuli can be decoded. In particular, we are interested in how different methods of responding to multiple stimuli affects decoding accuracy. As outlined above, the neural responses to multiple stimuli has been investigated experimentally in the context of the visual system by comparing the responses to stimuli presented singly with those presented simultaneously. A common question, when presenting two stimuli, A and B, is to ask what relates the neural response to stimulus A presented solo, R_A and the neural response to stimulus B presented solo, R_B, to the neural response when both stimuli are presented together, R_{AB}. A simple way of investigation this is to ask if there exists a function f such that R_{AB} is closely approximated by $f(R_A, R_B)$, and experimental evidence has suggested both the maximum and average as possible values of f Sato, 1989; Miller et al., 1993; Gawne and Martin, 2002; Lampl et al., 2004; Zoccolan et al., 2005. Whilst the use of $f(R_A, R_B) = \max(R_A, R_B)$ has been suggested as helping contribute to the formation of position - and scale - invariant neural responses (Riesenhuber and Poggio, 1999), in chapter 4 we will examine the theoretical implications of the choice of f from another angle - namely, how the method of responding to pairs of stimuli can impact on the coding capacity of a group of neurons. Using a simple model of neural noise and numerical simulations, we will show that linear functions are particularly poorly suited to use when combining the responses to stimuli, leading to higher error rates when decoding the responses to pairs of stimuli. We find this to be the case both when the responses to those stimuli occupy a parametrised space (such as orientation selectivity) and for a more general group of stimuli responses which show no structure. Using geometric arguments, we provide an intuition as to why the responses of networks employing linear combination functions are more vulnerable to classification errors than those of networks employing nonlinear combination functions, such as the maximum.

Completing this thesis, chapter 5 investigates spiking statistics and noise filtering in networks of noisy integrate - and - fire (IF) neurons featuring recurrent synaptic connections subject to short term depression. As discussed earlier, the onset latency of

a neural response in the visual system is often strongly affected by the contrast a stimulus is presented at, with high contrast stimuli eliciting responses in neurons earlier than low contrast stimuli. This *contrast dependent latency* is largest at deeper layers of the visual system, with the difference in latency potentially reaching several hundred milliseconds in STS (Oram et al., 2002) whilst in the early visual system, contrast dependent latency changes number tens of milliseconds (Gawne et al., 1996). A recent model sought to examine contrast dependent latencies using a multilayer network featuring recurrent connectivity and short - term synaptic depression (van Rossum et al., 2008). In this model, low amplitude inputs ensure that the recurrent connections depress slowly, causing the network to integrate its input over an extended time period. Meanwhile, a high amplitude input depressed out the synaptic connections rapidly, leading to short integration times and a rapid propagation of the neural response to the next layer. The model was proposed to act as a variable integrator - acting to integrate low contrast inputs over an extended time to improve the signal to noise ratio, whilst swiftly propagating high contrast inputs (the responses to which are high amplitude and thus less susceptible to noise) through each layer of the network (van Rossum et al., 2008).

We implemented this model using spiking neurons for two main reasons. Firstly, we wanted to examine the spiking statistics of such a model in light of recent experimental data which indicates that the onset of neural response to stimuli in the visual cortex is accompanied by a drop in local spike count variability (in the form of a reduced Fano factor) and the correlation of spike counts between pairs of neurons, as well as a response onset dependent decorrelation in the spike count leading up to the response onset with the spike count following stimulus response (Churchland et al., 2010, and Mike Oram, unpublished data - for the latter see Fig. 5.3 in chapter 5). Clearly, the earlier rate - based model is inadequate for considering spiking statistics. Our hypothesis was that the recurrent connections would contribute to the spike count variability and the correlation in spike counts between neurons - and that, following the depression of the recurrent connections which occurs during the response to a stimulus, the weaker connections will naturally lead to a reduction in both quantities. We will show that elements of this hypothesis are true, but that the increased firing rate during a neural response also has a strong effect on spike count variability.

The second reason for implementing the model from van Rossum et al. (2008) using spiking neurons was to further examine the ability of depressing, recurrent connections at filtering out high frequency noise, observed in the rate - based model. Net-

works of noisy IF neurons were used, with their noise current parameters varied over a wide range. We did this through use of a metric based on mutual information. We expected that the recurrent connectivity would act to filter out high frequency noise and improve the ability to detect the presence of an underlying, low contrast, stimulus just as the with the rate - based model in van Rossum et al. (2008). We found however, that recurrent depressing synapses did not improve the ability of a population of neurons to filter out high frequency noise from a low amplitude stimulus. In fact, over a range of noise parameters, the recurrent connectivity actually reduced the ability of a neural population to filter high frequency noise, indicating that care should be taken when generalising the results obtained from rate - based models. We finish the thesis with a short conclusions chapter.

Short - term synaptic depression

Much of this thesis involves examining networks featuring short - term synaptic depression. This is a phenomenon in which (chemical) synapses between neurons can decrease in efficacy in response to repeated stimulation. short - term synaptic depression can be viewed as part of group of phenomena known collectively short - term synaptic plasticity. In this section, we include a short review of the physiology of synaptic depression, as well as some of the methods which have been proposed to model it.

As a brief terminology background (readers interested in the properties and biochemistry of synaptic transmission in general should consult a secondary source such as Kandel et al., 2000) a synapse is a connection which carries signals from a presynaptic neuron to a postsynaptic neuron using intermediate messenger compounds called neurotransmitters[1]. An action potential (also referred to as a spike) initiated in the presynaptic neuron will travel up the axon to the presynaptic terminal, activating voltage gated calcium channels and leading to an influx of calcium ions. This causes some synaptic vesicles to release the neurotransmitter they contain into the synaptic cleft. This diffusing neurotransmitter then binds receptors in the membrane of the postsynaptic neuron which can, in turn, change the conductance of the postsynaptic neuron. Depending on the cell types involved, this typically leads to the depolarisation of the

[1]Throughout this thesis, we refer only to chemical synapses and not electrical synapses. Furthermore, the directionality of a chemical synapses is somewhat of a simplification - such synapses can also carry information from a postsynaptic neuron to a presynaptic neuron.

postsynaptic neuron (via an excitatory post synaptic potential, or EPSP), or to the hyperpolarisation of the postsynaptic neuron (via an inhibitory PSP, or IPSP). We refer to the *strength* or *weight* of the synapse as being the magnitude of the effect which a spike in the presynaptic neuron has on the postsynaptic neuron - typically, this term refers to the peak of an PSP or of the conductance change which leads to that PSP.

short - term synaptic plasticity involves changes in the strength of a synaptic connection due to ongoing spiking activity, and is the result of a number of distinct processes. The most rapidly acting of these processes are *facilitation* (which acts to enhance synaptic strength following repetitive stimulation) and *depression* (which acts to reduce synaptic strength following repetitive stimulation) - frequently, synapses are both facilitating and depressing, albeit with different timescales and strengths (see e.g. Zucker and Regehr, 2002 for an overall review). Whether depression or facilitation dominates can depend on the class of neurons involved and on the firing rate frequency of the presynaptic neuron. Typically, excitatory synapses between pyramidal neurons in the cortex are dominated by synaptic depression, whilst facilitation plays a larger role in synapses to and from interneurons (Thomson and Deuchars, 1994; Thomson, 1997; Varela et al., 1997; Tsodyks and Markram, 1997; Tsodyks et al., 1998; Mercer et al., 2005). The dynamics of cortical excitatory synapses in sensory areas can often be reasonably well modelled by depression alone (Abbott et al., 1997; Tsodyks and Markram, 1997; Varela et al., 1997).

In much of this thesis, we focus on the effect of short term synaptic depression. A key contributor to synaptic depression is thought to be the depletion, by activity, of a pool of vesicles which is 'available' for release. Suppose that there are N vesicles in this 'readily releasable pool' - according to the depletion model as first proposed (Liley and North, 1953; Betz, 1970), an action potential will release a fraction of these vesicles. For high N and fixed release probability p this is well matched by the mean reduction, $N \rightarrow N(1-p)$. The neurotransmitter in these vesicles will have an effect on the postsynaptic neuron proportional to pN, whilst the readily releasable pool of vesicles, now of size $N(1-p)$, recovers exponentially to its resting size N_0 by replenishment from a large reserve store of vesicles. This process naturally leads to depression in the amount of neurotransmitter released - should a further spike arrive presynaptically before any recovery can commence, then only $pN(1-p)$ vesicles can be released, less than the previous pN.

This simplified model does not account for all the observed properties of depression. For example, during repetitive stimulation, there is an increase in the rate of

recovery from depression (Kusano and Landau, 1975). This has been found to be dependent on inter - cellular calcium ion concentration in the presynaptic terminal, where high calcium ion concentration increases the magnitude of recovery from depression (Dittman and Regehr, 1998)[2]. Other sources of depression include the desensitisation of postsynaptic receptors (Wong et al., 2003) and the inactivation of presynaptic calcium channels (Forsythe et al., 1998). Finally, short term depression has been observed to be weaker in the presence of background activity - this could be important when examining how synaptic depression can impact network activity *in vivo* (Boudreau and Ferster, 2005; Reig et al., 2006).

We are particularly interested on how short term synaptic depression could impact the processing of inputs or ongoing network activity, and any possible advantages that synaptic depression might lead to (as opposed to being some sort of problem which sustained inputs must overcome). There have been range of interesting theoretical studies of synaptic depression. A common thread of enquiry is how synaptic depression leads to transient responses to input changes. For example, a classic paper by Abbott et al. (1997) postulates that synaptic depression (modelled with a simple depletion model[3]) allows for a certain amount of gain control in a neural system. They note that for a presynaptic neuron driven by a sustained firing rate r, the increase in the sustained presynaptic firing rate $r \rightarrow r + \delta r$ is communicated to the postsynaptic neuron as a transient increase in input asymptotic to $\delta r/r$. This allows a postsynaptic neuron to respond based on the *relative* changes of the firing frequency of its inputs, rather than being driven by the weighted average of their firing rates (which will be dominated by high frequency inputs which 'wash out' low frequency inputs). The concept that it is advantageous for input signals to be driven by firing rate changes has been also explored by other works - for instance, Puccini et al. (2007) combined short term synaptic depression with a form of spike frequency adaptation to model

[2]The multiple timeconstants of synaptic depression recovery (including calcium ion dependent recovery) is one of the things which point to the depletion model being overly simplistic. Rather than dividing vesicles into a readily releasable pool and a reserve pool, an alternative approach favours a three pool system (Dittman and Regehr, 1998; Rizzoli and Betz, 2005). In this framework, it is proposed that vesicles can be divided into three groups - the readily releasable pool, the recycling pool and the reserve pool. The readily releasable pool is replenished rapidly by the recycling pool which is replenished slowly by the reserve pool - the recycling pool can also be replenished directly through endocytosis.

[3]Most modelling studies investigating synaptic depression focus on variants of the depletion model (Abbott et al., 1997; Tsodyks and Markram, 1997) and its continuous, rate based version. The latter is articulated most clearly in Tsodyks et al. (1998). Note that the derivation of the depression dynamics for rate based neurons assumes that spike trains are well modelled as being generated by a non homogeneous Poisson process. This assumption does not fit the experimental data see, e.g. Teich (1989); Oram et al. (1999) - whether this has a big impact on the form of the resultant equation is unclear.

how a neural population could calculate the derivative of continuously varying signal, and anticipate its future values (through forward integration). Other studies examine how synaptic depression can aid in the processing of input signals, often in the light of experimental data. For instance, Goldman et al. (2002) explored how synaptic depression acts to reduce the auto - correlation observed in spike trains recorded from area V1 of a monkey freely viewing natural scenes, whilst a network model involving recurrent connectivity subject to synaptic depression has been proposed to help temporally integrate low contrast inputs, whilst also fitting data regarding the significant dependence on contrast of the onset latency of stimulus responses observed in higher visual areas (van Rossum et al., 2008). Another model driven by experimental observations was proposed by Carandini et al. (2002) in which a feed - forward model with depressing synapses can account for both the contrast saturation and cross - orientation suppression observed in V1. Meanwhile, a more theoretical approach by Pfister et al. (2010) proposed that the membrane potential induced by a synapse subject to short term depression can be viewed as behaving in a similar fashion to the optimal estimate of the presynaptic membrane potential given the previously observed spike history[4]. This range of investigations suggests diverse alternative roles for short term synaptic depression in the brain. We attempt, in this thesis, to suggest other possible roles for synaptic depression and try to connect further how short term synaptic depression can explain some observed stimulus response phenomena.

Whilst synaptic models incorporating both facilitation and multiple source of depression might better fit the precise properties of short term synaptic depression (Varela et al., 1997; Hennig et al., 2007), in this thesis we restrict ourselves to examining depressing synapse using the simple depletion model. This model can provide a reasonable fit to the dynamics of synaptic transmission in the cortex (Tsodyks and Markram, 1997) and helps to restrict the number of parameters that our models require to specify. We will talk in each section as to the appropriateness of each model which we use.

[4]Facilitation was included in the treatment by Pfister et al. (2010), but synaptic depression was the dominant feature involved when examining membrane potentials driven by OU processes, with facilitation more important with membranes existing in ‘on’ or ‘off’ states. Other work has been done examining the combined impact of facilitation with depression - see for instance the model of working memory proposed by Mongillo et al. (2008), in which calcium ion concentration takes the place of reverbatory synaptic activity as the source of working memory representations.

Chapter 2

Synaptic depression and ring attractor networks

Cortical circuitry shows an abundance of recurrent connections. A widely used model that relies on recurrence is the ring attractor network, which has been used to describe phenomena as diverse as working memory, visual processing and head direction cells. Commonly, the synapses in these models are static. In this chapter, we examine the behaviour of ring attractor networks when the recurrent connections are subject to short term synaptic depression, as observed in many brain regions. We find that in the presence of a uniform background current, the network activity can be in either of three states: a stationary attractor state, a uniform state, or a rotating attractor state. The rotation speed can be adjusted over a large range by changing the background current, opening the possibility to use the network as a variable frequency oscillator or pattern generator - this provides a possible mechanism for a sequential task, such as singing, to be carried out at different speeds. Finally, we extend the network to two - dimensional fields and find a rich range of possible behaviours. This chapter has previously appeared as a publication (York and van Rossum, 2009).

2.1 Introduction

The ring attractor network is a standard model in computational neuroscience. In the ring attractor network, neurons are arranged on a ring and have a rotation invariant connectivity. If the localised recurrence is strong enough, a weakly tuned input can lead to sharply tuned network activity (Ben-Yishai et al., 1995; Somers et al., 1995). In particular, with strong recurrent connections, the network can maintain a so called

attractor state or 'bump state' even in the absence of tuned input (Hansel and Sompolinsky, 1998). Bump states require that the pool of neurons keeps exciting itself while suppressing neurons that are not part of the pool. However, if not enough self - excitation can be maintained, the memory state will not be stable. This happens for instance when the neurons are subject to sufficient spike frequency adaptation. The adaptation will lead to a decay of the recurrent excitation, and as a result the bump state can become unstable and will start to propagate around the network (Hansel and Sompolinsky, 1998).

The ring model was originally introduced to model orientation selectivity in the primary visual system. Orientation selectivity in the primary visual system is contrast invariant - the width of a tuning curve is largely independent of the contrast with which stimuli are being presented (Sclar and Freeman, 1982; Skottun et al., 1987). This cannot be accounted for by the feed - forward model proposed by Hubel and Wiesel (1962), and several different models attempting to account for this property have been proposed (Ferster and Miller, 2000). The ring attractor network outlined by Ben-Yishai et al. (1995) features strong recurrent connectivity - neurons representing similar orientation preferences are connected to each other with strong excitation, whilst neurons of orthogonal orientation preferences strongly inhibit each other. When the network is presented with a weakly tuned input, it sharpens it through the recurrent connections until there is a central 'bump' of activity around the peak of the input, maintained through recurrent excitation, whilst this bump inhibits other orientations through strong surround inhibition. The width of the tuning curve is most determined by the aspect of the recurrent connections rather than the amplitude of the inputs, leading to contrast invariant tuning. (see also Somers et al., 1995 for a spiking neuron model incorporating similar ideas).

Although, as stated, the relevance of the ring model to the primary visual cortex is still under debate, with alternative mechanisms proposed which rely more on feed - forward connections (Ferster and Miller, 2000; Finn et al., 2007), the essential properties of the ring model - namely a periodic network with a centre - surround connectivity have been used to model many phenomena, such as prefrontal working memory (Compte et al., 2000) and head direction cells (Zhang, 1996; McNaughton et al., 2006; van der Meer et al., 2007). Its properties are thus an important target of investigation. One interesting avenue of research into ring attractor networks has been the study of the temporal dynamics of the ring attractor under different experimental setups. For instance, ring attractor networks with either split excitatory and inhibitory populations,

or those incorporating spike frequency adaptation can support the existence of travelling waves of activity in the absence of any tuned input (Ben-Yishai et al., 1997; Hansel and Sompolinsky, 1998).

In this study we consider the effect of incorporating short - term synaptic depression in the recurrent connections of a ring attractor network whilst the network receives a homogeneous background input. Synaptic depression has been observed in many of the synapses in the brain and is thought to be a major factor in the dynamics of neural systems (Tsodyks et al., 1998; van Rossum et al., 2008 - it should also be noted that in some areas synaptic responses can also have a strong facilitatory component Wang et al., 2006). Chapter 1 includes a review of some of the characteristics of synapses subject to short - term depression. Existing work has examined the effect of short - term depression on recurrently connected networks. One example is given by Tsodyks et al. (2000), in which a network of recurrently connected neurons with a range of different basal rates of firing has found that, when those recurrent connections are subject to short - term synaptic depression the network is capable of synchronising it's activity in the form of 'population spikes' - transient increases in firing encompassing the whole network. Attractor neural networks storing a number of patterns have also been examined - it was found that when their synapses are made subject to short - term depression, the memory states they store are now less stable, but that switching between memories is easier (Bibitchkov et al., 2002; Pantic et al., 2002). Furthermore, for some parameter regimes, the network activity spontaneously switches between the different memory states - we will see that such behaviour has relevance to our model. Finally, it has been suggested that short - term synaptic depression can allow recurrently connected networks to display rhythmic pattern generation without relying on input from pacemaker cells (Senn et al., 1998). What these studies have in common is that they indicate the temporal characteristics of networks recurrent connectivity can be significantly altered by short term synaptic depression leading to interesting modes of operation, which static synapses cannot account for. We add to this overall picture with our work in this chapter.

2.2 Method

We consider the model depicted in Fig. 2.1A. The neurons are labelled by an angular variable θ (between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$) representing their 'preferred orientation'. The neurons can be thought of as being placed on a ring. This does not imply the presence of

such ring structures in the neuroanatomy, merely that the neurons tuned to a periodic variable are functionally connected in a 'centre - surround' fashion. The activity is described by the neural firing rate m, with the time evolution

$$\tau_0 \frac{\partial m(\theta,t)}{\partial t} = -m(\theta,t) + g(I(\theta,t)), \tag{2.1}$$

where τ_0 is the neural time constant. The gain function g describes the relation between the synaptic input and the firing rate of the neuron. We use a rectifying non - linearity $g(I) = [I-T]_+$, with threshold T, which is set to zero without loss of generality. $I(\theta,t)$ is the synaptic input into the neuron with orientation preference θ. The input consist of two parts - recurrent synaptic input from the rest of the network, and the external input $I^{in}(\theta,t)$,

$$I(\theta,t) = \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} J(\theta-\theta')p(\theta',t)m(\theta',t)d\theta' + I^{in}(\theta,t). \tag{2.2}$$

The function J describes the synaptic weights between the different neurons in the network. J commonly takes the form of a "Mexican Hat", with an excitatory centre and inhibitory surround. As in previous work (Ben-Yishai et al., 1995, 1997; Hansel and Sompolinsky, 1998), the strength of interactions between neurons depends only on the difference in orientation

$$J(\theta-\theta') = J_0 + J_2\cos(2(\theta-\theta')). \tag{2.3}$$

We describe J_0 as the mean recurrent strength, and J_2 as the localised recurrent strength. We will assume that $J_2 > 0$ so that the weight profile has a centre - surround character.

In Eq. (2.2), $p(\theta,t)$ denotes the release probability and includes synaptic depression. Synaptic depression is a form of short - term plasticity in which the arrival of a spike at a synapse temporarily reduces the efficacy of signal transmission at that synapse (Markram and Tsodyks, 1996). Synaptic depression has been modelled for both spiking (Abbott et al., 1997; Tsodyks and Markram, 1997) and rate - based model neurons - we follow the latter formulation (Tsodyks et al., 1998). Let each synapse have a release probability $p(t)$ which governs the efficacy of the synapse and let U be the utilisation of synaptic efficacy, i.e. the reduction in the synaptic efficacy per spike. Under the assumption that the presynaptic spike train is Poisson, one has, from Tsodyks et al. (1998),

$$\tau_d \frac{dp(t)}{dt} = 1 - p(t) - \tau_d U p(t)m(t). \tag{2.4}$$

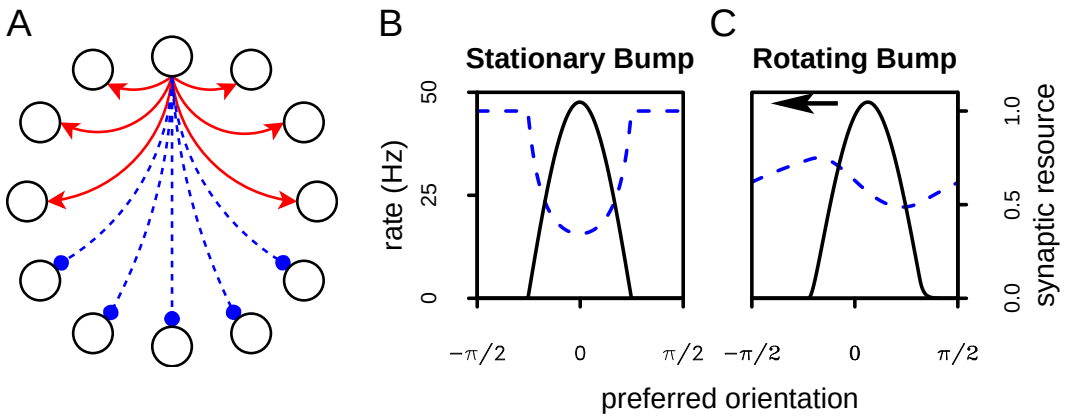

Figure 2.1: **A** Schematic of the network. The neurons are arranged in a ring, with a rotationally symmetric weight matrix which connects nearby neurons with excitatory connections (solid red line, stronger for neurons close to each other), and inhibitory connections to neurons on the opposite side of the ring (dashed blue line, strongest negative connections for neurons furthest from each other). **B+C** The firing rate (solid black line, left y - axis) and synaptic resource (dashed blue line, right y - axis) **B** The state corresponding to a narrow profile steady state solution called a 'bump state'. **C** The firing rate and synaptic resource from a single time - point corresponding to a rotating solution. In this example, the profile is moving to the left.

The effective output from a neuron with firing rate $m(t)$ and release probability $p(t)$ is proportional to $m(t)p(t)$ (Tsodyks et al., 1998).

We are interested in the behaviour of the system when it receives an isotropic input, i.e. $I^{in}(\theta-\theta_0)=B$, which we assume to be supra - threshold, i.e. $B>0$ (but see below). In summary, the following coupled differential equations describe the network

$$\begin{aligned}\tau_0\frac{\partial m(\theta,t)}{\partial t} &= -m(\theta,t)+\left[\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\{J_0+J_2\cos(2(\theta-\theta'))\}p(\theta',t)m(\theta',t)d\theta'+B\right]_+,\\ \tau_d\frac{\partial p(\theta,t)}{\partial t} &= 1-p(\theta,t)-\tau_d U p(\theta,t)m(\theta,t).\end{aligned} \tag{2.5}$$

Network states in the absence of synaptic depression

Before we study the effect of synaptic depression, we briefly review the behaviour of the model with static synapses. In the absence of synaptic depression ($U=0$, $p(\theta,t)=1$), the ring attractor network with constant, homogeneous input ends up in one of three possible regimes, depending on the parameters of the model and the initial conditions (Hansel and Sompolinsky, 1998). 1) The network undergoes amplitude instability -

the firing rates diverge for all neurons. 2) The network activity converges to a constant, homogeneous stable steady state solution, in which every neuron fires at the same, constant, rate (or all neurons fall silent, if the input is smaller than the threshold). 3) The network enters the marginal phase regime - and settles into a localised bump state. There is a continuum of such states, each with a different location, hence also the name ring attractor or line attractor. In this case the symmetry is broken and the location of the activity bump can be used as a memory storage. The shape of the activity bump is denoted $M(\theta - \theta_0)$ where θ_0 denotes the centre of the bump, and

$$M(\theta) = A[\cos(2\theta) - \cos(2\theta_c)]_+ \quad . \tag{2.6}$$

The bump location θ_0 is determined by the initial conditions. The angle θ_c represents the width of the activity bump.

Parameter choices

In the next section, we will consider the steady state solutions of the activity equations and their stability in the presence of synaptic depression. But before doing so, we discuss the various parameters and their default value. The system has two timeconstants: the neural firing rate time constant τ_0 and the depression recovery time constant τ_d. Given the typically observed values for these parameters (Wang et al., 2006; Thomson and Lamy, 2007), we can safely assume that $\tau_d > \tau_0$.

Eqs. (2.5) can be rescaled by introducing a normalised time t/τ_0, a normalised activity $m/(U\tau_d)$, and a normalised current $I/(U\tau_d)$. This reduces the effective number of parameters to four: J_0 and J_2, which together describe the connectivity profile; $\tau := \frac{\tau_d}{\tau_0}$, the ratio of the synaptic and depression recovery time constants; and $b := \tau_d UB$, the rescaled background input. In order to observe qualitatively different activity patterns we only need to vary these four parameters instead of the six parameters in Eqs. (2.5). This is still too many for most purposes, so we shall, in general, vary the rescaled background input b and the tuned part of the recurrent connections J_2, whilst leaving J_0 and τ fixed. For clarity, we continue to represent the model in the original parameters, and vary b by varying B and fixing $U = 0.2$. Unless denoted otherwise we fix $J_0 = 0$, $\tau_0 = 5\,\text{ms}$ and $\tau_d = 50\,\text{ms}$.

2.3 Results

2.3.1 Behaviours of the ring attractor with synaptic depression

In the presence of synaptic depression the network will eventually settle to one of three types of behaviour - a homogeneous steady state, a stationary bump state, or a rotating state. In the homogeneous steady state all neurons fire at the same, constant rate; the synaptic resource is also constant across the network. A stationary bump state is formed by a subset of the neurons firing in a self - reinforcing manner, causing the appearance of a bump of activity, Fig. 2.1B, similar to the bump state in the non - depressing network. Finally, in the rotating state a localised bump of activity propagates around the network, leaving a wake of replenishing synaptic resource, Fig. 2.1C. Note that in the presence of synaptic depression there is no longer a diverging state. A phase diagram illustrating these states is shown in Fig. 2.2, where we use analytic results to calculate the stability of the homogeneous solution, numerical stability analysis to analyse the stability of the bump state, and computational simulations to measure the rotation speed in rotating states. The three different regions will be described in more detail in the following subsections, while the stability calculations of the various regimes are presented in the appendix. For completeness, if we allow sub - threshold input ($B < 0$), a fourth possible state is the silent state - identical to the homogeneous steady state, except that the activity is zero, and the synaptic resource is equal to 1.

2.3.1.1 Homogeneous Steady State Solution

We consider first the broad profile steady state solution, 'broad' indicating that we look for steady state solutions such that $m(\theta,t) = M(\theta) > 0 \ \forall\theta$. Simulation results indicate that inhomogeneous solutions of this type are unstable. We thus consider only a homogeneous solution, i.e. one where all neurons have the same firing rate (appendix, Eq. 2.11), given by

$$M(\theta) = M_0 \quad = \quad \frac{1}{2\tau_d U}\left(\tau_d UB + J_0 - 1 + \sqrt{(\tau_d UB + J_0 - 1)^2 + 4\tau_d UB}\right).$$

Note that if $J_0 = 0$, the steady state solution reduces to $M_0 = B$, again assuming $B \geq 0$. In general, M_0 increases with increasing input B and increasing mean recurrence J_0.

The key question is whether the homogeneous solution is stable. In other words, if we displace network activity slightly away from the steady state, does the activity relaxes back to it or not? In section 2.A.1 we calculate the range of parameters for

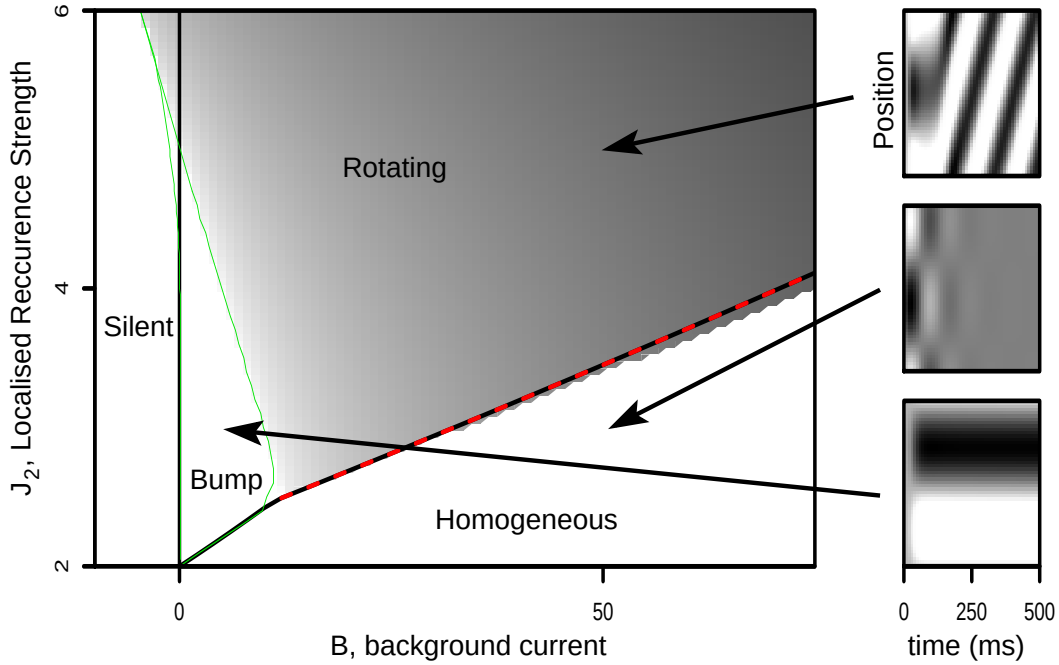

Figure 2.2: Left, the phase diagram of the network as a function of the background input B and the localised recurrent strength J_2. The various states are: a stationary bump, a rotating bump, a homogeneous state, and a homogeneous state where all neurons are silent. The region delimited by the green curve indicates where the stationary bump state is stable. The black line indicates the boundary of the regions where the homogeneous solution (right) or the silent state (left) are stable. The shaded region indicates the region where rotating solutions are stable (the darker the colour, the faster the rotation speed). This region overlaps slightly with the stability regions alongside the boundary lines, allowing for the coexistence of steady state solutions and rotating solutions. The red dashed line represents the segment of the instability line along which we make the analytic speed approximation in Fig. 2.3. Right, samples of the simulated network activity (y - axis), plotted against time (dark shade indicates high firing rate), illustrating each of the three activity regimes, starting from random initial conditions. The arrows indicate the parameters of B and J_2 which were used in each simulation.

which the homogeneous solution is stable, by linearising the equations under small perturbations. We find that the homogeneous steady state solution is always stable to homogeneous fluctuations in activity and release probability, but the network is not necessarily stable to localised fluctuations in activity and release probability. The precise condition depends on the value of the steady state solution M_0. For small inputs, for which $M_0 \leq \frac{\tau_0}{\tau_d U(\tau_d-\tau_0)}$, stability requires that (see Appendix 2.A.1, Eq. 2.14)

$$J_2 < 2(1+\tau_d U M_0)^2.$$

Alternatively, for larger values of M_0, $M_0 > \frac{\tau_0}{\tau_d U(\tau_d-\tau_0)}$, stability requires that

$$J_2 < 2(1+\tau_d U M_0) + 2\frac{\tau_0}{\tau_d}(1+\tau_d U M_0)^2.$$

These conditions are illustrated by the black line in Fig. 2.2 - the kink in the line occurs at the transition between the conditions for large, and small input. These conditions also imply that increasing input, which will increase M_0 (i.e. moving to the right in Fig. 2.2), will eventually always stabilise the constant solution. The synapses are in that case so strongly depressed that the recurrent connections become too weak to maintain an inhomogeneous steady state solution.

If $J_0 \neq 0$, then the stable homogeneous region, when plotted against J_2 and M_0 (in the place of B) has the same shape as in Fig. 2.2. However, in this case it can occur that the homogeneous solution is *always* stable. In particular, if $2J_0 > J_2$ it holds that $M_0 > \frac{J_2-1}{2\tau_d U}$ for all $B > 0$, which always satisfies the homogeneous stability criteria. In other words, the homogeneous steady state solutions are always stable in networks dominated by excitation.

Finally, it is interesting to see that before reaching the steady state, the activity oscillates, see Fig. 2.2, right - middle. This resembles the two - population oscillations studied in Senn et al. (1998).

2.3.1.2 Narrow Steady State Solution

The narrow profile steady state solution takes the form of a localised region of activity (a bump), Fig. 2.1B. For a profile centred at θ_0 we follow previous studies (Hansel and Sompolinsky, 1998), defining the width of the profile by a critical angle θ_c such that the activity is zero for $|\theta-\theta_0| > |\theta_c|$, and writing $M(\theta) = A\left[\cos(2(\theta-\theta_0)) - \cos(2\theta_c)\right]_+$. In general, the values of $A > 0$ and θ_c must be solved self - consistently and can be expressed only in implicit form. Again whether the network ends up in this state,

depends on the stability of this stationary state, which is analytically intractable. To determine the stability, we instead discretise the network into a finite number of neurons, and find the steady state solution and its stability using numerical methods (see Appendix 2.A.2). The narrow profile solution is typically stable for small background inputs, Fig. 2.2. At higher levels, the greater activity in the network leads to stronger depression, weakening the lateral connections necessary to maintain the localised profile, leading to rotating or homogeneous solutions.

An alternative approach to determine stability is to simplify network dynamics by assuming that activity reduces the release probability in proportion to $m(\theta,t)$, rather than $m(\theta,t)p(\theta,t)$ (detailed in section 2.A.3, Eq. 2.15). This linear approximation describes the network dynamics as

$$\begin{aligned}\tau_0\frac{\partial m(\theta,t)}{\partial t} &= -m(\theta,t)+\left[\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}J(\theta-\theta')p(\theta',t)m(\theta',t)d\theta'+B\right]_+\\ \tau_d\frac{\partial p(\theta,t)}{\partial t} &= 1-p(\theta,t)-\varepsilon m(\theta,t),\end{aligned}$$

which allows for an analytical treatment of the stability. We call this subtractive depression in contrast to the more common multiplicative depression model, Eq. (2.4), used above. The simplification is correct in the limit where the release probability is close to one for all θ, as happens at low firing rates, so that we may make the approximation $\tau_d U m(\theta,t)p(\theta,t)\approx\varepsilon m(\theta,t)$.

This simplification allows us to consider the stability of a transverse perturbation, where the activity and synaptic resource keep the same shape as the stationary state solution, but are displaced (see section 2.A.3). For our parameters, this approximation of the depression dynamics works reasonably well to predict stability, Fig. 2.3A. In addition, the shape of the stability boundary is a reasonable fit to that of the multiplicative depression. The approach becomes less accurate as the background input is increased and the release probability deviates much from 1.

2.3.1.3 Rotating Profile

In the regions of parameter space where neither the homogeneous steady state solution nor the narrow profile steady state solution are stable, the activity of the network starts to rotate. In the example of Fig. 2.1C the release probability of the neurons on the left side of the activity bump is higher than on the right - causing the neurons there to fire with higher efficacy. This means that the ‘centre of mass’ of the synaptic output from the activity profile lies to the left of the profile itself, and will draw the profile to the

left, continually consuming synaptic resource. As the profile moves away, the synaptic resource in depleted areas replenishes. The direction in which the profile moves is determined by the initial conditions, and obviously requires asymmetric initial conditions.

The rotation speed was measured using network simulations, grey shaded region in Fig. 2.2. However, near the transition between stability and instability of the homogeneous steady state solution, it can be calculated analytically by considering transverse fluctuations (black line in Fig. 2.2). Here the system is well approximated by the linearisation about the stationary state as described above. The boundary line for $M_0 \geq \frac{\tau_0}{\tau_d U(\tau_d - \tau_0)}$ is given by (appendix, Eq. 2.14),

$$J_2 = 2(1 + \tau_d U M_0) + \frac{2\tau_0(1 + \tau_d U M_0)^2}{\tau_d}. \tag{2.7}$$

The speed is given by the imaginary part of the eigenvalues obtained from the stability matrix (Appendix, Eq. 2.13)

$$v = \sqrt{\frac{\tau_d U M_0(\tau_d - \tau_0) - \tau_0}{\tau_d^2 \tau_0}}, \tag{2.8}$$

measured in rad/s. Eq. (2.8) and the speed from simulation are virtually identical, Fig. 2.3B.

Since the speed is increasing with increasing background current, the value of the background current for which Eq. (2.7) is true can be viewed as the point at which the maximum speed is attained for a given network. For higher background input, the homogeneous stationary state becomes stable. The rotation speed can thus be varied between 0 and the value given by Eq. (2.8) by adjusting the background current.

2.3.1.4 Multi - stability

Narrow parameter regions exist where more than one state is stable. First, for the case that $B > 0$ there is a bistable region of bump and rotating states, Fig. 2.3C, left. Secondly, a larger bistable region occurs with both rotating and homogeneous states, Fig. 2.3C right. This second region is also the cause that in Fig. 2.2 the red dashed line does not mark the extent of the rotating solution. In these bistable regions, the final activity is determined by the initial conditions. In contrast, the stability of the homogeneous and bump solutions do not appear to overlap when $B > 0$, although we did not show this analytically.

If we consider sub - threshold input, $B < 0$, other regions of multi - stability appear. Fig. 2.3D shows an enlarged plot of this region of phase space. In particular, for sub - threshold input the silent state, $M(\theta) = 0$, is always stable. This allows for an extra region of parameter space which supports multi - stability. There are two narrow regions of multi - stability - one where both the silent state and the rotating state are stable for large J_2 (marked **), and below that a region where both the stationary bump and the silent state coexist (marked *).

2.3.2 Two dimensional neural fields with synaptic depression

The model can be extended to situations in which the stimulus parameter has more than one dimension. Two dimensional models (either as a closed plane, or as an open manifold with a torus topology) have been proposed as a model for place cells, where the localised activity bumps code the location of the animal (Samsonovich and McNaughton, 1997; McNaughton et al., 2006). Our simulations of these networks show a behaviour similar to the 1d ring network - namely that varying the constant input B can switch between a number of different states and, as in the one dimensional case, increasing the background input tends to increase propagation speeds. In Fig. 2.4A, the activity profile from a two dimensional network presented with a constant background input. The neurons are arranged upon a torus, with periodic boundary conditions, and a centre - surround "Difference of Gaussian" connectivity profile

$$J(|x-y|) = J_0 + J_2 \left[\frac{1}{2\pi\sigma_c^2} \exp\left(-\frac{|x-y|^2}{2\sigma_c^2} \right) - \frac{1}{2\pi\sigma_s^2} \exp\left(-\frac{|x-y|^2}{2\sigma_s^2} \right) \right].$$

In the simulations presented here, we use a grid (connected in the form of a torus) of 70 by 70 neurons, with $\sigma_c = 4$, $\sigma_s = 8$, $J_0 = 0$ and $J_2 = 4$, and with $\tau_d = 200\,ms$. Fig. 2.4 shows the limiting behaviours this network settles to (after random initial conditions) for three different background currents.

In Fig. 2.4A, with a background current of 5 Hz, localised, intrinsically propagating, bumps of activity arise - an analogue to the rotating bumps which can be observed in the ring case. These activity bumps are bound together in a repeating grid and move in a common direction (which, in general, has a temporal dependence). At high background input, homogeneous, diagonally travelling wave solutions appear, Fig. 2.4B (background input 18 Hz). Even higher levels of background input leads to homogeneous activity, similar to the situation in the ring network. At intermediate inputs, grid - like patterns such as shown in Fig. 2.4C arise (background input 15 Hz).

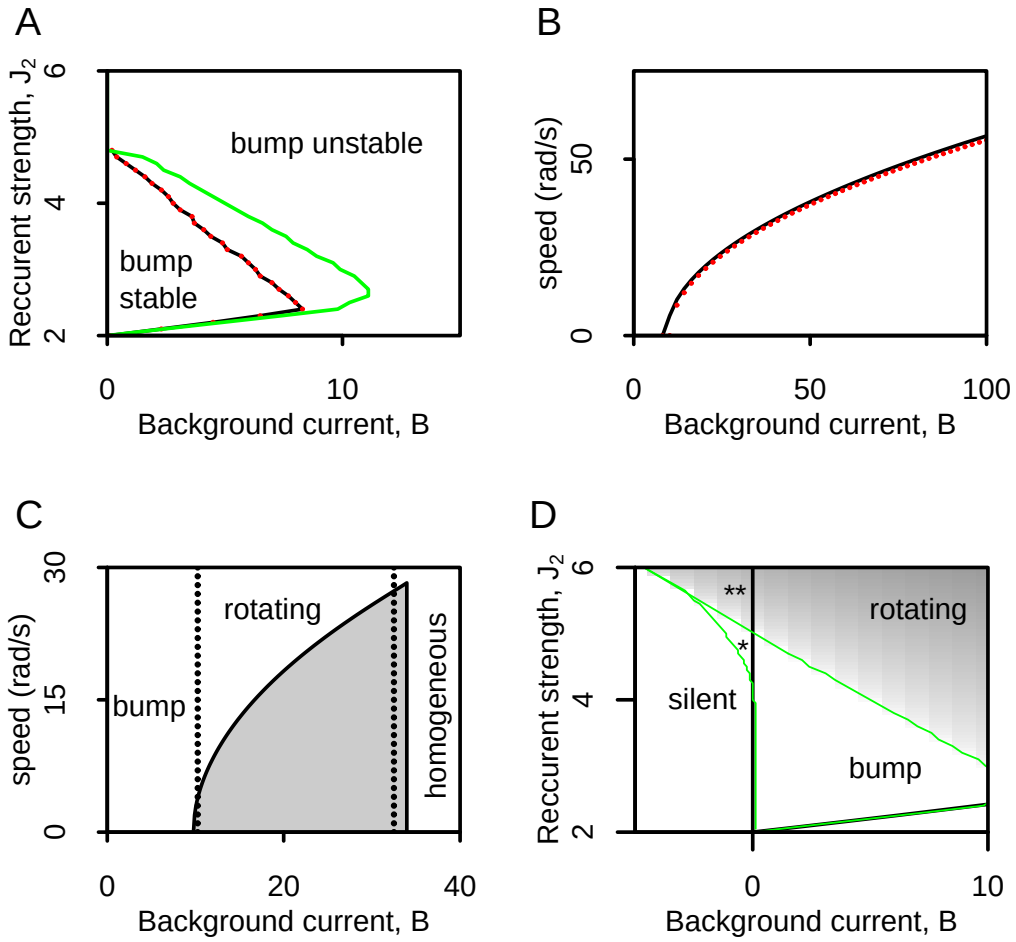

Figure 2.3: **A** The boundary between the region where the stationary bump profile is stable for the multiplicative depression (green, zoom of Fig. 2.2), and the subtractive depression approximation (black line and red dots). The boundaries have a similar shape. The green and black lines are created by comparing the eigenvalues of the full $2n \times 2n$ Jacobian for the multiplicative and subtractive variants respectively (n = 497 for green line, n = 200 for the black line). The red dots arise from the eigenvalues of the transverse fluctuation stability matrix, Eq. (2.17). **B** The rotation speed along the instability line given by Eq. (2.7), and marked by the red dashed line in Fig. 2.2. The simulated speed (red dots) matches the analytical expression (black line, Eq. 2.8). This expression is a good approximation for the maximum speed in the network where B is the corresponding input required for that maximum speed, and J_2 is given by the instability line in terms of B. **C** The rotation speed against the strength of the background input for the parameters $J_0 = 0$, $J_2 = 3$. As the background current is increased, the system transitions from stationary bump to a rotating solution, to a homogeneous solution. Increasing the background current will increase the propagation speed - there is an approximately linear relationship between the background current and the square of the speed. There are narrow bistable regions where both a rotating network state, as well as one of the stationary states are stable. The vertical dotted lines indicate, respectively, the maximum background input capable of maintaining a stable narrow profile steady state solution (left line), and the minimum background input required for the homogeneous steady state solution to remain stable (right line). **D** Zoomed in view of the regions of the parameter space from Fig. 2.2, showing a region of multi - stability for negative B. The lines and shading are as in Fig. 2.2. In the region marked ** both the silent state and the rotating solution are stable. In the region marked *, both the silent state and the bump state are stable.

For low amplitude background inputs (such as depicted in Fig. 2.4A), individual neurons spend most of the time with their inputs below threshold, punctuated by short deviations as a moving activity bump incorporates those neurons, Fig. 2.4D. The activity appears to be aperiodic. As the background input is increased (such as those in Fig. 2.4B,2.4C), the neurons in the network spend most of the time above threshold, Fig. 2.4E+F, and are active with a periodic firing pattern.

In Fig. 2.4G, networks with faster depression dynamics ($\tau_d = 50ms$) and a range of value of J_0 are displayed. For mean inhibition (J_0 negative, Fig. 2.4G left), there is only a single activity bump, whilst when J_0 is increased (to zero, Fig. 2.4G right), the number of bumps increases. In particular, the network in Fig. 2.4G left has a single bump for a range of input amplitudes, and can switch between a stationary and moving state by altering the background input, similarly to the one dimensional case (not shown). Such a network could plausibly be used to model dynamically changing two - dimensional data, such as the position of an animal, which is encoded in hippocampal place cells (Samsonovich and McNaughton, 1997; McNaughton et al., 2006).

These types of behaviour exist across a range of parameter values and network types, but some aspects are topology dependent. For instance, the diagonal travelling wave solution depicted in Fig. 2.4B does not appear when the network uses a spherical topology (the sphere has no equivalent to parallel lines, not shown). In addition, the common direction (and if such a common direction exists) of the moving bumps depicted in Fig. 2.4A depends on the topology.

