Modelling Divisive Inhibition in the Primary Visual Cortex from Optogenetic Circuit Mapping Data

Hong Noh

Linacre College

MSc by Research Thesis

Submitted Michaelmas Term 2011

Word Count: 5149 without appendix / 7550 with appendix

Abstract

Recurrent networks are abundant in the neocortex and are recognised as a means of amplifying feedforward thalamic sensory inputs. However, when operating at high gain, which is necessary for this signal amplification, the standard recurrent network firing rate model suffers from increased reaction times to rapidly changing stimuli. Divisive inhibition has been proposed as a means of bypassing this coupling of system gain and time constant. In my thesis I focus on the importance of inhibition in recurrent networks in visual information processing. This was motivated by a recent study where the presence and absence of translaminar inhibition distinguished cells in the primary visual cortex. I apply several divisive inhibition schemes to an existing recurrent network model of simple and complex cells. The schemes are studied analytically and also simulated to assess how well they can be integrated into this existing model whilst simultaneoulsy solving the coupled system gain and time constant problem. Though each scheme has its benefits, I propose that a mixture of schemes is likely in real physiology.

Acknowledgements

I would first like to thank Dr. Dennis Kätzel for his unwavering guidance throughout the project, not only in physiology but also the scientific thought process in general. I am grateful to Professor Gero Miesenböck for taking me in as a research student despite my shortcomings in neurobiology. I thank Professor Larry Abbott for the snippet of C code used to simulate recurrent networks, and Dr. Frances Chance for her suggestions about divisive inhibition. I am most grateful to Dr. Wyeth Bair, for meeting with me on several occasions to discuss models of the visual cortex, especially during his move. My time in the Miesenböck laboratory would not have been as gratifying without the company of the Miesenböck lab members: I thank Andrew Lin, Clara Ferreira, Diogo Pimentel, Eleftheria Vrontou, Evan Harrell, Fiona Woods, Gero, Jeff Donlea, the two Pauls (Overton and Coe), Shamik Dasgupta and Yan Tan for engaging conversations about neuronal circuits, academia, world affairs and squash. I particularly thank Diogo for his constant updates on the latest technology (neuroscience related or otherwise) and Jeff about news from conferences around the globe. The many discussions with Eleftheria at ungodly hours about eletrophysiology will be missed. I thank the EPSRC and the Doctoral Training Centre for funding my studies in Oxford. Finally, I thank my parents and my brother who have supported my endeavours in science.

Declaration

I declare that I have produced this thesis myself in its entirety, under supervision of Professor Gero Miesenböck. All analyses contained in this thesis have been done by myself unless stated otherwise.

Contents

List of Figures **5**

1 Introduction **6**

2 Firing Rate Network Models **8**

1 Visual Information Processing 12

2 Visual Stimuli 12

3 Simple Cells 14

4 Complex Cells 15

3 Optogenetic Mapping of Inhibitory Input Sources for Neocortical Excitatory Cells **18**

4 Divisive Inhibition Schemes **20**

1 Constant Divisive Inhibition 20

2 Dynamic Divisive Inhibition 21

2.1 The Feedback Model 22

2.2 The Disinhibition Model 22

2.3 The Feedforward Model 23

2.4 Feedback-Feedforward Model 24

5 Simulations **26**

1 F_0, F_1 and F_2 Components of the Network Response 27

6 Conclusions & Future Work **30**

7 Mathematical Appendix **32**

A Cosine Similarity Metric 32

B Useful Identities 32

B.1 Trigonometric Identities 32

B.2 Integral Identities 33

C Special Cases of Spatial Receptive Fields 33

C.1 Simple Cell Spatial Receptive Field with $\sigma_x = \sigma_y = \sigma$ and $\Theta = 0$ . . 33

C.2 Simple Cell Spatial Receptive Field with $\sigma_x = \sigma_y = \sigma$ and $\phi = \Phi = 0$ 36

D Stability Analysis & Conditions on Interneuron Activity 39

D.1 The Feedback Model 40

D.2 The Disinhibition Model 40

D.3 The Feedforward Model 41

D.4 The Feedback-Feedforward Model 41

E Short-Term Synaptic Plasticity 41

E.1 Synaptic Depression 43

E.2 Synaptic Facilitation 43

F MATLAB Simulations 44

F.1 Without Divisive Inhibition 44

F.2 With Divisive Inhibition 45

F.3 With Short-Term Synaptic Plasticity 45

F.4 MATLAB Code for Dynamics Simulation 46

F.5 Extracting F_0, F_1 and F_2 Components of the Network Response . . 56

Bibliography 58

List of Figures

2.1 Feedforward Firing Rate Networks 9

2.2 Recurrent Firing Rate Networks 10

2.3 Counterphase Grating Visual Stimuli 13

2.4 The Gabor Filter 14

2.5 Critical Slowing Down in Response to Counterphase and Drifting Grating 17

4.1 Various Divisive Inhibition Schemes 21

5.1 Response of a Model Neuron with and without Divisive Inhibition to Counterphase Grating 27

5.2 Response of a Model Neuron with and without Divisive Inhibition to Drifting Grating 28

5.3 F_1/F_0 and F_1/F_2 Measures of a Cell to Drifting and Counterphase Grating 29

7.1 Synaptic Depression Dynamic Variable 42

Chapter 1

Introduction

Recurrent networks are ubiquitous in the neocortex and its function in sensory information processing has been studied extensively in the literature. For example in the primary visual cortex (V1), recurrent excitatory connections have been shown to increase orientation selectivity [2, 9, 28] or decrease spatial phase sensitivity [4] by strongly interconnecting cells with similar orientation or dissimilar spatial phase preferences. Other studies have focussed on the mathematical properties of these networks whose attractor dynamics make them ideal as pattern storage facilities (e.g. [16]). However, a central problem with recurrent firing rate networks is the coupling between the effective time constant and the gain of the system where increasing the gain, which is necessary for many phenomena attributed to recurrent amplification, will result in the time constant increasing directly in proportion. Therefore, as the gain of the system increases, the network reacts increasingly sluggishly to rapidly changing stimuli. Here we consider a general approach to decoupling the time constant and gain of this system inspired by studies where divisive inhibition of the recurrent network activity has been suggested as a means to do exactly this [3]. We assess each scheme based on achievable system gain, network stability, effective system time constant and interneuron stability. We apply the inhibition schemes to an existing recurrent firing rate network model of cells in V1 [4] through simulations and measure any inconsistencies between analytical results and actual performance of the various models using Fourier decomposition.

The model communicated in [4] is an augmentation of the model proposed by Hubel and Wiesel in 1962 where the'edge-detector'-like paralinear simple receptive fields are

formed by sequential superposition of on- and off-centre receptive fields of the lateral geniculate nucleus (LGN) [15]. The authors of [4] show that complex receptive fields that do not have well-defined on and off regions can be formed by connecting cells with simple receptive fields and similar spatial frequency preferences but different spatial phase preferences in a recurrent network. The motivation to focus specifically on this system stems from a recent study using optogenetic mapping of V1, in which cells found to receive strong translaminar inhibition in layer VB and strong intralaminar inhibition in layers II/III corresponded to the complex cell population [1].

The report hereafter is divided into four main chapters and a mathematical appendix. In Chapter 2, we derive the coupling between the gain and effective time constant in recurrent firing rate networks as in [3] and introduce the model used in [4]. In Chapter 3, we briefly discuss the data in [1], from which this study is inspired. Subsequently in Chapter 4, divisive inhibition schemes both new and extant in the literature are analysed in detail. The divisive inhibition schemes from Chapter 4 are simulated in Chapter 5 and the results discussed. Conclusions from the study and future work suggestions are given in Chapter 6. The mathematical appendix is given in Chapter 7.

Chapter 2

Firing Rate Network Models

For the purpose of this study, we focus on firing rate network models that take the form

$$\tau_r \frac{d\boldsymbol{r}}{dt} = -\boldsymbol{r} + F(\boldsymbol{W}^{ff} \cdot \boldsymbol{u}) \tag{2.1}$$

where $\boldsymbol{u}$ is the thalamic input to the neurons, $\boldsymbol{r}$ represents the firing rate of all neurons considered, τ_r is the rate of recovery of the neuron without external input, $\boldsymbol{W}^{ff}$ is the weighting of the various inputs into the neurons and F is the activation function [6]. Throughout this report we use the notation that W_{ij} indicates the weight *from* neuron j *to* neuron i. This is the basic structure of a feedforward network, in which the neurons receive only feedforward thalamic input according to some weight matrix. Though F may take several forms, such as half-wave rectification or the logistic function, we investigate the case when it is a simple linear unit transformation, such that $F(x) = x$. Figure 2.1, top, shows the feedforward network with three neurons and two thalamic information sources. Here, $\boldsymbol{u} \in \mathbb{R}^2$, $\boldsymbol{r} \in \mathbb{R}^3$ and naturally the weight matrix $\boldsymbol{W}^{ff} \in \mathbb{R}^{2\times 3}$. The vector operation $\boldsymbol{W}^{ff} \cdot \boldsymbol{u}$ can be simplified to return $\boldsymbol{W}^{ff} \cdot \boldsymbol{u} = \boldsymbol{I} \in \mathbb{R}^3$: Now each neuron now receives a weighted input from a fictional input source $\boldsymbol{I}$ (Figure 2.1, bottom). By augmenting (2.1) such that the activation function has as input arguments both the feedforward thalamic inputs and the firing rates of the neurons themselves we arrive at the

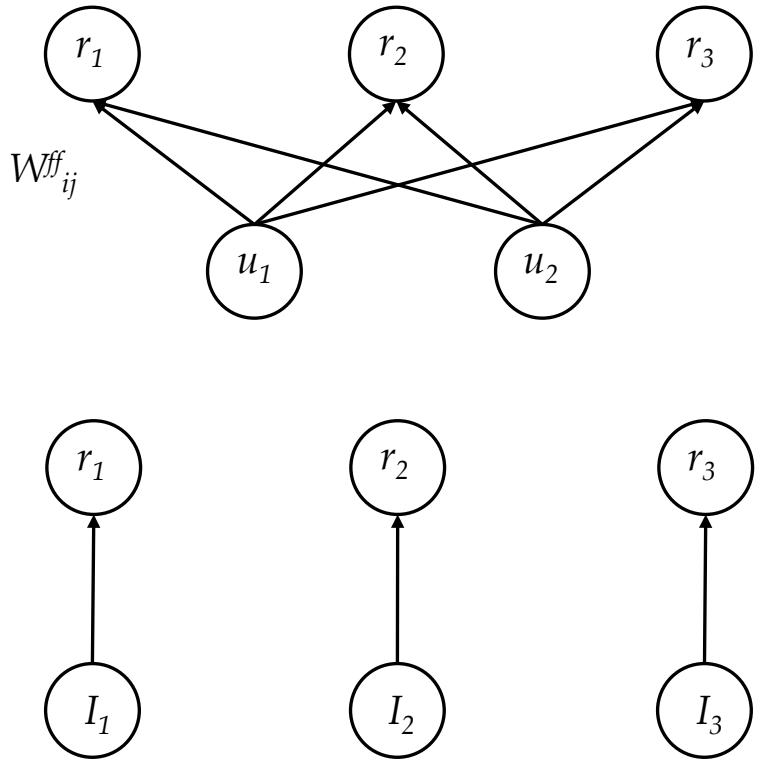


Figure 2.1: Feedforward firing rate networks described by (2.1). Top: There are three neurons and two thalamic information sources. The feedforward weight matrix, denoted $\boldsymbol{W}^{ff}$, determines the strength of the inputs from the thalamic source to the neurons. Bottom: The multiplication $\boldsymbol{W}^{ff} \cdot u$ of (2.1) can be represented by inputs $\boldsymbol{I}$ that are equal in number to the neurons.

mathematical form of a recurrent firing rate network

$$\begin{aligned} \tau_r \frac{d\boldsymbol{r}}{dt} &= -\boldsymbol{r} + F(\boldsymbol{W}^{ff} \cdot \boldsymbol{u} + \boldsymbol{W}^{rec} \cdot \boldsymbol{r}) \\ &= -\boldsymbol{r} + F(\boldsymbol{I} + \boldsymbol{W}^{rec} \cdot \boldsymbol{r}) \end{aligned} \tag{2.2}$$

This network architecture is shown in Figure 2.2 for which $\boldsymbol{W}^{rec} \in \mathbb{R}^{3\times 3}$. For a linear unit transformation activation function, (2.2) becomes

$$\tau_r \frac{d\boldsymbol{r}}{dt} = -\boldsymbol{r} + \boldsymbol{I} + \boldsymbol{W}^{rec} \cdot \boldsymbol{r} \tag{2.3}$$

Dale's Law states that postsynaptic synapses must be either excitatory or inhibitory but not both [6]. Mathematically this translates to the matrix $\boldsymbol{W}^{rec}$ ($\boldsymbol{W}$ hereafter) having columns that are of the same sign. However, this matrix is often taken to be symmetric

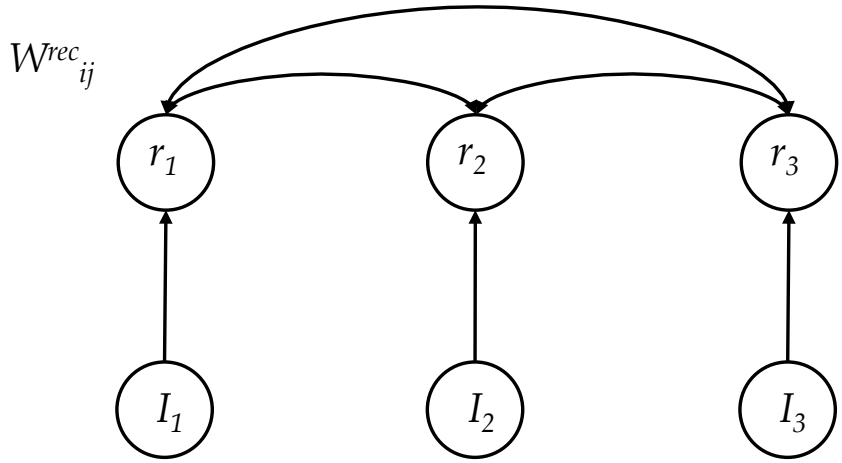


Figure 2.2: Recurrent firing rate network described by (2.2). To satisfy Dale's Law, the recurrent weight matrix $\boldsymbol{W}^{rec}$ ($\boldsymbol{W}$ hereafter) must have row elements that are all of the same sign. However, for mathematical convenience the matrix is often taken as symmetric: Symmetric matrices have eigenvalues that are all real and eigenvectors are orthogonal.

