THE UNIVERSITY
of EDINBURGH

DEVELOPMENT AND ENCODING OF VISUAL STATISTICS IN THE PRIMARY VISUAL CORTEX

PHILIPP JOHN FREDERIC RUDIGER

Doctor of Philosophy
School of Informatics
University of Edinburgh

2016

Philipp John Frederic Rudiger:
Development and encoding of visual statistics
in the primary visual cortex
Doctor of Philosophy, 2016

SUPERVISORS:
Dr. James A. Bednar
Dr. Alexander Thiele

LAY SUMMARY

The field of neuroscience has made tremendous progress over the last century. Neuroscientists have collected an overwhelming amount of data describing the brain at large and small levels, ranging from interactions between individual genes, proteins and cells to the communication between different brain regions. Yet we still understand very little about how the brain actually implements the major functions of a nervous system, such as learning about an animal's environment to solve even basic sensory and cognitive tasks. Fully explaining such processing requires linking across all these levels of description, explaining how the high-level behavior that can be observed externally is constructed from the operation of the lower-level elements that can be observed inside the brain.

Making such links is very difficult to do using laboratory experiments, which for practical reasons almost always focus on one or two levels at any one time. Computational models can be an important complementary tool, allowing researchers to express and then test hypotheses about how these levels relate to each other. Yet the vast majority of existing models focus on a very narrow range of phenomena, mirroring specific experiments rather than helping integrate experimental results into a coherent explanation. Unfortunately, any specific experiment could have a huge range of possible explanations, and only by considering all available evidence can we narrow in on the correct explanations for how the brain functions.

In this thesis we present a computational model of visual cortex that helps bridge across several disparate levels, using lower-level elements constrained by measurements of different cell types and their connections, grouped into neural regions that together form a sensory subsystem, and then tested using sensory stimuli used in behavioral experiments. The resulting model is a good match to low-level observations of neural connectivity, providing a specific and testable explanation for how that connectivity arises, and is also closely matched to the observed neural responses to a wide range of visual patterns used in experiments. The model demonstrates for the first time how specific connections between neurons in the cortex can store statistical information about the visual environment and then use that information to optimize perception of novel stimuli.

This work is clearly important for basic neuroscience, which seeks to understand how the brain functions. But it is also important for clinical neuroscience, because many neurological conditions have been associated with various defects in the devel-

opment and wiring of the specific cell types studied here. Future integrative models that can link neural machinery to behavior will help drive progress in both of these areas of research.

ABSTRACT

How do circuits in the mammalian cerebral cortex encode properties of the sensory environment in a way that can drive adaptive behavior? This question is fundamental to neuroscience, but it has been very difficult to approach directly. Various computational and theoretical models can explain a wide range of phenomena observed in the primary visual cortex (V1), including the anatomical organization of its circuits, the development of functional properties like orientation tuning, and behavioral effects like surround modulation. However, so far no model has been able to bridge these levels of description to explain how the machinery that develops directly affects behavior. Bridging these levels is important, because phenomena at any one specific level can have many possible explanations, but there are far fewer possibilities to consider once all of the available evidence is taken into account.

In this thesis we integrate the information gleaned about cortical development, circuit and cell-type specific interactions, and anatomical, behavioral and electrophysiological measurements, to develop a computational model of V1 that is constrained enough to make predictions across multiple levels of description. Through a series of models incorporating increasing levels of biophysical detail and becoming increasingly better constrained, we are able to make detailed predictions for the types of mechanistic interactions required for robust development of cortical maps that have a realistic anatomical organization, and thereby gain insight into the computations performed by the primary visual cortex.

The initial models focus on how existing anatomical and electrophysiological knowledge can be integrated into previously abstract models to give a well-grounded and highly constrained account of the emergence of pattern-specific tuning in the primary visual cortex. More detailed models then address the interactions between specific excitatory and inhibitory cell classes in V1, and what role each cell type may play during development and function. Finally, we demonstrate how these cell classes come together to form a circuit that gives rise not only to robust development but also the development of realistic lateral connectivity patterns. Crucially, these patterns reflect the statistics of the visual environment to which the model was exposed during development. This property allows us to explore how the model is able to capture higher-order information about the environment and use that information to optimize neural coding and aid the processing of complex visual tasks.

Using this model we can make a number of very specific predictions about the mechanistic workings of the brain. Specifically, the model predicts a crucial role of parvalbumin-expressing interneurons in robust development and divisive normalization, while it implicates somatostatin immunoreactive neurons in mediating longer range and feature-selective suppression. The model also makes predictions about the role of these cell classes in efficient neural coding and under what conditions the model fails to organize. In particular, we show that a tight coupling of activity between the principal excitatory population and the parvalbumin population is central to robust and stable responses and organization, which may have implications for a variety of diseases where parvalbumin interneuron function is impaired, such as schizophrenia and autism. Further the model explains the switch from facilitatory to suppressive surround modulation effects as a simple by-product of the facilitating response function of long-range excitatory connections targeting a specialized class of inhibitory interneurons. Finally, the model allows us to make predictions about the statistics that are encoded in the extensive network of long-range intra-areal connectivity in V1, suggesting that even V1 can capture high-level statistical dependencies in the visual environment.

The final model represents a comprehensive and well constrained model of the primary visual cortex, which for the first time can relate the physiological properties of individual cell classes to their role in development, learning and function. While the model is specifically tuned for V1, all mechanisms introduced are completely general, and can be used as a general cortical model, useful for studying phenomena across the visual cortex and even the cortex as a whole. This work is also highly relevant for clinical neuroscience, as the cell types studied here have been implicated in neurological disorders as wide ranging as autism, schizophrenia and Parkinson's disease.

ACKNOWLEDGEMENTS

First of all I would like to thank my supervisor Jim Bednar for his thoughtful advice and support throughout the course of this project and for hugely influencing my thinking about the field of neuroscience. My thanks also extend to my second supervisor Alexander Thiele who provided patient advice although the topic of the PhD quickly shifted away from his main interests. I am also deeply grateful to the other members of the Institute of Adaptive and Neural Computation for their invaluable support and help over the past few years. In particular I would like to thank Jean-Luc Stevens for discussions and feedback that have helped me refine the ideas that make up the core of this thesis.

I would also to thank my family who have been supportive throughout my studies and allowed me to pursue them in the first place. My father who has nurtured my intellectual curiosity since I was a child and my godmother who has been more supportive than anyone could hope for. Finally a very special thanks to my partner and friend Josephine for being patient while I worked day and night on this project.

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.

Edinburgh, 2016

Philipp John Frederic
Rudiger, December 8, 2016

CONTENTS

ABBREVIATIONS

LHI Local Homogeneity Index

OCSI Orientation Contrast Suppression Index

PV-ir Parvalbumin immunoreactive

Sst-ir Somatostatin immunoreactive

SOM Self-organizing map

GCAL Gain-control lateral adaptation lateral model

SCAL Spatially Calibrated GCAL model

SEPI Short-range excitation, parvalbumin inhibition model

LESPI Long-range excitation, somatostatin/parvalbumin inhibition model

DoG Difference-of-Gaussians

RoG Ratio-of-Gaussians

MSE Mean-squared-error

LGN Lateral geniculate nucleus

GABA γ-aminobutyric acid

GC Gain-Control

FB Feedback

mRF minimum receptive field

hsRF high-contrast summation receptive field

lsRF low-contrast summation receptive field

1

INTRODUCTION

One of the major experimental advances in visual neuroscience over the past decade has been the ability to manipulate and record from individual cells or cell classes by combining advanced genetic and imaging techniques. These techniques have given us a new window into the mammalian brain and have allowed us to begin hypothesizing what role different cell classes may play in development and behavior. However, so far very few attempts have been made to place these new findings into a greater context, particularly to investigate what makes different cell classes suited toward particular functions and how these neurons develop into a functional circuit that can efficiently extract and encode the statistics of the visual input.

The guiding hypothesis of this thesis is that the incredible complexity of the mammalian cortex is constructed through a process of self-organization giving rise to a developed circuit, which can robustly capture the statistics of the input and adaptively use this information to optimally encode the input for further processing and perform low-level inferences even when only a small amount of information is available. Over the course of this thesis I will outline how a stereotyped circuit of different neural cell classes develops into a specialized circuit that closely matches experimental observations of development and surround modulation. This model will for the first time bridge the early development of the circuits with a mechanistic account of how the learned statistics are used and finally play a role in specific cortical functions, which give rise to contextual modulation effects and higher level behaviors.

In particular I will adapt an existing developmental model of cortical development, which gives a good account of activity-dependent development in the primary visual cortex, demonstrating robust, yet adaptive organization of receptive fields and lateral connectivity. Once I have established how to relate the spatial scales in the model to experimental data, this thesis sets out to answer a number of questions:

- What type of inhibition is suited toward driving robust self-organization of V1 circuits?
- What are the roles of parvalbumin- and somatostatin-immunoreactive interneurons in the development and function of the primary visual cortex?
- What properties make a particular cell class suited for a particular function?

- How do the cell classes self-organize into a circuit that is robust and stable to the widely varying visual input that an animal's environment provides?
- What specific computational purpose, if any, do particular cell classes serve?
- What are the tradeoffs between long-range recurrent excitation and inhibition for stimulus detection and discriminability?
- Does the connectivity in the visual cortex store the visual statistics of the environment, and how can it begin using this information to improve neural coding and low-level inference?

Through an iterative process of analyzing the model and incorporating increasing anatomical and physiological detail, I make several concrete predictions about the role in both developmental and behavioral processes. Firstly I introduce the first developmental model which agrees with anatomical connectivity data, electrophysiological measurements and the known spatial scales of orientation maps in V1. Having established this I show that the fast, quickly spiking response properties of parvalbumin-expressing neurons can provide the fast local and largely isotropic inhibition of surrounding neurons that is crucial for robust orientation map development.

Through a further extension of the model I show that the facilitating and slow response profile of somatostatin-immunoreactive neurons give rise to narrower orientation tuning, which provides orientation-dependent surround suppression. In doing so the thesis will explore the tension between strong long-range recurrent excitation, useful for detection under uncertain or weak stimulus conditions, and suppression, useful to reduce the redundancy in the neural representation and enhancing discrimination under strong stimulus conditions. Finally, I show how the long-range lateral connectivity is able to capture the statistics of the input, acting to normalize higher-order statistics in the input, thereby optimizing the neural code, and relate these results to the known patchy connectivity in the primate visual cortex.

Not only does this thesis outline the first spatially calibrated, developmental model of the primary visual cortex, it also makes specific predictions about the role of different cell classes, which have been implicated in a variety of pathologies, most notably in autism, schizophrenia and Parkinson's. A better understanding of the developmental disruptions caused by malfunctioning parvalbumin neurons in particular will provide a better understanding of how to recover regular function and provides a solid starting point for future studies looking specifically at these disruptions. Furthermore, these results have clear implications for theoretical neuroscience and machine learning fields as they clearly predict that not only is the lateral connectivity in V1 shaped

by the visual statistics of the environment but also suggests that even low-level areas like V1 can make use of these statistics to optimize neural coding.

1.1 ORGANIZATION OF THE THESIS

After a thorough review of the existing literature, I present a workflow for easily and reproducibly exploring high-dimensional and temporally evolving models, allowing us to gain insights into non-linear interactions between different components and parameters in a model in ways that were not possible with previous tools.

The following chapter then introduces the initial developmental model of the visual cortex, summarizing what is known about the spatial scales of different types of connections in primary visual cortex and establishing a correspondence between the model and experimental data. In the process, we present various protocols to characterize the spatial tuning in the response of the model and various analyses applied to the connectivity in the model to confirm that model and experiment are indeed in close agreement.

In the second results chapter I extend this analysis to models of cortical development and function that takes into account the anatomical and physiological properties of different cell classes and begin elucidating their role in experimentally observed phenomena ranging from self-organization to particular computations performed by the fully developed cortex, including surround modulation, divisive gain control and sparse coding.

In the following chapter we introduce a second population of inhibitory interneurons demonstrating how the interactions between these neurons and long-range excitatory connections through a di-synaptic circuit can account for both feature specific facilitation and suppression.

In the final results chapter I further show that the long-range connections successfully capture higher-order correlations in natural images, explaining mechanistically how and why feature specific modulation arises in the cortex.

The thesis overall presents a unifying account of how different cell classes in the visual cortex self-organize into a robust circuit that can encode the natural image statistics of its rearing environment and make use of that information to improve how the cortex extracts, encodes and uses information from its environment.

2

LITERATURE REVIEW

The field of neuroscience has experienced an explosion in the size, quantity and quality of data being collected over the past few decades, but over the same time period very few novel theoretical advances have been made. Most of our theories for how the brain operates have been left unchanged by the radical improvements in techniques such as optogenetics and two-photon calcium imaging. A major obstacle to converting these breakthroughs in experimental neuroscience to genuine advances in our theoretical understanding of the brain is the difficulty in relating the many individual observations into a coherent picture of how these circuits work.

This problem is particularly evident in visual neuroscience, one of the most extensively researched areas in neuroscience, due to the relative ease with which the visual areas can be accessed, stimulated and studied. Since the first models of the function of the visual system were developed by the pioneers of visual neuroscience, namely Mountcastle and Hubel and Wiesel we have collected tremendous amounts of data and developed models to explain particular phenomena. However, we have made very little progress on explaining the fundamental computations performed in visual cortical areas; most of the core ideas such as receptive and association fields have been around since the early pioneers proposed them.

In this thesis we will argue that this failure to make significant advances in understanding the computations performed by the visual system is primarily a failure to attempt to integrate our knowledge across various levels of description. This issue is precisely what David Marr highlighted when stating that the brain, and the visual system in particular, should be analyzed at three distinct but complementary levels of analysis (Marr, 1982), later expanded to include a fourth level by Tomaso Poggio (Poggio, 2012). According to Marr and Poggio, models of cognitive systems such as the mammalian visual system should describe the system at one or more of the following levels of analysis:

- Learning: How does the system learn to perform the necessary computations?
- Computational: What does the system actually do? What computations does it perform?
- Algorithmic/Representational: How does the system represent the problem and solve it?

- Implementational/Physical: How is the system physically implemented?

Each of these levels of evidence when viewed in isolation can contribute to the overall understanding about the system, but to gain real insights, models should begin abstracting across these levels of analysis to find out where and how evidence from various sources fit together (or more interestingly, highlight potential discrepancies). In this chapter we will summarize the current knowledge about the circuits, cell classes and functions performed by the primary visual cortex at each of these levels before starting to integrate that information into models that will allow us to make novel predictions about the physical implementation, algorithmic operation, and computational function as well as learning and development. We will begin by reviewing the anatomy and physiology of the early visual system.

2.1 EARLY VISUAL SYSTEM: AREAS AND LAYERS

The earliest stage of the mammalian visual system (pictured in Figure 2.1) is the retina, where rod and cone photoreceptors convert incident photons into electrical and chemical signals. These signals are then further converted from analogue voltages into spike trains by retinal bipolar and ganglion cells and sent down the optic nerve to the lateral geniculate nucleus (LGN). Connections from the two eyes cross over at the optic chiasm to form projections of the right and left visual field contralaterally. The connections from the retina map retinotopically onto the LGN, meaning that nearby areas of the LGN respond to nearby portions of the visual field. After the initial processing in the retina and LGN, the visual stream is projected onto the primary visual cortex (V1).

In the retina, summation of various photoreceptor types gives rise to so called center-surround receptive fields. These ON- and OFF-center receptive fields can arise in bipolar cells but are more commonly associated with retinal ganglion cells (RGCs). This receptive field type responds most strongly to spots of light/dark moving through the visual field (as shown in Figure 2.2) but can be characterized as simple edge detectors.

In addition to the macro-organization of the visual system into distinct areas, the LGN and downstream structures are further broken down into individual layers and cell classes. The separation of LGN cells into layers also corresponds to functional separations, as layers 1, 4 and 6 usually receive contra-lateral input, while layers 2, 3 and 5 receive ipsi-lateral inputs from the retina. Furthermore, different layers consist of different cell types, with ventral layers 1 and 2 containing larger so called magno-cellular (M) neurons and dorsal layers 3, 4, 5 and 6 containing smaller parvo-

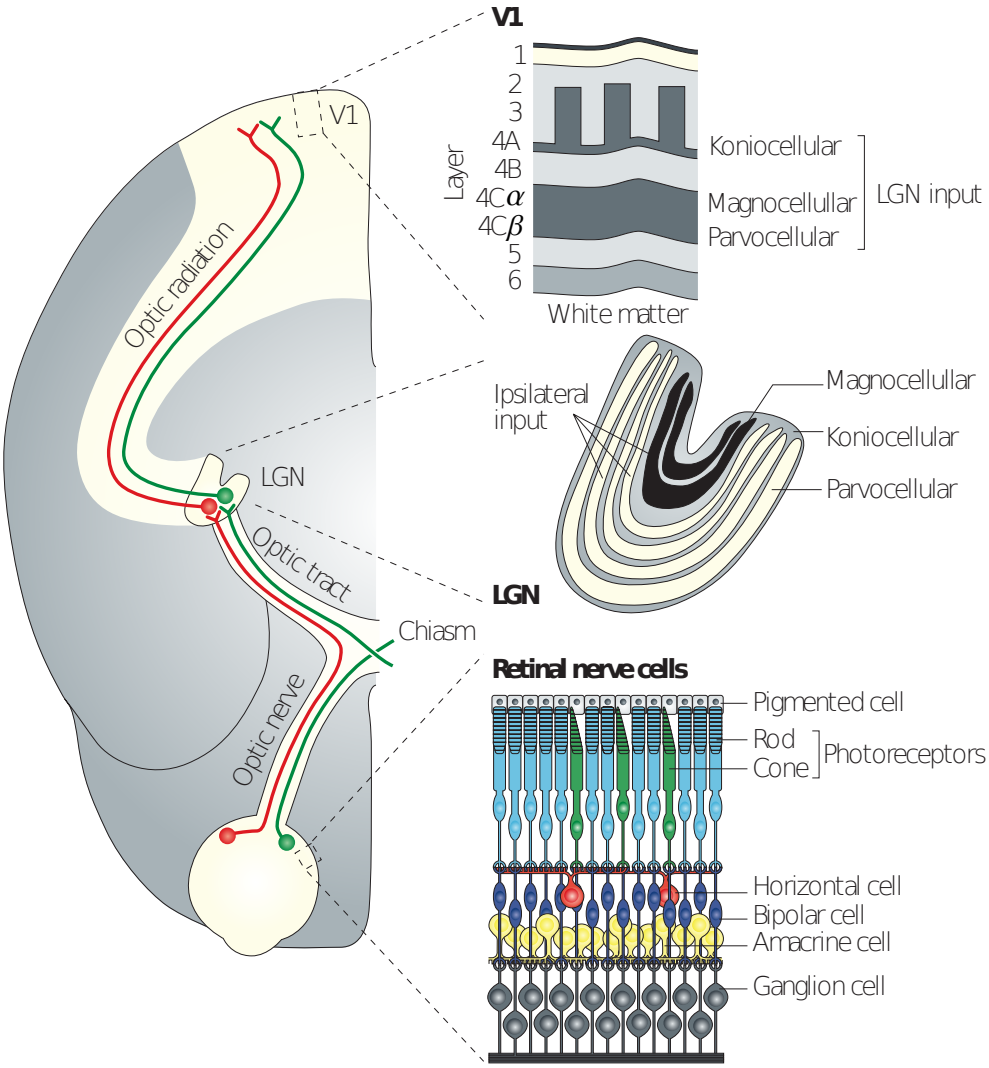


Figure 2.1: The early visual pathway in primates (also generally the same for most other mammals) from the retina to the primary visual cortex (V1) via the lateral geniculate nucleus (LGN) of the thalamus. The left panel shows the pathway, while the right panels highlight noteworthy sections including the structure of the retina, with the LGN and V1 broken down into their different layers and showing different cell types. Adapted from Solomon and Lennie (2007).

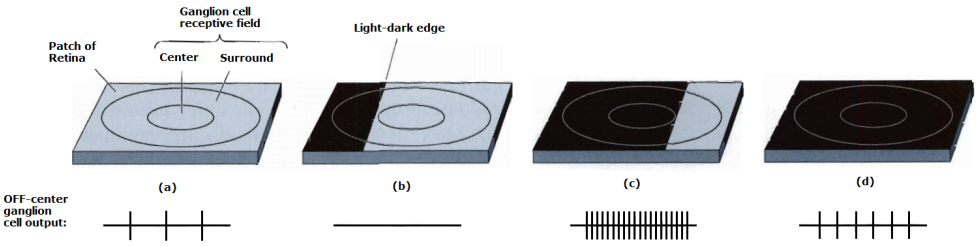


Figure 2.2: The canonical center-surround receptive field structure of retinal ganglion cells and LGN neurons, illustrating how a contrast edge activates different portions of the field and thereby results in different activation patterns. From left to right one can see that as the light-dark edge moves into the ON surround field, spontaneous activity is suppressed, and as it moves further over the OFF-center field the neuron becomes highly active. Adapted from Bear et al. (2006).

cellular (P) neurons with intra-laminar neurons being referred to as konio-cellular (K). Since these three cell types are also present in the retina and make connections mainly with their own cell type it is theorized that each carries its own parallel information stream. Functionally, P-cells have displayed greater sensitivity to chromatic contrast and higher spatial frequencies, linking them to the processing of detail and color, while M-cells have been shown to have greater sensitivity for high temporal frequencies associating them with motion processing.

2.1.1 *Primary Visual Cortex: Topographic Maps, Simple and Complex Cells*

The primary visual cortex (V1) or striate cortex provides the first cortical stage of processing of visual information. The cortex was classically divided into six layers but since then several subdivisions have been added after functional sub-groups were discovered. Feedforward input from the LGN is received in layer 4Cα and 4Cβ , which receive input most of their input from M- and P-cells respectively. The neurons in layer 4 then send projections up to layers 2/3, which has a diverse intra-laminar network of connections but also sends intracortical projections to a number of higher visual areas, while layers 5 and 6 provide feedback to the LGN.

Neurons in the primary visual cortex (V1) are tuned to respond to a variety of different features or complex combinations of such features, including orientation, spatial and temporal frequency, motion direction, color or ocular origin. In many mammalian species, including primates and carnivorans, these feature preferences map smoothly and topographically onto the cortical surface. This mapping extends vertically through the layers of the cortex, giving rise to the notion of distinct cortical columns. Retinotopy arises due to the mapping of visual information straight from the retina to the LGN and then to the cortex. Other response preferences such

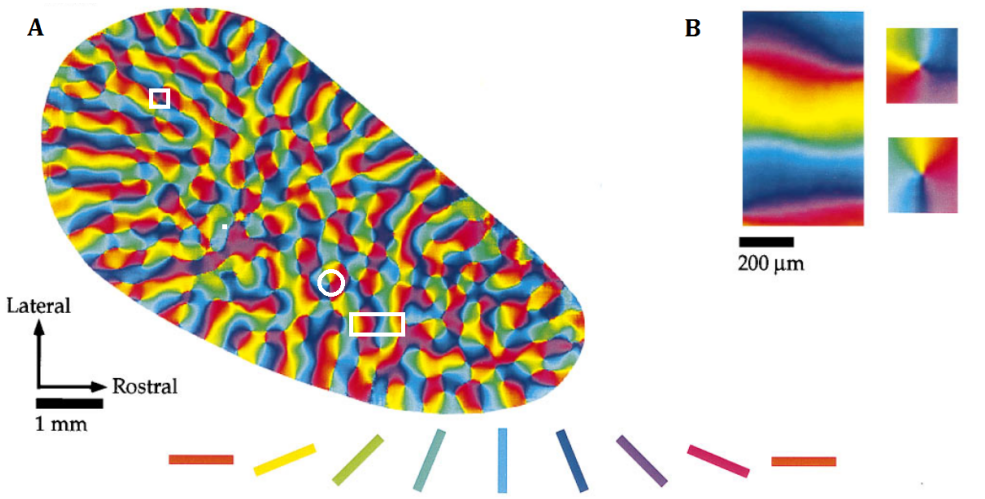


Figure 2.3: A) Orientation preference map in ferret V1 generated by overlaying the activity maps for different orientations and artificially coloring each area according to the orientation preference using the color key below. The image also highlights three common features of orientation maps in white. The square highlights a saddle point, where a patch of cortex selective for a particular direction is almost bisected by a patch selective to another direction. The circle highlights a pinwheel arrangement, where different orientations preference patches are arranged in a circular shape. Finally the rectangular shape highlights a linear zone in which orientation preference change continuously. B) Magnifications of a linear zone and two pinwheel arrangements. Adapted from Bosking et al. (1997).

as orientation and direction selectivity rarely arise in primate LGN and are usually thought of as emergent phenomena in the cortex.

The receptive fields of V1 neurons are different in that often there are no longer simple ON- or OFF-center surround fields, forming more complex spatio-temporal patterns. They are commonly modeled using Gabor filters as shown in Figure 2.4, which have elongated ON and OFF regions or lobes, generated by localizing a full-field sine grating by multiplying it with a Gaussian envelope. Orientation selectivity and spatial frequency preference are determined by the orientation and spacing of ON- and OFF-regions respectively. It is also possible for V1 cells to filter temporal patterns by employing spatio-temporal shifts in their ON and OFF lobes, giving rise to direction selectivity.

Orientation-selective neurons can generally be classed as simple or complex cells, depending on whether they display some form of spatial/phase invariance (Skottun et al., 1991). In reality this classification is not particularly clear, with cells being somewhere on a gradient from pure simple cells to a complex cell with the degree of phase invariance being the determining factor. Apart from phase invariance the neurons often exhibit contrast invariance, such that even at very low contrast they will respond

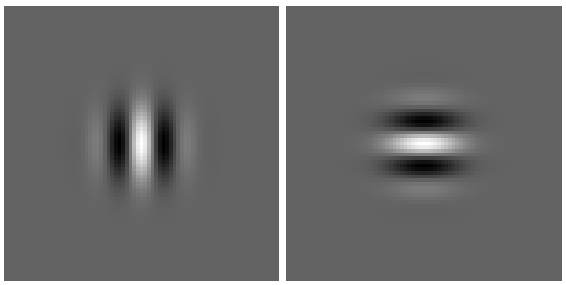


Figure 2.4: Gabor Patches at 0 degree and 90/180 degree orientations with clearly visible ON (white) and OFF (black) regions.

more strongly to their preferred orientation than to the orthogonal orientation (Sclar and Freeman, 1982).

In context of this project, topographic feature maps and RF interactions play a fundamental role as they provide the basis around which the neural circuit organizes. In fact one of the core assumptions of this thesis is that functional organization of V1 arises during the development of the animal and is mediated largely by activity dependent processes. While this assumption is founded on multiple lines of evidence it is still under heavy debate and we will describe the literature supporting this view in the next section.

2.1.2 *Development of Topographic Maps in V1*

The development and maturation of cortical topographic feature representations in the form of maps is closely linked to function and can reveal a lot about how the cortex is capable of capturing and encoding statistics of the natural world. Developmental studies involve imaging the same area of the cortex over a number of days and investigating what drives development of orientation maps and other topographic arrangements. Early developmental studies found, using relatively limited optical imaging techniques on ferrets shortly after eye opening, that the iso-orientation domains in the V1 develop very early in development and subsequently show very little change (Chapman et al., 1996) (pictured in Figure 2.5). These and other experiments (White and Fitzpatrick, 2007) showed that orientation preference develops even in absence of visual input, although the maps do not fully mature. Thus orientation maps and other topographic organizations develop initially even in absence of external visual input, through internally generated visual activity such as retinal waves (Cang et al., 2005), but then require external stimulation to fully mature.

In the initial stages of development various preprogrammed guidance cues set up the basic connectivity between the LGN and the cortical processing areas (Huberman et al., 2008). Some successful models of development have focused on the afferent

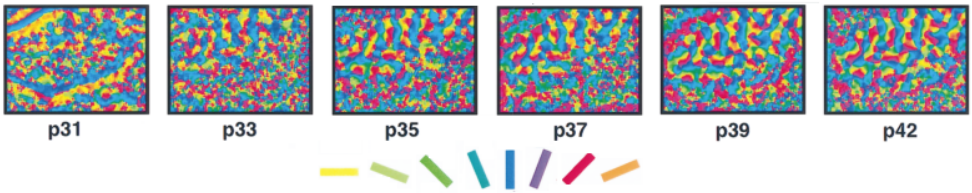


Figure 2.5: Development of orientation map in ferret primary visual cortex from post-natal day 31 to 42 revealed using chronic optical imaging of intrinsic signals. Adapted from Chapman et al. (1996).

connections between the LGN ON and OFF cells and their targets in the visual cortex as the driving force behind the development of orientation columns (Jin et al., 2011). This is likely to tell only part of the story as even during pre-natal development, retinal waves, consisting of periodic activity in retinal ganglion cells (RGCs), spread across the retina driving neighboring RGCs to fire in a correlated fashion, which allows the primary visual cortex to develop topographic feature maps (Firth et al., 2005). The most prominent proposals in theory and in models have been focused on the termination pattern of both the geniculocortical afferents in layer 4 (Katz et al., 2000; Ringach, 2007) and adjustments in feedforward and lateral connectivity through activity dependent, competitive processes (Bednar and Miikkulainen, 2003).

It is now generally accepted that the development of topographic maps is driven by activity dependent weight modification in some form of Hebbian learning (Shatz, 1996) and a variety of models have emerged to account for this idea (Bienenstock et al., 1982; Miller et al., 1989). The LISSOM and GCAL models (Bednar and Miikkulainen, 2003; Stevens et al., 2013b), which provide the basis of the modeling work proposed in this project, have shown that robust map development can be achieved with a small number of relatively simple mechanisms including homeostatic plasticity, lateral gain control in LGN and lateral connectivity in V1 (Stevens et al., 2013b) and will be considered in more detail at a later stage. These models can account for the development of orientation preference maps through both intrinsic activity such as retinal waves (Bednar and Miikkulainen, 2003) and visually induced retinal activity. Experiments show that at the very least activity dependent synaptic modification is required to achieve the finely tuned precision, which can now be observed at single cell resolution (Ohki et al., 2005; White and Fitzpatrick, 2007).

Lateral intra-areal connections in particular have been implicated in map development as their functional connectivity seems to be closely related to map structure (Gilbert and Wiesel, 1983). Experiments in layer 2/3 of the tree shrew involving orientation preference mapping and subsequent axonal staining have shown that although short range connections show no preference in their terminations, long range connections longer than 500 µm preferentially link neurons with co-oriented and co-axially

aligned receptive fields (Bosking et al., 1997). In iso-orientation regions, cells therefore make short-range connections largely with cells that prefer the same features as them, while at pinwheels short-range connections are made with cells with a wider range of orientation preferences (Yousef et al., 2001). However, it is known that the patchy lateral connectivity in V1 does not arise until after the orientation map has emerged (Ruthazer and Stryker, 1996), indicating that they may be involved in some higher-order processing not required during initial development.

The exact purpose of cortical maps is still under intense debate. Various hypotheses exist that they have emerged as a highly efficient solution to solve distinct constraints, firstly of providing optimal coverage of the feature space by mapping it onto a tiled two-dimensional surface and secondly to optimize the overall length of synaptic connections, which have high metabolic costs associated with them **?**. It is also possible that map formation is merely an epiphenomenon that arises due to the need for an evolutionary algorithm that can develop useful structure based merely on the interaction of locally interconnected neurons, eliminating the need for any complex interactions to be encoded genetically, making use of general pattern-formation principles to efficiently encode the structure of the cortex genetically (**?**). Cortical maps may also be important for the hierarchical organization of the cortex allowing high-level areas to pool over the full feature space within a relatively small region, e.g. allowing complex cells in layer 2/3 in V1 to pool over all different orientation preferences in layer 4. All of these explanations rely on the assumption that preferentially connecting neurons with similar feature preferences has some benefit. It is therefore unclear whether cortical maps have any concrete purpose or are merely the result of various constraints on cortical development.

The development of the early visual pathway is probably driven by a number of mechanisms complementing each other at various stages, starting with guidance cues setting up coarse predetermined connectivity patterns, which are then refined through Hebbian processes driven by internal and extrinsically stimulated activity. Although the structure of a neuron's receptive field is constantly changing and varies widely from cell to cell, a lot of work has gone into measuring the exact structure and functional properties of neural receptive fields and their underlying anatomical basis, the neurite arbors. The different time scales involved in the initial development of orientation maps and the maturation of long-range lateral connectivity indicates that the precise statistics of visual experience are not important in early development and only play a strong role once the rough classical receptive field structure has emerged.

2.1.3 *Functional and Anatomical Properties of Neural Receptive Fields*

The spatial properties of RFs in the LGN and V1 are of particular interest in the context of this thesis as they provide the basis of a realistic model of the para-foveal regions of the primary visual cortex in macaques, allowing direct comparisons between model in experiment in a system where the spatial scales are of great importance. Before attempting to model the recurrent, lateral inputs to V1, the spatial properties of each set of afferent connections entering V1 need to be thoroughly understood and incorporated into the model.

The extents and structure of neural receptive fields are defined by the axonal and dendritic arborization of afferent, horizontal and feedback connections, whether that is in the LGN, the V1 or higher up in the cortical architecture. While these extents can theoretically be measured physically by tracing the neurites of a number of cells, this has only systematically been possible for thalamocortical neurons, as tracing studies along the retinogeniculate pathway are generally infeasible. Therefore spatial properties of LGN RFs have been estimated through stimulus-driven protocols. The following sections will outline the methods employed in characterizing geniculate RFs and detail the relative contribution of afferent, lateral and feedback connections to para-foveal RFs in the LGN and V1 of macaques.

Receptive Fields in the Lateral Geniculate Nucleus

The spatio-temporal structure of receptive fields becomes increasingly more complex when moving up in the visual processing hierarchy. As described earlier, receptive fields in primate LGN are primarily made up of antagonistic center-surround regions, although Gabor-like lobes are also sometimes observed. Even at this early stage of processing, lateral and feedback connections can modulate neural responses and have been found to exert a suppressive effect (Hubel and Wiesel, 1961). More recent studies have concluded that this suppressive effect is mediated primarily by lateral connections and acts as a form of contrast-gain control allowing for the encoding of a high dynamic range of luminance values (Bonin et al., 2005).

As pointed out previously, the LGN receives a large proportion of its inputs from feedback cells originating in layer 6 of V1 (Sherman and Guillery, 2002). It is unclear how these feedback connections contribute to the RF properties of LGN neurons, but most evidence suggests they are mainly involved in higher level modulatory processes, especially in regard to the processing of motion (Sillito et al., 2006), and thus do not directly contribute to the RF structure.

Estimating the relative contribution and effect of the various LGN afferents on its neural receptive fields has been attempted using a variety of protocols. Unfortunately

little to no data is available from tracing studies, primarily because the LGN is embedded deep in the brain, which makes it incredibly difficult to trace individual neurites from and to their origin. Therefore a number of protocols have been developed by which the parameters of the center-surround fields could be estimated. After measuring the response of cat retinal ganglion cells (RGCs) to moving bar stimuli (Rodieck, 1965; Rodieck and Stone, 1965a), Rodieck and Stone (1965b) found that by fitting a Difference of Gaussian (DoG) model to the data it was possible to estimate the relative strength and size of the central and surround portions of the receptive field. It was not until later that systematic recordings of this kind were carried out in the LGN of macaques, at which point the moving bar stimuli were replaced with sine gratings of varying spatial frequency.

As not all papers can be reviewed here, the data from three different studies with protocols of increasing complexity will be considered. Derrington and Lennie (1984) was the first of such studies, attempting to characterize the spatial and temporal properties of parvocellular LGN neurons in *Macaca mulatta* by fitting DoG models to the responses. The analysis and confidence intervals of this first study were rather limited, so the first study whose results will be reported here is Spear et al. (1994), which also considered the effect of aging on receptive field properties. They found that the receptive field center radius only increased very weakly with eccentricity, that the smallest RF center radii were confined to parvocellular neurons, and that the RF surround was significantly smaller in parvocellular layers.

The next detailed study of LGN neural RFs was carried out by Levitt et al. (2001), investigating the effects of visual deprivation on their visual response properties. There was a tendency for parvocellular neurons to exhibit greater spatial resolution and the highest temporal resolution to be magnocellular. Finally, their analysis extended to earlier data from different species of macaque, which showed that there is some variation in the distribution of ON and OFF cells between *M. mulatta* and *M. fascularis*. Overall their results on spatial tuning match those found by previous studies quite closely, which is unsurprising as the measuring and fitting protocols were highly similar.

In an attempt to calculate the relative contributions of different neural connections, determine differences between the K-, M- and P-cellular pathways and measure contrast dependent tuning, a number of more recent studies have introduced more complex measurement and fitting protocols. In particular, these experiments for the first time attempted to separate out the influence of the non-classical or extra-classical surround (ECRF or nCRF), which is thought to be mediated by lateral and feedback connections. Therefore, a new measurement and fitting protocol, introduced by Sceniak et al. (1999) in form of the integrated DoG (iDoG) model to describe spa-

tial summation in the visual cortex, was used. Instead of measuring the response to varying spatial frequency, this protocol involves the presentation of drifting sine grating disks with varying apertures at the neuron's optimum spatial frequency. The rationale behind this new protocol was that the optimal spatial frequency would maximally drive the CRF excitatory center, while minimizing the influence of the CRF surround. Therefore the resulting area-summation tuning curve would represent only the response from the CRF excitatory center and the ECRF surround. The spatial frequency response measurement protocol was also modified by confining the drifting sine gratings to a circular aperture, reducing influences from beyond the CRF. While these assumptions do not necessarily hold for reasons that will be discussed later, they provide a first systematic attempt at separating the contributions from the CRF and ECRF.

Sceniak et al. (2006) were the first to study spatial RF properties of thalamocortical afferents in macaques by measuring both spatial frequency and area summation response functions and fitting the results with DoG and iDoG models (respectively) to estimate the spatial parameters of the probed neurons. These results represent the best estimates so far of the spatial properties of LGN receptive fields. The first thing to note is the clear discrepancy between the estimates of CRF excitatory center radius estimates in this paper compared to previous estimates. This difference may be explained by the more homogeneous distribution of cells, as the sample population was taken exclusively from layer 4. The older protocol also failed to confine the drifting sine grating to a disk, which may have resulted systematic underestimation of the excitatory component due to long-range suppressive effects. While excitatory extents vary hugely across the various studies, the suppressive surround estimates are fairly consistent. Furthermore, the spatial extent of the excitatory CRF centers was found to be contrast invariant, while both the ECRF and CRF suppressive surround extents were found to increase at lower contrast levels. In summary, looking back at all the studies considered here, excitatory CRF extents are generally distributed between 0.05-0.5° in radius, while inhibitory CRF and suppressive ECRF radii are distributed distributed anywhere between 0.6-1.5°, and the suppression index is quite high (SI>0.8) for 80% of cells.

While these results provide the best estimates that are currently available, the protocols used rely on a number of flawed assumptions. The DoG model fitted to the spatial frequency tuning curve relies on the assumption that no other components are contributing to the response. Although the limited size of the sine grating disk drive should *reduce* long-range influences on the response, and the ECRF surround is frequency invariant over a broader spectrum than the CRF, further contributing mechanisms cannot be excluded and may therefore affect the estimates. Similarly, the

iDoG model fitted to area summation tuning curves may be affected by a number of unaccounted mechanisms. The iDoG model (see equation 4.12) actually corresponds to an even-luminance disk of variable size, rather than the sine grating disks that were used by Sceniak et al. (2006). The decision to use technically incorrect stimuli was made to minimize the influence of the inhibitory CRF surround, which may itself still have some influence on the response. While all these limitations should be kept in mind, this study is still the best attempt at controlling unaccounted contributions, and thus provides the best data on the spatial properties of macaque LGN RFs until specific anatomical estimates are made available.

Receptive Fields in the Primary Visual Cortex

The receptive field properties of neurons in V1, in contrast to LGN neurons, have been characterized to a far greater extent, with a number of studies publishing direct anatomical data on neurite arborization in addition to studies involving stimulus protocols such as those employed to characterize LGN RFs. A number of reviews have been published in the past decade to classify different portions of the receptive field and link them to their physiological substrate in the form of feedforward (FF), lateral, and feedback (FB) connections. In order to attain a proper understanding of the spatial distribution of afferent neurites populations of V1 neurons are targeted by, this section will summarize the results.

Recent analyses have established a more complex model for the classification of the spatial properties of neural RFs in V1 than the simpler classical and extra-classical RF structure utilized in earlier work. The structure of a V1 receptive field has been visualized by Angelucci and Sainsbury (2006) and can be seen in Figure 2.6. It can be broken down into the minimum response field (mRF), the summation response field (sRF), which itself is broken down into the high-contrast and low-contrast summation RF (hsRF and lsRF), and the far surround. In particular, a distinction has been made between the near surround, which extends as far as the lsRF, and a suppressive far surround that extends beyond the lsRF. The distinction between the hsRF and lsRF has been introduced to account for the contrast dependence of size tuning. Problematically, not all studies nor even all the latest studies use this means of defining different portions of the RF, which has led to discrepancies in how the sizes of different RF areas have been reported. The following sections will attempt to integrate the results from studies employing different means of classifying RFs.