2.4 Discussion

We have shown how ring attractor networks behave when the synapses are subject to short term synaptic depression and the network is stimulated with a constant background current. The activity will settle into one of three states - (a) a flat homogeneous state in which the activity and synaptic resource settle to constant, limiting values, (b) a marginal state, i.e. a stationary, narrowly tuned bump of activity centred around an arbitrary centre, and (c) a rotating state which consists of bump - like travelling waves of activity and synaptic resource, which propagate around the network. We have confirmed that these key results also apply to networks of noise integrate - and - fire neurons (not shown). Experimental evidence for changing activity states in the absence of input has been found using imaging of spontaneous *in vivo* activity in cat (Kenet et al., 2003 - we are indebted to Dmitri Bibitchkov for pointing this out) - our

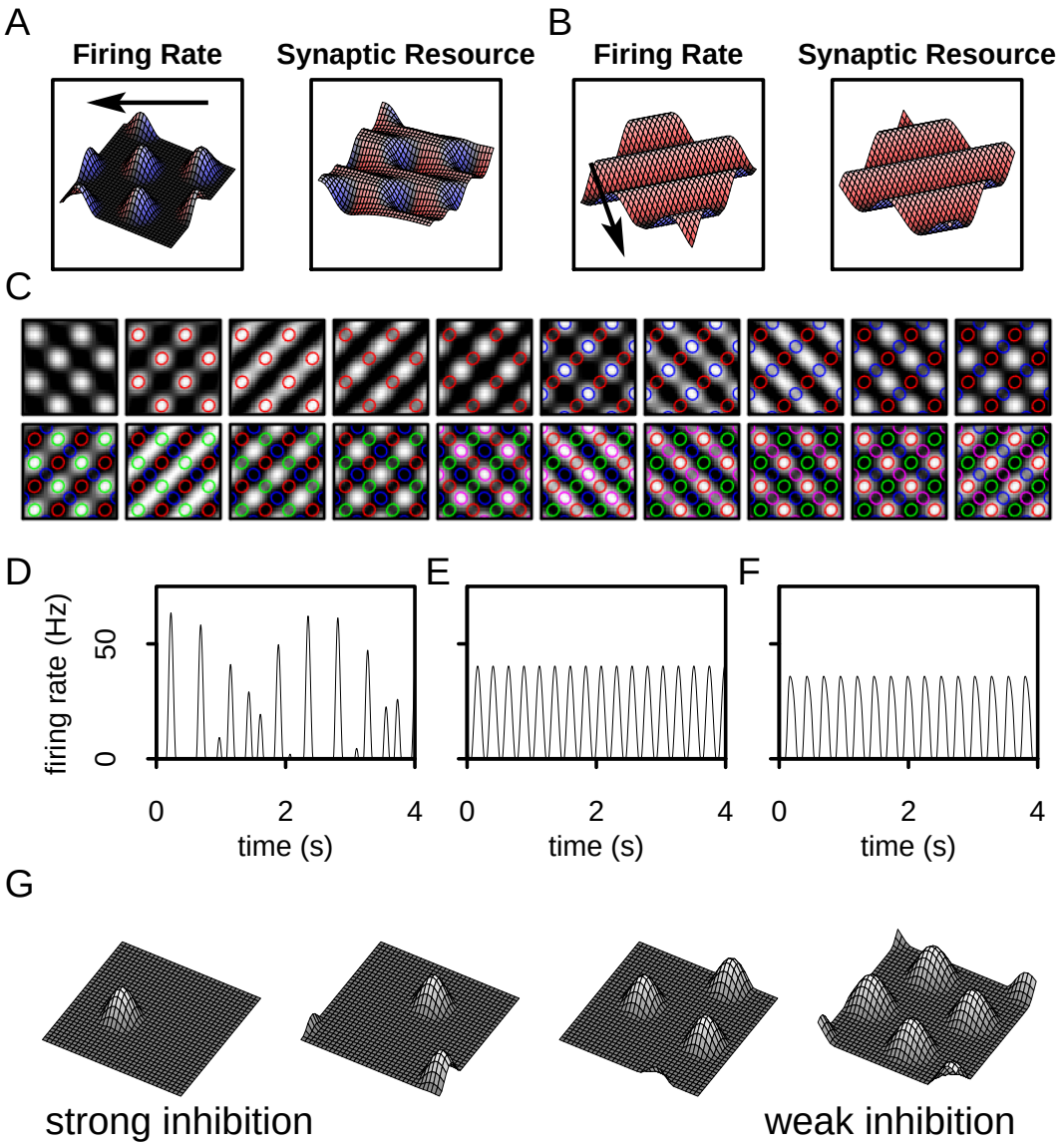

Figure 2.4: **A** The firing rate and synaptic resource in a two dimensional network with difference of Gaussian connectivity and a torus - equivalent topology stimulated with a background input of 5 Hz. The firing rate and synaptic resource graphs show the amount of firing rate or synaptic resource at a single time step, with the colour representing the rate of increase and decrease in the quantities. The greater the increase (resp. decrease), then the greater the red hue (resp. blue hue). The bumps of activity propagate across the surface of the network in common (although in this case not constant) direction; the arrow indicates the direction. **B** The same network, but now stimulated with a strong background input of 18 Hz, leading to a diagonally travelling wave solution. **C** The same network with an intermediate input of 15 Hz. The firing rate is plotted from twenty time steps, each 15 ms apart. The network makes transitions between four grid like pattens, which are identical up to translation. The activity is shown in grayscale (white means high activity), with the coloured lines added to show the location of the grid - like patterns as they appear. **D - F** The firing rate of a single neuron in the cases depicted in **A - C** over four seconds; the firing can either be periodic or not. **G** The number of bumps depends on the amount of inhibition J_0. When J_0 is strongly negative (left), only one bump of activity is supported. As J_0 is increased (to zero, right), a greater number of bumps of activity can be supported. All activity profiles here are drifting in a constant direction.

results indicate that recurrent, depressing, synaptic connectivity could naturally lead to the emergence of such dynamic patterns. An alternative model of these data assumes a non - adaptive ring attractor subject to noise (Goldberg et al., 2004). However, such a model leads to diffusion - like activity dynamics, very different from the mechanism proposed here, where activity propagates at a given speed. Careful analysis of the data should be able to distinguish between these different models.

The angular speed of the rotating state is controlled by varying the amplitude of the background current. This stands in contrast to ring attractor networks with firing rate adaptation where the rotation speed was constant for different background currents (Hansel and Sompolinsky, 1998). The ability to alter the speed of propagation by altering the background current opens up the possibility of such networks being used as variable speed oscillators, or as a mechanism for replaying a stored temporal sequence at a variable speed, such as happens in speech generation. Indeed, the ability to start or stop rotation based entirely on isotropic background activity allows for stationary bump to be moved to a new network position by transient increases in the isotropic input - allowing potentially for a mechanism in which temporal sequences can ‘pause’. Another mechanism which permits dynamically changing the rotation speed of a network is to separate excitatory and inhibitory populations, and to change the ratio between the background current to each of the populations. However, this approach requires spatially modulated inhibition, and will not work in networks with global inhibition (Ben-Yishai et al., 1997), in contrast to the mechanism proposed here.

The results we outline can also be examined in the context of previous work relating to networks subject to short - term plasticity. The intrinsically propagating wave that can emerge in our model changes its dynamics on the basis of the strength of isotropic input. Essentially, a higher background input causes the synapses to act on a faster timescale, a common feature of synaptic depression. This property is likely also to be present in other models. For instance, attractor networks subject to short - term depression also display, in some parameter regimes, spontaneous switching between multiple states, with slower recovery from depression leading to faster switching speeds (Bibitchkov et al., 2002; Pantic et al., 2002). It is possible that the speed at which this switching occurs will similarly increase if each neuron is supplied with a positive background current. Indeed, we suggest that the dependence on the speed of dynamics on the overall input amplitude is likely to be a general property of networks dominated by short - term depression - this may also apply experimentally.

Our approach can be criticised from the point of view of biological plausibility.

For example, the model of synaptic depression we use is a simple one (Tsodyks et al., 1998), in which there is only one timeconstant governing the speed of recovery from depression and no facilitation incorporated. In particular, synapses to and from inhibitory neurons are much more strongly facilitatory (Thomson and Deuchars, 1994; Thomson, 1997; Varela et al., 1997; Tsodyks and Markram, 1997; Tsodyks et al., 1998; Mercer et al., 2005). We don't take this into account - indeed it appears difficult to do so without separately incorporating excitatory populations and inhibitory populations of neurons - and the effect this might have on our results is not immediately apparent without extensive analysis. Furthermore, whilst we analyse the behaviour of a network across the full range of possible parameters, most the phase plane diagram in Fig. 2.2 is depicted when synaptic depression is weak ($U = 0.2$ and $\tau_d = 50$ ms). This is done to demonstrate all possible states in a single diagram, but for slower depression timeconstants the region where bump states are stable are much smaller, although weak synaptic depression is not necessarily un - physiological - see e.g. Boudreau and Ferster (2005); Reig et al. (2006). Finally, there is considerable heterogeneity observed in the properties of synaptic depression (Wang et al., 2006) - our approach of using only a single timeconstant and utilisation constant to describe every synapse is not necessarily realistic.

Our model contains other simplifications. For example, we incorporate rate based neurons, with a rectification firing rate non - linearity and with no noise incorporated, rather than spiking neurons. Previous work, involving networks of spiking neurons, has suggested that the addition of synaptic depression into a ring like model of working memory does not destabilise bump states (Barbieri and Brunel, 2007). Our study refines this view and finds that this is not always the case - parameter regimes exist in our model where the stationary bump is unstable and starts to rotate. Furthermore, the same study finds activity to be bistable, allowing for both a flat, homogeneous steady state and a bump state whilst in our much simpler model we only find a narrow bistable region when $B < 0$. It is possible that the bistability is due to the sigmoidal input - output functions in the spiking network, which ensures small localised perturbations from a low amplitude background state are dampened, whilst allowing for sufficient feedback at higher amplitudes in order to maintain a stable bump state (Barbieri and Brunel, 2007). This bistability is important for models of working memory, so that the memory can be turned off. Further investigation is needed to provide a full picture of how short - term depression interacts with other mechanisms in this kind of working memory model - but again we must note with caution that analysing conclusively the

behaviour of networks with lots of free parameters, compared to the limited number in our network, is non - trivial.

We have also examined networks arranged in two dimensional structures - with either a closed plane topology, or arranged upon the surface of a sphere, or a torus. A similar situation to the one dimensional model arises - namely that the limiting behaviour of the network presented with homogeneous background current can be either stationary, or moving, depending on the model parameters, and the amplitude of the background current. Localised activity bumps are a common occurrence, while other activity profiles were found to be topology dependent (such as travelling wave diagonals, not found on the sphere or in the closed plane). How different topologies interact with different behaviour types is an area requiring further examination. Qualitatively similar spatiotemporal patterns have been found in more mathematically orientated approaches on the infinite plane which incorporate spike frequency adaptation (Owen et al., 2007). Taking such an approach with synaptic depression is likely to be non - trivial, and analytical approaches may prove to be ultimately intractable.

Short term synaptic depression has been suggested to play an important role in cortical dynamics in a variety of network models (Abbott et al., 1997; Puccini et al., 2007; van Rossum et al., 2008) and has been proposed to underlie pattern generation in pre - cortical brain structures (Senn et al., 1998; Harris et al., 2002) and contribute to the rapid switching between memory states (Bibitchkov et al., 2002; Pantic et al., 2002). This chapter suggests that synaptic depression could play an important role in dynamic pattern generation in cortical areas as well and allow for variable speed rhythmic pattern generation. Our approach makes a number of simplifications compared with the situation observed experimentally, but we feel that the advantage of mathematical tractability offsets the associated drawbacks to this approach. The general conclusions that the simplified model has allowed us to make can be used to guide the investigation of networks with more biologically plausible synaptic models and parameter distributions, as well as to more closely investigate the impact of using different types of neuron model, such as spiking neurons, and the incorporation of background noise. Whilst we have focused on isotropic inputs, the impact of synaptic depression on tuned inputs is another potential avenue of further exploration. For the next chapter, we maintain our focus on synaptic depression, examining how networks featuring recurrent synaptic depression process pairs of stimuli presented asynchronously, in particular, to explain how the response to an early stimulus presented at high contrast can suppress the response to a later stimulus, and how anticipation in a sequence of

stimuli can occur.

2.A Appendix

In the appendices the details of the stability analysis are presented.

2.A.1 Stability of the homogeneous solution

We first analyse the broad solution, that is, the solution for which all $m(\theta,t)$ are positive and all neurons remain above threshold. Writing the functions in Fourier series form, $\hat{m}(\theta,t)=\sum_n \hat{m}_n(t)e^{2in\theta}$, $\hat{p}(\theta,t)=\sum_n \hat{p}_n(t)e^{2in\theta}$, and using that $2\,cos(2z)=e^{2iz}+e^{-2iz}$, we obtain from the governing equations, Eqs. (2.5), the following expansion:

$$\begin{aligned}
\tau_0\sum_n \hat{m}_n'(t)e^{2in\theta} &= -\sum_n \hat{m}_n(t)e^{2in\theta}+\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}(J_0+\frac{J_2}{2}(e^{i(\theta-\theta')}+e^{-i(\theta-\theta')}))\sum_n \hat{p}_n(t)e^{2in\theta'}\sum_j \hat{m}_j(t)e^{2ij\theta'}d\theta'+B,\\
&= -\sum_n \hat{m}_n(t)e^{2in\theta}+J_0\sum_n \hat{p}_{-n}(t)\hat{m}_n(t)+\frac{J_2}{2}\left(e^{2i\theta}\sum_n \hat{p}_{-n}(t)\hat{m}_{n+1}(t)+e^{-2i\theta}\sum_n \hat{p}_{-n}(t)\hat{m}_{n-1}(t)+B\right),\\
\tau_d\sum_n \hat{p}_n'(t)e^{2in\theta} &= 1-\sum_n \hat{p}_n(t)e^{2in\theta}-\tau_d U\sum_n \hat{m}_n(t)e^{2in\theta}\sum_j \hat{p}_j(t)e^{2ij\theta},\\
&= 1-\sum_n \hat{p}_n(t)e^{2in\theta}-\tau_d U\sum_n e^{2in\theta}\sum_j \hat{m}_{j+n}(t)\hat{p}_{-j}(t).
\end{aligned}$$

Since Fourier components are orthonormal, we can compare coefficients, to obtain

$$\begin{aligned}
\tau_0\hat{m}_0'(t) &= -\hat{m}_0(t)+J_0\sum_j \hat{p}_{-j}(t)\hat{m}_j(t)+B,\\
\tau_0\hat{m}_{\pm1}'(t) &= -\hat{m}_{\pm1}(t)+\frac{J_2}{2}\sum_j \hat{p}_{-j}(t)\hat{m}_{j\pm1}(t),\\
\tau_0\hat{m}_n'(t) &= -\hat{m}_n(t),\ |n|>1,\\
\tau_d\hat{p}_0'(t) &= 1-\hat{p}_0(t)-\tau_d U\sum_j \hat{m}_j(t)\hat{p}_{-j}(t),\\
\tau_d\hat{p}_n'(t) &= -\hat{p}_n(t)-\tau_d U\sum_j \hat{m}_{j+n}(t)\hat{p}_{-j}(t),\ |n|>0.
\end{aligned}\tag{2.9}$$

These equations are used below for the stability analysis. The equations for the steady state solution can be found by setting the derivative to be zero in Eq. (2.5), yielding

$$\begin{aligned}
M(\theta) &= \frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}[J_0+J_2\cos 2(\theta-\theta')]P(\theta')M(\theta')d\theta'+B,\\
P(\theta) &= \frac{1}{1+\tau_d U M(\theta)}.
\end{aligned}\tag{2.10}$$

From the Fourier series expansion, Eq. (2.9), we can see that the only possible solution for $M(\theta)$ is of the form $\alpha+\beta\cos(2\theta)+\gamma\sin(2\theta)$. Without loss of generality, we assume $M(\theta)=M_0+M_1\cos(2\theta)$, with $M_0,M_1\geq 0$. In order to ensure

that the steady state solution is in the broad profile regime, we additionally require $M(\theta) > 0, \forall \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, so $M_1 < M_0$. Using numerical simulations we found that inhomogeneous solutions, ($M_1 > 0$) where they exist, were unstable, although this is analytically intractable. Here, we only consider the homogeneous case, which is given by

$$\begin{aligned} M_0^{\pm} &= \frac{1}{2\tau_d U}\left(\tau_d UB + J_0 - 1 \pm \sqrt{(\tau_d UB + J_0 - 1)^2 + 4\tau_d UB}\right), \\ P_0^{\pm} &= \frac{1}{1+\tau_d U M_0^{\pm}}. \end{aligned} \tag{2.11}$$

Since we require that $M_0^{\pm} > 0$, we have that for supra - threshold input $B > 0$, only M_0^+ is a possible homogeneous steady state solution. Unless otherwise noted, M_0 always refers to M_0^+ (the negative root, where needed, is referred to as M_0^-). The existence of M_0^- as a valid stationary state requires $J_0 > 1$ and $B \leq 0$, although these conditions are not sufficient to ensure the stability of such a state. Finally note that for zero average synaptic connectivity, $J_0 = 0$, one has $M_0 = B$.

To examine the stability of the steady state solution, we linearise the Fourier components described in Eq. (2.9) about the fixed point M_0 and P_0 from Eq. (2.11), and consider small deviations away from the fixed point, in Fourier space, in order to assess whether or not the steady state is stable. We write the Fourier components as $\hat{m}_n(t) = m_n(t) + M_n$ and $\hat{p}_n(t) = p_n(t) + P_n$, where M_0, P_0 are the steady state values, and $M_n = P_n = 0, \forall n \neq 0$. We linearise the equations about the steady state solution (M_0, P_0), neglecting higher order terms to obtain

$$\begin{aligned} \tau_0 m_0'(t) &= m_0(t)\left[-1 + J_0 P_0\right] + J_0 p_0(t) M_0, \\ \tau_0 m_{\pm 1}'(t) &= m_{\pm 1}(t)\left[-1 + \frac{J_2}{2}P\right] + \frac{J_2}{2} p_{\pm 1}(t) M_0, \\ \tau_0 m_n'(t) &= -m_n(t),\ |n| > 1, \\ \tau_d p_n'(t) &= -p_n(t)\left[1 + \tau_d U M_0\right] - \tau_d U m_n(t) P_0. \end{aligned} \tag{2.12}$$

This is a particularly advantageous form, since for any n the time evolution of each pair of equations $m_n(t)$ and $p_n(t)$ decouple from all other pairs of equations. For the steady state solution (M_0, P_0) to be stable, small perturbations m_n and p_n must tend to zero as $t \to \infty$ for all n. We consider first the equations for $|n| > 1$, namely:

$$\begin{aligned} \tau_0 m_n'(t) &= -m_n(t), \\ \tau_d p_n'(t) &= -p_n(t)\left[1 + \tau_d U M_0\right] - \tau_d U m_n(t) P_0. \end{aligned}$$

Clearly, all small perturbations $m_n \to 0$ as $t \to \infty$. This also means that $p_n \to 0$ as $t \to \infty$ because $M_0 > -\frac{1}{\tau_d U}$ and $P_0 > 0$. Next, we consider the cases where $|n| \leq 1$.

For $n = 0$, we have the following equation, written in matrix form

$$\frac{d}{dt}\begin{bmatrix} m_0(t) \\ p_0(t) \end{bmatrix} = D\begin{bmatrix} m_0(t) \\ p_0(t) \end{bmatrix} := \begin{bmatrix} -\frac{1}{\tau_0} + \frac{1}{\tau_0}J_0P_0 & \frac{1}{\tau_0}J_0M_0 \\ -UP_0 & -\frac{1}{\tau_d} - UM_0 \end{bmatrix} \begin{bmatrix} m_0(t) \\ p_0(t) \end{bmatrix}.$$

The eigenvalues λ of the matrix D determine whether the solution is stable. If the two solutions for λ are negative, then perturbations m_0 and p_0 will die out, i.e. the steady state solutions will be stable to homogeneous fluctuations in firing rate and release probability. For $B > 0$, the steady state solution (M_0, P_0) is always stable to homogeneous fluctuations in firing rate and release probability, and so does not undergo amplitude instability.

Next, we consider the stability to perturbations m_1 and p_1 (the stability to perturbations $m_{-1}$and p_{-1} being identical). We call these perturbations "transverse fluctuations" in firing rate and release probability (Hansel and Sompolinsky, 1998), and, following the treatment for homogeneous fluctuations, we need to find the eigenvalues λ of the matrix

$$\begin{bmatrix} -\frac{1}{\tau_0} + \frac{1}{2\tau_0}J_2P_0 & \frac{1}{2\tau_0}J_2M_0 \\ -UP_0 & -\frac{1}{\tau_d} - UM_0 \end{bmatrix}.$$

This equation includes dependence on J_0 due to the inclusion of M_0 and P_0, and so is more complicated to analyse the stability, since it cannot be done in isolation of J_0. If we define $\mu = 1 + \tau_d U M_0$, then the values for the eigenvalues λ are given by

$$\lambda = \frac{\left[\tau_d(\frac{J_2}{2} - \mu) - \tau_0\mu^2\right] \pm \sqrt{\left[\tau_d(\frac{J_2}{2} - \mu) - \tau_0\mu^2\right]^2 + 4\tau_0\tau_d\mu(\frac{J_2}{2} - \mu^2)}}{2\tau_0\tau_d\mu}. \tag{2.13}$$

The region where the largest eigenvalue is less than zero, is given by satisfying one of two conditions. Either we require that $J_2 < 2\mu$, or we require that $2\mu < J_2 < 2\mu^2$ and $\tau_d(J_2 - 2\mu) < 2\tau_0\mu^2$. Since we are assuming that $\frac{\tau_d}{\tau_0} > 1$, this leads to the following conditions for the steady state solution (M_0, P_0) being stable to transverse fluctuations in the activity and release probability:

$$0 < M_0 \leq \frac{\tau_0}{\tau_d U(\tau_d - \tau_0)} \text{ AND } J_2 < 2\mu^2,$$

or,

$$M_0 > \frac{\tau_0}{\tau_d U(\tau_d - \tau_0)} \quad \text{AND} \quad J_2 < 2\mu + \frac{2\tau_0\mu^2}{\tau_d}. \tag{2.14}$$

We used these equations to draw the boundary line to the region of stable homogeneous solutions in the phase portrait in Fig. 2.2.

Eqs. (2.12) are only a good approximation when terms of size $O(m_n(t)p_j(t))$ can be neglected which requires, as much as possible, for the rectification to have minimal effect. In network simulations, these terms are small when the simulation operates in the broadly tuned regime - i.e. where the rectification has little effect. Thus when the real part of the eigenvalues from Eq. (2.13) is zero, $|m_{\pm 1}|, |p_{\pm 1}|$ remain unchanged, and the rotation speed is given by the imaginary part of Eq. (2.13).

2.A.2 Stability of the stationary bump

Another stationary state is the bump state, Fig. 2.1B. Following previous treatments (Hansel and Sompolinsky, 1998), we define its width θ_c to be the angle such that, $M(\theta) = 0$ for all $\theta > \theta_c$ (assuming, without loss of generality, that the profile is centred on $\theta_0 = 0$). We then can write the stationary state solution as

$$\begin{aligned} M(\theta) &:= A\left[\cos(2\theta) - \cos(2\theta_c)\right]_+ = \left[\frac{1}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} [J_0 + J_2 \cos 2(\theta - \theta')] P(\theta') M(\theta') d\theta' + B\right]_+, \\ P(\theta) &= \frac{1}{1 + \tau_d U A \left[\cos(2\theta) - \cos(2\theta_c)\right]_+}, \end{aligned}$$

where A and θ_c can be determined numerically. The stability condition of this state is intractable analytically, and so a numerical approach is used. Hereto we discretise the network into n neurons as

$$\begin{aligned} \tau_0 \frac{dm(i,t)}{dt} &= -m(i,t) + \left[\sum_{j=0}^{n} J_{ij} p(i,t) m(i,t) + B\right]_+, \ 1 \le i \le n, \\ \tau_d \frac{dp(i,t)}{dt} &= 1 - p(i,t) - \tau_d U p(i,t) m(i,t), \ 1 \le i \le n. \end{aligned}$$

Here, we have that $J_{ij} = J_0/n + (J_2/n)\cos((2\pi(i-j))/n)$, which is equivalent to the continuous formulation, with $m(i,t)$ and $p(i,t)$ defined on $\mathbb{N} \times \mathbb{R}$. Again, the steady state solution has the following form

$$\begin{aligned} M(i) &= A\left[\cos 2\left(\frac{2\pi i}{n} - \phi\right) - \cos 2\theta_c\right]_+, \\ P(i) &= \frac{1}{1 + \tau_d U M(i)}, \end{aligned}$$

where ϕ is the profile's centre. To examine the stability we consider the largest eigenvalue of the Jacobian of the system, evaluated at the stationary state. The Jacobian

matrix K is of size $2n \times 2n$, with matrix elements as follows, for $1 \leq i,k \leq n$,

$$\begin{aligned} K_{i,j} &= -\frac{1}{\tau_0}\delta_{i,j} + \frac{1}{\tau_0} H\left[\sum J_{ij}P(j)M(j) - B\right] J_{ij}P(j), \\ K_{i,j+n} &= \frac{1}{\tau_0} H\left[\sum J_{ij}P(j)M(j) - B\right] J_{ij}M(j), \\ K_{i+n,j} &= \delta_{i,j}\left(-\frac{1}{\tau_d} - UP(j)\right), \\ K_{i+n,j+n} &= -\delta_{i,j}UM(j), \end{aligned}$$

where $\delta_{i,j}$ is the Kronecker delta and H is the Heaviside step function. Unlike the continuous formulation, the stationary state is not fully rotationally symmetric, in particular, the stability of stationary states must be considered for $\phi \in [0, \frac{\pi}{n}]$. If the Jacobian has an eigenvalue less than zero for a specific ϕ, then the corresponding state is stable. The results are used to plot the parameter region where the bump profile is stable in Fig. 2.2 (green lines).

2.A.3 Subtractive depression approximation

At low activity levels, $p(\theta,t)$ is always close to one and does not vary greatly in comparison to its mean value. As a result we can approximate that the synaptic depression removes a fixed amount, rather than a fraction of synaptic resource. This 'subtractive depression' simplifies the governing equations. For ε small, we write the governing equations as follows,

$$\begin{aligned} \tau_0 \frac{\partial m(\theta,t)}{\partial t} &= -m(\theta,t) + \frac{1}{\pi}\left[\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} J(\theta-\theta')p(\theta',t)m(\theta',t)d\theta' + B\right]_+, \\ \tau_d \frac{\partial p(\theta,t)}{\partial t} &= 1 - p(\theta,t) - \varepsilon m(\theta,t). \end{aligned} \tag{2.15}$$

The steady state solution $P(\theta)$ now takes the form $P(\theta) = 1 - \varepsilon M(\theta)$. We note that as $M(\pm\theta_c) = 0$, we have that

$$M(\theta) = \frac{1}{\pi}\int_{-\theta_c}^{\theta_c} J(\theta-\theta')M(\theta')[1-\varepsilon M(\theta)]d\theta.$$

We use this to express $M'(\theta)$ as

$$M'(\theta) = \frac{1}{\pi}\int_{-\theta_c}^{\theta_c} J(\theta-\theta')M'(\theta')d\theta' - 2\varepsilon\frac{1}{\pi}\int_{-\theta_c}^{\theta_c} J(\theta-\theta')M(\theta')M'(\theta')d\theta'.$$

Following the example of Hansel and Sompolinsky (1998), we suppose that the activity and synaptic resource profiles keep their form, but are displaced by some small amount

$\delta\psi_m$ and $\delta\psi_P$ respectively, a so called transverse perturbation. We write:

$$\begin{aligned} M(\theta+\delta\psi_m(t)) &= M(\theta)+\delta\psi_m(t)M'(\theta)+O(\delta\psi_m^2(t)), \\ P(\theta+\delta\psi_P(t)) &= P(\theta)+\delta\psi_P(t)P'(\theta)+O(\delta\psi_P^2(t)). \end{aligned}$$

We want to consider how the small perturbations evolve over time. Neglecting terms of higher order, we can use Eq. (2.15) to write, for small $\delta\psi_m(t)$ and $\delta\psi_P(t)$, and for $|\theta|<\theta_c$ that

$$\begin{aligned} \tau_0\delta\psi'_M(t)M'(\theta) &= -\delta\psi_m(t)M'(\theta)+\delta\psi_m(t)\frac{1}{\pi}\int_{-\theta_c}^{\theta_c}J(\theta-\theta')P(\theta')M'(\theta')d\theta' \\ &\quad +\delta\psi_P(t)\frac{1}{\pi}\int_{-\theta_c}^{\theta_c}J(\theta-\theta')P'(\theta')M(\theta')d\theta', \\ \tau_d\delta\psi'_P(t)P'(\theta) &= -\delta\psi_P(t)P'(\theta)-\varepsilon\delta\psi_m(t)M'(\theta). \end{aligned}$$

Using $P(\theta)=1-\varepsilon M(\theta)$, we obtain

$$\begin{aligned} \tau_0\delta\psi'_M(t) &= \delta\psi_m(t)\left[-1+\frac{1}{M'(\theta)}x_1(\theta)-\varepsilon\frac{1}{M'(\theta)}x_2(\theta)\right]-\delta\psi_P(t)\varepsilon\frac{1}{M'(\theta)}x_2(\theta), \\ \tau_d\delta\psi'_P(t) &= \delta\psi_m(t)-\delta\psi_P(t), \end{aligned} \tag{2.16}$$

where $M'(\theta)\neq 0$, and

$$\begin{aligned} x_1(\theta) &= \frac{1}{\pi}\int_{-\theta_c}^{\theta_c}J(\theta-\theta')M'(\theta')d\theta', \\ x_2(\theta) &= \frac{1}{\pi}\int_{-\theta_c}^{\theta_c}J(\theta-\theta')M(\theta')M'(\theta')d\theta', \\ J(\theta) &= J_0+J_2\cos(2\theta), \text{ and} \\ M(\theta) &= A\left[\cos(2\theta)-\cos(2\theta_c)\right]_+. \end{aligned}$$

For $|\theta|<\theta_c$ we then have that

$$\begin{aligned} M'(\theta) &= -A\sin 2\theta, \\ x_1(\theta) &= \frac{J_2A(\sin 4\theta_c-4\theta_c)}{4\pi}\sin 2\theta, \\ x_2(\theta) &= \frac{J_2A^2(-24\theta_c\cos 2\theta_c+9\sin 2\theta_c+\sin 6\theta_c)}{24\pi}\sin 2\theta. \end{aligned}$$

Both $x_1(\theta)$ and $x_2(\theta)$ are proportional to $M'(\theta)$, leaving Eq. (2.16) independent of θ. We can examine the stability to transverse fluctuations of the system by writing

$$\frac{d}{dt}\begin{pmatrix}\delta\psi_m(t)\\ \delta\psi_P(t)\end{pmatrix}=D\begin{pmatrix}\delta\psi_m(t)\\ \delta\psi_P(t)\end{pmatrix},$$

where matrix D is evaluated at some $\theta \neq 0$, $|\theta| < \theta_c$ as

$$D = \begin{bmatrix} \frac{1}{\tau_0}\left(-1 + \frac{1}{M'(\theta)}x_1(\theta) - \varepsilon\frac{1}{M'(\theta)}x_2(\theta)\right) & \frac{1}{\tau_0}\left(\varepsilon\frac{1}{M'(\theta)}x_2(\theta)\right) \\ \frac{1}{\tau_d} & -\frac{1}{\tau_d} \end{bmatrix}. \tag{2.17}$$

The corresponding stability criterion is shown by the black line in Fig. 2.3A.

Chapter 3

Synaptic depression and asynchronously presented stimuli

Processing in the visual cortex can be quite non - linear; the response to two objects or object elements is quite often less than the sum of the responses to the individual objects, and a maximum function has in some cases been proposed to describe these competitive interactions. Recent data, obtained from presenting two different stimuli asynchronously, has emphasised that this competition has a temporal aspect as well, namely, that one stimulus presented at high contrast can suppress the response to another stimulus presented subsequently (Gawne, 2008). In the previous chapter, we detailed how a simple neural model, the ring model, is altered by the introduction of synaptic depression in its connections. In this chapter, we continue to investigate the impact of synaptic depression on the responses of networks, presenting another network that incorporates short - term synaptic depression. This model accounts for the data reported by Gawne (2008) whilst remaining consistent with older data and models. It also leads to strong predictions about the processing of rapid stimulus sequences. As in chapter 2, we emphasise the role that synaptic depression could play in the dynamics of neural activity.

3.1 Introduction

In visual processing information is transformed from local intensity information to high level concepts such as objects, texture, or motion. But although there is growing consensus about the representation of visual information at various stages in the cortex, the exact nature of the computation taking place between and within these stages is

less clear. In particular, to understand visual perception it is essential to know how localised, low - level information is integrated by higher regions of the visual system. In models of object recognition it has been realised that simple linear summation of local information is not sufficient as it does not lead to useful invariant representations because selectivity is washed out. Instead, the maximum operator - i.e. a response equal to the maximum of the inputs - has been proposed as a possible method of information integration (Riesenhuber and Poggio, 1999, 2002; Serre et al., 2007).

Neurophysiological evidence for a maximum - like computation has been gathered by comparing a neuron's response when two or more stimuli are presented simultaneously to the responses when the stimuli are presented alone. Some of the available data suggests that the responses to pairs/groups can be well approximated by the maximum of the responses to the individual stimuli, supporting the theoretical considerations (see Sato, 1989; Lampl et al., 2004; Gawne and Martin, 2002), although other work indicates that the mean of the individual responses is the better fit (see Miller et al., 1993; Zoccolan et al., 2005 and again Sato, 1989).

We study the question of how visual neurons integrate information in view of recent data obtained in area V4 (Gawne, 2008). This experiment used spatially separated stimuli, so that interaction before V4 were unlikely. The study compared the response of single stimuli to when both stimuli were presented together, however, in contrast to earlier studies (Sato, 1989; Lampl et al., 2004; Gawne and Martin, 2002) the presentation of each stimulus was asynchronous. The delay between the presentation onset times was systematically altered, as was the contrast of one of the stimuli (with the other always presented at full contrast). The data indicated that the response to the stimulus pair closely resembled the response to the stimulus with the shortest onset latency, while the response from the second stimulus was greatly reduced, almost as if it wasn't present. This is schematically illustrated in Fig. 3.1. The effect was especially prominent if the stimulus with the shortest onset latency had high contrast - a low contrast stimulus was less effective at suppressing the response to a later high contrast stimulus.

Similar behaviour has been observed in area STS, which is an area with object and face selective cells. STS neurons responding to an an effective image would show a suppression in their response if the presentation of the stimulus is immediately preceded by a mask (Perrett et al., 2009). This suppression, also referred to as forward masking, was found to be particularly pronounced when a briefly presented image (55 ms duration) immediately followed the presentation of another brief image as a mask

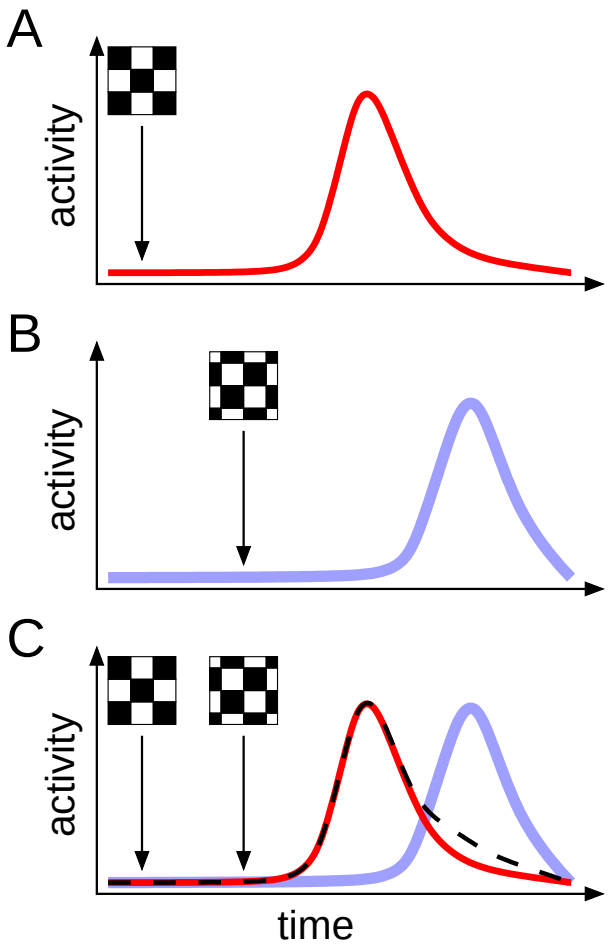

Figure 3.1: Schematic of the setup and result of the experiment in Gawne (2008). Two stimuli are presented separately (**A** and **B**, red and blue traces), or together with an arbitrary temporal delay (**C**). It was found that, in general, the response to both stimuli presented together (dashed black line) more closely resembles the response to the individual stimulus whose response has the shortest onset latency (red line), suppressing the response to the later arriving stimulus.

which is similar to that of the first (Perrett et al., 2009). This gave rise to an effect by which, when shown a sequence of images from a movie, neural responses would display anticipation, with peak responses coming *before* the image to which they are most strongly tuned is presented. Attention must also be drawn to repetition suppression in V1 (Nelson, 1991), in which a stimulus presented twice in quick succession can show a weakening of response to the second stimulus - this response suppression is stronger for higher contrast mask stimuli. In both these studies, however, and unlike the results reported by Gawne (2008), the stimuli are either presented in the same region of the visual field (Perrett et al., 2009) or were the same stimulus repeated (Nelson, 1991). This opens up the possibility that suppression in these situations is a feature which appears gradually over many layers through fatigue in upstream neurons.

These phenomena might be also related to the adaptation induced shifts in orientation tuning observed in V1. In these studies, V1 neurons presented with an 'adaptor' stimulus with orientation similar to that of their preferred orientation will see a repulsive tuning curve shift such that the preferred orientation moves away from the adaptor for a short period (Dragoi et al., 2002; Felsen et al., 2002) - although whilst the response to a stimulus of a preferred orientation can reduce the response to a subsequent stimulus of similar orientation, it appears that these are only relatively small reductions. This effect is also present in MT - for example in Kohn and Movshon (2004) the shift in direction tuning curves following prolonged adaptation is *towards* the adaptor stimulus, with the tuning curves narrowing, leading to much more dramatic reductions to some of the subsequent stimuli, although the adaptor stimulus is presented for a considerable time in these experiments.

Synaptic depression has been used model suppression before in an approach detailed in Carandini et al. (2002). This model is entirely feed - forward and features short - term synaptic depression in conjunction with a push - pull architecture (Hubel and Wiesel, 1962; Ferster and Miller, 2000). This model allows non - preferred stimuli to contribute to the depression of excitatory synaptic connections whilst the opponent inhibition prevents an overall activity contribution - such a non preferred stimulus will thus, when used in conjunction with a preferred stimulus suppress the overall response (cross - orientation suppression). This model also helps to account for the kind of forward suppression observed in Nelson (1991) - when stimuli at the preferred orientation are preceded by a mask stimulus, there is a response suppression which is strongest when the mask stimulus is presented at orientations close to the preferred orientation. Note however, the synapses in Carandini et al. (2002) are governed by a

form of depression which is very strong, and some evidence indicates that synapses in the cortex are weakly depressing when compared to previous investigations *in vitro* (Boudreau and Ferster, 2005; Reig et al., 2006). The model presented by Carandini et al. (2002) makes use of the idea that some synaptic pathways are *shared* to account for cross orientation suppression - we will further exploit this idea with our model.

Whilst we make use of shared feed - forward connections which are subject to short - term depression, we place a strong emphasis on convergence of inputs and shared recurrent connections to increase the suppression strength whilst the keeping the synapses more weakly depressing. The inclusion of recurrent connections subject to short - term synaptic depression is similar to a model introduced by van Rossum et al. (2008) to account for the dependence on contrast of the sizable response latencies that can be observed in deep layers of the cortex (Oram et al., 2002). In that study, each layer of a multi - layered network was subject to strong recurrent connections. A high contrast stimulus weakened the recurrent connections quickly, leading to a rapid response, whilst a low contrast stimulus weakened the recurrent connections slowly, leading to a slow response - we use these properties in the model we propose here.

In this chapter, we outline a model that accounts for the temporal interactions described in Gawne (2008) and Perrett et al. (2009) whilst still fitting with the spike count data of previous papers (Sato, 1989; Gawne and Martin, 2002; Lampl et al., 2004). We use a rate - based network model which features recurrent synaptic connectivity subject to short term synaptic depression (STD), which is believed to be present throughout the cortex (Wang et al., 2006; Thomson and Lamy, 2007). As we will show, the response to a high contrast stimulus suppresses the response to a second input, whilst the response to a low contrast stimulus leads to a much weaker suppressive effect. Moreover, we show that population of such networks can lead to stimulus anticipation.

3.2 Method

We model the activity of populations of neurons with similar tuning properties, such as a cortical column. Each population is represented by a unit which receives an average net current $I(t)$ at time t, and has an average instantaneous firing rate $r(t) = f(I(t))$ where f is the firing rate non - linearity. Throughout this work, we use the rectification non - linearity $f(x) = \max(x, 0)$, but similar results were obtained with sigmoidal firing rate non - linearities. The net current $I(t)$ that a unit receives obeys

$$\tau_0 \frac{dI(t)}{dt} = -I(t) + I_{input}(t) + I_{syn}(t).$$

In the absence of input, the current will decay exponentially to zero with time constant τ_0 - we use a time - constant of 5 ms to mimic the dynamics of fast excitatory AMPA synapses. The unit can receive input $I_{input}(t)$ from external sources and synaptic input $I_{syn}(t)$ from the other units.

Synaptic input from other units, $I_{syn}(t)$, is represented by the sum of component inputs, each the product of the synaptic weight W_i, the innervating unit's firing rate $r_i(t)$ and its associated synaptic resource $P_i(t)$ giving us the expression

$$I_{syn}(t) = \sum_{i=1}^{n} W_i P_i(t) r_i(t),$$

a term which can include recurrent connections. The synaptic connections between the units are subject to short term synaptic plasticity, which we model by varying the synaptic resource associated with each unit, $P_i(t)$. Following Tsodyks et al. (1998), and assuming Poisson firing statistics, we model the synaptic depression as

$$\tau_d \frac{dP(t)}{dt} = 1 - P(t) - \tau_d U P(t) r(t),$$

where the time - constant $\tau_d = 200$ ms describes how quickly the depression recovers to 1 in the absence of stimulation and the utilisation $U = 0.2$ controls how much of the available synaptic resource is used by each spike. These parameter values are at the low end of the range from the experimental literature (Abbott et al., 1997; Varela et al., 1997; Markram et al., 1998; Wang et al., 2006). This choice is deliberate, since some experimental evidence has indicated that synapses are less subject to synaptic depression in the presence of background activity (Boudreau and Ferster, 2005; Reig et al., 2006) and it is thus more challenging to account for a strong suppressive effect. The importance of the strength of depression is discussed in more detail in the results.

3.2.1 Shared depression network

The principal model that we investigate is depicted in Fig. 3.2A and called the *shared depression network*. The two *input units* (marked '1' and '2' in Fig. 3.2A) receive external inputs from input stimuli A and B, $I_A(t)$ and $I_B(t)$, respectively. When a stimuli is presented to the network, I_A or I_B jump to the desired amplitude for 50 ms. We use these short duration inputs to represent the transient part of neural responses, although

increased durations make little difference to our model as synaptic depression automatically leads to transient responses (Abbott et al., 1997; van Rossum et al., 2008). The maximum amplitude input we present is $I_A(t), I_B(t) = 200$, leading to peak firing rates in the input units of 200 Hz. These firing rates might seem high for cortical neurons but match those observed in V4 (Gawne and Martin, 2002; Gawne, 2008).

Apart from inhibiting each other, both input units innervate the *intermediate unit* (unit 3 in Fig. 3.2A). In addition, the intermediate unit receives recurrent, depressing synaptic input from its own output. The ratio between the strength of the feed - forward input and that of the recurrent connection is given by the parameter g_{rec}, which was set close to 1 to give recurrence and feed - forward connectivity approximately equal. The intermediate unit innervates the *output unit* (unit 4 in Fig. 3.2A), which also features recurrent connectivity. When measuring the *output* of a network, we are referring to the firing rate of the output unit.

In order to tune the network and its variants, we tune the weights to yield peak firing rates in the intermediate and output units of 200 Hz when an input of 200 amplitude is presented, matching peak firing rates observed in V4 (Gawne and Martin, 2002; Gawne, 2008). We also tune the strength of the mutual inhibition, such that when both inputs are presented simultaneously at the amplitude of 200 Hz lead to an output with a peak response of 200 Hz.

3.2.2 Network variants and spike frequency adaptation

We consider two different variant networks to examine different aspects of the proposed shared depression network. The first variant (Fig. 3.4A, top diagram) is a straightforward variant which removes both the synaptic depression and the recurrent connectivity by setting $g_{rec} = 0$ and fixing $P_i(t) = 1$ for all units. This necessitated re - tuning of the synaptic strengths. The second network variant (Fig. 3.4B, middle diagram) modifies the first variant by removing the mutual inhibition, and reintroducing synaptic depression. Again, the synaptic strengths of the network were re - tuned.

We also investigate a network incorporating spike frequency adaptation (Fig. 3.4C, bottom diagram). This network utilises non - depressing synaptic connectivity, and has units which are subject to spike frequency adaptation, with each receiving a negative feedback current $I_{a,i}(t)$ following the equation (from Hansel and Sompolinsky, 1998)

$$\tau_a \frac{dI_{a,i}(t)}{dt} = -I_{a,i}(t) + J_a r_i(t).$$

We use a fast timeconstant of 50 ms. We choose J_a equal to 0.2125 to ensure that the activity of a unit after a prolonged input is 87% of its peak response, in line with experimental evidence (Sanchez-Vives et al., 2000).

3.2.3 Population coding network

To investigate anticipation in population coding networks we construct a ring network with two layers, the intermediate layer and the output layer, each consisting n (= 31) units (Fig. 3.6A). The intermediate and output layers are arranged as in the shared depression model. We provide inputs to the network by current directly injected into the intermediate layer units in a sinusoidal fashion. The i'th intermediate unit receives current inputs of the form

$$I_{input}(t) = R(t)\left[1+\cos\left(\frac{2\pi i}{n}-\theta(t)\right)\right],$$

where $R(t)$ is the amplitude of the injected current, and $\theta(t)$ is the angular location of the input.

In order to investigate the network's anticipatory properties, we present it with a rotating input of constant amplitude and fixed rotation speed. We use the population vector method (Georgopoulos et al., 1986) to read out the position estimate from the output layer, measuring the average difference between the position estimate from the output layer, and the input position, over an extended stimulus presentation. We compare the anticipatory properties of this network with one where all synaptic connections are non - depressing.

3.3 Results

We aimed to find the simplest and most plausible network model that encompasses both the experimental data related to simultaneous stimulus presentation (Sato, 1989; Gawne and Martin, 2002; Lampl et al., 2004) as well as the data in Gawne (2008). In the context of this data we imagine that all units are neuronal populations in area V4. To model the cortical network we use a so called rate description, in which the activity of a population of neurons with similar tuning properties is described by their average firing rate.