for mathematical convenience: The eigenvalues of a symmetric matrix are all real (though not necessarily unique) and the eigenvectors are orthogonal, which can be normalised to be orthonormal. In general, for any matrix $\boldsymbol{W}$, we can carry out eigendecomposition such that

$$\boldsymbol{W}^T \boldsymbol{e}^\mu = \lambda^\mu \boldsymbol{e}^\mu \tag{2.4}$$

where $\boldsymbol{e}^\mu$ and λ^μ are the μ^{th} eigenvector and eigenvalue respectively. In other words, we can find the μ^{th} eigenproperties of a matrix by multiplying it by its μ^{th} eigenvector. Multiplying (2.3) by the μ^{th} eigenvector we have

$$\tau_r \frac{d(\boldsymbol{e}^\mu)^T \boldsymbol{r}}{dt} = -(\boldsymbol{e}^\mu)^T \boldsymbol{r} + (\boldsymbol{e}^\mu)^T \boldsymbol{I} + (\boldsymbol{e}^\mu)^T (\boldsymbol{W} \cdot \boldsymbol{r}) \tag{2.5}$$

Now we can write $(\boldsymbol{e}^\mu)^T (\boldsymbol{W} \cdot \boldsymbol{r}) = (\boldsymbol{W}^T \boldsymbol{e}^\mu)^T \boldsymbol{r}$. Following (2.4), $\boldsymbol{W}^T \boldsymbol{e}^\mu = \lambda^\mu \boldsymbol{e}^\mu$ and hence the above becomes

$$\tau_r \frac{d(\boldsymbol{e}^\mu)^T \boldsymbol{r}}{dt} = -(\boldsymbol{e}^\mu)^T \boldsymbol{r} + (\boldsymbol{e}^\mu)^T \boldsymbol{I} + \lambda^\mu (\boldsymbol{e}^\mu)^T \boldsymbol{r} \tag{2.6}$$

Keeping with the notation in [3], let us denote $\boldsymbol{r} = \sum_\nu a_\nu \boldsymbol{e}^\nu$ where a_ν are constants (the eigendecomposition of the rate vector $\boldsymbol{r}$) then

$$(\boldsymbol{e}^\mu)^T \boldsymbol{r} = (\boldsymbol{e}^\mu)^T \sum_\nu a_\nu \boldsymbol{e}^\nu = \sum_\nu a_\nu (\boldsymbol{e}^\mu)^T \boldsymbol{e}^\nu \tag{2.7}$$

Since the matrix $\boldsymbol{W}^{rec}$ is symmetric, eigenvectors $\boldsymbol{e}^\nu$ (and equally $\boldsymbol{e}^\mu$) are orthonormal and hence the above simplifies to the μth constant a_μ, that is,

$$(\boldsymbol{e}^\mu)^T \boldsymbol{r} = a_\mu \tag{2.8}$$

Substituting this into (2.6), and defining $h_\mu = \sum_i I_i e_i^\mu$, we are left with

$$\tau_r \frac{da_\mu}{dt} = -a_\mu + h_\mu + \lambda^\mu a_\mu \tag{2.9}$$

For such a linear system, let us check three characteristics:

- The gain, defined as the resulting ratio a_μ^*/h_μ where the * denotes steady-state activity, [3]
- The stability of the system, found by taking the derivative of the system at steady-state by the variable whose stability is in question: a_μ^*,
- The effective time constant, given by rearranging the equation in the form $\tau_{eff}(da_\mu/dt) = -a_\mu + \text{const.}$

We see that for (2.9), the gain is given by $1/(1-\lambda^\mu)$, the system is stable for $\lambda^\mu < 1$ and the effective time constant is given by $\tau_r/(1-\lambda^\mu)$ [3]. Herein lies the problem. For the system to operate at high gain, λ^μ must be less than but close to 1. However, having a recurrent weight matrix that satisfies this condition will result in the system operating close to instability and a large effective time constant. Essentially, we wish to decouple the gain and the time constant as well as ensure that the system does not operate dangerously close to being unstable. The pervasive nature of recurrent connections in the neocortex highlights the need to address this problem given the assumption that the mathematical formulation discussed thus far is close to true physiology. The sections hereafter describe how the presence of cells with simple and complex receptive fields in the primary visual

cortex can arise through feedforward and recurrent firing rate networks.

1 Visual Information Processing

In visual information processing, axons from the retina terminate in the LGN of the thalamus, from which cells project onto an area known as the visual cortex. Within this area, information is processed in a hierarchical manner starting from the primary visual cortex (area V1). Like other cortical areas, cells in V1 can be functionally classified into either excitatory or inhibitory neurons: Excitatory neurons use the neurotransmitter glutamate that depolarises, while many inhibitory neurons release γ-aminobutyric acid (GABA), which hyperpolarises, the postsynaptic cell. Throughout the cerebral cortex these cells are organised in layers labelled layer I to VI from the pial surface through to the white matter. Information processing in V1 starts at layer IV, where spiny stellate cells receive the majority of inputs from the LGN and distribute information to cells in layers II and III. Axons of excitatory pyramidal cells in layers II and III project to pyramidal cells in layer V, whose axons in turn project to pyramidal cells in layer VI. Cells in layers II and III also receive feedback from pyramidal cells in layer V as do inhibitory cells in layer IV from layer VI [17].

Neurons in both the retina and LGN are known have either on-centre or off-centre receptive fields, regions in the visual field that excite the cell, formed by concentric rings of inhibitory and excitatory regions [17]. In their seminal study, Hubel and Wiesel showed that neurons in the cat visual cortex can be functionally classified according to their receptive fields: Some neurons receive feedforward input from the LGN to exhibit 'simple' receptive fields that have parallel straight-line boundaries between the 'on' (excitatory) and 'off' (inhibitory) regions whereas other have 'complex' receptive fields with no clear excitatory or inhibitory zones [15].

2 Visual Stimuli

To define receptive fields, neurons are traditionally stimulated with a sinusoidal grating with varying orientation, spatial frequency, spatial phase and temporal frequency as its

activity is recorded. Visual stimuli can be described mathematically as

$$s(x, y, t) = A\cos\left(Kx\cos\Theta + Ky\sin\Theta - \Phi\right)\cos(\omega t) \tag{2.10}$$

where A, K, Θ, Φ and ω are the stimulation amplitude, spatial frequency, orientation, spatial phase and temporal frequency respectively (Figure 2.3). Variables x and y define the two-dimensional visual field [6]. With respect to these parameters, the aforementioned simple and complex receptive fields of cells in V1 correspond to cells' sensitivity and insensitivity to spatial-phase. Simple cells will only respond to stimuli with spatial phase that coincide with their receptive fields whereas complex cells will be less sensitive.

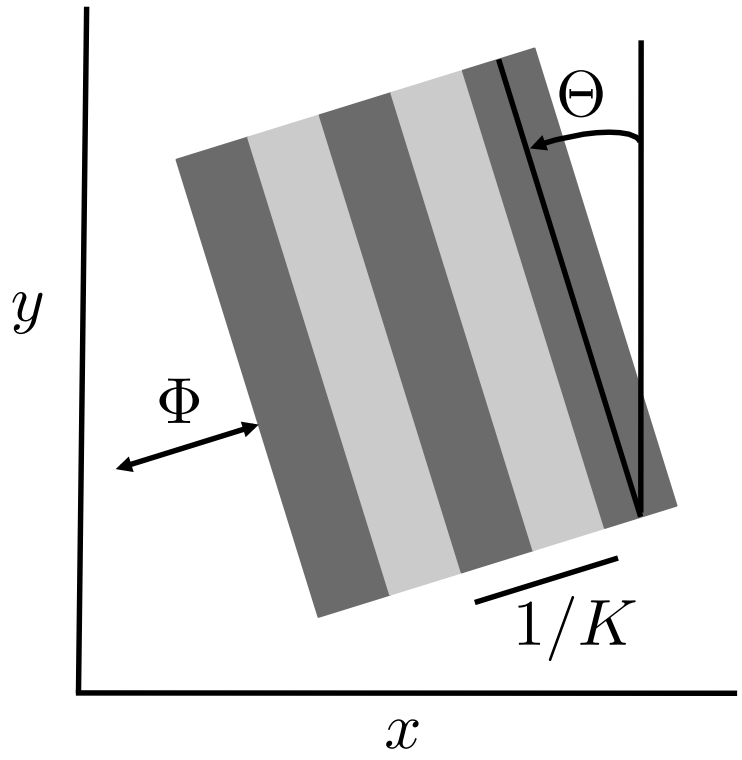


Figure 2.3: Counterphase grating described by (2.10) where light ($s > 0$) and dark ($s < 0$) regions of the sinusoid are shown as a square wave (adapted from [6]). Parameters Φ, K and Θ are the spatial phase, spatial frequency and orientation respectively. The temporal frequency ω controls the frequency with which light and dark areas oscillate at a specific spatial location (not shown). The spatial frequency K has units of cycles per degrees, which must be inverted to express in degrees per cycle. Simple cells exhibit spatial-phase sensitivity whereas complex cells do not.

3 Simple Cells

The receptive field of a neuron can be classified separable or nonseparable depending on whether the spatial and temporal receptive fields can or cannot be separated as independent functions [6]. In this study, we assume separable receptive fields of simple cells for mathematical simplicity. The response of a simple cell to spatial parameters, namely spatial frequency and spatial phase, is adequately described using a Gabor function of the form

$$G(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp\left(-\frac{x^2}{2\sigma_x^2}-\frac{y^2}{2\sigma_y^2}\right)\cos\left(kx-\phi\right) \tag{2.11}$$

Here, parameters σ_x and σ_y control the size of the receptive field in the x and y directions and k and ϕ are the preferred spatial frequency and spatial phase of the simple cell [6]. The Gabor function for parameters $\sigma_x = 1.4$, $\sigma_y = 1.6$, $k = 0.5$ cycles/degree, $\phi = 1.5$ cycles is shown in Figure 2.4. For the temporal response of simple cells, the following

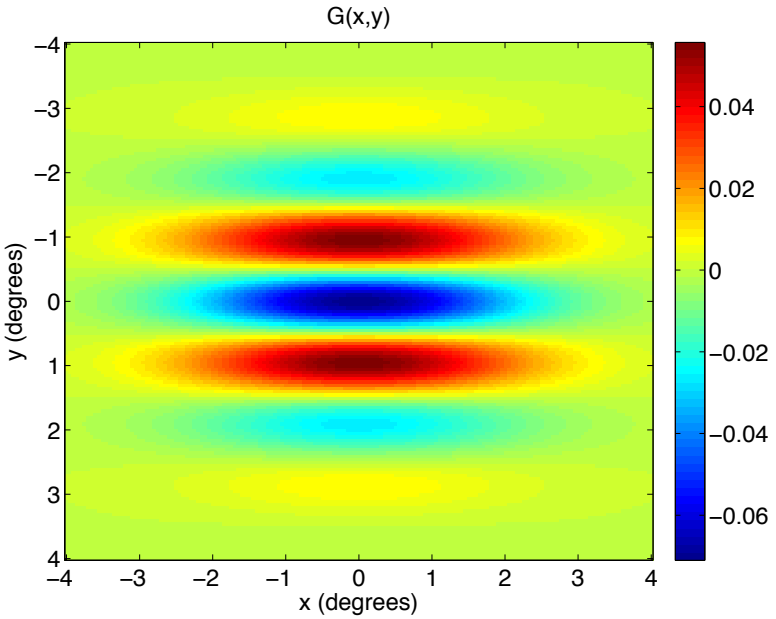


Figure 2.4: The Gabor function described by (2.11) with parameters $\sigma_x = 1.4$, $\sigma_y = 1.6$, $k = 0.5$ cycles/degree, $\phi = 1.5$ cycles for $-4^\circ \leq x, y \leq 4^\circ$. As expected, the plot shows the oscillatory function wrapped in a Gaussian-distribution envelope.

function is used

$$H(\tau) = \alpha\exp\left(-\alpha\tau\right)\left(\frac{(\alpha\tau)^5}{5!}-\frac{(\alpha\tau)^7}{7!}\right) \tag{2.12}$$

for $\tau \geq 0$. $H(\tau)$ is defined zero for any $\tau < 0$. The parameter α controls the reaction time of the cell. Since we have assumed separability of the receptive field, response of a simple cell to a counterphase grating is given by

$$R(t) = \int_0^\infty \left[\int G(x,y)s(x,y,t-\tau)\,dx\,dy \right] H(\tau)\,d\tau \tag{2.13}$$

In brief, the stimulus is filtered by the gabor function $G(x,y)$ and the resulting function is convoluted with (2.12), the temporal response function. In the special case where the receptive field has the same size in both the x and y directions such that $\sigma_x = \sigma_y = \sigma$ and if the spatial phase of the stimulus and of the receptive field are both zero ($\Phi = \phi = 0$), then the spatial receptive field simplifies to

$$A \exp\left(-\frac{\sigma^2(k^2+K^2)}{2} \right) \cosh\left(\sigma^2 kK \cos\Theta\right) \tag{2.14}$$

If instead, the grating is at orientation $\Theta = 0$, and if k is chosen such that $k\sigma$ is significantly large, the receptive field is now given by

$$\frac{A}{2} \exp\left(-\frac{\sigma^2(k-K)^2}{2} \right) \cos(\phi - \Phi) \tag{2.15}$$

Here simple cells show a bell-shaped response to spatial frequency and a cosine response to spatial phase as the stimuli diverge in these parameters from the cells' preferred values (from [6]). Derivations for (2.14) and (2.15) are provided in sections C.2 and C.1 of the appendix respectively.