The studies reviewed in Angelucci and Sainsbury (2006) suggest that geniculocortical FF projections linear integrate signals in the hsRF, while lateral connections underlie the lsRF and may thus account for the contrast dependence of spatial summation as well as modulatory effects in the near surround. In particular they argue that excita-

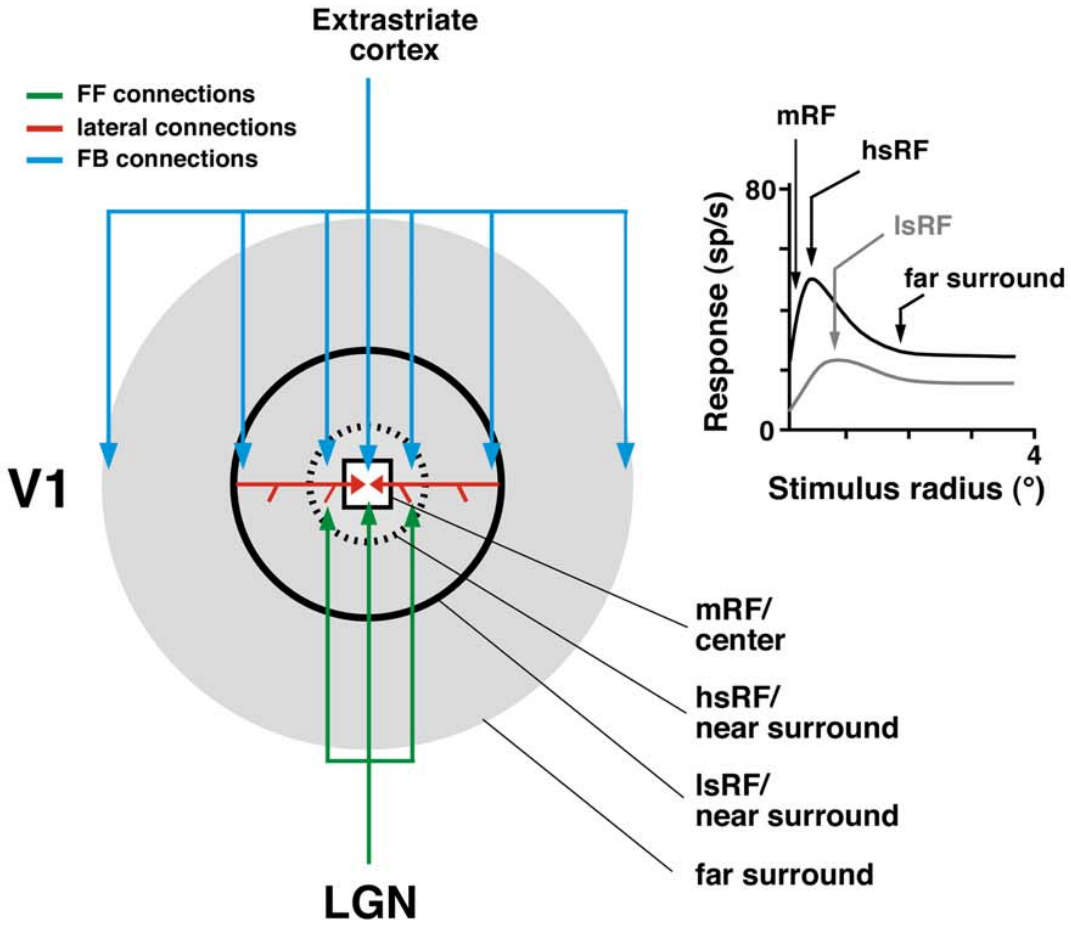


Figure 2.6: The receptive field structure of V1 neurons showing the minimum receptive field (mRF), high contrast summation RF (hsRF) and low contrast summation RF (lsRF). Reproduced from Angelucci and Bressloff (2006).

tory and inhibitory neurons differentially integrate their inputs, where the excitatory population integrates input linearly, inhibitory interneurons respond non-linearly responding only weakly with low excitatory drive but rapidly scale up their response to stronger input, effectively putting a brake on excitation, causing it to asymptote and eventually level off (Angelucci et al., 2002b). Other work by Sceniak et al. (1999) also suggests that the effective reduction of lateral integration under high-contrast conditions could be the result of greater depression of cortico-cortical synapses at higher levels of mean activity and the greater shunting of EPSPs resulting from strong tonic inhibition. In addition to these potential intra-areal interactions giving rise to the lsRF and hsRF under varying stimulus conditions, classification through measurement of spatial dimensions and onset latencies suggest that inter-areal FB connections are responsible for modulatory influences from the far surround. The influence and spatial properties of each of these projections will be detailed in the following sections.

Before considering what exact interactions might be giving rise to the lsRF and hsRF it is worth considering why it might be important to have differential signal integration under varying stimulus conditions. An effective reduction in the spatial integration area at high contrast makes functional sense; a strong signal generally provides a much better signal-to-noise ratio allowing a single neuron to precisely spatially localize image features even when they are only integrating over a small area. However under weaker stimulus conditions this spatial precision is sacrificed for greater sensitivity for detection, which is achieved by integrating over a larger area.

It is unclear whether this simple story attributing the different parts of the receptive to specific projections is realistic, as there will surely be overlap. Nonetheless, we will review the hypothesized contributions to the receptive field in detail so we can later confirm whether the different components in the model can actually account for the contrast-dependent changes in receptive field integration.

Geniculocortical Afferents and the Minimum and High Contrast Summation RF The primary visual cortex receives most of its driving inputs through the three previously detailed M-, P-, and K-cellular geniculocortical pathways. The M pathway principally terminates in layers 4Cα and 6, the Parvocellular afferents terminate in layers 4A, 4β and 6, while K cells primarily target layer 1 and some regions of layer 3. By combining anatomical-tracing with physiological recording the spatial extents of feedforward connections have been measured in detail and linked back to the different RF regions.

The minimum response field as defined above is commensurate to the classical RF and is usually mapped using drifting gratings masked to a small disk of optimal parameters for that particular neuron. It is surrounded by the high contrast summation RF, which is measured by increasing the size of a drifting grating disk at high contrast until the neuron reaches its peak response. Using a combination of tracing and electrophysiological recording (Angelucci and Sainsbury, 2006) found that the visuotopic extent of LGN afferents matches the hsRF size of the target V1 neuron. The diagram and bar chart in Figure 2.7 show how closely the estimates from tracing studies match the results from physiological classifications of RF areas for magno- and parvo-cellular pathways. The close match between these different experiments suggests geniculocortical afferents may underlie the extent of a V1 neuron's mRF. Recent evidence has also shown that the an LGN neurons hsRF is roughly commensurate with a V1 cell's mRF. This seems to suggest that the mRF of V1 cells is a product the summation of LGN cells at their peak spatial summation, while the hsRF region of V1 neurons is defined by the integration of excitatory inputs from partially sup-

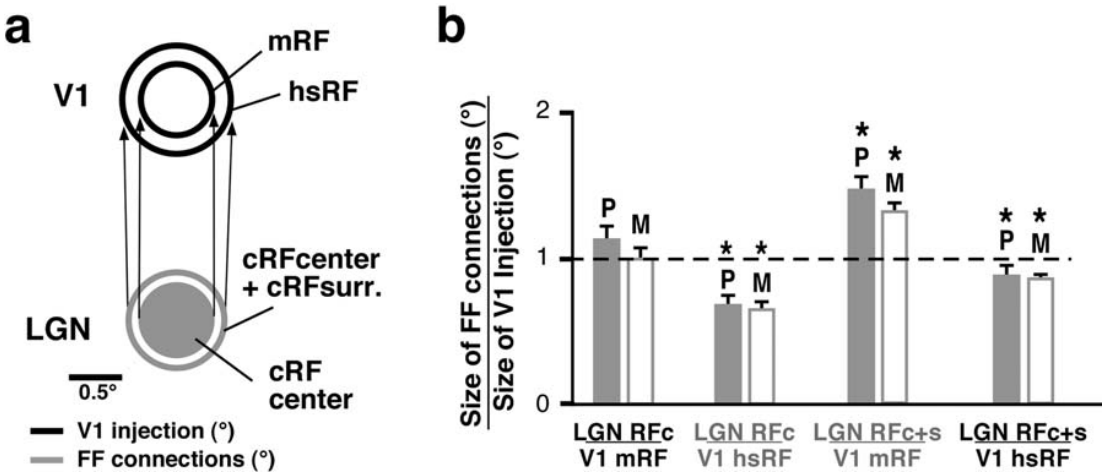


Figure 2.7: Comparisons between electrophysiological characterization of RF structure and the spatial structure of geniculocortical projections to V1 in (a) diagrammatic and (b) chart form. Both demonstrate that the mRF and hsRF are coextensive with the spatial extents of geniculocortical afferents to V1. Reproduced from Angelucci and Bressloff (2006).

pressed LGN cells. Beyond that it seems likely FF components partially contribute to surround suppression in V1, however the spatial scales of surround modulation as well as its orientation specificity seem to rule out LGN afferents as the major contributor to the modulatory surround (Angelucci et al., 2002b; Angelucci and Sainsbury, 2006).

Having established the contribution of geniculate afferents to the RF of V1 neurons, it is time to look at their spatial distribution. In their extensive studies and culminating review paper, Angelucci and Bressloff (2006) first fitted the iDoG to the spatial summation response curve of a number of V1 neurons, then injected the recording sites with tracers and measured the labeled connections and cell bodies. The linear extents of the labeled connections were converted to visual field coordinates using magnification factor (MF) estimates by Van Essen et al. (1984). The anatomical extent of the label was also measured in the LGN and then converted to visual space coordinates using MFs measured by Connolly and Van Essen (1984) and LGN RF size estimates by Derrington and Lennie (1984). These calculations were used to arrive at the aggregate receptive field (ARF) size, which takes the form:

$$ARF_{deg} = D^{\circ} + RF_{mean} \tag{2.1}$$

where RF_{mean} is the mean RF size of cells recorded at the edge of the injection site, which could reflect the mRF, hsRF or lsRF, and D° is defined as:

$$D^{\circ} = D_{mm}/MF_{mm/deg} + S_{deg} \tag{2.2}$$

where D_{mm} is the diameter, $MF_{mm/deg}$ is the magnification factor and S_{deg} is the RF scatter at the injection site. The results show a close match between mRF and hsRF sizes as estimated in V1 and sizes of the RF center and RF center + surround as measured in LGN, once again reaffirming the idea that the mRF and hsRF are primarily driven by geniculocortical afferents. The latest review (Angelucci and Bressloff, 2006) has measured the size of the hsRF in V1 neurons of macaques at 2-8° eccentricity as having mean of about $1° \pm 0.1$, which is roughly 2.2x larger than the mRF of the same cell based on results from Angelucci et al. (2002b) and Levitt and Lund (2002). Estimates from the latest anatomical study summarized in table 4.3 are slightly higher, with means of 1.09° and 1.41° in layer 4Cα and 4A/Cβ respectively.

In addition to the spatial extents of V1 RFs, Angelucci and Sainsbury (2006) also estimated the *number* of LGN afferents that would contact an individual V1 neuron. According to their estimates a single neuron in layer 4Cα can be expected to receive roughly 11 projections from LGN M-cells. Although they were not able to put their own estimates to the Parvocellular pathway, based on anatomical data from cats they determined that on average 10 geniculate cells converge on a V1 layer 4 cell, having observed only a maximum of 30. They conclude that the geniculocortical pathway in macaques exhibits an even lower level of convergence than in cats.

Lateral Connections and the Low Contrast Summation RF The patchy horizontal connections in the early visual cortex are one of the most striking features in the cortex and have been proposed as the mechanism for a number of observed phenomena, including the contrast dependence of size tuning, which is why it is thought they underlie the extent of the lsRF. Classically, however, developmental models have assumed that lateral connectivity manifests itself through short-range excitatory and longer range inhibitory connections (von der Malsburg, 1973; Obermayer et al., 1990). The most striking property of these connections is that they seem to link iso-orientation domains as can be seen in Figure 2.8, which has implicated them in a wide range of effects including long-range integration under low-contrast conditions.

The spatial scale of these connections has led several studies to conclude that they may mediate modulation of RF-center properties in the near surround (Angelucci et al., 2002b). Lateral connections could therefore provide a simultaneous mechanism for a number of observed effects, including the expansion of the summation RF at low stimulus contrast (Sceniak et al., 1999), co-linear facilitation (Mizobe et al., 2001), and suppression from the near surround outside the hsRF but within the lsRF (Sceniak and Hawken, 2001; Levitt and Lund, 2002). The previous section showed that such phenomena could not be adequately accounted for by geniculocortical afferents; measurement of spatial extents and response latencies of horizontal connections reaf-

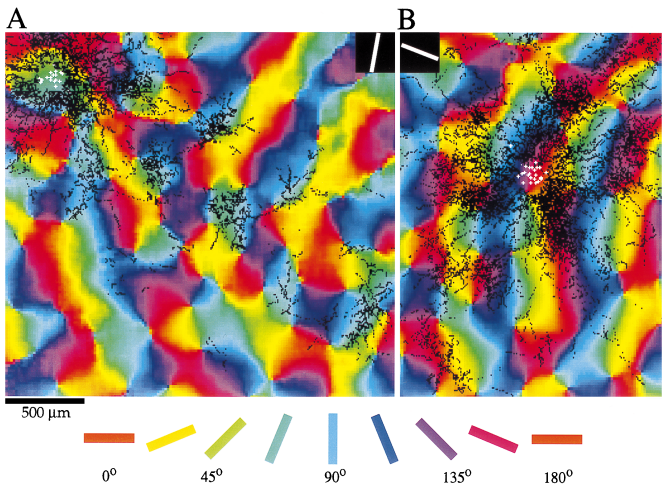


Figure 2.8: Illustrative examples of patchy lateral connections in the treeshrew visual cortex overlaid on an orientation map, clearly showing linking between iso-orientation domains. Reproduced from Bosking et al. (1997).

firm this view and have shown that the lsRF and lateral connections are coextensive (Angelucci et al., 2002b).

Apart from the exact spatial dimensions of horizontal connections, it is important to highlight several other functionally important properties. While layer 2/3 neurons display patchy connectivity, linking regions with similar functional properties such as orientation preference, this has been shown not to be the case in macaque layer 4B and upper 4Cα (Angelucci et al., 2002b). In addition, it was found that horizontal connections in macaque V1 are isotropic in visual space, unlike the anisotropy along the axis of preferred orientation observed in tree shrews (Bosking et al., 1997) and several other species. This could also indicate that orientation-specific contour completion in macaques is mediated by feedback connections. Long-range horizontal connections have been shown to elicit only subthreshold responses (Hirsch and Gilbert, 1991) and are thus limited to a modulatory influence. However, as surround modulation extends far beyond the monosynaptic spread of lateral connections, it is unlikely they account for modulation from the far surround. Polysynaptic chains of lateral connections are also an unlikely substrate for the far surround due to the slow conduction velocity of their axons, and their modulatory rather than driving nature. Bair et al. (2003) also showed that onset latencies of suppression from the far surround were almost equal to the delays from the near surround. This makes it likely that far-surround modulation is mediated primarily by inter-areal feedback connections, which we will look at in detail at a later stage.

Having established that spatial profile of lateral connections is commensurate to that of the lsRF and vice versa, the data from both sources will be laid out and analyzed. Anatomical data suggests that the spatial spread of lateral connections can be

anywhere between 3-10 mm (on average 6-7 mm) in total diameter (Angelucci et al., 2002b). Along its principal axis, the visuotopic monosynaptic spread of V1 horizontal connections has a mean of $2.47° \pm 0.3°$. This falls well within the range of estimates for the lsRF as published in a number of studies (Shushruth et al., 2009; Sceniak et al., 1999; Sceniak and Hawken, 2001), which were fit with the same integrated DoG model and stimulus protocol as used in the Sceniak et al. (2006) paper on the spatial properties of LGN neurons, reviewed previously.

In summary, there seems to be some agreement that lateral connections could underlie the extent of the lsRF and mediate a number of effects in the near surround, including contrast-dependent size tuning, co-linear facilitation and iso-orientation contrast suppression. The extents of horizontal connections range between 3-10 mm, which averages to around 2.5° in visual space.

Feedback from Higher Cortical Areas and the Far Surround As the previous two sections have shown, modulatory influences to V1 RFs extend well beyond the spatial spread of both geniculocortical afferents and horizontal connections. This extended modulatory field is known as the far surround and is thought to be mediated by feedback from higher cortical areas. The far surround has generally been characterized as suppressive, especially for iso-oriented gratings in the center and far surround. More detailed analysis has shown that the far surround can also exhibit response facilitation under some stimulus conditions. This section will characterize the function, termination patterns and spatial extents of feedback connections from higher cortical areas to V1.

The notion of a hierarchical organization of cortical visual areas has been around for quite some time and more recent analysis of feedforward and feedback connections has affirmed this view. At the bottom of this hierarchy is V1, sending partially segregated FF projections to areas V2, V3, V4 and the middle temporal (MT) visual area, which all send FB projections back to V1 (Felleman and Van Essen, 1991). Feedforward projections from V1 to V2 arise mainly from layer 4B and to a lesser degree from layer 2/3 and 6. Feedback connections, on the other hand, arise from layers 2/3A and 5/6 and terminate in the same layers from which FF connections are sent, which means the cells projecting up the hierarchy often overlap with the termination regions of FB projections being sent back down (Angelucci et al., 2002b).

Just like lateral connections, FB connections do not drive their target cells, exhibiting only modulatory influence on the RF center (Bullier et al., 2001). Inactivation of areas V2 and MT has been shown to reduce the firing rate of V1 neurons to stimuli in their RF center (Hupé et al., 1998), suggesting FB inputs are summed with FF inputs to increase activity. The exact balance between excitation and inhibition of FB con-

nections is so far not very well explored in macaques but evidence from rats suggest that they almost exclusively target excitatory cells. However Angelucci and Bressloff (2006) and Schwabe et al. (2006) have proposed a configuration where FB connections in the far surround target pyramidal neurons, which in turn send monosynaptic horizontal connections to excitatory and inhibitory neurons in the RF center and can thus mediate both suppressive and facilitatory effects depending on stimulus properties.

Feedback connections have been thought to underlie a number of top-down effects in V1, including attention (Treue, 2003) and the reverse hierarchy theory of visual learning (Ahissar and Hochstein, 2004), but more recent studies have suggested they contribute directly to the response of V1 neurons to simple visual patterns (Angelucci et al., 2002b; Angelucci and Bullier, 2003; Schwabe et al., 2006). Notably Schwabe et al. (2006) and Ichida et al. (2007) seem to have resolved the conflicting evidence about the far surround's dual suppressive and facilitatory role. Using both experimental and theoretical work they found that while the far surround is suppressive under high contrast conditions, the response of a neuron to a low contrast stimulus in the RF center is facilitated by a small annular stimulus in the far surround. This indicates that excitatory and inhibitory surround mechanisms have similar extents and that the sign of their contribution depends on changes in local excitation/inhibition balance brought about by surround stimulation.

Feedback projections from higher cortical areas to V1 mediate a number of important contextual effects and have been implicated in the early stages of visual attention, but also seem to be closely involved in the processing of simple visual stimuli. This section has summarized current knowledge on the spatial termination patterns of FB connections to V1, indicating how they could give rise to functional properties of V1 information processing. While the role of feedback connections always needs to be taken into consideration, it will not be the focus of this thesis, which will limit itself to afferent and lateral connectivity in the primary visual cortex. This limitation is primarily for practical reasons, in that until suitable models of extrastriate cortex are available, providing suitable feedback to V1 will remain very difficult and poorly constrained.

2.2 THE INTERNAL CIRCUITRY OF STRIATE CORTEX

The previous section outlined the overall structure of the early visual system, breaking down the contribution of inputs from various sources on the receptive field (RF) of neurons in primary visual cortex (V1). While this review provides general anatomical constraints and sheds some light on the functional circuitry underlying many of the contextual effects that have been observed in the primary visual cortex, it does

not address many of the fundamental questions about the functional significance of recurrent cortical processing. In particular, it does not address the delicate balance in both strength and spatial extent between excitation and inhibition that is required to halt runaway excitation, sparsify the inputs, and thereby allow for the formation of concentrated activity bubbles around which self-organization can take place. This section will cover how the understanding of intra-cortical connectivity has evolved and how this connectivity has been implemented in various developmental and non-developmental models.

2.2.1 *The Mexican Hat*

Before complex tracing and circuit reconstruction techniques became available, considerations about the connectivity profiles in the cortex were largely theoretical or based on data from peripheral regions, which were at the time more amenable to study. Anatomical data and electrophysiological studies in the retina had shown that there is a strong lateral inhibitory component involved in decorrelating photoreceptor activities, thereby enabling more efficient coding of the input (Atick and Redlich, 1992). Lateral inhibition was taken to be a general principle of sensory systems and, among others, Blakemore et al. (1970) suggested repulsive interactions between neighboring contours could account for psychophysical data. Further evidence of lateral inhibition as a general feature of sensory systems was provided by a variety of theoretical models of self-organization, highlighting the necessity of effective local excitatory and long-range inhibitory interactions for the formation of local activity bubbles, which in turn provided the basis for orderly map organization (von der Malsburg, 1973; Miller et al., 1989).

The connectivity profile employed in the various self-organizing map models became known as the Mexican hat profile due to its strong resemblance to a sombrero. A simulated Mexican hat profile is shown in Figure 2.9, generated through a simple difference of Gaussian (DoG) whereby a small excitatory Gaussian kernel is combined with a larger inhibitory kernel. Numerous cortical models have successfully employed this connection profile to explain a variety of effects ranging from topographic map organization and orientation tuning to contextual modulation. Problematically, it is unclear how biologically realistic the assumption of strong local excitation and broadly tuned and longer range inhibition really is.

The fundamental question then is whether there exists a set of connections that could account for the local excitatory and longer-range inhibitory component. A theoretical model by Kang et al. (2003) suggests that a relatively small isotropic but fast inhibitory projection could account for the inhibitory component of the Mexi-

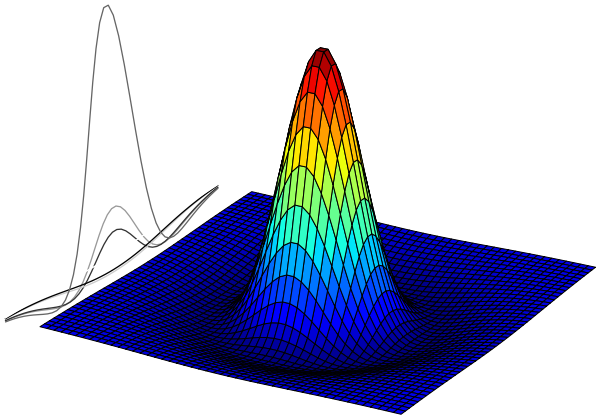


Figure 2.9: A 3D plot of stereotypical Mexican Hat connectivity.

can hat. To get a better understanding of how the Mexican hat might arise from a circuit where the excitatory connections are actually generally much longer, we will therefore investigate the specific circuits involved.

2.2.2 *Contrast dependence of suppression*

An interesting but related phenomenon is that suppression and facilitation can arise in the same neuron under under different stimulus conditions. Specifically the contrast levels seems to have complex effects on the overall sign of modulatory inputs to a neuron.

Electrophysiological and optical imaging have both confirmed that strongly driven cortical neurons receive strong net local excitation and long-range lateral inhibition (Grinvald et al., 1994; Sceniak and Hawken, 2001). At high contrasts Grinvald et al. (1994) showed that a neuron responding to a small, central grating stimulus in isolation exhibits far greater levels of activity than when presented with a co-linear surround stimulus alongside the central stimulus. This highlights an interaction between the center and surround RF that is not only dependent on the orientation statistics but also on the contrast levels in the input. In particular Hirsch and Gilbert (1991) and Weliky et al. (1995) showed that lateral connections impinging onto a neuron would exert a small excitatory effect, when embedded in a low contrast surround, while high contrast would flip the sign of these contextual influences and suppress the central neuron's activity. Additionally, Hirsch and Gilbert (1991) found that laterally evoked EPSPs, presumably underlying facilitatory effects, experienced strong

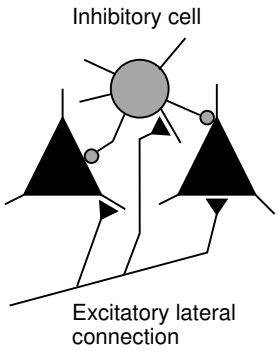


Figure 2.10: Diagram presenting a proposed local microcircuit for long-range suppression through di- or poly-synaptic circuit in V1. Reproduced from Miikkulainen et al. (2005) as adapted from Weliky et al. (1995).

voltage-dependent enhancement, speculating that this effect would allow stimuli in the surround to modulate cRF responses without driving the response on its own.

Precisely how these two input-dependent modes of contextual integration emerge is unclear. However, as anatomical tracing techniques have become more sophisticated and biomarkers for different cell types were discovered, attempts have been made at identifying the different cell and connection types involved in this circuit. These anatomical surveys showed that long-range connections extending beyond a single orientation column were almost exclusively excitatory, and 80% of these excitatory synapses target other excitatory pyramidal neurons (Hirsch and Gilbert, 1991; Kisvárday et al., 1997). The remainder of these connections were shown to target inhibitory interneurons, which would in turn contact pyramidal neurons, suggesting a di- or poly-synaptic mechanism for long range suppression. On the basis of some of this work Douglas and Martin (1991) developed what has become known as the canonical microcircuit for the neocortex. This circuit includes separate inhibitory and excitatory neurons, which are driven by thalamic afferents and recurrent connections. Further work has fleshed out the spatial profiles of these connections, which ultimately gave rise to the simplified circuit described in Figure 2.10. This proposal goes some way toward reconciling anatomy with the experimentally measured functional connectivity profile at high contrast levels.

More recent attempts at reconciling anatomy with function have been able to further resolve some aspects of the problem. In particular, there is clear evidence showing that excitatory synapses onto excitatory and inhibitory neurons differentially target their recipient neurons. Excitatory connections onto inhibitory interneurons seem to preferentially synapse perisomatically, in contrast with recurrent long-range excitatory connections which have been shown to target their recipient neurons dendritically (Gilbert and Wiesel, 1990; McGuire et al., 1991). Additionally, at least a subset of inhibitory interneurons seem to preferentially target the soma of pyramidal and stel-

late cells they inhibit (Markram et al., 2004). On that basis it is reasonable to assume that inhibitory connections are, in general, stronger and may act divisively.

Although these divergent properties of excitatory and inhibitory neurons were only discovered relatively recently, it had long been proposed that inhibitory interneurons are inherently more effective at suppressing activity than recurrent excitatory connections are at exciting the network, but due to a high threshold or some other related mechanism the inhibitory neurons are not strongly recruited unless there is strong afferent input (Sillito, 1979), as would be the case under high contrast conditions. Although it is now clear that network effects allow for strong long-range inhibition through di- or poly-synaptic connections under the right stimulus conditions, the mechanisms by which contrast-dependent behaviors emerge from the cortical circuit are still only vaguely characterized.

A number of models have been developed to explain contrast dependence of contextual effects on the basis of the general principle of asymmetry between the response properties of excitatory and inhibitory neurons. One of the first to publish such a model were Stemmler et al. (1995), who suggested inhibitory neurons require higher external input rates before activating because they receive significantly less spontaneous background input as compared to excitatory neurons, an effect known as stochastic resonance. Although this mechanism has at least been theoretically reaffirmed (Bezrukov and Vodyanoy, 1997), there is no experimental data establishing it as a functionally significant mechanism in V1. Other models hoping to account for a wider array of RF effects implement such a mechanism directly by setting a higher threshold in the inhibitory population and introducing very strong lateral excitation of inhibitory neurons (Schwabe et al., 2006). Another suggestion was made by Somers et al. (1998), who in addition to a simple threshold asymmetry, also point to the claim by Thomson and Deuchars (1994) and others (Abbott et al., 1997; Tsodyks and Markram, 1997), that synaptic depression causes recurrent excitation to quickly decline in efficacy during high frequency stimulation, while facilitation of excitatory synapses onto inhibitory interneurons increases transmission efficacy as presynaptic firing rates increase (Thomson et al., 1995).

The suggestion that inhibitory neurons have a higher contrast threshold has become very popular in the theoretical literature of the past 20 years. However, as of yet there is only limited evidence to support this core assumption, and there are a number of alternative or concurrent mechanisms that may explain all or at least some of the contrast-dependent effects.

The fact that the V1 circuit seems to exhibit both a Mexican hat profile and long-range facilitation, at least under some stimulus conditions, seems contradictory. In particular, dissociating the variety of different types of inhibitory interactions has

been very difficult from electrophysiological measurements alone. In order to begin separating the origins of these different influences on a V1 neuron's response, we will therefore more closely investigate the source of surround suppression in the cortex.

2.2.3 *Surround Suppression: Feedforward or Feedback?*

The last section highlighted how little we still know about the origin of surround suppression and inhibition. There is still significant controversy about whether surround suppression originates in feedforward or feedback pathways or whether both contribute over different spatial scales. The literature includes suggestions that suppression of the classical RF by stimuli in the surround is mediated through synaptic depression in the thalamo-cortical afferents (Carandini et al., 2002), broadly tuned inhibition by thalamo-cortical recipients, long-range excitation of local inhibitory interneurons, or even through various feedback mechanisms. This section will detail the evidence for each of these proposals and the possible anatomical origin of each of these mechanisms, teasing apart the circuit by looking at interactions between surround suppression, stimulus size, and contrast.

Since the circuitry of the cortex is so complex, the task of identifying feedforward and feedback contributions to surround suppression is difficult. Although only a starting point, one way of roughly separating these two possible contributions is to look at the time course of suppression. In the literature, early and late contributors to surround suppression have been identified (Webb et al., 2005). The early component is characterized as being driven by lower CRF contrasts with spatio-temporally broadband tuning and little adaptation (Levitt and Lund, 1997; Cavanaugh et al., 2002a). The late component, on the other hand, is driven more strongly by high-contrast stimuli in the CRF, has sharp spatio-temporal tuning, and can be strongly affected by adaptation (Levitt and Lund, 1997). Evidence suggests that the early, broadly tuned component originates in the LGN and the thalamocortical recipient layer of visual cortex (Blasdel and Fitzpatrick, 1984; Hawken et al., 2009). In monkey cortex in particular, this broadly tuned suppressive effect is only weakly evident in the LGN and is thought to arise much more strongly in layer 4 of striate cortex (Webb et al., 2005), which may have some correspondence to the broadly tuned inhibitory population identified by Hirsch et al. (2003).

There is also the possibility that the origins of surround suppression differ between species, which may arise due to the difference in orientation tuning of LGN neurons. (Ozeki et al., 2009a) found that $GABA^A$ blockade using bicuculline methiodide (BMI) in V1 did not abolish surround suppression in V1, suggesting it is the result of a reduction of thalamocortical inputs driving both excitatory and inhibitory in the cortex.

Carandini et al. (2002) on the other hand suggests that there is a synaptic explanation for suppression, primarily due to the speed with which the early component of suppression arrives, its immunity to cortical adaptation, and the fact that it is restricted to the CRF. However, they concede that synaptic depression does not account for gain control and the abolishment of cross-orientation suppression by $GABA^A$ blockade and the later component of suppression arriving from the surround, so a mechanism that can account for all these phenomena may still be preferable. In that vein, Webb et al. (2005) propose two inhibitory mechanisms, one of which sums local activity in a neuron's CRF and divides the response of the CRF, and a later component that receives inputs from a much larger area but provides narrowly tuned suppression. The broadly tuned component in particular has a strong relationship with contrast gain control, which has been firmly established to act divisively. Independent work by Xing et al. (2005) supports the suggestion of two inhibitory components and further expands on the size dependence of these two components. Specifically, Xing et al. (2005) conclude that the tuned component is recruited far more strongly for larger stimuli, which seems to confirm a contribution from beyond the CRF.

This recent work has identified two clear and distinct inhibitory components, but has not yet fully described which mechanisms and circuits by which they are mediated. The next section will attempt to address this shortcoming.

2.2.4 *Distinct Inhibitory Populations*

In order to begin to understand the origin of intracortical surround suppression mediated by various local inhibitory circuits, it is necessary to consider the different candidate cell classes. While there is a long list of different inhibitory cell types based on their morphology and spiking behavior, recent techniques have divided inhibitory cells into several broad, functionally distinct classes based on their immunoreactivity. The two cell types considered here are parvalbumin (Pv-ir) and somatostatin (Sst-ir) immunoreactive neurons, which are primarily differentiated by the cellular locus of their synaptic targets. While Pv-ir neurons seem to target pyramidal cells perisomatically, Sst-ir neurons target their recipients dendritically. Along with the Vasointestinal protein (VIP) expressing interneurons these cell classes make up the majority of inhibitory cells in the cortex. The following subsections will detail the anatomical, physiological and functional differences between these cell classes.

Parvalbumin Immunoreactive cells

The two main cell types that are immunoreactive to parvalbumin are the chandelier and basket cells (Binzegger et al., 2004). While chandelier cells make up only a small

fraction of GABAergic neurons in the cortex and are primarily found in layer 2/3, the fast-spiking (FS) basket cells are the predominant interneuron subtype in the mammalian cortex across all laminae, accounting for 42% in layer 2/3 and layer 5 and 78% in layer 4 of the cat (Hogan et al., 1992; Huxlin and Pasternak, 2001) and up to 74% across cortical layers in macaque (Van Brederode et al., 1990). The abundance in the thalamocortical-recipient layers and the fact that they preferentially target the soma and proximal dendrites of spiny neurons with multiple strong synapses, exhibiting high probability of GABA release (Freund and Katona, 2007; Markram et al., 2004), ensures that basket cells are of tremendous interest. On that basis it has been suggested that the perisomatic connectivity profile of basket cells gives them the ability to provide shunting inhibition to layer 4 spiny neurons, acting divisively to control their response gain (Wilson et al., 2012).

Basket cells can be further subdivided, primarily based on their size, into clutch and large basket cells. However, all basket cells can make multiple connections onto a target pyramidal neuron (Somogyi et al., 1983) and have a considerable spatial extent (Kisvárday et al., 2002). In particular, studies in cat area 17 and macaque V1 have identified basket cells 1-2 mm in extent (Somogyi et al., 1983; Lund, 1987; Lund and Yoshioka, 1991; Martin et al., 1983) and single-cell tracing studies have even identified large basket cells, which give off a roughly uniform number of boutons across a large diameter extending up to two hypercolumns (Buzás et al., 2001), which corresponds to about 1.5 mm in macaque. A schematic representation of the basket cell connectivity profile is summarized and compared against both the orientation column structure and the excitatory connection profile in Figure 2.11. Functionally such a connectivity profile may indicate that basket cells can suppress neurons with widely varying orientations, which may implicate basket cells as an important mechanism to sharpen orientation preference and would also lead to cross-orientation contextual effects.

In terms of their spiking behavior Pv-ir cells are typically characterized as being fast fast-spiking (FS) neurons, often firing in bursts with a very short response latency. It is important to note however that not all fast-spiking cells are immunoreactive for parvalbumin and that not all PV neurons are fast-spiking. Further, evidence from somatosensory cortex in rodent and lagomorph species suggests they receive strong input from thalamocortical afferents arriving in layer 4 and very effectively suppress sustained firing from spiny neurons receiving inputs from the thalamus (Swadlow, 2003), implicating them in feedforward inhibition. A possible feedforward inhibition circuit is shown in Figure 2.12. Their effectiveness in suppressing feedforward activity can be explained by the large number of thalamocortical axons they receive, which exhibit faster kinetics than those targeting spiny neurons (Cruikshank et al., 2007;

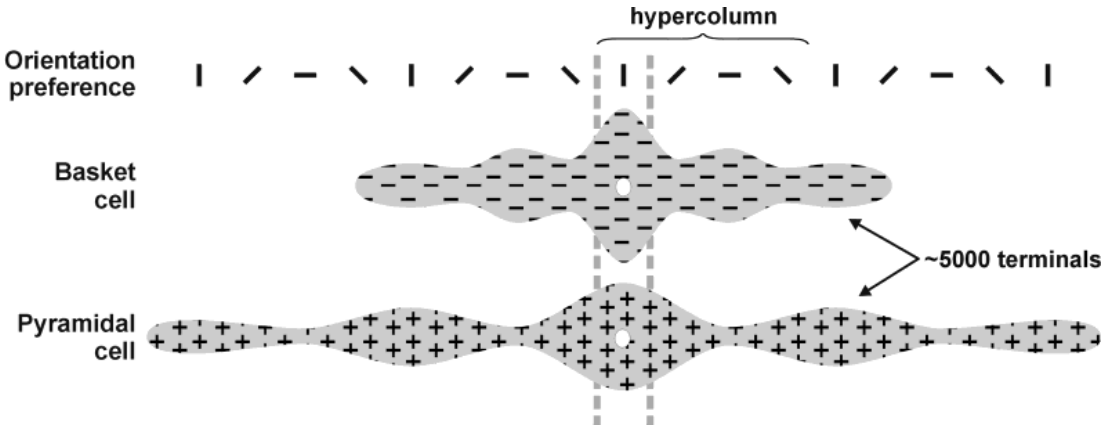


Figure 2.11: Summary schematic comparing relationship between long-range basket cell and excitatory connectivity with the underlying orientation preference structure. Upper legend represents different orientation domains in the cortical topographic map. Grey dotted lines indicate the orientation column within which the soma of the simplified basket cell and excitatory neuron are found. The grey field with minus signs indicates the extent of inhibition provided by the basket cell connections considered in the current study. The grey region with plus signs indicates the excitatory field of a stereotypical pyramidal cell, based on previous data by Bosking et al. (1997); Kisvárday et al. (1997) and others. The height of the grey plus/minus regions indicates the number of axon terminals provided in that column. While basket cell terminals show local maxima every half hypercolumn distance, pyramidal cell terminals are maximal at every full hypercolumn distance. Reproduced from Buzás et al. (2001).

Gabernet et al., 2005), and the fact that they evoke large inhibitory responses in spiny cells (Cruikshank et al., 2007; Gabernet et al., 2005).

It is also important to note that the thalamocortical synapses onto the Pv-ir population have been shown to be depressed by repetitive activation, resulting in weaker feedforward inhibition at high stimulation frequencies (Gabernet et al., 2005), a property which may indicate lower activation of the PV population at high contrasts. The Pv-ir population may also play an important role in network homeostasis, as activity blockade has been shown to decrease the efficacy of Pv-ir inhibition (Bartley et al., 2008), thereby indirectly up-regulating activity in excitatory cells. Further, selectively up-regulating Pv-ir cells using optogenetic stimulation was shown to have a similar effect as lowering the contrast, which is to increase preferred size and weakening surround suppression (Nienborg et al., 2013). This evidence is compatible with the idea that Pv-ir neurons provide strong feedforward inhibition such that the input drive in the cortex is decreased, as would be observed under low-contrast conditions. Overall then, Pv-ir neurons show strong interaction with stimulus contrast and may be involved in regulating the gain of the network, with complex implications for the contrast response of the network.

There is still considerable debate on the extent to which this subpopulation is tuned to a particular orientation. Most visuo-cortical models employ broad, non-specific GABAergic inhibition (Somers et al., 1998; Troyer et al., 1998). This seems to be supported by anatomical evidence, which has long shown that inhibitory projections are generally diffuse and display low specificity for specific stimulus features (Albus and Wahle, 1994; Kisvárday et al., 1997). Further, electrophysiological data paints a similar picture, revealing suppression that is broadly tuned for stimulus attributes, providing orientation-unspecific suppression from a visual region that is coextensive with the classical RF (DeAngelis and Robson, 1992). Recent attempts at studying the Pv-ir neurons at the single-neuron level have revealed a mixed picture. While the cells as a whole were broadly tuned to various stimulus features, individual branches often displayed very high specificity, which may underlie subfield antagonism and contribute to a push-pull configuration (Kisvárday et al., 2002).