The proposed model is depicted in Fig. 3.2A and we called it the *shared depression network*. We will first describe and examine the properties of the model before

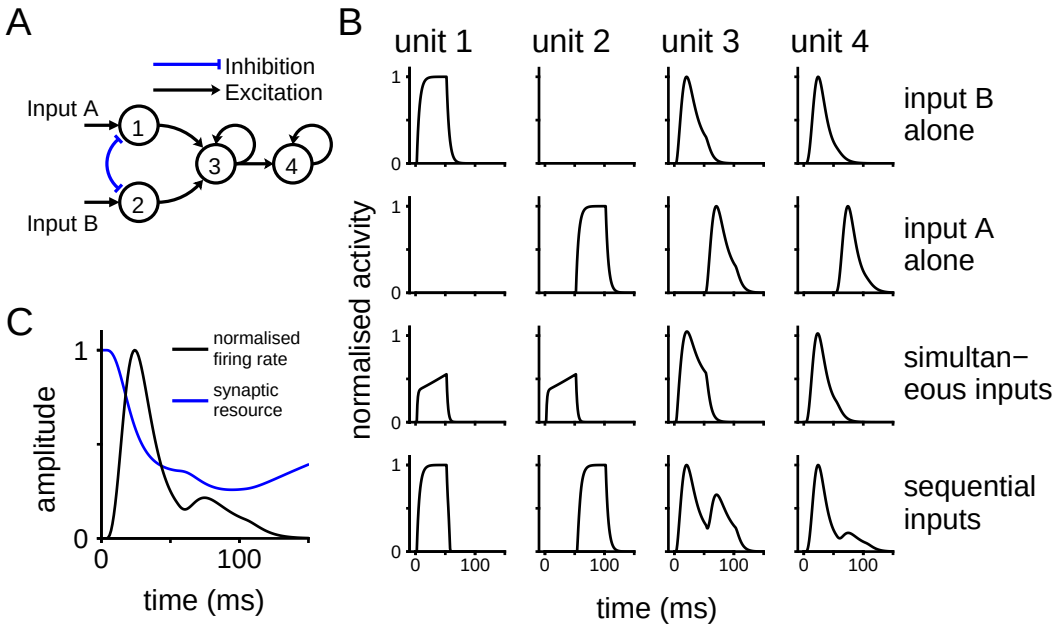

Figure 3.2: **A** The network diagram of the recurrent shared depression network. Each circle represents a unit. The input units, 1 and 2 receive the input to the model. These are then combined in an intermediate unit 3, which drives an output unit 4. With the exception of the recurrent inhibition between the input units, all connections are excitatory and subject to short term synaptic depression. **B** The network receiving two inputs, 200 Hz in amplitude, of 50 ms length. We show the firing rate of each unit when they are presented alone (top two rows) simultaneously (third row) and sequentially (bottom row). **C** The firing rate of the output unit (black dashed line) and the release probability of the synaptic input from the intermediate unit (blue trace).

comparing it less to satisfactory alternative models. The model has two input units, labelled 1 and 2, that receive input from a lower area, i.e. area V1. The input units mutually inhibit each other. Next, the two inputs are combined in an intermediate unit (unit 3 in Fig. 3.2A), and then passed to an output unit (unit 4), through a shared, depressing, excitatory connection. Both the intermediate unit and the output unit include recurrent, depressing, self - excitatory connections. There is strong recurrence, with the recurrent strength approximately as strong as the feed - forward inputs in both the intermediate and output units.

Fig. 3.2B demonstrates the shared depression network in action. Inputs presented alone (top two rows) cause identically shaped activity profiles in the output units. Note that is the signal propagates from the input unit to the output units, the activity profile becomes more transient, an effect of the synaptic depression. A similar effect is observed in neural recordings, with responses including a high amplitude transient, followed by a longer, lower amplitude, sustained part of the response.

When the inputs are presented simultaneously (third row), the input units compete with each other directly through their mutual inhibition - as a result the output unit implements an approximate maximum operation (see Yu et al., 2002 for similar methods to implement the maximum of simultaneous stimuli). However, when inputs are presented sequentially (bottom row), mutual inhibition cannot lead to competition in this network, since each input unit will inhibit the other only during its response - i.e. when the other input unit receives no stimulus input. In this case the competition between the stimuli is through synaptic depression. The response to the second input is weakened due to the depression of the synaptic connections that are shared by both stimuli. This occurs in two stages. Firstly, the recurrent connectivity of the intermediate unit depresses in response to the first stimulus, leading to a suppressed response to the second stimulus (Fig. 3.2B, bottom row, third column). This suppressive effect is magnified by the synaptic depression in the feed - forward connection from the intermediate unit to the output unit, and by the depression of the recurrent connectivity of the output unit. This leads to a much larger suppressive effect in the output unit (Fig. 3.2B, bottom right). Fig. 3.2C shows the activity of the output unit together with the synaptic resource of the intermediate unit, showing this explicitly.

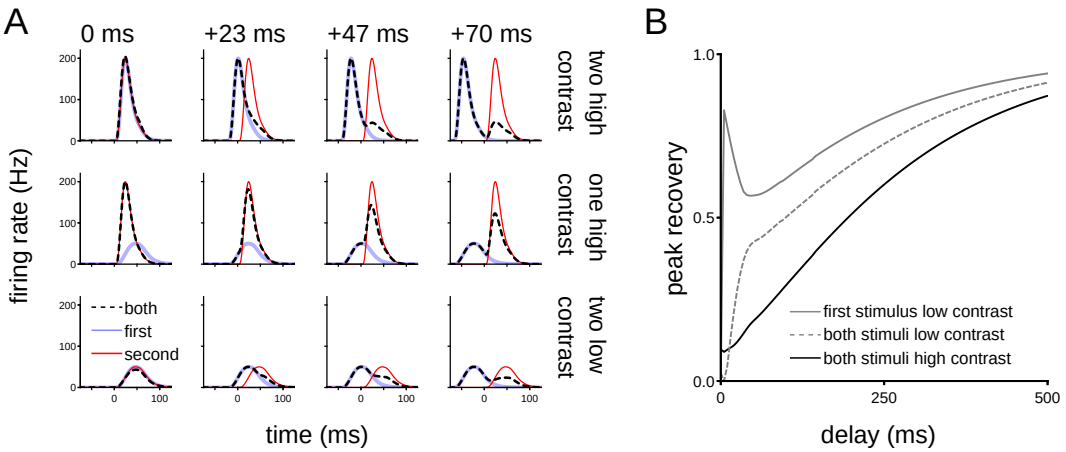

Figure 3.3: **A** The output when the model receives inputs from only one stimulus (red and blue traces) and both stimuli (dashed black trace), with various relative presentation onsets (left to right). In the top row we present two strong inputs, in the middle row we present one low contrast and one high contrast stimulus, and in the bottom row, we present two low low contrast stimuli. We find that a early high contrast stimulus can effectively suppress the response to a late stimulus (weak or strong), whereas an early low contrast stimulus is much less effective at suppressing a late arriving strong stimulus. **B** The relative peak response to a second stimulus after the first stimulus has been presented is plotted against the relative stimulus presentation latencies. The solid lines indicate that the first stimulus was a high contrast stimulus, with the second stimulus either high contrast (black) or low contrast (grey). The grey dashed line indicates that both stimuli are low contrast.

3.3.1 Response to asynchronous and mixed contrast stimuli

Fig. 3.3A shows the model's output when the presentation onset latencies of the two inputs to the model, and the contrast of those inputs are varied. In the top row, we present the input A in isolation, always with an onset latency of 0 ms, recording the output units firing rate (red line). Next, we present the input B in isolation with a range of relative onset latencies from -23 ms to $+70$ ms (blue line). These times were chosen according to the protocol used in Gawne (2008). Finally, we record the output when both inputs are presented together with the appropriate latencies (black dashed line). The response to the first input depresses the recurrent synaptic connections in the intermediate unit, the output unit, and the feed - forward connection from the intermediate unit to the output unit. This strongly reduces the response to the second input.

Fig. 3.3A, middle row - as top row, but one input is presented at low contrast, which we model by reducing the strength of the input presented in order to produce a peak output firing rate of 50 Hz. Now, if the high contrast input is presented first, it still suppresses the response to the low contrast stimulus. However, if the low contrast stimulus is presented first, the response does not depress the connections as much, and so the response to the strong input is only slightly reduced. Fig. 3.3A, bottom row - as top row, but now both inputs are low contrast. The result is that the suppressive effects are now weaker.

To examine the magnitude of the suppressive effect when the stimuli onsets are separated by longer time periods, Fig. 3.3B plots the peak response to stimulus B after stimulus A has already been presented. We normalise this by the peak response where stimulus A was not presented at all. To deal with overlapping stimuli presentations, we subtracted the response to stimulus A alone from the response to both stimuli, following the approach in Perrett et al. (2009). We consider the same stimuli contrast combinations as in Fig. 3.3A. When both stimuli are high contrast (black line, Fig. 3.3B), the suppressive effect is very strong. For instance, when stimulus B is presented 70 ms after stimulus A (traces plotted in Fig. 3.3A, top right), its peak firing rate is reduced to just 22% of its normal amplitude. Alternatively, a high contrast stimulus following an initial low contrast stimulus after the same delay, is only reduced to 58% of its normal amplitude (grey solid line Fig. 3.3B, traces plotted middle right Fig. 3.3A). The magnitude of the suppressive effect is not based just on the difference in contrast amplitudes between the two stimuli - when both stimuli are low contrast (grey dashed line Fig. 3.3B) the suppression is weaker than if both stimuli are high contrast. In such

a situation (traces plotted in Fig. 3.3A, bottom right), the peak response is reduced to 45% of its normal amplitude, a stronger relative response compared to where both stimuli are high contrast.

Two effects govern the magnitude of the suppressive effect. Firstly, whilst the first stimulus is being presented the magnitude of a response to a second stimulus is reduced by both the decreasing synaptic resource in the shared connections, and the inhibition from the first stimulus. In particular, the response to a low contrast stimulus reaches its peak activity later than that of a high contrast stimuli, and reaches its minimum synaptic resource later - a high contrast stimulus presented rapidly after a low contrast stimulus can avoid this, and respond more strongly than if it is presented slightly later (grey solid line Fig. 3.3B, short delays). In contrast, if both stimuli are low contrast, then the effect of the mutual inhibition when they are presented overlapping act to greatly increase suppression. When the stimuli are presented in a disjoint fashion, then the suppressive effect relies on the depressing synapses alone, and is much weaker. Secondly, the peak response recovers over time exponentially, in line with the exponential recovery of expended synaptic resource. This bears some similarity with the recovery of the magnitude of response suppression in STS to flashed stimuli following an (effective) masking stimulus (Perrett et al., 2009).

3.3.2 Model variants

The model which we propose, depicted in Fig. 3.2A, includes three synaptic pathways which are active when either input is presented - the recurrent excitation of the intermediate and output units as well as the feed - forward excitation from the intermediate unit to the output unit. Depression of these pathways during a response to an initial input causes suppressed responses to a subsequent input. The model also includes mutual inhibitory connections, to ensure that simultaneously presented inputs compete with each other. In this section, we examine importance of each of these components. Fig. 3.4 (top two rows) shows two variants of our model. We record the network responses to two strong inputs presented separately (red and blue traces) and together (black dashed trace) with increasing presentation synchrony, cf. Fig. 3.3.

In the first network variant (Fig. 3.4A, top) the synaptic connections are non - depressing, and corresponds to the simple case where the two inputs mutually inhibit each other and excite an intermediate and subsequently an output unit. This network acts like a point - wise max operator where the response to both inputs equals the

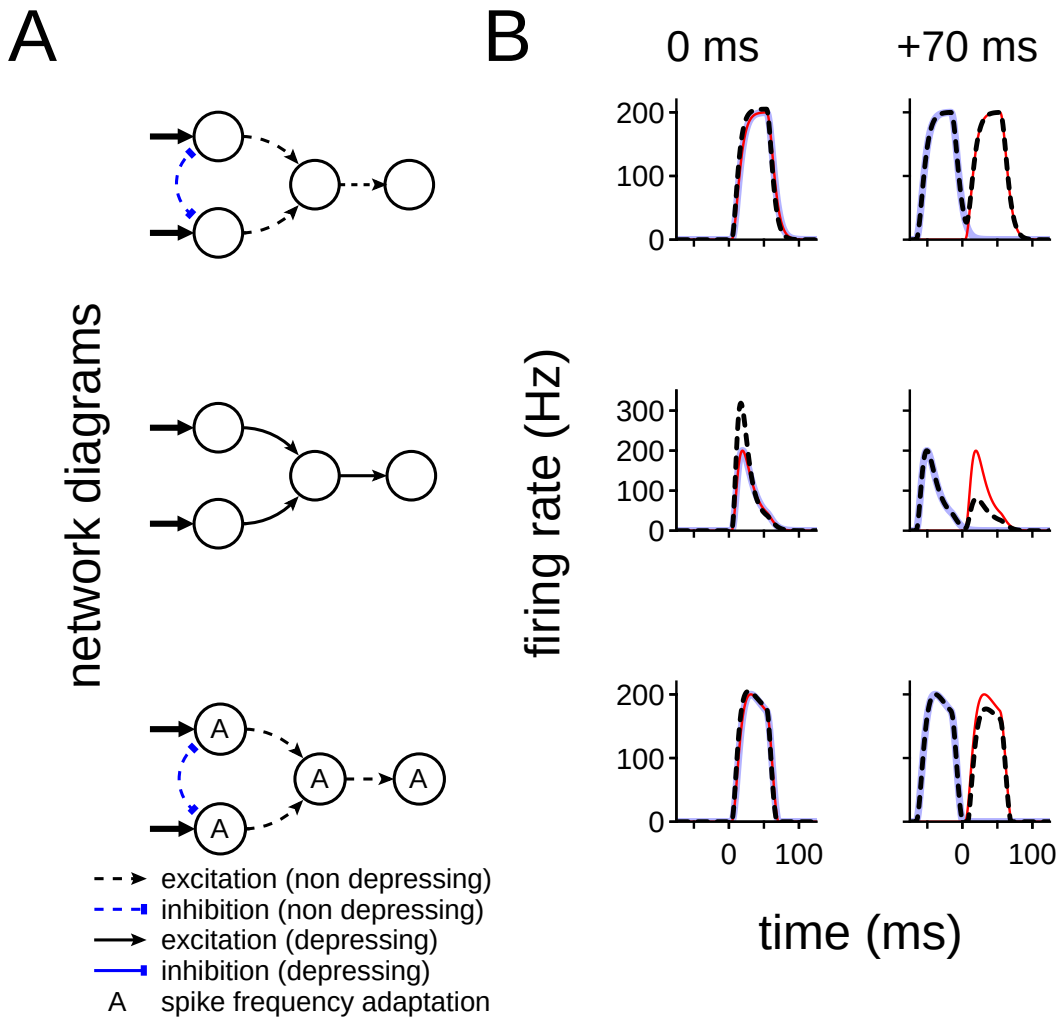

Figure 3.4: **A** Alternate networks, comparing how our model functions with various components removed. *Top row*, a basic network in which the input units inhibit each other but the synaptic pathways are not subject to short term depression. The input units are mutually inhibitory, and both excite an intermediate unit, which in turn excites an output unit. *Second row*, the network has no mutual inhibition between inputs, but incorporates synaptic depression in the connections. *Bottom row*, the network is connected as the top row, but all units are subject to spike frequency adaptation (SFA). **B** Traces of the firing rate in the final unit as the two inputs are presented at decreasing relative onsets (left to right) alone (red and blue), and together (black dashed trace) for the three different wiring diagrams.

maximum of the responses to the two inputs presented separately (Fig. 3.4B top, see Yu et al., 2002, for similar networks). Although the network response to simultaneously presented stimuli is a reasonable facsimile of the experimental data, it cannot reproduce the competition seen between temporally delayed stimuli, since the inhibition from an input unit is only effective when that unit is active. It should be noted that the activity of the intermediate and output units are almost identical - nothing is changed by the addition of another unit.

The second network variant (Fig. 3.4A middle) eschews both mutual inhibition between the input units, and the recurrent depressing connectivity of both the intermediate unit and the output unit. Now, when we present both inputs asynchronously (Fig. 3.4B, second row, right hand panels), the response of the intermediate unit to the first input acts to depress the synaptic connection innervating the output unit. This ensures that when the second stimulus is presented, the response is less than what it would have been when the stimulus is presented alone. However, when the inputs are presented simultaneously, the absence of inhibition between the inputs yields a supra - maximum response.

The amount of suppression the second network variant produces is limited by the amount of depression which the synaptic pathway can undergo. This is higher in the full model by the inclusion of recurrent connections, the depression of which greatly reduces their amplifying ability. Indeed, when only the feed - forward connections are present, the peak of the output response to a second high contrast stimulus presented 70 ms after the first is equal to 40% of the peak of the response when no prior stimulus is presented; whilst when recurrence is included, the peak of the response to a second stimulus is 22% of the peak response when no prior stimulus is presented. Alternate ways to increase the suppressive effect include increasing the utilisation proportion U and increasing the size of the recovery timeconstant τ_d. Overall then, the network requires both mutual inhibition and shared depressing synaptic connections in order to model the full phenomenon, however the specific form of the shared depressing connections is not necessarily fixed - with shared feed - forward and recurrent connections both contributing to the suppressing effect. The advantage of including recurrence is that in areas where peak firing rates are lower than those used in this chapter (in V4, the peak firing rate of approximately 200 Hz reported by Gawne, 2008 is much higher than those in STS - see, for example Perrett et al., 1982) then we need not increase U to physiologically unrealistic levels in order to model the same magnitude of suppression.

The last network variant we consider is one which incorporates spike - frequency

adaptation (SFA) in which the sustained firing of a unit leads to an inhibitory feedback (see Methods for implementation details). The network connectivity (Fig. 3.4A, bottom diagram) is as the first network variant above, but now all units are subject to SFA. We choose the parameters to ensure that the activity of a unit subjected to sustained input will reduce to 87% of its peak firing rate (Sanchez-Vives et al., 2000). With two high contrast stimuli, the response to the first stimulus suppresses the response to the second stimulus (Fig. 3.4B, bottom right panels), although only very weakly. Indeed using a much longer time - constant (5.7 s, from Sanchez-Vives et al., 2000, as opposed to the 50 ms we use here), the suppressive effect is virtually non - existent on these time scales. To get significant enough suppression, physiologically unrealistic levels of SFA must be used.

3.3.3 Relation to max model

Most experimental work investigating responses to pairs/groups of input stimuli has focused on coarser metrics of competition such as overall spike count (see, e.g. Sato, 1989; Lampl et al., 2004; Gawne and Martin, 2002). We test how well our model fits such data by presenting both stimuli simultaneously, varying the contrast of each, and comparing the result to when each stimulus is presented at the same contrast simultaneously. We compute the equivalent of a 'spike count' of the output unit as the area under the entire firing rate trace recorded.

We compare the spike count response of the output unit when presented with both inputs simultaneously, which we denote R_{A+B} with the spike count response of the output unit when the inputs are presented separately, which we denote R_A and R_B. In Fig. 3.5A, we plot the simulated spike count R_{A+B} against different predictions made from the simulated spike count responses to the individual inputs, R_A and R_B. The black dots graph R_{A+B} against the prediction from linear summation $R_A + R_B$, whilst the red triangles graphs R_{A+B} against the prediction made by the maximum operator, $MAX(R_A, R_B)$. We find, in accordance with Gawne and Martin (2002), that the maximum operator is a better predictor of response. This is expected, due to the inclusion of mutually inhibiting inputs, which will lead to a maximum type network - as outlined previously.

We also examined how our model performed using a metric that determines how effective a response is at suppressing a later response. This metric is given by the quantity

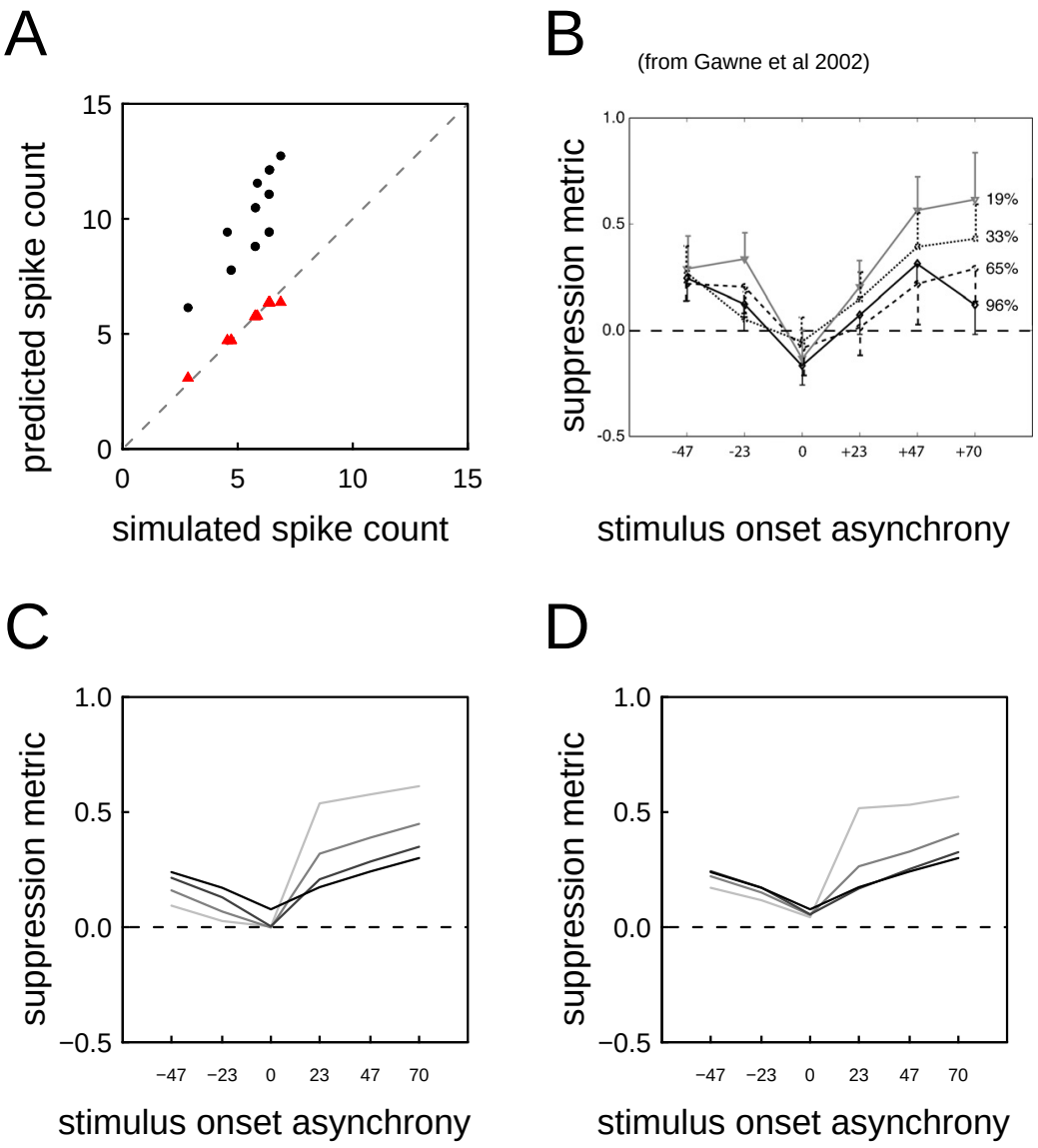

Figure 3.5: **A** The average firing rate over the whole trial when presented with two stimuli is plotted against the firing rate predicted from the responses to the stimuli when presented alone. The black dots predict the combined firing rate by summing the firing rates of the responses to the stimuli presented alone whilst the red triangles predict the combined firing rate by taking the maximum. The grey dashed line is the identity. **B** From Gawne (2008) - the response recorded from neurons in V4 to a second stimulus following the presentation of a first, divided by the response to a second stimulus presented alone. One of the stimuli is always high contrast. The line shading refers to the contrast of the other stimulus - lightest grey indicates lowest contrast, whilst darkest grey indicates that both stimuli are high contrast. **C** As B, but now our network is simulated, and the activity of the output unit is measured. **D** As C, but now the presentation onsets are altered to mimic the effects of contrast dependent latency by presenting the low contrast stimuli with additional delays.

$$\frac{R_{E+L} - R_E}{R_L},$$

where R_E is the response with the earlier latency, when measured alone, and R_L is the response with the longer latency, when measured alone, and R_{E+L} is the response to both stimuli being presented (from Gawne, 2008). We plot this metric against the stimulus presentation asynchrony in Fig. 3.5C. In cases where the low contrast response comes first, this metric gives a good match to the experimental data, with an initial low contrast stimulus much less effective at suppressing the response to a later high contrast stimulus than if the initial stimulus is presented at high contrast. Where the high contrast stimulus is presented first, there is a poor match to data (Fig. 3.5B, reproduced from Gawne, 2008 for comparison purposes). In particular, a high contrast input following a previous high contrast input will have a greater proportion of its response survive than if a low contrast stimulus follows a high contrast stimulus, whereas the data does not support this. This is due to weakening of the amplifying recurrent connections which has a larger effect on low contrast inputs than on high contrast inputs, which depend more strongly on the feed - forward pathways.

One possible explanation for this discrepancy is that our model fails to take into account the contrast dependent latency shifts observed experimentally (Gawne et al., 1996; Oram et al., 2002). We try to take this into account by presenting lower contrast stimuli at greater latencies. We delay the lowest contrast inputs after an additional 20 ms delay, the second lowest contrast inputs after an additional 13 ms delay, the second highest contrast inputs after an additional 7 ms delay and the highest contrast inputs at no additional delay. Fig. 3.5D displays the resultant metric against presentation onset asynchrony. Essentially, the additional length of time to allow the depressing synapses to recover more before a low contrast stimulus arrives, leading to closer values of the metric when the high contrast stimulus is presented first.

3.3.4 Anticipation in population model

The model explains existing experimental data obtained when two distinct stimuli are presented asynchronously. However, our network can also account for situations involving temporal suppression. Recent work recorded the responses of neurons in higher visual areas which are sensitive to complex objects when those objects are presented as part of a biologically plausible movie (such as a movie showing a head rotation, or hand movements). When a neurons response to the images presented as a

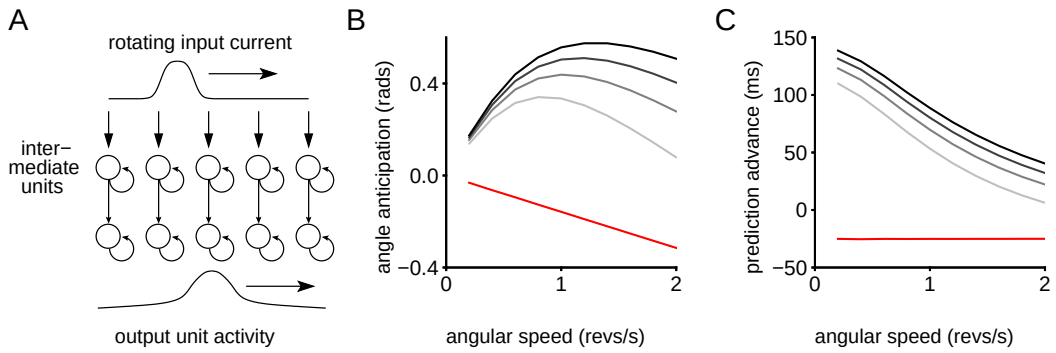

Figure 3.6: **A** The population network used to explain stimulus anticipation. Each intermediate unit receives an orientation - dependent injected current input, which in turns innervates an associated output unit. Apart from having inputs in common, each pair of intermediate and output units are disconnected from other pairs. We present a rotating input to the network at different speeds, and measure the output. **B** The difference in position (in radians) between the output activity and the input current for different contrasts (darker line indicate higher contrasts), a positive number indicates that the angle estimated from the output activity is in advance of the presentation angle (anticipation). The red line indicates the output from an identical network with non - depressing synapses. **C** As B, but we plot the prediction advance against the presentation speed.

movie is compared to its responses to the individual images presented separately, then its peak response occurs before the stimulus in the movie sequence which elicits the largest individual response is presented (Perrett et al., 2009). Thus, the maximally responding neuron from such a neural population could be one which is otherwise most responsive to an image which has yet to be presented, allowing the neural population to 'anticipate' future stimuli (Perrett et al., 2009). Various models have previously been proposed (van der Meer et al., 2007; Puccini et al., 2007) to explain anticipation. Whilst we do not specifically set out to model this phenomenon, it emerges naturally from our model.

We modified the model in order to accommodate a range of inputs parametrised by an angle. We use multiple intermediate and output units, arranged in a parallel populations. Each intermediate unit receives an injected current stimulus (Fig. 3.6A) of constant amplitude which rotates at a constant angular velocity for an extended period of time. By using a population vector readout we can determine the angle of the activity in the output layer, and compare it to the angle of the stimulus being presented. Fig. 3.6B graphs the difference between the angle of the output population vector and that of the input stimulus, against the presentation speed (measured in revolutions per

second) at different contrasts (black for high contrast, grey for lower contrasts). A positive value indicates anticipation - namely that the angle estimated from the output is ahead of that of the stimulus presentation. The depressing connectivity leads to a temporal 'biasing', with the output units responding more strongly to earlier inputs than to later ones. As a result the network responds in an anticipatory fashion - with maximal responses to a series of inputs arriving sooner than the most effective stimulus presentation. This effect is more pronounced at higher contrasts, leading to larger angle advances. Increasing angle advance comes naturally with increasing presentation speed, although this effect is limited - too fast a rotation speed ensures that not enough depression takes effect to sufficiently reduce the response to subsequent stimuli. Fig. 3.6C represent this data as a 'prediction advance' - the time lead or lag between input and output. We find that this increases with increasing contrast, and decreases with increasing presentation speed. We compare the anticipatory effects of this network with an identical network where the depressing synapses are replaced with non depressing synapses (red trace, Fig. 3.6B and 3.6C - for a network with non - depressing synapses all contrasts have identical behaviour). The output activity of such a network lags behind the inputs. In such a case, the recurrent architecture is actually disadvantageous in terms of anticipation.

3.3.5 Response latency and peak recovery

Our model assumes that the suppression of responses in a neural population to a second stimulus after the presentation of a first is due to synaptic depression in excitatory synaptic pathways which are active during the presentation of both stimuli. In particular, the output unit represents a population of neurons which receives input only from excitatory synapses which are utilised when either stimulus is presented. If depression in such shared synaptic pathways is the cause of an early, strong stimulus suppressing the response to a later stimulus, then we can use our model to make experimental predictions. We quantify the suppression in our network by measuring the peak response from the output unit in our network to stimulus B presented 70 ms after stimulus A (both at high contrast), normalised by the peak response to stimulus B presented alone, denoting such a quantity as the *peak recovery*. When the peak recovery is equal to 0, the response to stimulus B following the prior presentation of stimulus A peaks at zero - it is entirely suppressed. Meanwhile, a peak recovery equal to 1 indicates that the peak of the response to stimulus B is unaffected by prior presentation of stimulus

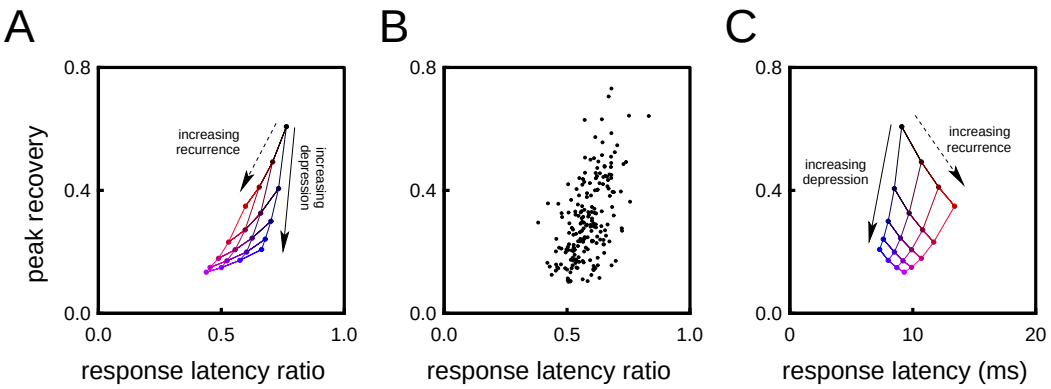

Figure 3.7: We examine the responses in the output unit of our network as we alter its parameters, measuring the *peak recovery* - the peak of the response to a high contrast stimulus following a stimulus presented 70 ms earlier as a proportion of the peak of the response to a stimulus in absence of a prior stimulus; and the *response latency ratio* - the (half - max) latency of the response to a high contrast stimulus divided by the latency of the response to a low contrast stimulus (both stimuli presented alone). **A** We plot the peak recovery against the the response latency ratio for response in the output unit of our network as we change the strength of the depression (the utilisation proportion U - blue dots indicate strong depression) and the strength of the recurrence in the intermediate and output units relative to the strength of the feed - forward connections (red dots indicate strong recurrence). **B** As A, but now the recurrence strength, synaptic depression strength and the depression recovery timeconstant are randomised for each unit - there remains a positive correlation between between the peak recovery and response latency ratio. **C** As A, but now we plot the peak recovery against the latency of the response to a high contrast stimulus presented alone - increasing the recurrence strength leads to increased latency whilst increasing the strength of the depression leads to decreased latency.

A - in such situations usually the entire firing rate trace is similarly unaffected by the presence of a prior stimulus. We also examine the half - max latency of the response to a single high contrast stimulus - when this is normalised by the latency of the response to a low contrast stimulus, we refer to this as the *response latency ratio*.

We focus first on the relationship between the peak recovery and the response latency ratio as different network parameters are varied. The shared depressing synaptic connections received by the output unit in our model have both feed - forward and recurrent components - we have found that increasing the strength of the recurrent connection relative to the strength of the feed - forward connection increases the magnitude of the suppression, as depressed recurrent connections are less able to amplify input signals. Previous work has also found that the latency of responses in recurrent networks featuring short - term synaptic depression depends greatly on the contrast of the input stimuli - and that for the most heavily recurrent networks, the difference between the latency of a response to a high contrast stimulus and the latency of the response to a low contrast stimulus is considerable (van Rossum et al., 2008). We illustrate this situation by plotting the peak recovery against the latency ratio (a measure of how much the latency depends on stimulus contrast) in Fig. 3.7A (red dots indicate networks with strong recurrence). Another parameter with a big impact on the magnitude of suppression is the strength of the synaptic depression - the utilisation proportion U. High values of U and stronger depression leads to strong suppression of subsequent responses - and so a lower peak recovery. We also find that low values of U leads to a increased response latency ratio - the latency of a response to high contrast stimulus is similar to the latency of the response to a low contrast stimulus when little depression occurs. This is also illustrated in Fig. 3.7A (blue dots indicate high values of U).

Overall then, when changing the relative strength of the recurrence or the strength of the synaptic depression, we find that the magnitude of the suppression (with low peak recovery indicating strong suppression) and the dependence on contrast (with low response latency ratio indicating strong contrast dependence) are correlated (Fig. 3.7A). Whilst we examine the strength of the recurrence and strength of the depression primarily - and assume that the ratio of recurrent input to feed - forward input is the same for the intermediate units and output units, this relationship persists when we randomise the relative strength of the recurrent strength and feed - forward strength separately for each unit, as well as randomising the synaptic depression utilisation proportion U and recovery timeconstant τ_d separately for unit (Fig. 3.7B - each point has random

parameters). This forms our prediction concerning experimental results, namely that variations of suppressive strength and the amount of contrast dependence in the response latency will be correlated on a cell by cell basis. The general framework which we outline predicts that the suppressive strength and the contrast dependence of the response latency will be correlated - although the contrast dependencies arising from prior processing layers might make this difficult to detect in the data.

We also examined how our model interacts with the absolute size of the response latency - as opposed to the situation above when we investigated what proportion of the response latency is influenced by stimulus contrast. As we have outlined, increasing either the relative strength of the recurrence or the strength of the synaptic depression increases the magnitude of the suppression, however the effect upon the response latency depends on which parameter is altered. If the relative strength recurrence is increased, the latency of the response to a high contrast stimulus is increased as the effective timeconstant of the intermediate and output units are increased (as with existing workvan Rossum et al. (2008)). In contrast, increasing the strength of the synaptic depression will decrease the latency of a response to a high contrast stimulus - the rapid depression ensuring that synaptic connections are strongest near the beginning of a response. This is illustrated in Fig. 3.7C. Such relationships can be used to make a qualified prediction as regards latency variations on a cell by cell basis *in vivo* - if the strengths of the synaptic depression of inputs to a cell vary little, or are bounded within a narrow range, we can predict that the strength of the suppression to a second input is correlated with the response latency.

3.4 Discussion

We argue that the combination of mutual inhibition and short term synaptic depression is a natural candidate to explain the observed neural responses of V4 neurons to two asynchronously presented stimuli (Gawne, 2008). A strong input to the shared depression model will depress the recurrent and feed - forward synapses that are active during the presentation of either stimulus - namely the connection from the intermediate unit to the output unit, and the recurrent connections amplifying the activity in both those units. Following an initial stimulus, the weakened synapses are less effective when a further stimulus is presented, which leads to a suppressed response to the second stimulus. If the first input presented is weak however, the network responds less vigorously, and the synapses will depress to a much lesser extent. This ensures that the response to

a second stimulus is not as strongly suppressed. For simultaneously presented stimuli, the mutual inhibition between input units leads to a maximum effect, but we should note that the results we present do not depend strongly on this - other suitable circuits can account for this, such as those featuring feed - forward inhibition suggested by Yu et al. (2002) (which has been observed experimentally, e.g. in Pouille et al., 2009), provided that the final stages of the network feature shared depressing pathways[1].

Due to the strong suppressive effect, even for brief stimuli, the model can also explain the data in STS relating to suppression of a briefly flashed image by a prior, equally brief, masking image (Perrett et al., 2009). However it should be noted that the work by Perrett et al. (2009) involves two similar stimuli presented sequentially at the same location. In such a situation, interactions in previous processing stages can contribute to the suppression. Additionally, following the presentation of a dissimilar mask, or non - preferred stimulus (which when presented alone leads to a reduced amplitude response), the suppressive effect was found to be reduced (Perrett et al., 2009), in a similar fashion to the weaker suppression following a low contrast stimulus found in V4 (Gawne, 2008) and modelled here. Care should be taken with such comparisons however, because while the presentation of a low contrast stimulus and a non - preferred stimulus might both lead to reduced amplitude responses and weaker suppression, they are not necessarily the result of the same circuitry - in a similar fashion to how response latency is strongly dependent on contrast, but not stimulus preference despite both leading to a low firing rate (Oram et al., 2002).

We have also shown how the shared depressing network, when arranged in parallel with others of its type and receiving a population of inputs, can lead to anticipation. We used a rather simple population network, and it should be noted that including lateral connections can complicate the dynamics considerably (York and van Rossum, 2009). Nonetheless the network produces a strong anticipatory effect, with a predicted position temporally in advance of a rotating stimulus - when the depressing synaptic connections were replaced with connections that do not depress, the predicted position always lagged behind the position of the rotating stimulus. It is possible that recurrent

[1]There were four models presented by Yu et al. (2002) which involve a variety of different approaches to calculate the maximum of a range of inputs - however all models involve mechanisms that act rapidly based on current inputs. In the presence of no input, they revert to zero output networks with no memory of any previous stimuli. Whilst in Gawne (2008), inputs overlap considerably, this is not the case for the situation involving sequentially presented images reported in Perrett et al. (2009) are so it is unlikely that any of these networks are suitable for modelling the suppression of inputs presented temporally - we further predict (see below) that the results in Gawne (2008) could be replicated by using flashed stimuli which do no overlap temporally.

connections could play a larger role in how this arises. For instance, it has been widely reported that neurons in V1, when presented with an orientation close to their preferred value will adapt such that the tuning curve shifting away from the orientation of the first stimulus (Dragoi et al., 2002; Felsen et al., 2002) - such a shift might be helpful to explain the some of the suppression following a stimulus presentation that we model in this chapter. Whether this stimulus anticipation leads to any behavioural effects is not clear - it is entirely possible that any readout mechanism (such as can be said to exist) in the brain would take this into account. This could possibly investigated psycophysically - no such investigation has yet been done, to the best of our knowledge. A model incorporating synaptic depression has previously been used to explain phase advance in V1 - in which a high contrast grating will respond more strongly and with a larger phase advance than a low contrast grating (Carandini and Heeger, 1994; Chance et al., 1998), so perhaps our results regarding anticipation are not so surprising.

Given the limitations and variability in the data, it is in general very difficult to exclude alternative models in biology. We show that spike frequency adaptation at a physiologically reasonable amplitude cannot account for the magnitude of suppression observed in the experimental data of Gawne (2008); Perrett et al. (2009). We believe that our model is physiologically plausible. For one we require only two stages rather than an multilayer network - convergence into the intermediate unit (perhaps corresponding to pyramidal neurons in layer 4 receiving inputs from the prior cortical or thalamic area), and a feed - forward projection to the output unit, with recurrent connections in both layers. An alternative model (which could explain repetition suppression for example) could arise as a result of weak spike frequency adaptation applied at each stage of a multi - layered model - but this is this is unlikely to be the case in Gawne (2008) as the data was collected from spatially separate locations in order to minimise the effects of interactions in earlier stages of the visual system (Gawne and Martin, 2002). Finally, the use of only feed - forward depressing synapses to explain suppressive behaviour has been considered (similar in spirit to the model proposed by Carandini et al., 2002), but to ensure a strong suppressive effect, strong synaptic depression is required. Our recurrent model does not rely on very strong synaptic depression to the same extent - which is in line with evidence that indicates synaptic depression *in vivo* may be weak (Boudreau and Ferster, 2005; Reig et al., 2006) Furthermore, an approach dominated by feed - forward connections might not fit with other properties observed in neural responses - in particular, successive layers of networks featuring only feed - forward, depressing synapses will lead to the responses in

each layer becoming more transient. This is at odds with the neural responses recorded in STS, which are often be relatively sustained, and even display a certain amount of invariance to the length of stimulus presentation. This behaviour, and the long, contrast dependent response latencies also observed in STS, can however be well captured by a network featuring strong recurrent connectivity (van Rossum et al., 2008) such as the one we outline - the tendency of the feed - forward connections to bias early responses is balanced by the effect of the feedback from the recurrent connections.

Our model of synaptic plasticity included only depression and not facilitation. Additionally, only one component of depression was taken into account. It would be interesting to examine what the potential role for short term facilitation could be. In particular, although we account the peak advance when responding to image sequences reported by Perrett et al. (2009), all of the responses to sequences in our model are reduced compared to the corresponding flashed stimuli. This is not surprising because the effect entirely rises due to synaptic depression - a flashed stimulus will always be stronger due to no depression having yet occurred. In contrast Perrett et al. (2009) displays peak responses to sequences which can often be greater than the response to any stimulus presented alone. Facilitation might help to account for this, although the effect that it would have on the response following a low contrast stimulus would no doubt also change. Regarding our model of depression - whilst we only incorporate one timeconstant of depression and assume that all our models connections are identical, we doubt whether introducing heterogeneity would change the central results - that depression of shared feed - forward and recurrent synaptic connections can account for the latency selection of inputs.

Our model makes a number of explicit, testable predictions. Firstly given that the sustained responses to briefly flashed inputs are commonly found in higher visual areas (Keysers et al., 2005), we predict that the phenomenon investigated in Gawne (2008) can be replicated using short duration ‘flashed’ stimuli, where stimuli presented in a completely non - overlapping fashion will still display the suppression noted in asynchronously presented stimuli of long duration. Were this to be confirmed, it would also provide additional evidence that the sort of network described by Yu et al. (2002) is not sufficient as an explanation of this phenomena. Our model also predicts that the dynamics of the network recovery is governed by the recovery of the depressing resource, and so will occur with a similar timeconstant and that, due to the nonlinear fashion in which synaptic depression scales with response amplitude, a low contrast stimulus will be proportionally less suppressed by a preceding low contrast stimulus

than if both stimuli are high contrast. For the same reason, using our model for stimuli anticipation, we predict that as the stimulus contrast is reduced, so is the magnitude of the anticipatory effect.

We also examined the latency of responses to high and low contrast stimuli, predicting that, for a given neuron, the strength of the suppression of the response to a second stimulus (following the presentation of an initial stimulus) correlates with the strength of the contrast dependency of the response latency (the proportion of the response latency to a low contrast stimulus which arises from the presentation contrast). We also make a more qualified prediction that the suppression strength correlates with the absolute value of the response latency, although this depends on the strength of the synaptic depression being relatively fixed across neurons. Either prediction regarding latencies might be difficult to measure experimentally, since neither takes into account the latency effects arising from earlier network layers.

In this chapter, we investigated the effects of synaptic depression on a network receiving the asynchronous presentation of two stimuli, trying to account for data which indicates that a high contrast stimulus can effectively suppress the response to a subsequently presented (and different) stimulus. It has been proposed that since responses with short latencies are often stronger than responses with long latencies, such latency selection could be a method of implementing a maximum function (Gawne, 2008). One reason the maximum operator has been proposed as an important computation is as a possible method of information integration. For instance consider a group of inputs which are tuned to the location of an object at different points of a receptive field - a neuron with a response equal to the maximum of these inputs will respond to the presence of the object whilst remaining invariant to the object location. Additionally, it will still respond weakly when each of the inputs are only weakly active (Riesenhuber and Poggio, 1999, 2002). In the next chapter, we examine a much more abstract model and ask what other advantages the maximum operation and other methods of responding non - linearly to pairs of inputs.

3.A Appendix

3.A.1 Feed - forward inhibition

We usually use recurrent, depressing connections or strong synaptic depression to model how the response to a second stimulus is greatly reduced compared to the re-

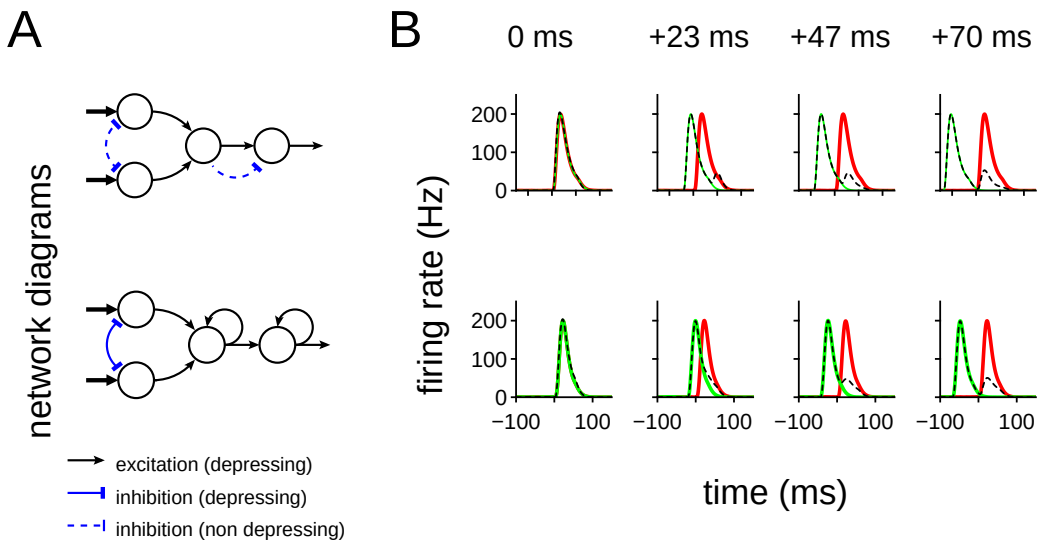

Figure 3.8: **A** *Top row*, the feed forward only architecture is supplemented by the addition of a non - depressing feed - forward inhibitory pathway, to counteract limit of the effect of depression. *Bottom row*, The full network. **B** Traces of the firing rate in the final unit as the two inputs are presented at decreasing relative onsets (left to right) alone (red and green), and together (black dashed trace) for the two different wiring diagrams.

sponse to the first stimulus. An alternate way of ensuring that there is a strong suppression of a subsequent stimulus is to incorporate feed - forward, non depressing, inhibition. In Fig. 3.8A, (top diagram) we include feed - forward inhibition, and change all the inhibitory connections to be non - depressing. To increase the amount of suppression of the second stimulus, this inhibition is tuned to counteract the fact that the depressing feed - forward pathway will not completely depress out, acting to cancel out the residual feed - forward excitation. The response to the second stimulus is now reduced to the same approximate level as the main model (3.8A and B, bottom row).