4 Complex Cells

Complex cell behaviour is suggested to arise from strong recurrent excitation of otherwise simple cells [4]. This is shown with a rate-model in which the firing rate of cell i, r_i, follows the differential equation identical in nature to (2.3)

$$\tau_r \frac{dr_i}{dt} = -r_i + I_i^{ff} + I_i^{rec} \tag{2.16}$$

where I_i^{ff} and I_i^{rec} are the feedforward and recurrent inputs. The feedforward input takes the form

$$I_i^{ff} = \mathcal{A}\left[\int G_i(x)\left\{\int_0^\infty H(\tau)s(x,t-\tau)d\tau\right\}dx\right]_+ \tag{2.17}$$

where $\mathcal{A}$ controls the feedforward input strength, $G_i(x)$, $H(\tau)$ and $s(x,t-\tau)$ are the Gabor spatial receptive field in the x direction for cell i only, the temporal response function (2.12) and the stimulus (2.10) again in the x direction only. The notation $[]_+$ indicates half-wave rectification. Without preferred orientations and with $\exp(-\sigma^2 kK) \approx 0$, the feedforward input becomes

$$I_i^{ff} = \mathcal{A}\left[M_i \cdot \int_0^\infty H(\tau)\cos(\omega(t-\tau))\,d\tau\right]_+ \tag{2.18}$$

where M_i is the constant from (C.6) for counterphase grating, and (C.11) for drifting grating stimuli for the ith cell, that is, the constant $\mathcal{M}$ calculated from (C.6) and (C.11) with spatial frequency and phase preferences $k = k_i$ and $\phi = \phi_i$. Recurrent input to model cell i is given by

$$I_i^{rec} = \frac{g_i}{N-1}\sum_{j\neq i}^{N}\left(2\exp\left(-\frac{(k_i-k_j)^2}{2\sigma_c^2}\right) - \exp\left(-\frac{(k_i-k_j)^2}{2\sigma_s^2}\right)\right)r_j \tag{2.19}$$

where $0 \leq g_i < g_{max}$ with g_{max} being the value at which the network is unstable. The spatial frequency tuning curve is essentially a difference-of-Gaussians (DoG) filter, the width of which is given by σ_c and $\sigma_s (> \sigma_c)$ the centre and surround variances. The DoG filter is designed so neurons with similar spatial frequency preferences excite each other, whilst those with different spatial frequency preferences inhibit each other [4]. The parameter g_i controls the strength of the recurrent connectivity. If g_i is zero, then the simple cells only receive the feedforward input and maintains spatial-phase sensitivity. As g_i increases closer to g_{max}, the strength of the recurrent connectivity increases and the simple cells will become increasingly spatial-phase invariant. In general, if the eigenvalues of a matrix W are given by λ^μ, then the eigenvalues of the matrix gW, where g is some constant multiplied to all entries of W, are given by $g\lambda^\mu$. By defining

$$\tilde{W}_{ij} = \frac{1}{N-1}\sum_{j\neq i}^{N}\left(2\exp\left(-\frac{(k_i-k_j)^2}{2\sigma_c^2}\right) - \exp\left(-\frac{(k_i-k_j)^2}{2\sigma_s^2}\right)\right) \tag{2.20}$$

the firing rate of cell i is given by

$$\tau_r \frac{dr_i}{dt} = -r_i + I_i^{ff} + g_i \sum_{j=1}^{N} \tilde{W}_{ij} r_j \tag{2.21}$$

As before, the system described by (2.21) is stable if $g_i \tilde{\lambda}^{\mu} < 1$ where $\tilde{\lambda}^{\mu}$ are eigenvalues of $\tilde{W}_{ij}$ and g_i is assumed to be equal for all i. Therefore, g_{max} where the network becomes unstable is given bys $g_{max} = 1/\tilde{\lambda}^{\mu}$. Figures 4.1 a) and b) show simple cells (black triangles) activated by the thalamic input only (orange arrows) and complex cells with both feedforward and recurrent activation. The effect of the coupled effective time constant

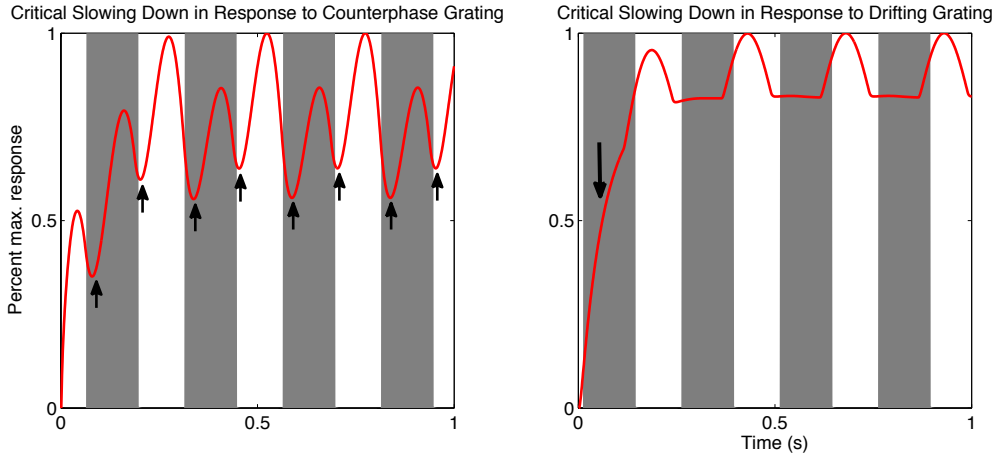


Figure 2.5: Critical slowing down in response to counterphase (left) and drifting (right) grating stimuli. The response of a cell in the recurrent network described by (2.3) is shown in red. Sinusoidal stimulus intensities are shown rectified. Areas of high ($s > 0$) and low ($s < 0$) intensity are shown as white and grey respectively. For counterphase grating, critical slowing down manifests itself as a phase lag between the stimuli (black arrows pointing up) whilst for drifting grating critical slowing down increases the time taken for the network to reach steady-state conditions (black arrow pointing down).

and gain for complex cells in recurrent firing rate networks is shown in Figure 2.5. For counterphase grating stimuli, the delay at high gain manifests as a phase lag between the stimuli and the response. For drifting grating stimuli, the increased effective time constant is shown as a delay in reaching steady-state activity.

Chapter 3

Optogenetic Mapping of Inhibitory Input Sources for Neocortical Excitatory Cells

Channelrhodopsin-2 (ChR2) is a light-activated cation conducting channel found in the alga *Chlamydomonas reinhardtii* [11]. Since ions involved in the electrical excitation of neurons are cations (e.g. Na^+ or Ca^{2+}) and the photosensitivity is encoded in a gene less than 1kb long, ChR2 is an ideal candidate to be used in light-activation of genetically specified excitable cells [32, 14]. Studies using ChR2 are prevalent in the literature, examples ranging from light activated behavioural analyses in the nematode *C. elegans* to photostimulation used to map neuronal connectivity in mice [23, 25]. In a recent study, a Gad65-Cre-driver was used to direct ChR2 expression to inhibitory neurons in V1[1]. These cells were activated using 473nm blue light, while monitoring inhibitory inputs into an excitatory postsynaptic cell via whole-cell recordings in voltage-clamp mode [1]. For all excitatory cells recorded, a cluster analysis using the cosine similarity metric (A.1) on the normalised inhibitory charge flow separated all pyramidal cells in V1, S1 and M1 into those that receive strong translaminar inhibition and those that receive strong intralaminar inhibition (Figure 7 of [1]). For all pyramidal cells in V1, the ratio of cells that receive strong putative feedback to intralaminar inhibition was found to be 4:11 in layers II/III and 8:4 in layer VB, which approximately match the proportion of simple to

complex cells in these layers (1:3 and 3:1 respectively [24]) raising the possibility that inhibition is indeed involved in formation of receptive fields. On the other hand, cells that receive the dominant proportion of inhibition from the home layer are inhibited from a broader lateral domain. Such intralaminar inhibition may be responsible for simple cell receptive field structure (though a feedforward model of simple cells without lateral inhibition is given in [26]). Nonspecific receptive fields of complex cells, on the contrary, are thought to result from loss of specificity by the recurrent connectivity of simple cells [4]. However, with many models of recurrent excitatory connectivity, as the ratio of network output to input, or the gain of the network, increases, cells in the network become sluggish in their responses to rapidly changing signals [3]. This problem of recurrent network models can be overcome by introducing divisive feedback inhibition mechanisms that reduces the recurrent excitation with increasing feedforward input [3]. It is conceivable that the translaminar putative feedback from the target layer to putative complex cells serve exactly this purpose, assuming it is in fact recurrent connectivity that gives rise to complex cell behaviour.

Chapter 4

Divisive Inhibition Schemes

1 Constant Divisive Inhibition

In [3], the authors propose divisive inhibition as a method of circumventing the coupled high gain and critical slowing down problem. The firing rate of a cell in a recurrent network with divisive inhibition is given by

$$\tau_r \frac{dr_i}{dt} = -r_i + I_i^{ff} + \frac{\sum_j W_{ij} r_j}{R} \tag{4.1}$$

where R is the firing rate of the inhibitory neuron. Taking eigendirections of (4.1),

$$\tau_r \frac{da_\mu}{dt} = -a_\mu + h_\mu + \frac{\lambda^\mu a_\mu}{R} \tag{4.2}$$

Upon observation, by simply substituting $\lambda^\mu = \lambda^\mu / R$, the same conditions for network stability, effective time constant and gain of the network apply for (4.2) as they did for (2.9). This is true for any multiplicative form of *constant* inhibition R,

$$\tau_r \frac{dr_i}{dt} = -r_i + I_i^{ff} + f(R) \sum_j W_{ij} r_j \tag{4.3}$$

since the factor $f(R)$ is simply multiplied with all elements of the recurrent weight matrix W_{ij}. For any constant divisive inhibition, $f(R)\lambda^\mu < 1$ will result in a stable system and $f(R)\lambda^\mu > 1$ in an unstable system and for $f(R)\lambda^\mu$ close to one, though at high gain, the effective time constant $\tau_r / (1 - f(R)\lambda^\mu)$ will increase, causing critical slowing down.

2 Dynamic Divisive Inhibition

The previous section showed that constant divisive inhibition cannot achieve the desired effect of solving the coupled high gain and critical slowing down problem. As in [3], treating R as a dynamic variable allows more flexibility. Since R is a function of time, in addition to considering the gain, stability and effective time constant of the recurrent network, we also consider the stability of the interneuron. In the following sections, we assess the feasibility of various interneuron dynamics of the general form $R(t) = v(a_1(t), h_1(t))$, that is, the interneuron activity is some function v of the principal eigendirection of the recurrent activity a_1 and feedforward input h_1.

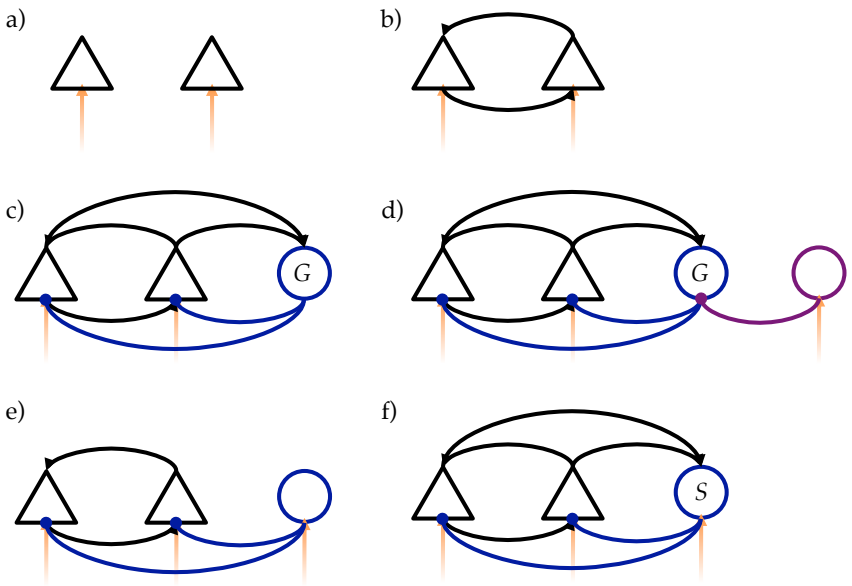


Figure 4.1: Various circuit architectures discussed herein. a) Simple cells (black triangles) receive thalamic feedforward input (orange arrows) only. b) In the model proposed by Chance *et al.*, simple cells connected in a recurrent network architecture gives rise to complex cell behaviour [4]. c) Interneuron connectivity described by (4.4) in the 'Feedback Model'. Again simple cells are connected in a recurrent network. However, unlike b) now the interneuron (blue circle) samples the weighted sum (weights controlled by parameter G) of the complex cells and inhibits appropriately. d) Interneuron connectivity described by (4.7) in the 'Disinhibition Model'. Recurrently connected simple cells are inhibited by an interneuron (blue circle) that again samples network behaviour but is further inhibited by another interneuron that receives feedforward input (violet circle). e) The 'Feedforward Model' described by (4.10), where an interneuron receives feedforward input and inhibits appropriately. f) The 'Feedback-Feedforward Model', in which the interneuron is activated by not only sampling the network activity but also from direct thalamic activation, coupled in a nonlinear manner as shown by (4.13).

2.1 The Feedback Model

The authors of [3] first propose R to follow

$$\tau_R \frac{dR}{dt} = -R + G \sum_{i=1}^{N} e_i^1 r_i \tag{4.4}$$

That is, the single inhibitory neuron is excited by the activity of cells in the recurrent network weighted by some constant G in the principal eigenvector direction. We refer to this model as the 'Feedback Model' and its network architecture is shown in Figure 4.1 c). The steady-state of (4.4) is $R^* = Ga_1$, which after substituting into (4.2) with $\mu = 1$ returns

$$\tau_r \frac{da_1}{dt} = -a_1 + h_1 + \frac{\lambda^1}{G} \tag{4.5}$$

Now, checking for critical slowing down

1. At steady-state, we have $a_1^* = h_1 + \lambda^1/G$, which results in unit gain.
2. $\tau_{eff} = \tau_r$ and hence there is no slowing down
3. In section D.1 of the appendix, we show that the interneuron is stable for

$$2(Ga_1^* + Q) - Ga_1^* + \frac{Q^2}{Ga_1^*} < \lambda^1 < 2(Ga_1^* + Q) \tag{4.6}$$

   The above requirement requires a balance between the network size (represented by λ^1) and the base interneuron activity (Q).

Whilst this circuit may not suffer from slowing down and remain stable for realistic conditions, it cannot achieve the necessary high gain.

2.2 The Disinhibition Model

In their second proposal, the authors suggest the following

$$\tau_R \frac{dR}{dt} = -R + \frac{G \sum_i e_i^1 r_i}{\sum_i e_i^1 I_i^{ff} + Q} \tag{4.7}$$

where Q is a constant added to ensure the denominator does not become zero. Physically, the inhibitory neuron with firing rate $R(t)$ is further inhibited by another inhibitory

neuron whose firing rate is the weighted sum of the feedforward inputs, with weights described by the principal eigenvector. We call this the 'Disinhibition Model', and its architecture is shown in Figure 4.1 d). The steady-state value of (4.7) is $R^* = Ga_1/(h_1 + Q)$, which after substituting into (4.2) and again with $\mu = 1$, returns

$$\tau_r \frac{da_1}{dt} = -a_1 + h_1 + \frac{\lambda^1(h_1 + Q)}{G} \tag{4.8}$$

Checking for critical slowing down

1. At steady-state, $a_1^* = h_1(1 + \lambda^1/G) + \text{const.}$, so high gain is achieved for large λ^1 or small G.

2. $\tau_{eff} = \tau_r$ and therefore the network does not exhibit slowing down

3. In section D.2 of the appendix we show that for interneuron stability we require

$$2\left(\frac{Ga_1^*}{h_1 + Q} + Q\right) - \frac{Ga_1^*}{h_1 + Q} + \frac{Q^2(h_1 + Q)}{Ga_1^*} < \lambda^1 < 2\left(\frac{Ga_1^*}{h_1 + Q} + Q\right) \tag{4.9}$$

   The condition above understandably balances the network size λ^1 with the base interneuron firing rate (Q) and feedforward input h_1.

For the 'Disinhibition Model', the decoupling of high gain and the effective time constant can only take place for large eigenvalues or reduced input strengths into the interneuron that inhibits the circuit divisively (G).

2.3 The Feedforward Model

In the simplest case consider a single inhibitory neuron, which receives the same feedforward input as the complex cells but weighted by the principal eigenvector of the weight matrix as shown in Figure 4.1 e), such that

$$\tau_R \frac{dR}{dt} = -R + h_1 \tag{4.10}$$

Rewriting the network activity for the principal eigendirection only substituting the steady state $R^* = h_1$ into (4.2) for $\mu = 1$,

$$\tau_r \frac{da_1}{dt} = -a_1 + h_1 + \frac{\lambda^1 a_1}{h_1} \tag{4.11}$$

Again assessing the above system for critical slowing down,

1. At steady-state, $a_1^* = h_1 \left(1/(1 - \lambda^1/h_1)\right)$. The gain is now dependent on the feedforward input and high gain is achieved for $\lambda^1 \sim h_1$ $(\lambda^1 < h_1)$.