Studies in mouse visual cortex seem to confirm such a dual purpose of Pv-ir neurons, although they also find higher heterogeneity in the Pv-ir population (Runyan et al., 2010). This result may indicate a laminar differentiation in function, as studies of Pv-ir neurons in the thalamocortical recipient layer 4 have characterized them to exhibit very broad tuning, due to their low spiking threshold and more convergent inputs (Ma et al., 2011). However, even in layer 2/3 most mouse Pv-ir cells displayed broad orientation tuning (Hofer et al., 2011). Further studies in cat area 17 find very similar results identifying a class of inhibitory complex cells in layer 4 exhibiting

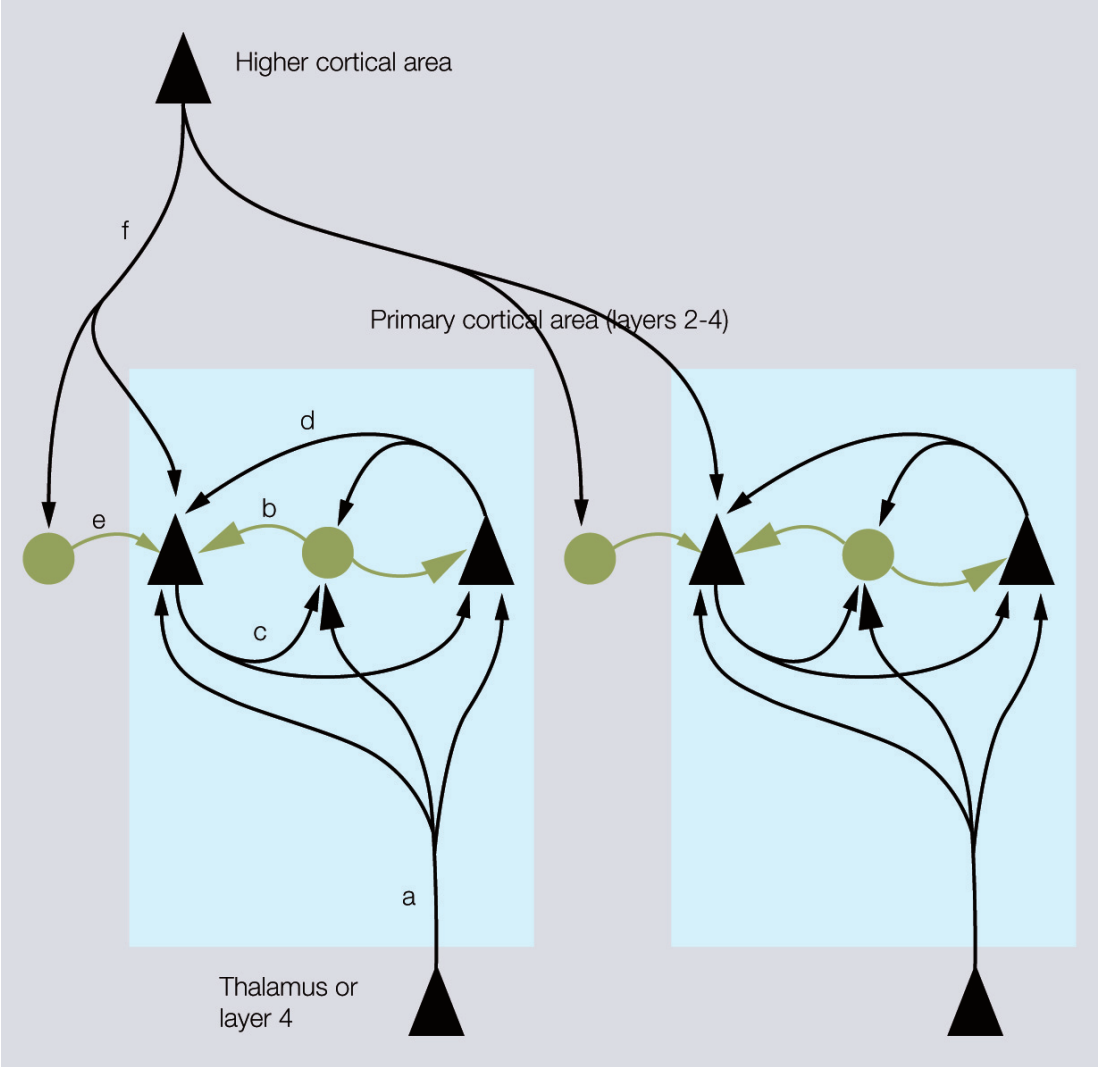


Figure 2.12: Inhibitory networks involving fast-spiking parvalbumin-immunoreactive neurons in thalamocortical, interlaminar, and interareal cortical circuits. Feedforward excitatory thalamocortical inputs to pyramidal cells, spiny stellate neurons (▲) and fast spiking interneurons (●) in layers 2-4 (a). Inputs to interneurons are stronger (large arrowheads) than inputs to spiny cells. PV neurons provide strong (large rectangular endings) feedforward inhibition (b) to spiny cells. Feedback inhibition (c) results from PV neurons that are excited by the same spiny neurons that they inhibit. These reciprocally connected spiny neuron/PV neuron pairs share common inputs (e.g., cells in layer 4 from thalamus or cells in layer 2/3 from layer 4) creating recurrent excitatory (d) and inhibitory subnetworks (contained within blue shaded boxes). 'Lateral' inhibition (e) of these subnetworks results from PV neurons that are driven by excitatory feedback connections (f) from outside the subnetworks (e.g., by layer 5 to layer 2/3 connections or feedback from higher cortical areas). Notice that 'lateral' inhibition is weaker (small rectangular endings) than feedforward and feedback inhibition and impinges on multiple subnetworks. Reproduced from Burkhalter (2008).

weak orientation tuning (Hirsch et al., 2003), which can primarily be accounted for by the tuning of synaptic responses and a lower spike threshold (Nowak et al., 2008). A more recent study in auditory cortex seems to affirm these conclusions, hypothesizing that 'while PV neurons may provide broadly tuned feedforward inhibition for a rapid control of ascending inputs to excitatory neurons, the delayed and more selective inhibition from SOM neurons may provide a specific modulation of feedback inputs on their distal dendrites' (Li et al., 2014). However, a more extensive study in layer 4 of the cat found no significant broadening in the synaptic orientation tuning of fast-spiking interneurons over that exhibited by regular spiking cells in the same region (Cardin et al., 2007); only when looking at spiking responses was a significant broadening found. As an explanation of this apparent difference in tuning between rodents, which exhibit no orientation maps, and primates or carnivorans, it has been suggested that these cells pool the input from nearby neurons and inherit their orientation preference (Isaacson and Scanziani, 2011). There may also be a layer-specific segregation in their function providing push-pull suppression in thalamocortical recipient layers and broader inhibition in upper layers.

Based on our current knowledge of Pv-ir cell population and more specifically the basket cells, it is clear that they provide a good candidate mechanism to account for a number of important phenomena. Their fast response profile, large spatial extent and seemingly broader orientation tuning may allow them to carry out fast, adaptive gain control and broadly tuned suppression, thereby sharpening the orientation preference of the PNs in their vicinity.

Somatostatin Immunoreactive cells

The Sst-ir population has been characterized to a lesser extent, but a general consensus is beginning to emerge around their function and electrophysiological properties. The Sst-ir cells account for around half of the non-PV-expressing neurons (Gonchar et al., 2007; Xu et al., 2010) and preferentially synapse onto distal dendrites and dendritic tufts of pyramidal neurons (Di Cristo et al., 2004; Silberberg and Markram, 2007), on the basis of which it has been suggested that Sst-ir neurons act subtractively (Wilson et al., 2012).

In trying to characterize the excitatory inputs to Sst-ir cells in layer 2/3 of the mouse, Xu and Callaway (2009) determined that unlike Pv-irs, the main source of excitation of Sst-irs were horizontal axons within layer 2/3, not the ascending layer 4 axons. This property also contributed to the size-dependent responses of the Sst-irs, which were shown to be recruited progressively more strongly when they were exposed to optogenetic photostimulation of increasing diameter. Importantly, the Sst-ir response grew larger even when photostimulation reached beyond its maximal den-

dritic extent, demonstrating that the recruitment of increasingly more distant pyramidal cells provided their main excitatory drive. This putative circuit is shown in Figure 2.13, exhibiting their pooling of tuned, excitatory input from pyramidal cells across a large area yet providing only very local inhibition. It also seems that Sst-ir interneurons are capable of disinhibiting the thalamocortical recipient layer 4 by targeting Pv-ir cells (Xu et al., 2013).

In terms of their response properties, mouse visual cortex Sst-ir neurons have been shown to exhibit much lower levels of spontaneous and evoked activity, stronger orientation and direction selectivity, and longer response latencies than Pv-ir neurons (Ma et al., 2011). These properties are consistent across both layer 2/3 and 4 and may point to a role in gating later-arriving intracortical excitatory inputs. In terms of their tuning properties, Sst-ir neurons display smaller On/Off subfields with less overlap than nearby spiny neurons and orientation tuning, on par with pyramidal neurons. While most data on their tuning properties comes from rodent visual cortex, which does not exhibit a topographic organization of orientation tuning, it seems the stronger orientation selectivity of Sst-ir cells can be accounted for by their preferential connectivity to neurons with similar orientation preference as well as a higher spiking threshold and weaker excitatory inputs (Bartley et al., 2008). Another interesting feature of Sst-ir response properties is the fact that although their excitatory inputs are weak, those inputs are facilitatory, resulting in delayed but strong activation under high frequency stimulation (Beierlein et al., 2003; Bartley et al., 2008; Tan et al., 2008). Based on these properties, Sst-ir neurons would only be recruited to provide significant inhibition if the stimulus contrast or size reaches a certain level (Adesnik et al., 2012). Thus they could provide a direct mechanism for contrast-dependent size tuning and surround-modulation effects, being only weakly recruited when contrast or stimulus size is low but becoming strongly activated at higher contrast levels and for larger stimulus sizes, providing a possible circuit mechanism for Sillito (1979) proposal for surround suppression.

Another suggested role for Sst-ir neurons is the gating of feedback signals on the distal dendrites of principal neurons. Martinotti cells, which provide strong axonal projections to layer 1 of the cortex, make up a large proportion of Sst-ir neurons in layer 2/3 of the cortex and are therefore well placed to suppress feedback signals arriving in the superficial layers of the cortex (Fanselow et al., 2008; Gentet et al., 2012). In the rodent somatosensory cortex, Gentet et al. (2012) showed that Sst-irs in layer 2/3 become spontaneously active during passive wakefulness but are strongly suppressed during active whisking behavior, presumably by the vasoactive intestinal peptide (Vip)-ir population discussed below. The Sst-ir cells may therefore

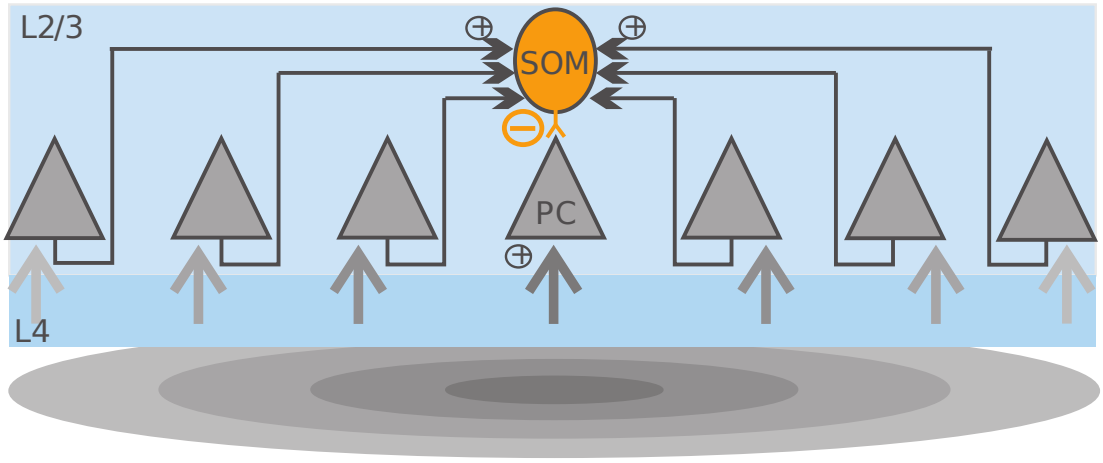


Figure 2.13: Schematic illustration of the cortical circuit in layer 2/3 contributing to surround suppression. As a visual stimulus expands (larger stimuli are shown in lighter grey), recruitment of adjacent principal cells (PCs) increases Sst-ir excitation through horizontal axons (horizontal arrows). Reproduced from Adesnik et al. (2012).

be involved in mediating top-down control of sensory processing, effectively gating context-dependent processing in a state-dependent manner.

In summary, Sst-ir neurons seem to provide delayed and feature-selective feedback inhibition, which puts them in a good position to effectively gate late-arriving intracortical excitatory inputs originating in either lateral or feedback connections, but may also be implicated in suppressing feedforward inhibition by inactivating layer 4 Pv-ir cells.

Vasointestinal peptide expressing interneurons

The material reviewed so far has focused primarily on the two most common types of inhibitory interneurons, the Parvalbumin (PV) and Somatostatin-expressing (Sst) cells. Since that initial work, a number of studies have focused on the role of 5HT3ar-expressing interneurons and particularly the vasointestinal peptide (Vip)-expressing subgroup (Lee et al., 2013; Fu et al., 2014; Higley, 2014; Kepecs and Fishell, 2014).

The Vip subgroup is particularly concentrated in upper, associative layers and feedback layers of the cortex, as shown in Figure 2.14 by Rudy et al. (2011). The most striking finding was their central role in state-dependent modulation during active whisking tasks in rodents. Lee et al. (2013) found that S1-projecting vM1 pyramidal neurons strongly recruited Vip-expressing interneurons in superficial layers of somatosensory barrel cortex, which in turn inhibited somatostatin-expressing interneurons causing effective disinhibition of cortical pyramidal cells. These results were then affirmed through optogenetic stimulation of Vip neurons in mouse V1, artificially mimicking the effects of locomotion (Fu et al., 2014). When considered in conjunction with pre-

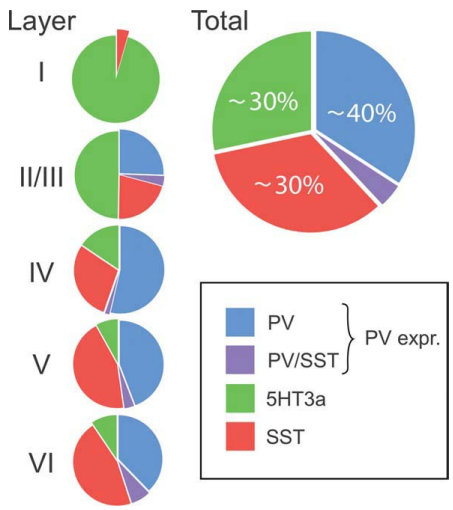


Figure 2.14: Distribution of GABAergic interneurons in mouse S1 cortex by immunohistological marker. Reproduced from Rudy et al. (2011).

vious studies that established strong cholinergic and nicotinic inputs to Vip neurons from the basal forebrain (Wickersham et al., 2007), the evidence suggests a strong involvement of Vip neurons in a cortical circuit responsible for the enhancement of activity in sensory cortex by behavioral state.

Connectivity between different cell types

In order to gain an understanding of the circuits the different interneuron cell types are involved in, it is important to consider their interconnectivity. Several studies have sought to determine the connectivity between Pv-ir, Sst-ir and other interneuron types. The core findings of these studies determined that Pv-ir cells preferentially inhibit one another, Sst-expressing cells avoid one another and inhibit all other types of interneurons, particularly the Pv-ir cells (Xu et al., 2013), while a third type, the Vip-ir cells preferentially inhibit Sst-ir cells (Pfeffer et al., 2013). This connectivity profile is schematically represented in Figure 2.15. In mouse cortex the Pv-ir, Sst-ir and Vip-ir cells accounted for about 40%, 18% and 8% of the GABAergic population, respectively (Xu et al., 2010), and although these percentages vary considerably across species Pv-ir and Sst-ir are always the two most commonly expressed GABAergic populations.

2.3 GABAERGIC REGULATION OF PLASTICITY AND COLUMN STRUCTURE

Experience-dependent plasticity has been shown to shape the organization of the sensory cortex during an initial critical period and beyond. Dark-rearing (Fregnac and Imbert, 1978) and monocular-deprivation (MD) experiments (Shatz et al., 1978) in par-

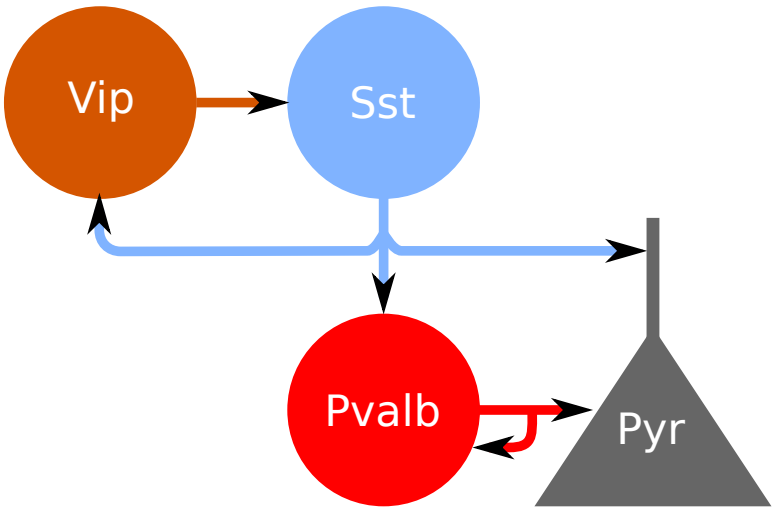


Figure 2.15: Connectivity between somatostatin (Sst), parvalbumin (Pv), vasoactive intestinal peptide (Vip) expressing and pyramidal (Pyr) cell types. Adapted from Pfeffer et al. (2013).

ticular have confirmed the fundamental importance of sensory experience in shaping the development of the cortex. The mechanisms controlling the onset of the critical period and regulation of plasticity thereafter have also been studied extensively. A large body of evidence points to the important role of the inhibitory neurotransmitter γ-aminobutyric acid (GABA) in regulating synaptic plasticity. However, as the above paragraphs have shown, the population of GABAergic neurons is highly heterogeneous with hugely divergent anatomical and functional profiles. Using specific pharmacological and genetic populations, it has been possible to narrow down the involvement of certain interneuron subtypes in shaping critical-period plasticity and column structure in the cortex.

One of the first indications that GABAergic circuits are involved in shaping plasticity came when it was shown that a gene-targeted disruption of the GABA synthetic enzyme glutamic acid decarboxylase 65 (GAD65) could delay critical period onset indefinitely (Fagiolini and Hensch, 2000). In order to further narrow down the specific GABA circuits underlying visual cortical plasticity, more specific pharmacological manipulations were required. On that basis Fagiolini et al. (2004) used benzodiazepine infusions, known to selectively enhance GABA type A ($GABA_A$) receptor-mediated currents through the $\alpha 1$ subunit (Rudolph et al., 1999), in conjunction with MD to prematurely trigger ocular dominance plasticity in mice. These $GABA_A$ receptor-$\alpha 1$ subunits are preferentially enriched at somatic synapses receiving input from Pv-ir large basket cell terminals (Klausberger et al., 2002), strongly implicating large basket cells in visual cortical plasticity.

Beyond controlling the timing of critical period plasticity, further experiments using benzodiazepines have shown strong effects on the columnar organization of the cortex. The experiment by Hensch and Stryker (2004) locally infused regions of cat area 17 with the $GABA_A$ agonist diazepam and an inverse agonist (DMCM) and studied the effects on ocular dominance columns. Chronic treatment with diazepam had little effect in the functional properties of mature cortical neurons in vivo apart from enhancing inhibitory postsynaptic currents. However, the treated hemisphere exhibited reduced binocularity of single unit responses and wider OD columns near the infusion site. Infusion with the benzodiazepine inverse agonist DMCM had the inverse effect, resulting in less discrete and narrower columns near the infusion site. These results suggest that the diazepam-mediated enhancement in competition reduces binocularity of single-unit responses, as well as sharpening and widening the anatomical segregation of monocular regions near the infusion site. This once again suggests that $GABA_A$ inhibitory currents, primarily originating from Pv-ir neurons in the cortex, are fundamentally important to shaping the plasticity and organization of the cortex.

In order to establish how ocular dominance plasticity emerges during monocular deprivation, Kuhlman et al. (2013) developed even more precisely targeted pharmacological manipulations. By selectively expressing specific receptors on Pv-ir cells they were able to selectively up- and down-regulate their activity. Their results indicate that a rapid but transient reduction in Pv-ir cell firing restores pyramidal cell firing to pre-deprivation levels, allowing competitive plasticity to occur. Pv-ir neurons therefore seem to play a permissive role in visual cortical plasticity. Interestingly, adult sensory plasticity such as reinforced associative learning occurs through a similar mechanism, where cholinergic activation of layer 1 interneurons suppresses Pv-ir neural activity allowing associative fear learning to occur (Letzkus et al., 2011). All this work suggests a crucial role for Pv-ir neurons in controlling cortical plasticity during the critical period and beyond.

2.4 CONTEXTUAL MODULATION AND ATTENTION

Arguably, the computational task in vision is to map visual experience to the cortical representation of that particular stimulus or if no such representation already exists, to extract lower level features in order to encode them for future reference. Using this statistical model the brain is then able to decide which visual features carry behavioral importance and which can be safely ignored. As such the neocortex has to combine prior information with the incoming information stream and quickly and reliably identify the most salient stimuli. It has often been argued that this process

is mediated by bottom-up and top-down processes, although it seems likely that there is close coupling between the two. This section will outline high-level models of attentional modulation, and attempts to understand the neurobiological processes behind them and more basic contextual modulation phenomena that may underlie many of these higher level effects.

2.4.1 *Contextual and Attentional Phenomena in V1*

A number of phenomena associated with attention and contextual modulation, including iso-orientation suppression or facilitation, boundary detection, contour completion and noise exclusion have been observed in V1. Although these phenomena are generally associated with bottom-up attention, they lay the foundation for higher level phenomena such as pop-out and figure-ground segregation and may reveal more about general mechanisms applying also to higher visual areas.

Basic contextual effects such as iso-orientation suppression have already been discussed and models have begun to suggest the functional connectivity mediating implicating both lateral and feedback connections. Li (2002) has proposed that pre-attentive bottom-up processes allow V1 to generate a saliency map of the visual input. However, the fact that higher cortical areas have also been associated with saliency signaling and the lack of long-range intra-areal connectivity in V1 suggest that while it can encode local saliency, feedback is required to globally integrate saliency across visual space.

Feedback modulation of V1 activity has been implicated in a number of effects, spatial attention being chief among them. Spatial attention is thought to be able to select multiple low- and high-level objects in the visual space across V1 and higher visual areas (McMains and Somers, 2004). Spatial attention is thought to underlie noise exclusion (Dosher and Lu, 2000) and may be explained by effects similar to what has been experimentally observed during iontophoretic application of ACh. Other effects that have been commonly associated with feedback in some form are the signaling of illusory contours, which have been shown to be negatively signaled or deemphasized in V1 (Ramsden et al., 2001) and boundary detection (Poort et al., 2012).

2.5 NATURAL IMAGE STATISTICS, SPARSITY AND HORIZONTAL CONNECTIONS

It has long been hypothesized that connectivity in the cortex captures the statistics of the sensory input in order to perform predictions and maintain sparse representations of subsequent inputs (Field et al., 1993; Simoncelli and Olshausen, 2001). These

effects cover both that the distribution of visual features that are thought to sparsely represent the visual features of world (Olshausen and Field, 1996), but also suggests that it captures the Gestalt law of good continuation in the form of so-called association fields, which preferentially link neurons which represent edges on a continuous contour (Field et al., 1993). Separately a wide range of work has explored the distributions of light intensities, color statistics and spatial correlations in natural images. In particular, the power-law distribution of spatial frequencies in natural images has been widely discussed in the literature, but ultimately this property largely seems to reflect scale invariance within natural images (Ruderman, 1997).

Numerous studies and models have since been devised to address whether the visual system takes advantage of the correlational structure of natural images. These types of normative models were able to show that surround inhibition, whether subtractive or divisive, could cancel out correlations, effectively whitening or decorrelating the activity in the visual system (Srinivasan et al., 1982; Atick and Redlich, 1992). In doing so they quickly found that simple decorrelation was not sufficient to optimally represent natural images, because whitening does not eliminate all structure in a natural image; e.g., edges and lines remain. By introducing an explicit sparsity constraint, Olshausen and Field (1996) were able to develop V1-like simple cell receptive fields with varying orientations, spatial frequencies and sizes. These models suggested that the sensory system was optimizing two constraints, sparsity and statistical independence. However, even these approaches cannot achieve complete statistical independence, since there are higher-order correlations even between non-overlapping receptive fields.

By introducing divisive normalization, Schwartz and Simoncelli (2001) were able to show that these types of dependencies could be further eliminated. Furthermore, the weights used in the computation of the normalization signal could be specifically optimized to maximize the independence of the normalized responses. Additionally, they demonstrated that the optimal weights were at least partly due to the prevalence of extended contours in natural images.

Attempting to quantify the co-occurence statistics of contours in natural images, Geisler et al. (2001) demonstrated that the performance in contour-detection tasks could be predicted by a local grouping rule derived from the co-occurence statistics. The first explicit link to horizontal connectivity was made by Sigman et al. (2001), who noted that the pattern of long-range patchy connectivity in the primary visual cortex linking iso-orientation columns has a close correspondence with the observation of co-circularity in natural image statistics. Noting the processes of iso-orientation suppression and contour integration, they argue that iso-orientation suppression may serve to further reduce redundancies in neural coding, thereby achieving greater sta-

tistical independence, which would explain why neural responses appear most sparse when presented with natural stimuli. Furthermore recent findings have shown that low-level co-occurrence are sufficient to classify images of animals (Perrinet and Bednar, 2015), suggesting that even early visual areas could be involved in high-level classification tasks. Secondly, observing that visual cortex can also exhibit co-linear facilitation under low-contrast conditions (Sceniak et al., 1999; Kapadia et al., 1999), they suggest that under low signal-to-noise conditions the cortex may act to enhance commonly encountered patterns to aid the identification of contours and form.

These theoretical studies have hugely advanced our thinking about the computations performed by the early visual cortex, yet very little work has been done to look at the actual structure of horizontal connectivity in V1, largely due to the difficulty in obtaining data from more than just a few cells. Even on the question of whether horizontal connections are anisotropic along the axis of preferred orientation of the neuron, as would be expected from theoretical studies, there is conflicting evidence. The result has been confirmed in monkey (Sincich and Blasdel, 2001), tree shrew (Bosking et al., 1997) and cat (Schmidt et al., 1997), but conflicting results have been reported for macaque (Angelucci et al., 2002b). By performing analyses on a tree-shrew dataset from 1997, Hunt et al. (2011) investigated whether horizontal connections captured the co-circularity of natural image statistics. Although they found neurons that exhibited co-circularity and anti-cocircularity and hypothesize a role for both, given the small number of lateral connections fields and the fact that second-order properties are highly sensitive to even small errors in the data, it is unclear how strong this result is. Further research in this area is desperately needed if we are to understand how the patterns of lateral connectivity relate to natural image statistics.

2.6 EXISTING LARGE-SCALE MODELS OF THE PRIMARY VISUAL CORTEX

In the previous sections we laid out a wide range of evidence relating specific empirical studies to each other, to begin teasing apart how low-level interactions can give rise to complex behaviors. A wide range of modeling work has been targeted to that very question, and there exists a wide breadth of models attempting to model the function of the visual cortex at the level of individually spiking neurons or modeling looking specifically at the surround modulation interactions between neurons. Here we will review some of these models, highlighting why although many models explain specific phenomena well, the lack of development and learning means that these models are static and cannot easily explain the dynamic, adaptive nature of computation in the cortex.

Large-scale spiking models

With the introduction of increasing computational power it has become feasible to model large networks of spiking neurons, allowing simplified models of the visual cortex to be created. In particular, large scale simulators such as NEURON (Hines, 1994), NEST (Gewaltig and Diesmann, 2007) and interfaces such as PyNN (Davison et al., 2008) have made it possible to specify complex neural network models at the level of individual neurons and even synapses and ion channels.

These models generally replicate the statistics and rough spatial scales of different types of connections, between different cell classes and investigate the response of the model at resting state or when provided with thalamic input (Shelley et al., 2002; Potjans and Diesmann, 2014). Additionally there are also even larger models attempting to model the visual cortex just as one area of a complex brain network (Eliasmith et al., 2012). What these models have in common is that they are not well suited towards addressing the kinds of question most vision scientists are interested in, particularly in regard to the way the visual cortex can extract and encode the statistics of the visual input and give rise to well-studied effects like contextual modulation. Instead these models are usually used to study only the overall statistics of the responses. As an example, Potjans and Diesmann (2014) develop a highly detailed and refined microcircuit model of the visual cortex, including different cell types and laminar distributions, which they can scale to large cortical networks. The resultant model can provide answers about the flow of information in the network between different layers, but because it is simply a statistical model of connectivity it cannot address questions at the functional, algorithmic, or computational levels.

Alternatively, there are large-scale models that disregard the architecture of different brain regions entirely, instead treating them as modules of excitatory and inhibitory cells serving a specific function. Eliasmith et al. (2012) present one such model, which models the entire cortical hierarchy and even performs specific tasks. Here the model of V1 is simplified to the point where it is simply a visual encoder, ignoring more complex interactions entirely.

As we have seen, models of this scale suffer of one of two problems. When incorporating a lot of biological detail, they are very useful for understanding the low-level interactions in a network and how information is communicated locally, but are not usually useful for linking low-level processes to the higher-level processes that these processes underlie. Models at the other extreme are relegated to be very high-level approximations of the visual cortex, without the full breadth of complex contextual effects that has been observed in the visual cortex.

Surround modulation models

Perhaps the precisely opposite approach to building in copious amounts of biophysical detail is to focus on a particular phenomenon and attempt to explain it with a reduced set of mechanisms. These models have traditionally been incredibly useful in explaining different aspects of the literature in V1. Specifically, point, rate-based neuron models have been used to explain effects ranging from surround suppression and to more complex contextual-modulation effects such as contour integration and pop-out.

The surround-modulation literature is one of the most complex topics in visual neuroscience, due to the high dependence of the observed phenomena not only on the precise stimulus conditions but also the behavioral state of the animal. As we discovered in Section 2.4, there are two main phenomena that are of interest: surround suppression, usually thought to reduce redundancy in the input, and contextual enhancement of responses through facilitatory effects. These two opposing mechanisms are thought to drive a wide range of the phenomena in the literature including pop-out, contour completion, as well as end stopping, flanker facilitation, and much more.

Various models have been proposed to explain how these contextual influences emerge. In one of the more straightforward models, Li (2002) proposes that through recurrent enhancement of responses, long-range lateral connections can give rise to a visual saliency map in V1, which underlies pop-out, enhancing features that stand out from the visual background. While this model explains how a particular surround modulation effect can emerge from the network interactions between neurons with different tuning properties, it does not account for both suppressive and facilitatory effects and since it is highly abstract cannot shed light on the precise interactions that could give rise to both in a circuit.

Other models such as the Schwabe et al. (2006) model break down the interactions from different sources in more detail demonstrating how long-range patchy horizontal connectivity can mediate both excitatory and inhibitory influences from the surround. Unlike previous models they suggest that long-range feedback connections could be the main driver of these contextual influences and that the drive to inhibitory neurons in the local circuit determines the sign of the effect. However, since this is a static model it provides no account of how the required lateral or feedback connectivity could emerge and since the model only represents 1D space cannot process realistic stimuli. Additionally this model still uses the assumption originally made by Somers et al. (1998), that inhibitory neurons have a higher threshold than excitatory neurons, to account for contrast dependent effects, which has received little experimental confirmation.

An alternative to this circuit-level analysis is to express these interactions in the more abstract framework of predictive coding. The work by Spratling (2010) in particular expresses the cortical circuit as predictive coding, modeling the response of V1 neurons using a competitive mechanism between neurons attempting to best represent the input. Like many others, this model provides a good account of various surround modulation effects including iso- and cross-orientation suppression but exhibits only weak facilitatory effects, suggesting that it does not fully capture the mechanism that is behind the contrast- and context-dependent interactions between suppression and facilitation. Furthermore, it is very difficult to relate the interactions between explicitly predictive or error-coding neurons to the real interactions taking place in the cortex.

Then there are a class of phenomenological models attempting to explain a variety of surround modulation effects based on the geometries found in visual fields and the Gestalt law of good continuation (Field et al., 1993). The Keemink and van Rossum (2015) model, for example, suggests that effects including the tilt illusion and contour detection can be explained by a single mechanism, which computes the elastica curvature energy of the smoothest contour connecting oriented bars. This suggests that the visual cortex can somehow compute the energy required to connect different contours but does not explain how this association field could emerge. It is therefore a good account of what the cortex is computing, but does not provide a canonical mechanism for how these interactions can emerge independent of the sensory modality that is involved. A general mechanism would allow the model to extract the statistical dependencies between arbitrary features, such that when trained on natural images it automatically extracts the geometric relationships within contours.

Another avenue of study has been the approach by the Schwartz lab, who have for a long time investigated the interaction between natural-image statistics and surround-modulation effects. This is perhaps one of the most interesting approaches, because it treats contextual phenomena not as a by-product of some underlying computational process, but rather supposes that they are driven by the statistics of the visual input and are therefore closely tied to the computational task that is actually being solved. In Coen-Cagli et al. (2015) they describe an extension of a simple normalization framework that computes the homogeneity in an image between center and surround, gating the divisive normalization component accordingly, which directly leads to feature-specific surround modulation effects observed in experiments. In this framework, surround modulation is simply the measurable effect of statistical inference, improving the efficient encoding of the stimulus.

All of these models capture important parts of the phenomena involved, highlighting how network interactions between neurons with different tuning properties can

give rise to complex contextual modulation effects (Li, 2002), where the sources of this input could originate from (Schwabe et al., 2006) and how these interactions might fit into a general computational paradigm that is concerned with prediction and inference (Spratling, 2010; Coen-Cagli et al., 2015). However, all of these models are static, presupposing an existing profile of connectivity, and thus cannot directly address how different cell types could interact to give rise to a circuit, mechanistically explaining how these effects emerge.

2.6.1 *Developmental models*

As it is thought that the cortex captures statistics about the sensory streams to construct internal models of the world, it is clear that function, structure and development are closely linked. Indeed it is now clear that V1 is a dynamic and adaptive circuit, which exhibits learning and learns to accurately encode and represent its inputs. In other words, V1 is not a static circuit, organizing solely based on genetic cues, but rather self-organizes through a complex interplay of genetic triggers, intrinsic mechanisms such as retinal waves and perhaps most importantly, external visual input. A number of models have been developed using these ideas, explaining the development and self-organization of the V1 circuit as fundamentally being driven by activity-dependent and correlation-based processes.

The original principles for what became the class of self-organizing map (SOM) models were first formulated by (von der Malsburg, 1973), who recognized this connection and sought to explain the self-organization of neurons into topographic maps via the simple process of Hebbian learning. The von der Malsburg (1973) model showed for the first time that very simple mechanisms such as Mexican Hat (Section 2.2.1) connectivity can give rise to the smooth organization of individual neurons into a topographic map.

Since then many models have sought to explain the self-organization of individual neurons into an overall functional circuit, which can extract and encode a wide range of features in sensory input. The paradigm has been applied not just to visual neuroscience but also the auditory, somatosensory, and olfactory cortices. There now exists a wide range of literature arguing to what degree the self-organizing processes are driven by the afferent, excitatory and inhibitory components. In particular the development of orientation-selective simple and complex cells has received significant attention.

Applying this correlation-based paradigm, Miller et al. (1989) were able to show that the development of realistic stripy ocular dominance columns could be explained, demonstrating the role of neurons competing to represent specific visual features in

the input. The Somers et al. (1995) model also applied this paradigm, once again to the development of orientation selectivity. They argued for a strong role for recurrent excitations to bind the responses of neurons together, which strengthens the weak orientation selective input from the thalamus, with strong unselective inhibition keeping the circuit in balance and sharpening the orientation tuning of the principal neurons.

The idea that the development of highly complex cortical circuits can be explained using a small set of mechanisms therefore has received a significant amount of attention and has been able to explain the development of not just visual receptive fields but also somatosensory (Wilson et al., 2010), auditory Khan (2011) and other feature maps. The fundamental principles underlying these computations are generally the same, with individual neurons competing to represent an input, leading to a smooth mapping of an underlying feature of the input onto the cortical surface. So while the precise process which allows neurons to compete and the learning rule that allows these maps to robustly, yet adaptively develop has not been conclusively established, a lot is now known about self-organization of the cortex.

In recent years some effort has been put into making the SOM-based models more closely represent the functional organization of the cortex, starting with the LISSOM model (Bednar and Miikkulainen, 2003) and more recently the simpler and more robust GCAL (Stevens et al., 2013b) model. In these models the retina, LGN, and V1 are modeled as a set of neuronal sheets with feedforward and lateral connectivity self-organizing into the complex topographic maps. These models provide a very good match to the known developmental time course and replicate a large range of phenomena observed in the mammalian cortex, yet work with only a small set of mechanisms including contrast-gain control, an adaptive per-neuron threshold, normalized Hebbian learning, and lateral connectivity (Stevens et al., 2013b). Extensions of this model explaining the development of complex cells and of contrast-dependent size tuning (Antolík and Bednar, 2011) as well as a continuous time model matching LGN/V1 temporal response functions have also been developed (Stevens, 2011).

The GCAL (gain control, adaptive, lateral) model is based on several core observations about information-processing in V1 and the cortex in general. As previous sections have shown, the primary visual cortex of primates responds strongly to specific low-level features in its visual input, including orientation, color and direction. Selectivity to these features is conserved across a wide range of contrasts, and neurons form topographic feature maps across the surface of V1 by virtue of self-organization.

While these experimentally confirmed findings are specific to vision, the concept of equipotentiality proposes that different areas in the cerebral cortex are originally cytoarchitectonically highly similar, becoming differentiated based largely on the statistical patterns in their inputs during development and are thus capable of capturing

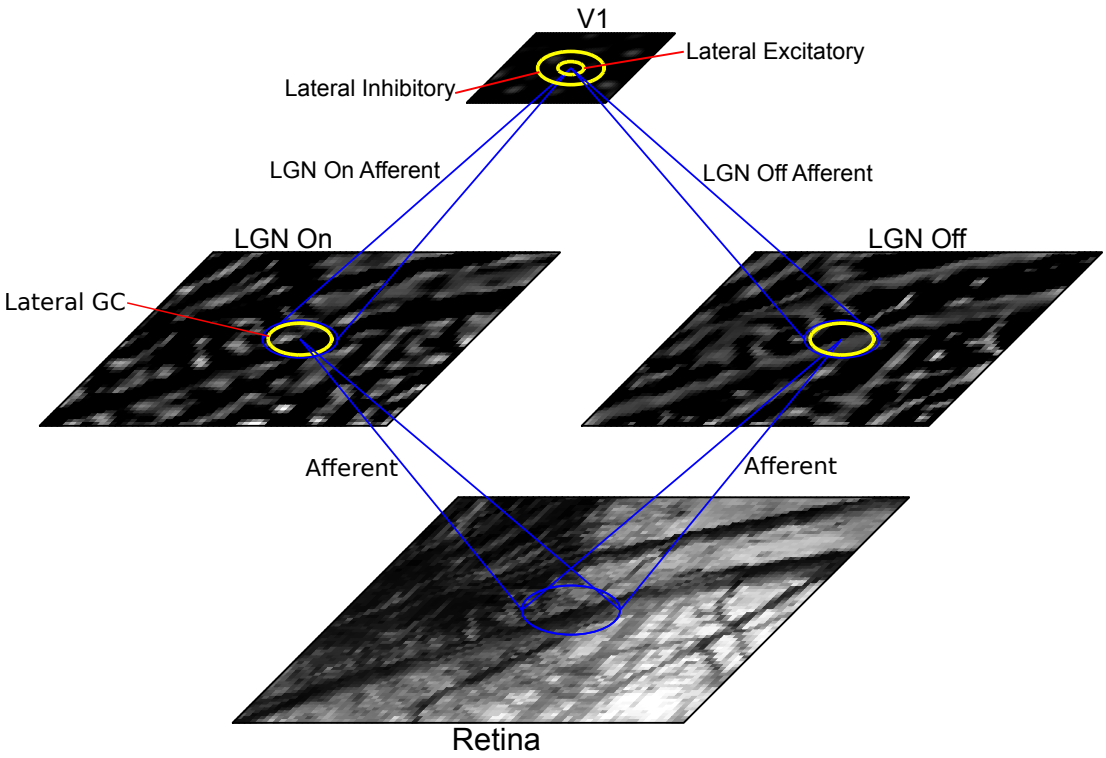


Figure 2.16: Schematic of simplest GCAL model for development of simple cells with surround modulation, retinotopic organization and orientation preference maps. It consists of a retinal sheet, two RGC/LGN sheets for ON and OFF cell responses and one V1 sheet, connected with intra- and inter-areal projections. The sheets are drawn to scale, with larger sheets for the RGC/LGN and retinal layers to avoid edge effects. Projections are illustrated with blue (feedforward connections) and yellow (lateral connections) ovals with cones converging on their target, all drawn to scale to show their spatial extents. RGC/LGN sheets consist of units with hardwired Difference of Gaussian RFs with ON and OFF center-surround regions. LGN Afferent projections to V1 are initially unspecific but develop Gabor-like RF structures through Hebbian learning of visual inputs, as observed experimentally. Reproduced from Stevens et al. (2013b).

any sensory modality. While the validity of this hypothesis remains controversial, experiments such as those by Sur et al. (1990) have at least partly validated this view. In this particular study, experimenters rewired optic nerve axons to interface with neurons in the medial geniculate nucleus (MGN), part of the pre-cortical auditory pathway, and subsequently showed that neurons in the primary auditory cortex (A1) would become receptive to visual features usually associated with V1. This idea is behind much of what has made V1 such a popular model area for neuroscience, and suggests that many of the insights that can be gleaned from the study of V1 could be applied to the cortex as a whole.