Chapter 4

Encoding pairs of stimuli

As we have discussed in chapter 3, the processing of multiple stimuli in the visual cortex can be quite nonlinear. In particular, studies have found that neural responses to pairs of stimuli are often less than the sum of the neural responses to each stimulus presented alone - often the response is more closely approximated by the maximum of the individual responses. In the previous chapter, we examined a specific neural model in which a strong response to a short latency input suppressed the response to another stimulus presented later, in line with recent experimental results (Gawne, 2008). This 'latency selection' of stimuli has been suggested as a possible method of ensuring that the response of a neuron to multiple stimuli follows the maximum of its inputs[1]. In this chapter we examine a more abstract model from a neural coding perspective. We will show that, under certain noise conditions, responding non linearly to pairs of stimuli can increase the coding capacity of a population compared to the situation when the response to a stimulus pair is created by a linear sum. We analyse the mathematics of the model to determine why responses to pairs of stimuli formed using the linear sum are vulnerable to classification errors. We also show that some of our assumptions can be weakened without changing the central thrust of our results.

4.1 Introduction

The study of neural coding investigates how information such as an external stimulus is encoded in neural responses. The number of different items that can be represented in a population of neurons is typically limited. For example, in absence of noise,

[1]This relies, to some extent, on high amplitude responses having short onset latencies (Gawne, 2008); note that the relationship between the latency and amplitude of responses is not always so straightforward - see e.g. Oram and Perrett (1992); Oram et al. (2002).

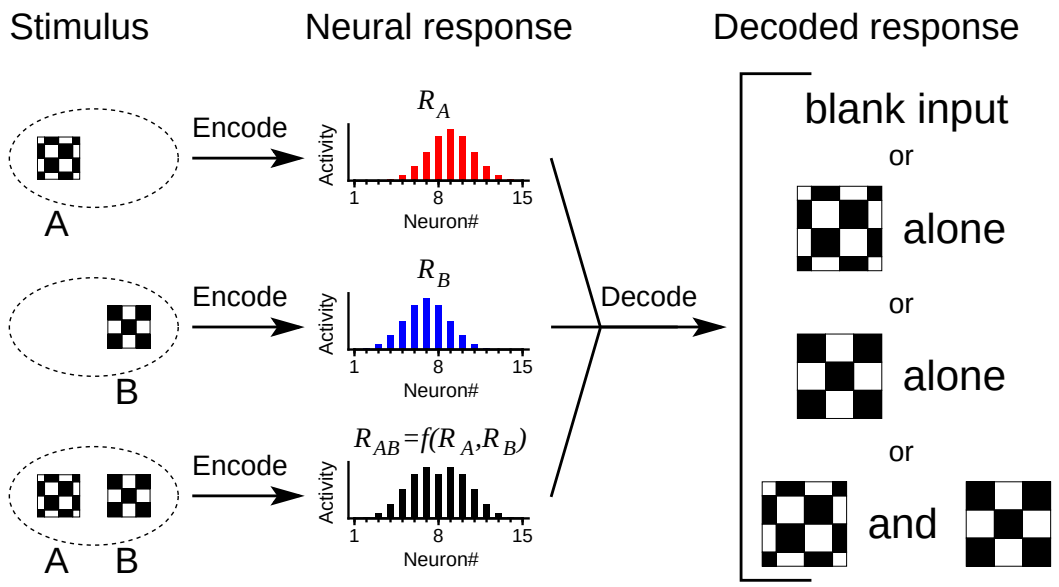

Figure 4.1: The setup for encoding and decoding neural activity. *Left column,* the visual field (dotted oval) containing one or two different stimuli (A and B, checker - board patterns). The stimuli are presented as inputs to the system, either separately, or simultaneously (denoted as stimulus pair AB), over many trials. *Central column,* the population response to each stimulus configuration. The mean responses to the single stimuli presented alone (R_A and R_B, with red and blue population responses respectively) are compared with the mean response to both stimuli presented together (R_{AB}, black responses). Throughout this work we assume that $R_{AB} = f(R_A, R_B)$ for some function f - which choice of function would best fit experimental data is still unclear although the maximum and average functions have both been proposed. *Right column,* the accuracy with which the population of responses are decoded is of paramount importance. When investigating the effectiveness of different types of combination function, we will use the error rate of various decoders as a performance measure.

N binary units can represent maximally 2^N distinct stimuli. When neural responses are continuous the number of distinct responses is often still limited because of the presence of noise. In general, if noise is additive, the efficiency of an encoding scheme is maximal when the encoded responses use the response space maximally (see, for instance, Ringach, 2010). This is complicated by the fact that neural populations must frequently respond to multiple distinct stimuli presented simultaneously. This raises the question of how multiple stimuli are be encoded into neural responses, how the chosen encoding method affects the decoding accuracy and thus the encoding capacity, and what type of errors can be expected when decoding.

Whilst the responses of neurons to single stimuli have been extensively investigated in the visual system, less attention has been paid to the responses of neurons to multiple stimuli presented simultaneously. However, experimental studies have compared the neural responses of single units presented with two (or more) stimuli presented simultaneously to the responses to the stimuli presented singly (Sato, 1989). The general approach (when pairs of stimuli are used) is illustrated in Fig. 4.1. There are two different stimuli, A and B, which we illustrate as checker - board patterns, presented at disjoint locations (similar stimuli have been noted to effectively drive neurons in V4 Gawne and Martin, 2002), giving three non trivial inputs - either A or B presented alone, or the stimulus pair AB, in which both stimuli are presented simultaneously.

Experimental evidence regarding how to best describe the response to stimulus pair AB, in terms of the responses to single stimuli A and B, is mixed. One of the earliest studies, by Sato (1989), examined the response in IT to two stimuli presented in different locations. The key finding was that the response to the pair of stimuli was less than the sum of the responses to each stimulus when presented alone. For some neurons tested, it was found that the response to the pair of stimuli was approximately equal to the maximum of the responses to each stimulus presented singly - whilst in other neurons the pair of stimuli when presented produced a response less than either response to one of the stimuli presented singly. The response to the pair of stimuli was not well matched by the linear sum of the responses to each stimulus separately - and other studies have also reached this conclusion. For complex cells in V1 and cells in IT, stimulated with stimuli in disjoint regions of the receptive field (oriented bars in V1 or complex patterns or objects in V4), some of the evidence suggests that the mean response to a stimulus pair AB is best approximated by the maximum of the mean responses to either stimulus presented alone (Gawne and Martin, 2002; Lampl et al., 2004).

Other experimental work suggests that a suppressive effect is more important, in which the responses to pairs of stimuli are less than the largest response to the component stimuli presented alone. For example cells in V1 exhibit cross orientation suppression in which the response to a drifting grating of optimal orientation is reduced when another drifting grating of sub - optimal orientation (the mask) is presented simultaneously (Bonds, 1989; Bauman and Bonds, 1991). There are other sources of suppression in V1 - for instance that of surround suppression. It has been observed that the response to a drifting grating effective at driving a V1 neuron is reduced in the presence of a surround grating. This type of suppression is stronger when the surround grating is the same orientation as the central grating (Cavanaugh et al., 2002), with orthogonal masks having the largest suppressive effect - this indicates that the suppression is not wholly based on the amplitude of the response to the mask in question. Evidence for a suppressive effect has also been found in IT, including from recordings of the responses to triplets of stimuli (Miller et al., 1993; Zoccolan et al., 2005), and in MT, with the response of motion sensitive neurons to multiple stimuli with different directions of motion well approximated to the average of the responses to the component moving stimuli presented singly (van Wezel et al., 1996; Recanzone et al., 1997; Treue et al., 2000).

Suppression and possible response normalisation have long been important target for experimental and theoretical studies, with normalisation thought to enhance the ability of a group of neurons to make the most of their limited dynamic range and improve the accuracy of encoding that can be carried out with a limited number of neurons (Heeger, 1992; Carandini and Heeger, 1994; Carandini et al., 2002; Ringach, 2010). There is not necessarily a disconnect between experimental evidence indicating that the response to a pair of stimuli could be will fit by the maximum of the responses to the individual stimuli, and those studies which indicate that an average model is more appropriate - it is conceivable that both types of neuron exist, perhaps due to heterogeneity in the different properties of the neurons measured (Zoccolan et al., 2005), or perhaps due to a more fundamental split between types of neuron into the simple cell / complex cell paradigm (Lampl et al., 2004). Indeed, investigations of different methods of responding to multiple stimuli which have focused on the ‘maximum’ function have often done so is motivated by the suggestion that complex cells responding equally to the maximum of their inputs might account for the invariance observed in higher visual areas, whilst simple cells remain ‘feature detectors’ (Riesenhuber and Poggio, 1999; Serre et al., 2007).

Here, we approach the problem from a coding efficiency point of view, examining how accurately population responses to single stimuli and pairs of stimuli can be decoded. We will show, using an abstract model, that combining the responses using maximum and sub - maximum functions can often lead to an improvement in both coding capacity and error resistance when compared to the linear sum. This applies both in cases where the mean responses to single stimulus inputs are independent and randomly generated, with no additional assumptions about the underlying stimuli space, as well as cases in which the stimuli are tuned to some parameter, such as orientation. We will identify the key sources of error which contributes to this differential for both types of stimuli, providing an intuition for the results we find. We believe that focusing on the combination non - linearity is a novel approach, although existing work focusing on uncertainty, and in which multiple stimuli can be combined linearly exists (Zemel et al., 1998; Zemel and Dayan, 1999; Sahani and Dayan, 2003).

4.2 Method

Our goal is to measure the coding capacity of a model population of n units. We will do this by encoding input stimuli, attempting to read them out, and calculating the error rate. We first need to define how stimuli are encoded, which describes how the population responds when presented with a input. Secondly we need to set a decoding method which determines how responses to inputs are read out.

4.2.1 Population model

We write that the activity of the population in response to an input is a vector of continuous values r, such as the spike count over a small time period. We restrict ourselves to three classes of input - no stimuli (the zero input), a single stimulus, or a stimulus pair. Throughout this work, the mean response to a full contrast input, X, is called a *response vector* and is labelled R_X. The three classes of input are as follows

- $X = 0$, when no stimulus is present, called the zero input, with response vector $R_0 = 0$, indicating that the mean population response in such a situation is equal to its base rate (contrast is thus not relevant for such an input).

- $X = i$ when the input to the population is the i'th single stimulus, from m possible stimuli, with corresponding response vector R_i.

- $X = ij$ when the input to the population is a pair of stimuli, the i'th and j'th stimuli presented simultaneously, with corresponding response vector R_{ij} - any combination $i < j$ is allowed (R_{ij} is assumed equal to R_{ji}) so there are $m(m-1)/2$ such stimulus pairs.

Our noise model corrupts the response to an input with additive Gaussian noise (see below). As a result, the (Euclidean) length of the response vector associated with a given input directly relates to how accurate responses to such inputs can be decoded - in particular, our model predicts that a long response vector will produce responses easier to decode than a short response vector (see the decoder section below for further information). To investigate the impact of factors such as the shape of the response vectors rather than the length, we normalise all response vectors to have length 1, save for the zero response vector. Noise with $\sigma = 0.1$ is thus equivalent to a noise with standard deviation 10% of the length of the response vectors.

4.2.2 Encoding single stimuli

To choose the response vectors associated with single stimuli we follow one of two methods. The first method of generating response vectors we use is to generate them randomly for each single stimulus. We choose the components of each response vector independently from an identical exponential distribution (neurons in the visual system stimulated by physiological images have been observed to have an exponential firing rate distribution (Baddeley et al., 1997). As we then normalise the length of the response vectors to be length 1, we can assume, without loss of generality that the mean of the exponential distribution is $1/\sqrt{2n}$ - note that normalisation introduces a slight dependency between different components of the response vector. This type of 'unstructured' stimulus space ensures that minimum assumptions are made about the underlying stimuli and provides a good comparison to stimuli generated using a tuning curve (see later).

We sometimes wish to alter the sparseness of the response vectors of single stimuli - to do this we generate the response vector as before, but subtract a constant cutoff term from each component, rectifying any values less than zero. We then normalise the response vector as usual. When describing such stimuli, we do so using a common sparseness measure, equal to $\mathrm{E}[R]^2/\mathrm{E}[R^2]$ (Treves and Rolls, 1991; Rolls and Tovee, 1995), with high cutoff leading to many zero elements in the response vector and a low sparseness value. A negative cutoff value means that a positive term is added to each

response vector component, and so after normalisation the components are closer in size.

The second method we use to generate the response vectors associated with single stimuli is to assume that mean responses are tuned to some angular variable (such as the orientation of a presented bar, the direction of motion, or head rotation angle). Such tuning has been observed experimentally and neurons are often tuned to multiple variables simultaneously (Hubel and Wiesel, 1962; Maunsell and Essen, 1983; Perrett et al., 1991; Pasupathy and Connor, 2001), which we shall label *orientation* without loss of generality. The population units can be visualised as forming a ring - like structure, with each given a *preferred orientation*. Units respond strongly to inputs of single stimuli only when the presentation orientation is similar to its preferred orientation (we ignore bimodal distributions and such). Mathematically, after denoting that j'th unit's orientation preference $\phi_j = 2\pi j/n - \pi$ and stimulus k has orientation $\theta_k = 2\pi k/m - \pi$, $1 \leq j \leq n$, $1 \leq k \leq m$. We define the j'th element of the response vector R_k solely by the absolute angular difference $|\phi_j - \theta_k|$, and is equal to $F(|\phi_j - \theta_k|)$ for tuning curve F, which takes the shape of a circular normal distribution,

$$F(x) = \exp(\kappa \cos x). \tag{4.1}$$

Note that we interpret orientations, in the region $[-\pi, \pi]$, as distinct - for example as a direction of motion, or a head rotation angle, rather than orientations that range from 0 to π only. For the work here, we choose $\kappa = 2$. Note we normalise the lengths of the response vectors as usual.

4.2.3 Encoding stimulus pairs

As the purpose of this paper is to investigate the impact of different methods of responding to pairs of stimuli presented simultaneous, we must also generate the response vectors associated with inputs consisting of pairs of stimuli. Although, from a coding point of view, it would be best that the responses to stimulus pairs are as dissimilar as possible to the responses to single stimuli, in the real system, that might not always be desirable. In particular, if adding an object to a scene would completely change the population response, the decoding process would likely become more complex.

To formalise the coding of pairs, we introduce the *combination function* f, which describes the response vector of stimulus pairs in terms of the single stimulus response vector R_A and R_B.

- We assume that the interactions can be described for each unit independently (component - wise). For f mapping $\mathbb{R}^+ \cup \{0\} \times \mathbb{R}^+ \cup \{0\}$ onto $\mathbb{R}^+ \cup \{0\}$ we extend it a map $f : (\mathbb{R}^+ \cup \{0\})^n \times (\mathbb{R}^+ \cup \{0\})^n \to (\mathbb{R}^+ \cup \{0\})^n$ by writing the i'th component of $f(R_A, R_B)$ as follows

$$\{f(R_A, R_B)\}_i \quad = \quad f(\{R_A\}_i, \{R_B\}_i)$$

  In principle, coordinated changes in the code might exist and are even likely, however, there is no experimental data that accurately describes this, while at the same time analysis seems infeasible as the number of possible interactions explodes.

- We assume that the function f is symmetric, so $f(R_A, R_B) = f(R_B, R_A)$.

- Finally, we assume that $f(R_A, 0) = R_A$, imposing trivially that the single object response is preserved by the combination function.

A final condition is that the response vector R_{AB} is normalised in length, so we have $R_{AB} = f(R_A, R_B)/|f(R_A, R_B)|$. Since we corrupt responses with additive noise *after* the individual response vectors are created (see below for details), the length of the response vector determines how accurate responses can be decoded. Without normalisation, we find that a combination function which produces strong responses will be better than a combination function which produces weak responses. The larger the response, the more accurate the decoding can be, irrespective of the shape of the population response produced. The normalisation gets rid off this artefact.

For concreteness we focus one of three different forms - a *linear sum* in which $R_{AB} = R_A + R_B$, after normalisation equivalent to taking the average; a *maximum* $R_{AB} = \max(R_A, R_B)$; and a competitive function we denote as *suppression* in which $R_{AB} = \max(R_A, R_B) - \min(R_A, R_B)/2$, where a unit will respond most strongly to pairs where only one of the stimuli alone would lead to a strong response, whilst when both stimuli alone lead to a strong response, the stimulus pair presented together leads to a weaker response. Note that the normalisation of response vectors ensures that these combination functions actually stand for an equivalence class of functions, all proportional to each other.

4.2.4 Contrast effects and noise model

Importantly, we assume that stimuli can be presented at any contrast. We assume that the responses to stimuli are contrast invariant (Sclar and Freeman, 1982; Sclar et al.,

1990) and write that the mean response to stimulus X at contrast c is equal to $g(c)R_X$, where g (which maps $[0,1]$ to $[0,1]$ bijectivly, with $g(0) = 0$ and $g(1) = 1$) describes the effect of contrast scaling. For stimulus pairs we only consider situations where both stimuli are presented at the same contrast. These assumptions are necessary for our analytical and numerical calculations. For the simulations we typically assume that $g(c) = c$, but many aspects of the mathematics do not depend strictly on this choice.

The fact that stimuli can occur at different contrasts, severely complicates the decoding problem. If the stimuli were presented at a known contrast, the best combination function would be the sum (Mike Oram, personnel communication). In such a case the overall population activity can be used to estimate the total number of stimuli, reducing the possibility of any confusion between a single stimulus and a pair of stimuli. However, when the contrast is unknown, the overall population response size provides less information about the number of stimuli present and the shape of the response becomes more important.

Note that in the absence of noise, decoding responses is trivial (provided that the contrast is non - zero and that no two response vectors are equal). We assume that the responses to input stimuli are corrupted by additive, Gaussian noise. We can thus express the response to an input in the following form

$$r = g(c)R + \sigma\nu, \tag{4.2}$$

where ν is a random variable generated from an n - dimensional normal distribution with zero mean and a covariance matrix I. Note that the response vector R has typically has length 1 - see above. The length of the randomly generated response r is not itself normalised.

4.2.5 Decoding mechanisms

When assessing the performance of a particular method of encoding responses in a population, we present a large number of both single stimuli and stimulus pairs at a (uniformly distributed) random contrast, and then measure the percentage of errors made by applying our chosen decoding method - we call this the *error rate*. A related quantity is the *coding capacity* which we define as the maximum number single stimuli encoded in a population, which also encodes all possible stimulus pairs, that results in a pre - determined target classification error rate (set arbitrarily to 10%). When estimating the error rate, we present equal numbers of single stimuli as we do pairs.

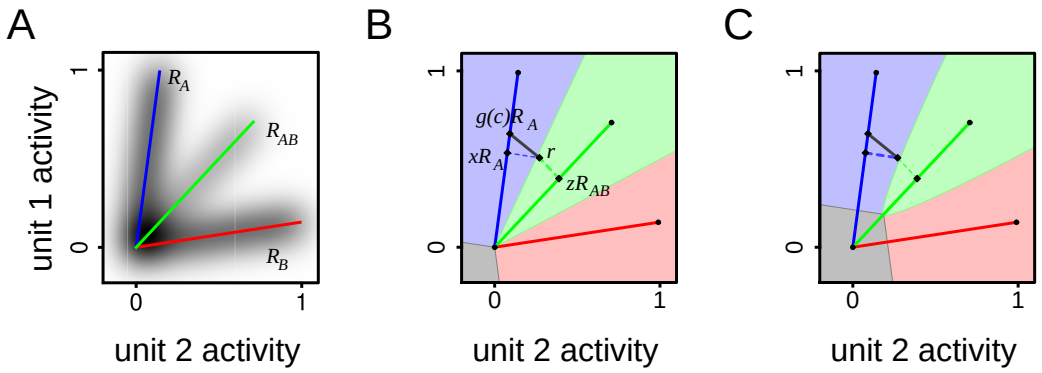

Figure 4.2: **A** Probability density function for the responses of a toy two - element population given that an input is presented - black indicates high density. The blue and red vectors are the response vectors R_A and R_B to inputs of a single one of the stimuli alone, whilst the green vector is the response vector R_{AB} to an input of both stimuli simultaneously. When a non - zero input is presented, the stimulus pair is presented one fifth of the time, with each single stimulus being presented two fifths of the time. **B** Stimulus A is presented alone, at contrast c and it generates a population response $r = g(c).R_A + \sigma\nu$, with the noise amplitude $\sigma = 0.1$. The ML decoder will classify it as being generated by the input whose response vector (R_A, R_B or R_{AB}) is closest - here AB. The contrast estimated is $g(c_{est}) = z$ where $(r - z.R_{AB})$ is orthogonal to R_{AB}. The shaded areas indicate the regions of phase space in which a response will be classified as having being generated by one single stimulus presented as an input (light blue and light red regions) or the pair of stimuli together (light green), or the zero input (grey). **C** The MAP decoder classifies the same response as being generated by the correct input - the MAP decoder weights against infrequently presented inputs, such as the stimulus pair AB. In this case, the single stimuli are each inputted alone with probability of $2/9$, the input is both with probability $1/9$ and the zero input is chosen with probability $4/9$ - chosen for illustrative reasons.

Although we do not present zero inputs, they are a possible, erroneous, decoding output. In cases where there is a random element in generating the responses to stimuli, we average the error rate over a range of random seeds making use of the resources provided by the Edinburgh Compute and Data Facility.

Fig. 4.2 displays the situation in a simple population consisting of only two units and encoding two stimuli, A and B. The blue and red vectors in Fig. 4.2A are the response vectors R_A and R_B to the inputs of each of the single stimuli, whilst the green vector is the response vector R_{AB} to an input of both stimuli simultaneously. The shading indicates the probability density function of the response, r, given that a non - zero input is presented - we present one of the stimuli A or B each two fifths of this time, with the pair AB presented one fifth of the time - the dark tint indicating the regions most likely for the response to be generated in. Fig. 4.2B and C show the two different decoders we can choose from in order to classify the responses, the operation of which we detail next.

4.2.6 Maximum likelihood decoder

The first readout method which we consider is the *maximum likelihood* decoder (ML decoder). For a recorded response r_{rec}, parameters x and c, and the conditional probability distribution $p(r = r_{rec}|x,c)$, the ML decoder tries to find the parameters which maximise the likelihood $p(r = r_{rec}|x,c)$ (Paradiso, 1988; Seung and Sompolinsky, 1993; Pouget et al., 2000). In our case, our two parameters are the identity of the input being presented, x, which can be the zero input, a single stimulus or a stimulus pair, and the contrast, c. The decoder then estimates the input x_{est} and the contrast $c_{est} \in [0,1]$ as follows

$$\{x_{est}, c_{est}\} = argmax_{x,c} p(r = r_{rec}|x,c) = argmax_x \left(p(r = r_{rec}|x, argmax_{c\in[0,1]} p(r = r_{rec}|x,c))\right). \tag{4.3}$$

Eq. (4.2) gives us the following form for the conditional probability distribution

$$p(r = r_{rec}|x,c) = \frac{1}{(2\pi)^{n/2}\sigma} \exp\left(\frac{-(g(c)R_x - r_{rec})^2}{2\sigma^2}\right), \tag{4.4}$$

where R_x is the response vector associated with input x. Using Eq. (4.4) and Eq. (4.3) we first estimate the contrast $c_{x,est}$ for a given input estimate x which maximises the likelihood of the response r_{rec}

$$c_{x,est} = argmax_{c\in[0,1]} p(r = r_{rec}|x,c) \quad = \quad argmin_{c\in[0,1]} (g(c)R_x - r_{rec})^2. \tag{4.5}$$

Eq. (4.5) indicates that for a given input x, the contrast which maximises the likelihood is that which minimises the distance between $g(c)R_x$ and r_{rec}. We can find $g(c_{x,est})$ by differentiating $(g(c)R_x - r_{rec})^2$ with respect to $g(c)$. The resulting value gives us the following formula, for $R(x) \neq 0$

$$g(c_{x,est}) = \frac{R_x.r_{rec}}{R_x^2}. \tag{4.6}$$

This allows us to determine $c_{x,est}$ since g is bijective and continuous (and hence always has an inverse on $[0,1]$, although in practice we only need to determine $g(c_{x,est})$, since we are interested primarily in the identity of the input). If Eq. (4.6) produces a value outside of $[0,1]$ then there is no value in $[0,1]$ which is a local minimum of $(g(c)R_x - r_{rec})^2 = g(c)^2R_x^2 - 2g(c)R_x.r_{rec} + r_{rec}^2$ from Eq. (4.6) and so $c_{x,est}$ is equal either to 0 or to 1. In such cases, if $R_x.r_{rec} < 0$, then the smallest possible value of the quantity minimised in Eq. (4.5) is r_{rec}^2, giving us $c_{x,est} = 0$. On the other hand, if $R_x.r_{rec} \geq 0$, then the smallest possible value of the quantity minimised in Eq. (4.6) is $(R_x - r_{rec})^2$, giving us $c_{x,est} = 1$. This gives us the following estimates for contrast $c_{x,est}$:

$$g(c_{x,est}) = \max\left(\min\left(\frac{R_x.r_{rec}}{R_x^2}, 0\right), 1\right). \tag{4.7}$$

We can now return to Eq. (4.5) to find our parameter estimate for x_{est} and c_{est}:

$$\begin{aligned}
\{x_{est}, c_{est}\} &= \{argmin_x(R_x(R_x.r_{rec})/R_x^2 - r_{rec})^2, c_{x_{est},est}\} \\
&= \{argmin_x((R_x.r_{rec})^2/R_x^2 - 2(R_x.r_{rec})^2/R_x^2 + r_{rec}^2), c_{x_{est},est}\} \\
&= \{argmax_x((R_x.r_{rec})^2/R_x^2), c_{x_{est},est}\}.
\end{aligned}$$

Using Eq. (4.7), we determine $g(c_{x,est})$ for each possible input x. If $c_{x,est} = 0$ for all x, then clearly $c_{est} = 0$, and x_{est} is indeterminate. In this case, we set $x_{est} = 0$ to represent the zero input since there is no specific non - zero input x which maximises the likelihood (as they all have the same estimate of zero contrast). The choice to classify the decoder output as $x_{est} = 0$ can be seen as a 'default' decoding output. On the other hand, if there exists at least one non - zero contrast estimate $g(c_{x,est})$, then a definite estimate x_{est} can be found.

When the response vector lengths are the same length (unless otherwise stated we normalise response vectors from both single stimuli and stimulus pairs to have length 1 - see above), we can write the input estimate as follows

$$\begin{aligned}
x_{est} &= argmax_x((R_x.r_{rec})^2/(R_x^2 r_{rec}^2)) \\
&= argmax_x \cos\theta_x = argmin_x|\theta_x|,
\end{aligned}$$

where θ_x is the angle between the response vector R_x and the recorded response r_{rec}. This leads to a simple geometric interpretation of the ML decoder - the maximum likelihood input x is that which minimises the angle between the recorded response r_{rec} and the response vector R_x, with the maximum likelihood contrast estimate equal to the length of the projection of r_{rec} in the direction of $R_{x_{est}}$ normalised by $|R_{x_{est}}|$ (bounded below by 0 and above by 1). As before, if the maximum likelihood contrast estimate is zero, we set the output of the of the decoder to be the zero input - this occurs when the size of the angle $|\theta_x|$ is greater than 90 degrees for all stimuli/stimulus pairs x.

Fig. 4.2B shows the ML decoder in action in a toy example with our two - unit population. The response to the single stimulus A, $r = g(c)R_A + \sigma\nu$ must be classified. The ML decoder compares the distance between the nearest point on the vector R_A (denoted xR_A) and the response r with the distance between r and the other input's response vectors. In this case, the point zR_{AB} on the line R_{AB} is closer to r than xR_A is so the ML classifier estimates that the stimulus pair AB generated r. Responses in the blue region are classified as being generated by single stimulus A, whilst those in the red or green regions are classified as having been generated by single stimulus B or stimulus pair AB respectively. Responses which lie in the grey region are classified as being generated by the zero input - no potential non - zero input has an estimated contrast greater than zero to responses in this region. Note that the ML decoder is entirely insensitive to the choice of g, or the presentation probabilities.

4.2.7 *Maximum a posteriori* decoder

The maximum likelihood decoder classifies the population responses without reference to the frequency at which different inputs are presented. An alternate method of decoding is to use the *maximum a posteriori* (MAP decoder, see e.g. Salinas and Abbott, 1994; Pouget et al., 2000), which takes into account the relative presentation frequency of different inputs by ensuring that infrequent distractors have a limited ability to distract the responses of frequently presented inputs. This method attempts to maximise the posterior distribution $p(x,c|r_{rec})$, i.e. it finds the input presented x_{est}, and the contrast at which the presentation is made, c_{est}, which has the highest probability, conditioned on the recorded response r_{rec}. As in other treatments, Bayes' rule can be used to rewrite this as

$$\{x_{est}, c_{est}\} = argmax_{x,c} p(r_{rec}|x,c)p(x,c)$$

where $p(x,c)$ is the prior distribution which details the probability with which inputs are presented and what contrast they are presented at. If it is a uniform distribution, then the MAP decoder reduces to the ML decoder.

In order to examine the performance the MAP decoder with as little information about the input statistics as possible, we will fix the prior distribution $p(x,c)$ before testing its performance. This distribution can be written $p(x,c) = p(x)p(c)$ with the input being presented and the contrast at which is presented being independently drawn. We further assume that presentation contrast is drawn uniformly from the range $[0,1]$ and that the prior distribution for the inputs is as follows:

$$p(x = x_{input}) = \begin{cases} P_{zero}, & x_{input} \text{ is the zero input (see previously),} \\ \frac{P_{single}}{m}, & x_{input} \text{ is a single stimulus, from } m \text{ possibilities,} \\ \frac{2P_{pair}}{m(m-1)}, & x_{input} \text{ is a stimuli pair, from } m(m-1)/2 \text{ possibilities,} \end{cases} \tag{4.8}$$

where P_{zero} is the probability that no input is presented, P_{single} is the probability that only one input stimulus is presented and P_{pair} is the probability that a pair of input stimuli is presented - clearly $P_{zero} + P_{single} + P_{pair} = 1$. Unlike with the ML decoder, it is possible for the zero input to be outputted as a classification result other than the default choice from an otherwise indeterminate list of zero - contrast possibilities.

In a similar fashion to the ML decoder, we can use Eq. (4.4) and Eq. (4.3) to determine the contrast $c_{x,est}$ which maximises $p(r = r_{rec}|x,c)p(x,c)$ for a given input x, and find that it is identical to that for the ML decoder. We now return to finding x_{est} and can write that

$$\{x_{est}, c_{est}\} \quad = \quad \{argmin_x((g(c_{x,est})R_x - r_{rec})^2 - 2\sigma^2 \log p(x)), c_{x_{est},est}\}.$$

Essentially then, the approach is to first find, for each possible input, an optimal presentation contrast using Eq. (4.7). The input which maximises the posterior distribution x_{est} is that which minimises the square of the distance from its rescaled response vector $g(c_{x,est})R_x$, adjusted by the associated prior presentation probability and noise amplitude. When using the decoder, we set P_{single} and P_{pair} equal to each other (since, for inputs, we present equal number of single stimuli as we do stimulus pairs) and set P_{zero} to be slightly larger than P_{single}/m (this ensures that the decoder has a non - unique output with zero probability provided no response vectors are equal to each other).

Fig. 4.2C shows the MAP decoder in action in a toy example with our two - unit population. It compares the distance between the nearest point on the vector R_A (denoted xR_A) and the response r with the distance between r and the response vectors of

the other inputs, as well as the distance to the origin (the response vector of the zero input). These distances are then modified to take into account the probability at which the input in question is presented - the higher the probability, the greater the reduction in distance. Now, the MAP decoder correctly classifies the response as being generated by stimulus A. As before, the regions of classification are colour coded. Compared to the ML decoder, the regions associated with inputs with low presentation probabilities are smaller whilst those associated with large presentation probabilities are larger. Note that we do not actually present the zero input but for decoding with the MAP decoder, we assign a nominal presentation probability to the zero input - it is usually slightly larger than the probability for presenting a single stimulus, but in the figure we exaggerate it to twice that probability for illustrative purposes.

4.2.8 Correlated noise

We will mostly consider only uncorrelated noise in this chapter. However, when we consider mean population responses which are tuned, it is reasonable to consider that the corrupting noise might also be tuned in some fashion. When using correlated noise, we must adjust each decoder to counteract this. We assume that the covariance matrix for the additive Gaussian noise is positive definite (as opposed to just positive semi positive definite) - we can thus write it as $\sigma^2\Sigma$, where σ is a scalar and Σ is a positive definite matrix with the diagonal elements equal to one - σ is thus the standard deviation of the response of a single unit. Since Σ is real and positive definite, it can be written in the form with $\Sigma = AA^T$, where A is another real matrix, which is, in general, non unique. For ν normally distributed with zero mean and a covariance structure equal to the identity matrix (as used above), then $\sigma A\nu$ is normally distributed with mean zero and covariance $\sigma^2 AA^T = \sigma^2\Sigma$. So, to generate a response to an input with response vector R at contrast c and with covariance $\sigma^2\Sigma$, we can use the equation

$$r = g(c)R + \sigma A\nu. \tag{4.9}$$

Whilst this approach works for any such matrix A, we use the (unique) matrix square root s.t. $\Sigma = A^2$ (the results do not depend on this choice - using the Cholesky decomposition to generate a suitable matrix gives similar results). Clearly, the larger the number of matrix manipulations we perform, the greater the computational cost - to speed this up we multiply Eq. (4.9) by the inverse of A, to obtain the whitened response z, given by

$$z := (A^{-1}r) = g(c)(A^{-1}R) + \sigma\nu,$$

where z is the transformed response. This then is identical to the method described above with uncorrelated noise, except that rather than using a response vector R when generating response r, we use the transformed response vector $A^{-1}r$ to generate the transformed response $z := (A^{-1}r)$. Classification of these responses can be done using the MAP and ML decoders as outlined above to obtain the error rate to a model using full, correlated noise. By transforming all the response curves prior to presenting stimuli, we can thus reduce the computational cost - this is known as whitening (a common technique in signal processing - see e.g. Hyvarinen et al., 2001).

In order to create the shape of the covariance matrix, Σ, we must make some assumptions. For units arranged in a ring structure, we assume that the covariance between two units, i and j, can be defined purely as a function of the distance between them (measured relative to the ring size) $d(i,j) := Min(|i-j|/n, 1-|i-j|/n)$. We define a covariance shape function K, such that $Cov(x_i,x_j) = \sigma^2\Sigma_{i,j} = \sigma^2 K(d(i,j))$. This means that the matrix Σ is circulant, as well as strictly positive definite, and we can thus write it as having the inverse square root given by

$$(A^{-1})_{ij} = \frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{\sqrt{\lambda_k}}\cos(2\pi(i-j)k/n),$$

where λ_k is the k'th eigenvalue of the matrix Σ with the form $\lambda_k = \sum_{j=0}^{n-1}\cos(2\pi jk/n)K(j/n)$. For the covariance shape function K we choose an exponential function, so $K(x) \propto e^{-x/\mu}$, with μ, which we denote the *correlation width*, controlling how the correlations decay with distance.

4.3 Results

As discussed in the introduction, we investigate how different methods of responding to pairs of stimuli presented simultaneously impact on coding capacity. To do so, we use a simplified population of n units which responds to inputs in a contrast invariant manner. Each input, presented at full contrast has mean response R - we call this the *response vector*. The mean response is scaled by the input contrast - when the input is presented at contrast c, the population response has mean cR. Responses are then corrupted by additive, uncorrelated normally distributed noise (with covariance matrix $\sigma^2 I$ where σ^2 is the variance and I is the n - dimensional identity matrix). There are advantages and disadvantages to using additive noise rather than any other noise model. The main advantage is with the mathematics of the system (see later) - additive noise

allows us to build up a more intuitive understanding of our results. The disadvantage is that it is not necessarily physiologically realistic, and that using additive noise ensures that stimuli with long response vectors are decoded with greater accuracy than the responses to stimuli with short response vectors. We therefore normalised the response vector lengths such that $|R| = 1$.

We consider two types of single stimuli, one type which have randomly generated response vectors, and another type in which the stimuli display orientation tuning. In order to generate response vectors associated with pairs of stimuli presented simultaneously (and at identical contrast), we use a *combination function* f to combine the response vectors, R_A and R_B, of each of the presented single stimuli A and B, into the response vector $R_{AB} = f(R_A, R_B)/|f(R_A, R_B)|$ of the stimulus pair AB. We use three such functions - *linear sum* ($f(R_A, R_B) = R_A + R_B$), *maximum* ($f(R_A, R_B) = \max(R_a, R_b)$) and a competitive *suppression* function ($f(R_A, R_B) = \max(R_A, R_B) - \min(R_A, R_B)/2$). The normalisation ensures that the response vectors all have length 1, as discussed above - this means that we focus on the 'shape' rather than the amplitude of how combined responses are generated. In particular the linear sum combination function leads to the same response vectors as an averaging combination function.

To assess how effectively a population can encode stimuli, we present equal numbers of single stimuli and stimulus pairs to the population at random contrasts and use our chosen decoding method to the classify the responses which are generated. The proportion of incorrect classifications is called the *error rate*. Changing the number of stimuli encoded by a population will change the error rate. For a fixed target error rate (we use 10% unless otherwise noted) we call the maximum number of single stimuli which can be encoded (along with all possible stimulus pairs) by a population the *coding capacity*. We estimate the error rate using computational simulations. In these simulations, we first generate the response vectors associated with the single stimuli and stimulus pairs. Then, we generate the responses to 10'000 single stimuli inputs and 10'000 stimulus pairs inputs with uniformly random contrast, using our decoding method to classify the results which are generated - we measure the error rate of this decoding. We repeat each simulation run using 100 different random seeds (in situations with randomly generated response vectors, this allow different sets of response vectors to be used), averaging the results (using the resources provided by the Edinburgh Compute and Data Facility).

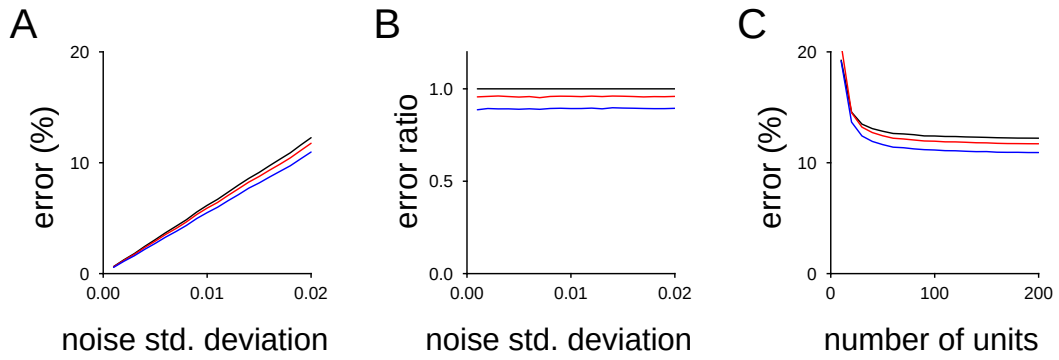

Figure 4.3: **A** The classification error percentage (using the MAP decoder) is plotted against the standard deviation of the noise for a population of 200 units where the response vectors of 100 different single stimuli are generated (iid) from an exponential distribution, and the response vectors associated with stimulus pairs are produced using one of three different combination functions - linear sum (black line), maximum (red line) and suppression (blue line). **B** As A, but we now plot classification errors from the MAP decoder where the maximum (red) and suppression (blue) combination functions are used as a proportion of the classification errors when the linear sum combination function is used against the amplitude of the noise. **C** We plot the error percentage when 100 single stimuli are encoded using a different number of units subject to noise with a standard deviation of 0.02.

4.3.1 Non - linearly combining responses to multiple inputs with random population responses increases coding capacity

We first investigate the situation where the single stimuli response vectors are independent of each other, with random components drawn independently from an exponential distribution (before normalising the response vector's length). Fig. 4.3A shows the average error rate of a population of 200 units which each encode 100 single stimuli and all possible stimulus pairs using one of the three different combination functions (linear sum - black line, maximum - red line, suppression - blue line) in the presence of noise of increasing standard deviation. We measure the error rate when using the *maximum a posteriori* (MAP) decoder (note the results here are broadly similar to the results when we use the ML decoder, albeit using the ML decoder tends to result in a greater error rate) and find that the use of the maximum and suppression combination functions provide modest decreases in error rate for identical noise levels (with suppression the least susceptible to errors). Fig. 4.3B shows the error rate when the maximum or suppression combination functions are used, as a proportion of the errors

produced when the linear sum combination function is used. The proportions are approximately constant for increasing noise - although obviously, they converge at high noise to 1 (not visible on plot).

When we fix the noise level and instead vary the number of units in the population, we find that increasing the number of units beyond a certain size leads to a limited improvement only (Fig. 4.3C, which shows the error rate when the noise has a standard deviation of 0.02 for the different combination function types as the population size n is changed). As n increases, the distribution of distances between any pair of response vectors will sharpen around the mean distance - likewise the distribution of angles between the various response vectors will also become more sharply peaked - see the mathematical analysis below for more details. As the error rate in decoding responses is based on the geometry of the response vectors, the sharpening of the distribution of angles between response vectors as n increases will ensure that the error rate will approach some limit. Note that this depends on response vector normalisation to length 1. In absence of this, the average length of the response vector and the average distance between response vectors distance grows with $\sqrt{n}$ and so for fixed noise the error rate will drop. This effect has been remarked upon before as a possible advantage of encoding responses to stimuli in very high dimensional spaces (Kanerva, 2009).

The source of errors to the system can also be examined. When the input is a single stimulus, the simulations show that the most effective distractor, for each of the different combination functions, is most probably another single stimulus - such a source of errors is relatively invariant over the different combination functions. However, when the input is a stimulus pair, the frequency of errors depends more strongly on the combination function used. In particular, simulations show that the most effective distractors (relative to their presentation frequency) to decoding responses to a stimulus pair AB are the single stimuli A or B, which are the *component stimuli* of the pair AB, and other stimulus pairs, AX or BX, which share a *component stimulus* with AB. These two sources of errors are the most dependent on the combination function used, and are the source of much of the relative differences in decoding performance for different combination functions. We return to this issue below, when considering the mathematics of the model which we define. Simulations also show that such errors also typically occur when the input is presented at mid to high contrast - at lower presentation contrasts, the probability of classifying the response as being generated by the input of a single stimulus (with little bias in favour of one of the input's component stimuli) dominates (the MAP decoder biases in favour of frequently presented inputs

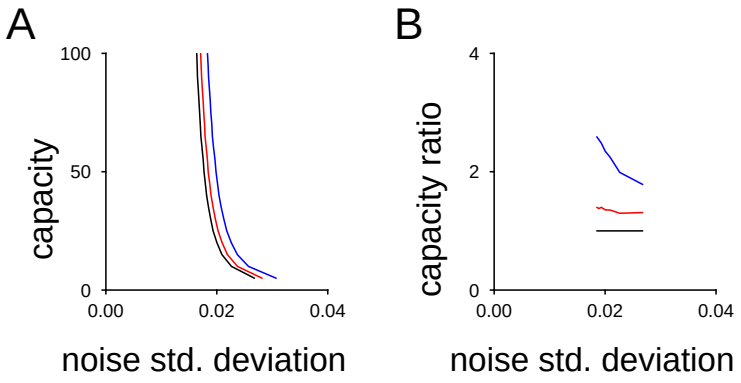

Figure 4.4: **A** The coding capacity (measured by the largest number of single stimulus inputs which can be encoded - along with all stimulus pair inputs - whilst the error rate remains below 10%) of a 200 unit population using one of three different kinds of combination function (linear sum - black line, maximum - red line, suppression - blue line) plotted against the standard deviation of the corrupting noise. **B** As A but we plot the coding capacity of the population using either the maximum or suppression combination functions normalised by the coding capacity when the linear sum combination function.

such as single stimuli, see methods for details) - this type of error is largely independent of the choice of combination function. This ensures that errors are much more likely with linear combination functions when the responses are generated at high contrast although overall error rates are low. At low contrast meanwhile, there are high error rates for all combination functions.

We also measured the number of stimuli which each population of units can encode (including all possible stimulus pair inputs) whilst maintaining an error rate below 10% (simulations show error rates from 5% to 20% provide similar results). Fig. 4.4A plots the coding capacity of a population of 200 units, with one of three different combination functions (linear sum - black line, maximum - red line, suppression - blue line) against the standard deviation of the noise. Despite the limited relative improvement in error rate provided by the maximum and suppression combination functions over the linear sum, at low noise levels clearly this can translate as a large relative improvement in coding capacity. Fig. 4.4B shows this, plotting the coding capacity of the population using the maximum (red line) and suppression (blue line) combination functions as a multiple of the coding capacity when the linear sum combination function is used (black line, for comparison). In particular, the suppression combination function allows, for a range of noise levels, more than double the coding capacity as when using a linear sum combination function - this increases as noise is decreased, although com-

putational limitations makes exploring this impractical for low noise levels when the number of different stimuli increases (the number of calculations the decoder requires increases with m^2).

4.3.2 Mathematical approximation indicates that response vector geometry plays major part in combination - function - - dependent error rate

Simulations show that the types of error most dependent on the combination function used are where the responses to stimulus pairs are misclassified by the decoder as having been generated by one of the stimulus pairs' component stimuli, or by another stimulus pair which shares a component stimulus with the original input. When the response vector is formed using the linear sum combination function, this leads to a higher error rate contribution from of one of these types of error than when one of the nonlinear combination functions is used instead. In this section, we examine the first of these error types in more detail to give an intuition into the difference in error rates between the different combination functions. We show how responses to stimulus pair AB (with response vector Z) can be erroneously decoded as stimulus A or stimulus B (with response vectors X and Y respectively) and, in accordance with simulation results, why this type of error rate is more likely when the response vector Z is formed by the linear sum combination function. We approach this problem in two stages. First, we come up with a reduced dimension approximation of the system we are considering, and then we use a simplified ML decoder (rather than the MAP decoder, which is mathematically more difficult) to examine the error rates we are considering.