2. $\tau_{eff} = \tau_r/(1 - \lambda^1/h_1)$. Treating h_1 as a constant, critical slowing down is prevented by having $\lambda^1/h_1 \ll 1$.

3. In section D.3 of the appendix we show that the interneuron is stable for

$$\lambda^1 < h_1 + Q \tag{4.12}$$

   This restricts the size of the network (λ^1) relative to the feedforward input (h_1) and the base interneuron firing rate (Q).

Despite easily achieving a stable interneuron, the 'Feedforward Model' cannot decouple the gain and effective time constant.

2.4 Feedback-Feedforward Model

Extending the 'Feedforward Model' is the 'Feedback-Feedforward Model' where the feedforward thalamic input and the feedback from the recurrent network are combined multiplicatively as shown in Figure 4.1 f), such that

$$\tau_R \frac{dR}{dt} = -R + a_1 h_1 \tag{4.13}$$

In the principal eigendirection, after substituting $R^* = a_1 h_1$, we have for the network

$$\tau_r \frac{da_1}{dt} = -a_1 + h_1 + \frac{\lambda^1}{h_1} \tag{4.14}$$

Assessing (4.14) for critical slowing down,

1. At steady-state, $a_1^* = h_1\left(1 + \lambda^1/h_1^2\right)$. The gain is again feedforward input-dependent and high gain is achieved for low h_1.

2. $\tau_{eff} = \tau_r$ and therefore the network does not exhibit slowing down.

3. In section D.4 of the appendix we show that the interneuron is stable for

$$2(a_1^* h_1 + Q) - a_1^* h_1 + \frac{Q^2}{a_1^* h_1} < \lambda^1 < 2(a_1^* h_1 + Q) \tag{4.15}$$

The 'Feedback-Feedforward Model' appears to meet all the conditions without any caveats. However, the feedforward input-dependence on the gain is unrealistic and raises doubts about the model representing real physiology. A summary of the above systems is given in Table 6.1. The 'Feedback Model' is immediately disqualified as a plausible scheme since the network cannot exhibit gains greater than one. The 'Disinhibition Model' requires a large principal eigenvalue, that is, a λ^1 large. This is achieved in large networks. The 'Feedforward Model' not only imposes contradictory conditions on the size of the network with respect to the thalamic input, it also restricts the size in exchange for stability and is therefore an unrealistic network architecture for large neuronal network structures. The requirement that λ^1 must be similar to but smaller than h_1 for high gain yet $\lambda^1/h_1 \ll 1$ are contradictory and therefore this architecture cannot decouple the gain and effective time constant. The 'Feedback-Feedforward Model' limits the feedforward input such that h_1 is kept low, which may not be possible in physiology.

Chapter 5

Simulations

In order to effectively include divisive inhibition to the firing rate models described in [4], the response of $N = 256$ cells to drifting and counterphase grating of 4Hz was simulated in MATLAB as described in section F of the appendix. For simplicity simulations were carried out with $\tilde{W}_{ij} = 1/(N-1)$ ($\tilde{W}_{ij}$ of (2.20) is now a block matrix) such that $\tilde{\lambda}^1 = N/(N-1)$ and $\lambda^1 = g_i\tilde{\lambda}^1 = g_iN/(N-1)$. For each divisive inhibition scheme, the system of equations to be simulated is given by

$$\begin{aligned} \tau_r \frac{dr_i}{dt} &= -r_i + I_i^{ff} + \frac{\sum_j W_{ij} r_j}{UR + Q} \\ \tau_R \frac{dR}{dt} &= -R + Sv(a_1, h_1) \end{aligned} \tag{5.1}$$

Now, common parameters in all systems took the following values

- Q = 1: This is the baseline firing rate of the interneuron.
- $\tau_R = 5$ms (as in [4, 3])
- G = 0.1 (as in [3])
- $\tau_D = 300$ms (as in [4])
- $\tau_F = 80$ms
- $d = 0.95$ (as in [4])
- $S = 1\text{s}^{-2}$ for the 'Feedback-Feedforward Model'. This is needed in non-dimensionalising the firing rate of the interneuron in (4.13).

- $U = 1\text{s}^{-1}$ for the 'Feedback' and 'Feedforward' models. This is needed in non-dimensionalising the firing rate of the interneuron in (4.4) and (4.10).

For the 'Feedback Model', studies have shown that somatostatin-positive cells in L2/3 of the somatosensory cortex inhibit recurrent networks of pyramidal cells in a supralinear manner [18]. We assume that similar mechanisms are in place in the visual cortex and therefore we augment (4.4) with nonlinear synaptic faciliation. The mathematical description of this is shown in section E.2. The resulting system is shown in (E.8). The effect of divisive inhibition on the spatial phase invariance of complex cells is shown in Figures 5.1 and 5.2 for counterphase and drifting grating stimuli respectively. In all simulations we record the normalised firing rate.

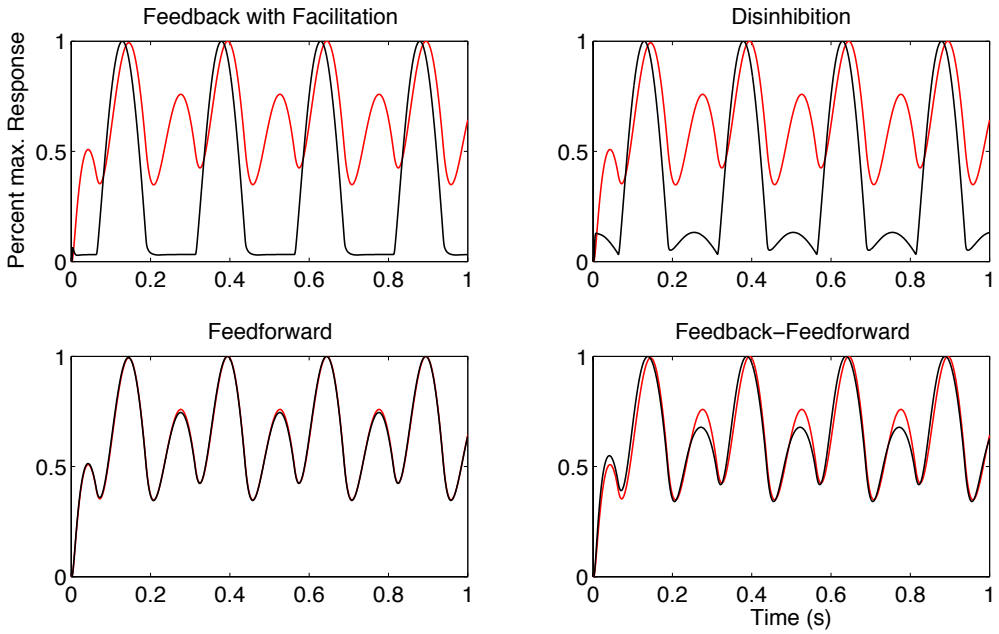


Figure 5.1: Response of a model neuron with (black) and without (red) divisive inhibition in a network of $N = 256$ simple cells to a counterphase grating stimuli of 4Hz at gain 16 ($g = 15/16$) for the 'Feedback', 'Disinhibition', 'Feedforward' and 'Feedback-Feedforward' models. The 'Feedback' model is augmented with facilitation as described in (E.8).

1 F_0, F_1 and F_2 Components of the Network Response

A traditional method of quantifying the 'complexness' of a V1 cell for drifting grating stimuli is to calculate the ratio between the amplitude of the fundamental frequency

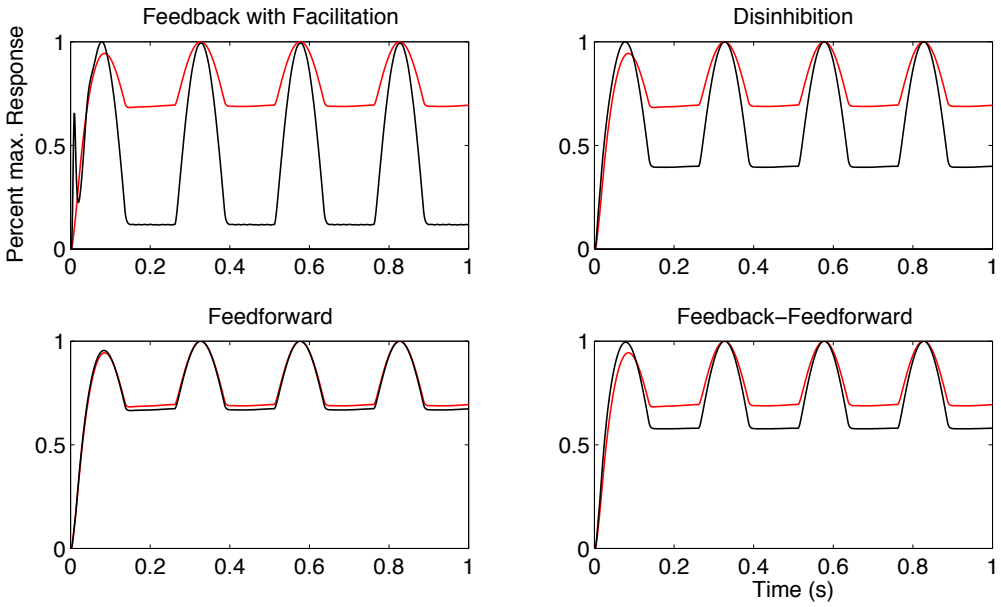


Figure 5.2: Response of a model neuron with (black) and without (red) divisive inhibition in a network of $N = 256$ simple cells to a drifting grating stimuli of 4Hz at gain 16 ($g = 15/16$) for the 'Feedback', 'Disinhibition', 'Feedforward' and 'Feedback-Feedforward' models. The 'Feedback' model is augmented with facilitation as described in (E.8).

component (F_1) and the mean (F_0) of its firing rate response [27, 22]. In essence, this measures the size of the cell's response at the same frequency as the drifting stimulus relative to the mean of this response. Skottun *et al.* showed that this ratio, also known as the *relative modulation*, forms a bimodal distribution with dip at unity where simple cells constitute the mode greater than unity and complex cells constitute the mode less than unity [27]. Also of interest is the second harmonic component, which has been used by several studies to measure the extent of frequency doubling in a cell's response (for example [19, 8, 31]). Dean and Tolhurst found that for counterphase grating stimuli the ratio of the second harmonic component to the fundamental frequency component (F_2/F_1) resulted in a weakly bimodal distribution, again with dip at unity: Complex cells have $F_2/F_1 > 1$ and simple cells $F_2/F_1 < 1$ [7] (see [13] for review). Authors of [4] use the inverse of the F_2/F_1 ratio such that, in similar fashion to relative modulation, simple cells have $F_1/F_2 > 1$ and complex cells have $F_1/F_2 < 1$. Here we follow [4] by using relative modulation and F_1/F_2 to assess the 'complexness' of cells in architectures that incorporate the divisive inhibition schemes discussed in chapter 4. The result of this analysis is shown in Figure 5.3.

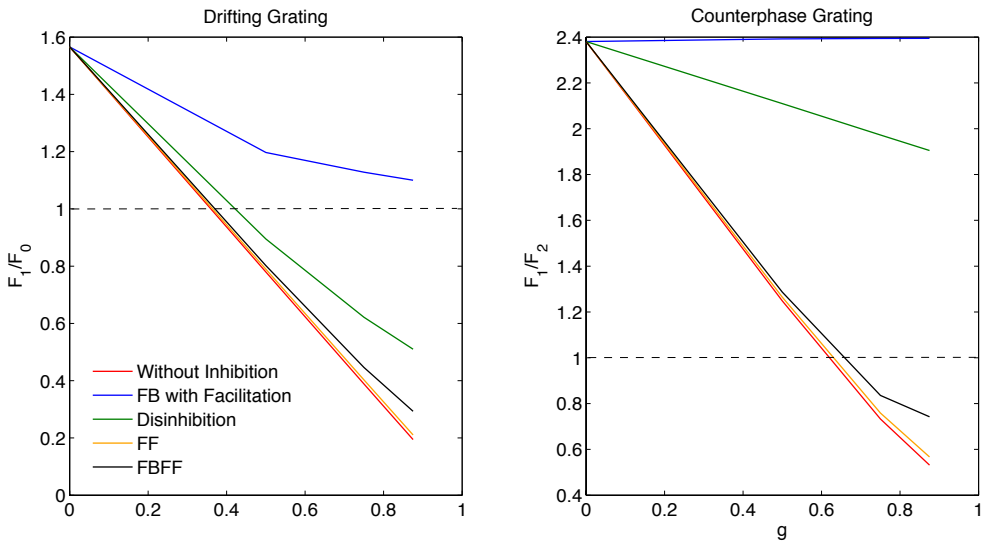


Figure 5.3: F_1/F_0 and F_1/F_2 for response of a cell to drifting and counterphase grating respectively for the 'Feedback' with facilitation (blue), 'Disinhibition' (green), 'Feedforward' (orange) and 'Feedback-Feedforward' (black) models as well as the recurrent network without inhibition (red) plotted against the parameter $g_i = g, \forall i$, which controls the strength of recurrent connectivity. When $g = 0$, cells are fully simple and as $g \to 1$ cells become more complex and unstable. The network becomes unstable at $g = 1$. Only the 'Feedback-Feedforward' and 'Feedforward' models cross the $F_1/F_0 = F_1/F_2 = 1$ thresholds where a cell is deemed to be complex ([22, 27] for F_1/F_0 and [4, 7] for F_1/F_2).

Chapter 6

Conclusions & Future Work

In this report, we introduced the problem of coupled effective time constant and gain of a typical recurrent firing rate network in the neocortex and examined both existing and new divisive inhibition mechanisms that may solve this problem. Based on optogenetic mapping data from a recent study that showed translaminar inhibition may play a role specifically in the formation of complex receptive fields of cells in V1 [1], we applied the aforementioned divisive inhibition schemes on an existing recurrent network model of complex cell behaviour [4]. Table 6.1 shows a summary of results from the analysis provided in section 2. The 'Feedback Model' is disqualified as high gain cannot be achieved and the 'Feedforward Model' restricts the size of the network in exchange for stability as well as imposing contradictory conditions for decoupling the gain and effective time constant. The 'Feedback-Feedforward Model' can solve the problem for low thalamic inputs. However, the feedforward input-dependence of the system gain does not demonstrate the threshold linear activation function recurrent networks typically display [10]. The 'Disinhibition Model' takes a close form to the inhibition scheme necessary to achieve this linear input-output relationship and only requires that the prin-

Table 6.1: Summary of Divisive Inhibition Schemes

Inhibition Scheme	Gain	Network Stability	Slowing Down
Feedback	NO	YES	YES
Disinhibition	large λ^1/small G	YES	YES
Feedforward	$\lambda^1 \sim h_1$	$\lambda^1 < h_1 + Q$	$\lambda^1 \ll h_1$
Feedback-Feedforward	low h_1	YES	YES

cipal eigenvalue is large - a condition easily accomplished with large networks. Yet, how much of real physiology the disinhibition system represents remains unclear.