The architecture of the GCAL model in its simplest form and as it will initially be used in this project relies on only four sets of neurons organized into 2D grids, which

will be referred to as sheets. These sheets consist of the retinal sheet for the presentation of stimuli, two RGC/LGN sheets representing ON and OFF center receptive fields and a V1 sheet. Neurons in these layers are connected with different intra- and inter-areal connection fields. This simple model can only be used to demonstrate retinotopy, orientation preference and the emergence of simple cell-like RFs, but more complex models have been shown to additionally account for complex cells, ocular dominance, motion direction, spatial frequency, temporal frequency, disparity and color (Bednar, 2012). All these models are trained by presenting a visual input on the retina, allowing the response to propagate through the different sheets and then adjusting the connections weights to V1 neurons based on a local learning rule.

This limited number of mechanisms already gives rise to an incredibly robust model of development of topographic map development and generates different experimentally observed RF profiles. To add further robustness to the LISSOM model and to allow it to respond across a wide range of contrasts, in GCAL contrast-gain control was introduced in the LGN sheets (marked as Lateral GC). This mechanism was closely modeled on the divisive normalization model introduced and validated by Bonin et al. (2005). Finally, the lateral excitatory and inhibitory fields in the V1 sheet give rise not only to the topographic map structure but also to some limited surround-modulation effects.

Indeed there are existing models based on this architecture designed to account for both complex cells and surround modulation Antolík (2010). This model mirrors the approach taken in this thesis in some ways, however modeling complex cells adds significant additional computational overhead and makes it more difficult to reason about the roles of different inhibitory components. Additionally, the model is only coarsely spatially calibrated making it more difficult about the precise extents of individual connections. The major real difference is in the profile of activation of inhibitory neurons, as excitatory neurons activate inhibitory cells in a larger radius than other excitatory cells, while the models here assume that excitatory cells project over the same distance and it is the the inhibitory neurons that target a larger area, mirroring the properties of large basket cells. Finally it is not quite clear to what extent the model exhibits long-range facilitation. However, the model should be seen as complementary to the models that will be introduced in this thesis and may be combined in future.

Overall, due to the simplicity of its mechanisms and its explanatory strength, GCAL provides an ideal starting point to explore the contribution of feedforward and lateral components to the response and function of V1. Most importantly it provides an alternative to static models of the visual cortex, which are unconstrained by the requirement of having to developing into a functional circuit that can actually capture

features of the visual input. Therefore the GCAL model sits at the three-way intersection between anatomical organization, circuit-level and algorithmic implementation, and the actual functional consequences on the response and computation performed by this brain area. However, while existing models have focused on the overall principles of self-organization, they have not been used to study specific interactions between cell classes, how those contribute to development and how the cortical circuit can adapt to the specific computational requirements driven by a task. These topics will be the focus of the remaining chapters.

3

REPRODUCIBLE SCIENCE AND VISUALIZATION

Although the subject of this thesis is centered around modeling of the visual cortex, a reproducible workflow is crucial to all scientific disciplines and especially so in the computational sciences. As part of the work presented in this thesis, we describe a new workflow to make the exploration, analysis and publication of complex models easier and more reproducible at every step. In particular this chapter describes how the HoloViews library developed to aid the work as part of the PhD achieves these goals and provides a general solution to data visualization, storage and analysis that is now used in a number of research projects. Every result in this thesis was generated using this workflow and a fully reproducible record of the work is available at `thesis.philippjfr.com`.

Developing such a workflow was essential to allow quick exploration of the highly non-linear models developed as part of this research. Specifically, when dealing with complex models with a large number of parameters, which also evolve over time, existing tools were not suited for the job. This is because a major component of this work and work in other fields is an iterative process exploring some parameter space, gaining insight into the behavior of the model and then refining the model with the insights gained by careful analysis. For this purpose we present a workflow, starting with the declaration of a parameter space to explore, demonstrating how the simulations can easily be launched locally or on a distributed cluster and then be collected together for analysis. Finally we demonstrate how the HoloViews library lets you load the data as it is generated, monitor the progress, and then provides powerful tools to visualize and analyze the resulting data up to and including publication-quality figures.

We will begin by outlining the overall workflow and the philosophy behind the individual components, before providing a case study on the usage of the workflow as employed in this thesis.

3.1 HOLOVIEWS: BUILDING COMPLEX VISUALIZATIONS EASILY FOR REPRODUCIBLE SCIENCE

One of the major contributions developed as part of this thesis is development of the HoloViews data analysis and visualization library. The library was developed to

support this research in collaboration with Jean-Luc Stevens and was published by Jean-Luc Stevens and myself as joint first authors in the SciPy Conference Proceedings (Jean-Luc R. Stevens et al., 2015). Since then the library has become a popular open-source library used in a wide range of scientific disciplines. This section will outline the core design principles behind HoloViews before applying it to a case study taken from one of the results sections.

3.1.1 *Motivation*

In modern data analysis in scientific and engineering fields there is often significantly more data available than a researcher can easily review using standard plotting tools, in particular when the data is dependent on a large number of variables, which cannot easily be represented by the standard 2D or 3D plot types. The other barrier to exploring these kind of datasets is that most plotting libraries are not very well suited towards interactive analysis. The process of visualizing is usually one that requires a lot of back and forth between writing custom plotting code, viewing the results, and then customizing the plot or writing more plotting code to view the data in a different way. All this back and forth between the code and what the researcher actually wants to do, which is to view the data, can kill productivity and is a major obstacle in analyzing complex datasets. Furthermore, it results in a disjointed, inefficient workflow, accumulating custom code, which has to be maintained and often becomes unreproducible.

HoloViews solves these fundamental problems by making the data immediately visualizable. Unlike standard plotting approaches, the data is always available on HoloViews objects. Thus a plot does not become a dead end; each piece of data is annotated with the appropriate metadata so it can be composed and processed into further plots. Further, HoloViews objects can easily be converted between each other, allowing very quickly iterating over different visualizations. Finally HoloViews makes interactivity central to the workflow—data with higher dimensionality than can be meaningfully represented on a simple plot can be explored through interactive sliders or by laying the data out in complex but highly informative faceted plots.

HoloViews also improves the reproducibility of the results because dependency on external plotting code is massively reduced, the metadata is tightly coupled with the actual data, and HoloViews' declarative style provides a readable description of what exactly is being plotted. Next we will highlight the core design principles of HoloViews that makes it so powerful as a tool for analysis, visualization and publication.

3.1.2 *Design Principles*

The overriding design principle of HoloViews is that the user should not have to write a complex recipe of how they want their data to be displayed. Instead, HoloViews provides a set of declarative data structures, which allow annotating the data with a minimal amount of metadata to display a visual representation automatically and transparently—the data should simply reveal itself.

The second core principle of HoloViews is that there should be a very clear distinction between the details of visualization and actual data and metadata required to describe a dataset. This means that the atomic data objects (or Elements) should be very lightweight wrappers around the actual data structure, along with a small amount of semantic metadata. This allows creating, working with and storing the data independent of the plotting code or any concerns about the visual aesthetics of the plots.

The third design goal is the composability of the Elements. By ensuring that different components can easily be composed together it becomes very easy to build even complex plot types from simple components. HoloViews provides operators and methods to make overlaying and composing plots in layouts and grids very straightforward.

These three guiding philosophies make HoloViews into an extremely powerful tool to explore complex datasets easily while also providing the flexibility to generate publication quality figures. In the next section we will outline how this fits into an overall workflow and demonstrate what these principles mean in practice. For a more in depth review of the philosophy and principles behind the library refer to the HoloViews paper in Appendix A.1.

3.1.3 *Contributions*

Since the HoloViews library was a joined project between Jean-Luc Stevens it is important to delineate our contributions to the library. The project was initially started as a way to explore the temporally evolving models that were central to our research projects. Soon after we came to realize the generality of the approach and one of my major contributions was to push for a general system to represent and explore data of any dimensionality and provide a system not only to visualize the data but also apply basic statistical operations on the data easily such as sampling and aggregating it.

Additionally my major contributions to the project were developing a plotting backend for the bokeh interactive visualization library and data backends for data analysis

libraries such as pandas and seaborn. Over the course of the project my contribution accounted for around two thirds of all commits to the project and extended to every aspect of the software.

3.2 A UNIFIED WORKFLOW FOR THE ANALYSIS OF COMPLEX COMPUTATIONAL MODELS

When working with complex computational models or analyses it is often necessary to launch large scale, parallel parameter explorations, gather the results and apply analyses to the results. This often involves separate scripts to launch the simulations and analysis either locally and on a cluster and separate tools to gather this data up from a shared storage location to visualize the results or apply further analyses. This often results in collections of different scripts and data files, which stitch together a disjointed workflow, providing a serious barrier to reproducibility. As an alternative we developed an integrated workflow based on the LANCET and HOLOVIEWS libraries, which lets you trivially launch parameter explorations either locally or on a cluster, monitor the progress and then lazily load just the required subset of data for further analysis.

Through a concrete example we will explore how to specify each stage in this analysis pipeline, highlighting the ease with which we can launch, analyze and revise a model using this system. For this purpose we will use a parameter exploration that is part of one of the results chapters, skipping over the evaluation of the actual results for the time being. While this is just one example, all the results produced as part of this thesis are made available in the same format to provide a fully reproducible record of the work done as part of this project.

3.2.1 *The Jupyter notebook environment*

In order to generate truly reproducible research, it is not enough simply to make sure that the analysis code is re-runnable. Instead, one should guide someone who is trying to reproduce the results through each step so that they can not only replicate the results but also understand each step and can begin modifying the workflow to take the research further. This distinction separates mere replicability from true reproducibility. In the past achieving this standard required documenting the code and results meticulously in separate files, but even then this workflow usually devolves into chaining script after script until the workflow becomes difficult or impossible to follow for an external researcher. The Jupyter notebook environment, spawned through the IPython project (Pérez and Granger, 2007), promises a solution to this

problem, by providing a platform that allows interleaving code with rich media output and exposition, providing a detailed account of the steps required to generate results and figures in a publication.

The entire workflow presented here is centered around the Jupyter notebook, while not depending on it. It is what allows us to specify an experiment to run, monitor progress, and present the results, all accompanied by detailed exposition describing each stage. Each chapter of this thesis will therefore be accompanied by one or more Jupyter notebooks, which provide a fully reproducible account of how each figure was generated, along with supplemental materials and results.

3.2.2 *Step 1: Specifying a parameter space*

The usual process of launching a parameter space analysis is to write scripts to launch the jobs either locally or on a remote cluster. The resulting files then have to be gathered up appropriately and should preferably be accompanied by a log of the precise environment used to launch the jobs. LANCET makes it trivial to specify jobs to execute and will handle launching these jobs both locally and on a cluster, keeping track of the files that were generated, the environment used and any other extraneous information like version control data. In Stevens et al. (2013a) the fundamental ideas behind LANCET are outlined in detail, so here we will focus on how it integrates with the overall workflow outlined in this chapter.

In this outline of the workflow we will be working with one of the models that will be explored and developed as part of this thesis. To get a better understanding of the model we want to vary multiple parameters, apply measurements and analysis to the model, and then evaluate the results. Defining the parameter space to explore is incredibly straightforward. First we declare constant parameters, in this case just the `area` of the model and the times in the development of the model at which we want to apply measurements to the model. Secondly we declare the varying parameters, in this case we simply combine a `Range` of lateral excitatory strength (`latexc_strength`) with another `Range` of contrasts. The multiplication operator automatically expands this parameter space by taking the Cartesian product of the two varying sets of parameters. Finally we combine the constant and varying parameters into the overall parameter set for the batched experiment:

```
1 times = [1000*i for i in range(21)]
  constants = lancet.Args(area=3.0)
  parameter_space = lancet.Range('latexc_strength', 0, 3.0, 11) * lancet.Range('
      contrast', 10, 100, 10)
  batch_arguments =  constants * parameter_space * lancet.Args(times=times)
```

3.2.3 *Step 2: Specifying analysis*

Having defined the parameter space, the next step is defining the actual model to run any measurements or analysis that are applied to it. In this example we are working with the Topographica neural simulator, which allows training large scale rate-based models on visual input patterns. Here we load the SCAL model, which we will present in the next chapter, and define a measurement of the orientation map with subsequent analysis applied to the measured orientation maps. This analysis will now be performed for each of the parameters defined above, at each of the declared times. This approach makes it incredibly easy to set up deferred execution of any number of measurements and analyses, and even allows setting up complex chains of measurements and analysis by declaring the output of one analysis as input to the next.

The `Collector` object will then apply the measurements at the times defined as part of the parameter space, storing them as small individual files, which can be loaded independently or all together. Another benefit is that the Collector automatically records the precise measurements that were defined, ensuring that the results stay reproducible. Since the results are stored inside HoloViews objects, they are instantly visualizable. All of these features ensure that with very minimal effort by the user all the specifications to set up the experiment are recorded, the results can easily be accessed, and the entire process can easily be replicated from the logs that are output by Lancet.

```
c = Collector()

# Measurement
c.collect(measure_or_pref, frequencies=[1.4, 1.6, 1.8])

# Analysis
c.Pinwheels.V1   = c.analyze(c.ref.OrientationPreference.V1 *\
                   c.ref.OrientationSelectivity.V1, PinwheelAnalysis)
c.FFTAnalysis.V1 = c.analyze(c.ref.OrientationPreference.V1,
    PowerSpectrumAnalysis)
```

3.2.4 *Step 3: Launching jobs*

Having declared the parameter space we want to explore and what exactly we want to measure, Lancet makes it trivial to launch the jobs either locally or on a cluster, along with declarations about the resources that each job should use. In most workflows

this would require distinct scripts for local and cluster execution. Using LANCET you simply switch out the particular launcher that should be used. In this case we declare that the jobs should be run on GridEngine via the `QLauncher` but simply by toggling `QSUB` to False we can switch to local execution.

By providing this flexibility, the analysis and simulations can be debugged locally and then trivially be executed on a cluster to launch large-scale analyses or parameter explorations.

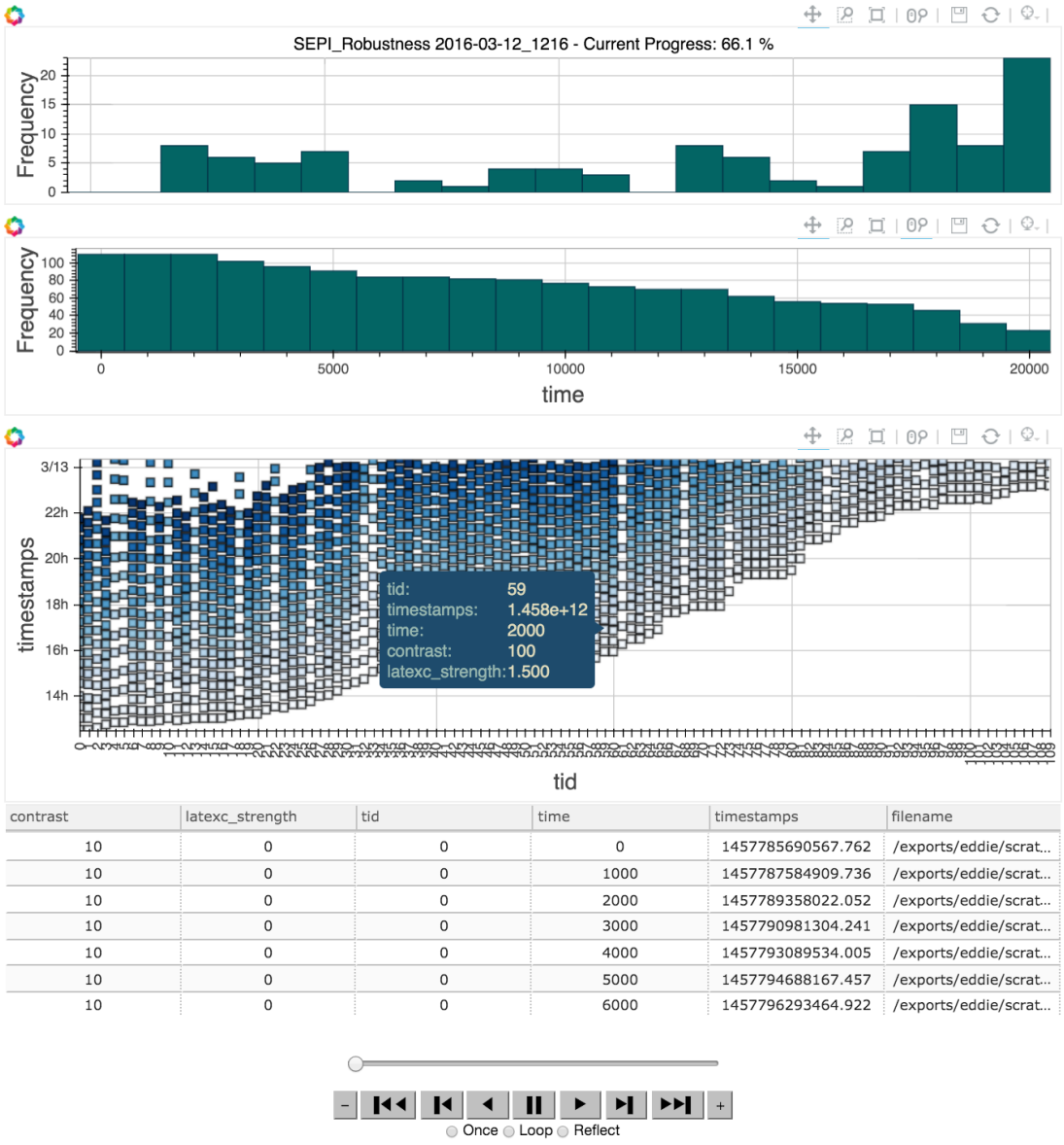
```
QSUB = True

\@lancet.review_and_launch()
def launch():
    qsub_options = (b='y', pe=('sharedmem', '4'))
    launch_options = {'qsub_flag_optionsdict': qsub_options} if QSUB else {}
    runbatch_cmd = RunBatchCommand(ty_file, c, snapshot=False,
                                   metadata=batch_arguments.varying_keys)
    Launcher = lancet.QLauncher if QSUB else lancet.Launcher
    return Launcher(batch_name, batch_arguments, runbatch_cmd,
                    metadata=metadata(), **launch_options))
launch()
```

3.2.5 *Step 4: Monitoring progress*

When launching large numbers of jobs it is important to keep track of their progress to identify any problems and to ensure individual jobs are actually completed. The modular file structure output by LANCET can easily be loaded progressively, and using the live-interaction features of HOLOVIEWS we can monitor the progress in real time. Through various interactive tools it is easy to check on individual jobs and find out which jobs are running slow or have failed, so they can easily be relaunched if needed.

By building a simple dashboard using dynamic features in HOLOVIEWS, we can keep track of the overall progress and each individual job. We also provide tools to relaunch any jobs that fail to complete. These features are crucial to keeping track of large numbers of jobs, and ensure that the workflow is robust to various modes of failure.



contrast	latexc_strength	tid	time	timestamps	filename
10	0	0	0	1457785690567.762	/exports/eddie/scrat...
10	0	0	1000	1457787584909.736	/exports/eddie/scrat...
10	0	0	2000	1457789358022.052	/exports/eddie/scrat...
10	0	0	3000	1457790981304.241	/exports/eddie/scrat...
10	0	0	4000	1457793089534.005	/exports/eddie/scrat...
10	0	0	5000	1457794688167.457	/exports/eddie/scrat...
10	0	0	6000	1457796293464.922	/exports/eddie/scrat...

Figure 3.1: Interactive real time dashboard to monitor the progress of an experiment and catch any problems or failures in individual jobs. Top panel shows current progress in developmental time, second panel shows completion at each time. Third panel plots the times at which each part of a job completed.

dataset	tid	time	timestamps	filename
Gaussian	0	0	2016-03-28T05:20:...	/exports/eddie/sc...
Gaussian	0	1000	2016-03-28T06:26:...	/exports/eddie/sc...
Gaussian	0	2000	2016-03-28T07:26:...	/exports/eddie/sc...
Gaussian	0	3000	2016-03-28T08:27:...	/exports/eddie/sc...
Gaussian	0	4000	2016-03-28T09:27:...	/exports/eddie/sc...
Gaussian	0	5000	2016-03-28T10:27:...	/exports/eddie/sc...
...	...	...	...	...
treeshrew	2	14000	2016-03-28T20:45:...	/exports/eddie/sc...
treeshrew	2	15000	2016-03-28T21:55:...	/exports/eddie/sc...
treeshrew	2	16000	2016-03-28T23:12:...	/exports/eddie/sc...
treeshrew	2	17000	2016-03-29T00:29:...	/exports/eddie/sc...
treeshrew	2	18000	2016-03-29T01:35:...	/exports/eddie/sc...
treeshrew	2	19000	2016-03-29T02:50:...	/exports/eddie/sc...
treeshrew	2	20000	2016-03-29T07:16:...	/exports/eddie/sc...

Figure 3.2: Table showing the output of a parameter exploration using Lancet and HoloViews. Each row represents one file of measurement results. The Table allows selecting just the subset of files to be loaded.

3.2.6 *Step 5: Loading results*

A core part of the workflow is linking up the output generated by Lancet with the data structures provided by HoloViews. In order to make this easy the user can specify a file pattern to search for generated files, which can be loaded into a Table. The Table can then be expanded into another HoloViews data structure called a Layout, which makes it possible to access individual measurements or analyses without having to load the whole file. This means that with just a single line of code we can specify precisely which measurement and which part of the parameter space to load, then visualize it immediately using interactive widgets and complex plot arrangements.

Each measurement or analysis can be accessed via a tree-like structure, letting us access individual measurements by accessing them with attribute access and selecting by the parameter values.

```
:Layout
 .CFs.LateralInhibitory     :Collator   [Index,contrast,latexc_strength,tid,
    time,timestamps]   (filename,entries)
 .CFs.LateralExcitatory     :Collator   [Index,contrast,latexc_strength,tid,
    time,timestamps]   (filename,entries)
 .OrientationPreference.V1  :Collator   [Index,contrast,latexc_strength,tid,
    time,timestamps]   (filename,entries)
 .OrientationSelectivity.V1 :Collator   [Index,contrast,latexc_strength,tid,
    time,timestamps]   (filename,entries)
```

3.2.7 *Step 6: Analysis and Visualization*

The exploratory analysis and visualization of a dataset is one of the most important aspect of a scientist's workflow. However, traditional analysis and visualization libraries split those two aspects apart, which provides a significant bottleneck to what the user actually wants to do, namely to understand their data. By giving data elements their own visual representation, we can provide immediate visual feedback, while still making all the data available for further numerical processing. HoloViews also provides containers and other datastructures that automatically get translated into complex and powerful plot arrangements and widgets, making it trivial to explore datasets of any dimensionality and complexity.

In the previous step we saw how to load a dataset. Here we will quickly demonstrate how we can work with this data once it is loaded. As an example we load some of the weights that represent synaptic neural connections between different neurons in our model. In Figure 3.3 we can see how we can load two sets of weights for a specific set of parameters and compose them together by overlaying them. The resulting visualization demonstrates the complexity this approach allows as we can represent the X/Y positions of the neurons in the V1 sheet on a grid, while representing other dimensions such as the lateral excitatory strength and contrast as sliders. Such a plot would usually be very complex to construct, but here it is just a side-effect of the way the data has been structured and annotated and we could easily transform the dataset to slice and visualize it in any number of ways. Furthermore all the underlying weights are easily accessible and we can therefore feed this data structure straight into further analysis functions.

Combining visualization and analysis in this proves to be an incredibly powerful paradigm to work with large datasets. The main benefits are that the user no longer has to manually keep track of metadata, which means that plots automatically generate appropriate labels and the analysis code always has access to all the information it requires. This also means that instead of storing a plot in an unreproducible format, the full specification of the plot including data, metadata, and styling details can be stored as a single object, which can be loaded later to extract the raw data or merely to reproduce the plot.

3.2.8 *Step 7: Visual aesthetics and publication-quality plotting*

The flexibility of HoloViews not only allows for quick data exploration but even extends to generating publication-quality figures. As stated previously, all plots used as part of this thesis were generated using HoloViews and each chapter is accompa-

```
%%output fig='svg'
lgn_on  = data.CFs.LGNOnAfferent.select(time=20000, contrast=(10, 50), latexc_strength=(0, 1))()
lgn_off = data.CFs.LGNOffAfferent.select(time=20000, contrast=(10, 50), latexc_strength=(0, 1))()
lgn_on * lgn_off
```

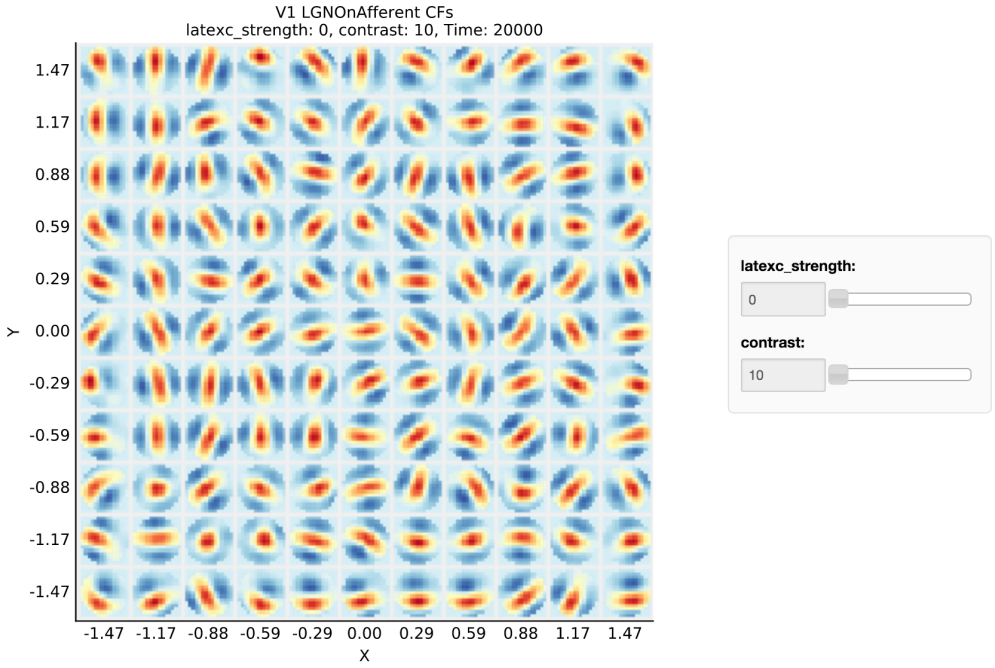


Figure 3.3: Demonstration of loading a measurement from the parameter exploration showing the afferent weights of the model varying across the parameter space we defined for the experiment. Through the use of multi-dimensional datastructures we can achieve complex plotting arrangements and explore any further data dimensions using widgets, which appear automatically.

nied by notebooks, which replicate the entire workflow. This is made possible because HoloViews allows for huge flexibility in customizing the automatically generated visual representations.

3.3 SUMMARY

In this chapter we have demonstrated a complete workflow for the specification, launching, collection, exploration and visualization of a computational experiment developed for this thesis. It provides huge benefits in terms of usability and reproducibility over previous approaches, since it captures the full breadth of an individual research project, keeping a complete record of everything required to reproduce both the data and the actual published figures. It is already in use for research projects ranging from neuroscience (Keemink and van Rossum, 2015) to physics (Nijholt and Akhmerov, 2015; Tenner et al., 2016), astronomy, meteorology, and chemistry.

4

A SPATIALLY CALIBRATED MODEL OF VISUAL CORTEX

One of the major obstacles in modern neuroscience is integrating the vast amount of experimental data that has been generated, highlighting where different sources of evidence is and is not in agreement and offering testable hypothesis to resolve such discrepancies. The primary visual cortex is one of the most well studied areas in the mammalian brain and we have previously seen that it has been extensively explored at varying levels of description, from development, circuits and anatomy to surround modulation, behavioral studies and theoretical models of computation.

In order to provide a better account on how all this information fits together in a generalized model describing the organization and computations performed by the cortex, a unified reference frame regarding the various spatial scales is desperately needed. A careful reading of the literature highlights just how dependent various effects are on the spatial profiles of the underlying neurite projections. Here I will present a developmental model that takes these levels of evidence into account, allowing us to cross-validate known anatomical properties by confirming them against the response properties of the model after development. This will allow bridging between known measurements of anatomy and circuitry and electro-physiological or even behavioral experiments performed on visual cortex.

So far only very few attempts have been made at developing models that take into account the various spatial properties that have been described in the literature ranging from anatomy to electro-physiological measurements. This chapter will demonstrate how existing models of cortical development, specifically the Gain Control Adaptation Lateral model (GCAL) (Stevens et al., 2013b) can be calibrated to match known measurements of spatial extents more closely in a new **S**patially **CAL**ibrated (SCAL) model.

The analysis will focus on various experimental assessments of the spatial properties of the visual pathway and describe how we can use these to build a model that achieves a high-level of consistency with experimental results across a wide range of measures. Specifically, the model will be calibrated with experimental measurements in the parafoveal regions of the visual system of macaque. The macaque has long been a experimental model for in visual neuroscience and the literature surrounding contextual modulation in particular. In doing so this chapter will critically evaluate the literature surrounding spatial tuning properties in the mammalian cortex, high-

lighting discrepancies and specifically assessing various models used to characterize the spatial tuning properties of neurons in the visual pathway.

Since there are wide number of constraints the analyses and results in this chapter represent a highly complex optimization problem. The concrete parameters being optimized are the spatial extents of the various projections to attempt to match experimentally verified extents as closely as possible, while also maintaining a fixed hypercolumn distance and closely matching the electrophysiologically measured response properties of individual V1 neurons.

Once we have collected the data we will provide a full characterization of the spatial response properties, receptive fields and synaptic weights in the model confirming they closely match experimental data. At the same time we will outline in which ways the model falls short and discuss some ways in which these shortcomings might be remedied, which we will pick up in the following chapters. In particular we will highlight problems with the classical GCAL architecture, which makes no distinction between excitatory and inhibitory cell classes and how that relates to the spatial calibration and long-range surround suppression, which is thought to be mediated by disynaptic, long-range excitatory mechanisms.

4.1 METHODS

In this chapter we first introduce the developmental models of the primary visual cortex the more complex models are based on. We begin by outlining the equations and mechanisms underpinning these models and then describe various analyses we can apply to these models to replicate experimental protocols and compare the model against experimental results.

4.1.1 *A spatially calibrated model of cortical development*

The GCAL model put forth in Stevens et al. (2013b) will serve as the starting point for the models in this thesis. As discussed previously (see Section 2.6.1), it provides the first model that develops robust and stable orientation maps independent of visual contrast and for a wide range of training inputs. In this section we describe the equations governing this model, how it is structured and will highlight the modifications that were made to achieve a more consistent spatial calibration.

The major changes introduced in this model are the replacement of subtractive with divisive inhibition and large changes to the spatial profile of connections. One of the major issues we will attempt to settle here is whether the formation of orientation maps requires real long-range inhibitory connections or whether the extents

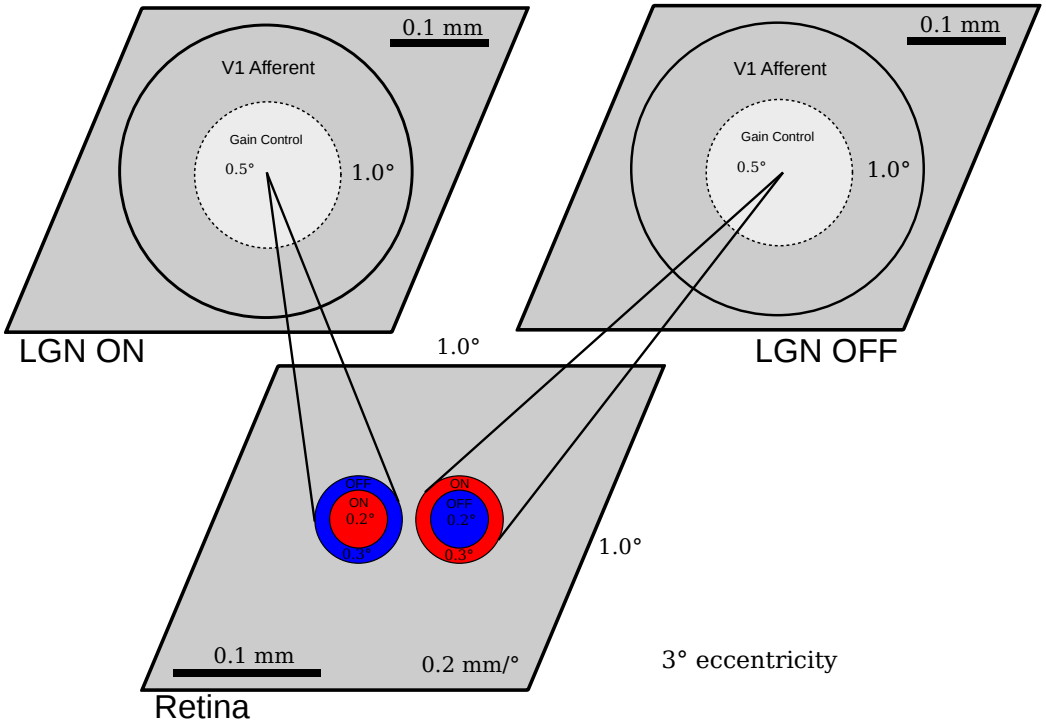


Figure 4.1: Diagram of the retinogeniculate stages of the spatially calibrated (SCAL) model. At the LGN level the excitatory and inhibitory center-surround components are shown in red and blue respectively. In the LGN the absolute size of the divisive gain control connection is indicated by a solid line while the diameter of the Gaussian used to initialize the lateral weights is marked by a dotted line. The given spatial scales hold for an eccentricity of 3°in macaque retina and LGN.

of known inhibitory cell classes is sufficient to account the development of smooth and robust orientation maps. Additionally we add an optional long-range excitatory connection to the model to determine if the model develops in a realistic manner. Before analyzing the behavior of the model we describe its overall architecture and the equations governing its behavior.

Architecture

The architecture of this family of models builds on two main concepts, the idea of 2D sheets of firing-rate point neurons and projections between them, representing the synaptic connections between the neurons. All models we will introduce share the same basic organization at the retinal and lateral geniculate nucleus level, but we will introduce increasingly complex models of the interactions in the primary visual cortex. In Figure 4.1 you can see the organization of the retinal and lateral geniculate nucleus ON and OFF sheets.

The model operates by presenting patterns on the retinal sheet, which then get filtered through difference-of-gaussian connection fields, which give rise to the response the ON and OFF sheets. There a lateral gain control projection applies some pooling normalization to the response. This early stage of processing represents a crude model of retinal ganglion and LGN function and provides the input to the various cortical models introduced here.

The architecture of the retinal ganglion cell and lateral geniculate nucleus layers will remain unchanged for all the models presented in this thesis, while the V1 stage will be progressively refined.

The model diagram for the SCAL V1 stage is shown in Figure 4.2 consisting of a single neural sheet with a comparatively large afferent connection when considered against the GCAL model. The local excitatory and inhibitory kernels however have changed only slightly in their size and finally the model optionally includes a long-range excitatory connection, which represents the long-range patchy lateral connectivity that is so characteristic of layer 2/3 in the visual cortex of higher mammals.

Input patterns

The organization of the self-organizing map (SOM) based developmental models of the cortex, is determined by a complex interplay between the model and the input patterns it is trained on. Just as in the developing brain connections are formed depending on the statistics of the sensory experience of the strongly influencing the spatial organization of the model. In order to accurately assess how the models respond to and self-organize in different visual environments we will use three different visual stimuli to present to the model, however the results in this chapter will be based purely on the simple Gaussian stimuli.

These patterns, used as the baseline for most measurements, are simply elongated Gaussian shapes matching the length of the integrative area of a V1 neuron and with a spatial frequency that would allow three distinct lobes to form within this area. The training patterns are given by:

$$exp(-x^2/(2\sigma_x^2) - y^2/(2\sigma_y^2) \tag{4.1}$$

These simplistic patterns are then presented at random orientations and locations on the retinal layer. Additionally two image datasets will be employed. The first taken from a database of natural images and the other recorded from within the rearing environment of ferrets in a laboratory, which is dominated by the long colinear statistics of the cage bars. These datasets will allow us to explore the effect of

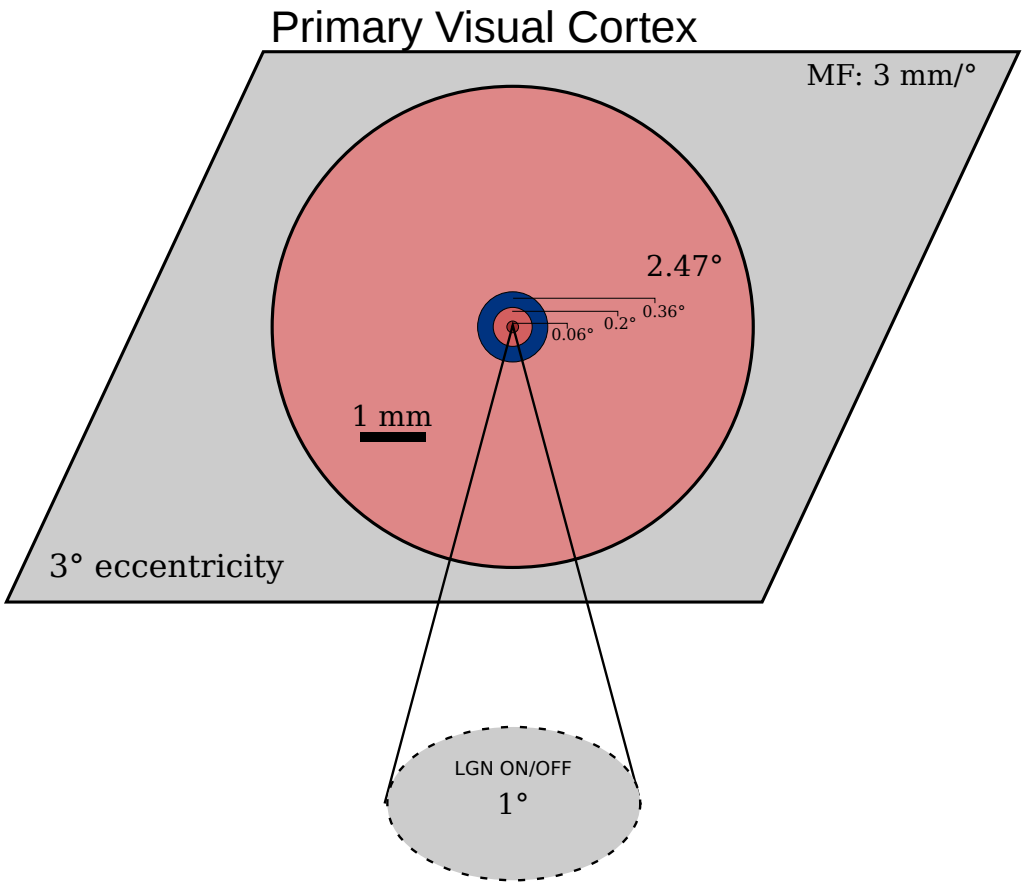


Figure 4.2: Diagram of the SCAL V1 stage of the model showing the spatial scales of the various excitatory (red) and inhibitory (blue) connections. The innermost red circle represents the initial extent of the excitatory kernel, while the next circle represents the full maximum extent of the local excitatory kernel. The blue circle represents the maximum extent of the inhibitory kernel and the large red circle the cut-off extent of the long-range excitatory projection. A fixed cut-off for the kernel is set such that it has no effect on early organization but limits the long-range projections to realistic extents arising due to biophysical constraints.

the natural image statistics on the organization of the models and confirm robustness against a wide range of visual input.