4.3.2.1 Approximation of response vector geometry

We want to examine how responses to stimulus pair AB can be successfully distracted by stimulus A or stimulus B. To examine this, we consider a reduced system in which only the response vectors of stimuli A, B and stimulus pair AB (denoted X, Y and Z) are possible decoder outputs, along with the zero input. To examine how the choice of combination function used affects such errors in a systematic manner, we use a parametrised combination function to generate response vector Z with the following form

$$Z = \hat{f}(X,Y,\alpha) = \kappa(\alpha)[\max(X,Y) + \alpha \min(X,Y)].$$

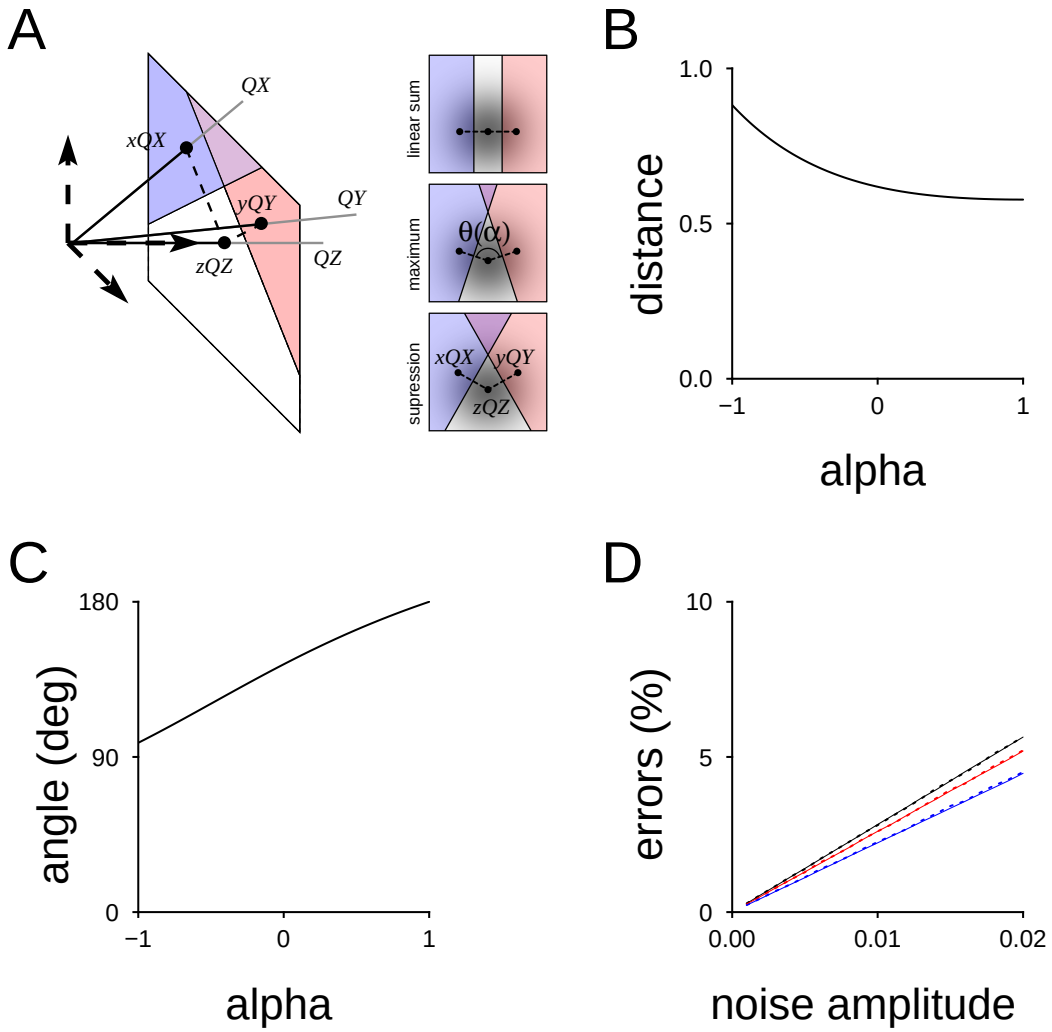

Figure 4.5: **A** The diagram of the reduced geometry of the response vectors X,Y and Z associated with the input of single stimuli A and B and the input of stimulus pair AB in the three dimensional subspace - one of the dimensions of which is defined as the direction of the vector Z - which contains the three response vectors. The plane perpendicular to point zZ, where z is the distance from the origin, (which intersects X and Y at points xX and yY, where x and y are scalars) is plotted - the coloured regions indicate those which are closer to X (light blue), Y (light red) or both (light purple) than to Z and are thus classified erroneously by the ML decoder. *Inset* The plane perpendicular to Z for different combination functions. The angle between $(xX - zZ)$ and $(yY - zZ)$ is defined as $\theta(\alpha)$, and varies with the combination function. **B** The theoretical mean distance between zZ and xX is plotted for $z = 1$ as a function of α. **C** as B, but we now plot the theoretical mean angle between $(xX - zZ)$ and $(yY - zZ)$, $\theta(\alpha)$. **D** The response to a presented stimulus pair AB is classified with only single stimuli A and B, and the zero input as distractors, using the ML decoder. The classification error is plotted as a percentage (dashed lines - black for linear sum, red for maximum, blue for classification error) and compared to the error rates obtained from the approximate approach (solid lines).

Here, $\kappa(\alpha)$ is a scaling factor (defined below) and α (in $[-1,1]$) controls the type of function - $\alpha = 1$ leads to the linear sum, $\alpha = 0$ leads to the maximum function, and $\alpha = -1/2$ leads to the suppression function. We simplify the system by effectively reducing the dimensionality of the problem to a manageable level.

We first try to ensure the response vectors X, Y and Z have length 1. We start with X, assuming, as above, that components X_i, $i \in [1,n]$, are independently drawn from an exponential distribution with probability density function $\mu \exp(-\mu x)$. We have that the square of the length of $X^2 = \sum_{i=1}^{n} X_i^2$ is a random variable with the following expression for the mean squared length

$$\mathrm{E}[X^2] = \sum_{i=1}^{n} \mathrm{E}[X_i^2] = n\mathrm{E}[X_i^2] = \frac{2n}{\mu^2}.$$

and that the variance of the squared length is given by the following expression

$$\mathrm{Var}(X^2) = \sum_{i=1}^{n} \mathrm{Var}(X_i^2) = \sum_{i=1}^{n} (\mathrm{E}[X_i^2]^2 - \mathrm{E}[X_i^4]) = \frac{20n}{\mu^4}.$$

Since we normalise the length of the stimuli response vectors we generate, we choose $\mu = \sqrt{2n}$, so the square of the response vector length, X^2, now has expected value 1, and its variance is $5/n$. As n increases then, the variance of the length of X will decrease (law of large numbers) whilst the mean length of X will remain 1. In order to simplify the mathematical approach, we assume that n is high, and that the length $|X|$ is thus close to 1 - we can thus omit the normalisation step and be satisfied with the approximate relation $|X|,|Y| \approx 1$.

We follow a similar approach with Z. The mean of the squared length of Z can be written as $\mathrm{E}[Z^2] = n\mathrm{E}[Z_1^2]$. We calculate this, obtaining

$$\begin{aligned}
\mathrm{E}[Z^2] &= n\mathrm{E}[Z_1^2] = n\mathrm{E}[\hat{f}(X,Y,\alpha)^2] \\
&= n\int_{x=0}^{\infty}\int_{y=0}^{\infty} \kappa^2(\alpha)(\max(x,y) + \alpha\min(x,y))^2(2n\exp(-\sqrt{2n}(x+y)))dxdy \\
&= \frac{(7+\alpha(4+\alpha))\kappa^2(\alpha)}{4}
\end{aligned}$$

We then choose $\kappa(\alpha)$ such that the average squared length of Z is 1, giving us the combination function

$$\hat{f}(X,Y,\alpha) = [2/\sqrt{(7+\alpha(4+\alpha))}](\max(x,y) + \alpha\min(x,y)).$$

As before, the variance of the length decreases with increasing n - and we again assume that we are using n high enough such that omitting the usual response vector normalisation step has little overall effect on the vector length.

The next step is to transform the system to more easily examine how responses to an input of the stimulus pair AB can be incorrectly classified by the decoder as having been generated by single stimuli A or B. The response vectors X, Y and Z can be contained entirely within a three dimensional subspace of $\mathbb{R}^n$ (this is unique if X, Y and Z are linearly independent, but uniqueness is not required - we merely assume that none of the response vectors are equal to each other). We can always rotate X, Y and Z so that they are present only in the the first three dimensions of $\mathbb{R}^n$. In particular, we choose rotation Q such that:

- QZ has one non - zero component - the first one (which is equal to 1),
- QX has only two non - zero components - the first two (with the second positive),
- QY has at most three non - zero components - the first three (with the third non - negative).

Without loss of generality, we can write $\nu = Q^T\omega$ where ω has independent components drawn from a normal distribution with zero mean and unit variance - this satisfies the requirement that the additive noise has zero mean and covariance matrix $\sigma^2 I$. We now transform the responses using Q to give us $\bar{r} = Qr = g(c)QZ + \sigma\omega$ - the components of transformed responses $\bar{r}$ are thus independent to each other. Now the difference in squared distances from transformed response $\bar{r}$ to the nearest points on the transformed vectors QX, QY and QZ depend only on the first three dimensions - the contributions from the fourth dimension to the n'th dimension to the squared distance from $\bar{r}$ to points on QX, QY or QZ are the same at $\sum_{i=4}^{n}\bar{r}_i^2$. Since both the ML decoder and the MAP decoder are based on such differences in distance and since the noise components in the transformed space are independent to one another, we can classify response r using only the first three components in transformed response $\bar{r}$, denoted $(\bar{r}_1, \bar{r}_2, \bar{r}_3)$, where $\bar{r}_1 = g(c) + \sigma\nu_1$, $\bar{r}_2 = \sigma\nu_2$ and $\bar{r}_3 = \sigma\nu_3$ for ν_i normally distributed with zero mean and unit variance.

4.3.2.2 Approximation of the ML decoder

Finally, we examine how the ML decoder operates in this subspace of the reduced system with inputs A,B, AB and the zero input. If we assume that X, Y and Z are of length 1 (a good approximation for high enough n, as we have discussed), then the ML decoder (applied only on inputs A, B, AB and the zero input) will classify the response r as having been generated by the input whose transformed response vector, from QX,

QY and QZ, has the smallest angle between itself and $\bar{r}$, provided the associated contrast estimate is positive. Fig. 4.5A shows this situation. The plane plotted in Fig. 4.5A is perpendicular to QZ, and intersects QZ at the point zQZ (intersecting QX and QY at xQX and yQY respectively). Noting that $z = g(c) + \sigma\nu_1$, we have that if $z < 0$ then the ML decoder's contrast estimate associated with stimulus pair AB is always zero - the decoder will thus always prefer the zero input to the pair AB, ensuring a classification error. On the other hand, if $z \geq 0$, then whether the ML decoder makes a classification error depends on which region of the plane in Fig. 4.5A that $(\bar{r}_2, \bar{r}_3)$ is in:

- For $(\bar{r}_2, \bar{r}_3)$ in the light blue region, the ML decoder will always classify the response as being generated by stimulus A ahead of stimulus AB since the angle between X and $\bar{r}$ is smaller than the angle between Z and $\bar{r}$.

- For $(\bar{r}_2, \bar{r}_3)$ in the light blue region, the ML decoder will always classify the response as being generated by stimulus B ahead of stimulus AB since the angle between Y and $\bar{r}$ is smaller than the angle between Z and $\bar{r}$.

- For $(\bar{r}_2, \bar{r}_3)$ in the purple region, the ML decoder will always classify the response as being generated by either stimulus A or stimulus B ahead of stimulus AB since the angle between X or Y and $\bar{r}$ is smaller than the angle between Z and $\bar{r}$.

- For $(\bar{r}_2, \bar{r}_3)$ in the white region, the ML decoder will correctly classify the response r as having been generated by stimulus pair AB.

The boundary lines on the plane scale linearly with z with smaller values of z leading to much closer boundaries, and thus a lower probability of $(\bar{r}_2, \bar{r}_3)$ lying in the white region (since the generation of $\bar{r}_2$ and. The error rate then depends on three principle values - the distances from zQZ to xQX and to yQY, which control how vary the boundary lines are from zQZ, and the angle $\theta(\alpha)$ between $xQX - zQZ$ and $yQY - zQZ$, which controls to what extent the red and blue error regions overlap, in the form of the purple region. The further zQZ is from xQX and yQY, the less likely an error is to occur, and the smaller the angle $\theta(\alpha)$ the more the two error regions overlap, reducing the probability that an error occurs.

Using methods similar to those we use to approximately normalise the lengths of X, Y and Z, we find approximations for the mean distance (multiplied by z), $zD(\alpha)$, between zQZ and xQX (the distance from zQZ and xQX has the same mean) and the

mean angle $\theta(\alpha)$, which is independent of z. The expressions are as follows

$$\begin{aligned} D(\alpha) &= \sqrt{-1+\frac{16(7+\alpha(4+\alpha)}{9(3+\alpha)^2}}, \\ \theta(\alpha) &= \arccos\left(\frac{25+\alpha(22+\alpha)}{31+\alpha(10+7\alpha)}\right). \end{aligned}$$

We plot these approximations in Fig. 4.5B and Fig. 4.5C. We find that the more linear the combination function is (α approaching 1), the smaller the distance $D(\alpha)$ and the larger the angle $\theta(\alpha)$, both of which increase the classification errors. This then is a big source of the increased error rate when using the linear combination function as opposed to one of the nonlinear alternatives.

Fig. 4.5A, inset, shows the plane from the main portion of Fig. 4.5A for a fixed value of z with three different combination functions - linear sum, maximum and suppressive. The grey shaded region is probability density function for the point position $(\bar{r}_2, \bar{r}_3)$ in this plane - darker indicates greater probability. We can see the effect on the position and angle of the boundary lines clearly. In the top panel, where the combination function is that of linear sum, the red and blue regions are non - overlapping, ensuring a smaller probability that the point $(\bar{r}_2, \bar{r}_3)$ will lie within the white, error - free region. The bottom panel on the other hand shows a much larger overlapping error region (purple region). This increases the probability of the point $(\bar{r}_2, \bar{r}_3)$ lying within the error - free region, as the more the red and blue error regions overlap, the lower proportion of space which their union occupies. The integrals over these regions are done stochastically, following a similar procedure to calculating the error rate we use on the full network. Fig. 4.5D compares the error rate approximation calculated in such a fashion with the percentage of response to stimulus pairs in which the likelihood of the response given the stimulus pair (at optimal contrast) is less than the likelihood of the response given either of its component stimuli, or the zero input (all at optimal contrast). The close match shows that the mathematical approximations we have implemented here lead to accurate results. Increasing n further increases the accuracy of this approach as the greater number of components in X, Y and Z lead to better accuracy of the expressions we use to approximate the means.

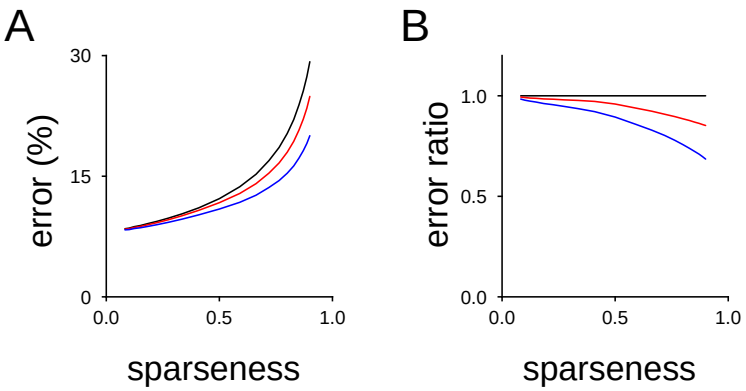

Figure 4.6: **A** The classification error percentage (using the MAP decoder) of the population responses plotted against the sparseness of the response vectors associated with the 100 single stimulus response vectors for a population of 200 units where the response vectors associated with inputs of stimulus pairs are produced using one of three different combination functions - linear sum (black line), maximum (red line) and suppression (blue line), for noise with standard deviation equal to 0.02. **B** As A, but we now plot classification errors when the maximum (red) and suppression (blue) combination functions are used as a proportion of the classification errors when using the linear sum combination function against the sparseness of the single stimulus response vectors.

4.3.3 Sparse responses have reduced error - rate dependence on combination functions

We have examined the situation where the components of response vectors to single stimulus inputs are exponentially distributed, and found that using the maximum or suppression combination functions to generate the response vectors for stimulus pairs from those of the single stimuli inputs reduced the error rate slightly, for fixed population size and number of stimuli, and substantially increased coding capacity. Although it is impossible to test every possible response vector distribution, we here alter the single stimuli responses to indicate how sparseness affects the situation we describe. We generate response vectors of different sparseness by generating components from an exponential distribution and then subtracting or adding a fixed cutoff value from each (a value from zero to twice the mean of the exponential distribution, positive and negative), before normalising the response vector as usual. We use a common measure of sparseness, $s = \mathrm{E}[R]^2/\mathrm{E}[R^2]$, to describe the response vectors which are generated by this method (Treves and Rolls, 1991; Rolls and Tovee, 1995). For sparse response vectors s is close to 0, whilst for dense response vectors s is closer to 1. The original exponentially drawn response vectors have $s = 1/2$. A high, positive cutoff will lead to many zero elements and thus a low value of s whilst a negative cutoff value adds a positive term to each component, leading to a higher value of s.

Fig. 4.6A plots the average error rate of a 200 unit population encoding 100 single stimulus inputs and using one of the three different combination functions (linear sum - black line, maximum - red line, suppression - blue line) as we vary sparseness measure s measured from the single stimuli response vectors. Again, we present equal numbers of single stimuli and stimulus pairs as inputs, at uniformly random contrasts, and classify the population responses using the *maximum a posteriori* (MAP) decoder in the presence of fixed noise of standard deviation 0.02. We find that errors are less common for sparse single stimulus response vector distributions for all type of combination functions. This is to be expected - both the MAP decoder and the ML decoder depend, to a certain degree, on the distance between response vectors, with those response vectors which are further apart leading to fewer classification errors. It is trivial to show that for the form of single stimuli we choose, the sparse response vectors are further, on average, from each other - as sparseness falls, the components of any two response vectors become more similar, on average, leading to a low angle between such vectors. This is likely true for other types of response vector distributions - dense

response vectors have components which are more similar to each other in value - for fixed length response vectors, this will likely lead to high dot products, and thus low distances between any two such vectors. Fig. 4.6B shows the error rate from Fig. 4.6A, where the error rates when the maximum and suppression combination functions are used are plotted as a ratio to the error rate when the the linear combination function is used. As the sparseness measure s decreases (i.e. for response vectors with few non - zero elements), the error rates converge. Again, this is to be expected - the sparse response vectors (i.e. those with low sparseness values) contain large numbers of zero components, and both the maximum and suppression combination functions behave identically (up to a proportionality constant, which is eliminated by the normalisation constraint) to the linear sum when presented with one zero input.

4.3.4 Non - linearly combining the responses to pairs of input orientations leads to reduced error rates and increased coding capacity

We have so far assumed that the mean responses of our population of units to different single stimuli are independent from each other (save for the normalisation constraint). However, experimental evidence in both low - level and high level areas of the visual system has indicated that this is not the case. Indeed, the response of a neuron to different stimuli can often display evidence of tuning to a given parameter, such as orientation in V1 (Hubel and Wiesel, 1962), or head rotation in STS (Perrett et al., 1991). We therefore examine the problem of decoding and encoding pairs of stimuli using response vectors which display such parametrised tuning (we denote this parameter as *orientation* but other than its modulo property, this label has no bearing on our results). Each unit has a preferred orientation, and each single stimulus input has an orientation value - both of these are arranged evenly in a ring structure. We assume that the mean response of a given unit to a single stimulus is entirely determined by the difference between that unit's preferred orientation and the orientation of the presented stimulus, using a von Mises *tuning curve* (see method for details). As before the response vector of a stimulus pair input are created by combining, and then normalising, the response vectors of the component single stimulus inputs using a combination function, either the linear sum, the maximum, or the suppression function.

We presented equal numbers of single stimulus inputs and stimulus pair inputs at random contrasts to a population of 200 units and encoding 100 separate single stim-

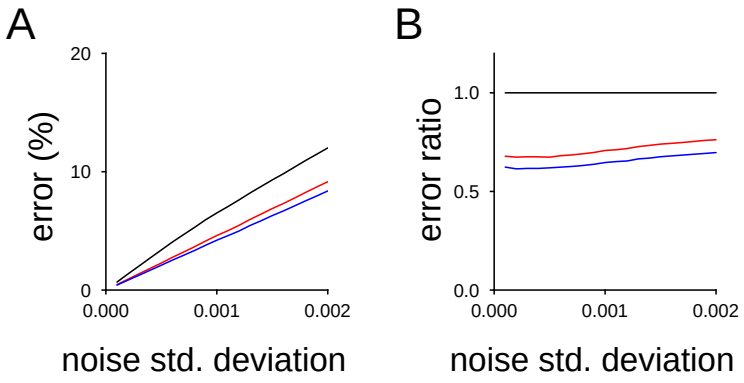

Figure 4.7: **A** The classification error percentage (using the MAP decoder) is plotted against the standard deviation of the noise for a population of 200 units where the response vectors of the 100 different single stimuli are tuned with a von Mises tuning profile, and the response vectors to stimulus pairs are produced using one of the three different combination functions - linear sum (black line), maximum (red line) and suppression (blue line). **B** As A, but we now plot classification errors when the maximum (red) and suppression (blue) combination functions are used as a proportion of the classification errors when the linear sum combination function is used, against the amplitude of the noise.

uli (and all possible pairwise combinations of stimulus pairs) at different noise levels. Fig. 4.7A shows the resultant error rate as a percentage (using the MAP decoding method) against the standard deviation of the noise input for the three different combination functions - linear sum (black line), maximum (red line) and suppression (blue line). When the response vectors to single stimuli have tuned properties, the error rates are much higher than when they are randomly generated. Similar orientation inputs have response vectors which are close together, and as the MAP (and MLE) decoders are distance based, these neighbouring stimuli form effective potential distractors. By contrast, randomly generated response vectors are rarely close together, especially for large networks. Additionally, using the linear sum as the method of combination produces many more errors than either of the alternate combination functions (which have similar performance to each other), whereas when the response vectors were generated randomly, the error percentage differences between combination functions were much more similar to each other. This is shown more clearly in Fig. 4.7B, which plots the classification errors when using either the maximum or the suppression combination function as a proportion of the classification errors when the linear sum combination function is used, with the ratio around 0.65 for both at low noise levels. We also find

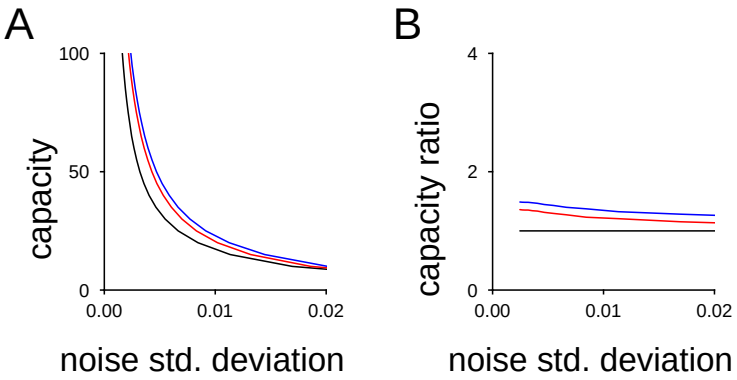

Figure 4.8: **A** The coding capacity of a 200 unit population (measured as the number of distinct stimuli with response functions which display von Mises tuning) using one of the three different kinds of combination function (linear sum - black line, maximum - red line, suppression - blue line) operating with a 10% error rate is plotted against the standard deviation of the noise. **B** As A but we plot the coding capacity when using either the maximum or suppression combination functions normalised by the coding capacity when using the linear sum combination function.

that changing the number of units in the population (whilst keeping the width of the tuning curve, as a proportion of the population size, fixed) does not impact the decoding performance in the presence of noise for a fixed number of input stimuli (not pictured). This is to be expected, as high n for a population simply means that the single stimulus response vectors are much smoother representations of the underlying Gaussian tuning function, and thus the distance between each of them is similar to those of the 'low resolution' response vectors which are associated with low n.

As before, we can examine the coding capacity of a 200 unit population where the response vectors to single stimuli inputs are tuned. We again present equal numbers of single stimulus inputs as we do stimulus pair inputs, and look for the number of single stimuli which can be encoded whilst not exceeding a 10% error rate (other error rate percentages lead to similar results as we present here). Again, the linear combination function is the least effective method of generating response vectors to stimulus pair inputs, with the suppression function the most effective in terms of coding capacity. There is a greater relative difference in resistance to errors displayed across the different combination functions when using stimuli with tuned response vectors compared with the situation, discussed earlier, in which the response vectors are randomly generated with no overall structure. This does not translate across to coding capacity, with the combination function choice not having as large effect as might be expected. This

is because there is a much sharper decrease in coding capacity, as noise amplitude increases, when the response vectors are unstructured (at least, for the noise amplitude to which we measured relative performances) which ensures that the relative differences are larger than when the response vectors are tuned.

4.3.5 Smooth tuning curves arranged with rotational/translational symmetry combined linearly have effective approximations

When the single stimuli have response vectors which are tuned, the populations responses to inputs are, in general, more susceptible to errors, than when single stimuli response vectors are randomly, and independently, drawn from an exponential distribution. The tuning constraint on the response vectors of single stimuli ensures that they can be very close together, particularly for stimuli which have similar presentation orientations - the response to an input is much more likely to be distracted by another input which has a nearby response vector. In contrast, when the response vectors of single stimuli are randomly generated, the distances between stimuli are narrowly peaked around an average distance, ensuring that no two response vectors are close enough to each other to act as effective distractors. This likely accounts for much of the difference in overall error rates (compare the noise standard deviation in Fig. 4.8A with that in Fig. 4.4A). The relative difference in decoding error rates across combination functions is also larger when the single stimulus response vectors are tuned. Here, we examine this issue in more detail.

For stimulus pair AB, with response vector R_{AB}, we want to examine single stimuli distractors which could be big sources of decoding error. Using the ML decoder, we can see that effective distractors of responses to stimulus pair AB will be inputs which have response vectors close to R_{AB}. Suppose that stimulus C and contrast c maximise the likelihood of a response equal to R_{AB} over all possible single stimuli and contrasts (i.e. the angle between its response R_C and R_{AB} is smaller than with any other single stimulus) - then we denote C as the *optimal distractor*. A response to stimulus pair AB is much more likely to be misclassified as being generated by single stimulus D when R_D and R_{AB} are close to each other in angle, such as the optimal distractor C.

This situation is plotted in Fig. 4.9A, which shows the optimal single stimulus distractor for different stimulus pair input response vectors. The top row shows the response vectors of the two component stimuli A and B, as well as the response vector of the optimal distractor stimulus C (with the orientation C intermediate to those

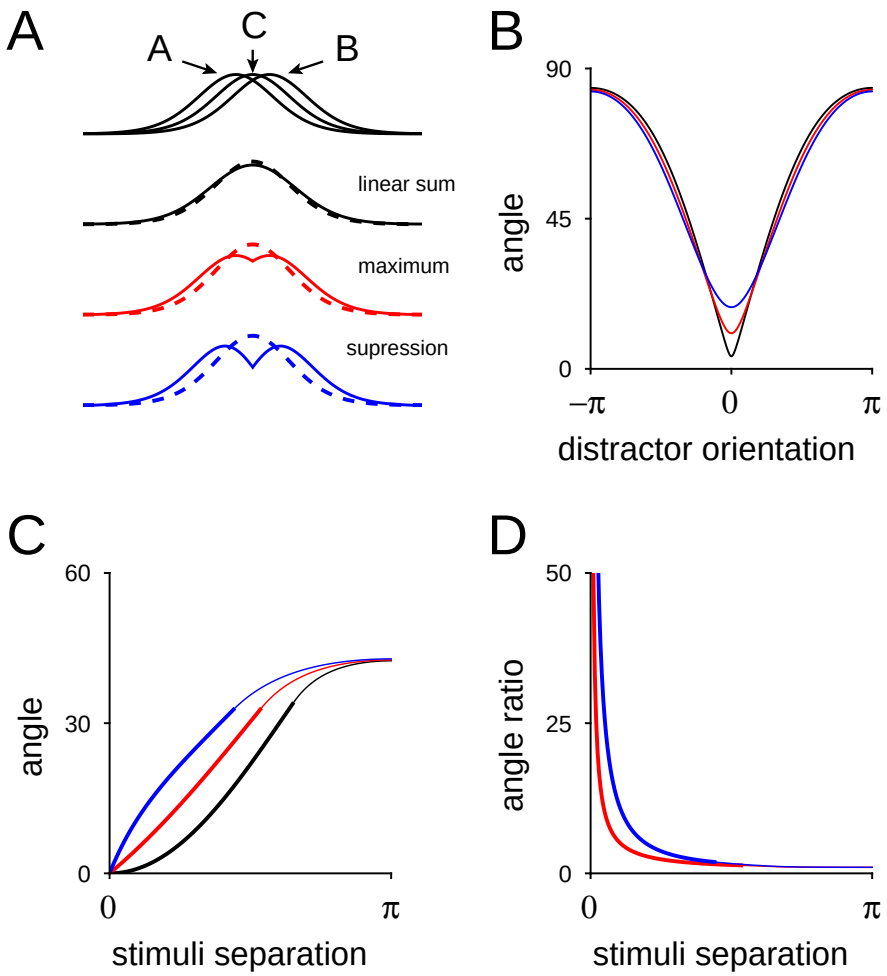

Figure 4.9: **A** *Top row* Response vectors associated with three separate single stimuli inputs, where the stimuli each have different orientations and are denoted A, B and C (which lies directly in between A and B). **B** *Second row to bottom row* The response vectors R_{AB} of stimulus pair AB presented together using three different combination functions - linear sum (black solid line), maximum (red solid line) and suppression (blue solid line). Also plotted is the response vector R_C associated with the input of single stimulus C, scaled to minimised its distance from the response vector of the stimulus pair (dashed lines). **B** We plot the angle between the response vector of stimulus pair AB and the response vector of the stimulus D, against the orientation of stimulus D, relative to the orientation of stimulus C. **C** The angle between the response vector of the optimal distractor stimulus and the response vector of stimulus pair AB, against the difference in orientation between A and B. The thick lines indicate regions for which optimal distractor is stimulus C, lying between A and B. **D** As C, but with the angle when using the nonlinear combination functions plotted as plotted as the ratio w.r.t. the corresponding angle when using the linear combination function.

of A and B). We create, using the three combination functions - linear sum (black), maximum (red) and suppression (blue), the response vectors to stimulus pair AB, R_{AB} (solid lines, bottom three rows). We plot $g(c)R_C$, where c minimises the distance from $g(c)R_C$ to R_{AB} (essentially maximising the likelihood of R_{AB} given contrast c and stimulus C - dashed lines). We clearly see that the response vector of the stimulus pair AB produced by the linear sum combination function is close to the rescaled response vector of the single stimulus C, unlike those response vectors produced by the nonlinear combination functions - responses to stimulus pair AB constructed using the linear sum combination function will thus be more susceptible to errors.

Fig. 4.9B plots the angle between the response vector R_{AB} (constructed from the response vector of stimuli A and B using the linear sum - black, maximum - red and suppression - blue, combination functions) and the the response vector R_D of stimulus D against the orientation of stimulus D relative to that of stimulus C, halfway between A and B (and the optimal distractor in this case for all three combination functions). When the linear combination function is used, the angle between the response vector R_{AB} and the response vector of the optimal distractor is much smaller than when nonlinear combination functions are used. We also examine this effect as the orientation distance between A and B is changed. Fig. 4.9C plots the angle between R_{AB} and the response vector of its optimal distractor, as the difference in orientation between A and B varies - the thick lines indicate that the optimal distractor is indeed C, with orientation intermediate between those A and B. Fig. 4.9D plots the ratios of the angle between the response vector R_{AB} and the response vector of the optimal distractor when the nonlinear combination functions are used compared to the same angle when the linear combination function is used. We find that for small separations between A and B, not only is the angle between the response vector R_{AB} and the response vector of the optimal distractor smaller, but that when the linear sum is used, the angle is dramatically smaller than the same angle when the nonlinear combination functions are used.

Altogether, when two similar stimuli are combined using the linear sum combination function, the resultant response vector has a small angle to that of an intermediate single stimulus - equivalent to being close in distance. Any response to such inputs is thus prone to miss - classification as such distractor stimuli (since both the MAP decoder and the ML decoder depend, in part, on distance). A similar situation exists in reverse also - responses to single stimuli inputs can be misclassified as having been generated by stimulus pair inputs with similarly orientated response vectors. We can

see the reason for this by using some basic mathematics. Suppose that stimulus A and B have response vectors $F(\theta-\phi)$ and $F(\theta-\phi+2\psi)$ where θ is the orientation of a given unit, ϕ is the orientation of stimulus A, ψ is equal to half of the separation between them, and$F$ describes the shape of the tuning curve. By applying Taylor's theorem we obtain the following equations

$$\begin{aligned} F(\theta-\phi) &= F(\theta-\phi+\psi)-\psi F'(\theta-\phi+\psi)+\mathscr{O}(\psi^2), \\ F(\theta-\phi+2\psi) &= F(\theta-\phi+\psi)+\psi F'(\theta-\phi+\psi)+\mathscr{O}(\psi^2). \end{aligned}$$

Adding these equations together, we obtain that $F(\theta-\phi)+F(\theta-\phi+2\psi)=2F(\theta-\phi+\psi)+\mathscr{O}(\psi^2)$. Since the response vector associated with the distractor stimulus with orientation $\theta-\phi+\psi$ is proportional to $F(\theta-\phi+\psi)$, then for small enough ψ the response vector of stimulus pair AB is close to the response vector of the distractor single stimulus with intermediate orientation. This feature of the linear sum ensures that the response vectors of stimulus pairs, when the orientation separation of their components is small, will have distractor stimuli with close response vectors, causing high error rates. The mathematics is the same for any differentiable F (or indeed, any F differentiable on $[-\pi,\pi]$), and several different forms of tuning functions (log - normal and rectified cosine) which are (almost always) differentiable give similar results. The sharper the tuning curves, the lower the range of orientation separations (and thus the lower the proportion of available stimulus pair inputs) for which this effect is strong, which leads to reduced differences between the different types of combination functions.

4.3.6 Model assumptions have a critical effect on absolute and relative error rates

The model we investigate makes several assumptions and simplifications. In this section, we show how weakening these assumptions effects the error rate and coding capacity of our model populations.

The most obvious problem with the model above is the length normalisation of response vectors - the requirement that all response vectors have the same length. This was introduced because of our noise model - with additive Gaussian noise, a stimulus with an associated response vector of considerable length would be much more resistant to noise than a stimulus associated with a short response vector. Since we wished to investigate the effect of changing the 'shape' of the combination function,

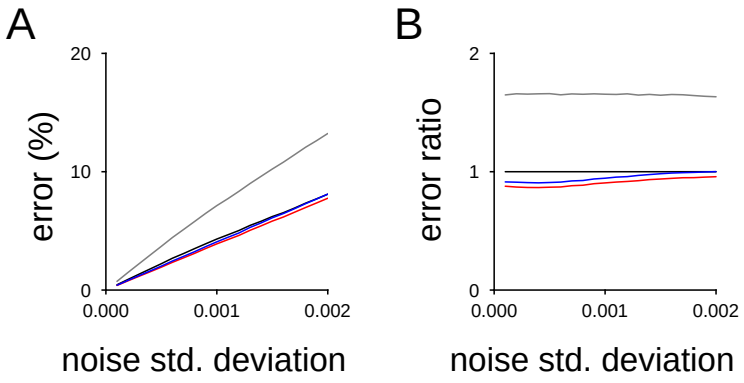

Figure 4.10: **A** The classification error percentage (using the MAP decoder) is plotted against the standard deviation of the noise for a 200 unit population where the response vectors of the 100 different single stimuli have a von Mises tuning profile, and the responses to stimulus pairs are produced using one of four different combination functions - sum (black line), average (grey line), maximum (red line) and suppression (blue line). **B** As A, but we now plot then classification errors where the maximum (red) and suppression (blue) combination functions are used as a proportion of the classification errors when using the linear sum combination function, against the amplitude of the noise.

we chose to normalise the response vector lengths to remove this length based factor. Whilst causing no problem in terms of our response vector choice for single stimuli (since the tuned stimuli have fixed length response vectors anyway, and the length of the randomly generated response vectors is tightly distributed around the mean value - for a large enough number of units), it does mean that our combination function equivalence classes don't quite match with those defined in the experimental literature, with the sum in particular being indistinguishable from the average. Now we have a better understanding of what effect the different combination function shapes have, we could also put that in context by removing this constraint and examining how the performance, in terms of error rate, changes.

We look at the implications of removing the constraint on response vector length for orientation tuned single stimuli. Fig. 4.10A, shows the average error rate, plotted against noise amplitude, for populations presented with equal numbers of single stimuli inputs as stimulus pair inputs, in the same fashion as previously. Now however, the response vectors to pairs of stimuli are not normalised - so the *summation* combination function produces response vectors to a pair of stimuli by simply adding the response vectors of the component stimuli together. By the same token, we define the *average*

combination function as that which produces the mean response vector from the two inputs (we refer to the maximum and suppression combination functions as before, but we now omit the requirement that the response vectors they produce be normalised). Clearly, as a result of these changes, use of the suppression combination function lead to a much higher error rate due to the average response vector length produced by such a function becoming a lot shorter, to the extent that the classification of responses occurs at a higher error rate when presented with stimuli than if the the maximum combination function was used, with the error rate more comparable to when the summation combination function is used. In contrast, although the maximum function now provides less of an improvement over the sum, it still provides an improvement, with the 'more effective' shape it leads to able to make up for the shorter length response vectors it leads to. The grey line shows the average operator - as expected, the reduced response vector lengths that it produces (compared to the sum) leads to increased error rates. The error rate is thus a trade off between the shape of the combination function and the magnitude of the response vectors it produces. Overall, the normalisation is an important factor when it comes to the improvements offered by nonlinear combination functions, with the maximum combination function having an advantage under this type of noise model. We should note that for single stimuli response vectors with independently drawn components, the response vector length has a much larger impact, to the point where the sum is the most advantageous method of response vector combination.

Removing constraints on response vector length is not necessarily more realistic - whilst the combination functions become easier to relate to experimental work when we no longer require normalisation - i.e. the sum of two response vectors is no longer in the same equivalence class as an average - how the increase in performance provided by a high amplitude response relates to experimental data is not so clear cut. Neural spiking data has indicated that the ratio of the variance of spike counts in a fixed window to the mean spike count remains approximately constant, i.e. multiplicative noise (Buracas et al., 1998). Incorporating noise which scales in such a fashion with response length might decrease the dependence of errors rate on response vector length (with high amplitude inputs much more susceptible to noise than when using additive noise). In general, a more realistic noise model will be needed to conclude which combination function is superior, but with the geometric arguments we have made here, we feel it likely that the combination function shape will still play a big role in coding capacity.

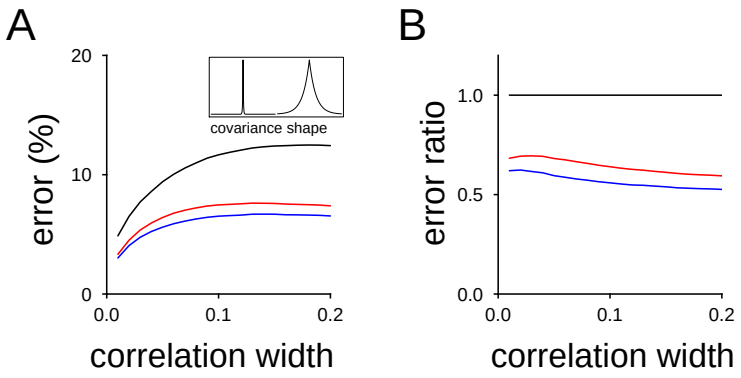

Figure 4.11: **A** The classification error percentage (using the MAP decoder) is plotted against the width of the correlation function for a 200 unit population where the responses to the 100 different single stimuli have response vectors with a von Mises tuning profile, and the responses to stimulus pairs are produced using one of three different combination functions - linear sum (black line), maximum (red line) and suppression (blue line). Noise is fixed with standard deviation 0.0005. *Inset* The covariance shape function (*left curve* - smallest correlation width, *right curve* - largest correlation width) **B** As A, but we now plot classification errors from where the maximum (red) and suppression (blue) combination functions are used as a proportion of the classification errors when using the linear sum combination function against the amplitude of the noise.

All the results above assume that the additive Gaussian noise is uncorrelated. This is clearly a simplification - correlations have long been observed to be present in the visual system (Smith and Kohn, 2008). We can incorporate this into our model to see if this simplification is critical in the results we observe. Again, we arrange the units in a ring, with tuned response vectors. We define the covariance matrix which we use to be circulant, with the covariance of two components of the noise dependent only on (and decaying exponentially with) the (shortest) distance between them (see methods for details). Fig. 4.11A shows the error rate of a 200 unit population with 100 single stimuli with tuned response vectors and one of the usual three combination functions, with Fig. 4.11B showing the error rate when using the maximum and suppression combination functions as a multiple of the error rate when using the linear sum combination function as we vary the width of the covariance shape function, whilst keeping the diagonal units constant (we call the diagonal element the variance, as they are all identical). The inset shows the covariance shape function - for low width (left inset curve, left of main figure) the noise is sharply peaked, with only close units exhibiting strong noise correlation, for high width (right inset curve, right of main figure) the noise is broadly peaked. We find, as expected that the wider the covariance shape function, the higher the error rate is. However, we note that, whilst the absolute error rate increases, the relative error rate as we use different combination functions is unchanged - we find this is approximately true for a range of different covariance determinant amplitudes and single stimulus numbers, indicating that the use of uncorrelated noise is not critical to the results which we report here.

4.4 Discussion

Using a simple model, we have investigated how different methods of encoding responses to stimulus pairs impacts upon the error rate and coding capacity. To generate the mean response to a pair of stimuli presented simultaneously, we combine the mean responses to those stimuli presented separately using combination functions corresponding to the *linear sum*, the *maximum*, and a competitive *suppression* function. We used two different types of stimuli set and found broadly similar results, namely that for inputs of random contrast, the decoding error rate when the linear combination function is used to generate stimulus pair responses is higher than when the maximum or suppression functions is used, with the suppression function leading to the lowest error rate. Accordingly, we found that the coding capacity - the maximum number of

single stimuli a population of units can encode (along with all possible stimulus pairs) whilst the error rate is less than 10% - increased when the maximum and suppression functions were used instead of the linear sum.

When the mean response to a single stimulus is generated independently from an exponential distribution, the difference in error rates over different combination functions is relatively small, but the relative differences in coding capacity are much larger, with the coding capacity decreasing steeply as the standard deviation of the noise increases contributing to large relative differences in coding capacity. Examining the sources of classification error which depended the most strongly on the choice of response function, we identified errors in which the response to a pair of stimuli presented simultaneously was classified as being generated either by one of the single stimuli which make up the pair components, or by another pair of stimuli which includes one of those components. Investigating the errors which see inputs of stimulus pairs misclassified as being generated by one of their component single stimuli, we find that the linear combination function leads to the mean responses to pairs which are closer (in terms of angles, and distances for fixed contrast) to the mean responses to the individual single stimuli than when the maximum or suppression functions are used. The linear summation combination function also leads to the geometric situation most favourable for classification errors to occur. We also investigated the effect of varying sparseness of the mean responses, finding that the sparser the stimuli, to lower the relative difference in error rate and coding capacity as the combination function is changed.

In order to examine the generality of our results, the second set of stimuli was tuned, with mean responses of a unit to a single stimulus input based on the difference between the orientation of the stimulus and the preferred orientation of the unit, tuned with a von Mises distribution. In contrast to the other stimuli set, the tuned responses led to much greater differences in error rates between the different types of combination function, although the overall error rate was greatly increased across the board. These results do not critically depend on the choice of tuning function. Again, we identified key sources of differences in the error rate, finding that responses to inputs of stimulus pairs were frequently misclassified as being generated by single stimuli inputs, and that these errors were the most dependent on combination function choice. Examining the situation in more detail, we find that the structure of the stimuli space leads to the response vectors of pairs of stimuli with small orientation separation distances being very close to the response vector of a single stimulus with an intermediate

orientation, with the responses to stimulus pairs generated by the linear sum combination function much closer to such optimal distractors than those generated by the maximum or suppression combination functions - we provide a mathematical intuition as to why this occurs.

Several issues can be identified with the assumptions made in our model, most of them due to the type of noise we choose. In order to ensure the model is mathematically tractable, we assume that the response to an input is corrupted by noise which is additive, uncorrelated and normally distributed. Such noise will obviously introduce a response vector length dependence on decoding accuracy - the longer the response vector, the less affected by additive noise the final response will be. Introducing the normalisation constraint removed this constraint, forcing the combination function's effectiveness to be more dependent on the 'shape' of the response function and not on its magnitude. This normalisation constraint has the effect of making the combination functions into equivalence classes up to a constant of proportionality, making direct comparisons with experimental results more difficult; in particular, the sum and average functions become inseparable in such a model. Other aspects of the model can be criticised. Whilst we assume that the additive noise is uncorrelated, experimental evidence indicates that this is not the case (Smith and Kohn, 2008). We modified the model to examine correlated noise in the case of a tuned stimulus space, verifying that uncorrelated noise was not a critically important assumption with respect our results - although other covariance structures might introduce decoding problems.

When we alter the model to remove the assumption that the response vectors are of equal length, we found that since response vectors created using a linear sum are longer than those created using maximum, suppressive or average functions, additive noise was less likely to lead to erroneous decoding of said response vectors. Given such a result, we should be still wary about concluding that the magnitude of the response vector is the the most important factor in decoding, since such a conclusion depends on the trial - by - trial variability in responses being well modelled by additive noise. Indeed, experimental evidence from spike count Fano factors generally indicate that the response variance will increase with response magnitude (Buracas et al., 1998) and a different noise model (such as a multiplicative noise model) might be more appropriate. This is likely to somewhat de - emphasise the importance of response magnitudes, since high amplitude mean responses will lead to high amplitude noise - although response vector length will likely still have some effect[2]. Furthermore, combining the

[2]If the noise used incorporates fixed Fano factors instead of multiplicative noise, then that would en-

responses to stimuli using summation could lead to problems if we incorporated some form of firing rate saturation - if the model was expanded to larger numbers of stimuli, the sort of 'washing out' that it would lead to difficulties in discriminating between different stimuli groups (although normalised, or averaged responses would not suffer from this). Nonetheless, given the geometric arguments we make, we predict the effect of combination function 'shape' will be to benefit maximum and other nonlinear combination functions whilst hampering the effectiveness of linear combination functions. A more detailed noise model might ultimately have to take into account the specific networks which could implement each function; such additional complexity would likely change the results we present here. In particular the architecture of a neural circuit which has to compute the linear sum, maximum or other type of combination structure will undoubtedly introduce complicated correlations in the noise which will make decoding more difficult.