Upon observing the results of the F_1/F_0 and F_1/F_2 ratio extractions from simulating the aforementioned inhibition systems (Figure 5.3), it appears that the 'Feedback-Feedforward' divisive inhibition scheme is most optimal in maintaining the 'complexity' of the cells whilst solving the coupled time constant and gain problem. In essence, the 'Feedback-Feedforward Model' sacrifices the threshold linear activation function characteristic of the recurrent network but can show 'complex' behaviour as measured by ratios F_1/F_0 and F_1/F_2. On the other hand, the 'Disinhibition Model' maintains a linear activation function but does not exhibit enough 'complexity'. Which is the correct model? The models in this thesis are by no means mutually exclusive, and neuronal networks in the neocortex may be a weighted combination of several models discussed here and otherwise. Furthermore, statistical re-evaluation of data has shown weak bimodality of the F_2/F_1 measure and that even the stronger bimodality of the F_1/F_0 measure need not indicate a simple/complex dichotomy [21]. Therefore, in order to accurately model the role inhibition plays in recurrent networks, not only must more data be collected but also detailed *in vivo* investigations that question the fundamental definition of 'simple' and 'complex' cells must take place. Future work based on this report also requires *in vivo* experiments to investigate further the exact form of connection the interneuron(s) make onto the pyramidal cells in V1. This could be done by designing experiments where interneuron activity is blocked (possibly using genetically targeted optical inhibition [12, 5]) during presentation of counterphase and grating stimuli and the dynamics of complex cells recorded. The purpose of such an experiment would be to assess whether inhibition does indeed play a role in complex receptive field formation. Theoretically, the network models could be expanded to include spiking neurons similar in nature to [20] or [29], which may elucidate mechanisms of inhibition inaccessible with firing rate models.

Chapter 7

Mathematical Appendix

A Cosine Similarity Metric

The cosine similarity metric between two vectors **a** and **b** both necessarily of length N is given by

$$\cos\theta = \frac{\mathbf{a}^T\mathbf{b}}{\sqrt{\mathbf{a}^T\mathbf{a}}\sqrt{\mathbf{b}^T\mathbf{b}}} \tag{A.1}$$

In essence the metric finds the angle between the two vectors in N-dimensional space. If the two vectors are identical in direction (since they are both of length N, this also means they are identical), $\cos\theta = \mathbf{a}^T\mathbf{a}/\mathbf{a}^T\mathbf{a} = 1$. If the two vectors are opposite in direction, $\cos\theta = -\mathbf{a}^T\mathbf{a}/\mathbf{a}^T\mathbf{a} = -1$.

B Useful Identities

The following identities will be used throughout the next section.

B.1 Trigonometric Identities

$$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B \tag{B.1}$$

$$\cos A \cos B = \frac{1}{2}\left[\cos(A - B) + \cos(A + B)\right] \tag{B.2}$$

B.2 Integral Identities

$$\begin{aligned}\int_{-\infty}^{\infty} \exp(-ax^2)\cos(bx)\,dx &= \mathcal{R}\left\{\int_{-\infty}^{\infty} \exp(-ax^2)\exp(ibx)\right\} \\ &= \mathcal{R}\left\{\int_{-\infty}^{\infty} \exp(-ax^2+ibx)\right\} \\ &= \exp\left(-\frac{b^2}{4a}\right)\mathcal{R}\left\{\int_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\left(\sqrt{2a}x-\frac{ib}{\sqrt{2a}}\right)^2\right)\,dx\right\}\end{aligned}$$

where $\mathcal{R}\{\cdot\}$ indicates taking the real part. The last equality results from completing the square. Now since

$$\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2}{2\sigma^2}\right)\,dx = 1$$

we have, for $\sigma = 1$,

$$\int_{-\infty}^{\infty}\exp\left(-ax^2\right)\,\cos(bx)\,dx = \sqrt{\frac{\pi}{a}}\exp\left(\frac{-b^2}{4a}\right) \tag{B.3}$$

C Special Cases of Spatial Receptive Fields

C.1 Simple Cell Spatial Receptive Field with $\sigma_x = \sigma_y = \sigma$ and $\Theta = 0$

Counterphase Grating

Since we are not concerned with time we omit the time-dependent terms of (2.10). Now, the spatial receptive field of a simple cell, which we assign M, is given by

$$\begin{aligned}M &= \frac{1}{2\pi\sigma^2}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2-y^2}{2\sigma^2}\right)\,\cos(kx-\phi)\cdot A\cos(Kx-\Phi)\,dx\,dy \\ &= A\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\,\cos(kx-\phi)\cos(Kx-\Phi)\int_{-\infty}^{\infty}\frac{1}{2\pi\sigma^2}\exp\left(\frac{-y^2}{2\sigma^2}\right)\,dy\,dx \\ &= \frac{A}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\,\cos(kx-\phi)\cos(Kx-\Phi)\,dx\end{aligned} \tag{C.1}$$

We use (B.2), with $A = kx-\phi$ and $B = Kx-\Phi$. Therefore,

$$\begin{aligned}M &= \frac{A}{2\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2}{2\sigma^2}\right)\cdot\cos(kx-\phi-Kx+\Phi)\,dx \\ &+ \frac{A}{2\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2}{2\sigma^2}\right)\cdot\cos(kx-\phi+Kx-\Phi)\,dx = I+J\end{aligned} \tag{C.2}$$

For I we have

$$I = \frac{A}{2\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) \cos((k-K)x + \Phi - \phi)\, dx$$

Now using the identity $\cos(A+B) = \cos A \cos B - \sin A \sin B$ with $A = (k-K)x$ and $B = \Phi - \phi$,

$$\begin{aligned} I = &\frac{A}{2\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) \cos((k-K)x) \cos(\Phi - \phi)\, dx \\ &- \frac{A}{2\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) \sin((k-K)x) \sin(\Phi - \phi)\, dx \end{aligned} \tag{C.3}$$

However, for the second term in (C.3), since $\sin((k-K)x)$ is an odd function, $\int \exp(-x^2/2\sigma^2)\sin((k-K)x) = 0$ and hence

$$I = \frac{A}{2\sqrt{2\pi\sigma^2}} \cos(\Phi - \phi) \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) \cos((k-K)x)\, dx$$

We can use (B.3) for the integral above, which returns

$$I = \frac{A}{2} \exp\left(-\frac{(k-K)^2\sigma^2}{2}\right) \cos(\Phi - \phi) \tag{C.4}$$

For J, following a similar procedure we find

$$J = \frac{A}{2\sqrt{2\pi\sigma^2}} \cos(-\Phi - \phi) \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right) \cos((k+K)x)\, dx$$

Again using (B.3) we have

$$J = \frac{A}{2} \exp\left(-\frac{(k+K)^2\sigma^2}{2}\right) \cos(-\Phi - \phi) \tag{C.5}$$

Substituting back (C.4) and (C.5) into (C.2),

$$M = I + J = \frac{A}{2} \exp\left(-\frac{(k-K)^2\sigma^2}{2}\right) \cos(\Phi - \phi) + \frac{A}{2} \exp\left(-\frac{(k+K)^2\sigma^2}{2}\right) \cos(-\Phi - \phi)$$

We now assume $\exp(-\sigma^2 kK) \approx 0$ such that $J \approx 0$. The resulting expression for M is

$$M = \frac{A}{2}\exp\left(-\frac{(k-K)^2\sigma^2}{2}\right)\cos(\Phi-\phi) = \frac{A}{2}\exp\left(-\frac{(k-K)^2\sigma^2}{2}\right)\cos(\phi-\Phi) \quad \text{(C.6)}$$

since $\cos(-A) = \cos(A)$.

Drifting Grating

In the case of drifting grating, M is given by

$$M = \int G(x,y)\, A\cos(Kx-\omega t)\, dx\, dy$$

Setting $\sigma_x = \sigma_y = \sigma$,

$$\begin{aligned} M &= \frac{A}{2\pi\sigma^2}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2-y^2}{2\sigma^2}\right)\cos(kx-\phi)\cdot\cos(Kx-\omega t)\, dx\, dy \\ &= \frac{A}{2\pi\sigma^2}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos(kx-\phi)\cdot\cos(Kx-\omega t)\left[\int_{-\infty}^{\infty}\exp\left(\frac{-y^2}{2\sigma^2}\right)dy\right]dx \\ &= \frac{A}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos(kx-\phi)\cdot\cos(Kx-\omega t)\, dx \end{aligned}$$

Using (B.1) on $\cos(Kx-\omega t)$,

$$\begin{aligned} M = &\frac{A}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos(kx-\phi)\cos(Kx)\cos(\omega t)\, dx \\ &+\frac{A}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos(kx-\phi)\sin(Kx)\sin(\omega t)\, dx \end{aligned}$$

Since $\sin(Kx)$ is an odd function, the second integral above is zero and we have

$$M = \frac{A\cos(\omega t)}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos(kx-\phi)\cos(Kx)\, dx \quad \text{(C.7)}$$

(B.2) is used on (C.7) to return

$$M = \frac{A\cos(\omega t)}{2\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\left[\cos(kx-\phi-Kx)-\cos(kx-\phi+Kx)\right]\, dx \quad \text{(C.8)}$$

Now, again from (B.1)

$$\cos((k \pm K)x - \phi) = \cos((k \pm K)x)\cos\phi + \sin((k \pm K)x)\sin\phi \tag{C.9}$$

Substituting (C.9) into (C.8) and noting that $\sin((k-K)x)$ is odd with respect to x, we have for M

$$M = \frac{A\cos(\omega t)}{2\sqrt{2\pi\sigma^2}} \left[\int_{-\infty}^{\infty} \exp\left(\frac{-x^2}{2\sigma^2}\right) \cos((k-K)x)\cos\phi\, dx - \int_{-\infty}^{\infty} \exp\left(\frac{-x^2}{2\sigma^2}\right) \cos((k+K)x)\cos\phi\, dx \right]$$

In general,

$$\int_{-\infty}^{\infty} \exp(-ax^2)\cos(bx)\, dx = \sqrt{\frac{\pi}{a}} \exp\left(\frac{-b^2}{4a}\right)$$

Hence,

$$M = A\cos(\omega t)\cos\phi \left[\exp\left(\frac{-(k-K)^2\sigma^2}{2}\right) - \exp\left(\frac{-(k+K)^2\sigma^2}{2}\right) \right] \tag{C.10}$$

Again, assuming $\exp(-\sigma^2 kK) \approx 0$,

$$M = A\cos(\omega t)\cos\phi \left[\exp\left(\frac{-(k-K)^2\sigma^2}{2}\right) \right] \tag{C.11}$$

This is identical to (2.15)

C.2 Simple Cell Spatial Receptive Field with $\sigma_x = \sigma_y = \sigma$ and $\phi = \Phi = 0$

Counterphase Grating

Now, in (B.1) set $\sigma_x = \sigma_y = \sigma$ and $\Theta = 0$. The full expression for M is then

$$M = \int G(x,y) A \cos(Kx\cos\Theta + Ky\sin\Theta - \Phi)\, dxdy \tag{C.12}$$

where $G(x,y)$ is as shown in (2.11). Now, setting $\Phi = \phi = 0$ and $\sigma_x = \sigma_y = \sigma$, (B.1) becomes

$$\begin{aligned} M &= \frac{A}{2\pi\sigma^2}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2-y^2}{2\sigma^2}\right)\cos kx\cdot\cos\left(Kx\cos\Theta + Ky\sin\Theta\right)dxdy \\ &= \frac{A}{2\pi\sigma^2}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos kx\left[\int_{-\infty}^{\infty}\exp\left(\frac{-y^2}{2\sigma^2}\right)\cos(Kx\cos\Theta + Ky\sin\Theta)\,dy\right]dx \end{aligned} \tag{C.13}$$

If we define $Q = Kx\cos\Theta$, $P = K\sin\Theta$ and hence

$$I = \int_{-\infty}^{\infty}\exp\left(\frac{-y^2}{2\sigma^2}\right)\cos(Q+Py)\,dy$$

Using (B.1) we have

$$\begin{aligned} I &= \int_{-\infty}^{\infty}\exp\left(\frac{-y^2}{2\sigma^2}\right)(\cos Q\cos Py - \sin Q\sin Py)\,dy \\ &= \cos Q\int_{-\infty}^{\infty}\exp\left(\frac{-y^2}{2\sigma^2}\right)\cos Py\,dy - \sin Q\int_{-\infty}^{\infty}\exp\left(\frac{-y^2}{2\sigma^2}\right)\sin Py\,dy \end{aligned} \tag{C.14}$$

The function $\exp\left(-y^2/2\sigma^2\right)$ is even and $\sin Py$ is odd, meaning the product $\exp\left(-y^2/2\sigma^2\right)\sin Py$ is odd. Therefore the integral of their product $\int_{-\infty}^{\infty}\exp(-y^2/2\sigma^2)\sin Py\,dy = 0$. (C.14) now simplifies to

$$I = \cos Q\int_{-\infty}^{\infty}\exp\left(\frac{-y^2}{2\sigma^2}\right)\cos Py\,dy \tag{C.15}$$

Now, note that the integral in (C.15) takes the form of (B.3), with $a = 1/2\sigma^2$ and $b = P$ and so

$$I = \cos(Kx\cos\Theta)\cdot\sigma\sqrt{2\pi}\cdot\exp\left(\frac{-K^2\sin^2\Theta\,\sigma^2}{2}\right)$$

where we have substituted back $Q = Kx\cos\Theta$. Now, M from (C.13) can be written

$$M = \frac{A}{2\pi\sigma^2}\int_{-\infty}^{\infty}\exp\left(-\frac{x^2}{2\sigma^2}\right)\cos kx\cdot\cos(Kx\cos\Theta)\cdot\sigma\sqrt{2\pi}\cdot\exp\left(\frac{-K^2\sin^2\Theta\,\sigma^2}{2}\right)dx$$

We now define $R = K\cos\Theta$, such that

$$M = \frac{A}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos kx\cos Rx\cdot\exp\left(\frac{-K^2\sin^2\Theta\sigma^2}{2}\right)dx$$

The identity (B.2) is used to give

$$M = \frac{A}{2\sigma\sqrt{2\pi}} \cdot \exp\left(\frac{-K^2\sin^2\Theta\sigma^2}{2}\right)\left[\int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos(k-R)x\,dx + \int_{-\infty}^{\infty}\exp\left(\frac{-x^2}{2\sigma^2}\right)\cos(k+R)x\,dx\right]$$

Again using (B.3) the above becomes

$$\begin{aligned} M &= \frac{A}{2\sigma\sqrt{2\pi}} \cdot \exp\left(\frac{-K^2\sin^2\Theta\sigma^2}{2}\right)\left[\sigma\sqrt{2\pi}\exp\left(\frac{-(k-R)^2\sigma^2}{2}\right) + \sigma\sqrt{2\pi}\exp\left(\frac{-(k+R)^2\sigma^2}{2}\right)\right] \\ &= \frac{A}{2}\exp\left(\frac{-K^2\sin^2\Theta\sigma^2}{2}\right)\left[\exp\left(\frac{-k^2\sigma^2+2kR\sigma^2-R^2\sigma^2}{2}\right) + \exp\left(\frac{-k^2\sigma^2-2kR\sigma^2-R^2\sigma^2}{2}\right)\right] \\ &= \frac{A}{2}\exp\left(\frac{-K^2\sin^2\Theta\sigma^2}{2}\right)\cdot\exp\left(\frac{-k^2\sigma^2-R^2\sigma^2}{2}\right)\cdot\left[\exp\left(\frac{2kR\sigma^2}{2}\right) + \exp\left(\frac{-2kR\sigma^2}{2}\right)\right] \end{aligned} \tag{C.16}$$

Substituting back for R, the first two exponentials of (C.16) are better written as

$$\exp\left(\frac{-K^2\sin^2\Theta\sigma^2}{2}\right)\cdot\exp\left(\frac{-K^2\cos^2\Theta\sigma^2}{2}\right)\cdot\exp\left(\frac{-k^2\sigma^2}{2}\right)$$

Because $\sin^2\Theta + \cos^2\Theta = 1$, the above expression simplifies to

$$\exp\left(\frac{-K^2\sigma^2}{2}\right)\cdot\exp\left(\frac{-k^2\sigma^2}{2}\right) = \exp\left(\frac{-\sigma^2(K^2+k^2)}{2}\right) \tag{C.17}$$

Now, replacing the first two exponentials with (C.17) returns

$$M = \frac{A}{2}\cdot\exp\left(\frac{-\sigma^2(K^2+k^2)}{2}\right)\cdot\left(\exp\left(kR\sigma^2\right) + \exp\left(-kR\sigma^2\right)\right) \tag{C.18}$$

Noting that $\cosh x = (1/2)(\exp(x) + \exp(-x))$, (C.18) is written

$$M = A\cdot\exp\left(\frac{-\sigma^2(K^2+k^2)}{2}\right)\cdot\cosh\left(kK\cos\Theta\sigma^2\right)$$

which is identical to (2.14).