Activation

All the models operate by presenting a new retinal input at each iteration updating the activation of each unit in each sheet. The neurons in the sheets are firing-rate point neurons, with the main state being a floating point activation value. For all models, the activation level η for a unit at position j in an ON/OFF sheet O at time $t + \delta t$ is defined as:

$$\eta_{j,O}(t+\delta t) = f\left(\frac{\gamma_O \sum_{i\in F_{j,P}} \Psi_i(t)\omega_{ij}}{k+\gamma_S \sum_{i\in F_{j,S}} \eta_{i,O}(t)\omega_{ij,S}}\right) \tag{4.2}$$

The constant $\gamma_O = 14.0$ is an arbitrary multiplier for the overall strength of connections from the photoreceptor sheet to the ON/OFF sheets, chosen to give typical activations in the range 0.0 to 1.0, while γ_S is the strength of the feed-forward contrast-gain control. Ψ_i is the activation of unit i in the two-dimensional array of neurons on the photoreceptor sheet from which ON/OFF unit j receives input (its afferent connection field $F_{j,P}$) and $\eta_{i,O}(t)$ is the activation of other ON/OFF units on the previous time step (received over the suppressive connection field $F_{j,S}$). κ is a fixed constant to avoid zero division errors. The activation function f is a half-wave rectifying function that ensures the activation of ON/OFF units is always positive.

The weights ω_{ij} represent the fixed connection weights from photoreceptor i to the ON or OFF unit j defined with a standard difference-of-Gaussians (DoG) kernel. The connection fields for ON units have a positive center and negative surround, and vice versa for OFF units. More precisely, the weight ω_{ij} from an ON-center cell at location (0,0) in the ON sheet and a photoreceptor sheet in location (x, y) on the photoreceptor sheet is given by:

$$\omega_{ij} = \frac{1}{Z_c}\exp\left(-\frac{x^2+y^2}{2\sigma_c^2}\right) - \frac{1}{Z_s}\exp\left(-\frac{x^2+y^2}{2\sigma_s^2}\right) \tag{4.3}$$

The kernel sizes of the central Gaussian σ_c and surround mechanism σ_s are what we will be determining here. Unlike simple DoG kernels, the center-surround are jointly normalized to 1.0 using Z_c and Z_s. The weights for an OFF-center cell are the negative of the ON-center weights (i.e., surround minus center). The center of the connection field of each ON/OFF unit is mapped to the location in the photoreceptor sheet corresponding to the location of that unit in sheet coordinates, making the projection retinotopic.

The weights $\omega_{ij,S}$ in the denominator of equation 4.2 specify the spatial profile of the lateral inhibition received from other ON/OFF units when contrast-gain control is active. The weights of these connections have a fixed, circular Gaussian profile so that for a neuron located at (0,0) in either the ON or OFF sheet:

$$\omega_{ij,S} = \frac{1}{Z_S} \exp\left(-\frac{x^2 + y^2}{2\sigma_S^2}\right) \tag{4.4}$$

where (x, y) is the location of the pre-synaptic neuron, σ_S determines the width of the Gaussian, and Z_S is a normalizing constant that ensures that the total of all the lateral inhibitory weights ω_{ij} to neuron j sum to 1.0. This gain-control projection is activated once per iteration before activity is sent to the V1 sheet.

Except for the various spatial constants and the strengths of the different projections in the LGN layer, the equations are inherited directly from the GCAL model, including the divisive normalization stage. It is only in the V1 layer that we introduce significant changes to the model.

The V1 model

As we saw above in the LGN section the model described here is heavily based on the GCAL model (Stevens et al., 2013b), however it does differ in one major respect—it employs divisive rather than subtractive inhibition. Here we will describe the equations that govern the model and point out where they differ to previous models.

Each V1 neuron in each model receives connections from three different connection types or 'projections' (p), i.e., the afferent projection from the ON/OFF sheets (both channels concatenated into one input vector; $p = A$), the recurrent lateral excitatory projection ($p = E$), and the recurrent lateral inhibitory projection ($p = I$) from other V1 neurons.

The contribution $C_{j,p}$ to the activation of unit j from each projection type ($p = A, E, I$) is calculated as:

$$C_{j,p}(t + \delta t) = \sum_{i \in F_{j,p}} \eta_{i,p}(t) \omega_{ij,p} \tag{4.5}$$

where $\eta_{i,p}$ is the activation of unit i taken from the set of neurons in V1 to which unit j is connected (its connection field F_j) and $w_{ij,p}$ is the connection weight from unit i in V1 to unit j in V1 for the projection p. Afferent activity ($p = A$) remains constant after the first update from the retina, but the other contributions change over 16 settling steps, depending on the activity in V1.

The contributions from all three projections to V1 (afferent (p_A), excitatory p_E and inhibitory p_I) described above are combined using equation 4.6 to calculate the activation of a neuron j in V1 at time t:

$$\eta_{j,V}(t) = f\left(\frac{\sum_{p=\{E,A\}} \gamma_p C_{jp}(t)}{1 + \sum_{p=\{I\}} \gamma_p C_{jp}(t)}\right) \tag{4.6}$$

The projection strength scaling factors γ are defined for each projection type set to provide a balance between excitation and inhibition, and between afferent and lateral influences, to provide robust formation of activity bubbles that allows smooth maps to form. This does not mean that the model is a balanced excitation/inhibition model in the formal sense, indeed the default parameters ensure that inhibition exceeds excitation in the absence of strong afferent input. The function f defines an adaptive threshold point (θ) dependent on the average activity of the unit as described in the next subsection, but in all cases the gain is fixed at unity. Note that unlike GCAL the inhibitory projection acts divisively rather than subtractively.

At the end of the 16 settling steps, the settled V1 activation pattern is deemed to be the V1 response to the presented pattern. Since this is a recurrent model there is no guarantee a steady state is reached within the settling period but through manual inspection it was confirmed this was almost always the case. At this point we use the V1 response to update the threshold point (θ) of V1 neurons (using the adaptation process described below) and to update the afferent and lateral inhibitory weights via Hebbian learning. Unlike the regular GCAL model the V1 activity is not reset to zero instead being allowed to decay until the onset of the next visual input pattern, which had no direct effect but provides a pathway to a more continuous model of visual responses.

In order to assess the development of long-range patchy connections in this model we additionally add a long-range excitatory connection, which we let develop specific synaptic weights but which is given an overall strength of zero so that this projection has no effect on the actual model. In this way we can evaluate how long-range excitatory connectivity would develop within this model without any of the disruptions to development the introduction of long-range correlations might cause. It also provides a unified means of expressing the influence of surround modulation from the surround that will stay consistent between models. This term is modeled as an additional multiplicative component with a constant offset of 1, ensuring only facilitatory modulation from the far surround. The activation of a model neuron is therefore given by:

$$\eta_{exc} = \frac{(\eta_A + \eta_E)(1 + \eta_{LR})}{1 + \eta_I} \tag{4.7}$$

where η_A is the afferent activity, η_E is the local excitatory input, η_I is the divisive inhibitory input and η_{LR} is the long-range excitatory input. These parameters have no absolute scale or units and merely express activity levels of the different projections. In general activity levels among the excitatory population are kept in the range from zero to one.

Adaptation

In order to set the threshold for activation, each neuron unit j in V1 calculates a smoothed exponential average of its settled activity patterns ($\overline{\eta_j}$):

$$\overline{\eta_j}(t) = (1-\beta)\eta_j(t) + \beta\overline{\eta_j}(t-1) \tag{4.8}$$

The smoothing parameter ($\beta = 0.991$) determines the degree of smoothing in the calculation of the average. $\overline{\eta_j}$ is initialized to the target average V1 unit activity (μ), which for all simulations is $\overline{\eta_{jA}}(0) = \mu = 0.024$. The threshold is updated using:

$$\theta(t) = \theta(t-1) + \lambda(\overline{\eta_j}(t) - \mu) \tag{4.9}$$

where $\lambda = 0.01$ is the homeostatic learning rate. The effect of this scaling mechanism is to bring the average activity of each V1 unit closer to the specified target. If the activity in a V1 unit moves away from the target during training, the threshold for activation is thus automatically raised or lowered in order to bring it closer to the target. Note that an alternative rule with only a single smoothing parameter (rather than β and λ) could be formulated, but the rule as presented here makes it simple for the modeler to set a desired target activity μ.

Learning

Initial connection field weights are isotropic 2D Gaussians for the lateral excitatory projection and uniformly random within a Gaussian envelope for afferent and lateral inhibitory projections. Specifically, for a neuron located at (0,0):

$$\omega_{ij} = \frac{1}{Z_p} u \exp\left(-\frac{x^2+y^2}{2\sigma_p^2}\right) \tag{4.10}$$

where (x, y) is the sheet-coordinate location of the pre-synaptic neuron, $u = 1$ for the lateral excitatory projection ($p = E$) and u is a scalar value drawn from a uniform random distribution for the afferent and lateral inhibitory projections ($p = A, I$),

σ_p determines the width of the Gaussian in sheet coordinates, which correspond to visual degrees in the spatially calibrated model. The choice of a uniform random distribution does not have a major effect but avoids introducing spatial biases in the initial model. For a full summary for the parameters of the model see Appendix A.2.

In the model, as images are presented to the photoreceptors, V1 afferent connection weights $\omega_{ij,A}$ from the ON/OFF sheets are adjusted once per iteration (after V1 settling is completed) using a simple Hebbian learning rule. This rule results in connections that reflect correlations between the pre-synaptic ON/OFF unit activities and the post-synaptic V1 response. Hebbian connection weight adjustment at each iteration is dependent on the pre-synaptic activity, the post-synaptic response, and the Hebbian learning rate:

$$\omega_{ij,p}(t) = \frac{\omega_{ij,p}(t-1) + \alpha\eta_j\eta_i}{\sum_k \left(\omega_{kj,p}(t-1) + \alpha\eta_j\eta_k\right)} \tag{4.11}$$

where for unit j, α is the Hebbian learning rate for the afferent connection field F_j. Unless it is constrained, Hebbian learning will lead to ever-increasing (and thus unstable) values of the weights (Rochester et al., 1956). In all the models the weights are constrained using divisive post-synaptic weight normalization (equation 4.11), which is a simple and well understood mechanism. Afferent connection weights from ON and OFF units are normalized together in the model, which has been shown to extract PCA like weights. Symmetry breaking using this learning rule occurs as a result of the initial patterns introducing small biases in the spatial weight distribution, which become reinforced through recurrent interactions over time. We expect that a more biologically motivated homeostatic mechanism for normalization such as multiplicative synaptic scaling (Turrigiano, 1999; Turrigiano and Nelson, 2004; Sullivan and De Sa, 2006) or a sliding threshold for plasticity (Bienenstock et al., 1982) would achieve similar results, but have not tested these.

The learning rates α are defined separately for the afferent, lateral excitatory and lateral inhibitory projections. The density-specific value used in the equation above is then calculated as $\alpha = \frac{\alpha_A}{\tau_A}$, where τ_A is the number of connections per connection field in the afferent projection.

4.1.2 *Measurements*

Measurements in the model are performed by presenting patterns varying by one or more features and computing the maximum response across one or more of those features. This closely matches the protocol usually employed in experiments ranging

from optical imaging to measure orientation maps and electrophysiological measurements to capture individual tuning curves.

Orientation tuning

The orientation map measurements in the model follow the procedure by Blasdel (1992), presenting sine gratings at 12 different orientations and 18 different phases computing computing the vector average across the phases and returning the phase averaged orientation preference and selectivity maps. Since the model does not have direction tuning or complex cells each phase is presented individually rather than as directionally drifting gratings as is usually done in experiments.

Size tuning

The size-tuning measurements are modeled on the measurements by Sceniak et al. (1999) and Sceniak and Hawken (2001) presenting sine grating stimuli of increasing area at the preferred position, orientation and spatial frequency of the central neuron. Due to the uncertainty in experimental measurements we can include all neurons within 0.1° of central neurons, which respond significantly and have an orientation preference varying from the central neuron by no more than $\pi/16$ radians.

The low and high contrast values were chosen to fall into the lower linear region and just below the saturation point of the contrast response curve respectively.

Frequency tuning

The frequency tuning measurements were performed the same way as the size-tuning measurements, but instead of optimizing the spatial frequency of the pattern it was varied. The size of the stimulus was restricted to a 2x2° region in the retina centered around the measured neurons.

4.1.3 *Analyses*

In order to analyze the spatial organization of the developed model we borrow from a variety of techniques used to analyze electrophysiological, optical imaging and anatomical data obtained in vivo allowing us to compare directly between experimental results and our model.

Area summation curves

The area summation curves were measured in the model by presenting the model with disks of sine gratings of increasing size and varying phase and at the optimal

spatial frequency of each neuron. The size-tuning curves obtained in this way were then fitted with the integrated Difference-of-Gaussian model described by the following equation:

$$R(s) = R_0 + K_e \int\int re^{-\frac{r^2}{a}}\, drd\theta - K_i \int\int re^{-\frac{r^2}{b}}\, drd\theta \tag{4.12}$$

where R_0 is the spontaneous response rate, K_e the excitatory gain, K_i the inhibitory gain, a the excitatory space constant and b the inhibitory space constant. By separating the inhibitory and excitatory components in equation 4.12 and define them as R_e and R_i, we can formulate a subtractive and divisive version of this equation:

$$R = R_e - R_i \tag{4.13}$$

$$R = \frac{R_e}{1 + R_i} \tag{4.14}$$

Using the estimated spatial constants and gain parameters we can also calculate the suppression index of each neuron from the parameters of both the DoG models:

$$SI_s = \frac{K_i b}{K_e a} \tag{4.15}$$

$$SI_d = 1 - \frac{1}{1 + K_i b} \tag{4.16}$$

Unlike in the original experiment (Sceniak and Hawken, 2001) we eliminate the R_0 constant offset and β exponent from both models since they represent the baseline activity and spiking threshold respectively, which does not apply to a firing-rate model.

Hypercolumn distance

One of the major features of the development of the visual cortex in higher mammals is the organization of neurons into orderly topographic maps, forming smoothly varying feature preferences across the cortical surface. In particular the LISSOM family of models has been able to develop highly realistic orientation maps. These maps have been studied in detail in various animal species and precise estimates for their spatial scales are known. Kaschube et al. (2010) were able to provide a detailed account of the hypercolumn distances in different species. The hypercolumn distance is defined as the distance on the cortical surface at which the orientation preference repeats a cycle.

The hypercolumn distance is determined by taking the 2D fast Fourier transform of the orientation map and collapse the real component to 1D, allowing us to determine the distance of the peak by fitting a simple kernel to the collapsed FFT and finding the frequency of the strongest component.

Modeling lateral connectivity

Assessing the organization and extent of lateral connectivity in a way that is consistent with experimental measures is difficult as they variously use the maximum extent, which is highly dependent on the process used for tracing the axonal projections. The only quantitatively rigorous assessment of the extent of long-range patchy connectivity in primary visual cortex was performed by Buzas et al. (2006) in cat.

By injecting pre-synaptic neurons with a tracer, which stains synaptic boutons, they were able to identify the location of each bouton relative to the injected neuron and correlate it with the orientation preference map at pre- and post-synaptic locations. This allowed them to build a model of lateral connectivity with both spatial and orientation dependent components, expressed as the combination of von Mises and Gaussian distributions respectively. The orientation dependent component is expressed as the von Mises distribution:

$$V(\phi, \kappa, \mu) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa cos2(\phi-\mu)} \tag{4.17}$$

where ϕ is the difference in the orientation preference between the pre- and post-synaptic neuron, μ is the orientation preference of the post-synaptic neuron, κ is the concentration parameter and $I_o(\kappa)$ is the modified Bessel function of the first kind of zero order.

The Gaussian component on the other hand is a simple 2D Gaussian function, where x and y are the cortical coordinates and $\sigma_{x/y}$ to control the standard deviation and aspect ratio of the Gaussian:

$$G(x, y, \sigma_x, \sigma_y) = e^(- x^2/(2\sigma_x^2) - y^2/(2\sigma_y^2) \tag{4.18}$$

These two components can be combined into a single spatially weighted von Mises distribution in the orientation map by simply multiplying the components:

$$D_1(x, y, \phi) = s_1[G_{11}(x, y, \sigma_{11})V_1(\phi, \kappa_1, \mu_1)] \tag{4.19}$$

where s_1 represents the weighting factor of the combined Gaussian and von Mises functions representing long-range connectivity.

In order to accurately estimate the local isotropic kernel an additional Gaussian component is added, such that the full model is described by:

$$D_2(x, y, \phi) = s_1[G_{11}(x, y, \sigma_{11})V_1(\phi, \kappa_1, \mu_1) + G_{22}(x, y, \sigma_{22})] \tag{4.20}$$

The weights generated by the full model are then divisively normalized to sum to one just like the actual weights. This means that no weighting factor for the G_{22} component is needed because the strength of the first and second component are scaled relative to each other by the s_1 scaling factor.

Based on this model we can obtain estimates of the spatial extent of the local isotropic component and separately measures of the orientation dependence and size of the long-range lateral connectivity kernel. We fitted this model to the long-range lateral connections in the model using the optimization routines built into the SciPy Python library, reporting the mean square error (MSE) and explained variance (R^2) of the final fit when compared to a naive model ignoring the orientation dependent component.

We further extend this model with an additional variable controlling the aspect ratio of the long-range lateral Gaussian field relative to the axis of preferred orientation of the pre-synaptic neuron. This additional component allows us to capture the effect of highly co-linear training patterns on the model connectivity.

Receptive field fitting

The receptive fields of neurons in primary visual cortex are often described in terms of simple Gabor functions characterized by the spatial position, phase, frequency and orientation (Jones and Palmer, 1987; Ringach, 2002). The literature also confirms that this simple function provides a good approximation for the spatial properties of simple cell receptive fields in the visual cortex.

The Gabor function in the real domain is defined as:

$$g(x, y, \lambda, \theta, \phi, \sigma, \gamma) = exp(-\frac{x'^2 + \gamma y'^2}{2\sigma^2})cos(2\pi\frac{x'}{\lambda} + \phi) \tag{4.21}$$

where

$$x' = x\cos(\theta) + y\sin(\theta) \tag{4.22}$$

and

$$y' = -x\sin(\theta) + y\cos(\theta) \tag{4.23}$$

The spatial properties of each receptive field can be summarized by the width of receptive field relative to frequency of the periodic grating (n_x) and the elongation of the Gabor subregions relative to the period (n_y). The distribution of these values, at least in experiment, approximately follows a 1D curve indicating a shared principle in receptive field organization that is conserved across species.

4.2 RESULTS

4.2.1 *Spatially calibrating LGN receptive fields*

The retinal ganglion cells and lateral geniculate nucleus were long treated as simple relay stages, which apply only minimal processing to their inputs. It is now known that is not the case and they are much more actively involved in gating and modulating the incoming information. Since the models we are working with are primarily concerned with the response of V1 neurons, we still use the simplified RGC/LGN model used in GCAL (Stevens et al., 2013b). However to achieve a realistic orientation tuning we adjust the size of the feedforward center-surround kernel and the gain-control connection in V1 to match the known spatial constants more closely.

The first step toward spatially calibrating a model of the visual system is to declare how the modeled space relates to the physical space in the retina, LGN and V1. We will, throughout this thesis, define one unit area in the model as one visual degree of arc. This makes conversion very simple and provides a starting point, which we can build on. This conversion factor is of course arbitrary, but the overall aim will be is to set parameter values, which are a) consistent with known values and b) through various measurements and experiments are confirmed to be internally consistent, i.e. if a particular measurement provides inconsistent result it should be clear why it diverges.

Spatial Tuning

Unfortunately the measurements of LGN center and surround components in the literature have a huge degree of variance largely due to different measurement protocols. The latest and likely most reliable measurements come from (Sceniak et al., 2006), who measured the responses in macaque V1 afferent connections, which should match the responses of LGN neurons themselves very closely. In Table 4.1 we summarize population estimates from a number of studies, measured by presenting disk masked sine gratings of varying sizes and fitting the responses with the previously described Difference of Gaussian model.

A set of area-summation curves measured at varying contrast levels in the SCAL model can be seen in Figure 4.3. These curves were then fitted using the subtractive and divisive integrated Difference-of-Gaussian model described in Section 4.1.3. After fitting this model we calculate the mean-squared error (MSE) and the explained variance of the model (R^2) to characterize the goodness of fit and compare how well the two models can capture the data. The goodness fit analysis of the model across contrasts and a sample subtractive and divisive fit are shown in Figure 4.4. Both mod-

Connection	Literature	Ecc. (°)	Layer	$R_{c/s}$
LGN Center	Sceniak et al. (2006)	2-5	parvo	$median = 0.46°$ $mean = 0.5°$
	Levitt et al. (2001)	0-10	parvo	$0.069 \pm 0.076°$
	Spear et al. (1994)	0-10	parvo	$0.087 \pm 0.046°$
	Bonin et al. (2005)	13.9	parvo	$0.6 \pm 0.4°$
				$0.4 \pm 0.2°$
LGN Surround	Sceniak et al. (2006)	2-5	parvo	$median = 0.51°$ (0.15-0.85)
	Levitt et al. (2001)	0-10	parvo	$0.33 \pm 0.076°$
	Spear et al. (1994)	0-10	parvo	$0.53 \pm 0.39°$
	Bonin et al. (2005)	13.9	parvo	$2.0 \pm 1.1°$
				$1.8 \pm 2.6°$

Table 4.1: Estimates of macaque LGN neuron spatial tuning properties fitted using Difference-of-Gaussian models with either subtractive or divisive inhibitory components. Variability arises from differences in eccentricity, layer and stimulus protocol.

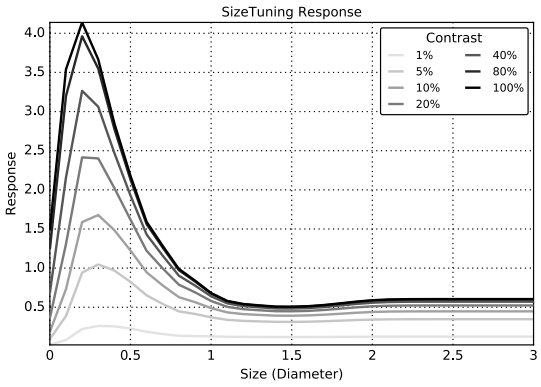


Figure 4.3: LGN neuron size tuning representing the response of a stereotyped LGN neuron to a sinusoidal grating disk stimulus with increasing contrast. The tuning curves demonstrate a small but noticeable shift toward larger size preferences at lower contrasts.

els capture the data well but the subtractive model performs significantly better than the divisive model with mean R^2 values of 0.876 and 0.96 respectively. Therefore all analyses going forward will use the subtractive DoG model.

Since LGN responses in this model do not have any source of variability we cannot compare the distributions directly. To compare the responses with the experimental results we will treat the 10% contrast response as the low contrast response and 70% contrast as the high contrast response. This matches the experiment where the contrast with 20-50% and 70-90% of the maximal response are taken as low- and high-contrast responses respectively. It is worth noting that because there is no real experimental setup and the retinal sheet represents the spiking responses of RGC cells we use an operational definition of contrast. This means the range of contrasts

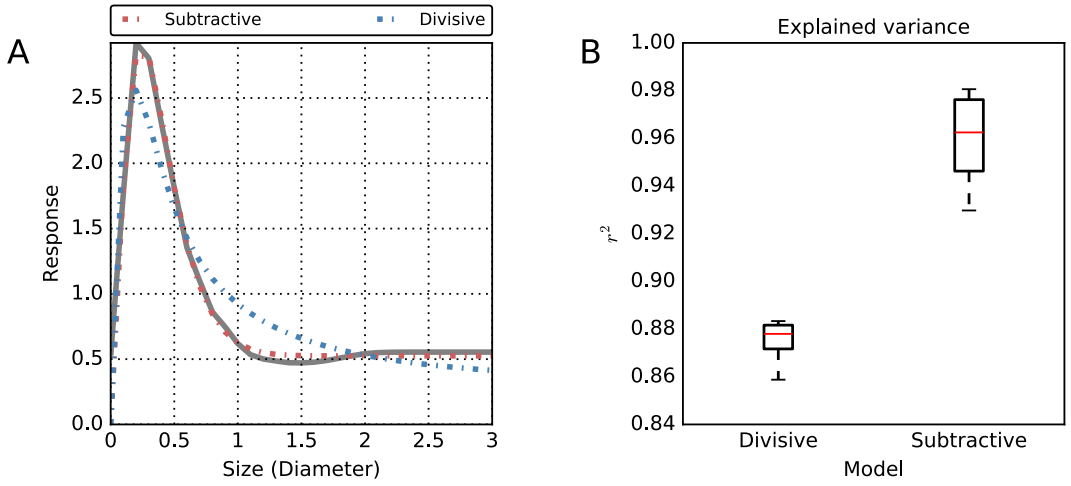


Figure 4.4: Difference-of-Gaussian model goodness of fit comparison for SCAL size-tuning curves. A) LGN area-summation curve (gray) fitted using a subtractive (blue) and divisive (red) integrated DoG model. B) Fraction of explained variance (R^2) of the subtractive and divisive DoG models.

are scaled to match the dynamic range of responses by the network. To compare the data directly to the experimental data, we have overlaid the low and high contrast estimates of the excitatory and inhibitory constant directly on experimental data from Sceniak et al. (2006) (see Figure 4.5).

The excitatory space constant α was estimated at 0.16° at high contrast, expanding to 0.294° at low contrast. The inhibitory or surround space constant on the other hand was 0.62° at low contrast and 0.56° at high contrast. These values fall well in the distributions that have been measured in experiments, even if they are reaching toward the smaller end of what was measured in thalamocortical afferents in the visual cortex.

Frequency Tuning

The frequency tuning of LGN neurons in macaque V1 is strongly correlated with the size tuning, nonetheless we want to confirm the frequency tuning curve matches what is seen in experiment. Therefore we replicate the frequency tuning measurements performed by Levitt et al. (2001). The sinusoidal gratings used for measurement were presented in a 2x2° area, well beyond the size of the surround field. The resulting frequency tuning curves are shown in Figure 4.6, largely displaying contrast invariance.

We compare the spatial frequency preference of the model LGN directly with a scatter plot of spatial frequency preferences against eccentricities, again demonstrat-

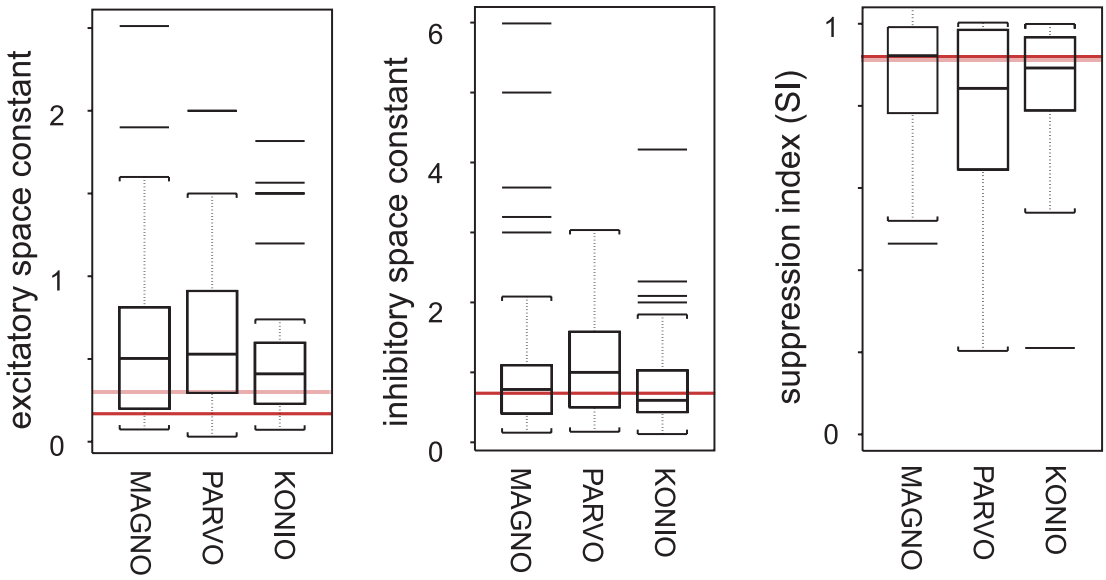


Figure 4.5: Distribution of excitatory and inhibitory space constants in thalamocortical afferents reproduced from Sceniak et al. (2006) (in macaque) and overlaid with estimates obtained from the SCAL LGN model at low contrast (pink) and high contrast (dark red).

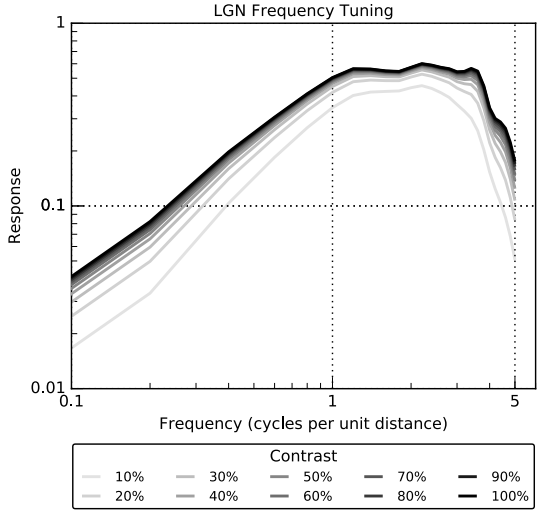


Figure 4.6: LGN neuron frequency tuning representing the response of a stereotyped LGN neuron to a 2x2° sinusoidal grating stimulus at a wide range of contrasts.

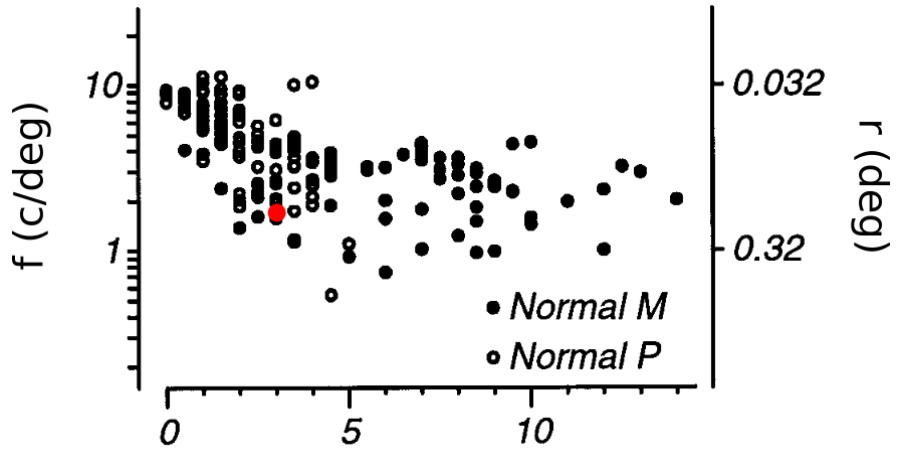


Figure 4.7: Plot of preferred frequency tuning of neurons in the LGN of normally reared and visually deprived macaque monkeys overlaid with the optimal spatial frequency of the stereotyped SCAL LGN neuron (red circle). Preferred spatial frequency is within the range of preferred spatial frequencies and is well adapted to provide appropriate input for the V1 models. Adapted from Levitt et al. (2001).

ing that we are well within the empirically validated distribution of values (see Figure 4.7).

4.2.2 *The V1 model*

Before going into the spatial calibration of the V1 stage of SCAL model, we will first review how the model self-organizes into a smooth, high-quality orientation map. The model was in most cases simulated as a 4°x4° area, which corresponds to around $12mm^2$ of cortical area. After training a wide range of measurements were performed to confirm the model developed correctly and exhibited the same properties as GCAL.

The orientation map, receptive fields and orientation tuning curve for the fully developed model are shown in Figure 4.8. These results demonstrate that the final spatially calibrated model exhibit high quality orientation maps, contrast independent orientation tuning and very clear oriented lobes in individual receptive fields. Comparing the robustness of map organization to the GCAL model will be performed later on, when we are looking at the properties that make a model robust to a wide range of inputs.

Additionally, to get a better understanding of how the model processes visual inputs in a single presentation and over development, we have plotted the response of the neurons to a simple Gaussian pattern drawn on the retina before and after development in 4.9. The figure highlights how the model initially responds with a diffuse

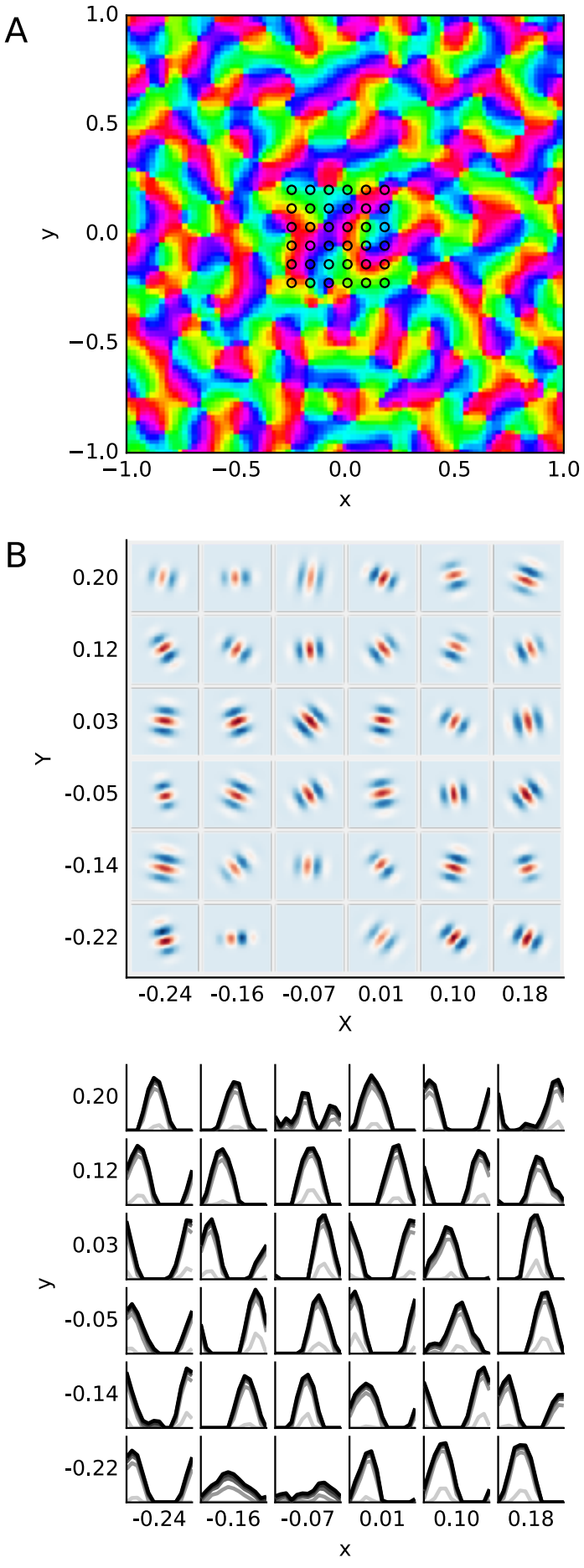


Figure 4.8: SCAL model orientation tuning properties after presenting 20,000 oriented Gaussian patterns. A) Orientation map measured by presenting sine gratings at the optimal spatial frequency to the model, along with the locations of the receptive fields shown in B. B) Gabor fits to receptive fields measured using sparse random noise. C) Orientation tuning curve of a single neuron across contrasts, demonstrating largely contrast-independent orientation tuning. The model exhibits some diversity in tuning properties with different tuning bandwidths and even very broadly tuned neurons.

activity pattern driven, while forming so called activity bubbles corresponding to individual orientation columns after development.

4.2.3 *V1 Spatial Calibration*

A neuron in primary visual cortex receives input from a variety of sources, including feedforward connections from the LGN, horizontal connections from within V1 and feedback connections from extrastriate cortex as seen in Figure 2.6. Translating the fully complexity of the known spatial profiles of these different synaptic inputs to the visual cortex therefore requires integrating a wide range of information from different sources including anatomical, electrophysiological and optical imaging. Not only will this let us build a model that corresponds more closely to the macaque visual cortex in vivo, but also let's us cross-validate the experimental data, highlighting potential discrepancies.

The first step to bridge measurements from different sources will be to establish a well defined mapping from visual space to cortical space. From there we will evaluate the different sources of measurement, replicating the measurements in the model and comparing the results. Finally we will summarize the model parameters that were chosen and will be carried over to later models.

Magnification factors & Hypercolumn Distance

As we outlined above most studies of V1 particularly in the surround modulation literature focus on parafoveal regions between $2-5°$ in eccentricity. Therefore we have chosen to model the region around $3°$ in eccentricity. This already gives us a number of constraints, first of all giving us an approximate V1 magnification factor of 3 mm/°as described by Van Essen et al. (1984) and shown in Table 4.2. This means that each degree of visual space corresponds to 3 mm on the cortical surface at this particular eccentricity. However without an independent measure of the cortical space this conversion factor is still entirely arbitrary. Therefore we make use of the fact that the hypercolumn distance between orientation columns in V1 is well defined by numerous studies.

To give actual scale to our model, we can therefore measure the orientation map hypercolumn distance. Using estimates provided by the Wolf group the hypercolumn distance in macaque V1 has been estimated at roughly $(710 \pm 50\mu m)$. By combining this information with the magnification factor, we can establish that we would expect roughly 4.2 hypercolumns per visual degree $(3mm/° \cdot 710\mu m)$. We will also define an acceptable range of hypercolumn cycles per degree to ensure later models do not diverge too far from the spatial tuning implemented here. Taking the confidence

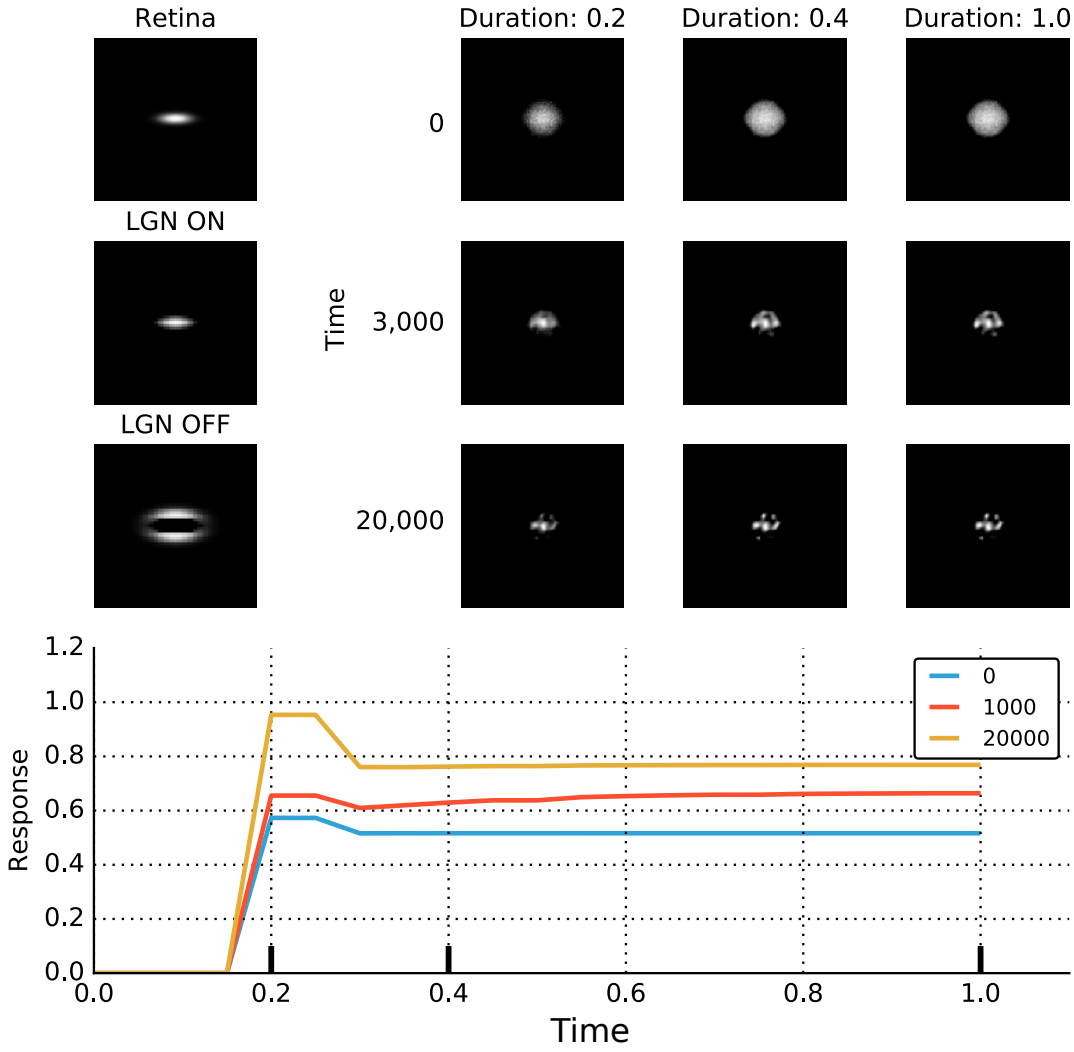


Figure 4.9: SCAL Responses to a simple Gaussian pattern before and after self-organization. Top left shows the input pattern and the LGN ON/OFF response to the pattern. Top right shows the evolving patterns of activity before and after development. The response of the central neuron at various stages of development is shown at the bottom. Demonstrates initial randomness of connections, the process of settling, and finally a selective response reflecting the process of self-organization.