We focused throughout this work on the MAP decoder, but qualitatively similar results were obtained with the ML decoder - note that in the limit of low noise, the decoders are identical. In general, the margins of difference in error rate across different combination functions tends to be slightly larger when the ML decoder is used than when the MAP decoder is used - this is due the rarity of the response to a single stimulus being misclassified as another single stimulus (a type of error whose rate is largely invariant across combination function choice) when the ML decoder is used. Other decoder types such as one which outputs the posterior distribution could also be considered.

Some elements of our approach could be expanded if computational resources allow it. We limit our model to consider only cases where pairs of stimuli are presented at equal contrast - this simplifies the mathematical analysis considerably. Other model setups which allow the contrasts to vary independently would be interesting to examine, although such setups could require more computational resources. We also only consider single stimuli and stimulus pairs as inputs, omitting larger groupings due to reasons of computational complexity. With suitable approximations, it might be possible to investigate groupings of more than two stimuli to see how the results we present here change. We only considered the situation in which combination functions are identical for each neuron, however, there is likely to be considerable heterogeneity in how the response to stimuli are combined from neuron to neuron (see the data col-

sure that response variance remains in constant proportion to the mean response, the standard deviation of responses will decrease relative as a proportion of the mean.

lected by Zoccolan et al., 2005 for instance). This will be challenging to investigate in an exhaustive fashion, but perhaps some sort of gradient descent process could alter the combination function of individual units in a given population step by step to try and improve a networks ability to encode different stimuli.

Finally, note that we consider only the efficacy of a coding method. We take no account as to any other possible roles for different computational roles (for instance, the model by Riesenhuber and Poggio, 1999 suggests the maximum function for introducing invariances) - it is entirely likely that coding in the brain is driven more by utilitarian considerations than coding efficiency. Consider the experimental work from Treue et al. (2000), in which it was found that the response to two overlaid moving dot patterns is equal to the sum of the responses to each moving dot pattern presented alone. They then combined three or five moving dot patterns in such a way that the scaled sum of the patterns was very close to a combination of two moving dot patterns, presenting them to subjects in a psycophysics experiment. They found that subjects reported two directions of motion present which matched the two component stimulus rather than any of the directions actually in the three of five component stimulus they were presented with. This provides strong indication that the directions of motion of moving dot patterns are represented by the same population of neurons, and that the responses to each component are combined linearly. This appears to go against what we suggest here, and could possibly be explained by a more realistic noise model as outlined above, but equally, there may be downstream considerations which require a particular form of stimulus encoding. It would be interesting to see if psycophysics methods could be used to probe other types of stimulus to investigate whether coding efficiency is the driving force behind how different types of encoding are carried out.

Despite some technical issues, which we discuss above, we feel confident that, due to the effect on the geometry of the response space, the choice of combination functions we outline will play a large part in determining the coding capacity of a population. The maximum operation has previously been proposed as a possible method of pooling inputs to aid with object recognition (Riesenhuber and Poggio, 1999), responding to the 'best match' of a range of inputs and helping to introduce response invariance over a stimulus parameter. Our work suggests another possibility, namely that the use of maximum or suppressive operations in neural populations might increase coding capacity and help keep error rates down. Whilst further work is needed to find which combination functions are most effective in this regard - a choice which depends on the stimulus space and noise model - our results make the case that the linear sum is

poorly suited as a choice. A natural prediction is that, in brain regions where there appears to be a certain degree of response normalisation, the response to pairs of stimuli presented simultaneously will be poorly modelled as being proportional to the sum of the responses to each stimulus presented singly, in comparison to a suitably chosen non liner combination such as the maximum function. This prediction relies on both the normalisation requirement and additive noise - a more detailed noise model might lead to different predictions.

Chapter 5

Synaptic depression and layered networks of spiking neurons

Neural responses in deep layers of the visual system often have onset latencies which are highly susceptible to stimulus contrast. These latencies are somewhat difficult to account for, with the latency of a response to a low contrast input in STS often of the order of several hundred milliseconds, and thus too long, and too dependent on contrast, to be explained by common neural delays such as axon propagation times, synaptic onset delays or membrane timeconstants. One possible source of such contrast - dependent latencies is the recurrent connectivity within each cortical layer, and its susceptibility to synaptic depression, a hypothesis illustrated by a recent model featuring layered networks of rate - based neurons (van Rossum et al., 2008) (the networks we present in chapter 2 and chapter 3 have similar features). In this chapter, we consider layered networks similar to those from van Rossum et al. (2008), albeit we use integrate - and - fire neurons to construct them. We verify that the results obtained from the rate - based network are still valid in the spiking network. We then examine the spiking statistics of single layer networks in light of experimental evidence indicating a response onset dependent drop in spike count variability (Churchland et al., 2010, Mike Oram, unpublished data), something which is not possible when using the existing rate - based model. We find that the spiking statistics in our model depend considerably on the parameter choices, and that the relationship between spiking statistics and synaptic depression is non - trivial. We also examine in more detail the information processing properties of single layer recurrent networks, comparing them to a purely feed - forward network. We find that recurrent connectivity does not improve a network's ability to filter out high frequency noise corrupting the input signal,

in contrast to the rate - based model, indicating that we should be careful not to assume that results obtained using rate - based neurons are always valid with similar models featuring spiking neurons.

5.1 Introduction

The visual system is often modelled in a feed - forward fashion, especially in the earliest visual areas. In V1 for example, a common approach is to model the response of a neuron to a presented image by pooling inputs over some 'receptive field', before applying a non - linearity (often a rectification) to obtain a firing rate (Hubel and Wiesel, 1962). Such feed - forward models have been extended to multiple layers - for example to model visual neurons which respond to specific objects whilst remaining invariant to size and position (Riesenhuber and Poggio, 1999). Nonetheless, the cortex is known to include high levels of recurrent connectivity.

The effects of recurrent connectivity on the responses to stimulus inputs has recently been modelled in one particular context (van Rossum et al., 2008). In particular, when a layered network, composed of rate - based units and incorporating strong recurrent connections which were subject to short term synaptic depression, the response of a deep - layer unit had a significantly larger onset latency when the network was presented with a weak input than it did when presented with a strong input. This result matches experimental data from the visual system of the Macaque monkey which finds that the response of neurons in the anterior superior temporal sulcus (STSa - in the inferotemporal cortex) presented with a low contrast stimulus (the neurons are characterised as responding selectively to faces, see e.g. Perrett et al., 1982; Oram and Perrett, 1992) has an onset latency significantly greater than that of the response to a high contrast stimulus (Oram et al., 2002; van Rossum et al., 2008). The *contrast dependent latency* of the responses generated by this model varied in magnitude according to layer, with deeper layers having much greater onset latencies, matching experimental data observed in the Macaque, in which the onset latency of neural responses in V1 are small, whilst those in deeper layers are larger (Gawne et al., 1996). The purpose of this chapter is to re - examine the model presented by van Rossum et al. (2008) when it incorporates spiking neurons as opposed to rate - based neurons. We initially confirm that the contrast dependent latency effects observed with rate - based neurons is still valid in the spiking network - provided sensible noise parameters are chosen. We then use our spiking model to investigate spike count variability, some-

thing which cannot be extracted from a rate - based model. We compare the spiking statistics to those observed experimentally, in which the variability of the spike count is found to decrease (and not in a purely rate - dependent manner) as the response to a presented stimulus begins (Churchland et al., 2010, Mike Oram, unpublished data - see also examine aspects of the spike - count matched model - see Oram et al., 1999).

Noise has long been considered of importance when investigating networks of integrate - and - fire neurons, particularly given the high variability in firing rates of neurons observed in the brain (Teich, 1989; Teich et al., 1990, 1996; Oram et al., 1999). In particular, the introduction of random fluctuations to a population of (leaky) IF neurons has been shown to allow such a population to respond to rapidly varying stimuli Knight (1972). Furthermore, injecting a noise current into populations of IF neurons has been shown to allow for the linear propagation of a rapidly fluctuating firing - rate signal through a layered network with only a small number of neurons in each layer (van Rossum et al., 2002). The possible source of noise in the nervous system has also been examined using integrate - and - fire neurons. In particular, 'balanced inputs' (Shadlen and Newsome, 1998) and the related 'balanced networks' (van Vreeswijk and Sompolinsky, 1996, 1998; Brunel, 2000), in which neurons receive both strong excitation and inhibition of equal mean, and are driven by temporary imbalances between the two sources due to stochastic fluctuations, have been found to lead to highly variable spike counts. A balanced network generating asynchronous spiking activity can use such activity to aid signal transmission and simple computations without requiring an outside noise source (Vogels and Abbott, 2005). Here however, we take the approach of examining noise as if it was externally injected (as in van Rossum et al., 2002), and ask how our network will affect such a noise source. Our hypothesis is that synaptic depression of the recurrent connections in our network will lead to reduced variability with the correlation between the spike counts of neurons in the same layer decreasing as the recurrence strength weakens. The effects on response statistics of synaptic depression has previously been examined - in particular Goldman et al. (1999, 2002) found that depressing synapses acted to reduce the positive auto - correlation present in the firing statistics of V1 neurons in a monkey viewing natural stimuli - we note such a property arise from our model when handling correlated noise (despite a different structure).

We find that whilst the depression of the recurrent synaptic connections does indeed lead to a fall in spike count variability and cross - correlation after a neural response, certain choices of noise parameter can often lead to an onset - dependent increase in

spike count variability, an issue which we discuss. We also investigate the ability of a single layer network with recurrent connectivity to filter out high frequency noise corrupting an input signal. For rate - based neurons, we find that single units are more effective at filtering out high frequency noise when they have recurrent connections. In contrast, over a wide range of parameters, when a single layer spiking network is used we find this advantage is no longer present. This finding indicates that we must take care when applying the results derived from rate - based models to networks composed of spiking neurons.

5.2 Method

5.2.1 Spiking network model

For much of this chapter, we investigate noisy, leaky, integrate - and - fire (IF) neurons (Dayan and Abbott, 2001). The IF neuron is a single compartment model with a threshold spiking mechanism, usually with a leak term which ensures its membrane potential decays to a resting value in the absence of input. The membrane potential $V_m(t)$ obeys the equation,

$$C_m \frac{dV_m}{dt}(t) = -\frac{1}{R_m}[V_m(t) - V_{rest}] + I_{inj}(t), \tag{5.1}$$

where C_m denotes the membrane capacitance, R_m the membrane resistance, V_{rest} the resting membrane potential and $I_{inj}(t)$ the injected current. If the membrane potential exceeds the threshold voltage V_{thr}, the membrane potential is reset to the reset potential V_{reset}, and a spike is emitted. The injected current consists of an input current $I_{input}(t)$, a noise current, $I_{noise}(t)$ and a synaptic current, $I_{syn}(t) = \sum_{i=1}^{n} I_i(t)$. The input current allows us to send external signals into the network. The noise currents are generated by Ornstein - Uhlenbeck processes. Each neuron can receive a noise current from two sources - an independent source and a shared source. We write that

$$I_{noise}(t) = \mu_{noise} + \sqrt{1-s}\sigma_{noise}x(t) + \sqrt{s}\sigma_{noise}z(t),$$

where μ_{noise} is the mean current, σ_{noise} is the noise standard deviation and $x(t)$, $z(t)$ are OU processes with zero mean, unit variance and a time - constant τ_{noise}. In most networks, and unless otherwise stated, s, the *shared current proportion* is equal to 0 and $x(t)$ is an OU process unique to the IF neuron in question - each neuron's noise current is thus independent, and can be written as $I_{noise}(t) = \mu_{noise} + \sigma_{noise}x(t)$. $z(t)$ is

an OU process which is identical across multiple neurons, and so setting s to be non - zero allows us to consider the impact of correlated noise (in a similar fashion to de la Rocha et al., 2007).

5.2.1.1 Depressing synapses

We investigated synapses which feature short term synaptic depression, a phenomenon in which the arrival of a spike at a synapse acts to temporarily reduce its synaptic efficacy (Markram and Tsodyks, 1996). It is modelled by a synaptic resource $P(t) \in [0,1]$ which can be conceptually thought of as the amount of available pre - synaptic neurotransmitter. An arriving action potential at time t will release a set proportion, U (the utilisation proportion), of the resource, $UP(t)$, which is then used to determine the synaptic output. The available pre - synaptic resource $P(t)$ is then decremented by this amount, $P(t) \to (1-U)P(t)$. Outside of the spike transmission times, the resource recovers exponentially to its base value P_0 with recovery timeconstant τ_{rec}. The resource thus obeys the following equation (Abbott et al., 1997; Tsodyks and Markram, 1997)

$$\tau_{rec}\frac{dP}{dt}(t) = P_0 - P(t) - \tau_{rec}UP(t)\sum_{j=1}^{n_i}\delta(t-t_j), \tag{5.2}$$

where t_j, $j \in [1,n_i]$, are the spike transmission times. The amount of neurotransmitter released by an incoming spike, UP, helps determine the strength of the post synaptic response. When we wish to investigate synapses not subject to synaptic depression, we fix P equal to 1.

5.2.1.2 Conductance synapse models

The inputs to our noisy IF neurons are provided by conductance based synapses. As we have described, the strength of the synaptic input to an IF neuron is subject to short term synaptic depression. We now detail how such inputs translate into a post - synaptic response. For the i'th such synapse, the synaptic current contribution can be written in the form

$$I_i(t) = g_i(t)[V_{rev,i} - V_m(t)], \tag{5.3}$$

where $g_i(t)$ is the synaptic conductance and $V_{rev,i}$ is the synaptic reversal potential. The synapse is one of two types - *exponential* or *difference of exponential*. For exponential synapses, g_i decays exponentially to zero with timeconstant τ_i in the absence of a spike,

and is incremented by the *synaptic weight* w_i multiplied by the amount of released neurotransmitter $U_i P_i(t)$ upon the arrival of a spike, leading to the following equation for the synaptic conductance dynamics

$$\tau_i \frac{dg_i}{dt}(t) = -g_i(t) + \sum_{j=1}^{n_i} w_i U_i P_i(t) \delta(t - t_j), \tag{5.4}$$

where $\{t_1, ..., t_{n_i}\}$ are the times of the spike arrivals at the i'th synapse, and δ is the delta functional.

For difference of exponential synapses, the synaptic conductance rise from an incoming spike is instead non - instantaneous. We split the synaptic conductance into two components, $g_i(t) = g_i^+(t) - g_i^-(t)$, which obey the following dynamical equation

$$\tau_i^{\pm} \frac{dg_i^{\pm}}{dt}(t) = -g_i^{\pm}(t) + \sum_{j=1}^{n_i} w_i U_i P_i(t) \delta(t - t_j). \tag{5.5}$$

Setting $\tau_i^+ > \tau_i^-$ ensures that an input spike will cause the conductance g_i to rise initially with timeconstant τ_i^-, and, over longer time periods, decay exponentially with timeconstant τ_i^+. Switching AMPA and GABA synapses from single exponential models to double exponential models, with an appropriate rising timeconstant, does not qualitatively affect the results we present below.

5.2.1.3 Poisson input units

Inputs to our network take the form of spikes generated by a number of rate modulated Poisson processes. In such a process, the probability of a spike being generated in the time interval $[t, t + \delta t]$ approaches $r(t)\delta t$ as $\delta t \to 0$ for firing rate $r(t)$. These *Poisson input units* make synaptic connections to noisy IF neurons with $r(t)$ representing the input signal. When we incorporate absolute refractory periods of 1 ms into the Poisson process, results are similar to those displayed below - likely since the firing rates used are relatively low.

5.2.1.4 Model implementation

A few techniques were used to speed up computational implementation of network of noisy integrate and fire neurons connected by depressing synapses, the most notable of which we describe here. Simulations requiring many parallel jobs to be run simultaneously made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF). (http://www.ecdf.ed.ac.uk/). The ECDF is partially supported by the eDIKT initiative (http://www.edikt.org.uk).

Grouping synapses Recall that upon receipt of a spike, the conductance of an exponential synapse is increased instantaneously by some input. It then decays exponentially with time. Assume now that one has a collection of such synapses, all impinging upon the same IF neuron - we can use a common approach to speed up the simulation of such synapses if we collect synapses with identical reversal potentials and timeconstants. For a group of exponential synapses with identical reversal potential V_{rev}, we can use Eq. (5.4) to write down the total synaptic current contribution

$$I_{group}(t) = [V_{rev} - V_m(t)] \sum_{i=1}^{m} g_i(t).$$

We then use Eq. (5.4) to write down how $g(t) := \sum_{i=1}^{m} g_i(t)$ varies with time - if we assume that $\tau_i = \tau$ for all such synapses, this simplifies to the following equation

$$\tau \frac{dg}{dt}(t) = -g(t) + \sum_{i=1}^{m} \sum_{j=1}^{n_i} w_i U_i P_i(t) \delta(t - t_j).$$

This allows use to simulate just one conductance variable in the place of several. A similar approach can be used for double exponential synapses. The synaptic groupings we choose emerge naturally:

- We simulate AMPA activated excitatory synapses using the exponential model with a reversal potential of 0 mV and a 5 ms timeconstant.
- We simulate GABA activated inhibitory synapses using the exponential model with a reversal potential of - 70 mV and a 5 ms timeconstant.

In a similar fashion, we group the output from one neuron's synapses together where possible - in particular, if an IF neuron makes synaptic connections with all the same depression parameters and initial synaptic resource, then the calculation of the synaptic depression across all synapses can by reduced to a single update - this dramatically reduces the number of times an exponential function needs calling, providing a large boost in simulation speed.

Network layout, tuning and inputs We investigate spiking neurons arranged in layers, following the models previously used to investigate contrast dependent latency effects with rate - based neurons presented in van Rossum et al. (2008). Each network contains *layers*, each of which contains one or more *populations* of IF neurons which correspond to rate - based units. Each neuron has resting membrane potentials of - 60 mV, reset membrane potentials the same, absolute refractory periods of 1 ms, threshold

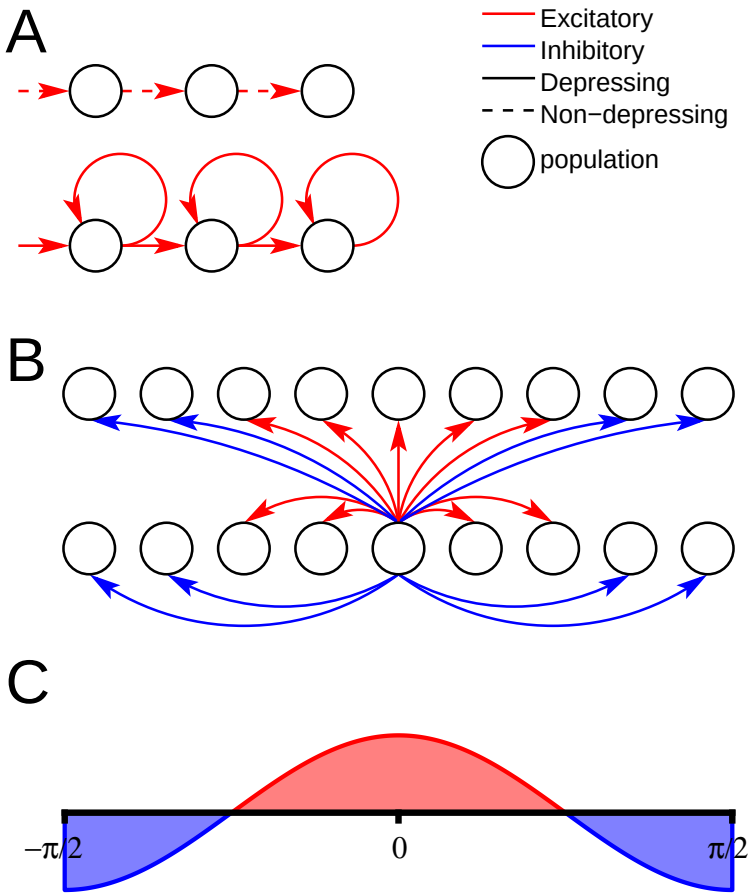

Figure 5.1: Model setup. **A** Two simple networks - the *feed - forward* network (top) in which three populations are connected in a layered structure, with each layer innervating the next using a non - depressing synaptic pathway, and the *recurrent* network (bottom) in which each population also innervates itself by an excitatory synaptic pathway, and all synapses are now subject to short term synaptic depression. **B** Synaptic connections for a *population coding network* - each population of neurons has an orientation label and excites populations in the same layer and the target layer with similar orientations (red arrows) whilst (directly) inhibiting those with dissimilar orientations (blue arrows). 80% of the neurons in each population are excitatory and 20% are inhibitory. **C** The recurrent connectivity in each layer of the population coding network uses a cosine - shaped weight profile.

membrane voltages of - 50 mV, membrane resistances of 100 MΩ and a membrane capacitance's of 200 pF. We use the term *synaptic pathway* to refer to a group of synapses linking one layer of neurons to another. The network structures used are as follows:

- ***Feed - forward network:*** The feed - forward network consists of three or more layers, each containing a single population of 200 IF neurons. The first layer receives input from either a direct current injected into each neuron or from 200 Poisson input units with 10% connection probability via non - depressing excitatory AMPA synapses. Avoiding all - to - all connectivity helps with computational complexity when considering larger models or extended time period simulations, but connection percentages of 5% and 20%, as well as all - to - all connections have also been considered, with few apparent changes - this also applies to recurrent networks. Each subsequent layer receives inputs from the previous layer in a similar way. The synaptic weights (and, where present, the strength of the input current) are adjusted to ensure that a 50 Hz input (or, equivalently, a 100 pA current injection) leads to a peak firing rate of 50 Hz in each layer. This layout is pictured in Fig. 5.1A.

- ***Recurrent network:*** The recurrent network (pictured in Fig. 5.1B) is identical to the feed - forward network except that each synapse is now subject to short term synaptic depression (with $U = 0.2$ and $\tau_{rec} = 200$ ms) and recurrent synaptic pathways have been added to each layer. These recurrent pathways are formed with 10% connection probability (self - recurrence is not permitted). Again the synaptic weights (and, where present, the strength of the input current) are adjusted to ensure that a 50 Hz input (or, equivalently, a 100 pA current injection) leads to a peak firing rate of 50 Hz in each layer - the strength of the synaptic weight of the feed - forward pathway arriving at each layer is kept in proportion to the synaptic strength of the recurrent pathway entering that same layer - the synaptic weight in the recurrent pathway is equal to either 80%, 160% or 200% of the synaptic weight in the feed - forward pathway. Note that, without synaptic depression, the maximum strength of the recurrent pathway is limited, as otherwise the network is vulnerable to runaway excitation - with depressing synapses this is not an issue since high firing rates automatically reduce the efficacy of the synaptic strength.

- ***Population coding network:*** A population coding network is also arranged in layers. Each layer contains 20 populations of neurons, each of which contains

12 excitatory neurons and 3 inhibitory neurons (80% excitatory - larger population sizes give similar results, but are more computationally challenging to simulate). Each population is assigned an orientation, taking a position in a ring - like structure. Again, the synaptic pathways from one layer to the next (and to itself) are formed with 10% connection probability, only now the weight of these connections varies. Fig. 5.1B illustrates how these connections are formed - excitatory neurons form strong connections (red arrows) when their targets are in populations with a similar orientation label whilst strong inhibitory connections (blue arrows) are formed when the orientations are maximally different. To scale the connection weights, the cosine of the orientation difference is used, pictured in Fig. 5.1C (negative indicates inhibition). The inhibitory synapses were further strengthened by a factor of 20 to take into account the relative number of inhibitory neurons and the small difference between the inhibitory reversal potential and the IF neuron membrane potential range. Again, the weights were scaled to ensure a 50 Hz input leads to a 50 Hz peak firing rate in each subsequent layer - the recurrent strength were 80% of the strength of the feed - forward inputs. We simulated feed - forward networks by removing the recurrent connections and using non - depressing synapses in the feed - forward pathways.

The first layer of each network receives inputs from either Poisson input units - of the same size as the first layer - or via a direct current injection. Inputs are presented for 100 ms at different strengths (and, where relevant, orientations) in randomly generated order. There are 900 ms of recovery time between input pulses during which the input trains are set at a 5 Hz background firing rate (or the equivalent amplitude for direct current inputs).

5.2.2 Signal propagation and noise filtering measures

For layered networks of spiking neurons, we often need a measure of how effectively an input signal is propagated through to the its final layer. Previously used examples are *dissimilarity* and *similarity*, both of which provide measures of linearity (van Rossum et al., 2002; Vogels and Abbott, 2005). For a time varying, non - negative input signal $s(t)$, an output firing rate $r(t)$ (calculated from the binned spiking activity of several neurons), the dissimilarity $D(r,s,\triangle t)$, for a set time delay $\triangle t$, is defined by the following equation

$$D(r,s,\triangle t) = \left| \frac{r(t-\triangle t)}{|r(t)|_2} - \frac{s(t)}{|s(t)|_2} \right|_2^2.$$

Here $|...|_2$ is the L^2 norm (defined by $|f|_2^2 := \int_{-\infty}^{\infty} |f(x)|^2 dx$) and so $D(r,s,\triangle t) = 0$ iff $r(t-\triangle t) \propto s(t)$ is true for almost all t. Typically, $\triangle t$ is varied to find the optimal dissimilarity, $\min_{\Delta t} D(r,s,\Delta t)$, (essentially Δt corresponds to the time taken to propagate the signal through the network) - the smaller $\min_{\Delta t} D(r,s,\Delta t)$ is, the better the network is judged to perform (van Rossum et al., 2002).

Another method of judging effectiveness is the similarity measure, $\max_{\Delta t} \mathrm{corr}(r,s,\triangle t)$ which is defined as the maximum over Δt of the correlation between the variable $r(t-\Delta t)$ and $s(t)$, as t varies. Clearly $\mathrm{corr}(r,s,\triangle t) = 1$ iff $r(t-\triangle t) - \bar{r} \propto s(t) - \bar{s}$ (with positive proportionality constant) almost everywhere, for $\bar{r}$ and $\bar{s}$ the temporal averages of $r(t)$ and $s(t)$. Again, $\triangle t$ is taken to be the signal propagation delay (Vogels and Abbott, 2005).

An alternate measure of network effectiveness when it comes to signal propagation makes use of information theory. We are interested in knowing how much the output signal from a network can tell us about the input which is presented. The key quantity used is *mutual information*. Consider a pair of random variables, R and S, then the mutual information between them $I(R;S)$ is given by the following formula (Shannon, 1948),

$$I(R;S) = \sum_{r,s} p(r,s) \log_2 \left(\frac{p(r,s)}{p(r)p(s)} \right), \tag{5.6}$$

where $p(r,s)$ is the probability density function for R and S, $p(r)$ is the probability density function for R and $p(s)$ is the probability density function for S. The mutual information between two variables is a measure of how much knowledge of one variable informs you about the other - a high value indicates that we can use a recorded response to recover an input accurately. Since the input signal $s(t)$ has a probability distribution independent on t it is reasonable to treat $r(t-\Delta t)$ and $s(t)$ as the observations of random variables $R_{\Delta t}$ and S (the response of the network and the input signal) over 'trial' t. We estimate $p(r,s)$ and thus define the *delayed mutual information* measure $I_{delay}(r,s,\Delta t)$ by the equation

$$I_{delay}(r,s,\Delta t) := I(R_{\Delta t};S).$$

Typically, we require a long simulation run to obtain a smooth estimate of $p(r,s)$. We are typically interested in the mutual information, maximised across all possible delay

times Δt, $\max_{\Delta t} I_{delay}(r,s,\Delta t)$. Again, the optimal delay time $\text{argmax}_{\Delta t} I_{delay}(r,s,\Delta t)$ is taken to be the signal propagation time.

When modelling the presentation of a stimulus for a fixed time period we are interested not in the amplitude (or contrast) of the input, but rather whether we can detect its presence or absence from the output signal; we represent this by the binary variable S, and set $S = 0$ if no input stimulus is presented during a trial and $S = 1$ if an input signal is presented during the trial. In such situations, rather than measuring the mutual information between a continuous input signal and the firing rate response, we instead measure the mutual information between the firing rate and the presence of the stimulus S - we choose to examine S rather than the contrast to de - prioritise the linearity of the input - output relationship. We denote mutual information between the random variable S and the parametrised random variable $r_{\delta t}(t)$ - the average firing rate of the network over $[t-\delta t, t]$, for t the stimulus relative time, as $I(r_{\delta t}(t);S)$. When $I(r_{\delta t}(t);S) = 1$, then the local firing rate $r_{\delta t}(t)$ can entirely inform about whether a stimulus has been presented to the network or not in a given trial. We expect $I(r_{\delta t}(t);S) = 0$ for $t < 0$ and $t \gg 100\,\text{ms}$ (for stimulus presentations of 100 ms duration) - the firing rate will clearly only inform about the presence or absence of an input stimulus during the network response. Note also that the size δt of the bin which we use to calculate $r_{\delta t}(t)$ can be important - when it is large, the extended integration period can help average out fast noise fluctuations - we talk about the effects of using a different bin size in the results.

5.2.3 Latency measures and statistics

We measure several aspects of the output of the network simulations we run. The *PSTH*, or peri - stimulus time histogram, in response to an input at stimulus onset relative time t, $r_{PSTH}(t)$, is the average over all trials of the spike count of a given neuron from t to $t + \Delta t$ (we use $\Delta t = 1$ ms unless otherwise stated) in response to that stimulus. The *latency* in response to an input stimulus is measured using the half - max method - i.e. it is defined as the time after stimulus presentation it takes for the neuron to respond at half its maximum firing rate over the trial, after the background firing rate has been subtracted. For the latency of the response of a population of neurons to an input stimulus, we construct PSTHs for each neuron, before taking the average across neurons to give a population PSTH. The population PSTH is then used to read the latency using the half - max method. We note that measuring the half - max latency from each individual PSTH and then averaging these latencies across the population

give similar results.

In addition to measuring the mean firing rate profiles and response latencies, we also measure three spiking statistics - Fano factor, sequential, or auto, correlation and cross - correlation. Suppose neuron i has $C_i(J_j)$ spikes in time interval J_j from interval set $J = \{J_1, ..., J_m\}$.

- The *Fano factor* of the spike count of neuron i over the interval set J is the variance of the spike count divided by the mean, written

$$Fano(i,J) = \frac{\text{var}_j(C_i(J_j))}{\langle C_i(J_j)\rangle_j},$$

where $\text{var}_j(C_i(J_j))$ is the variance of the spike count and $\langle C_i(J_j)\rangle_j$ is the spike count mean.

- The *sequential (auto) correlation* of the spike count of neuron i over the interval set J for delay T is equal to correlation between the set of spike counts $\{C_i(J_j)\}_j$ and the set $\{C_i(J_j - T)\}_j$, and is written

$$Seq(i,J,T) = \frac{\text{cov}_j(C_i(J_j), C_i(J_j - T))}{\sqrt{\text{var}_j(C_i(J_j))\text{var}_j(C_i(J_j - T))}},$$

where $\text{cov}_j(C_i(J_j), C_i(J_j - T))$ is the covariance between the spike count measured using interval set J and the spike count measured using intervals $J - T$.

- The *cross - correlation* between the spike count of neuron i and the spike count of neuron k over the interval set J is equal to correlation between the set of spike counts $\{C_i(J_j)\}_j$ and the set $\{C_k(J_j)\}_j$and is written as

$$Cross(i,j,J,T) = \frac{\text{cov}_{\hat{j}}(C_i(J_{\hat{j}}), C_j(J_{\hat{j}}))}{\sqrt{\text{var}_j(C_i(J_{\hat{j}}))\text{var}_j(C_j(J_{\hat{j}}))}}.$$

These statistical measures depend on the set of time intervals J over which they are measured. In general, we will align these interval to stimulus presentations. For stimuli presented at $\{t_1, ..., t_m\}$, post - stimulus time t and bin size Δt we define the set $J_{\Delta t}(t)$ of stimulus relative intervals as follows:

$$J_{\Delta t}(t) = \{[t_j + t, t_j + t + \Delta t) \,\text{s.t.}\, 1 \leq j \leq m\}.$$

The Fano factor, sequential correlation and cross - correlation of the spike count of neuron i over interval set $J_{\Delta t}(t)$ are denoted $Fano_{i,\Delta t}(t)$, $Seq_{i,\Delta t}(t)$ and $Cross_{ij,\Delta t}(t)$ respectively - we denote statistics measured from such windows as being generated by a *sliding window*. Typically, we use bin size of $\Delta t = 50\,\text{ms}$ for such statistics.

5.2.4 Rate based model

We also use a rate based model of a single neural population receiving a current injection input which features short term synaptic depression. Following previous approaches to rate coded models featuring synaptic depression (see chapter 3 or Tsodyks et al., 1998; van Rossum et al., 2008), its dynamics are described by the following dynamical system

$$\begin{aligned} \tau_0 \frac{dI}{dt}(t) &= -I(t) + w_{ff} I_{input}(t) + w_{rec} P(t) r(t), \\ \tau_{rec} \frac{dP}{dt}(t) &= 1 - P(t) - \tau_{rec} U P(t) r(t), \\ r(t) &= \max(I(t), 0). \end{aligned} \tag{5.7}$$

where $I(t)$ is the unit's net current which recovers with timeconstant τ_0, $I_{input}(t)$ is the input current, w_{ff} is the strength of the feed - forward input current, w_{rec} is the strength of recurrent connections, $P(t)$ is the synaptic resource which recovers with timeconstant τ_{rec}, U is the utilisation proportion and $r(t)$ is the instantaneous firing rate - we set $\tau_0 = 5\,\text{ms}$, $\tau_{rec} = 200\,\text{ms}$ and $U = 0.2$. We have two typical setups for this network - as a feed - forward network with $w_{ff} = 1$ and $w_{rec} = 0$; and a recurrent network with $w_{ff} = 0.3$ and $w_{rec} = 1$. The parameter choices ensure that a sustained input of 50 Hz leads to a peak output firing rate of 50 Hz.

5.3 Results

5.3.1 Contrast dependent latencies

We first investigate how onset latency of the response of layered networks of IF neurons to an input stimulus depends on contrast and stimulus identity, comparing the results we find to those of an existing rate - based model (van Rossum et al., 2008). We use a layered network of integrate - and - fire (IF) neurons and compare two principle types of network; feed - forward networks with non - depressing synaptic pathways, and recurrent networks which includes synapses subject to short - term synaptic depression and in which each layer innervates itself as well as the next layer. These networks are described in detail above, and illustrated in Fig. 5.1A. Each noisy IF neuron was injected with a noise current generated by an independent OU - process with standard deviation of 70 pA and a 2 ms time constant. The mean current is adjusted to ensure a background firing rate of 5 Hz - the resultant mean current compares well with noise

parameters previously observed to lead to effective signal propagation over multiple layers (van Rossum et al., 2002).

The first layer of the network receives inputs from a population of 200 Poisson input units with pulses of activity of varying input strengths, each of 100 ms in length - we use different amplitude inputs as a stand in for different stimulus contrasts. The strengths of the synaptic connections entering each layer are adjusted to ensure that a 50 Hz input pulse leads to a peak firing rate of 50 Hz in each layer (see Methods section for more details).

We presented inputs at a range of input strengths (from 20 Hz to 50 Hz in 10 Hz steps), to the feed - forward network and two recurrent networks (one with the recurrent strength equal 80% of the feed - forward strength and the other with the recurrent strength equal to 160% of the feed - forward strength). Fig. 5.2A, B and C show the PSTH in response to these input strengths, averaged over neurons in the third layer of these networks. The relationship between the strength of the input and the peak firing rate of the responses in the third layer of either recurrent network is highly nonlinear (not shown), in contrast to the feed - forward network where the amplitude of the responses in the third layer are approximately linearly related to the input strengths. The non - linearity of the recurrent network becomes more pronounced as the relative strength of the recurrence in strengthened and motivates the choice of the number of layers in our networks - in deeper layers, the response becomes all - or - nothing.

The latency of the responses in the third layer of the recurrent networks varies with input strength (Fig. 5.2E and F), with a response to a weaker input showing an increased onset latency than the response to a stronger input, in a similar fashion to the corresponding rate - based model and the experimental evidence (van Rossum et al., 2008). For strong inputs, a rapid response onset is accompanied by an early peak in the firing rate, as the recurrent connections depress out rapidly, leading to relatively short latency responses. In contrast, the lower amplitude response to a weak input depresses the feedback connections more slowly - these recurrent connections act to integrate the incoming stimulus and amplify it over time, leading to a longer response latency. These contrast dependent latency effects are also present in a weaker form when using feed - forward networks, Fig. 5.2D. This is probably due to the filtering effect from both the synapses and the membrane potential - the effect is much larger than with the equivalent rate - based network - this is possibly due to the longer membrane timeconstant.

Experimental evidence suggesting that response latencies in the visual system were

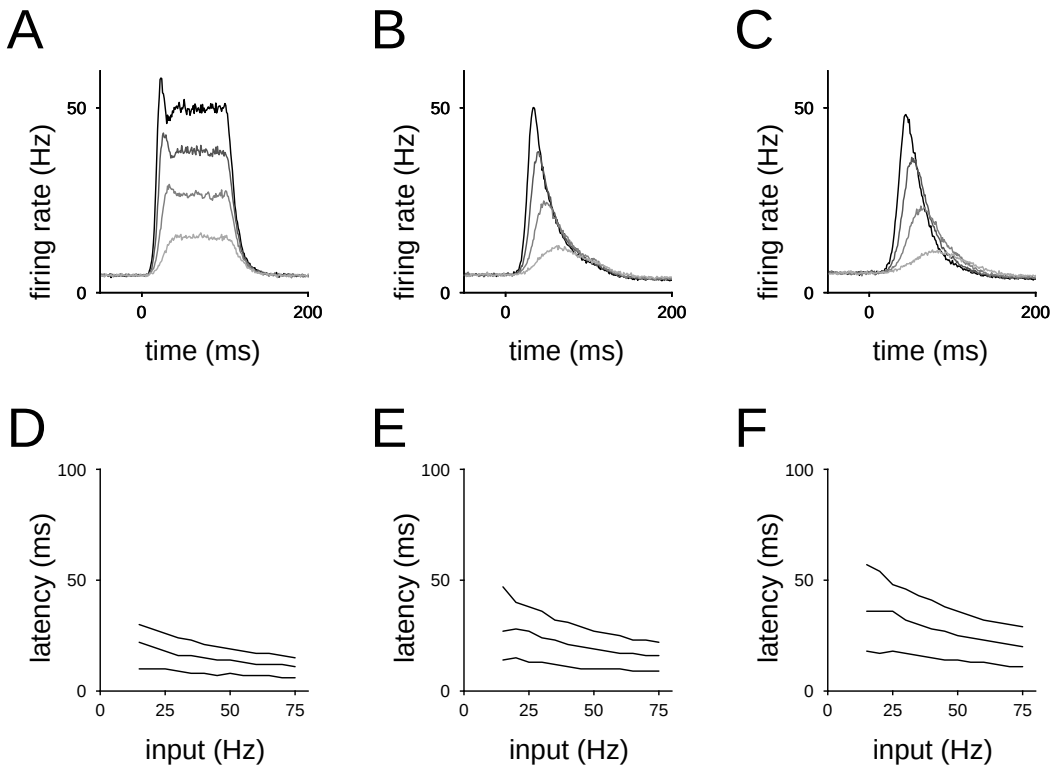

Figure 5.2: **A** The PSTHs in the third layer of the feed - forward network in response to inputs of different strengths (of 20 Hz, 30 Hz, 40 Hz and 50 Hz) created using bins of 1 ms duration. **B** As A but with the PSTHs created using the third layer of the recurrent network in which the recurrent strength is equal to 80% of that the feed - forward strength. **C** As B, but now the recurrent strength is equal to 160% of the feed - forward strength. **D** The onset latency (measured using the half - max approach) of the population responses in the first three layers of the feed - forward networks is plotted against the input stimulus strength - curves for earlier layers are plotted lower, with earlier response latencies. **E** as D but now the recurrent network from panel B is used. **F** as E but now the network from panel C is used.

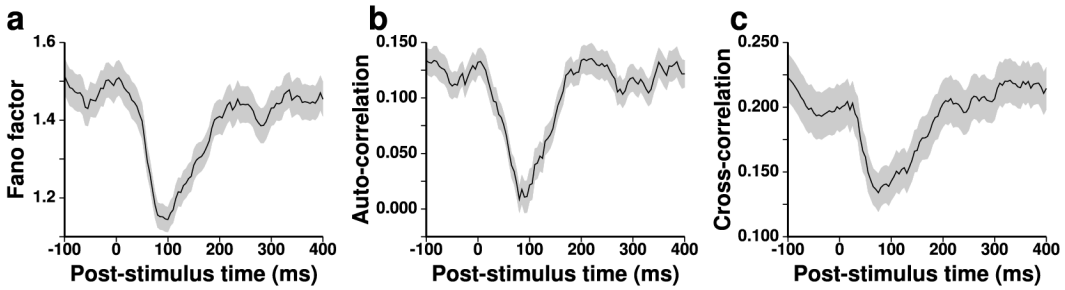

Figure 5.3: The Fano factor (**a**), sequential (auto -) correlation (**b**) and cross - correlation (**c**) between neuron pairs for the spike count in 50 ms windows, relative to the stimulus presentation time of neurons in the macaque STS (data and graph from from Mike Oram - unpublished).

strongly dependent on stimulus contrast, but only weakly dependent on stimulus choice (Carandini and Heeger, 1994; Albrecht, 1995; Gawne et al., 1996), was modelled in van Rossum et al. (2008) by a rate - based population coding network. We verified that a spiking analogue of this network leads to similar results (see Appendix section 5.A.1 for details). Additionally, responses of the rate - based model in van Rossum et al. (2008) to flashed stimuli were similar to the responses to stimuli presented for longer time periods - this was also true for the spiking networks presented here (not shown). Overall, with respect to the response latency, the results from the spiking model are are broadly comparable with those obtained when using the rate - based model, at least with noise parameters that put the system in the 'rate mode' (described in van Rossum et al., 2002).

5.3.2 Spike count statistics

Experimental evidence has indicated that the *Fano factor*, *sequential correlation* and *cross - correlation* of the neural spike counts collected by a *sliding window*, (which we denote $Fano_{i,\Delta t}(t)$, $Seq_{i,\Delta t}(t)$ and $Cross_{ij,\Delta t}(t)$, for t the time relative to stimulus presentation and Δt the bin size from which the spike counts are collected), falls with the stimulus response onset (see Fig. 5.3 - Mike Oram, unpublished data; also reported by Churchland et al., 2010). We examine the spike count statistics recorded from networks of spiking neurons featuring depressing recurrent connectivity with the hypothesis that depression in the recurrent connections as a result of the response to a stimulus will reduce spike count variability and cross - correlation measured from pairs of neurons in the network. This should follow naturally, as the reduced recurrent

strength following depression will weaken the ability of firing rate fluctuations in individual neurons to contribute to variability in the rest of the network. We focus on the first layer of the recurrent network depicted in Fig. 5.1A, using the first layer of the feed - forward network as a comparison[1]. As a prelude to investigation the spike count statistics of recurrent networks, we first examine noisy IF neurons in the absence of synaptic connectivity.

5.3.2.1 Spike count statistics in noisy IF neurons

To provide context when we examine the spiking statistics of networks of noisy IF neurons, it is useful to first examine the spiking statistics of pairs of noisy IF neurons in the absence of any network interactions - i.e. with no synapses active and activity driven only by the injected noise currents. We inject pairs of IF neurons with noise currents generated by OU processes with mean μ_{noise}, standard deviation σ_{noise} and timeconstant τ_{noise} as we measure the Fano factor, sequential correlation and cross - correlation of their spike counts. In order to examine cross - correlation in a meaningful fashion, the two noise currents injected into each IF neuron are composed of individual current contributions $I_{noise}(i,t)$, for neuron $i = 1,2$, and a noise current shared between both neurons, $I_{shared}(t)$ - a situation illustrated in Fig. 5.4A and which has been used elsewhere (de la Rocha et al., 2007). The proportion of noise current variance which is shared between neurons is defined by the value s and is investigated for positive values whilst keeping the total variance of the noise currents fixed (see Methods for details).

In the left panels of Fig. 5.4B, C and D, we plot the mean Fano factor (across all neurons) of the spike counts against the average firing rate[2] for different magnitudes of σ_{noise} (black indicates σ_{noise} is equal to 10 pA; each lighter curve increments this by 10 pA). In a similar fashion, we plot the mean sequential correlation of the spike counts (the correlation between the number of spikes in consecutive windows) in the

[1] In the deeper layers of the spiking networks which we investigate, spike counts across multiple neurons become more correlated as the layer depth increases (not shown) - this is natural, as small fluctuations in initial layers can be amplified by subsequent layers, since the structure of the models' connectivity from one layer to the next ensures each neuron receiving inputs only from neurons in the previous layer with similar response properties and whose spike count statistics are positively correlated. A more natural structure for deep layers might be to ensure that inputs from a previous layer arise from a range of sources (such as with the model proposed by Riesenhuber and Poggio, 1999), in which the correlations between inputs to a neuron and the relationship between the response characteristics of those inputs are likely more complex.

[2] There is a monotonically increasing relationship between the IF neuron's average firing rate and the noise current mean μ_{noise}. When we plot the Fano factor, sequential correlation and cross - correlation, we do so against the average firing rate rather than the mean noise current μ_{noise}, to help with comparisons across different noise parameters.