Drifting Grating

From (C.11) we have, for the response of a simple cell to drifting grating,

$$M = A\cos(\omega t)\cos\phi \left[\exp\left(\frac{-(k-K)^2\sigma^2}{2}\right)\right]$$

Setting $\phi = 0$, we have

$$M = A\cos(\omega t) \left[\exp\left(\frac{-(k-K)^2\sigma^2}{2}\right)\right] \tag{C.19}$$

D Stability Analysis & Conditions on Interneuron Activity

Here we consider the stability requirements for the various dynamical systems discussed in the main text. For simplicity, we take the recurrent weight matrix $\tilde{W}_{ij}$ of (2.20) to be a block matrix, that is, $\tilde{W}_{ij} = 1/(N-1)$, $\forall i, j$. This returns the convenient result that the principal eigenvector is simply $e_i^1 = 1$, $\forall i$ and therefore

$$a_1 = e_i^1 r_i = r_1 + \ldots + r_N \tag{D.1}$$

and

$$h_1 = e_i^1 I_i^{ff} = I_1^{ff} + \ldots + I_N^{ff} \tag{D.2}$$

However, the analysis to follow can be extended for other weight matrices. We wish to find the stability requirements of the coupled system given by the first eigendirection of (2.21) with divisive inhibition and the interneuron dynamics:

$$\begin{aligned} \tau_r \frac{da_1}{dt} &= -a_1 + h_1 + \frac{\lambda^1 a_1}{R+Q} \\ \tau_R \frac{dR}{dt} &= -R + v(a_1, h_1) \end{aligned} \tag{D.3}$$

where we have combined the parameter g into $\tilde{W}$ such that $W = g\tilde{W}$. Taking the Jacobian of (D.3) we have

$$J = \begin{pmatrix} -1 + \frac{\lambda^1}{R^*+Q} & -\frac{\lambda^1 a_1}{(R^*+Q)^2} \\ \frac{\partial v(a_1,h_1)}{\partial a_1} & -1 \end{pmatrix} \tag{D.4}$$

For a stable system we require that its trace be negative and its determinant be greater than zero, that is,

- **Trace**: $\lambda^1/(R^* + Q) - 2 < 0$
- **Determinant**: $(R^* + Q)^2 - \lambda^1(R^* + Q) + \lambda^1 a_1^* \cdot (\partial v(a_1, h_1)/\partial a_1) > 0$

where the superscript * indicates steady-state solutions.

D.1 The Feedback Model

For interneuron activity described in (4.4), $v(a_1, h_1) = Ga_1$. We then have

- Trace: $\lambda^1 < 2(R^* + Q)$
- Determinant: $Ga_1^*\lambda^1 > (R^* + Q)^2$

Substituting $R^* = Ga_1$ we have the condition

$$2(Ga_1^* + Q) - Ga_1^* + \frac{Q^2}{Ga_1^*} < \lambda^1 < 2(Ga_1^* + Q) \tag{D.5}$$

which is easily met with a fine balance between the network size (λ^1) and base interneuron activity (Q).

D.2 The Disinhibition Model

For the 'Disinhibition Model' where the interneuron activity is described by (4.7), $v(a_1, h_1) = Ga_1/(h_1 + Q)$. We then have

- Trace: $\lambda^1 < 2(R^* + Q)$
- Determinant: $Ga_1^*\lambda^1 > (R^* + Q)^2(h_1 + Q)$

Substituting $R^* = Ga_1/(h_1 + Q)$ we have

$$2\left(\frac{Ga_1^*}{h_1 + Q} + Q\right) - \frac{Ga_1^*}{h_1 + Q} + \frac{Q^2(h_1 + Q)}{Ga_1^*} < \lambda^1 < 2\left(\frac{Ga_1^*}{h_1 + Q} + Q\right) \tag{D.6}$$

which again dictates the fine balance between the network size (λ^1), the base interneuron activity (Q) and the feedforward input h_1.

D.3 The Feedforward Model

For the 'Feedforward Model', (4.10) describes the interneuron activity where $v(a_1, h_1)$ is simply h_1. We then have

- Trace: $\lambda^1 < 2(R^* + Q)$
- Determinant: $\lambda^1 < R^* + Q$

Substituting $R^* = h_1$ we have

$$\lambda^1 < h_1 + Q \tag{D.7}$$

which restricts the size of the network (λ^1) in relation to the feedforward input (h_1).

D.4 The Feedback-Feedforward Model

In the case of the 'Feedback-Feedforward Model', (4.13) shows that $v(a_1, h_1)$ takes the form of the product $a_1 h_1$. We then have

- Trace: $\lambda^1 < 2(R^* + Q)$
- Determinant: $\lambda^1 a_1^* h_1 > (R^* + Q)^2$

Substituting $R^* = a_1^* h_1$ we have

$$2(a_1^* h_1 + Q) - a_1^* h_1 + \frac{Q^2}{a_1^* h_1} < \lambda^1 < 2(a_1^* h_1 + Q) \tag{D.8}$$

which can be achieved simply by controlling the base interneuron activity (Q) and the size of the network (λ^1).

E Short-Term Synaptic Plasticity

In both synaptic depression and facilitation, we use a dynamic variable $0 \leq D \leq 1$ that obeys the following kinetics:

- Between spikes the dynamic variable follows a simple asymptotic return to unity:

$$\tau_D \frac{dD}{dt} = 1 - D \tag{E.1}$$

The solution to which is given by

$$D = 1 - \exp(-t/\tau_D) \tag{E.2}$$

This recovery to unity is shown in Figure 7.1.

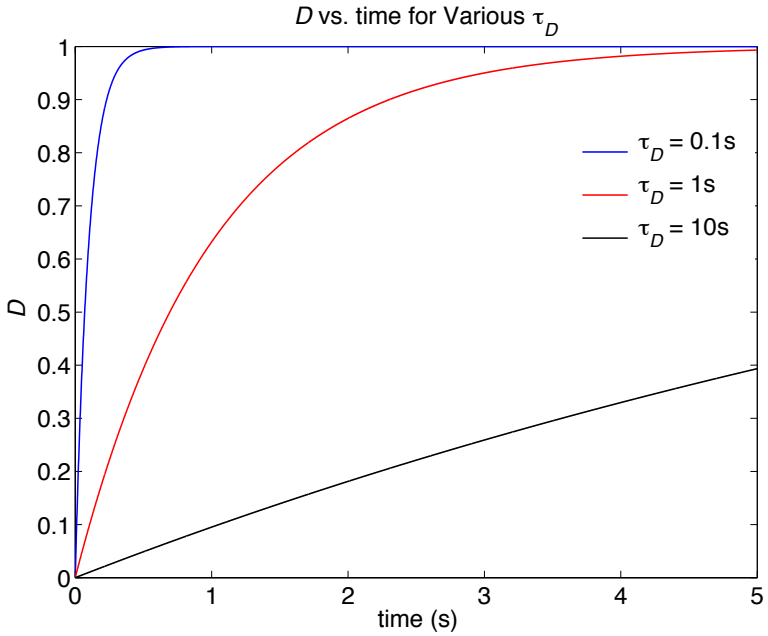


Figure 7.1: Synaptic depression variable D in time for various time constants τ_D = 0.1s (blue), 1s (red) and 10s (black) representing the recovery from synaptic depression after each presynaptic action potential as described in (E.2).

- At the instance where there is a spike, the dynamic variable is multiplied by factor $(d-1)$ such that D obeys

$$\frac{dD}{dt} = D(d-1) \tag{E.3}$$

where synaptic depression and facilitation is brought about by $0 < d < 1$ and $d > 1$ respectively.

Combining (E.1) and (E.3) as a linear system we have the overall short-term synaptic plasticity dynamics as

$$\tau_D \frac{dD}{dt} = 1 - D(1 + \tau_D(1-d)) \tag{E.4}$$

E.1 Synaptic Depression

In [4] synaptic depression is incorporated into the model by defining dynamic variables D^{ff} and D^{rec} for the feedforward and recurrent inputs that follow

$$\tau_D \frac{dD_i^{ff}}{dt} = 1 - D_i^{ff}(1 + \tau_D I_i^{ff}(1-d)) \tag{E.5}$$

and

$$\tau_D \frac{dD_{ij}^{rec}}{dt} = 1 - D_{ij}^{rec}(1 + \tau_D r_j(1-d)) \tag{E.6}$$

respectively. The feedforward input is now $D_i^{ff} I_i^{ff}$ and the recurrent input is given by

$$I_i^{rec} = \frac{g_i}{N-1} \sum_{j \neq i}^{N} D_{ij}^{rec} \left(2\exp\left(-\frac{(k_i - k_j)^2}{2\sigma_c^2} \right) - \exp\left(-\frac{(k_i - k_j)^2}{2\sigma_s^2} \right) \right) r_j \tag{E.7}$$

which is to be compared with (2.19) [4].

E.2 Synaptic Facilitation

We model synaptic facilitation in a very similar manner to synaptic depression. Though the authors of [30] model facilitation with additive rather than multiplicative updates, we use the multiplicative model based on [33], where the ratio of synaptic currents after paired stimulation is described by a single exponential fit that asymptotes to unity. In the network model, facilitation takes place at the synapse from the complex cells onto somatostatin expressing interneurons that receive input from the recurrent circuit only. This gives the resulting system

$$\begin{aligned} \tau_r \frac{dr_i}{dt} &= -r_i + I_i^{ff} + \frac{1}{(N-1)} \frac{\sum_{j \neq i} g_i r_j}{(R+Q)} \\ \tau_R \frac{dR}{dt} &= -R + 0.1 \sum_{i=1}^{N} F_i e_i^1 r_i \end{aligned} \tag{E.8}$$

where we have included a multiplicative factor of 0.1.

F MATLAB Simulations

F.1 Without Divisive Inhibition

To simulate (2.16), the authors of [4] first calculate I_i^{ff} and I_i^{rec} at some given time and thereafter treat their sum $C = I_i^{ff} + I_i^{rec}$ as a constant to solve the differential equation

$$\tau_r \frac{dr_i}{dt} = -r_i + C \tag{F.1}$$

With $\tau_r = 1$ and substituting $P = C - r_i$ we have

$$\frac{dr_i}{dt} = \frac{dr_i}{dP} \cdot \frac{dP}{dt} = -1 \cdot \frac{dP}{dt} = P \tag{F.2}$$

which yields the simple solution of

$$P = \exp(-t) \tag{F.3}$$

Substituting back for P and rearranging,

$$r_i = C - \exp(-t) \tag{F.4}$$

Now discretising the time evolution of r_i in (F.4)

$$\begin{aligned} r_i(t + \Delta t) &= C - \exp(-t - \Delta t) \\ &= C - \exp(-t)\exp(-\Delta t) \\ &= C\exp(-\Delta t) - C\exp(-\Delta t) - \exp(-t)\exp(-\Delta t) + C \qquad \text{(F.5)} \\ &= \exp(-\Delta t)(C - \exp(-t) - C) + C \qquad \text{(F.6)} \\ &= \exp(-\Delta t)(r_i(t) - C) + C \qquad \text{(F.7)} \end{aligned}$$

where we have simply added and subtracted $C\exp(-\Delta t)$ in (F.5), collected terms in (F.6) and used (F.4) in (F.7).

F.2 With Divisive Inhibition

Interneuron dynamics is simulated in a similar manner as described previously for (2.16). The firing rate of the interneuron given by

$$\tau_R \frac{dR}{dt} = -R + v(a_1, h_1) \tag{F.8}$$

yields the solution

$$R = v(a_1, h_1) - \tau_R \exp(-\tau_R t) \tag{F.9}$$

Discretising the above and rearranging, the time evolution of R is described by

$$R(t + \Delta t) = \exp(-\Delta t \cdot \tau_R)(R(t) - v(a_1, h_1)) + v(a_1, h_1) \tag{F.10}$$

The term $v(a_1, h_1)$ is time-dependent and hence calculated for each t to be used in finding $R(t + \Delta t)$.