Visual Area	Magnification Factor (mm/°)	Anisotropy Index
Retina[1]	0.223	-
LGN[2]	0.324	1.0-2.0
V1[3]	2.54-3.545	1.0-3.0

Table 4.2: Magnification Factors and Anisotropy Index associated with different visual areas at 3° eccentricity estimated from areal and linear magnification factor equations. Footnotes: [1] - Perry and Cowey (1985), [2] - Connolly and Van Essen (1984), [3] - Van Essen et al. (1984)

intervals for both the magnification factor and hypercolumn distance into account the acceptable range of hypercolumns per sheet coordinate is between 3.29 and 5.3.

The hypercolumn distance in the model was calculated by taking the 2D Fourier transform of the orientation map, reducing it to one dimension and applying a least-squares fit of a Gaussian curve with additional linear and quadratic terms (see Kaschube et al. (2010) for more details). A sample fit to an SCAL orientation map can be seen in Figure 4.10. The final model has 3.86 hypercolumns per visual degree, which falls well within the allowed range. Models trained on natural images not shown here, generally fall in the range of 4-4.5 hypercolumns per visual degree.

Based on that measurement we can now explicitly state that 1 visual degree corresponds to 3 mm in the model, which will let us independently confirm that the extent of the final model weights are consistent with experimental measurements but also internally consistent with the size-tuning response of the model. The next steps will be to evaluate feedforward and lateral connections independently.

Feedforward

The afferent input from the lateral geniculate nucleus is the main driver of responses in the primary visual cortex. It is therefore crucial in determining the size-tuning profile of the visual cortex. In section 2.1.3 we discussed how the integrative area of a neuron in V1 is made up of the combined integrative area of thalamocortical projections and retinogeniculate projections.

Area summation The first step in the fitting procedure was to repeat the protocols applied to the LGN, i.e. measuring area summation curves and fitting DoG models to the results. Using this approach we obtained a large number of size estimates for the excitatory and inhibitory kernels contributing to the V1 response, which we can compare to the results obtained from comparable studies in macaque visual cortex.

The Table 4.3 summarizes the mean results obtained by these studies, with a general agreement of an afferent integration field that spans about 1° in diameter, with inhibitory surround that is on average about two times larger at about 2°. The most thorough analysis being performed by Cavanaugh et al. (2002b), who fitted a number of variants of the model. In addition to measuring integrative area of V1 neurons in visual space they converted the results into cortical space using the known magnification factor. In Figure 4.11 the results from the model, also converted to cortical space are compared to these distributions, showing less diversity but general agreement on mean and median values.

Additionally we also compare the shift in size-tuning preference between low and high contrast conditions, between experiment (Sceniak et al., 1999) and both SCAL

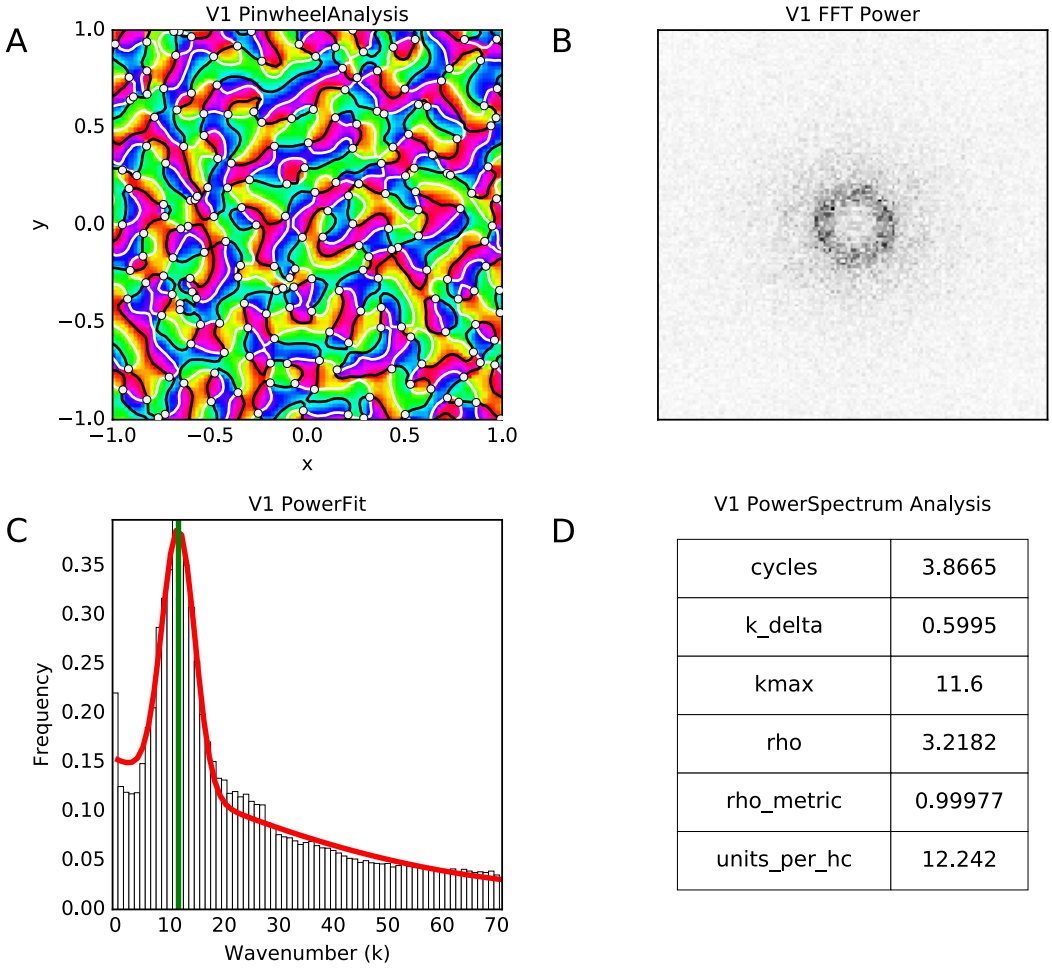


Figure 4.10: Hypercolumn and pinwheel density fitting procedure. A) Orientation map in V1 overlaid with real and imaginary contours and pinwheels at their intersections. B) 2D FFT of the orientation map showing a ring identifying the periodicity of the map. C) 1D histogram of the FFT along with Gaussian fit marking the best fit hypercolumn distance. D) Summary table showing various parameters of the fit, including the hypercolumn `cycles`, along with pinwheel density (`rho`) and the pinwheel density metric (`rho_metric`) which classifies the quality of the map. The SCAL model exhibits near π pinwheel density, a measure that has been associated with realistic orientation maps (Kaschube et al., 2010; Stevens et al., 2013b).

Measurement	Literature	Layer	Diameter
V1 hsRF	Levitt and Lund (2002)	2-6	$1.0 \pm 0.1°$ (0.15 - 1.1)
V1 Excitatory DoG fit	Levitt and Lund (2002)	2-6	0.9°
	Sceniak and Hawken (2001)	2-6	2.0°
	Cavanaugh et al. (2002b)	2-6	1.4°
	Solomon et al. (2002)	not stated	0.94°
V1 Inhibitory DoG fit	Levitt and Lund (2002)	2-6	1.9°
	Sceniak and Hawken (2001)	2-6	4.4°
	Cavanaugh et al. (2002b)	2-6	2.7°
	Solomon et al. (2002)	not stated	2.97°

Table 4.3: Functional estimates of V1 receptive field size using Difference-of-Gaussian models. Measurements differ due to measurement protocols used. All measurements come from macaque V1.

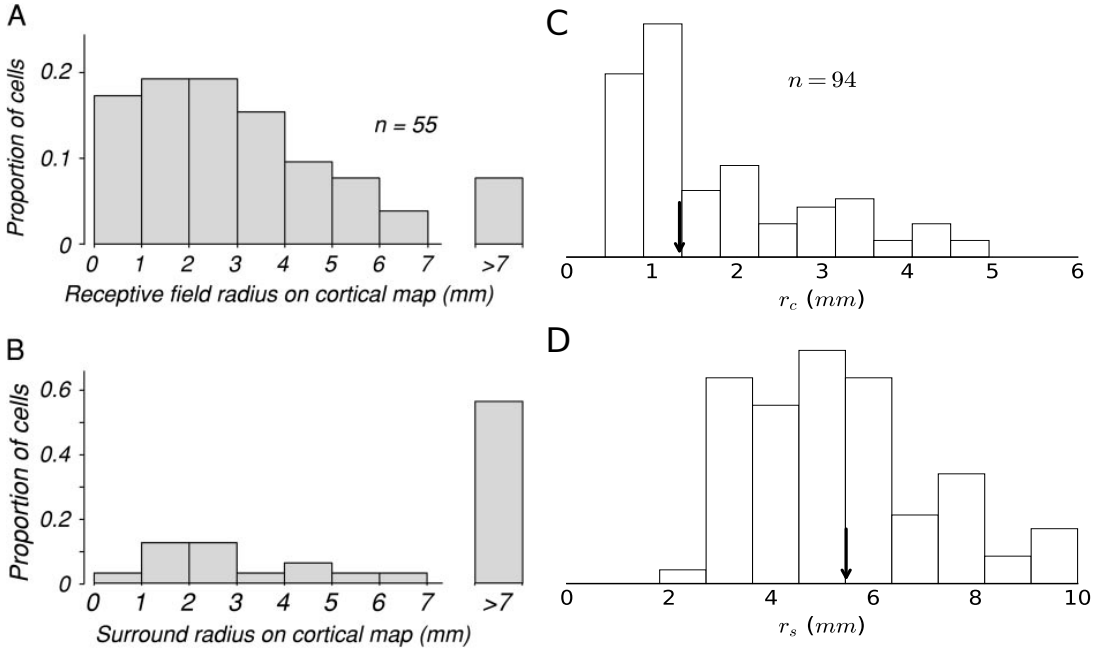


Figure 4.11: Comparison between the distributions of space constants estimated by Cavanaugh et al. (2002b), converted from visual angle to cortical space and the equivalent measurements applied to the model. Distributions of maximal integration area radii for the from experiment in A) and the model in B). Distributions of suppressive surround area radii from experiment in A) and the model in D).

model and GCAL model, which employ divisive and subtractive inhibition respectively. The results shown in 4.12 highlight that the GCAL model does not really exhibit any meaningful size-tuning shift between low and high contrast. Additionally both model scatter plots clearly highlight the lack of diversity in the model. Generally however, the results are qualitatively very similar between SCAL and the experimental data, showing a mean shift and distribution of shifts that is very similar (2.2x shift in experiment vs. 1.8x in the model).

The difference in the contrast dependent size tuning shifts between GCAL and SCAL may be partially explained by the difference in the LGN normalization pool size, as we can already see a small amount of shift in Figure 4.3, which is not evident in the GCAL model. This may allow the additive term k in the normalization pool (see Equation 4.2) to exert more influence under low contrast conditions. Additionally as indicated by Cavanaugh et al. (2002b) and (Solomon et al., 2006), a divisive contrast-gain control model can capture contrast-dependent changes in the shape of the size tuning response function simply by shifting the balance of the excitatory and inhibitory components. Therefore both the LGN and V1 divisive components likely contribute to the increase in size tuning shifts observed in the SCAL model compared to the GCAL model.

Finally we investigate the distribution of suppression indices, which shows a similar mean value to experiments ($SI = 0.65$ in the model vs. $SI = 0.8$ in experiment (Sceniak et al., 2006)) but also highlights the complete lack of unsuppressed cells, which may be explained by the fact that only excitatory neurons were modeled and therefore there is considerably less diversity in connection profiles.

Receptive Fields In addition to measuring the area summation responses of V1 neurons we can also assess the spatial profile of V1 receptive fields by mapping them using reverse correlation techniques. After mapping the receptive fields using 5000 sparse, random stimuli the results were fit using the simple Gabor model described in Section 4.1.3. In addition to once again confirming the spatial frequency preference of the neurons, we compared the $n_x = \sigma_x f$ and $n_y = \sigma_y f$ ratios of the neurons to equivalent measurements in macaque (Ringach, 2002) and cat (Jones and Palmer, 1987). These values can be thought of as representing the number of sinusoidal cycles of the Gabor function fitting in a segment of length σ_x and σ_y respectively, allowing us to compare the number and elongation of Gabor subfields in a way that is independent of phase, translation or orientation.

The results show a similar general shape but also demonstrate a shift toward greater n_x values as compared to the experimental data.

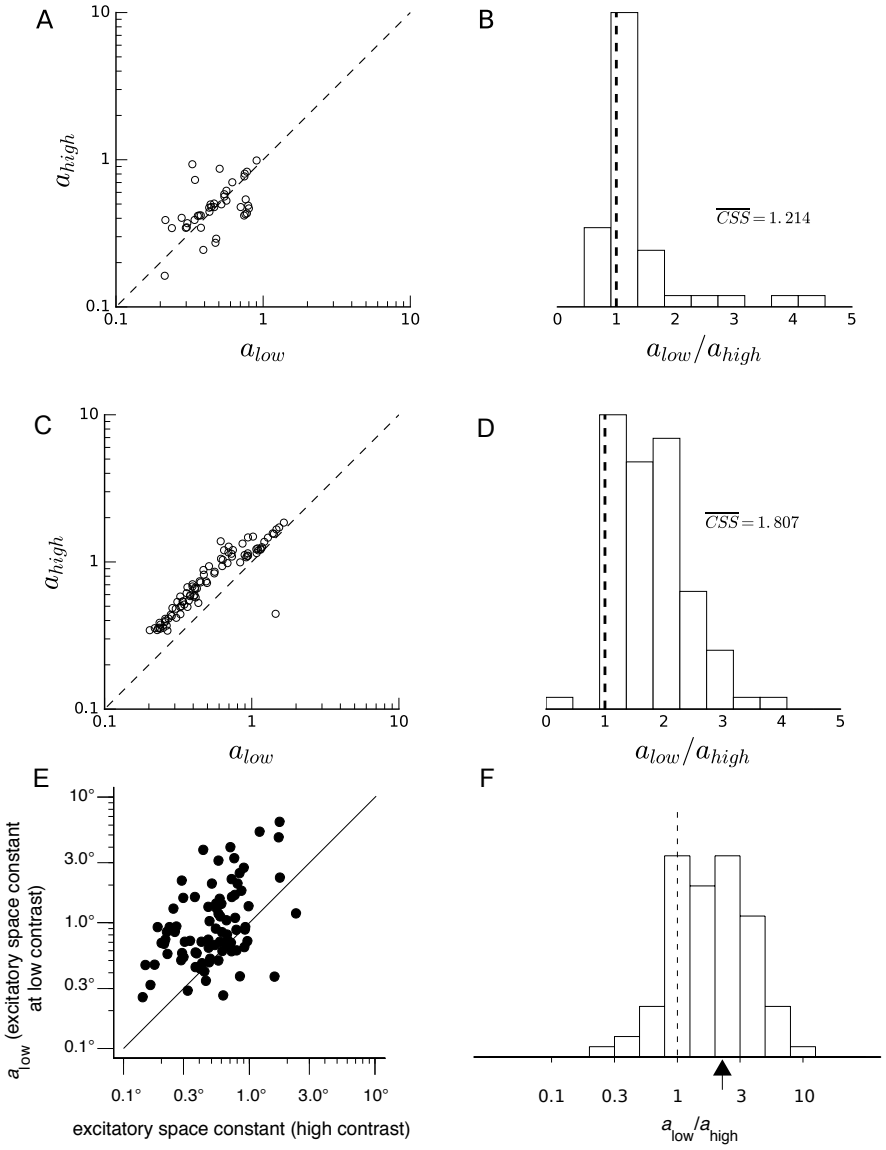


Figure 4.12: Comparison between contrast-dependent size-tuning shifts between the GCAL model (A, B), the SCAL model (C, D) and experimental results from Sceniak et al. (1999) (E, F). A, C, E) Scatter plot of DoG excitatory constant at low and high contrast in experiment and the model respectively. B, D, F) Distribution of contrast-dependent shift in excitatory constant. The introduction of divisive gain control results in a significant contrast-dependent size-tuning shift.

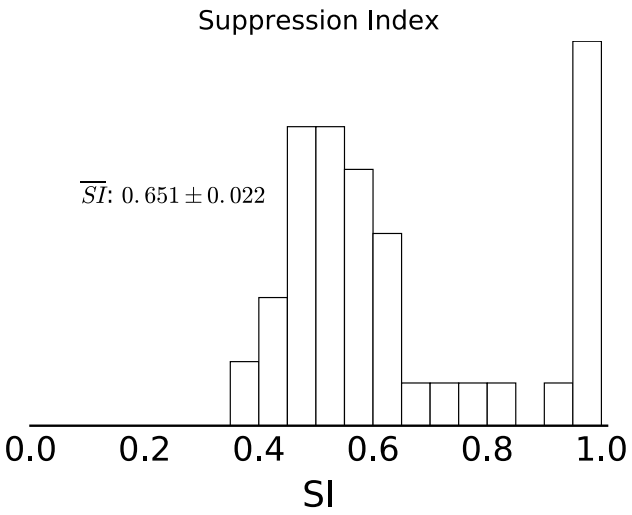


Figure 4.13: Distribution of Suppression Index (SI) values under high-contrast stimulus condition.

Intracortical connectivity

Although they make up a small fraction of connections in V1 thalamocortical afferents are a major driver of activity in primary visual cortex (Binzegger et al., 2004). However intracortical, recurrent connections between neurons also provide significant modulatory influences on V1 responses. Precisely breaking down the contributions of specific connections, particularly because experimental data is much more scarce when taking into account the different cell types. In Section 2.2.4 we summarized the potential role of different inhibitory subclasses, however beyond some estimates of the maximum extent of the axonal projections from different cell classes (Kisvárday et al., 1993, 1997; Budd and Kisvárday, 2001; Buzás et al., 2001), very little is known about their spatial profile. The extent of excitatory projections between V1 neurons are better studied, with a range of estimates for the maximum extents (Angelucci et al., 2002b) and even some models of the spatial profile of connections (Buzas et al., 2006) being available.

We summarize various anatomical estimates of the extents and spatial distribution of intracortical connections in Table 4.4. Note that due to the lack of data a lot of these estimates were obtained in cat V1, which means we have to extrapolate these results to macaque V1.

Excitatory Connections The literature has had a much harder time of picking apart the contribution of intracortical and particularly the patchy lateral connectivity found in V1 so to confirm that these connections have developed as expected is to compare it

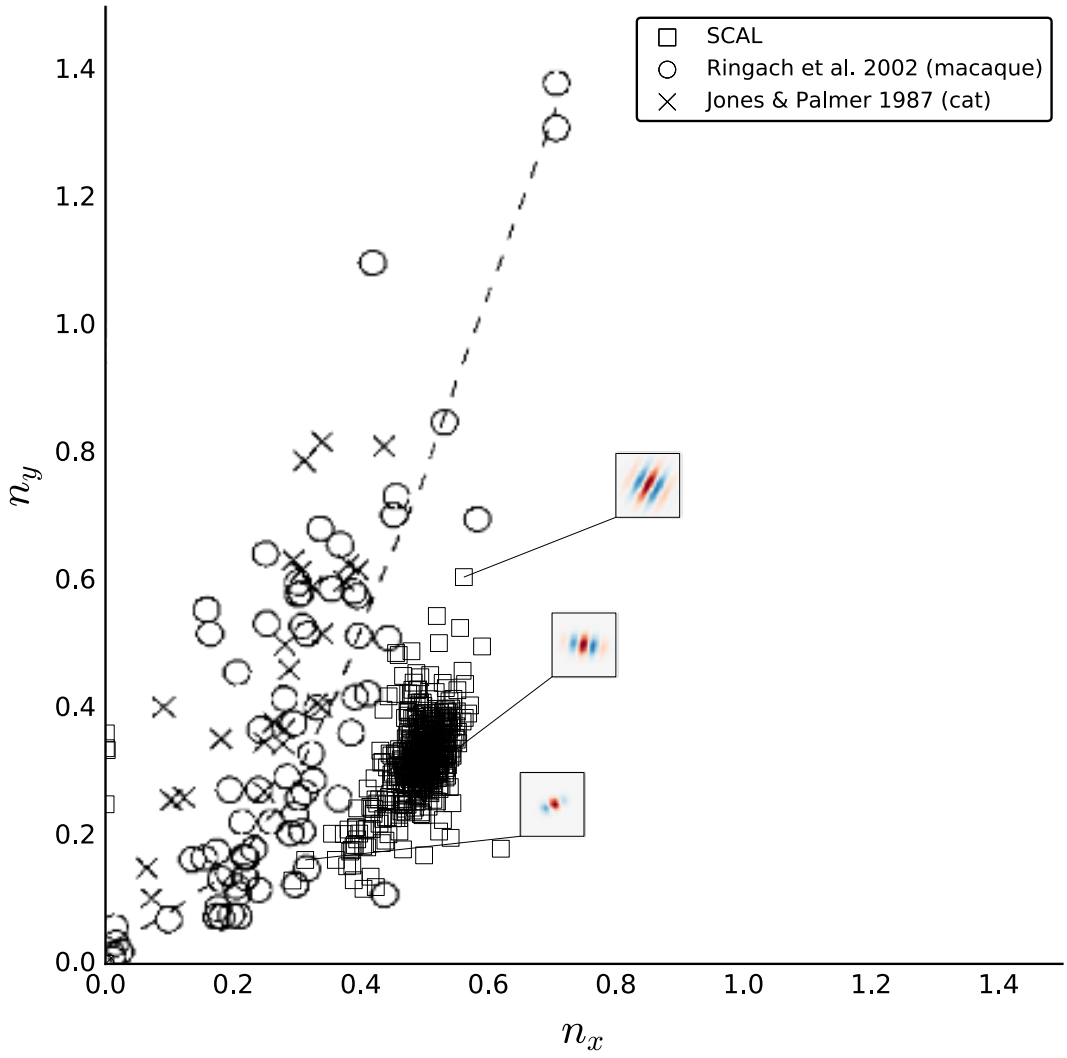


Figure 4.14: Scatter plot of receptive field n_x and n_y values, representing the width and length of the receptive field relative to the period of the grating underlying the Gabor fit respectively. Comparison between results from Ringach (2002) in macaque V1 (circles), Jones and Palmer (1987) in cat V1 (crosses) and in the SCAL model (squares).

Connection	Literature	Species	Layer	Diameter
LGN-V1 Afferents	Angelucci et al. (2002a)	macaque	4Cα	$0.8-1.6°$
	Angelucci and Sainsbury (2006)	macaque	4A/4Cβ	$0.91 \pm 0.041°$
V1 local excitation	Buzas et al. (2006)	cat	2-4 single cell	288µm
	Buzas et al. (2006)	cat	2-4 population	520µm
V1 basket cells	Buzás et al. (2001)	cat	2-6	$0.7-1.9°$
	Buzás et al. (2001)	cat	2-6	$0.76-2.6$mm
V1 long-range excitation	Angelucci et al. (2002b)	macaque	2/3	6 ± 0.7mm (3-9)
			4B/4Cα	6.7 ± 0.7mm (4.7-10)
			population	$2.47 \pm 0.3°$
	Buzas et al. (2006)	cat	2/3	6mm

Table 4.4: Anatomical estimates of the spatial profiles of V1 connectivity from macaque and cat.

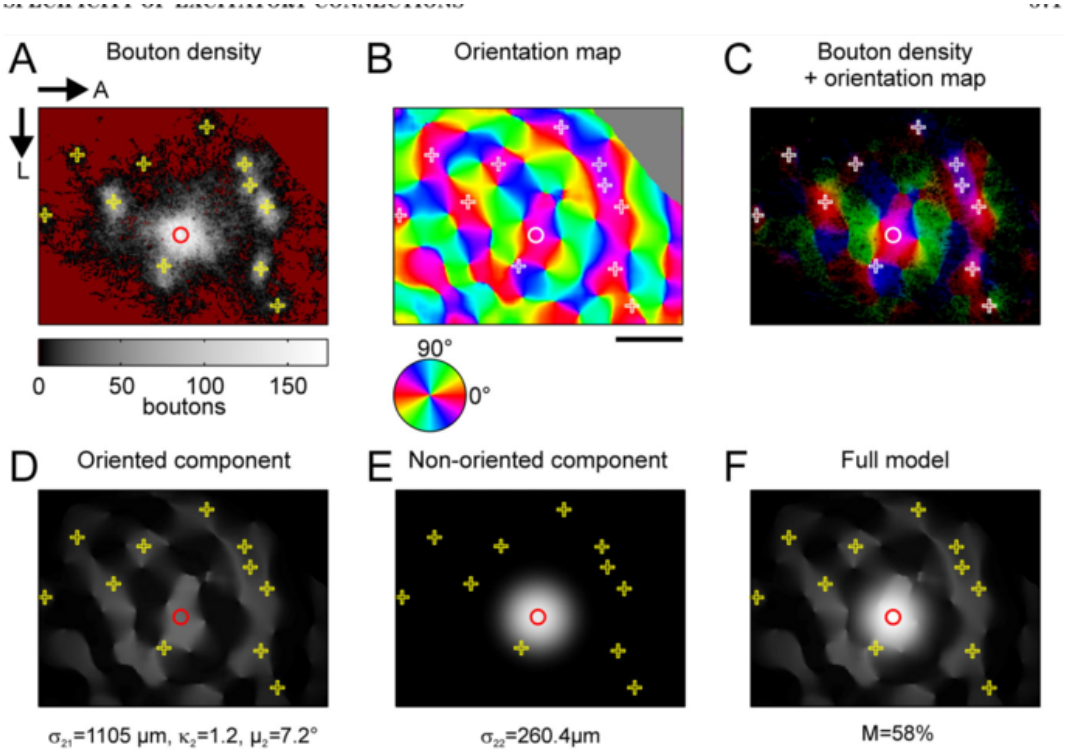


Figure 4.15: Lateral excitatory projection bouton density and orientation maps in layer 2/3 of cat V1 demonstrating the application of a Gaussian and von Mises model of lateral connectivity. Reproduced from Buzas et al. (2006).

to anatomical measurements. For this purpose we will be fitting a descriptive model, developed by Buzas et al. (2006) to the lateral connectivity data.

The distance of long-range connectivity varies considerably across species so using some anatomical estimates from macaque we will attempt to refine our estimates of the long-range oriented component. Anatomic data suggests that the spatial spread of lateral connections can be anywhere between 3-10 mm (on average 6-7 mm) in total length (Angelucci et al., 2002b). Along its principal axis the visuotopic mono-synaptic spread of V1 horizontal connections has a mean of $2.47° \pm 0.3°$. This falls well within the range of estimates for the lsRF as published in a number of studies (Sceniak et al., 1999; Sceniak and Hawken, 2001; Shushruth et al., 2009), which employed the iDoG protocol.

The model that was used to fit the lateral weights describes the patchy lateral connectivity found in layer 2/3 of V1 as a function of two distinct components. A short range isotropic Gaussian pattern and a long range pattern, defined as a von Mises function, which is combined with the orientation map. The model therefore assumes that lateral connectivity develops as a function of both the proximity in space but also along a particular feature dimension, in this case the orientation. The

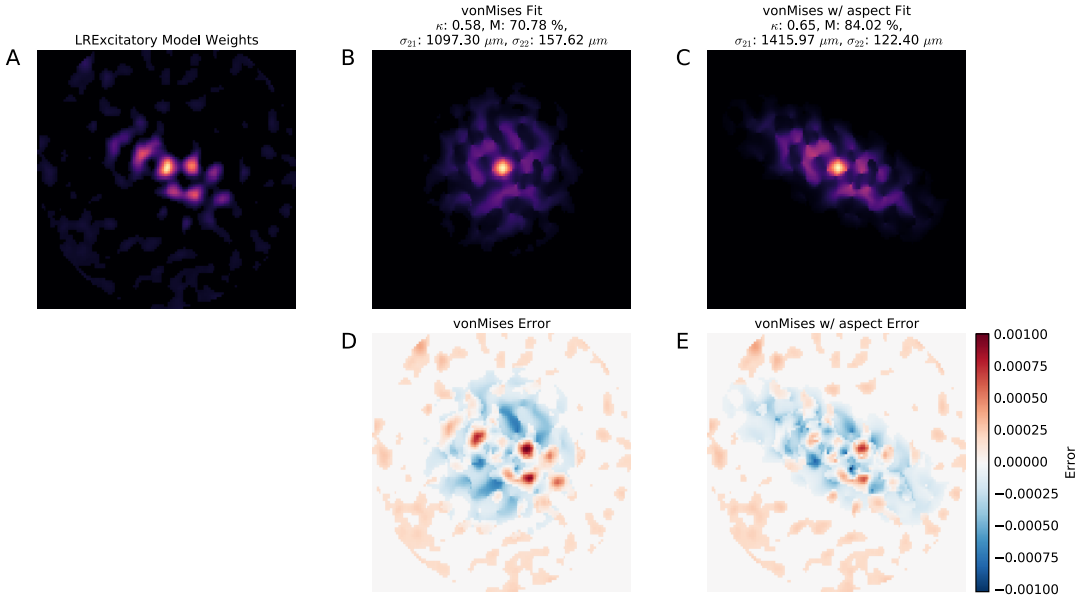


Figure 4.16: Example fits obtained by fitting Gaussian and von Mises distributions to the thresholded long-range lateral weights. A) The thresholded long-range lateral weight matrix (at the 70th percentile). B, C) Example fits using the von Mises orientation model with and without an aspect parameter. D, E) Error between the fits and the thresholded weight matrix pattern.

equations underlying this model and an extension, which also exposes the aspect ratio of the long-range Gaussian kernel are described in Section 4.1.3.

The full model fitting procedure for an experimentally traced lateral connection field is shown in Figure 4.15. By applying this fitting procedure we can effectively estimate the spatial extents of both the local isotropic local kernel and the long-range excitatory kernel. The Figure 4.16 demonstrates what one such fit looks like for the SCAL model, while the full distribution of local and long-range kernel values is shown in Figure 4.17. Note that the weight array that is being fit has been thresholded at the 70th percentile of non-zero values to ensure the weight matrix more closely resembles the bouton density maps obtained by Buzas et al. (2006).

The results of our fitting procedure shown in Figure 4.17 show good correspondence with these experimental estimates with a mean long-range connectivity that has a spatial constant of around 1645µm but extends beyond that with our cut-off defined at 2.5° or 7.5mm. The model with the additional aspect parameter fits the experimentally obtained long-range spatial constant more closely, but neither extends to the furthest reaches of the lateral connection field. The local excitatory kernel provides a reasonable fit to experimental estimates with a mean local excitatory kernel with a spatial constant of around 221µm, compared to the 280µm estimated in cat V1

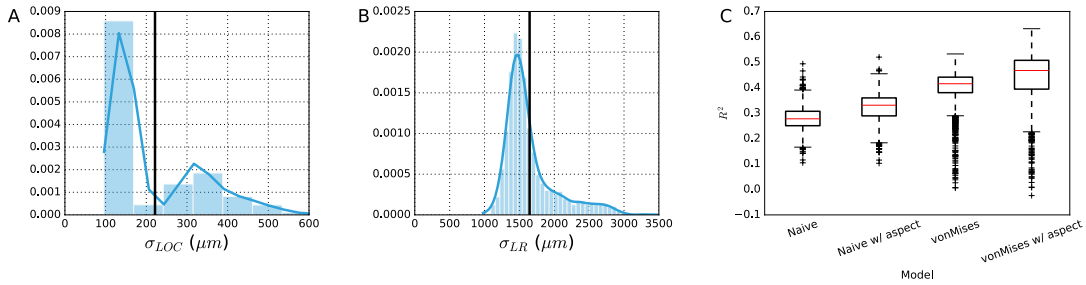


Figure 4.17: Distribution of spatial constant obtained by fitting the Buzas et al. (2006) von Mises and Gaussian model to long-range lateral excitatory connections developed as part of the SCAL model. A) Distribution of kernel sizes of the local excitatory component. B) Distribution of kernel sizes of the long-range excitatory component C) Comparison between naive model without an orientation dependent component (i.e. pure Gaussian), the von Mises model including the orientation dependent component and additionally versions of both with an additional term for the anisotropy in the connectivity.

(Buzas et al., 2006). However most neurons exhibit much smaller local kernels and the distribution seems to be distributed bi-modally.

Finally we can evaluate in how far the orientation preference and selectivity of the post-synaptic neuron predicts the long-range lateral connections it receives. Specifically we can evaluate how closely the κ value of the von Mises fit, which represents the bandwidth of the kernel in the orientation domain, is correlated with the orientation selectivity of the neuron. The relationship between these two variables is shown in Figure 4.18, and clearly predicts that the orientation specificity of the lateral connections received by a V1 neuron is highly dependent on the orientation selectivity of the neuron (Spearman's $\rho = 0.59$, $p < 10^{-184}$).

Inhibitory connectivity

Since the SCAL model does not have distinct populations of V1 we will consider the joint spatial and orientation distribution of basket cells in order to determine the inhibitory profile and compare model to experiment. Kisvárday et al. (1997) provide the best known estimates for the spatial distribution of inhibitory basket cells, having injected tracers much more conservatively than previous studies. Their estimates indicate that inhibitory projection in cat V1 extend no further than 1.5-2 mm with most synaptic boutons being distributed in a central regions approximately 1 mm in diameter. In the absence of any precise fits, equivalent to the Gaussian and von Mises model available for the excitatory connectivity, we restrict inhibitory connectivity to a conservative 1.2 mm in diameter, which excludes the long tail of these connections.

By binning the strength of connections by their distance and the orientation difference between pre- and post-synaptic neurons we can at least qualitatively assess

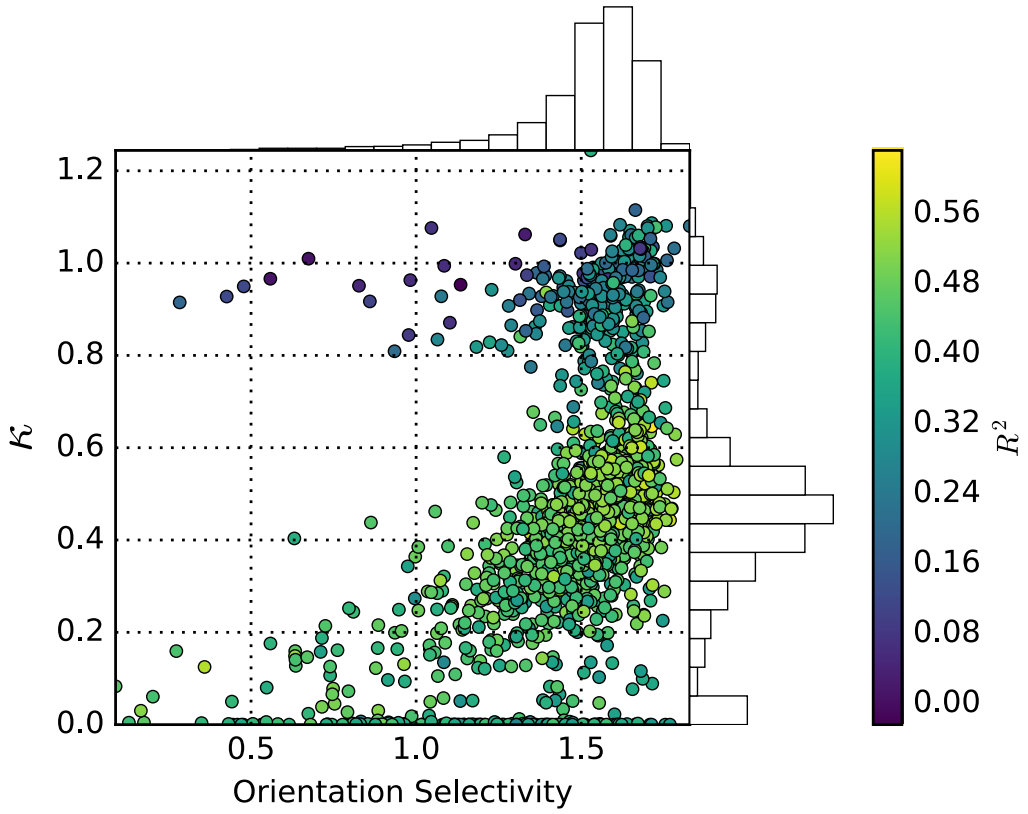


Figure 4.18: Scatter plot of the κ variable of the von Mises distribution and the orientation selectivity of each neuron, showing a clear relationship between the orientation specificity of patchy lateral connections and the post-synaptic neurons orientation selectivity. The points are additionally colored by their goodness of fit and marginal histograms of the distributions of the κ and orientation selectivity are provided. Fits with an R^2 value below 0.3 were rejected.

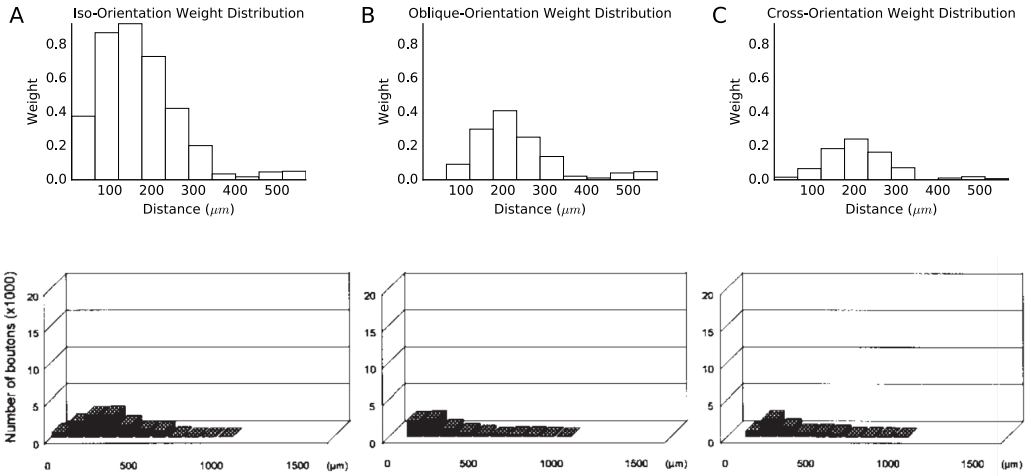


Figure 4.19: Distributions of inhibitory projections as a function of space broken down into iso-orientation, oblique orientation, and cross orientations relative to the orientation preference of the post-synaptic neuron. Top row shows model results bottom row reproduces results from Kisvárday et al. (1997) in cat area 17. The iso-oriented connections are most numerous and strongest overall and follow a similar distribution although they drop off more quickly at larger distances. Oblique- and cross-orientation connections start at further distances and are overall less frequent. However since only the central region is simulated the long tail of the experimental distributions can't be replicated.

how closely the model distribution matches experimental bouton density maps. The analysis breaks down the spatial distribution of connections targeting iso-, oblique and cross-orientation regions, comparing against the equivalent analysis performed by Kisvárday et al. (1997) in cat area 17. While we are restricting the inhibitory connectivity to a smaller region the experimental results highlight again that most weights fall into the central region. Additionally we can see that the distributions of weights broken down by distance are qualitatively very similar, with most of the weight in the iso-orientation case being centered near the center, while oblique and cross-orientations are strongest at medium ranges.

Sparsity of connections

The implementation of the GCAL and SCAL models has generally meant that neurons densely innervate neurons within the defined spatial extent. However from experimental studies it is known that neurons actually only make fairly sparse contact to neurons around them and once the initial sprouting and pruning phase of development is over it is generally much harder to form new connections. In particular we know that a V1 neuron is only contacted by a fairly small number of afferent pro-

jections from the LGN and that lateral projections while very dense innervate only a small fraction of neurons around them. This has profound implication for development because neurons have much sparser input than is assumed in these models, secondly it can strongly affect the development of the model because once connections have been formed it is much more difficult to shift the synaptic weights to some other location.

A detailed investigation into the processes underlying sprouting and pruning and the competition for afferent inputs would be a thesis in itself. However for computational purposes but also to test the robustness of the model we implemented a simple form of sprouting and pruning, which starts out with a densely connected network and slowly prunes away connections while also generating a small number of new weights.