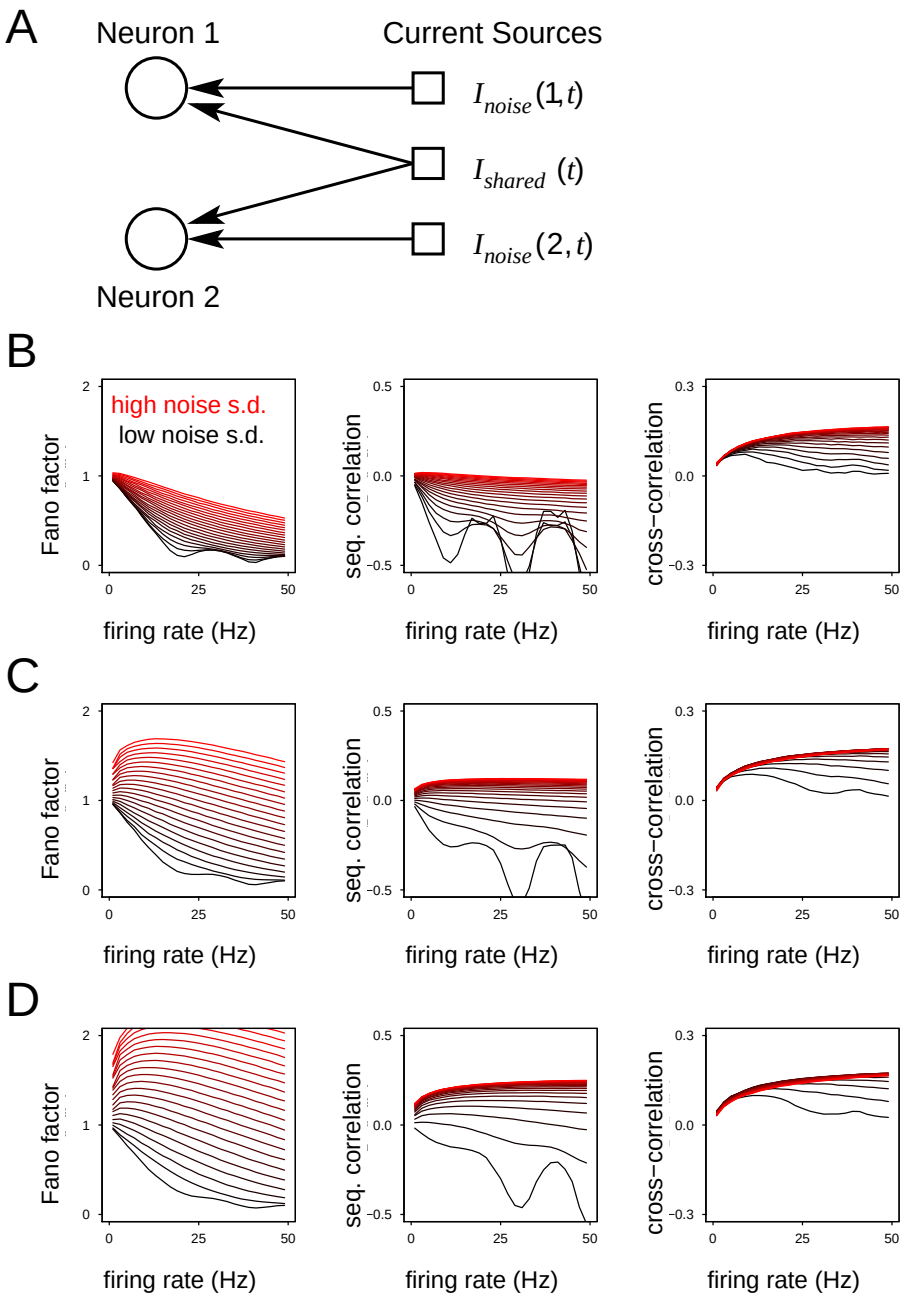

Figure 5.4: **A** A diagram of the different noise current sources received by two IF neurons (circles). The i'th neuron receives two noise current contributions - its own individual current $I_{noise}(i,t)$, with variance $0.8\sigma^2_{noise}$ and a common current shared across all neurons, $I_{shared}(t)$, with variance $0.2\sigma^2_{noise}$, for total noise current variance σ^2_{noise}. **B** The Fano factor (*left panel*), sequential correlation (*central panel*) and cross - correlation (*right panel*) of spikes produced by noisy IF measured using 50 ms windows over a range of firing rates. The black line indicates that the noise current has a low standard deviation $\sigma_{noise} = 10$pA - this is increased in 10 pA steps, plotter with lighter lines, with the timeconstant of the noise currents (generated by OU - processes) set at 2 ms. **C** as B, but the timeconstant of the noise is 10 ms. **D** as B, but the timeconstant of the noise is 20 ms.

central panels and the mean cross - correlation of spike counts between neuron pairs (we set $s = 0.2$, indicating that 20% of the noise current variance is shared between each neuron pair) in the right hand panels. The timeconstant of the OU process used to generate the noise was 2 ms in row B, 10 ms in row C and 20 ms in row D.

In general, we find that raising the standard deviation σ_{noise} of the noise also increases the Fano factor - which is unsurprising as increasing the variability in the noise current whilst holding the firing rate fixed will naturally lead to an increase in spike count variability. Additionally we find that increasing the firing rate (by increasing μ_{noise}) decreases the Fano factor, whilst increasing τ_{noise} increases the Fano factor:

- To generate a high firing rate requires a noise current with a high mean, μ_{noise}. When μ_{noise} is increased and σ_{noise} remains constant, the noise current becomes dominated by its mean, leading to more regular spike trains, with spike counts thus of reduced variability. For the parameters which we consider (Fig. 5.4B, C and D, left panels), for high firing rate, (i.e. as the firing rate is increased towards 50 Hz), the Fano factor decreases with increasing firing rate. This also remains true for a variety of other parameter regimes (not shown). It is also worth noting that as firing rates decrease towards 0, the dependence of one spike on the time gap from the previous one will drop, leading to Poisson - like statistics (and Fano factors of 1).

- When the timeconstant is fast ($\tau_{noise} = 2\,\text{ms}$) the refractory effect after a spike (arising both from the absolute refractory period, and the effect of resetting the membrane potential to its resting state) greatly reduces the probability with which clusters of spikes appear - a high noise current which leads to a spike will typically have decayed by the time the refractory period has ended. In contrast, we find that as the timeconstant is increased, the Fano factor increases - slower varying noise currents lead to greater opportunity for clustering of spikes and thus greater variability in spike counts.

We turn to the sequential correlation, finding that increasing the firing rate of the IF neurons, which acts to increase the impact of spike train refractoriness, will decrease the spike count sequential correlation. We find also that the increased clustering effect that accompanies larger τ_{noise} will act to increase sequential correlation:

- For high firing rates the presence of a spike close to the end of a time window can delay the first spike in an immediately following second time window due

to refractory effects - this effect acts to anti - correlate the spike count in subsequent windows. Additionally, the periodicity of the sequential correlation for low σ_{noise} (as firing rate is varied) likely arises from the spike train regularity which occurs when the noise current varies little. This is indicated by the fact that when the time windows we use to count spikes is equal to a multiple of the mean inter - spike interval (such as at 20 Hz or 40 Hz when windows of 50 ms are used) the sequential correlation is higher than when the counting windows are either slightly shorter or slightly longer.

- For large τ_{noise}, especially for high σ_{noise}, spikes tend to be generated in clusters - if the noise current is strong enough to depolarise an IF neuron and trigger a spike, then the slow recovery dynamics can keep the current high enough to cause repeat spikes after the absolute refractory period has worn off and the membrane potential has risen from its reset value. This clustering effect acts to correlate the spike counts in consecutive counting bins. Over the range of firing rates we consider (i.e. 50 Hz or less), when τ_{noise} and σ_{noise} are both high (Fig. 5.4C and D, lighter lines), this effect more than counteracts the tendency of spike trains to regularise. As the firing rate is increased beyond 50 Hz (up towards 300 Hz), for fixed τ_{noise} and σ_{noise}, the sequential correlation will eventually begin to decrease with increasing firing rate.

We also verify that the cross - correlation between spike counts from pairs of IF neurons rises with firing rate, in line with existing evidence (de la Rocha et al., 2007). As σ_{noise} is increased for a fixed firing rate, the cross - correlation of spike counts from neuron pairs appears to approach some limiting value - whilst we plot this for pairs of IF neurons with $s = 0.2$, similar results appear to be true as the percentage of shared variance is changed. Overall, for a fixed firing rate and τ_{noise}, increasing the noise standard deviation σ_{noise} increases all three spike count statistics - Fano factor, sequential correlation and cross correlation.

As well as considering noisy IF neurons, it is also helpful to examine the spike count statistics of spike trains generated from an existing model - the spike - count matched model (Oram et al., 1999)[3]. The spike - count matched model was introduced in an attempt to explain the high frequency of precise spike patterns observed in experimental data - that is, higher frequency than can be explained by a non - homogeneous

[3]We should note that the most common model for generating spike trains, the Poisson process, leads to Fano factors always equal to 1, with sequential correlation fixed at 0 due to the memory - less - ness of the Poisson process.

Poisson processes. The key feature of the spike - count matched model is to match the overall Fano factor from each neuron (by using actual spike counts in the model) - the spikes in each trial are assigned to time points using that neuron's recorded PSTH as a probability density function. This process is repeated twice to ensure that the generated spike train has the same inter - spike interval (ISI) distribution as the recorded spike train - see Oram et al. (1999) for details. We consider a simplified version (which enables some straightforward mathematical analysis) in Appendix 5.A.2 which generates the total number of spike in a trial using a random variable (rather than using recorded data) and omitting the stage dealing with the ISI correction. We find that when the rate is kept constant, the sequential correlation increases monotonically with increasing Fano factor. Additionally, when a slowly varying PSTH is used, a high firing rate is accompanied by a high windowed sequential correlation and windowed Fano factor (for an overall Fano factor greater than 1), unlike the situation observed experimentally which features stimulus onset dependent drops in variability and sequential correlation (Churchland et al., 2010, Mike Oram, unpublished data Fig. 5.3). We conclude that the spike - count matched model is unable to explain the observed variability and sequential correlation changes. A modified version, which treats the number of spikes before stimulus onset and after stimulus onset separately (rather than only counting spikes over the entire trial) could account for these differences (Mike Oram, in preparation).

5.3.2.2 Spike count statistics in recurrent networks

We now consider how recurrent synaptic connectivity subject to short term synaptic depression impacts on spike - count statistics. To do this, we investigate single layer networks of IF neurons featuring recurrent connectivity, equivalent to the first layer of the recurrent network plotted in Fig. 5.1. We examined the Fano factor, sequential correlation and cross correlation of spike counts collected using a sliding window / consecutive sliding windows from the neurons in these networks (denoted, respectively, $Fano_{i,50\text{ ms}}(t)$, $Seq_{i,50\text{ ms}}(t)$ and $Cross_{ij,50\text{ ms}}(t)$ for t the time relative to stimulus presentation - we set the sliding window size $\Delta t = 50$ ms), in response to a strong input stimulus (50 Hz amplitude for 100 ms duration). We compared the windowed spike count statistics to those generated by a single layer of neurons receiving (non - depressing) feed - forward input only, equivalent to the first layer of the feed - forward network plotted in Fig. 5.1[4]. We consider two types of recurrent network - one network

[4]The first layer of the feed - forward network depicted in Fig. 5.1, containing IF neurons with no synaptic inter - connections, is similar to the situation outlined in section 5.3.2.1 - the only difference

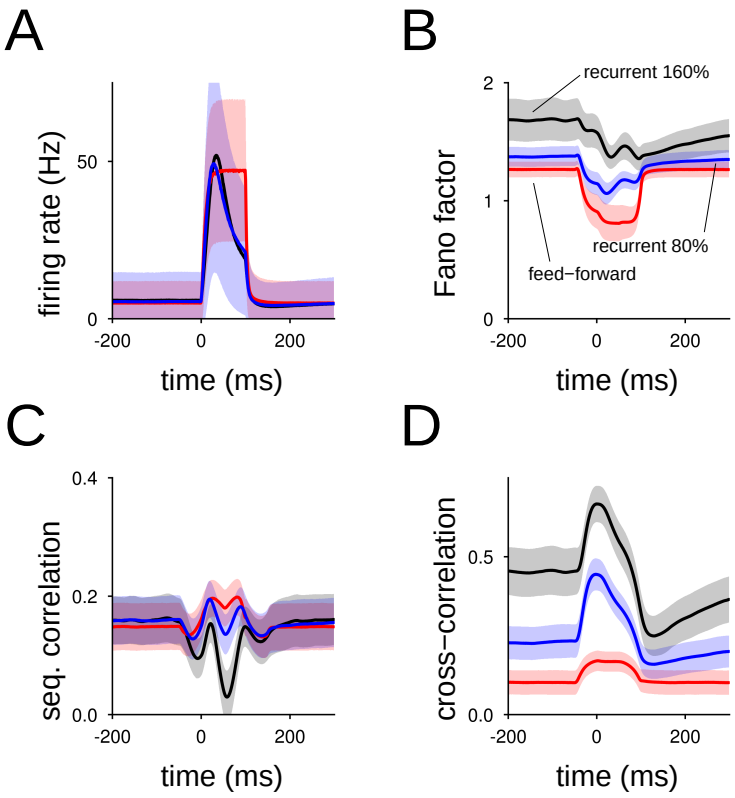

Figure 5.5: **A** The mean PSTH (created using 1 ms bins) from a population of 200 noisy IF neurons in response to an input from 200 Poisson spike trains (from 0 ms to 100 ms) - the red line indicates a single layer feed - forward network, receiving inputs using non - depressing synapses, the blue line indicates a single layer recurrent network, where the recurrence strength equal to 80% of the feed - forward strength whilst the black line indicates a single layer recurrent network in which the recurrence is 160% of the feed - forward strength, with all synapses subject to short term synaptic depression. Note that the shaded regions represents one standard deviation - the standard error in this panel (and subsequent panels) is only slightly larger than the traces. **B** As A, but the mean Fano factor of the spike count generated from a sliding window positioned in all trials at $[t, t + 50 \text{ ms}]$, plotted against t. **C** The mean sequential correlation between the spike counts from two sliding windows, $[t - 50 \text{ ms}, t]$ and $[t, t + 50 \text{ ms}]$. **D** The mean cross correlation between the spike counts collected from neuron pairs is plotted.

in which strength of the recurrent synapses are 80% the strength of the input synapses and another, more strongly recurrent network in which the strength of the recurrent synapses are 160% of the strength of the input synapses.

We wish to investigate high Fano factors and sequential correlation so we use noise parameters $\sigma_{noise} = 80$ pA and $\tau_{noise} = 20$ ms. Because we are interested in IF neurons with positive spike count cross correlations, we follow the example from section 5.3.2.1 (and Fig. 5.4), in which 20% of the noise received by each IF neuron in a layer comes from a shared noise current (i.e. $s = 0.2$). These are the same parameters used in the middle line in Fig. 5.4D which show, as firing rate is increased, spike counts from IF neuron pairs receiving the same parameter noisy input currents decrease in Fano factor, slightly increase in sequential correlation, and slightly increase in cross - correlation. By examining a purely feed - forward input receiving a stimulus presentation, we can use these results to help determine whether any spike - count statistics we observe are solely the result of the network firing rate[5].

In Fig. 5.5A , we plot the mean PSTH recorded from the weakly recurrent network (with recurrent strength 80% as strong as the feed - forward strength, blue line), the strongly recurrent network (with recurrent strength 160% as strong as the feed - forward strength, black line) and the feed - forward network (red line) in response to a strong input stimulus. Due to the differences between the recurrent network responses and that of the feed - forward network, we would expect the spike - count statistics during the response to be correspondingly different due to any rate - dependent effects. At the high strength input used here, the response profiles of the two different recurrent networks plotted in Fig. 5.5A are very similar, which will aid us in investigating to what extent the recurrent connectivity contributes to spike count statistics beyond a simple rate - dependent effect. When examining the response statistics of the spike counts generated by a sliding window, we find several differences between the recurrent networks and the feed - forward network:

The mean $Fano_{i,\Delta t}(t)$ was measured from the weakly recurrent network (Fig. 5.5B, blue line), the strongly recurrent network (Fig. 5.5B, black line) and the feed - for-

being the input type (Poisson input units for the feed - forward model as opposed to direct current injection).

[5]Although we focus on using Poisson spiking units as inputs to our networks, we also simulated a feed - forward network receiving direct current input. We find that if we do so, that $Fano_{i,50\text{ ms}}(t)$, $Seq_{i,50\text{ ms}}(t)$ and $Cross_{ij,50\text{ ms}}(t)$ measured from the feed - forward network using current input do not differ from those same quantities measured from the feed - forward network receiving inputs from Poisson spiking units (not shown). We conclude that any extra variability which is generated by the use of Poisson input trains is small, compared to the variability arising from the IF neuron noise currents, at least for the parameters we choose (the situation holds for a range of other parameters as well).

ward network (Fig. 5.5B, red line). We find that $Fano_{i,\Delta t}(t)$ measured from either recurrent network to be higher than for the feed - forward network, with $Fano_{i,\Delta t}(t)$ highest for the strongly recurrent network. During the stimulus response (i.e. $t \in$ [-50 ms, 150 ms]), $Fano_{i,\Delta t}(t)$ in all three types of network falls, as we expect from the increased firing rate. Fano factor changes are not purely rate dependent however, with $Fano_{i,\Delta t}(t)$ for t between - 50 ms and 150 ms in the heavily recurrent network still greater than that of the weakly recurrent network despite both recurrent networks having similar firing rate responses over the same period. For both recurrent networks, after the response to an input ($t > 150$ ms), depression in the recurrent connectivity results in a drop in $Fano_{i,\Delta t}(t)$ as we have predicted - this is particularly noticeable for the strongly recurrent network.

The mean $Seq_{i,\Delta t}(t)$ was also measured from the three network variants (Fig. 5.5C). In the absence of an input, $Seq_{i,\Delta t}(t)$ is similar across each network. However, once an input is presented, small differences begin to appear. In particular, the strongly recurrent network shows a fall in $Seq_{i,\Delta t}(t)$ at the mid - point of a stimulus response (t close to 50 ms), unlike the feed - forward network which shows a slight rise (in line with the simulation results reported in Fig. 5.4D, central panel). This is likely due to the effects of synaptic depression - a particularly strong common noise input during the initial response to a stimulus will depress the recurrent connections greatly, a phenomena which has been observed previously (Goldman et al., 1999, 2002). These reduced strength recurrence will then act to counteract the high noise input, leading to lower $Seq_{i,\Delta t}(t)$.

The mean $Cross_{ij,\Delta t}(t)$ was also measured from the three network variants (Fig. 5.5D). We find that $Cross_{ij,\Delta t}(t)$ in the first layer of either recurrent network to be higher than that measured from the first layer of the feed - forward network, with the first layer of the strongly recurrent network having the highest $Cross_{ij,\Delta t}(t)$. In a similar fashion to the Fano factor, the increase in $Cross_{ij,\Delta t}(t)$ is dependent on the recurrent connectivity - for sliding windows aligned after the stimulus response has ended ($t > 150$ ms), the cross correlation between spike counts is reduced below the level observed prior to the response onset. During the stimulus response, cross correlation between neurons is higher, as expected from the higher firing rate (and commensurate with the results reported in Fig. 5.4D, right panel).

We have found that $Fano_{i,\Delta t}(t)$ and $Cross_{ij,\Delta t}(t)$ in the recurrent networks are both higher than in the feed - forward network, as the recurrent connectivity acts to increase variability by amplifying stochastic fluctuations in firing rate. This depends greatly

on the noise source - in particular the shared noise current received by each neuron. If we remove these shared currents (by setting $s = 0$), we find that the previously observed increases in $Fano_{i,\Delta t}(t)$ and $Cross_{ij,\Delta t}(t)$ disappear when the sliding windows are not overlapping with the stimulus response, and also greatly reduced during stimulus presentation (see Appendix 5.A.3, Fig. 5.11). With uncorrelated noise, and thus with largely uncorrelated spike counts, the feedback from any network recurrence has a small effect on variability compared to an IF neuron's independent noise current, since the synaptic contributions from a group of neurons are easier to 'average out'.

We have discussed only the first layers of feed - forward and recurrent networks here. As the signal propagates through deeper layers of the network, small stochastic fluctuations in firing rate tend to be amplified, increasing the Fano factor, sequential correlation and cross - correlation of the spike counts measured from neurons in deeper layers. This is to be expected, but it is worth noting that the propagation models we consider, in which all inputs to a given neuron come from neurons in the previous layer which have similar stimulus response properties, is not necessarily biologically relevant. The results we present here are at odds with the experimental evidence which show a stimulus onset dependent fall of $Fano_{i,\Delta t}(t)$, $Seq_{i,\Delta t}(t)$ and $Cross_{ij,\Delta t}(t)$ (see Fig. 5.3, also observed in Churchland et al., 2010). We predicted that spike count variability would reduce due to the depression of recurrent connectivity. Whilst we did observe reductions in $Fano_{i,\Delta t}(t)$ and $Cross_{ij,\Delta t}(t)$ for $t > 150$ ms, the parameters of noise currents injected into each IF neuron also have a strong impact on variability, sequential correlation and cross correlation as well.

5.3.3 Signal propagation and noise filtering

5.3.3.1 Signal propagation

When tuning the parameters of our network (especially the noise parameters), we typically employ the largely arbitrary requirement which prioritises networks in which the output responses to different amplitude inputs varies linearly with the input amplitude. Such approaches have previously been considered for layered networks of IF neurons which receive a noisy rapidly varying input signal, in which network performance is measured by how linear the relationship between the output firing rate and the input signal is. To do this, one can use measures such as *dissimilarity* (van Rossum et al., 2002) or *similarity* (Vogels and Abbott, 2005) as metrics measuring linearity (see methods section for details on these metrics). This is not necessarily the best approach

- a network could propagate a signal in a non - linear fashion whilst still retaining all information about it. In the networks we consider, low contrast inputs are filtered by the recurrent synaptic connectivity - whilst this may lead to a nonlinear response, it has been conjectured that such connections can help filter out high frequency noise in the input (van Rossum et al., 2008). Therefore, we examined alternate metrics which might be better suited to examining such network properties.

Our approach to the problem of tuning is to use information theory. Essentially, the idea is to judge the efficacy of a network by the *mutual information* between the input signal S and an output signal R. We first examine how this metric compares with those of the dissimilarity and similarity metrics. We simulated layered networks of integrate and fire neurons, with each of the five layers containing 20 neurons connected all - to - all (using non - depressing synapses) to the next layer. The network receives an input signal in the form of a time varying current $s(t)$ (generated by an OU process with a 50 ms timeconstant and zero mean, rectified to be non - negative), which is injected into each of the IF neurons in the first layer. The output of the network is the firing rate, $r(t)$, measured from the final layer by counting spikes using 5 ms bins (see methods for further simulation details). Fig. 5.6A and B shows the dissimilarity and similarity measures, $\min_{\Delta t} D(r,s,\triangle t)$ and $\max_{\Delta t} \mathrm{corr}(r,s,\triangle t)$, applied to the signal s and the response r for different noise parameters - mean μ_{noise} and standard deviation σ_{noise} (left hand panels, noise timeconstant $\tau_{noise} = 2\,\mathrm{ms}$). We find that the similarity between r and s is high, then the dissimilarity between r and s will be low, and vice - versa - this is trivially expected since they both measure linearity. By adjusting the mean noise current μ_{noise} we also change the background firing rate - the dissimilarity and similarity are thus also plotted for different levels of background firing rates (Fig. 5.6A and B, central panels).

In addition to calculating similarity and dissimilarity, we measured the mutual information between the firing rate $R_{\Delta t}$, Δt ms after the stimulus S using $r(t-\Delta t)$ and $s(t)$ as t varies as observations of $R_{\Delta t}$ and S. We maximise this mutual information metric, $I_{delay}(r,s,\Delta t)$, over Δt, recording the result in Fig. 5.6C. We find that parameter regimes which give rise to low dissimilarity and high similarity also lead to conditions in which the mutual information between r and s is high. This result is not trivial, since information transmission is not necessarily reduced by nonlinearities[6], however this assumes that the response to the signal is best measured by the firing rate. It is

[6]For instance, a bijective distortion $R = f(S)$ will not lead to a loss of information regarding S, although no function f can *increase* information about S - i.e. $I(S,R) \leq I(S,S)$ (Shannon, 1948).

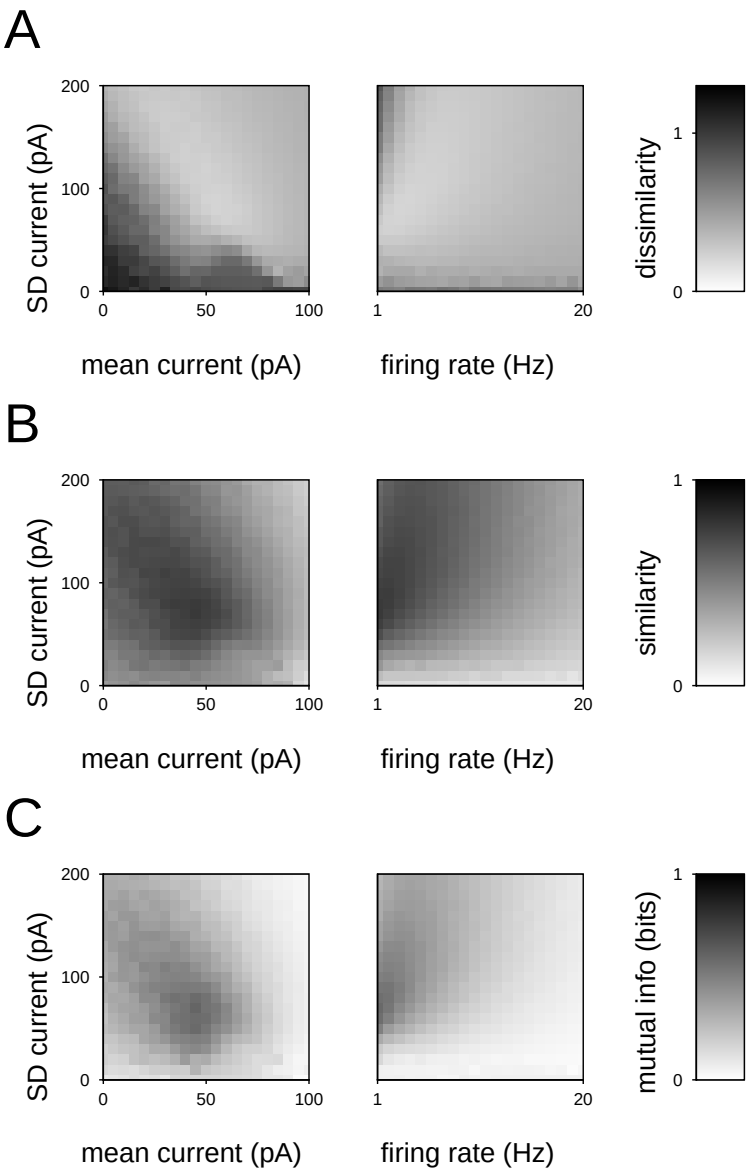

Figure 5.6: **A** We present a time varying signal $s(t)$ to a five layer network, measuring the firing rate $r(t)$ in the final layer by binning spikes into 5 ms bins, measuring the *dissimilarity* between them. This dissimilarity is plotted against the mean and standard deviation of the noise current injected into the noisy IF neurons (left panel - key is plotted in the right panel). Alternatively (centre panel), we vary the standard deviation of the noise current and the background firing rate of the network (which we adhere to by altering the noise current mean until the network has the required background firing rate).**B** As A, but now the *similarity* between r and s is measured. **C** As A, but now the *mutual information* (for an optimal delay) between r and s is measured.

possible that an output spike train might be informative about the input signal in ways other than the localised firing rate, such as by spike times or inter - spike intervals. The parameter regimes favoured by both linear and informative signal propagation incorporate low firing rates, with sufficient noise to desynchronise spike times of the IF neurons whilst not overwhelming the signal - the 'rate mode' (van Rossum et al., 2002).

5.3.3.2 Noise filtering by a recurrently connected rate - based network

One of the properties of the original rate based network incorporating recurrent, depressing connectivity, was its ability to filter out high - frequency input noise more effectively than the purely feed - forward network (Van Rossum et al., 2008). First, we examine this property of the rate - based networks using mutual information - we later apply the same approach to spiking networks. This approach is illustrated in Fig. 5.7A. We represent the input signal into our networks by a random variable, S, which indicates the presence or absence of a stimulus during a given trial. When $S = 1$, with probability 1/2, a stimulus is presented from 0 ms to 100 ms - the input signal is equal to a set stimulus strength for that interval whilst the rest of the time it is equal to the background strength of 5 Hz (Fig. 5.7A left, upper trace). When $S = 0$, no stimulus is presented to the network and the input signal is set to the constant background strength of 5 Hz (Fig. 5.7A left, lower trace). The input signal is then corrupted by additive noise (generated by an OU process, Fig. 5.7A, central traces, with 5 Hz standard deviation and a 5 ms timeconstant). Finally, the corrupted input signal is presented to our networks (Fig. 5.7A, right) from which we record the firing rate averaged over $[t - \delta t, t]$, denoted $r_{\delta t}(t)$. Repeating this procedure over many trials allows us to measure the mutual information $I(r_{\delta t}(t); S)$ between the average firing rate $r_{\delta t}(t)$ and S, the presence or absence of a stimulus in the current trial. A high value of $I(r_{\delta t}(t); S)$ indicates that $r_{\delta t}(t)$ is very informative about whether or not a stimulus has been presented. We usually set $\delta t = 5$ ms - note also that the stimulus strength and positioning within the trial are always fixed. The rate - based networks we use are depicted on the right hand side of Fig. 5.7A) - the feed - forward network has no recurrent connectivity, whilst the recurrent network includes depressing recurrent feedback connections (both networks obey Eqs. 5.7). The relative strengths of the feed - forward and recurrent connections are kept in a fixed proportion to each other - their overall amplitudes were adjusted simultaneously to ensure that a 50 Hz input led to a peak firing rate of 50 Hz.

We show here that the firing rate of the recurrent network provides more informa-

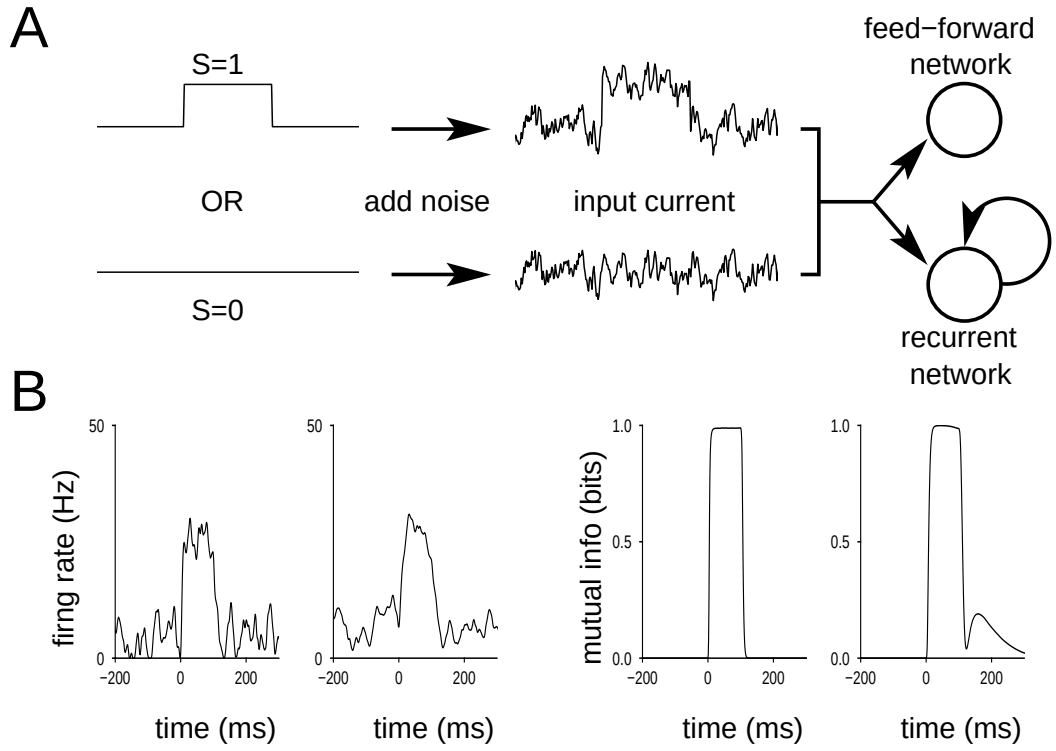

Figure 5.7: **A** *Left*, there are two possible input signals to our networks. When random variable $S = 1$ (with 50% probability), a stimulus is presented, so the input signal contains a square pulse of activity (of constant amplitude) from 0 ms to 100 ms, whilst when $S = 0$ the input signal is constant, equal to the background rate. *Centre*, the input signal is corrupted by additive noise, generated by an OU process with a short timeconstant. *Right*, the corrupted signal is then injected into either a feed - forward network consisting of a single unit, or a recurrent network which consists of a single unit featuring recurrent, depressing, synaptic connectivity. **B** *Left hand panel*s, Firing rate traces from both the feed - forward network (far left panel) and the recurrent network (centre left panel) in response to an input signal containing a stimulus presented at 25 Hz (with a 5 Hz background rate), corrupted by additive noise (generated by an OU process with a standard deviation of 5 Hz and a 5 ms relaxation timeconstant). *Right hand panels*, the mutual information between the firing rate averaged over 5 ms bins, $r_{5\text{ ms}}(t)$, and S, plotted against t for the feed - forward network (centre right panel) and recurrent network (far right panel). When the mutual information equals 1 bit, the presence or absence of a stimuli can be detected from the firing rate with 100% accuracy.

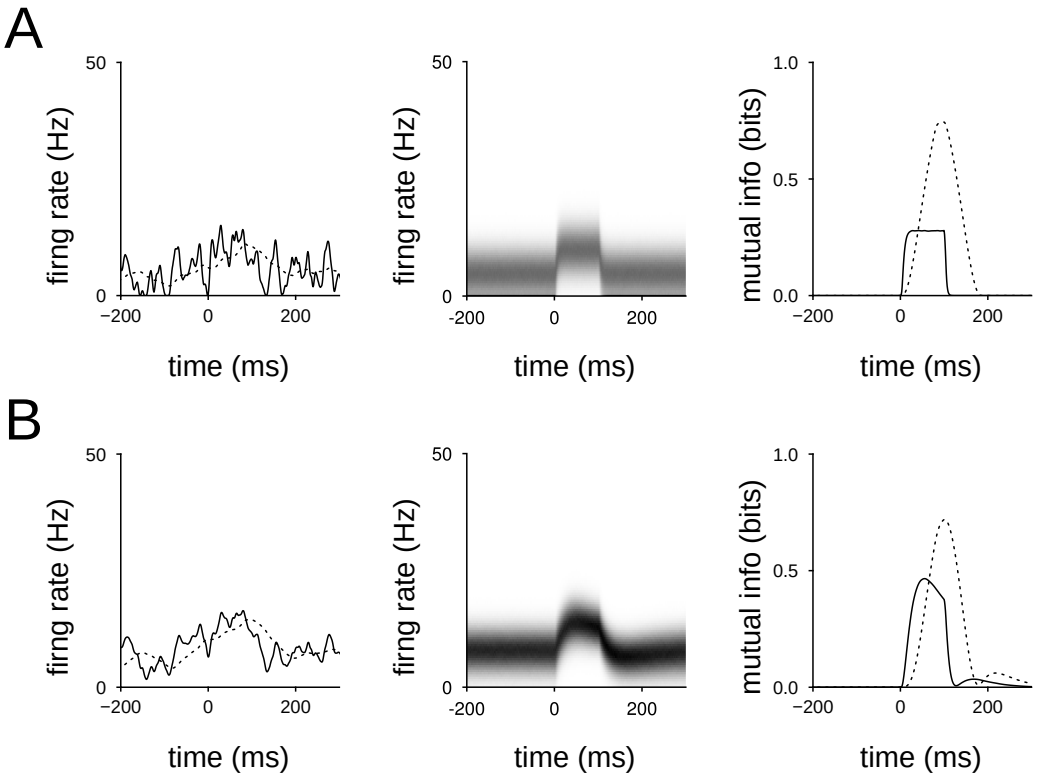

Figure 5.8: **A** *Left panel*, $r_{5\text{ ms}}(t)$ from a rate - based feed - forward network (see Fig. 5.7A), presented with a 10 Hz stimulus (with a 5 Hz background rate) corrupted by additive noise is plotted against t (solid line). We also plot the firing rate averaged over bins of 80 ms, $r_{80\text{ ms}}(t)$, as a comparison (dashed line). *Centre panel*, the probability density functions for $r_{5\text{ ms}}(t)$ plotted for time t - black indicates highest densities. *Right panel*, the mutual information between $r_{5\text{ ms}}(t)$ and S (solid line). For comparison, the mutual information between the presence or absence of a stimulus, and the firing rate averaged over 80 ms bins is also plotted (dashed line, plotted against the bin end - point). **B** As A, but the recurrent network is now used.

tion about the presence or absence of a stimulus in the current trial than the firing rate of the feed - forward network. We also find that this difference between feed - forward and recurrent networks is exacerbated when the stimuli presented are of low amplitude and the firing rate is estimated using small time bins. For high strength inputs, the differences between networks are small. In the left hand panels of Fig. 5.7B, we plot the firing rate (averaged over 5 ms bins) $r_{5\text{ ms}}(t)$ recorded from the feed - forward (far left panel) and recurrent (centre left panel) networks in response to a 25 Hz stimulus corrupted by additive noise. In the right hand panels of Fig. 5.7B, we plot the mutual information between $r_{5\text{ ms}}(t)$ and S, the presence or absence of an input stimulus in the current trial, for the feed - forward network (centre right panel) and the recurrent network (far right panel). We can see that for both networks, during the stimulus interval [0 ms, 100 ms], the mutual information between $r_{5\text{ ms}}(t)$ and S is close to 1 - the maximum value the quantity can reach (the information entropy of S is only one bit). Although the firing rate of the recurrent network is less susceptible to high frequency noise fluctuations than the feed - forward network, the high stimulus amplitude ensures that the average of the firing rate of either network over a 5 ms interval during this period can be used to predict whether a stimulus has been presented during the current trial with near - perfect accuracy. After the stimulus has been removed, there is a second, lower, peak in $I(r_{5\text{ ms}}(t);S)$ for the recurrent network - this is due to the firing rate after a stimulus presentation decreasing below the networks natural background activity (which comes from the recurrent connectivity being depressed) - this effect means that $r_{5\text{ ms}}(t)$ just after 100 ms can provide some information about the presence or absence of a stimulus.

As the stimulus strength is reduced, the input signal is more vulnerable to corruption by additive noise. For such weak inputs, we find that the firing rate from the recurrent network is much better at discerning the presence or absence of a stimulus than the firing rate from the feed - forward network. We examine this by reducing the fixed stimulus strength from 25 Hz to 10 Hz. In the left hand panel of Fig. 5.8A, we plot $r_{5\text{ ms}}(t)$ as before (solid line) as well as the rate averaged over 80 ms bins, $r_{80\text{ ms}}(t)$ (dashed line). The combination of the noise and the low strength stimulus ensures that $r_{5\text{ ms}}(t)$ is highly variable - it is hard to detect that an input has be presented at all. This is further demonstrated in the central panel of Fig. 5.8A, in which we plot the probability density function for $r_{5\text{ ms}}(t)$ given $S = 1$, against t. In contrast, $r_{80\text{ ms}}(t)$ is much less variable - the larger time bin helping to average out the rapid noise fluctuations. In the right hand panel of Fig. 5.8A, we plot the mutual information $I(r_{5\text{ ms}}(t);S)$ against

t (solid line) - with the low amplitude input, the mutual information is much lower and a 5 ms average of the firing rate will tell you only a limited amount about the presence or absence of an input stimulus. We repeat this procedure with the recurrent network in Fig. 5.8B. We find that the recurrent connectivity acts to filter the input signal over time and ensures that $r_{5\text{ ms}}(t)$ is less vulnerable to high frequency noise fluctuations. The firing rate measured from the recurrent network can thus provide us with more information about the presence or absence of a stimulus, with $I(r_{5\text{ ms}}(t);S)$ much higher than when the feed - forward network is used. If we increase the recurrent strength relative to the feed - forward strength (up until some sensible limit - recurrence as much as five times the input strength), the mutual information between $r_{5\text{ ms}}(t)$ and S rises further (not shown). The difference between $I(r_{\delta t}(t);S)$ measured from feed - forward and recurrent networks decreases as δt is increased - as the rate is averaged over larger windows, the filtering from any recurrent feedback becomes less important, as rapid noise fluctuations are filtered out by the extended integration windows. This is illustrated in Fig. 5.8A and Fig. 5.8B, right hand panels, which plot $I(r_{80\text{ ms}}(r);S)$ for the feed - forward and recurrent networks. The disadvantage of using a long integration time window is that to obtain the maximum information using this method, the firing rate must be integrated over a longer bin.

5.3.3.3 Noise filtering by a recurrently connected spiking network

Having established the advantages of having recurrent connectivity for the single rate - based unit, we consider the corresponding network modelled with IF neurons. We use feed - forward and recurrent networks similar to those spiking networks presented earlier, with a few alterations. Firstly, each network was reduce to a single layer. Secondly, the input to each network is changed to a current injection to each IF neuron, instead of Poisson spiking units, to maintain similarity with the rate - based networks. Thirdly, the mean μ_{noise} and standard deviation σ_{noise} of the OU process noise current injected separately into each noise IF neuron were systematically altered (we set the time - constant of those processes at 2 ms). When we record $r_{\delta t}(t)$ for such a network, we use all the IF neurons to contribute to it - so $r_{\delta t}(t)$ is the number of spikes produced by the an IF neuron in the interval $[t-\delta t, t]$ averaged over all neurons and divided by δt. The input current was corrupted by additive noise before being injected into the network and was weighted by a constant gain, tuned such that a 100 pA input pulse for 100 ms would lead to a peak firing rate in the network of 50 Hz.

As before, we use the random variable S to indicate the presence of a stimulus in

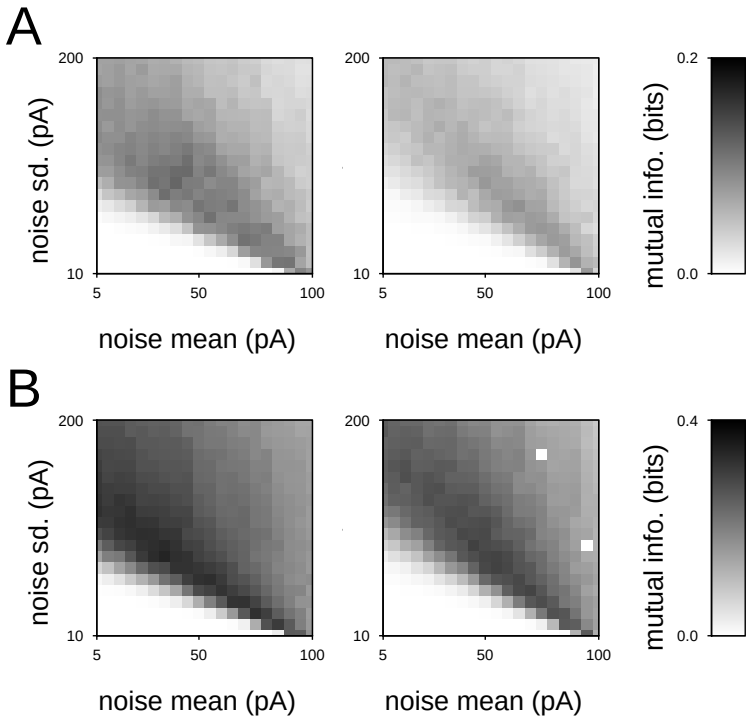

Figure 5.9: **A** We repeat the procedure in Fig. 5.7 with single layer networks of 200 IF neurons either with or without recurrent connections. The input, corrupted by OU noise, is injected into the network. We plot the peak over t of the mutual information between the firing rate averaged over 5 ms bins, $r_{5\,\mathrm{ms}}(t)$, and S, representing the presence or absence of a stimulus, for the feed - forward network (left panel) and the recurrent network (right panel), against the mean and standard deviation of the noise currents which are independently injected into each neuron. **B** As A, but the single layer networks now contain 2000 neurons each (the two isolated white squares in right hand panel a result of computational errors).

a given trial, and look at the mutual information between S and $r_{\delta t}(t)$. Here $r_{\delta t}(t)$ is now the mean network firing rate, averaged over the time interval $[t-\delta t, t]$. In trial containing the presence of a stimulus ($S=1$), the input current is a square pulse of 20 pA on the interval [0 ms, 100 ms] (with background current of 0 pA), corrupted by additive noise generated by an OU process (with 20 pA standard deviation and 5 ms relaxation time - constant). We systematically vary IF neuron noise parameters μ_{noise} and σ_{noise} whilst measuring the maximum over t of $I(r_{5\text{ ms}}(t);S)$. We plot this result in Fig. 5.9A for both the feed - forward network (left panel) and the recurrent network (right panel).

Immediately, we can see that $r_{5\text{ ms}}(t)$ recorded from the feed - forward network feed - forward network is more informative about S over a wider range of IF neuron noise parameters than $r_{5\text{ ms}}(t)$ recorded from the recurrent network. There are few parameter regimes where the firing rate recorded from the recurrent network is more informative than the firing rate recorded from the feed - forward network, and none where the improvement is as large as was observed with the rate - based networks. One possibility for this discrepancy is the spiking network size - with our single layer containing just 200 neurons, the contribution of the individual noise currents to the network is still likely to be strong, outweighing any positive effect from the filtering. To test this, we increased the network size by a factor of ten to 2000 neurons (the recurrent connections still form with probability 10%, although each is now only one tenth as strong) - now the population should behave more rate - like, with the recurrent input for each neuron less dependent on individual spikes. This increased network size indeed removes a lot of the noise from the firing rate $r_{5\text{ ms}}(t)$, with the peak mutual information $I(r_{5\text{ ms}}(t);S)$ higher across a range of parameters for both networks. However, the firing rate of the recurrent network still provides less information about the presence or absence of a stimulus than the feed - forward network. Isolated parameter regimes when networks of 5000 neurons and upwards have been used - none lead to a situation in which the firing rate of a recurrent network is a great deal more informative about the presence or absence of an stimulus in the input signal than the firing rate from a feed - forward network, although computational cost quickly makes exhaustively exploring such regimes difficult.

Several alternate reasons exist which could explain why the addition of recurrent feedback connections aids in the filtering of high frequency noise for rate - based networks, but not for networks of spiking neurons. Firstly, the only source of noise in the rate - based model is the noise distorting the input - in contrast, when using a network

of noisy IF neurons, the noise currents of the individual neurons are present. Whilst this should be mitigated by large network sizes, the fact remains that each neuron in the feed - forward network receives an independent noise source, whilst in recurrent networks, a noise current can affect several neurons. Secondly, the act of spiking, in which the membrane potential is reset, can lead to a refractory - like effect whose dynamics are not captured by the simple input - output function used in the rate - based network. A more accurate rate - based model featuring recurrent connectivity might not lead to the same advantage in filtering out high frequency noise. We did, on the basis of some simulations, consider that the recurrent networks might be much more susceptible to synchronisation - whilst this is true in deeper layers of the multilayer network, further investigations indicated that this was not the case here in a one layer network for the majority of noise parameters.

5.4 Discussion

We have investigated how the response onset latency in layered networks of IF neurons, featuring recurrent, depressing synaptic connectivity, depends on contrast and stimulus identity. We have found the results to be broadly comparable with those of an existing rate - based model and experimental evidence (van Rossum et al., 2008). In particular, stimulus contrast plays a large role in determining response latency in recurrent networks, with the latency of responses to high contrast stimuli shorter than the latency of responses to low contrast stimuli. These results applied only when the networks were operating in ‘rate mode’ in which the noise currents injected independently into each IF neuron were of sufficient variability to counteract the tendency of spikes to synchronise over multiple layers.