F.3 With Short-Term Synaptic Plasticity

Depression

We have already described synaptic depression in (E.5) and (E.6) for the feedforward and recurrent inputs to the circuit. To implement (E.5), first define $E = 1 + \tau_D I_i^{ff}(1 - d)$, such that

$$\tau_D \frac{D_i^{ff}}{dt} = 1 - E D_i^{ff} \tag{F.11}$$

Now, defining $P = 1 - E D_i^{ff}$, we have

$$\frac{dD_i^{ff}}{dt} = \frac{dD_i^{ff}}{dP} \cdot \frac{dP}{dt} = -\frac{1}{E}\frac{dP}{dt} \tag{F.12}$$

such that (F.11) is now

$$\frac{dP}{dt} = -\frac{E}{\tau_D} \cdot P \tag{F.13}$$

Rearranging,

$$\int \frac{1}{P} dP = -\frac{E}{\tau_D} \int dt \tag{F.14}$$

and hence

$$\ln P = -\frac{E}{\tau_D} \cdot t \tag{F.15}$$

Substituting back $P = 1 - ED_i^{ff}$ and rearranging,

$$D_i^{ff} = \frac{1 - \exp\left(-Et/\tau_D\right)}{E} \tag{F.16}$$

(F.16) is now in the form of (F.4) and (F.9) and hence following the procedure of section F.3 we have for the evolution of the dynamic variable D_i^{ff},

$$D_i^{ff}(t + \Delta t) = \exp\left(\frac{-E\Delta t}{\tau_D}\right)\left(D_i^{ff}(t) - \frac{1}{E}\right) + \frac{1}{E} \tag{F.17}$$

where recall $E = 1 + \tau_D I_i^{ff}(1 - d)$. Implementing (E.6) is a similar process and returns

$$D_{ij}^{rec}(t + \Delta t) = \exp\left(-\frac{H\Delta t}{\tau_D}\right)\left(D_{ij}^{rec}(t) - \frac{1}{H}\right) + \frac{1}{H} \tag{F.18}$$

here $H = 1 + \tau_D r_j(1 - d)$.

Facilitation

Using a similar approach to depression, simulating the dynamics of the interneuron of system (E.8) follows

$$R(t + \Delta t) = \exp\left(-\Delta t \cdot \tau_R\right)\left(R(t) - \sum_{i=1}^{N} F_i e_i^1 r_i\right) + \sum_{i=1}^{N} F_i e_i^1 r_i \tag{F.19}$$

where, in a similar manner to $D_i^{ff}(t)$ and $D_{ij}^{rec}(t)$, F is described by

$$F_i(t + \Delta t) = \exp\left(\frac{-J\Delta t}{\tau_F}\right)\left(F_i(t) + \frac{1}{J}\right) + \frac{1}{J} \tag{F.20}$$

F.4 MATLAB Code for Dynamics Simulation

The code used to simulate the various divisive inhibition systems above is given below.

```
1 clear; close all; clc;
```

```
%% Adaptation of code from Larry Abbott.
%% initialisation
N = 256;
da = 180/(N-1);
a = linspace(-90,90,N)'; % neurons' preferred spatial phase

%% parameters
c = 1.5;
b = 1.0;
wFac = 1;
phi = 0;
dt = 0.3;
expFac = exp(-dt);
w_dg = 4;       % Hz.
w_cg = 8*pi;    % divide by 2*pi for Hz.

%% simulation (without divisive inhibition)
% preallocation
for i = 1
gain = [1 2 4 8];
gmax = 1;   % all neurons have identical spatial frequency preference.

r_cgwA = struct([]);    r_dgwA = struct([]);    % vanilla model with A
r_cgwoA = struct([]);   r_dgwoA = struct([]);   % vanilla model without A

% with divisive inhibition preallocation
rd_cg1 = struct([]); rd_cg2 = struct([]); rd_cg3 = struct([]);
rd_cg4 = struct([]); rd_cg5 = struct([]); rd_cg6 = struct([]);

rd_dg1 = struct([]); rd_dg2 = struct([]); rd_dg3 = struct([]);
rd_dg4 = struct([]); rd_dg5 = struct([]); rd_dg6 = struct([]);

R_dg1 = struct([]); R_dg2 = struct([]); R_dg3 = struct([]);
R_dg4 = struct([]); R_dg5 = struct([]); R_dg6 = struct([]);

R_cg1 = struct([]); R_cg2 = struct([]); R_cg3 = struct([]);
R_cg4 = struct([]); R_cg5 = struct([]); R_cg6 = struct([]);

```

```
Q = 1;
tau_R = 5;
expFac2 = exp(-dt*tau_R);
tmax = 1e3;

G = 0.1;
S = 1e-6;
U = 1e-3;

D_ff_cg = struct([]);   D_rec_cg = struct([]);  % Depression variables
D_ff_dg = struct([]);   D_rec_dg = struct([]);
F_dg = struct([]);  F_cg = struct([]);          % Facilitation variables
tau_D = 300;    % 300ms
tau_F = 80;    % 100ms
d = 0.95;
end

g = zeros(1,length(gain));

for i = 1:length(gain)
    % preallocation
    g(i) = (gain(i)-1)*gmax./gain(i);

    r_cgwA{i} = zeros(N,tmax);  r_dgwA{i} = zeros(N,tmax);
    r_cgwoA{i} = zeros(N,tmax); r_dgwoA{i} = zeros(N,tmax);

    rd_cg1{i} = zeros(N,tmax); rd_cg2{i} = zeros(N,tmax);
    rd_cg3{i} = zeros(N,tmax); rd_cg4{i} = zeros(N,tmax);
    rd_cg5{i} = zeros(N,tmax); rd_cg6{i} = zeros(N,tmax);

    rd_dg1{i} = zeros(N,tmax); rd_dg2{i} = zeros(N,tmax);
    rd_dg3{i} = zeros(N,tmax); rd_dg4{i} = zeros(N,tmax);
    rd_dg5{i} = zeros(N,tmax); rd_dg6{i} = zeros(N,tmax);

    R_cg1{i} = zeros(1,tmax); R_cg2{i} = zeros(1,tmax);
    R_cg3{i} = zeros(1,tmax); R_cg4{i} = zeros(1,tmax);
    R_cg5{i} = zeros(1,tmax); R_cg6{i} = zeros(1,tmax);

```

```
78      R_dg1{i} = zeros(1,tmax); R_dg2{i} = zeros(1,tmax);
79      R_dg3{i} = zeros(1,tmax); R_dg4{i} = zeros(1,tmax);
80      R_dg5{i} = zeros(1,tmax); R_dg6{i} = zeros(1,tmax);
81
82      D_ff_cg{i} = zeros(N,tmax); D_rec_cg{i} = zeros(N,tmax);
83      D_ff_dg{i} = zeros(N,tmax); D_rec_dg{i} = zeros(N,tmax);
84      F_cg{i} = zeros(N,tmax);    F_dg{i} = zeros(N,tmax);
85
86      for t = 1:tmax-1
87          %% Without Divisive Inhibition (VANILLA,A+)
88              %% counterphase grating
89          % recurrent input
90          hRec_cgwA = (g(i)/N)*sum(r_cgwA{i}(:,t));
91
92          % feedforward input
93          hIn_cg = wFac*c*cos(2*pi*(a-phi)/180)*cos(w_cg*t/1000);
94          hIn_cg(hIn_cg<0) = 0;
95
96          % total input
97          C_cgwA = b*(1-g(i))*hIn_cg + hRec_cgwA;    % with A
98          C_cgwA(C_cgwA<0) = 0;
99          r_cgwA{i}(:,t+1) = C_cgwA + (r_cgwA{i}(:,t)-C_cgwA)*expFac;
100
101             %% drifting grating
102         % recurrent input
103         hRec_dgwA = (g(i)/N)*sum(r_dgwA{i}(:,t));
104
105         % feedforward input
106         hIn_dg = wFac*c*cos(2*pi*(a-w_dg*t/1000));
107         hIn_dg(hIn_dg<0) = 0;
108
109         % total input
110         C_dgwA = b*(1-g(i))*hIn_dg + hRec_dgwA;    % with A
111         C_dgwA(C_dgwA<0) = 0;
112         r_dgwA{i}(:,t+1) = C_dgwA + (r_dgwA{i}(:,t)-C_dgwA)*expFac;
113
114         %% Without Divisive Inhibition (VANILLA,A-)
115             %% counterphase grating
```

```
116         % recurrent input
117         hRec_cgwoA = (g(i)/N)*sum(r_cgwoA{i}(:,t));
118
119         % total input
120         C_cgwoA = b*hIn_cg + hRec_cgwoA;    % without A
121         C_cgwoA(C_cgwoA<0) = 0;
122         r_cgwoA{i}(:,t+1) = C_cgwoA + (r_cgwoA{i}(:,t)-C_cgwoA)*expFac;
123
124             %% drifting grating
125         % recurrent input
126         hRec_dgwoA = (g(i)/N)*sum(r_dgwoA{i}(:,t));
127
128         % total input
129         C_dgwoA = b*hIn_dg + hRec_dgwoA;    % without A
130         C_dgwoA(C_dgwoA<0) = 0;
131         r_dgwoA{i}(:,t+1) = C_dgwoA + (r_dgwoA{i}(:,t)-C_dgwoA)*expFac;
132
133         %% Without Divisive Inhibition but with depression (SYSTEM 1)
134             %% Dynamic variables for counterphase grating
135         E_cg = 1 + tau_D*hIn_cg*(1-d);
136         D_ff_cg{i}(:,t+1) = exp(-E_cg*dt/tau_D).*(D_ff_cg{i}(:,t)- ...
137             1./E_cg)+ 1./E_cg;
138
139         H_cg = 1 + tau_D*r_cgwA{i}(:,t)*(1-d);
140         D_rec_cg{i}(:,t+1) = exp(-H_cg*dt/tau_D).*(D_rec_cg{i}(:,t)- ...
141             1./H_cg)+ 1./H_cg;
142
143         % recurrent input
144         hRecd_cg = (g(i)/N)*sum(D_rec_cg{i}(:,t).*rd_cg1{i}(:,t));
145
146         % feedforward input
147         hInd_cg = D_ff_cg{i}(:,t).*hIn_cg;
148
149         % total input
150         Cd_cg = b*hInd_cg + hRecd_cg;
151         Cd_cg(Cd_cg<0) = 0;
152         rd_cg1{i}(:,t+1) = Cd_cg + (rd_cg1{i}(:,t)-Cd_cg)*expFac;
153
```

```
154             %% Dynamic variables for drifting grating
155         E_dg = 1 + tau_D*hIn_dg*(1-d);
156         D_ff_dg{i}(:,t+1) = exp(-E_dg*dt/tau_D).*(D_ff_dg{i}(:,t)- ...
157             1./E_dg) + 1./E_dg;
158 
159         H_dg = 1 + tau_D*r_dgwA{i}(:,t)*(1-d);
160         D_rec_dg{i}(:,t+1) = exp(-H_dg*dt/tau_D).*(D_rec_dg{i}(:,t)- ...
161             1./H_dg) + 1./H_dg;
162 
163         % recurrent input
164         hRecd_dg = (g(i)/N)*sum(D_rec_dg{i}(:,t).*rd_dg1{i}(:,t));
165 
166         % feedforward input
167         hInd_dg = D_ff_dg{i}(:,t).*hIn_dg;
168 
169         % total input
170         Cd_dg = b*hInd_dg + hRecd_dg;
171         Cd_dg(Cd_dg<0) = 0;
172         rd_dg1{i}(:,t+1) = Cd_dg + (rd_dg1{i}(:,t)-Cd_dg)*expFac;
173 
174         %% With Divisive Inhibition %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
175         %% With Divisive Inhibition and facilitation (SYSTEM 2)
176             %% Dynamic variables for counterphase grating
177         J_cg = 1 + tau_F*rd_cg2{i}(:,t);
178         F_cg{i}(:,t+1) = (1./J_cg).*exp(-J_cg*dt/tau_F).* ...
179             (J_cg.*F_cg{i}(:,t)+1) + 1./J_cg;
180 
181         % recurrent input
182         hRec_cg2_2 = (g(i)/N)*sum(rd_cg2{i}(:,t).*F_cg{i}(:,t));
183 
184         % total input
185         C_cg2_2 = b*hIn_cg + hRec_cg2_2/(R_cg2{i}(t)+Q);
186         C_cg2_2(C_cg2_2<0) = 0;
187 
188         % firing rate dynamics
189         rd_cg2{i}(:,t+1) = C_cg2_2 + (rd_cg2{i}(:,t)-C_cg2_2)*expFac;
190 
191         tmp_cg = 0.1*sum(F_cg{i}(:,t).*rd_cg2{i}(:,t));
```

```
192         R_cg2{i}(t+1) = tmp_cg + ...
193             (R_cg2{i}(t)-tmp_cg)*expFac2;
194             %% Dynamic variables for drifting grating
195         J_dg = 1 + tau_F*rd_dg2{i}(:,t);
196         F_dg{i}(:,t+1) = (1./J_dg).*exp(-J_dg*dt/tau_F).* ...
197             (J_dg.*F_dg{i}(:,t)+1) + 1./J_dg;
198
199         % recurrent input
200         hRec_dg2_2 = (g(i)/N)*sum(rd_dg2{i}(:,t).*F_dg{i}(:,t));
201
202         % total input
203         C_dg2_2 = b*hIn_dg + hRec_dg2_2/(U*R_dg2{i}(t)+Q);
204         C_dg2_2(C_dg2_2<0) = 0;
205
206         % firing rate dynamics
207         rd_dg2{i}(:,t+1) = C_dg2_2 + (rd_dg2{i}(:,t)-C_dg2_2)*expFac;
208
209         tmp_dg = 0.1*sum(F_dg{i}(:,t).*rd_dg2{i}(:,t));
210         R_dg2{i}(t+1) = tmp_dg + ...
211             (R_dg2{i}(t)-tmp_dg)*expFac2;
212
213             %% counterphase grating
214             %% SYSTEM 3 (FEEDBACK)
215         % recurrent input
216         hRec_cg2_3 = (g(i)/N)*sum(rd_cg3{i}(:,t));
217
218         % total input
219         C_cg2_3 = b*hIn_cg + hRec_cg2_3/(U*R_cg3{i}(t)+Q);
220         C_cg2_3(C_cg2_3<0) = 0;
221
222         % firing rate dynamics
223         rd_cg3{i}(:,t+1) = C_cg2_3 + (rd_cg3{i}(:,t)-C_cg2_3)*expFac;
224
225         tmp_cg = G*sum(rd_cg3{i}(:,t));
226         R_cg3{i}(t+1) =  tmp_cg + ...
227             (R_cg3{i}(t)-tmp_cg)*expFac2;
228
229             %% SYSTEM 4 (DISINHIBITION)
```

```
230         % recurrent input
231         hRec_cg2_4 = (g(i)/N)*sum(rd_cg4{i}(:,t));
232         % only the recurrent input changes with different systems
233
234         % total input
235         C_cg2_4 = b*hIn_cg + hRec_cg2_4/(R_cg4{i}(t)+Q);
236         C_cg2_4(C_cg2_4<0) = 0;
237
238         % firing rate dynamics
239         rd_cg4{i}(:,t+1) = C_cg2_4 + (rd_cg2{i}(:,t)-C_cg2_4)*expFac;
240
241         tmp_cg = G*sum(rd_cg4{i}(:,t))/(sum(hIn_cg)+Q);
242         R_cg4{i}(t+1) =  tmp_cg + ...
243             (R_cg4{i}(t)-tmp_cg)*expFac2;
244
245             %% SYSTEM 5 (FEEDFORWARD)
246         % recurrent input
247         hRec_cg2_5 = (g(i)/N)*sum(rd_cg5{i}(:,t));
248
249         % total input
250         C_cg2_5 = b*hIn_cg + hRec_cg2_5/(U*R_cg5{i}(t)+Q);
251         C_cg2_5(C_cg2_5<0) = 0;
252
253         % firing rate dynamics
254         rd_cg5{i}(:,t+1) = C_cg2_5 + (rd_cg5{i}(:,t)-C_cg2_5)*expFac;
255
256         tmp_cg = G*sum(hIn_cg);    % SYSTEM 5
257         R_cg5{i}(t+1) = tmp_cg + ...
258             (R_cg5{i}(t)-tmp_cg)*expFac2;
259
260             %% SYSTEM 6 (FEEDBACK-FEEDFORWARD)
261         % recurrent input
262         hRec_cg2_6 = (g(i)/N)*sum(rd_cg6{i}(:,t));
263
264         % total input
265         C_cg2_6 = b*hIn_cg + hRec_cg2_6/(R_cg6{i}(t)+Q);
266         C_cg2_6(C_cg2_6<0) = 0;
267
```

```
268         % firing rate dynamics
269         rd_cg6{i}(:,t+1) = C_cg2_6 + (rd_cg6{i}(:,t)-C_cg2_6)*expFac;
270
271         tmp_cg = S*sum(hIn_cg)*sum(rd_cg6{i}(:,t));
272         R_cg6{i}(t+1) =  tmp_cg + ...
273             (R_cg6{i}(t)-tmp_cg)*expFac2;
274
275             %% drifting grating
276             %% SYSTEM 3 (FEEDBACK)
277         % recurrent input
278         hRec_dg2_3 = (g(i)/N)*sum(rd_dg3{i}(:,t));
279
280         % total input
281         C_dg2_3 = b*hIn_dg + hRec_dg2_3/(U*R_dg3{i}(t)+Q);
282         C_dg2_3(C_dg2_3<0) = 0;
283
284         % firing rate dynamics
285         rd_dg3{i}(:,t+1) = C_dg2_3 + (rd_dg3{i}(:,t)-C_dg2_3)*expFac;
286
287         tmp_dg = G*sum(rd_dg3{i}(:,t));
288         R_dg3{i}(t+1) = tmp_dg + ...
289             (R_dg3{i}(t)-tmp_dg)*expFac2;
290
291             %% SYSTEM 4 (DISINHIBITION)
292         hRec_dg2_4 = (g(i)/N)*sum(rd_dg4{i}(:,t));
293
294         % total input
295         C_dg2_4 = b*hIn_dg + hRec_dg2_4/(R_dg4{i}(t)+Q);
296         C_dg2_4(C_dg2_4<0) = 0;
297
298         % firing rate dynamics
299         rd_dg4{i}(:,t+1) = C_dg2_4 + (rd_dg4{i}(:,t)-C_dg2_4)*expFac;
300
301         tmp_dg = G*sum(rd_dg4{i}(:,t))/(sum(hIn_dg)+Q);
302         R_dg4{i}(t+1) = tmp_dg + ...
303             (R_dg4{i}(t)-tmp_dg)*expFac2;
304
305             %% SYSTEM 5 (FEEDFORWARD)
```

```
306         % recurrent
307         hRec_dg2_5 = (g(i)/N)*sum(rd_dg5{i}(:,t));
308
309         % total input
310         C_dg2_5 = b*hIn_dg + hRec_dg2_5/(U*R_dg5{i}(t)+Q);
311         C_dg2_5(C_dg2_5<0) = 0;
312
313         % firing rate dynamics
314         rd_dg5{i}(:,t+1) = C_dg2_5 + (rd_dg5{i}(:,t)-C_dg2_5)*expFac;
315
316         tmp_dg = G*sum(hIn_dg);
317         R_dg5{i}(t+1) = tmp_dg + ...
318             (R_dg5{i}(t)-tmp_dg)*expFac2;
319
320             %% SYSTEM 6 (FEEDBACK-FEEDFORWARD)
321         % recurrent
322         hRec_dg2_6 = (g(i)/N)*sum(rd_dg6{i}(:,t));
323
324         % total input
325         C_dg2_6 = b*hIn_dg + hRec_dg2_6/(R_dg6{i}(t)+Q);
326         C_dg2_6(C_dg2_6<0) = 0;
327
328         % firing rate dynamics
329         rd_dg6{i}(:,t+1) = C_dg2_6 + (rd_dg6{i}(:,t)-C_dg2_6)*expFac;
330
331         tmp_dg = S*sum(hIn_dg)*sum(rd_dg6{i}(:,t));
332         R_dg6{i}(t+1) = tmp_dg + ...
333             (R_dg6{i}(t)-tmp_dg)*expFac2;
334
335     end
336
337     for j = 1:N
338        r_cgwA{i}(j,:) = r_cgwA{i}(j,:)/max(r_cgwA{i}(j,:)); % normalise
339        r_dgwA{i}(j,:) = r_dgwA{i}(j,:)/max(r_dgwA{i}(j,:));
340
341        r_cgwoA{i}(j,:) = r_cgwoA{i}(j,:)/max(r_cgwoA{i}(j,:));
342        r_dgwoA{i}(j,:) = r_dgwoA{i}(j,:)/max(r_dgwoA{i}(j,:));
343
```

```
344         rd_cg1{i}(j,:) = rd_cg1{i}(j,:)/max(rd_cg1{i}(j,:));
345         rd_dg1{i}(j,:) = rd_dg1{i}(j,:)/max(rd_dg1{i}(j,:));
346
347         rd_cg2{i}(j,:) = rd_cg2{i}(j,:)/max(rd_cg2{i}(j,:));
348         rd_dg2{i}(j,:) = rd_dg2{i}(j,:)/max(rd_dg2{i}(j,:));
349
350         rd_cg3{i}(j,:) = rd_cg3{i}(j,:)/max(rd_cg3{i}(j,:));
351         rd_dg3{i}(j,:) = rd_dg3{i}(j,:)/max(rd_dg3{i}(j,:));
352
353         rd_cg4{i}(j,:) = rd_cg4{i}(j,:)/max(rd_cg4{i}(j,:));
354         rd_dg4{i}(j,:) = rd_dg4{i}(j,:)/max(rd_dg4{i}(j,:));
355
356         rd_cg5{i}(j,:) = rd_cg5{i}(j,:)/max(rd_cg5{i}(j,:));
357         rd_dg5{i}(j,:) = rd_dg5{i}(j,:)/max(rd_dg5{i}(j,:));
358
359         rd_cg6{i}(j,:) = rd_cg6{i}(j,:)/max(rd_cg6{i}(j,:));
360         rd_dg6{i}(j,:) = rd_dg6{i}(j,:)/max(rd_dg6{i}(j,:));
361     end
362 end
```