In this way we can demonstrate that sparsity of connections can indeed increase the robustness of the model and that how true long-range patchy connectivity akin to the thresholded version used in the von Mises model fitting procedure can emerge from the model.

4.3 DISCUSSION

In this chapter we set out to spatially calibrate the GCAL developmental model in order to cross-validate known measurements of the spatial profile of afferent, excitatory and inhibitory connections in the primary visual cortex. Additionally we wanted to explore how the intra-areal connectivity in the cortex can give rise to a Mexican hat like profile, which is thought to underlie the organization of the cortical neurons into smooth topographic maps, which optimally map features in the sensory input onto the surface of the cortex.

In the literature review we had established that long-range connections in the cortex are almost exclusively excitatory. However the spatial scales involved in mediating Mexican-hat–like connectivity are well below the long-range excitatory connections, which can often span many hypercolumns. By incorporating the empirically determined spatial scales of the model we were able to establish that the relatively long inhibitory projections originating from the basket cell population in V1 are sufficient to account for the Mexican hat profile. At the same time we show that long-range surround suppression cannot be implemented directly by these projections and must instead by mediated by a di-synaptic mechanism, via the long-range excitatory projections.

First we will review how we determined the spatial parameters of the SCAL model, before critically reviewing the experimental measurements given in the light of the

model results. This will let us highlight issues with DoG fitting procedures employed to fit area summation curves in visual cortex and call into questions some conclusions made about the source of contrast-dependent size-tuning shifts. Finally we will discuss the model of lateral excitatory connectivity and how the addition of an aspect parameter can help us quantify the isotropy of patchy lateral connectivity in a way that can easily be applied to experiments.

4.3.1 *The LGN Model*

In spatially calibrating the early stages of the model, roughly corresponding to the retinal ganglion cell and lateral geniculate nucleus, we relied mostly on electrophysiological estimates of the size and area summation curves of the neurons. In doing so we had two primary aims, first of all since we chose only a single spatial filter rather than a distribution for computational reasons. Later models can easily add further diversity to the model by drawing the the spatial constants of the center-surround filters from a distribution. In the absence of this diversity we chose a spatial filter size that had to have several important properties.

The first was that the chosen values should not diverge too far from the values obtained by fitting a DoG model to area summation curves. We confirmed this was the case and plotted the results compared to the empirical data obtained from Sceniak et al. (2006), finding the values at the lower end of the distribution.

The second constraint was that the spatial frequency preference of the neurons could not be too low in order to allow V1 neurons to integrate over at least one cycle of the main spatial frequency preference. This meant that the preferred spatial frequency of the network should fall somewhere between 1 and 2 cycles per degree of visual angle. Again we confirmed this was the case and compared the result to similar results obtained by Levitt et al. (2001) in the visual cortex of macaques.

The extent of the gain control projection in LGN could not be experimentally verified as there is little to no data about the precise anatomical profile of inhibitory neurons in the LGN. Therefore we chose a value that would provide divisive gain-control over the span of the receptive field of a single V1 neuron.

The final model parameters are summarized in Appendix A.2 and visualized in Figure 4.1. In future this model may be extended by drawing the extent of these afferent connections from a distribution. For the purposes of this thesis, this level of calibration is sufficient however and will provide the V1 model with input that is roughly at the desired spatial scale.

4.3.2 *The V1 Model*

The V1 model that was introduced in this chapter is able to replicate a wide range of experimental results and provides a much closer match to experimental results than the existing GCAL model. In addition to spatially calibrating the extents of afferent and lateral connections, the model is also the first developmental model that uses divisive rather than subtractive inhibition. The divisive model provides a much better account of size-tuning responses, which shift to larger integrative areas depending on the contrast, which agrees with a wide range of existing models. We also showed that the extents of inhibitory basket cells is sufficient to account for the Mexican hat profile that has long been used to explain the organization of neurons into topographic feature maps by pulling together activity locally and suppressing it in nearby regions, which leads to a smooth mapping of different feature preferences across the cortical surface.

However, in exploring this model we have also encountered clear limitations, the most important of which is the lack of distinct excitatory and inhibitory populations. This violation of Dale's law, which states that neurons release the same neurotransmitters at all of its synapses and therefore have either an excitatory or a inhibitory influence on post-synaptic cells, poses a serious problem for a developmental model. This is because Hebbian learning will ensure the strength and spatial profile of the neural connections is dependent on the response properties of the neuron, yet in animals different cell types may respond very differently. This has made it difficult to directly compare the lateral connectivity in the model to experimental data.

Differences to GCAL

There are a number of major changes between the GCAL model and the SCAL model that was introduced in this chapter. As we have already emphasized the SCAL model operates on divisive rather than subtractive inhibition. While inhibition in the brain generally has both divisive and subtractive effects, the divisive normalization paradigm has shown tremendous success in explaining a wide range of effects including modulatory effects, contrast-dependent changes in responses, decorrelating the inputs and even effects associated with visual attention (Cavanaugh et al., 2002a; Graham, 2011; Carandini and Heeger, 2012; Reynaud et al., 2012; Coen-Cagli et al., 2015).

The other aim was to move away from other unrealistic mechanisms. While the model does not learn continuously it no longer artificially resets the activity between stimulus presentations or saccades. The model now behaves more continuously with no hard reset between stimulus presentations. Instead the model activity from a pre-

vious step decays at the start of a new presentation and is usually extinguished completely by the time it reaches V1.

While all the spatial profiles have changed in some way or another the most significant change was the size of the afferent field relative to the lateral connections in V1. While the receptive fields of neurons in GCAL generally only subtended only a single hypercolumn, the receptive fields in SCAL span somewhere between 3-4 hypercolumns. Long-range connectivity in the GCAL model therefore generally connected neurons with non-overlapping receptive fields, while the SCAL model highlights that this is generally not the case in parafoveal regions of macaque V1. A good portion of the lateral connectivity field actually connects neurons with significantly overlapping connection fields and only a portion of long-range patchy connections reaches outside the neurons receptive field. This has significant implications for the function of these connections, rather than purely reflecting the co-occurence of stimuli the patchy connections will strongly mirror the response of the neuron itself. This suggests that lateral connections are involved in improving the encoding of the local stimulus in addition to mediating true surround influences.

Despite some of these limitations the model exhibits very robust development under varying stimulus conditions and still exhibits all the nice properties of the GCAL model, including contrast invariant orientation tuning as well as robust, yet adaptive self-organization.

Size tuning

A major component of the analysis we applied here was in comparing Difference-of-Gaussian (DoG) and Ratio-of-Gaussian (RoG) models of the area summation curves from both models and comparing the results to equivalent experimental data. Since the model itself does not exert any long-range influences any changes in the fit is most likely mediated by the relatively short-range connections in the model.

Using the subtractive DoG model we did find significant shifts in the integration area depending on the contrast level in the SCAL model but not in the GCAL model. Since the actual integration area is fixed in both models this suggests the contrast-dependent shifts in effective integration area can be explained by changes in the relative gain of the center and surround component in a ratio-of-Gaussians model, just as was suggested by Cavanaugh et al. (2002b) and (Solomon et al., 2006). However it is also worth noting that the value of k in equation 4.2 describing the LGN activation may have a role here. Under low contrast conditions the additive k component in the LGN normalization term can dominate relative to the projection strength and result in shifts in size tuning.

The divisive RoG models were not able to fit the data quite as well (see Figure 4.4) and have therefore not been shown in as much detail. However comparing these two models of size tuning shows that the divisive model performs significantly better when the spatial constants of the center and surround are not fixed, i.e. the ratio in the gain between the center and surround components explains part of the size tuning shift the subtractive DoG model captures through a shift in the spatial constants of center and surround. This suggests that subtractive suppression is not a good model of surround suppression in the LGN or V1 and that contrast-gain control can explain a large portion of the contrast dependent effects observed in experiment and in the SCAL model.

It is therefore not necessary to invoke integration over a larger area via long-range lateral connections as was suggested by the experimental work by Levitt and Lund (2002) and Sceniak and Hawken (2001) and translated into a model by Schwabe et al. (2006). This does not mean however that long-range lateral connections do not have any influence. Indeed it is difficult to envision how the model can mediate any longer range feature specific excitation and inhibition without corresponding long-range projections between neurons, which are themselves tuned for those specific features.

Lateral Connectivity

Excitatory In addition to fitting the response properties of the V1 neurons this chapter also demonstrated how the anatomical distribution of synaptic boutons can be compared to the lateral connectivity in the model. By fitting a model made up of a local Gaussian kernel and an orientation dependent long-range kernel, first devised by Buzas et al. (2006), to the data, we are able to achieve good fits demonstrating that the lateral connectivity captures the statistics of the visual input. In addition the results indicated that at least the extent of the long-range connectivity is consistent with experimental results. However, the short-range kernels were significantly smaller in scale than what we would expect from experiment. There are multiple factors that could account for this difference. The main issue being that the model only captures one population of cells, which generally display a high level of feature selectivity. More diversity, such as the introduction of unselective inhibitory cells could explain larger and more weakly tuned local connections. Another factor that could play a role is that neurons in macaque V1 are generally much more dense and the hypercolumn distance considerably shorter. With relatively larger hypercolumns we could also expect that the excitatory connectivity grows in size. Finally, the development of connectivity in the model is driven entirely by Hebbian rules, and therefore cannot account for any intrinsic, genetically mediated programming, which could lead to a dense local network of connections.

We will revisit some of these issues in later chapters to investigate if some of these factors can significantly affect these results and whether we can reach a more harmonious fit of the local excitatory connectivity.

Inhibitory The literature surrounding inhibitory connectivity is much more limited and no detailed estimates of cell-type specific spatial profiles for the primate visual cortex currently exist. Therefore we have had to extrapolate from the limited existing data. In the literature review we explored the known properties of various cell classes and identified fast-spiking Parvalbumin-expressing interneurons as the most likely source of connectivity to drive developmental processes, particularly due to their broad tuning profile (Albus and Wahle, 1994; Kisvárday et al., 1997; Hofer et al., 2011; Ma et al., 2011) and high abundance in the thalamocortical recipient layers (Van Brederode et al., 1990; Hogan et al., 1992; Huxlin and Pasternak, 2001) but also because they have been implicated in developmental disruptions (Fagiolini et al., 2004; Hensch and Stryker, 2004).

In a Hebbian learning model lacking distinct excitatory and inhibitory projections the distribution of the inhibitory connections is determined by the responses of the joint population. Therefore it is not possible to directly compare the inhibitory connectivity profile to any anatomical equivalents. The comparison of the spatial and orientation distribution of connections does show significant similarity however, something we will investigate in more detail when we have a model that actually expresses distinct excitatory and inhibitory populations.

Diversity of responses

One of the major design decisions in creating this model was to decide on whether to match the average response of neurons in V1 or to try to incorporate the huge amount of variability that is observed in the real cortex. Due to computational tractability issues we have focused on matching the average response rather than capturing the full level of detail immediately. This means that the LGN stage filters the visual input at just one spatial frequency, which severely limits the variability in receptive field properties. This is particularly evident when comparing the distribution of spatial constants, the size-tuning shifts and indeed the measured receptive fields themselves.

The receptive field Gabor analysis in particular highlights the lack of diversity in the receptive fields. This is a direct result of the lack of LGN filter diversity and a better fit may be achieved by incorporating a wider range of spatial frequency filters at that level. Additionally the circular bounding region of the model connection fields will prevent neurons with very elongated receptive fields from forming.

There are almost infinite sources of variability in the cortex, including the distributions of different cell classes, variability in the number and strength of received connections, and the spatial and frequency preferences of the feedforward filters. Nonetheless the SCAL model performs well as a first order approximation and will slowly be refined as we construct more complex models in later chapters.

4.4 CONCLUSIONS

In summary, we have presented the first developmental model of the primary visual cortex that captures the spatial extents of various connections in the cortex. Unlike previous static models, performing this calibration and analysis in a developmental model has several major benefits. Specifically, because we calibrate only the initial profile of the connectivity the developmental processes themselves provide a form of validation for the final results. This property has allowed us to draw several concrete conclusions.

First of all we were able to establish that the network of local excitatory and inhibitory connections is sufficient to account for the Mexican hat profile, which drives the organization of cortical neurons into smooth topographic maps. Indeed the estimates used for the extent of inhibitory projections was deliberately chosen to be very conservative to demonstrate that implausibly long-range inhibitory connections are not needed to drive development of an orientation map.

Secondly we showed that introducing divisive inhibition is sufficient to account for contrast-dependent changes in the spatial integration of V1 neurons. Divisive normalization has been described as a canonical computation performed by the cortex and indeed we are able to account for contrast-dependent changes in the response of V1 neurons using the divisive model. This fits the findings by Carandini and Heeger (2012) and many others and suggests that long-range lateral connections are not necessary to account for increases in the integration area at low contrast as has been suggested by some models (Levitt and Lund, 2002; Schwabe et al., 2006).

Finally we were able to show that the model develops realistic long-range patchy connectivity that we were able to fit with an adapted version of the Buzas et al. (2006) model. This provides a framework to be able to assess in how far the similarity in feature preferences between pre- and post-synaptic drives strong connectivity between the two cells. Additionally it provides a way to quantitatively assess the spatial extent of these connections.

However we have also identified a number of problems with this model, the most important of which is the violation of Dale's law, which states that a single neuron can only employ one main neurotransmitter and therefore not mediate excitatory

and inhibitory effects at the same time. This has profound effects on the final model because the connectivity is shaped by responses in a Hebbian model. In the following chapter we will therefore explore how we can separate these populations and what properties make different inhibitory cell classes suited towards specific functions such as sparsifying the response, divisive normalization and feature specific contextual effects.

5

EXPLORING THE ROLE OF INHIBITION IN CORTICAL DEVELOPMENT

In the previous chapter we explored the spatially calibrated SCAL model, establishing how various sources of evidence about the spatial structure of V1 projections relate to each other. While this model was able to account for the size-tuning properties of the visual cortex quite well, it also exhibited two major shortcomings. Most importantly it highlighted that a model lacking distinct excitatory and inhibitory populations will not be able to capture the diversity in connectivity and response profiles that are observed in the mammalian cortex. Secondly we showed that based on the known lack of truly long-range inhibitory projections, long-range orientation-specific suppression must be mediated by a di-synaptic mechanism mediated by long-range excitatory connectivity.

Additionally, in the literature section we described the known properties of various inhibitory cell classes and what roles they might perform. In particular we discovered that PV and SOM-expressing interneurons exhibit highly distinct response properties and layer-specific expression patterns. With recent techniques enabling the targeting of specific neural populations, there is now huge interest in understanding their role both in development and in mediating and gating both contextual and attentional modulation phenomena.

In this chapter we will propose a model that incorporate the distinct response properties of an excitatory pyramidal cell population and the Parvalbumin expressing interneuron population, allowing us to make concrete predictions about their role in development. We will establish that the fast and linear response of the PV-ir, fast-spiking interneuron population makes them ideally suited towards controlling feedforward activity, sparsifying activity and thereby driving map formation. Additionally we will show that this model exhibits more robust and stable map formation even in the presence of strong lateral excitation than the previous model.

Specifically we will assess how the model performs when manipulating the response properties of the PV population, under varying levels of contrast and strong lateral excitation. The SOM inhibitory population will be disregarded we are primarily focused on development of the orientation map, which is known to occur before long-range lateral connectivity emerges, which is known to drive the SOM interneurons. This inhibitory cell class will be revisited in the following chapters.

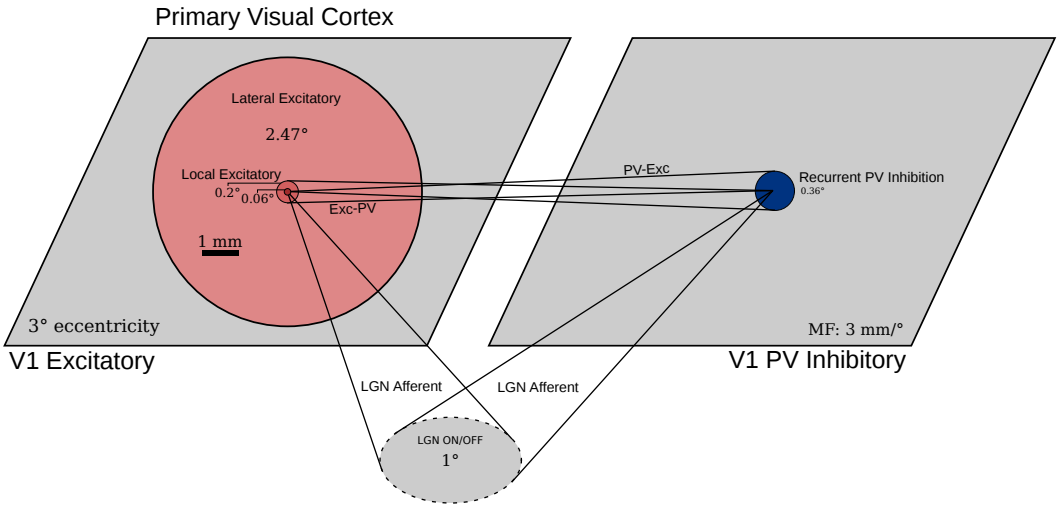


Figure 5.1: Diagram of the SEPI V1 stage of the model showing the spatial scales of the various excitatory (red) and inhibitory (blue) connections. Saturated colors indicate the kernel radii, while lightly shaded regions indicate kernel cut-off extents.

5.1 METHODS

5.1.1 *The SEPI Model*

The **S**hort-**R**ange **E**xcitation **PV** **Inhibition** (SEPI) model is based on the same principles, spatial profiles and equations that were described for the SCAL model in the previous chapter. However, unlike SCAL, the model has a distinct excitatory and inhibitory V1 population, which both receive afferent input and are connected to each other recurrently. The architecture diagram (Figure 5.1) shows the two sheets and projections between them. When comparing this against the SCAL model diagram (Figure 4.2), we can see that the spatial scales of the retinotopy, which generally indicates poor orientation map organization various projections are unchanged—they have simply been split up between the populations.

In the literature review we identified the parvalbumin (PV) expressing interneurons, which include fast-spiking basket and clutch cells, as the most likely candidate to provide feedforward inhibition and controlling the gain of the response in of the pyramidal cells. Their high abundance in the thalamocortical recipient layer 4 and layers 2/3 (making up over half the population in each; Van Brederode et al. 1990) as well as their involvement in the onset of the critical period (Fagiolini and Hensch, 2000) and their effect on the columnar organization of the visual cortex (Hensch and Stryker, 2004) makes them of particular interest. The other defining characteristics of the PV population is their fast response (Cruikshank et al., 2007; Gabernet et al.,

2005), and their ability to linearly match the activity in the excitatory population (Atallah et al., 2012). These properties describe the role of the inhibitory projection in the SCAL model, which provides divisive gain control and is directly coupled to the response of the excitatory population.

The equations governing the responses of both the excitatory and inhibitory population are also unchanged, driven by summing the excitatory input and divided by the inhibitory activity. This closely matches what we know about PV neurons in the cortex, which provide strong peri-somatic synaptic inputs to spiny neurons, providing shunting inhibition (Atallah et al., 2012; Wilson et al., 2012). Based on tracing studies we know that excitatory and inhibitory neurons and specifically the Parvalbumin expressing subpopulation strongly and densely innervate each other and themselves (Buzás et al., 2001; Ma et al., 2011; Pfeffer et al., 2013). Additionally these PV population receive strong afferent input (Burkhalter, 2008). All of these properties are captured by the model.

There is some limited evidence the learning rules on parvalbumin expressing neurons may differ in input-cell-specific ways, with some anti-Hebbian phenomena having been found in hippocampal circuits (Le Roux et al., 2013). However for the sake of simplicity the learning rule applied in the inhibitory population is the same as in the excitatory population described in equation 4.11.

In order to model the response of the PV population we apply only half-rectification to the inhibitory population, rather than a floating homeostatically determined threshold, because basket cells generally have a very low spiking threshold (Ma et al., 2011). It would be possible to use homeostatic plasticity to set the thresholds instead as suggested by studies in the retina (Hennig et al., 2011), but it would be complicated to set up mechanisms to balance the thresholds between the populations to ensure suitable overall levels of activity in the network. Secondly, we keep the effect of the inhibition divisive, both for each other and when targeting pyramidal cells.

Activation

The activation for both the excitatory and inhibitory population is given by:

$$\eta_E = \eta_I = \frac{\eta_A + \eta_L}{1 + \eta_P} \tag{5.1}$$

where η_A is the LGN afferent activity, η_L the local excitatory contribution and η_P the PV inhibitory contribution. In other words, both populations integrate over afferent and local excitatory input and are then divisively normalized by the PV population. To allow testing its effects, we will also allow an optional additional long-range ex-

citation term to modulate the neuron activity depending on long-range excitatory input:

$$\eta_E = \frac{\eta_A + \eta_L}{1 + \eta_P}(1 + \eta_{HE}) \tag{5.2}$$

where η_{HE} is the input from the horizontal excitatory projection. The full set of parameters for the SEPI model are listed in Appendix A.2. Modeling the long-range input as an additional multiplicative component has a number of biophysical explanations, which we will revisit in more detail in the next chapter, however there is strong evidence that the influence of long-range excitatory input is strongly voltage dependent (Hirsch and Gilbert, 1991), having little effect when afferent input is weak but strongly modulating the response under stronger input conditions.

The combined activation of the various projections is still combined in the same way as in the previous models (shown in equation 4.6 and as stated above a half-rectifying function (f) is used in the inhibitory population, while a homeostatic function is used among the excitatory population. The response function of the neurons is therefore given by:

$$\eta_{j,V}(t) = f\left(\frac{\sum_{p=\{E,A\}} \gamma_p C_{jp}(t)}{1 + \sum_{p=\{I\}} \gamma_p C_{jp}(t)}\right)^{\beta} \tag{5.3}$$

where β is the only new term controlling the linearity of the response, which will be used throughout this chapter.

Hysteresis

In order to control the temporal properties of the PV population we introduce a hysteresis term, which will be disabled in the final model but will in the meantime allow us to test the importance of fast inhibition in the model. The hysteresis function is defined as a discrete exponential weighting between the previous activation and the new activity, expressed as:

$$\eta_h(t + \delta\eta) = \eta(t) + \tau\big[\eta(t + \delta t) - \eta(t)\big] \tag{5.4}$$

where τ is the time constant. If $\tau < 1$, hysteresis will effectively slow down PV responses relative to the excitatory population. In the final model, the hysteresis term is eliminated simply by setting $\tau = 1$.

5.1.2 *Assessing the quality of orientation maps*

A major component of the analysis applied in this section is to assess the sensitivity of the model to changes in the response properties of the different cell classes. Therefore we have to establish a number of metrics to assess the robustness of the model to changes in contrast by assessing the quality of orientation map organization. To measure the maps, we will focus on the excitatory V1 neurons, though similar results would be found if we were to pool all the neural types together as would be the case in optical imaging experiment. There are a variety of measures to assess map properties, but we will focus on four specific measures in particular—smoothness, stability, selectivity and pinwheel density.

Pinwheel density

Interestingly, it has been found that experimentally measured orientation maps share a fundamental property across species: pinwheel count scales linearly with hypercolumn size. More specifically, there are π orientation pinwheels within the area of one hypercolumn, when averaging over sufficiently large cortical areas. This dimensionless, statistical measure of pinwheel distribution is thought to reflect a universal constant of map organization, converging to π across carnivorans, primates, cats, and tree shrews (Kaschube et al., 2010; Keil et al., 2012). This value was predicted by a theoretical model of map organization and has strong empirical evidence, with a mean pinwheel density across four species (tree shrew, galago, cat, and ferret) statistically indistinguishable from π.

Pinwheel density provides such a good measure of map quality because it is strongly influenced by any number of disruptions of map organization. Analyzing poor quality maps often reveals other disruptions in development including poor retinotopy and selectivity. While it is possible to generate maps which have a deceptively good pinwheel density metric, this problem can be eliminated by initializing the same model with multiple seeds as was demonstrated by Stevens et al. (2013b).

To determine the pinwheel density for any given map, the hypercolumn distance is computed and all pinwheel centers are found. Pinwheel centers are located at the intersection of the zero contours of the real and imaginary components in the polar representation of orientation preference (Löwel et al., 1998). Then we simply divide the number of pinwheels in the modeled area by the number of hypercolumns to arrive at the pinwheel density. Stevens et al. (2013b) introduced pinwheel density as a numerical metric for the biological accuracy of a map, for which it works well because it has a clear target value.

Orientation Selectivity

When measuring an orientation map using the protocol outlined in section 4.1.2, two components are returned. The first is the orientation preference computed as the vector averaged estimate of the neurons preference. Additionally there is a selectivity component, which provides an estimate of how much the neuron will respond to the preferred orientation relative to all other orientations. This value provides a simple metric to assess how selective a neuron is, and can easily be averaged across all neurons in a central region to get an average selectivity value.

Stability

The stability of orientation maps is determined by determining the average orientation similarity index of the orientation map over the course of development. To assess similarity quantitatively, Chapman et al. (1996) computed the difference of orientation preference at each developmental age with the organized preference map observed in the final recording for that animal. As in Stevens et al. (2013b) we normalize these similarity values to fall between 0 (completely uncorrelated) to 1 (identical orientation preference). The orientation similarity index is therefore defined as:

$$\mathrm{OSI} = 1 - \frac{4}{n\pi} \sum_i |(F_i - O_i) \bmod \frac{\pi}{2}| \tag{5.5}$$

where i iterates over each neuron in the orientation map F_i represents the final orientation preference and O_i is the orientation preference at an earlier time. By averaging the OSI at every 1000 development steps we can numerically compare the stability of the model over time, which provides a measure of how quickly and how smoothly the map is becoming organized. Note that there is no objective criterion for "best" stability, since a perfectly stable map that never achieves selectivity would be a failure both as a model and as a visual system, but in conjunction with other measures the stability metric is helpful for evaluating the effect of changing specific parameters for a given model.

Local Homogeneity Index

The local homogeneity index (LHI) was first introduced by Nauhaus et al. (2008) as measure of the local smoothness of the map. It is therefore useful to determine whether a neuron is in a smooth region of the orientation map or near a pinwheel region where the orientation preference varies considerably. The LHI is defined based on a specific σ value, which should reflect the spatial properties of the map. There-

fore we can use the mean LHI across the entire map as a measure of the scale of homogeneous regions of the map. The LHI at location x is computed as follows:

$$\mathrm{LHI}(x) = \frac{1}{2\pi\sigma^2}\left|\int \exp\left(\frac{-\|x-y\|^2}{2\sigma^2}\right)\exp(i2\theta_y)dy\right| \tag{5.6}$$

where θ_y is the orientation preference at site y and σ determines the spatial scale of the analysis. In their experiments Nauhaus et al. (2008) used $\sigma = 180\mu m$ to match the spatial extent of dendritic fields in both cat and monkey cortex.

Center-of-Gravity Distortion

Since we have access to all the weights in the model, we can easily compute the center-of-gravity (CoG) of the afferent weights relative to the position of the neuron, to determine how far the neuron's preferred location is from its perfectly retinotopic mapping. Some retinotopic distortion will always be present, due to locally smooth organization for features like orientation, but extreme distortions are not typical for biological maps, and are thus an indication of inappropriate parameter settings or training patterns for a model network.

The CoG along one axis is computed as the mean of the normalized vector product:

$$G(p,\omega) = \frac{1}{\sum_{i,j}\omega_{i,j}}\sum_{i,j}\omega_{i,j}p_{i,j} \tag{5.7}$$

where p is the position along either the x- or y-axis and ω is the weight at that location. The distortion in the CoG is then given by Euclidean distance between the center of gravity and the actual retinotopic location of the neuron:

$$G_d(x,y) = \sqrt{(\mathrm{CoG}_x - x)^2 + (\mathrm{CoG}_y - y)^2} \tag{5.8}$$

By computing this distance for all neurons in a central region, we can use this metric as a measure of the quality of the retinotopic mapping in the model, highlighting any distortions.

5.2 RESULTS

In this section we will present results analyzing the results from both the SCAL and SEPI models. First establishing the role of inhibitory neurons in development in the SEPI model and then comparing how robust the models are to varying levels of contrast and long-range excitation. These results will primarily be based on four

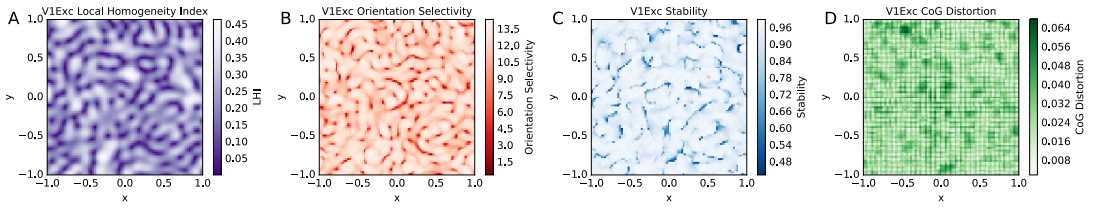


Figure 5.2: Metrics used to evaluate orientation map development taken from a 2x2° region in the model. A) Local homogeneity index representing how similar neurons in a local region are, poor organization will lead to large high-homogeneity regions or low global homogeneity. B) Orientation selectivity measures the overall selectivity of the neurons. Low mean selectivity indicates poor and unstable development. C) Stability measures the self-similarity of the orientation preference over time. Low stability (<0.8) indicate continuous reorganization of the orientation map. D) Center-of-gravity distortion represents a measure of how precise the retinotopy in the model is. Significant amounts of CoG distortion indicates poor retinotopy, which generally indicate poor orientation map organization.

different metrics the local homogeneity index (LHI), orientation selectivity, stability, Center-of-Gravity distortion and pinwheel density.

Before using these metrics to evaluate the different models it is important to understand what they actually measure and how it relates to the quality of the model. The pinwheel density is the only scale-free measurement and has been discussed sufficiently in the previous chapter. The other metrics are shown for each unit in the model in Figure 5.2 and will be averaged to get a single value to assess the model.

5.2.1 *Development*

The role of inhibition in development is hugely important, as evidenced by the fact that a variety of pathological conditions including autism (Wöhr et al., 2015) and schizophrenia (Lewis et al., 2012) have been tied to changes in PV-expressing interneurons in the cortex. Modeling the role of these cell classes in development and adult visual processing requires accurately modeling their response properties. Since the synaptic projections develop as a result of these properties, the tuning properties of the modeled neurons in the developed model will let us make concrete predictions about the role of these neurons in shaping V1 development.

Accurately capturing the response of these neurons in a rate-based model presents a number of challenges because we cannot explicitly model fast-spiking kinetics, which means we are limited to modeling higher absolute firing rates. Additionally, much of the data about these cell classes comes from mouse experiments, because of the new, rich techniques for genetic manipulation of specific cell classes, rather than primates or carnivorans as for most of the functional data.

The spatial extent of the PV neurons was already calibrated for the SCAL model, and so the next step is to calibrate their inputs and response properties. The literature is very clear that PV neurons in the visual cortex receive strong inputs both from thalamocortical afferents and from recurrent inputs, so the main question to settle is how closely these neurons are coupled to the excitatory population. It has been hypothesized that these neurons can quickly and robustly balance out the excitation arriving from the LGN (Swadlow, 2003; Burkhalter, 2008), by linearly matching the response of the pyramidal cell population and providing contrast gain control to sparsify the activity and thereby sharpen the orientation tuning (Wilson et al., 2012). If so, then a non-linear and very slow response among this population should disrupt map development, which is a prediction that we will test in the model.

Linearity and time profile of inhibitory responses

To help understand how the PV population affects the behavior of the model, variants of the model were launched using different PV activation functions determined by two parameters. The parameter determines the linearity of the response, by adding an exponential term after the half-rectifier defined as the β term in equation 5.3. By varying this exponential term we could vary the response of the neuron from sub- to supra-linear values. To control the time profile of the response, a hysteresis term was added as well. After training, these model variants were assessed based on the metrics outlined in Section 5.1.2.

The results of the parameter exploration are shown in figure 5.3. They include a measure of stability over time, orientation selectivity, distortion of the retinotopy, the mean local homogeneity index (LHI) and finally the pinwheel density (ρ), used to assess the quality of the maps.

The immediately striking result is the narrow band of good (i.e., near π) pinwheel densities in the linear region of the response. Apart from that, we can clearly see that orientation selectivity, stability and the CoG distortion are strongly correlated with each other and all anti-correlated with the LHI. In particular, we can see increasing distortion of the retinotopy as the inhibitory population responds more strongly, which causes serious disruptions to the map quality. At the same time, sub-linear responses lead to very weak selectivity, which causes relatively low stability and high homogeneity.

In order to get a better overview of the relationships between these variables, Figure 5.4 shows the correlation matrix between the independent variables we are exploring and the metrics. The results highlight once again the strong relationship between orientation selectivity, stability, and the CoG shift and supra-linear PV cell responses. This view makes it clear that the PV cell exponent has a much greater effect on the

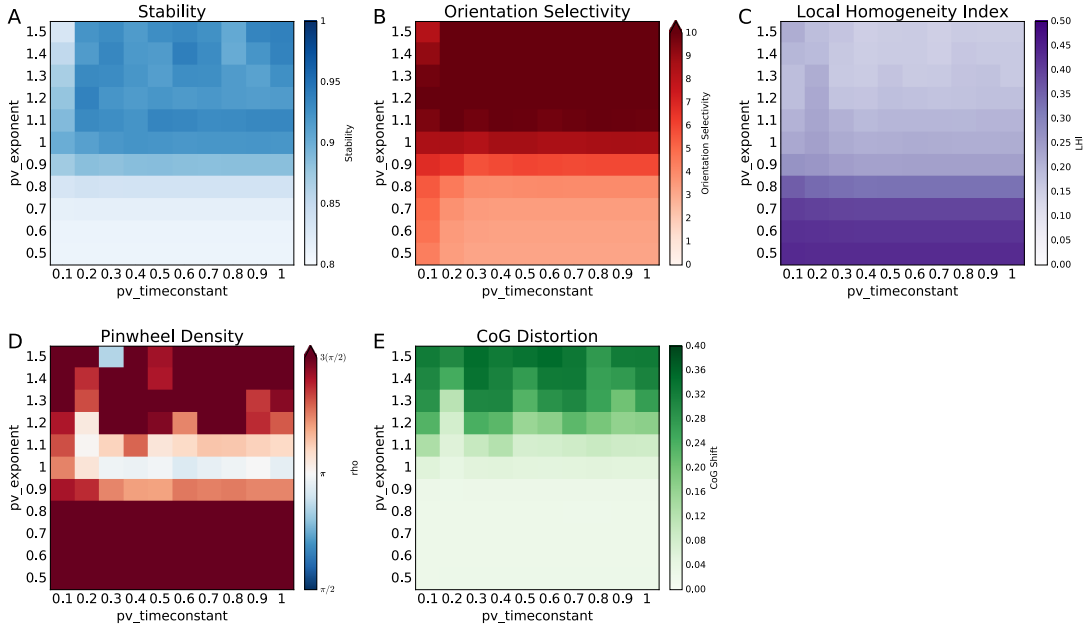


Figure 5.3: Parameter exploration to determine how the linearity and time constant of inhibitory responses affect development. A) A measure of stability of the model over time, integrating the similarity index of the model over time. B) Orientation selectivity in the model measured using orientation map measurements. C) The average local homogeneity index measuring the smoothness of the map at a particular spatial scale. D) Pinwheel density of the model, which should approach π for a perfect map. E) The Center-of-gravity (CoG) distortion provides a measure of how much the retinotopy in the model has been disrupted. The color mapping was chosen to match Figure 5.2 and simplify comparison. The model maps match biological results only when the PV exponent is 1 (linear response) and does not respond too slowly.

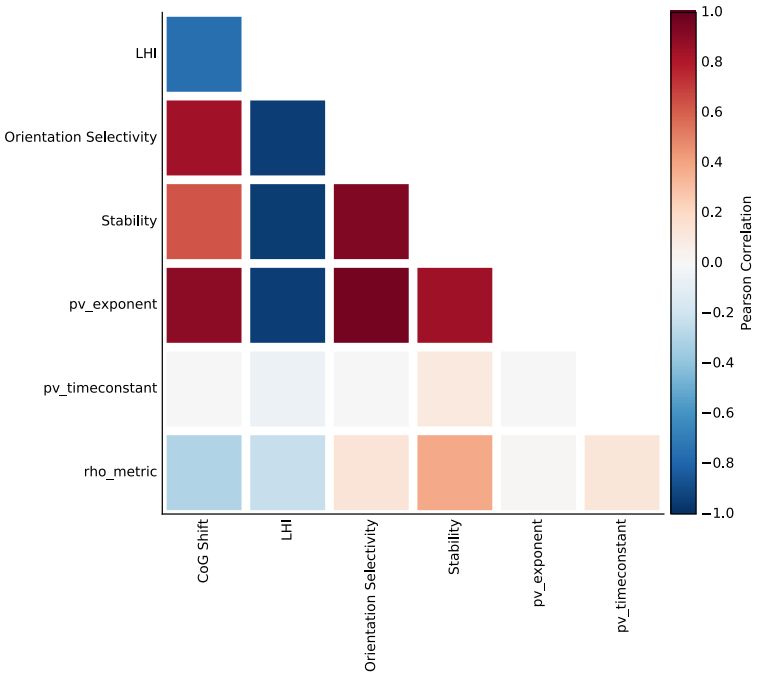


Figure 5.4: A correlation matrix of the various dependent and independent variables when varying the linearity and time constant of PV neuron responses. Plots the Pearson correlation of the various metrics with each other, showing how these metrics are related. Highlights the strong correlation between a super-linear response and high selectivity, stability and retinotopic distortions.

various metrics than the time constant of the response, which does not show significant effects until the response is very slow.

Overall we can conclude that a linear response, but not necessarily a very fast response, is required to drive the development of a high-quality orientation map. We will attempt to disentangle the causal relationships between these variables in the discussion. Meanwhile, however, we have a strong claim that we can develop high-quality orientation maps in the model. As a next step we will investigate how this model compares to the SCAL model.

5.2.2 *The SEPI Model*

The idea behind developing a model with distinct excitatory and inhibitory populations was based on three main observations. The first is straightforward: we know that in biology a single class of neurons will essentially never provide excitation at some synapses and inhibition at others, a principle known as Dale's law. From this perspective the SEPI model already provides an advantage over the SCAL model, showing how SCAL could be implemented in neural circuitry. However we also wanted to see whether this model could improve the diversity in responses and more

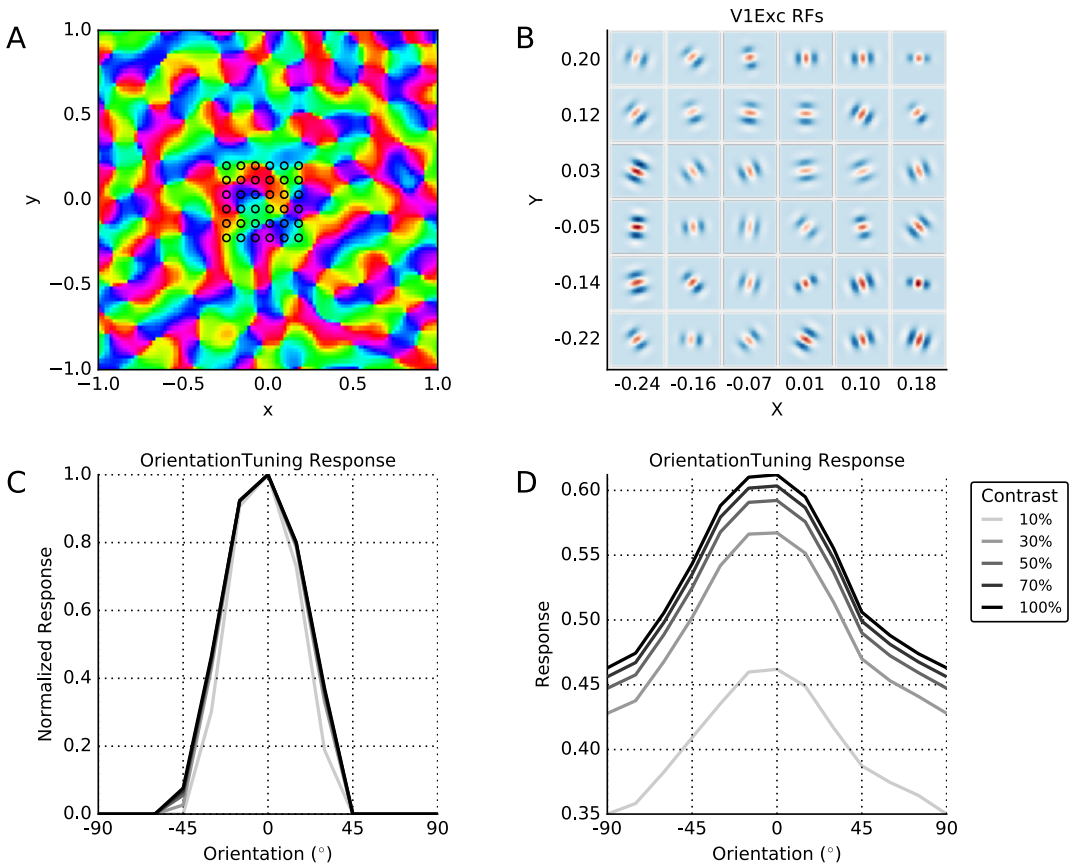


Figure 5.5: SEPI model orientation tuning properties after presenting 20,000 oriented Gaussian patterns. A) Orientation map measured by presenting sine gratings at the optimal spatial frequency to the model, along with the locations of the receptive fields shown in B. B) Gabor-fits to receptive fields measured using sparse random noise. C) Orientation tuning curve of a single excitatory neuron across contrasts, demonstrating largely contrast-independent orientation tuning. D) Orientation tuning curve of a PV neuron displaying a weakly tuned response, with a certain baseline of activity which is only weakly modulated by orientation.

accurately capture contrast-dependent responses in the model, while continuing to develop high quality maps and exhibiting robust and stable self-organization.