We next considered the spiking statistics of our models when they were injected with correlated noise. Experimental evidence has found that the Fano factor, sequential correlation and cross - correlation of spike counts falls in line with the onset of a response to stimulus presentation (Churchland et al., 2010, Mike Oram, unpublished data). We found that the Fano factor and cross - correlation of spike counts measured from the first layer of the recurrent network were both higher than those statistics measured from a single layer of the feed - forward network prior to stimulus onset. However, during the response to a stimulus, whilst the Fano factor fell (in line with experimental evidence), the sequential correlation changed little and the cross - correlation of spike counts from neuron pairs increased. We had predicted that depression

in the recurrent connectivity would lead to a fall in cross - correlation (as well as in Fano factor) - we found that this effect was only noticeable after a response to a stimulus finished, a phenomenon most likely due to the strong impact of the noise current parameters and neural firing rate on spike count statistics.

When investigating spiking statistics, several factors were not considered here. Firstly, most of the contributions to spike count noise came from external current sources injected into each neuron - whilst the purely excitatory recurrent connections can amplify noise fluctuations, they are not the primary initiator of them. This is not the case in, for example, balanced networks, which feature both excitatory and inhibitory connections. Such a network type is very different from the network we present here - rather than using recurrent connections to attempt to filter the input, a balanced network with strong enough synaptic connections can generate chaotic activity (van Vreeswijk and Sompolinsky, 1996, 1998). When such a network receives an input strong enough to suppress the chaotic fluctuations it generates, a fall in variability and cross - correlation likely occurs (Rajan et al., 2010)[7]. Another possible factor in statistical differences is that the input trains are poorly modelled by the Poisson processes - physiological results from a number of parts of the nervous system, including cortical neurons, indicate non - Poissonian statistics in spike trains (Teich, 1989; Teich et al., 1990, 1996; Oram et al., 1999). However this will not have a big impact unless we assume that the inputs to a given neuron from the input layer are themselves correlated. It is not clear that this is the case - inputs to cortical neurons arise from a diverse range of sources - the structure of correlations between such sources is likely to be non - trivial. Further investigation of how different network types and parameter regimes interact to lead to different spiking statistics is needed.

We also examined tuning recurrent networks using mutual information. Initially, we examined rate - based networks with single units, to which we presented stimuli with a fixed amplitude, corrupting such inputs with high frequency, additive, noise. We found that the firing rate, when measured over short time bins, of a single unit with recurrent connectivity was more informative about the presence or absence of a stimulus than the same firing rate measured from a single unit with no recurrence - in line with the improved signal - to - noise ratio of such firing rates reported by van Rossum et al. (2008). The recurrent connectivity helps to filter out high frequency noise by

[7] Any examination of networks featuring both excitation and inhibition would be a good opportunity to integrate short - term facilitation into our model - known to be strongly present in inhibitory neurons (Thomson, 1997). In particular, facilitating inhibition of sufficient length might induce positive sequential correlations without the need for an external noise current.

integrating the input signal over time - performance is improved if the recurrence is strengthened (to a degree) relative to the feed - forward input. Next, we used the same approach to examine how a spiking network with recurrent feedback connectivity acts to filter incoming signals by translating the single unit rate - based model into a single layer network of 200 IF neurons. In contrast to the firing rate model, we found that the recurrent spiking network did not, typically, outperform the feed - forward spiking network. There are several possible reasons for this difference. One is that the noise injected into each IF neuron contributes to the noise which the recurrent network must deal with through the recurrent connections. Using large networks should counteract this, but the differences between the feed - forward and recurrent networks remained when the number of neurons in our single layer networks was increased from 200 to 2000 (limited simulations of 5000+ neuron networks also revealed no substantial difference in mutual information across network types). Additionally, the membrane potential resetting and resultant refractory - like effect following a spike might act to reduce corruption by high frequency noise, a property not captured by rate - based neurons. These findings illustrate that care needs to be taken when comparing the results obtained from rate - based models with those obtained from spiking models, especially when the impact of noise is important.

5.A Appendix

5.A.1 Latency dependence on stimulus contrast and preference

As an extension of the single stimuli networks, a layered network of rate based neurons with recurrent, depressing centre - surround connectivity has previously been used to examine the relationship between stimulus preference and response latency, finding that the latency of a response to a presented stimulus depends more strongly on contrast than on the identity of the stimulus presented, in line with experimental evidence (van Rossum et al., 2008). We investigated whether this situation changes when networks of noisy IF neurons are used instead. For this, we use a three layer network, with each of the three layers consisting of 20 sub - populations, each of which contains 12 excitatory neurons and 3 inhibitory neurons. Each sub - population is assigned an orientation and are arranged in a ring - like structure. The excitatory neurons within a sub - population project strongly to neurons in sub - populations which have similar orientation labels (for target neurons in both the same layer and the next layer), and

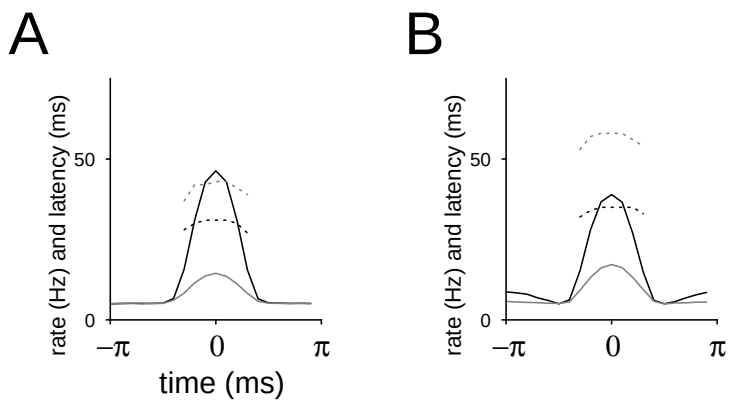

Figure 5.10: **A** The maximum response to a high contrast input to a feed - forward network coding orientation (solid lines) two two different strength inputs plotted against the orientation relative to the presentation stimulus (darker for high contrast inputs) and the half maximum latency of the response (dashed lines). **B** As A but using a recurrent network using recurrent, depressing connectivity.

form no connections with neurons in sub - populations which have dissimilar orientation labels. Meanwhile, the inhibitory neurons in a sub - population project strongly to neurons in sub - populations which have the opposite orientation label (again, for target neurons in both the same layer and the next layer), and form no connections with neurons in sub - populations which have similar orientation labels (see Methods for details). The strength of the recurrence was limited to 80% of the strength of the feed - forward connections - this prevents a rotating activity profile forming in absence of input.

Fig. 5.10 shows the peak response to inputs of high and low contrasts (dark and light solid lines) along with the half - maximum latency (dashed lines) of the response, plotted against the difference between the orientation of a presented stimulus and the orientation preference of the IF neurons the response is recorded from for a purely feed - forward network (with no depression - panel A) and a recurrent network (panel B) confirming the results obtained previously using a rate based network that response latency depends more strongly on contrast than stimulus preference (van Rossum et al., 2008).

5.A.2 Spike - count matched model

In addition to investigating how the firing statistics of IF neurons can be influenced by the noise current parameters, we also investigate the firing statistics of the spike -

count matched model, previously used to fit a number of statistics of real neurons in which the number of spikes in a given trial is matched to experimentally recorded spike counts (Oram et al., 1999). We used a simplified version of the spike - count matched model in order to carry out some mathematical analysis. We find that the Fano factor and sequential correlation of spike counts produced by this model and measured using sliding windows can be expected to co - vary - with high Fano factor leading to high, positive, sequential correlation.

For trial of length T, let the number of spikes it contains be equal to S. The first stage of the spike count matched model is to distribute these spikes in the trial window $[0,T]$ according to a probability density function $p(t)$ - proportional to the mean firing rate. The next stage concerns a correction to account for the refractory nature of real neurons - since we are interested in a mathematical result here, we omit this stage. Additionally, rather than using real spike counts, we assume that S is a random variable with mean rT and variance FrT, where r is the mean firing rate and F is the Fano factor of the spike count over the trial window. We call this model the *simplified spike - count matched model*.

Constant Rate Basic Spike - Count Matched Model Initially, we assume that $p(t)$ is flat - i.e. our hypothetical neuron fires with constant rate. We want to measure the Fano factor, $Fano_{i,\Delta t}(t)$, and sequential correlation, $Seq_{i,\Delta t}(t)$, when they are collected over sliding windows (we may assume these to be located at any point in the trial, due to the constant firing rate).

To calculate the Fano factor $Fano_{i,\Delta t}$, suppose that the measurement window has length l and contains $C = C_i$ spikes. With constant $p(t)$, the probability a spike will be contained within our counting window is $P = l/T$. Since these spikes are allocated independently from each other, the number of spikes in our counting window is thus bionomially distributed, $C|S \sim B(S, l/T)$. By the law of total expectation, we have that

$$E[C] = E[E[C|S]] = E[SP] = rTP. \tag{5.8}$$

Also, by the law of total variance, we have that

$$\begin{aligned} var(C) &= var(E[C|S]) + E[var(C|S)], \\ &= var(SP) + E[SP(1-P)] = (F-1)rTP^2 + rTP, \\ &= \frac{(F-1)l^2 r}{T} + rl. \end{aligned} \tag{5.9}$$

Combining Eq. (5.8) and Eq. (5.9) gives us the Fano factor $Fano_{i,\Delta t}(t)$ as

$$Fano_{i,\Delta t}(t) = 1 + (F-1)P = 1 + \frac{(F-1)l}{T}. \tag{5.10}$$

Likewise, it is straightforward to calculate the sequential correlation. Suppose that we have two, non - overlapping, windows, of length l with spike counts C and K. The probability of a spike being assigned to the first window, P_C, and the probability a spike is assigned to the second window, P_K, are both equal to l/T. The spike counts conditioned on S, $(C,K)|S$ thus form a multinomial distribution. Clearly the mean and variance of each spike count is the same as before - we must also consider the covariance. For any two random variables, X and Y, conditioned on a third, Z, clearly, we have, from the law of total expectation that

$$\begin{aligned} cov(X,Y) &= E[(X-E[X])(Y-E[Y])], \\ &= E[XY] - E[X]E[Y], \\ &= E[E[XY|Z]] - E[E[X|Z]]E[E[Y|Z]], \\ &= E[cov(X,Y|Z)] + cov(E[X|Z], E[Y|Z]). \end{aligned} \tag{5.11}$$

Using this equation, we obtain that:

$$\begin{aligned} cov(C,K) &= E[cov(C,K|S)] + cov(E[C|S], E[K|S]), \\ &= E[-SP_CP_K] + cov(SP_C, SP_K), \\ &= (F-1)rTP_CP_K = \frac{(F-1)l^2r}{T}. \end{aligned} \tag{5.12}$$

When the counts are from adjacent windows, one gets the sequential correlation at that time point:

$$Seq_{i,\Delta t}(t) := corr(C,K) = \frac{(F-1)P_CP_K}{\sqrt{((F-1)P_C^2 + P_C)((F-1)P_S^2 + P_S)}} = \frac{(F-1)l}{(F-1)l + T}. \tag{5.13}$$

Note that this correlation depends entirely on the fact that the spike counts are linked through the total number of spikes S in a trial - there is no dependence on where in the trial the two sliding windows are placed, unlike with experimentally collected data.

For constant rate simplified spike - count matched models, raising F, the Fano factor of the spike count over trials of length T, raises both the Fano factor and sequential correlation collected by sliding windows. The situation can be more easily visualised by considering extreme examples of the Fano factor F:

$F = 0$**:** A Fano factor of zero over the trial length does not lead to a sliding window Fano factor of zero - the random assignment of spikes within the trial window leaves some variability. It does, however, have a strong anti - correlating effect on spike counts in consecutive windows within the trial - since the total number of spikes is fixed, every spike in one window is one that cannot be in the other. The most extreme example is when the two consecutive windows are $[0, T/2]$ and $[T/2, T]$ - the spike counts in these are perfectly anti - correlated.

$F \gg 1$**:** A Fano factor much greater than 1 leads to a much different situation. Now, the greatest source of variability in each sliding window is not whether or not a spike is assigned within it, as before, but on the total number of spikes in each trial. When two sliding windows are placed within the trial, although a spike assigned to one can no longer be assigned to another (the anti - correlate effect for low Fano factors) this is less important than the total, highly variable, number of spikes assigned to the trial overall - a high spike count in one window is indicative of a high spike count over the trial (and thus a high spike count in the other window) leading to strong, positive correlation between spike counts.

These examples illustrate the general rule of Fano factor and sequential correlation being interdependent in the spike - count matched model.

Varying Rate Basic Spike - Count Matched Model To model the rate dependence of the Fano factor and sequential correlation generated from spike counts in sliding windows, it is natural to ask what would happen if we simply used a time - varying spike probability (as is done for the full spike count matched model, albeit here without the ISI distribution correction stages). Now, all that has changed now is that $P(t)$, the probability of a spike falling within a given sliding window $[t, t+l]$, is different, depending of where the window is placed in the trial window. Suppose now that each spike is allocated in proportion to a trial independent profile $p(t)$, then we have that:

$$P(t) = \int_t^{t+l} p(t)dt \tag{5.14}$$

From Eq. (5.8), Eq. (5.9) and Eq. (5.14), we have the following formula for the windowed Fano factor:

$$Fano_{i,\Delta t}(t) \quad = \quad 1 + (F-1)P(t).$$

We also find, from Eq. (5.13) that,

$$Seq_{slide} = \frac{(F-1)P(t)P(t+l)}{\sqrt{(P(t)+(F-1)P^2(t))(P(t+l)+(F-1)P^2(t+l))}}. \quad (5.15)$$

We can see then that an increase in firing rate rate (i.e, a rate increase in which $P(t)$ and $P(t+l)$ both increase) will increase both the Fano factor and sequential correlation provided $F > 1$. This is at odds with the experimental evidence (Churchland et al., 2010, Mike Oram, unpublished data). More recently, a modified version of the spike - count matched model has been used to account for this (Mike Oram, personnel communication).

5.A.3 Spike count statistics in networks with uncorrelated noise

In section 5.3.2.2, we examined the spike count statistics of networks in which the noise currents injected into each neuron have a shared component. In Fig. 5.11, we show what happens when uncorrelated noise currents are used. In all respects, the parameters used are identical to those used in Fig. 5.5, except that the shared current proportion s is set to zero - all IF neurons thus receive independently generated noise currents. When comparing networks whose neurons receive correlated noise currents with network whose neurons receive independent noise currents, we find that the mean firing rate profiles remain similar (Fig. 5.5 and 5.11, panel A). For purely feed - forward networks, the windowed Fano factor and sequential correlation remain unchanged by the presence/absence of correlation in the IF neuron noise currents (Fig. 5.5 and 5.11, panels B and C, red traces). This is not surprising - the lack of recurrent connections, and the fact that Fano factors and sequential correlations are measured from single neurons would lead us to suspect this.

For networks which contain recurrent connections, the situation is different. Networks which have a shared noise current show larger windowed Fano factors throughout the trial (Fig. 5.5 and 5.11, blue/black traces, panel B). This shows how the recurrent connections amplify the effect that the shared noise current has on the network. We see how this additional variability is dependent on the recurrent connections when we examine the sequential correlation (Fig. 5.5 and 5.11, blue/black traces, panel C). The sequential correlation during stimulus onset is lower for the recurrent networks receiving correlated noise currents than for those those which do not. As discussed in section 5.3.2.2, a strong increase in the shared noise current during the first part of the

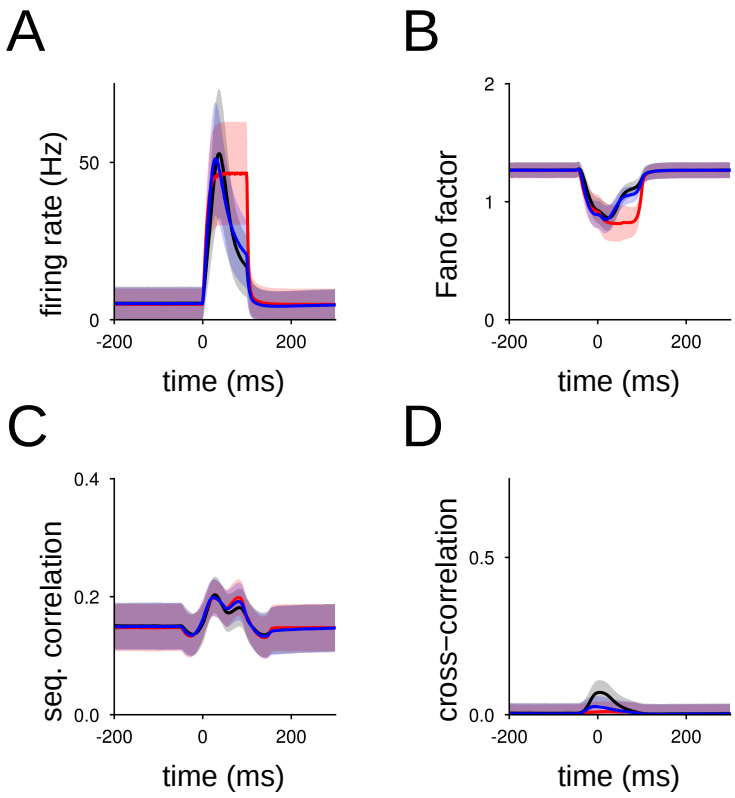

Figure 5.11: This figure replicates Fig. 5.5, except here, each of the population of 200 noisy IF neurons receive independent noise currents. **A** The mean PSTH (created using 1 ms bins) from a population of 200 noisy IF neurons in response to an input from 200 Poisson spike trains (from 0 ms to 100 ms) - the red line indicates a single layer feed - forward network, receiving inputs using non - depressing synapses, the blue line indicates a single layer recurrent network, where the recurrence strength equal to 80% of the feed - forward strength whilst the black line indicates a single layer recurrent network in which the recurrence is 160% of the feed - forward strength, with all synapses subject to short term synaptic depression. Note that the shaded regions represents one standard deviation - the standard error in this panel (and subsequent panels) is only slightly larger than the traces. **B** As A, but the mean Fano factor of the spike count generated from a sliding window positioned in all trials at $[t, t+50 \text{ ms}]$, plotted against t. **C** The mean sequential correlation between the spike counts from two sliding windows, $[t-50 \text{ ms}, t]$ and $[t, t+50 \text{ ms}]$. **D** The mean cross correlation between the spike counts collected from neuron pairs is plotted.

response will strongly increase depression in all recurrent connections and lead to reduction in the second part of the response (Goldman et al., 1999, 2002) - this is absent when there is no shared noise current, since any stochastic rises in the firing rate of one neuron are unlikely to be replicated in all of the others. Unsurprisingly, in the absence of a shared noise current, cross - correlation in the spike counts of pairs of neurons are dramatically lower (Fig. 5.5 and 5.11, panel D). This is because the bulk of the variability in the networks comes from the noise currents injected into each neuron - when they are independent, even the presence of recurrent connections will lead only limited cross - correlation.

Chapter 6

Concluding remarks

Over the course of this thesis, we have investigated the implication of short - term synaptic depression on the activity of neural models, as well as examining a related issue of how neural coding is affected by the encoding of multiple stimuli. We have seen that synaptic dynamics can have a strong influence on network activity, both when that network is responding to a presented stimulus and when it is acting spontaneously, in the absence of input. In particular, the inclusion of synaptic depression undermines the ability of sustained high firing rates in a neuron to have more than a transient effect on the network, which can ensure previously stable activity patterns become unstable (as in chapter 2) or limit the ability of a network to undergo runaway excitatory activity. Also, synaptic depression leaves a activity dependent history which can influence future dynamics by suppressing subsequent activity (as in chapter 2 - the depressing 'wake' ensuring that the rotating solution travels in one direction, chapter 3 - the suppression of responses to subsequent stimuli, and chapter 5 - the reduction in the ability of recurrent connections to contribute to spike count variability). Furthermore, short term depression acts in a contrast dependent fashion, differentiating its effects from those of other mechanisms, such as spike frequency adaptation, thought to act more linearly, and allowing us to account for many behaviours which have a strong contrast dependence.

Summary of results

In chapter 2, we considered a commonly used model - the ring attractor network (Ben-Yishai et al., 1995) - and find that its dynamics can change greatly when we introduce synaptic depression. Our main finding is that for some levels of background

input, static firing rate profiles are no longer stable (in contrast to the non - depressing model), and the network activity instead converses to a rotating profile of activity. This rotating profile is one of three possible limiting behaviours - the other two being a flat homogeneous state, in which every neuron has equal firing rate and narrowly tuned, stationary 'bump' of activity - due to the synaptic depression, runaway excitation leading to diverging activity is no longer possible. Such a rotating profile (not possible in the non - depressing model first proposed by Ben-Yishai et al., 1995) suggests a possible source of the moving / switching activity states observed experimentally (Kenet et al., 2003), although alternate models incorporating separate excitatory and inhibitory neuron populations (Ben-Yishai et al., 1997) or spike frequency adaptation (Hansel and Sompolinsky, 1998) could also account for this. When the ring attractors activity converges to a rotating profile, we find that the propagation speed of this profile monotonically increases with increasing amplitude background input, in contrast to a ring attractor network with neurons subject to spike frequency adaptation (Hansel and Sompolinsky, 1998). Using mathematical analysis we are able to determine a close approximation of the maximum possible rotation speed for a given network, and the background input amplitude for which such a rotation speed can be attained.

The variable rotation speed, controlled by the the amplitude of the background current, which we see in the ring attractor network with depressing synapses, suggests that such a network could play a role in the generation of temporal sequences at a variable speed, such as is required in speech generation. We did examine, in a limited fashion, similar networks featuring centre - surround connections arranged on the surfaces of a sphere or a torus, finding that the amplitude of the background input again controls the limiting behaviour of the network activity - with homogeneous states, stationary bumps of activity and moving bumps of activity all potentially types of behaviour. Such higher dimensional structures might be relevant for networks which encode multiple feature maps, but further investigation is needed to provide a better picture as to how synaptic depression interacts with the activity of such networks.

In chapter 3, we examined a network which processes two input stimuli presented asynchronously. This was motivated by experimental evidence indicating that some neurons in V4 implement 'latency selection' on their inputs. Such a neurons response to two stimuli presented together most closely resembles its response to the first stimulus, had it been presented alone (i.e. the stimulus which, when presented separately, leads to responses with the shortest onset latency), with the contribution from the second stimulus greatly suppressed (Gawne, 2008, with somewhat related results reported

by Perrett et al., 2009). We argue that short term synaptic depression is a natural candidate to explain this behaviour. A key property of the model is that it features several *shared connections*, both feed - forward and recurrent, which are always activated during the response to a stimulus. As a result of the neural response to a high strength input, these shared connections will depress, and so reduce the ability of a subsequent stimulus to generate a neural response - this suppression is weaker when low strength inputs are presented first to the network, in line with the experimental evidence (Gawne, 2008). Such suppression also fits with experimental evidence regarding the processing of sequences of briefly - flashed stimuli where it was found that the neural response to an effective stimulus could be reduced by prior presentation of another effective stimulus (Perrett et al., 2009). We used our network to make predictions that the magnitude of the contrast dependence of the response latency will correlate, on a cell by cell basis, with the strength of suppression which arises from the latency selection effect.

Synaptic depression in our model was also found to be vital when modelling how stimulus anticipation might arise. We arranged in parallel a number of networks featuring shared recurrent connectivity, and showed how such a group of networks can act to anticipate the future position of a moving stimulus. This model was used to make the prediction that the magnitude of anticipation will increases for high contrast stimuli. Such anticipation was dependent on the depressing nature of the synapses - replacing the depressing synapses with static ones led instead to a lag in responding to sequences of inputs. There are some similarities between the ring attractor network presented in chapter 2 and these parallel networks. Initial simulations of the ring attractor network when presented with moving stimuli indicates that anticipatory behaviour by such a network is also possible, but the natural frequency of rotation arising from the recurrent connections interacts with the input rotation speed in a complicated manner that requires further elucidation.

The latency selection mechanism observed in Gawne (2008) was suggested to be one possible mechanism for implementing a maximum function - the stimulus which produced the largest response would also produce the response with the shortest latency. Neurons which respond as the maximum of their inputs have an important role to play in some proposed models of object recognition (Riesenhuber and Poggio, 1999). In chapter 4, we investigated methods of encoding multiple stimuli from a different theoretical angle - namely we asked how different methods of encoding responses to stimulus pairs affects the coding capacity of a neural population. Using a simple, addi-

tive noise model and two different classes of stimuli, we found that if the responses to pairs of stimuli were generated by a linear sum, decoding errors were higher than if the responses to pairs of stimuli were generated by the maximum function or a suppressive function. We provided a geometric intuition for this - the response produced by a linear sum of two component response vectors is closer to the individual responses, and thus more susceptible to misclassification errors.

Some issues can be taken with the approach we took to neural coding. In particular, we imposed a normalisation constraint on the overall population responses in response to any stimulus / pair of stimuli. Removing this constraint causes the linear sum to improve as a method for combining responses. However, such performance improvement is dependent on the additive noise model we use (chosen for mathematical tractability) - we argue that a more realistic noise model in which the amplitude of the noise changes with the overall activity would mitigate this effect somewhat. Repeating this approach with a more realistic noise model would aid in the construction of testable predictions regarding the coding of multiple stimuli. Another flaw in the approach we took was that only pairs of stimuli and solo stimuli were considered - the simultaneous presentation of three or more stimuli was computationally intractable for the approach we took. Perhaps suitable simplifications of our approach might make this more computationally feasible. Finally, we considered only stimuli presented at identical contrasts - allowing the contrast of each stimulus to vary separately will obviously have some impact on the results we have reported.

In chapter 5 of this thesis, we returned to models incorporating short - term synaptic depression when we investigated a spiking analogue of the model presented in van Rossum et al. (2008) which features a layered structure and recurrent, depressing, connectivity. The original model exhibited large contrast dependent latencies - when operating in the the so - called ‘rate mode’ our spiking model replicated this result. We then examined the spiking statistics of our model when the neurons were injected with correlated noise currents. Experimental evidence has found that in the visual system, the localised spike count Fano factor, sequential correlation and cross - correlation all drop at the onset of the neural response to a presented stimulus (Churchland et al., 2010, Mike Oram, unpublished data). Our spiking model provided the opportunity to test these statistics with the hypothesis that synaptic depression of the recurrent connections would lead to stimulus response onset dependent falls in spike count variability. We found this to be true, but other factors, such as a high firing rate, also played a large role in the models spike count statistics. Additionally, changing the parameters

of the noise currents led to vastly different spike count statistics. In general, without further parameter constraint, it will be difficult to identify how the different elements of spiking networks interact to produce observed spike count statistics, but synaptic depression in the recurrent connections will doubtless play a major role.

Partly motivated by a desire to place performance - inspired constraints on the choice of noise parameters which our network could adopt, we also investigated the ability of a spiking network featuring recurrent, depressing synaptic connectivity to filter out high frequency noise from a low strength input signal. To do so we used information theory, measuring the mutual information between the output firing rate and the presence or absence of an input stimulus. We found that whilst the output of rate - based unit with recurrent, depressing feedback was more informative about the presence or absence of an underlying low amplitude input signal than a rate - based unit with no recurrent connections, this was not true when considering a group of spiking neurons. In contrast, we found that, when using spiking neurons, that feed - forward networks outperform recurrent networks over a wide range of parameters. One possible reason for the discrepancy between the spiking model and the rate - based model might be that recurrent spiking networks have the tendency to synchronise. This result illustrates the danger of assuming that results obtained from rate - based models necessarily generalise to spiking models.

Future Directions

This work shows that something as seemingly innocuous as synaptic depression can have a strong influence on neural dynamics. This is particularly true when the time - scales of depression and neural processing are of the same order (i.e. approximately 100 ms). There are many topics in computational neuroscience that would be worth reconsidering by taking synaptic dynamics into account. Most promising are attractor networks, perceptual dynamics and long term plasticity.

Bibliography

Abbott, L. F., Varela, J. A., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, 275(5297):220–224.

Albrecht, D. G. (1995). Visual cortex neurons in monkey and cat: effect of contrast on the spatial and temporal phase transfer functions. *Vis Neurosci*, 12(6):1191–1210.

Baddeley, R., Abbott, L. F., Booth, M. C., Sengpiel, F., Freeman, T., Wakeman, E. A., and Rolls, E. T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc Biol Sci*, 264(1389):1775–1783.

Barbieri, F. and Brunel, N. (2007). Irregular persistent activity induced by synaptic excitatory feedback. *Front Comput Neurosci*, 1:5.

Bauman, L. A. and Bonds, A. B. (1991). Inhibitory refinement of spatial frequency selectivity in single cells of the cat striate cortex. *Vision Res*, 31(6):933–944.

Ben-Yishai, R., Bar-Or, R. L., and Sompolinsky, H. (1995). Theory of orientation tuning in visual cortex. *Proc Natl Acad Sci U S A*, 92:3844–3848.

Ben-Yishai, R., Hansel, D., and Sompolinsky, H. (1997). Traveling waves and the processing of weakly tuned inputs in a cortical network model. *Journal of Computational Neuroscience*, 4:57–77.

Betz, W. J. (1970). Depression of transmitter release at the neuromuscular junction of the frog. *J Physiol*, 206(3):629–644.

Bibitchkov, D., Herrmann, J. M., and Geisel, T. (2002). Pattern storage and processing in attractor networks with short-time synaptic dynamics. *Network*, 13(1):115–129.

Bonds, A. B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. *Visual Neuroscience*, 2(01):41–55.

Boudreau, C. E. and Ferster, D. (2005). Short-term depression in thalamocortical synapses of cat primary visual cortex. *J Neurosci*, 25(31):7179–7190.

Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J Comput Neurosci*, 8(3):183–208.

Buracas, G. T., Zador, A. M., DeWeese, M. R., and Albright, T. D. (1998). Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron*, 20(5):959–969.

Carandini, M. and Heeger, D. J. (1994). Summation and division by neurons in primate visual cortex. *Science*, 264(5163):1333–1336.

Carandini, M., Heeger, D. J., and Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *J Neurosci*, 22(22):10053–10065.

Cavanaugh, J. R., Bair, W., and Movshon, J. A. (2002). Selectivity and spatial distribution of signals from the receptive field surround in macaque v1 neurons. *J Neurophysiol*, 88(5):2547–2556.

Chance, F. S., Nelson, S. B., and Abbott, L. F. (1998). Synaptic depression and the temporal response characteristics of v1 cells. *J Neurosci*, 18(12):4785–4799.

Churchland, M. M., Yu, B. M., Cunningham, J. P., Sugrue, L. P., Cohen, M. R., Corrado, G. S., Newsome, W. T., Clark, A. M., Hosseini, P., Scott, B. B., Bradley, D. C., Smith, M. A., Kohn, A., Movshon, J. A., Armstrong, K. M., Moore, T., Chang, S. W., Snyder, L. H., Lisberger, S. G., Priebe, N. J., Finn, I. M., Ferster, D., Ryu, S. I., Santhanam, G., Sahani, M., and Shenoy, K. V. (2010). Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci*, 13(3):369–378.

Compte, A., Brunel, N., Goldman-Rakic, P. S., and Wang, X. J. (2000). Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex*, 10(9):910–923.

Dayan, P. and Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. MIT Press.

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

Dittman, J. S. and Regehr, W. G. (1998). Calcium dependence and recovery kinetics of presynaptic depression at the climbing fiber to purkinje cell synapse. *J Neurosci*, 18(16):6147–6162.

Dragoi, V., Sharma, J., Miller, E. K., and Sur, M. (2002). Dynamics of neuronal sensitivity in visual cortex and local feature discrimination. *Nat Neurosci*, 5(9):883–891.

Felsen, G., song Shen, Y., Yao, H., Spor, G., Li, C., and Dan, Y. (2002). Dynamic modification of cortical orientation tuning mediated by recurrent connections. *Neuron*, 36(5):945–954.

Ferster, D. and Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annu Rev Neurosci*, 23:441–471.

Finn, I. M., Priebe, N. J., and Ferster, D. (2007). The emergence of contrast-invariant orientation tuning in simple cells of cat visual cortex. *Neuron*, 54(1):137–152.

Forsythe, I. D., Tsujimoto, T., Barnes-Davies, M., Cuttle, M. F., and Takahashi, T. (1998). Inactivation of presynaptic calcium current contributes to synaptic depression at a fast central synapse. *Neuron*, 20(4):797–807.

Gawne, T. J. (2008). Stimulus selection via differential response latencies in visual cortical area v4. *Neurosci Lett*, 435(3):198–203.

Gawne, T. J., Kjaer, T. W., and Richmond, B. J. (1996). Latency: another potential code for feature binding in striate cortex. *J Neurophysiol*, 76(2):1356–1360.

Gawne, T. J. and Martin, J. M. (2002). Responses of primate visual cortical v4 neurons to simultaneously presented stimuli. *J Neurophysiol*, 88(3):1128–1135.

Georgopoulos, A. P., Schwartz, A. B., and Kettner, R. E. (1986). Neuronal population coding of movement direction. *Science*, 233(4771):1416–1419.

Goldberg, J. A., Rokni, U., and Sompolinsky, H. (2004). Patterns of ongoing activity and the functional architecture of the primary visual cortex. *Neuron*, 42(3):489–500.

Goldman, M. S., Maldonado, P., and Abbott, L. F. (2002). Redundancy reduction and sustained firing with stochastic depressing synapses. *The Journal of Neuroscience*, 22(2):584–591.

Goldman, M. S., Nelson, S. B., and Abbott, L. (1999). Decorrelation of spike trains by synaptic depression. *Neurocomputing*, 26-27:147 – 153.

Hansel, D. and Sompolinsky, H. (1998). *Methods in Neuronal Modeling: From Ions to Networks*, chapter 13, pages 499–567. MIT Press.

Harris, R. E., Coulombe, M. G., and Feller, M. B. (2002). Dissociated retinal neurons form periodically active synaptic circuits. *J Neurophysiol*, 88(1):188–195.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, 9(02):181–197.

Hennig, M., Postlethwaite, M., Forsythe, I., and Graham, B. (2007). A biophysical model of short-term plasticity at the calyx of held. *Neurocomputing*, 70:1626–1629.

Hennig, M. H., Postlethwaite, M., Forsythe, I. D., and Graham, B. P. (2008). Interactions between multiple sources of short-term plasticity during evoked and spontaneous activity at the rat calyx of held. *J Physiol*, 586(13):3129–3146.

Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, 160:106–154.

Hyvarinen, A., Karhunen, J., and Oja, E. (2001). *Independent Component Analysis*. Wiley Interscience.

Kandel, E., Schwartz, J., and Jessell, T. (2000). *Principles of Neural Science, 4th ed*. McGraw-Hill.

Kanerva, P. (2009). Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive Computation*, 1:2:1866–9956.

Kenet, T., Bibitchkov, D., Tsodyks, M., Grinvald, A., and Arieli, A. (2003). Spontaneously emerging cortical representations of visual attributes. *Nature*, 425(6961):954–956.

Keysers, C., Xiao, D. K., FÃ¶ldiÃ¡k, P., and Perrett, D. I. (2005). Out of sight but not out of mind: the neurophysiology of iconic memory in the superior temporal sulcus. *Cognitive Neuropsychology*, 22:316–332.

Knight, B. W. (1972). Dynamics of encoding in a population of neurons. *J Gen Physiol*, 59(6):734–766.

Kohn, A. and Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque mt neurons. *Nat Neurosci*, 7(7):764–772.

Kusano, K. and Landau, E. M. (1975). Depression and recovery of transmission at the squid giant synapse. *J Physiol*, 245(1):13–32.

Lampl, I., Ferster, D., Poggio, T., and Riesenhuber, M. (2004). Intracellular measurements of spatial integration and the max operation in complex cells of the cat primary visual cortex. *J Neurophysiol*, 92(5):2704–2713.

Liley, A. W. and North, K. A. (1953). An electrical investigation of effects of repetitive stimulation on mammalian neuromuscular junction. *J Neurophysiol*, 16(5):509–527.

Markram, H. and Tsodyks, M. (1996). Redistribution of synaptic efficacy between neocortical pyramidal neurons. *Nature*, 382(6594):807–810.

Markram, H., Wang, Y., and Tsodyks, M. (1998). Differential signaling via the same axon of neocortical pyramidal neurons. *Proc Natl Acad Sci U S A*, 95(9):5323–5328.

Maunsell, J. H. and Essen, D. C. V. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. i. selectivity for stimulus direction, speed, and orientation. *J Neurophysiol*, 49(5):1127–1147.

McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., and Moser, M.-B. (2006). Path integration and the neural basis of the 'cognitive map'. *Nat Rev Neurosci*, 7(8):663–678.

Mercer, A., West, D. C., Morris, O. T., Kirchhecker, S., Kerkhoff, J. E., and Thomson, A. M. (2005). Excitatory connections made by presynaptic cortico-cortical pyramidal cells in layer 6 of the neocortex. *Cereb Cortex*, 15(10):1485–1496.

Miller, E. K., Gochin, P. M., and Gross, C. G. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Res*, 616(1-2):25–29.

Mongillo, G., Barak, O., and Tsodyks, M. (2008). Synaptic theory of working memory. *Science*, 319(5869):1543–1546.

Nelson, S. B. (1991). Temporal interactions in the cat visual system. i. orientation-selective suppression in the visual cortex. *J Neurosci*, 11(2):344–356.

Oram, M. W. and Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *J Neurophysiol*, 68(1):70–84.

Oram, M. W., Wiener, M. C., Lestienne, R., and Richmond, B. J. (1999). Stochastic nature of precisely timed spike patterns in visual system neuronal responses. *J Neurophysiol*, 81(6):3021–3033.

Oram, M. W., Xiao, D., Dritschel, B., and Payne, K. R. (2002). The temporal resolution of neural codes: does response latency have a unique role? *Philos Trans R Soc Lond B Biol Sci*, 357(1424):987–1001.

Owen, M. R., Laing, C. R., and Coombes, S. (2007). Bumps and rings in a two-dimensional neural field: splitting and rotational instabilities. *New Journal of Physics*, 9:378.

Pantic, L., Torres, J. J., Kappen, H. J., and Gielen, S. C. A. M. (2002). Associative memory with dynamic synapses. *Neural Comput*, 14(12):2903–2923.

Paradiso, M. A. (1988). A theory for the use of visual orientation information which exploits the columnar structure of striate cortex. *Biol Cybern*, 58(1):35–49.

Pasupathy, A. and Connor, C. E. (2001). Shape representation in area v4: position-specific tuning for boundary conformation. *J Neurophysiol*, 86(5):2505–2519.

Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., and Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Exp Brain Res*, 86(1):159–173.

Perrett, D. I., Rolls, E. T., and Caan, W. (1982). Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res*, 47(3):329–342.

Perrett, D. I., Xiao, D., Barraclough, N. E., Keysers, C., and Oram, M. W. (2009). Seeing the future: Natural image sequences produce "anticipatory" neuronal activity and bias perceptual report. *Q J Exp Psychol (Colchester)*, 62(11):2081–2104.

Pfister, J.-P., Dayan, P., and Lengyel, M. (2010). Synapses with short-term plasticity are optimal estimators of presynaptic membrane potentials. *Nat Neurosci*, 13(10):1271–1275.

Pouget, A., Dayan, P., and Zemel, R. (2000). Information processing with population codes. *Nat Rev Neurosci*, 1(2):125–132.

Pouille, F., Marin-Burgin, A., Adesnik, H., Atallah, B. V., and Scanziani, M. (2009). Input normalization by global feedforward inhibition expands cortical dynamic range. *Nat Neurosci*, 12(12):1577–1585.

Puccini, G. D., Sanchez-Vives, M. V., and Compte, A. (2007). Integrated mechanisms of anticipation and rate-of-change computations in cortical circuits. *PLoS Comput Biol*, 3(5):e82.

Rajan, K., Abbott, L. F., and Sompolinsky, H. (2010). Stimulus-dependent suppression of chaos in recurrent neural networks. *Phys. Rev. E*, 82(1):011903.

Recanzone, G. H., Wurtz, R. H., and Schwarz, U. (1997). Responses of mt and mst neurons to one and two moving objects in the receptive field. *J Neurophysiol*, 78(6):2904–2915.

Reig, R., Gallego, R., Nowak, L. G., and Sanchez-Vives, M. V. (2006). Impact of cortical network activity on short-term synaptic depression. *Cereb Cortex*, 16(5):688–695.

Riesenhuber, M. and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat Neurosci*, 2(11):1019–1025.

Riesenhuber, M. and Poggio, T. (2002). Neural mechanisms of object recognition. *Curr Opin Neurobiol*, 12(2):162–168.

Ringach, D. L. (2010). Population coding under normalization. *Vision Res*, 50:2223–2232.

Rizzoli, S. O. and Betz, W. J. (2005). Synaptic vesicle pools. *Nat Rev Neurosci*, 6(1):57–69.

Rolls, E. T. and Tovee, M. J. (1995). Sparseness of the neuronal representation of stimuli in the primate temporal visual cortex. *J Neurophysiol*, 73(2):713–726.

Sahani, M. and Dayan, P. (2003). Doubly distributional population codes: simultaneous representation of uncertainty and multiplicity. *Neural Comput*, 15(10):2255–2279.

Salinas, E. and Abbott, L. F. (1994). Vector reconstruction from firing rates. *J Comput Neurosci*, 1(1-2):89–107.

Samsonovich, A. and McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *J Neurosci*, 17(15):5900–5920.

Sanchez-Vives, M. V., Nowak, L. G., and McCormick, D. A. (2000). Membrane mechanisms underlying contrast adaptation in cat area 17 in vivo. *J Neurosci*, 20(11):4267–4285.

Sato, T. (1989). Interactions of visual stimuli in the receptive fields of inferior temporal neurons in awake macaques. *Exp Brain Res*, 77(1):23–30.

Sclar, G. and Freeman, R. D. (1982). Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Exp Brain Res*, 46(3):457–461.

Sclar, G., Maunsell, J. H., and Lennie, P. (1990). Coding of image contrast in central visual pathways of the macaque monkey. *Vision Res*, 30(1):1–10.

Senn, W., Wannier, T., Kleinle, J., LÃŒscher, H. R., MÃŒller, L., Streit, J., and Wyler, K. (1998). Pattern generation by two coupled time-discrete neural networks with synaptic depression. *Neural Comput*, 10(5):1251–1275.

Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., and Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:411–426.

Seung, H. S. and Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proc Natl Acad Sci U S A*, 90(22):10749–10753.

Shadlen, M. N. and Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci*, 18(10):3870–3896.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656.

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. *J Neurophysiol*, 57(3):773–786.

Smith, M. A. and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. *J Neurosci*, 28(48):12591–12603.

Somers, D. C., Nelson, S. B., and Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *J Neurosci*, 15(8):5448–5465.

Teich, M. C. (1989). Fractal character of the auditory neural spike train. *IEEE Transactions on Biomedical Engineering*, 36:150–160.

Teich, M. C., Johnson, D. H., Kumar, A. R., and Turcott, R. G. (1990). Rate fluctuations and fractional power-law noise recorded from cells in the lower auditory pathway of the cat. *Hearing Reasearch*, 46:41–52.

Teich, M. C., Turcott, R. G., and Siegel, R. M. (1996). Temporal correlation in cat striate-cortex neural spike trains. *IEEE Engineering in Medicine and Biology*, 96:79–87.

Thomson, A. M. (1997). Activity-dependent properties of synaptic transmission at two classes of connections made by rat neocortical pyramidal axons in vitro. *J Physiol*, 502 (Pt 1):131–147.

Thomson, A. M. and Deuchars, J. (1994). Temporal and spatial properties of local circuits in neocortex. *Trends Neurosci*, 17(3):119–126.

Thomson, A. M. and Lamy, C. (2007). Functional maps of neocortical local circuitry. *Front Neurosci*, 1(1):19–42.

Treue, S., Hol, K., and Rauber, H. J. (2000). Seeing multiple directions of motion-physiology and psychophysics. *Nat Neurosci*, 3(3):270–276.

Treves, A. and Rolls, E. T. (1991). What determines the capacity of autoassociative memories in the brain? *Network: Computation in Neural Systems*, 2:4:371–397.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput*, 10(4):821–835.

Tsodyks, M., Uziel, A., and Markram, H. (2000). Synchrony generation in recurrent networks with frequency-dependent synapses. *J Neurosci*, 20(1):RC50.

Tsodyks, M. V. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc Natl Acad Sci U S A*, 94(2):719–723.

van der Meer, M. A. A., Knierim, J. J., Yoganarasimha, D., Wood, E. R., and van Rossum, M. C. W. (2007). Anticipation in the rodent head direction system can be explained by an interaction of head movements and vestibular firing properties. *J Neurophysiol*, 98(4):1883–1897.

van Rossum, M. C. W., Turrigiano, G. G., and Nelson, S. B. (2002). Fast propagation of firing rates through layered networks of noisy neurons. *J Neurosci*, 22(5):1956–1966.

van Rossum, M. C. W., van der Meer, M. A. A., Xiao, D., and Oram, M. W. (2008). Adaptive integration in the visual cortex by depressing recurrent cortical circuits. *Neural Comput*, 20(7):1847–1872.

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

van Vreeswijk, C. and Sompolinsky, H. (1998). Chaotic balanced state in a model of cortical circuits. *Neural Comput*, 10(6):1321–1371.

van Wezel, R. J., Lankheet, M. J., Verstraten, F. A., Maree, A. F., and van de Grind, W. A. (1996). Responses of complex cells in area 17 of the cat to bi-vectorial transparent motion. *Vision Res*, 36(18):2805–2813.

Varela, J. A., Sen, K., Gibson, J., Fost, J., Abbott, L. F., and Nelson, S. B. (1997). A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *J Neurosci*, 17(20):7926–7940.

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

Wang, Y., Markram, H., Goodman, P. H., Berger, T. K., Ma, J., and Goldman-Rakic, P. S. (2006). Heterogeneity in the pyramidal network of the medial prefrontal cortex. *Nat Neurosci*, 9(4):534–542.

Wong, A. Y. C., Graham, B. P., Billups, B., and Forsythe, I. D. (2003). Distinguishing between presynaptic and postsynaptic mechanisms of short-term depression during action potential trains. *The Journal of Neuroscience*, 23(12):4868–4877.

York, L. C. and van Rossum, M. C. W. (2009). Recurrent networks with short term synaptic depression. *J Comput Neurosci*, 27(3):607–620.

Yu, A. J., Giese, M. A., and Poggio, T. A. (2002). Biophysiologically plausible implementations of the maximum operation. *Neural Comput*, 14(12):2857–2881.

Zemel, R. S. and Dayan, P. (1999). Distributional population codes and multiple motion models. In *Proceedings of the 1998 conference on Advances in neural information processing systems II*, pages 174–180, Cambridge, MA, USA. MIT Press.

Zemel, R. S., Dayan, P., and Pouget, A. (1998). Probabilistic interpretation of population codes. *Neural Computation*, 10(2):403–430.

Zhang, K. (1996). Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J Neurosci*, 16(6):2112–2126.

Zoccolan, D., Cox, D. D., and DiCarlo, J. J. (2005). Multiple object response normalization in monkey inferotemporal cortex. *J Neurosci*, 25(36):8150–8164.

Zucker, R. S. and Regehr, W. G. (2002). Short-term synaptic plasticity. *Annu Rev Physiol*, 64:355–405.