F.5 Extracting F_0, F_1 and F_2 Components of the Network Response

We use the discrete Fourier transform (DFT) to extract the F_0, F_1 and F_2 components of the network firing rate response. Using the DFT on a response $r[t]$, $t = 0, \ldots, N-1$, (hereafter square brackets are used to denote the elements of a discrete signal), returns

$$F_n = \sum_{t=0}^{N-1} r[t] \exp\left(-2i\pi nk/N\right) \tag{F.21}$$

Computationally, we first separately find the normalised real and imaginary components of (F.21) and find the magnitude of the resulting complex number. The MATLAB code used to find the F_0, F_1 and F_2 components are given below.

```
1 function [ramp rtheta] = get_F_components(data,order,omega)
2 % GET_F_COMPONENTS retrieves the 'order'th component of a cell's rate
3 % response (data). OMEGA is the frequency of the stimulus. The code is a
```

```
4  % result of modifications made to GET_FOURIER_HARMONIC_ARRAY.c originally
5  % by Wyeth Bair.
6
7  n = length(data);
8  T = 1/omega;     % omega is in Hz.
9
10 todd = 0;   teven = 0;
11
12 for i = 1:n % units of ms.
13     x = ((i-1)/1000)/(T/order)*2*pi;
14     % 1000ms = 1s; [s]/[s] (this will go into the sin)
15     % there is a reason for keeping the (T/order). If order is 2, then we
16     % halve the time period as the frequency is doubled.
17     todd = todd + 2*sin(x)*data(i);
18     teven = teven + 2*cos(x)*data(i);
19     % above we multiply the components by 2 since the sum of todd and teven
20     % are only half of the rectified signal.
21 end
22
23 teven = teven/n;
24 todd = todd/n;
25 % in the above, we normalise for the steps below
26 ramp = sqrt(teven^2 + todd^2);
27 rtheta = atan2(todd,teven)*180/pi;
```

Bibliography

[1] D. Kätzel , B. V. Zemelman, C. Bütfering, M. Wölfel, and G. Miesenböck. The columnar and laminar organization of inhibitory connections to neocortical excitatory cells. *Nature Neuroscience*, 14:100 – 107, 2011.

[2] R. Ben-Yishai, R. L. Bar-Or, and H. Sompolinsky. Theory of orientation tuning in visual cortex. *Proceedings of the National Academy of Sciences USA*, 92:3844 – 3848, 1995.

[3] F. S. Chance and L. F. Abbott. Divisive inhibition in recurrent networks. *Network*, 11:119 – 129, 2000.

[4] F. S. Chance, S. B. Nelson, and L. F. Abbott. Complex cells as cortically amplified simple cells. *Nature Neuroscience*, 2:277 – 282, 1999.

[5] B. Y. Chow, X. Han, A. S. Dobry, X. Qian, A. S. Chuong, M. Li, M. A. Henninger, G. M. Belfort, Y. Lin, P. E. Monahan, and E. S. Boyden. High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature*, 463:98 – 102, 2010.

[6] P. Dayan and L. F. Abbott. *Theoretical Neuroscience - Computatinal and Mathematical Modeling of Neural Systems*. MIT Press, 2001.

[7] A. F. Dean and D. J. Tolhurst. On the distinctness of simple and complex cells in the visual cortex of the cat. *Journal of Physiology*, 344:305 – 325, 1983.

[8] N. T. Dhruv, C. Tailby, S. H. Sokol, N. J. Majaj, and P. Lennie. Nonlinear signal summation in magnocellular neurons of the macaque lateral geniculate nucleus. *Journal of Neurophysiology*, 102:1921 – 1929, 2009.

[9] R. J. Douglas, C. Koch, M. Mahowald, K. A. Martin, and H. H. Suarez. Recurrent excitation in neocortical circuits. *Science*, 269:981 – 985, 1995.

[10] R. J. Douglas and K. A. C. Martin. Recurrent neuronal circuits in the neocortex. *Current Biology*, 17:R496 – R500, 2007.

[11] M. Fuhrmann, A. Stahlberg, E. Govorunova, S. Rank, and P. Hegemann. The Abundant Retinal Protein of the *Chlamydomonas* eye is not the Photoreceptor for Phototaxis and Photophobic Responses. *Journal of Cell Science*, 114:3857 – 3863, 2001.

[12] X. Han and E. S. Boyden. Multiple-color optical activation, silencing, and desynchronization of neural activity, with single-spike temporal resolution. *PLoSOne*, 2(3):e299, 2007.

[13] M. J. Hawken and A. J. Parker. Spatial properties of neurons in the monkey striate cortex. *Proceedings of the Royal Society of London Series-B Biological Sciences*, 231:251 – 288, 1987.

[14] P. Hegemann, S. Ehlenbeck, and D. Gradmann. Multiple photocycles of channelrhodopsin. *Biophysical Journal*, 89:3911 – 3918, 2005.

[15] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160:106 – 154, 1962.

[16] C. Johansson and A. Lansner. Imposing biological constraints onto an abstract neocortical attractor network model. *Neural Computation*, 19:1871 – 1896, 2007.

[17] E. R. Kandel, J. H. Schwartz, and T. M. Jessell. *Principles of Neural Science*. McGraw-Hill, fourth edition, 2000.

[18] C. Kapfer, L. L. Glickfeld, B. V. Atallah, and M. Scanziani. Supralinear increase of recurrent inhibition during sparse activity in the somatosensory cortex. *Nature Neuroscience*, 10:743 – 753, 2007.

[19] Y-J. Kim, M. Grabowecky, K. A. Paller, and S. Suzuki. Differential roles of frequency-following and frequency-doubling visual responses revealed by evoked neural harmonics. *Journal of Cognitive Neuroscience*, 23(8):1875 – 1886, 2011.

[20] J. Lücke. Receptive field self-organization in a model of the fine structure in v1 cortical columns. *Neural Computation*, 21:2805 – 2845, 2009.

[21] F. Mechler and D. L. Ringach. On the classification of simple and complex cells. *Vision Research*, 42:1017 – 1033, 2002.

[22] J. A. Movshon, I. D. Thompson, and D. J. Tolhurst. Receptive field organisation of complex cells in the cat's striate cortex. *Journal of Physiology*, 283:79 – 99, 1978a.

[23] G. Nagel, M. Braunner, J. F. Liewald, N. Adeishvili, E. Bamberg, and A. Gottschalk. Light Activation of Channelrhodopsin-2 in Excitable Cells of *Caenorhabditis elegans* Triggers Rapid Behavioral Responses. *Current Biology*, 15:2279 – 2284, 2005.

[24] C. M. Niell and M. P. Stryker. Highly selective receptive fields in mouse visual cortex. *Journal of Neuroscience*, 28:7520 – 7536, 2008.

[25] L. Petreanu, D. Huber, A. Sobczyk, and K. Svoboda. Channelrhodopsin-2-assisted circuit mapping of long-range callosal projections. *Nature Neuroscience*, 10:663 – 668, 2007.

[26] N. J. Priebe and D. Ferster. Inhibition, spike threshold, and stimulus selectivity in primary visual cortex. *Neuron*, 57:482 – 497, 2008.

[27] B. C. Skottun, R. L. Valois, D. H. Grosof, J. A. Movshon, D. G. Albrecht, and A. B. Bonds. Classifying simple and complex cells on the basis of response classifying simple and complex cells on the basis of response modulation. *Vision Research*, 31:1079 – 1086, 1991.

[28] D. C. Somers, S. B. Nelson, and M. Sur. An emergent model of orientation selectivity in cat visual cortical simple cells. *Journal of Neuroscience*, 15:5448 – 5465, 1995.

[29] L. Tao, M. Shelley, D. McLaughlin, and R. Shapley. An egalitarian network model for the emergence of simple and complex cells in visual cortex. *Proceedings of the National Academy of Sciences*, 101:366 – 371, 2004.

[30] J. A. Varela, K. Sen, J. Gibson, J. Frost, L. F. Abbott, and S. B. Nelson. A quantitative description of short-term plasticity at excitatory synapses in layer 2/3 of rat primary visual cortex. *The Journal of Neuroscience*, 17(20):7925 – 7940, 1997.

[31] X. Xu, J. M. Ichida, J. D. Allison, J. D. Boyd, A. B. Bonds, and V. A. Casagrande. A comparison of koniocellular, magnocellular and parvocellular receptive field properties in the lateral geniculate nucleus of the owl monkey. *Journal of Physiology*, 531.1:203 – 218, 2001.

[32] F. Zhang, L.-P. Wang, E. S. Boyden, and K. Deisseroth. Channelrhodopsin-2 and optical control of excitable cells. *Nature Methods*, 3(10):785 – 792, 2006.

[33] R. S. Zucker and W. G. Regehr. Short-term synaptic plasticity. *Annual Review of Physiology*, 64:355 – 405, 2002.