The linearity analysis already confirmed the model could produce high-quality orientation maps. In 5.5 this is further confirmed with a high-quality map, and clearly defined receptive fields, which exhibit contrast invariant orientation tuning, which rivals previous models. Without going into detail, we can also confirm that the spatial calibration still holds, with 3.5 cycles per visual degree, well within the acceptable margins that were established in Figure 4.2.3.

The main thing we are interested in, however, is how decoupling the two populations affects contrast-dependent size-tuning shifts, which rely on strong divisive inhibition. Recall for example that the GCAL model showed very little contrast-

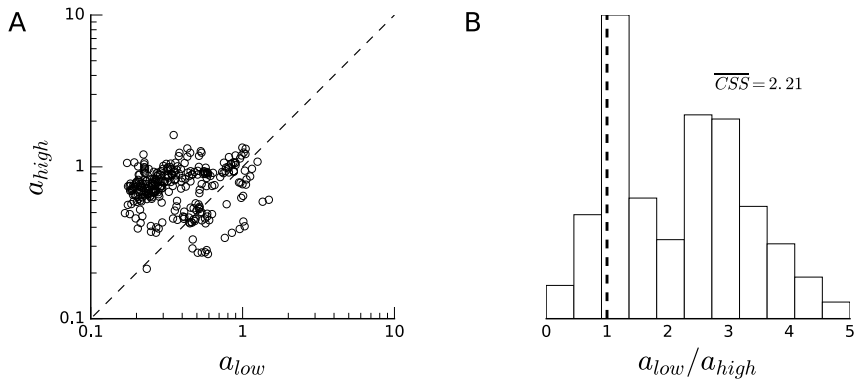


Figure 5.6: Contrast-dependent shifts in size tuning as estimated by the DoG model. A) Spatial constant of the excitatory center of the DoG model at low vs. high contrast. B) Distribution of contrast-dependent shifts, showing a much better match to the experimental data than GCAL or SCAL (see figure 4.12).

dependent size-tuning changes due to the inhibition being purely subtractive. The SCAL model demonstrated a strong shift, but it lacked the diversity observed in the experimental data. Repeating this same protocol in the SEPI model as is shown in figure 5.6, demonstrates a slightly larger shift in contrast-dependent size tuning, which is in fact exactly in line with other measurements performed in macaque ($\approx 2.2\times$). Even more strikingly, the model now exhibits diversity in size tuning shifts that provide a far better match to experimental results (Figure 4.12). This likely can be attributed to the greater strength and variation of inhibition that is made possible by partially decoupling the excitatory and inhibitory population.

Introducing a separate inhibitory population also allows us to compare the spatial and orientation profile of inhibitory neurons more realistically, since they no longer reflect the activation of the joint excitatory/inhibitory population. When comparing Figure 5.7 to 4.19 we can see that the neurons are no longer as strongly biased toward the preferred orientation of the excitatory neuron, which is a direct result of much broader orientation tuning in the inhibitory PV population. Additionally the inhibitory connections seem to exhibit longer tails, particularly in the oblique and cross-orientation distributions, which more closely matches the experimental data. The major difference when compared to the experimental data seems to be in the relative weights of connections targeting iso-, oblique- and cross-orientation domains. Since the model distributions represent the actual weights while experimental data just uses the total count of boutons it is unclear whether this difference arises simply because it is not possible to accurately measure connection strengths in experiments. Overall these orientation distribution of inhibitory connections in the SEPI model more closely reflects the experimental data when compared to the SCAL model.

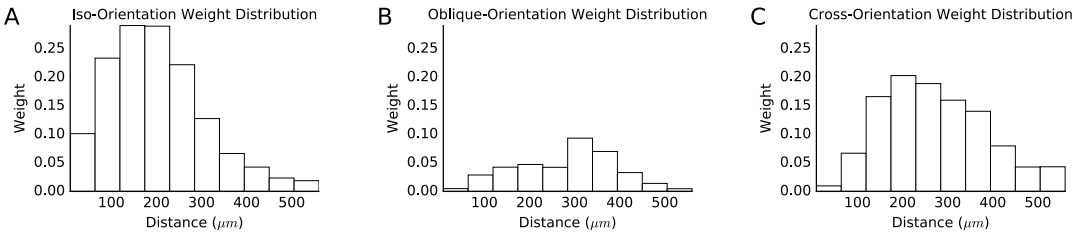


Figure 5.7: Spatial distribution of weights targeting A) iso-orientation regions, B) oblique-orientation regions and C) cross-orientation regions. Inhibitory neurons in the SEPI model have a much broader distribution of connections in the orientation domain, more closely matching Kisvárday et al. (1997) than the SCAL model.

Finally, we can confirm the spatial and orientation tuning properties of the SEPI model in another way. By computing the local homogeneity index as we did above, we can compute the smoothness of the map at a specific spatial scale. Since we know the spatial scales in the model, we can directly compare the distribution of the LHI to experimental data. The σ value was therefore set to 180μm and plotted against the orientation tuning width of all neurons in a central region of the simulated map. The resulting scatter plot is shown in Figure 5.8, compared directly to equivalent measurements in macaque and cat visual cortex. Both the distribution of tuning widths and LHI values closely match the experimental data from macaque, but not cat, which is exactly what was desired.

While the model now exhibits many of the properties that were lacking in the SCAL model, one last issue remains. A large reason why the SEPI model was introduced in the first place was to come up with a model that would exhibit realistic long-range connectivity, which is presumed to underlie various contextual modulation effects. In this next section the effect of introducing long-range excitatory on the development of the model will be explored to see whether this model can robustly develop even in the presence of the long-range patchy connections that are so characteristic of pyramidal cells in layer 2/3 of the cortex.

Robustness to long-range excitation

At this point two separate but very similar models have been introduced, SCAL and SEPI, which differ mostly in that the latter has distinct excitatory and inhibitory populations. The major reason why the populations were split in this way is to match the anatomy, but does this change actually buy us anything from a functional perspective? In order for long-range excitation to mediate contextual modulation phenomena, these projections need to be able to meaningfully modulate neuron activity at long distances. However, from a theoretical perspective, introducing strong long-range excitatory connectivity in a model seems problematic for a number of reasons.

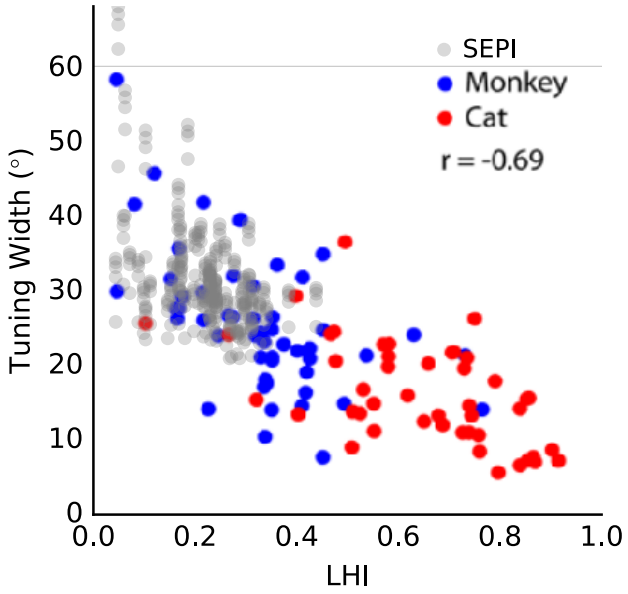


Figure 5.8: Local homogeneity index against tuning width in the SEPI model compared to experimental results from macaque monkey and cat. Reproduced from Nauhaus et al. (2008). Provides another confirmation of appropriate spatial calibration, as the LHI is dependent on the scale of orientation columns.

The most obvious and problematic result of strong excitation is that it can lead to runaway, 'epileptic' activity in the cortex, which results in severely disrupted development as we already saw when inhibitory neurons responded sub-linearly. More subtly however, by introducing long-range correlations into the activity we can disrupt the retinotopy in the model, which alters the sparsity constraints on the model and can have profound effects on self-organization.

Therefore assessing how the models hold up in the presence of strong lateral excitation under varying contrast levels will help us establish how robust the model is and lay the groundwork for a model that exhibits both long-range excitation and suppression. Thus we will compare how the SCAL and SEPI models deal with varying levels of contrast and lateral excitation to investigate whether separating the populations has any benefits in addition to being more anatomically realistic.

SCAL The results of this analysis, once again assessed by the list of metrics described above, are shown above in 5.9. The first thing to note is that increasing lateral excitation leads to immediate disruptions to the self-organization of the model until we reach a transition point at which the model fails to self-organize completely, presumably the point where the additional excitation can no longer be balanced by increases in the gain of inhibition. Indeed, the CoG distortion increases with increasing lateral excitation, suggesting that activity is allowed to propagate horizontally from

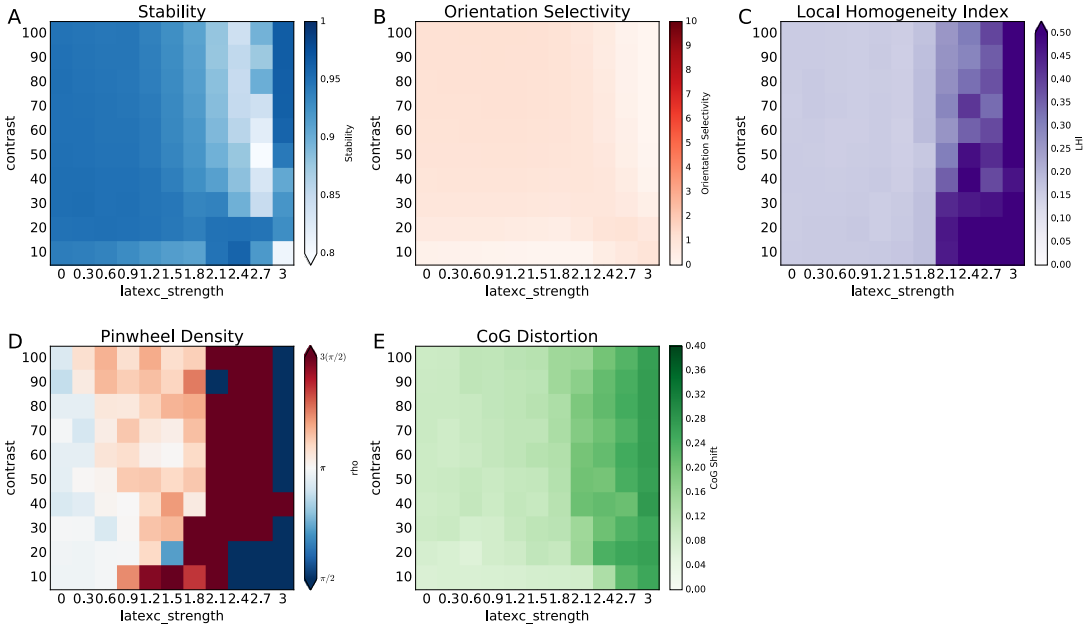


Figure 5.9: Parameter explorations of three separate metrics of orientation map development using the SCAL model from Chapter 4. Plotted on x and y are the strength of long-range lateral excitation and the stimulus contrast during training. The metrics show **A** the orientation map stability metric, **B** the average selectivity over the time course of development, **C** the average local homogeneity index providing a measure of map quality, **D** the ρ pinwheel density metric, which characterizes the quality of the map, and **D** the mean amount of Center-of-Gravity distortion providing a measure of retinotopic distortions. The model is highly robust to varying stimulus contrasts, at low levels of lateral excitation, but quickly deteriorates with increasing levels of excitation.

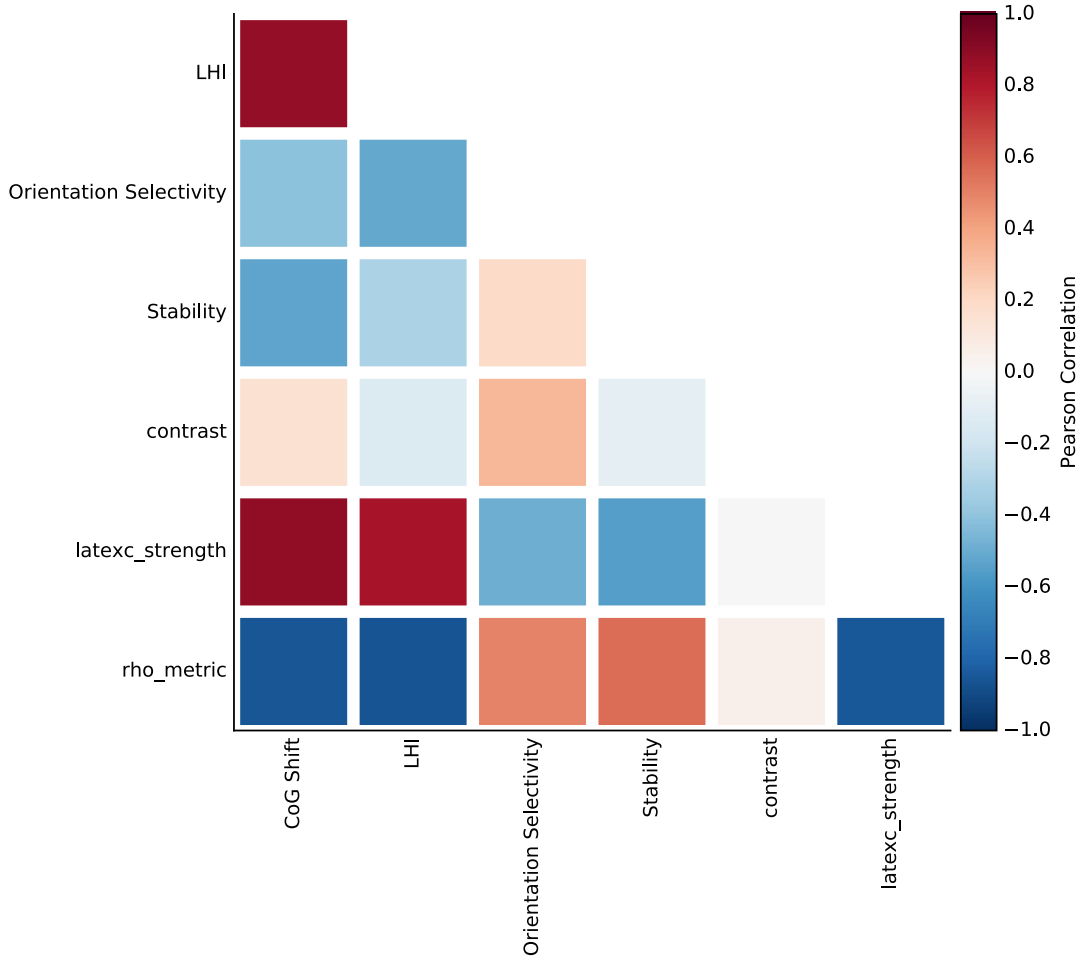


Figure 5.10: A correlation matrix of the various dependent and independent variables of the linearity analysis. Plots the Pearson correlation of each variable against each other variable, for the SCAL model from Chapter 4.

where it arrives in V1, leading to disruptions in retinotopy. Overall, we can see that adding strong lateral excitation to SCAL causes disruptions to the stability, orientation selectivity, pinwheel density, and CoG distortion, causing the eventual breakage of the entire model.

The correlation matrix shown in Figure 5.10 highlights what is going wrong as we increase the lateral excitation. It demonstrates a strong relation to CoG distortion and increasing homogeneity, with clear anti-correlation with orientation selectivity and stability. As a result we can conclude that the lateral excitation leads to disruptions in retinotopy and lower selectivity, which in turn disrupt the stability and pinwheel density of the model.

SEPI When applying the same analysis to the SEPI model (Figure 5.11), we can immediately see that the model shows very little disruption at equivalent levels of

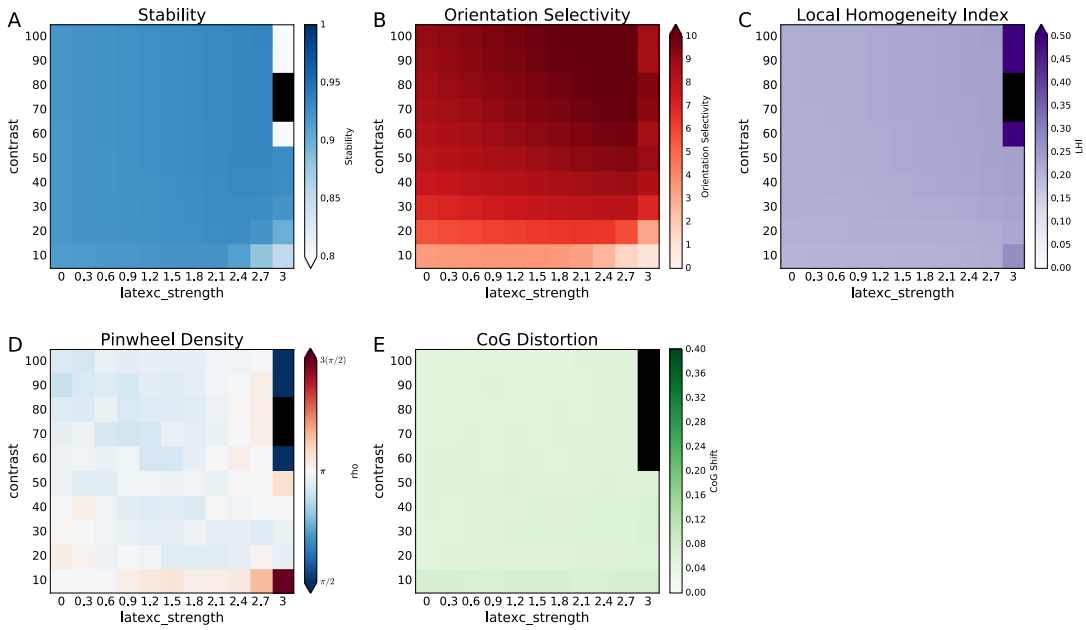


Figure 5.11: Parameter explorations of three separate metrics of orientation map development using the SEPI model. In comparison to the SCAL model, the model is far more robust to strong lateral excitation and maintains almost uniform stability across almost all explored parameter values. As in Figure 5.9 panel **A** shows the orientation map stability metric, **B** the average selectivity over the time course of development, **C** the average LHI providing a measure of map quality, **D** the ρ pinwheel density metric and **D** the mean CoG distortion. Uncoupling of excitation and inhibition allows the model to handle changes in parameter strengths in a robust and responsive way.

lateral excitation. The model shows basically no CoG distortion, and very little variation in pinwheel density, stability, and LHI. The orientation selectivity, conversely, actually increases with both excitation and contrast, opposite to what was observed in SCAL. Of course, even this model begins to break down under very high contrast and strong lateral excitation.

Overall, however, the SEPI model is significantly more robust to variations in the strength of afferent input and even lateral excitation, which is usually thought to cause problems for models that rely on competition between neurons to represent the visual space.

5.3 DISCUSSION

In this chapter we have explored how distinct excitatory and inhibitory modeled after the pyramidal and basket cells in the visual cortex interact to give rise to robust and

stable map formation and then confirmed the model still matches the other spatial and orientation tuning properties we carefully evaluated in the SCAL model.

5.3.1 *The role of inhibition in development*

In the first experiment we asked what happens when we decouple the responses of the inhibitory population from that of the excitatory population. In the SCAL model and all most previous SOM based models both excitatory and inhibitory cells were represented by the same neurons, which means their responses were coupled directly. While the time constant had very little effect on the organization of the model, the analysis highlighted how strong an effect a non-linearity in the activation function of the inhibitory population had. A sub-linear response was associated lower selectivity and stability, while a supra-linear response was associated with very high orientation selectivity but also greater distortion in the retinotopy.

From a theoretical perspective these results make sense: a linearly responding PV population is able to match the level of excitation that is arriving in the cortex from thalamocortical afferents and recurrent intra-cortical connections; shifting that balance in either direction will disrupt the development. Indeed, if we compare the effect of up-regulating the PV responses, we can see some parallels to what happens during pharmacological up-regulation of $GABA_A$ neurons by administering the $GABA_A$ agonist benzodiazepine (diazepam) (Fagiolini et al., 2004; Hensch and Stryker, 2004) and looked at ocular dominance columns rather than orientation columns. Much like in those experiments, the neurons become more segregated and selective when the inhibitory population was up-regulated in this way. The results predict that the orientation column spacing narrows with increasing inhibition, which does not match the results obtained for ocular dominance columns. This may be explained by the fact that the development of ocular dominance columns precedes that of an orientation map, and the diffuse and unorganized initial pattern of GABAergic connections acts to patches of correlated activity further apart rather than concentrating it in a smaller column.

The hysteresis term in the responses of PV neurons was not used in the final model for two reasons. Firstly it had little effect on the organization of the model unless very low time constants were used resulting in highly sluggish responses and potential decoupling of the excitatory and inhibitory populations. Additionally, since this population is modeled after the fast-spiking population, which exhibit very fast kinetics (Cruikshank et al., 2007; Gabernet et al., 2005), a more instantaneous response fits more closely with experimental observations about this population of interneurons.

Finally, we compared how the SCAL model compared to the SEPI model in terms of robustness to lateral excitation and varying levels of contrast. These results clearly showed that the SEPI model self-organizes for a much wider range of lateral excitation values, indicating it can deal with the destabilizing introduction of long-range excitation much better. This property is likely a direct result of separating the two populations. Although separating the populations allows these neurons to become uncoupled, it also ensures that the inhibitory population can effectively respond and cancel out the additional excitation arriving via long-range connections. In this way high activity among the excitatory population results in more effective suppression rather than complete destabilization of the network. However, at the very highest lateral excitation and contrast values tested here, even this model had trouble canceling out this additional excitation.

Based on the overall observations, it seems that the model's PV population provide a means for the neurons in a local region to compete to represent the stimulus, repelling neurons with similar but somewhat different feature preferences. This process results in a smooth local mapping of the feature preferences into an orientation map. Having neurons that respond so quickly and robustly ensures the development of a realistic orientation map even in the presence of other influences, though it also means that pathological changes in the PV response properties could severely disrupt the organization of the cortex.

5.3.2 *Feature tuning and inhibition*

When analyzing the model size and contrast response, we found very good agreement with experiment, and even found that the variability in responses was much more realistic compared to the SCAL model. By decoupling the excitatory and inhibitory neurons, we were able to achieve a greater contrast-dependent size-tuning shift, both qualitatively and and quantitatively much closer to the kinds of shifts that are seen in experiment (see Figure 5.6 and 4.12). This result fits very well with the idea that the PV neurons exert strong control over the response gain of principal neurons via a divisive mechanism (Wilson et al., 2012). Indeed, through artificial upregulation of the PV population, not currently shown, we could mirror the effects of decreasing the contrast, matching results found by Nienborg et al. (2013) in mouse visual cortex.

It is worth noting that the average magnitude of contrast dependent size-tuning shifts has not changed significantly between the SCAL and SEPI models. However there is far greater diversity observed between different neurons than was present in the simpler SCAL model. This diversity can also be observed in experiments, sug-

gesting something about the SEPI model captures this phenomenon more accurately. It is likely the result in a matching increase in the diversity of connectivity between neurons. Due to the separation of excitatory and inhibitory neurons into distinct cell classes, neurons can now receive varying levels of excitation and inhibition. If we accept the argument made in the previous chapter and put forth by Cavanaugh et al. (2002b) and Solomon et al. (2006), that the shift in size tuning emerges as a result of different excitatory and inhibitory gain values, this explanation makes perfect sense. Since neurons now have greater diversity in the excitatory and inhibitory inputs they receive, the ratio between these components also becomes more diverse, resulting in greater variance in the observed size-tuning shift.

One major question in the literature has been the tuning properties of PV neurons. A number of theoretical studies had suggested that untuned or broadly tuned inhibition in the visual cortex was required for the development of orientation selectivity or its contrast invariance (Somers et al., 1995; Troyer et al., 1998). The work by Cardin et al. (2007) challenged this assumption, because they found that the fast-spiking inhibitory cells in layer 4 of cat area 17 did not exhibit significantly broader tuning in the subthreshold response than the regular spiking excitatory population. However, they did find broader tuning in the spiking response in this population, and similar studies had found conflicting results (Hirsch et al., 2003; Nowak et al., 2008). In the SEPI model, the PV cells generally display orientation tuning, but they are overall significantly more broadly tuned than the excitatory population, which is largely driven by the fact that the modeled cells do not have any spiking threshold. Cardin et al. (2007) argue that studies such as Nelson et al. (1994), which demonstrated no widening in orientation tuning of excitatory cells under GABAergic inactivation, suggest that broad inhibitory tuning does not significantly contribute to sharpening of orientation tuning. However, this argument ignores the role of inhibitory neurons in shaping the responses and tuning of neurons during development, such that the instantaneous responses of the inhibitory population are no longer needed to significantly sharpen the tuning of excitatory cells, as their tuning has already been encoded in the feedforward connections (as noted by Miikkulainen et al. 2005).

Another difficulty is in determining layer-specific effects; e.g., Cardin et al. (2007) primarily measured from fast-spiking cells in layer 4. It may be the case that fast-spiking neurons in layer 4 are more closely involved in a push-pull inhibitory regime, which would require sharper orientation tuning (Hirsch et al., 2003; Hirsch and Martinez, 2006). Further investigation will be required to determine the precise properties of the PV-ir neuron population across layers in species with maps, in order to determine whether there are functionally distinct subgroups.

5.3.3 *Conclusion*

Overall, the results here indicate that the PV population is very well suited towards mediating the kinds of tasks that are important for robust development. Specifically, they can linearly track the responses of pyramidal cells in their vicinity by integrating feedforward and recurrent input and matching the levels of excitation in these neurons. Including PV neurons leads to a model that is highly robust to long-range excitation, but still exhibits all the desirable properties of the SCAL model, including contrast-invariant orientation tuning and the same spatial tuning.

At the same time, however, the model still has no way of mediating any true long-range orientation specific inhibitory effects. This is because the PV neurons are very broadly tuned and if allowed to develop only receive very weakly tuned long-range excitatory connections. Additionally, due to the strong inputs they receive and their internal dynamics, PV neurons are very fast in their response and adapt very quickly, making it difficult to envision how they could integrate activity that arrives over longer spatial and temporal scales.

6

LONG-RANGE INTERACTIONS IN THE VISUAL CORTEX

Patchy long-range lateral connections that link regions with similar visual preferences are one of the most intriguing features of the early visual cortex. These connections have been suggested to drive a wide range of contextual-modulation effects including pop-out, iso-orientation suppression, contour completion, and more (Gilbert and Wiesel, 1983; Hirsch and Gilbert, 1991; McGuire et al., 1991; Grinvald et al., 1994; Fitzpatrick, 2000; Hupé et al., 2001; Stettler et al., 2002). The most commonly reported effects particularly at high contrast are suppressive, and even though the connections are known to be excitatory and to primarily target other excitatory cells, some connections target local inhibitory neurons that provide strong di- or poly-synaptic inhibition (Hirsch and Gilbert, 1991; Weliky et al., 1995). While both the SCAL and SEPI models optionally support long-range excitation that could provide such suppression di-synaptically via the PV inhibitory cells, the suppression would not be specific for features like orientation, because the PV cell population is so weakly tuned.

On theoretical grounds, it has been hypothesized that the neural code in the cortex achieves a neural representation that is both sparse and reduces redundancy in the input (Barlow and Földiák, 1989; Olshausen and Field, 1996). Indeed, modulation from the non-classical receptive field has been shown to suppress redundant information in the input, sparsifying the representation (Vinje and Gallant, 2000). Here it is crucial that the target of the suppression is *redundant* activity; achieving sparsity by simply globally or randomly reducing activity will increase sparseness but lose information. For instance, in previous developmental models of V1 with long-range lateral inhibition, it was shown that both isotropic and feature-specific connections could achieve a given level of sparsity, but the isotropic connections did so by reducing information about the input patterns (Miikkulainen et al., 2005). For instance, if two unrelated contours are nearby in an image, isotropic suppression will eliminate the representation of the weaker one, achieving sparsity by losing information about the input, while feature-specific suppression will reduce the number of neurons responding to each contour separately.

These considerations suggest a difference between short-range interactions, where the cortex is trying to ensure it achieves maximal coverage of the feature space, and long-range interactions, where the neuron is competing to represent the stimulus efficiently, i.e., by using a sparse code. In the previous models, PV inhibition drove the

local competition, resulting in a smooth coverage of the visual space with orientation feature detectors, forming an orientation map. Since the PV neurons act over a local range and are not highly tuned in the orientation domain, they are generally isotropic in the spatial distribution and large enough to cover neurons with a wide range of feature preferences, and thus cannot mediate feature-specific modulation.

Interestingly, the next-most-common type of inhibitory interneurons has properties that do seem suitable for mediating specific modulation. The Somatostatin immunoreactive (Sst) neurons (Gonchar et al., 2007; Xu et al., 2010) are more numerous in layer 2/3, where the longest-range patchy connectivity is found, and they seem to respond strongly only under strong, sustained stimulation (Ma et al., 2011) or for very large stimuli (Adesnik et al., 2012). Additionally, their input scaling does not seem to be linear, with an accelerating response function driving much stronger responses when recruited consistently and robustly (as they would be for large or high contrast stimuli; Beierlein et al. 2003; Bartley et al. 2008; Tan et al. 2008).

On the basis of what we have learned about the PV population from the previous models and what we know about Somatostatin neurons from the literature, we therefore propose a general theory of how the circuit performs specific computations. Afferent input provides strong, low-latency excitation to the Pv-ir neurons in the thalamocortical recipient layer 4, which in turn act as both a feedforward inhibition and dynamic gain control mechanism on the broadly activated excitatory cell population. This process results in medium-range decorrelation of the neural activity, which allows short-range recurrent excitation to amplify the activity in a local column. The smooth local blobs of activity that emerge as the response settles then lead to the formation of a well-organized orientation map.

At the same time, Sst-ir neurons begin to integrate the local activity through the local and long-range orientation-specific lateral connections. If their inputs are sufficiently strong they will activate strongly, allowing this polysynaptic circuit to reduce long-range correlation in the input activity, further reducing redundancy. If they are only weakly activated, as would occur when presented with low-contrast or very sparse input patterns, long-range lateral excitatory connections are not outcompeted and the circuit can fill in weak or missing information based on past statistics. In such a regime, the differential recruitment of the two separate inhibitory populations would be responsible for a shift in cortical state from a mode of redundancy-reduction and feature discrimination to one of visual inference and feature completion. Note that because Hebbian learning is proportional to the activity levels, strong inputs should dominate the learning process, avoiding having the network learn significantly from the filled-in activations typical of responses to weak inputs.

In this chapter we extend the SEPI model by adding this second population of inhibitory neurons modeled after the Sst interneuron population. In doing so we will demonstrate how this second population modulates responses to large and high-contrast stimuli, providing feature-specific inhibition under those conditions, while only very weakly responding in low contrast conditions. This circuit provides a considerable advance over previous models of surround modulation, which typically hard-code the statistical relationships encoded in the lateral connections while explaining contrast-dependent changes through a higher inhibitory threshold that conflicts with the experimental evidence (Li, 2002; Schwabe et al., 2006). In this model, the contextual-modulation effects are an emergent phenomenon, arising directly from the correlations present in the training patterns (as transformed sub-cortically). Identifying and understanding the circuits involved in these processes is a fundamental challenge for neuroscience, and will hugely contribute to extending our knowledge of cortical information processing.

6.1 METHODS

In this section we will first introduce the organization and equations describing the LESPI model and then describe various measurements to apply to the model to characterize surround modulation effects.

6.1.1 *The LESPI model*

The **L**ong-**R**ange **E**xcitation, **S**st, and **PV** **I**nhibition (LESPI) model again builds on the previous models, this time introducing an additional population of inhibitory neurons. These neurons receive no direct afferent input, being entirely recurrently driven via short- and long-range projections from the two other V1 cell types. The diagram in Figure 6.1 shows how the three populations of V1 neurons are connected.

The newly introduced Sst population differs from the PV population in several important respects. First of all, it receives no afferent connections from the LGN, as this population is most numerous in the supra-granular layers, which do not typically receive direct afferent input. Secondly, this population has a much more restricted spatial profile, being not nearly as extensive laterally as the PV-expressing basket cells. Most importantly, however, to model the facilitating response properties of lateral synaptic connections contacting this population, the model Sst neurons have an exponential non-linearity in their response. This property allows Sst neurons to activate weakly in low-contrast conditions, while providing strong suppression when the input is strong or dense enough. Finally, to model the more sluggish responses

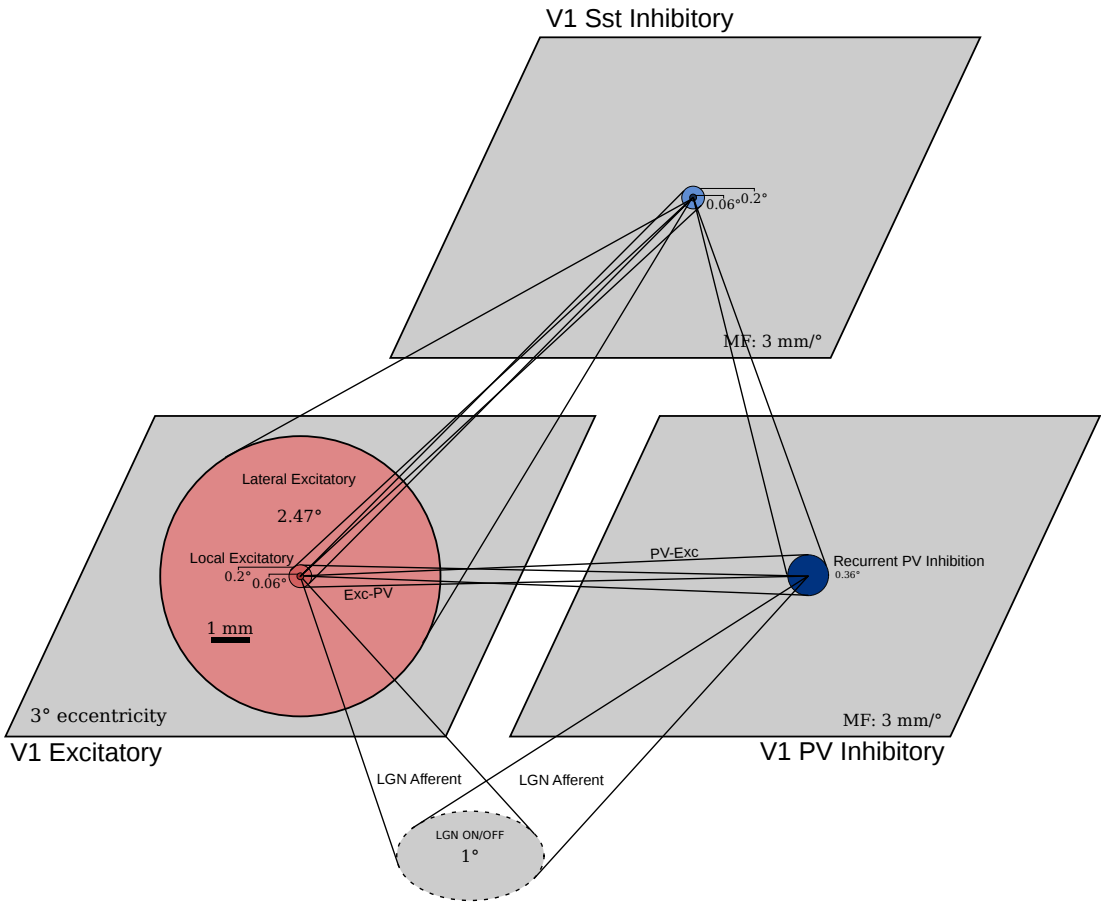


Figure 6.1: Diagram of the LESPI V1 stage of the model showing the spatial scales of the various excitatory (red) and inhibitory (blue) connections. Dark colors indicate the kernel radii, while lightly shaded regions indicate kernel cut-off extents.

typical of Sst neurons, a hysteresis term is added. In this way this population will respond strongly after sustained and strong local and long-range stimulation.

The full set of parameters for the SEPI model are listed in Appendix A.2.

Excitatory Activation

The activity of the excitatory population is computed from all of the incoming projections, with the Sst population acting as a multiplicative factor modulating the response of the excitatory neurons:

$$\eta_E = \frac{\eta_A + \eta_L}{1 + \eta_P}\eta_{SM} \tag{6.1}$$

where η_A is the LGN afferent activity, η_L the local excitatory contribution, η_P the PV inhibitory contribution, and the surround-modulation term η_{SM} is defined as:

$$\eta_{SM} = 1 + \eta_{HE} - \eta_S \tag{6.2}$$

where η_H represents the long-range horizontal excitatory contribution and η_S is the Sst inhibitory contribution. In the standard SEPI model the η_{SM} term simply reduces to 1, eliminating all long-range interactions, while in the long-range SEPI model $\eta_{SM} = 1 + \eta_{HE}$, with only long-range excitation. Here LESPI's surround-modulation term provides gain when excitation exceeds inhibition, and shunting inhibition when the reverse is true.

Inhibitory Sst Activation

The activation of the inhibitory Sst population is given simply by the summation of the local and lateral excitatory projection activity:

$$\eta_S = \eta_L + \eta_{HE} \tag{6.3}$$

The η_{HE} term also has hysteresis and an exponential activation function applied to it, such that its contribution is given by:

$$\eta_{HE}(t + \delta\eta) = (\eta(t) + \tau\big[\eta(t + \delta\eta) - \eta(t)\big])^{\beta} \tag{6.4}$$

where $\tau = 0.2$ and $\beta = 3$. This combination of hysteresis and an exponential nonlinearity allows the lateral connections to integrate over time, accelerating its response under strong and prolonged activation, which closely models what has been observed in experiments (Beierlein et al., 2003; Bartley et al., 2008; Tan et al., 2008).

Mechanisms

The mathematical effect of long-range horizontal connections is controversial and complex. The connections typically target distal dendrites rather than the soma, which would suggest a subtractive model. However, there is evidence for multiplicative effects as well (Wilson et al., 2012), including strong voltage dependence (Hirsch and Gilbert, 1991). The connections may act multiplicatively by locally gating excitatory horizontal and feedback inputs at the dendrites (Ma et al., 2011; Gentet et al., 2012). Theoretical studies also suggest that active dendritic spike backpropagation can lead to multiplicative increases in gain, while reduction in spike backpropagation can lead to divisive scaling of the firing rate (Mehaffey et al., 2005). Modeling the net effect of long-range input as a single multiplicative term, which can act divisively given sufficient inhibitory input, therefore provides a convenient but highly simplified abstraction of long-range interactions.

The overall operation of this circuit can be summarized by the schematic in Figure 6.2. The diagram shows the preferential linking of columns with similar orientation preference via patchy long-range connections and the long-range integration of Sst neurons, which receive both local and lateral input from pyramidal neurons, providing di-synaptic inhibition to the local region. Furthermore it suggests that under low-contrast conditions the Sst population is effectively disabled, activating only under high contrast input.

Sparsity

In the current form of the model, both afferent and lateral connection fields are simulated as dense projections of weights in a local region. However, real lateral connections only form patchy connections in the cortex. Having these diffuse connections means that only a fraction of the synaptic weight is concentrated at the patchy terminals. Informally, we have confirmed that the model self-organizes well using a simple sparse sprouting-and-retraction algorithm, but this sparsification algorithm would need to be documented, explained, analyzed, and validated separately before being applied here. Since that complexity would be distracting for our current modeling goals, we instead train the model with dense projections and then subsequently sparsify it by thresholding all weights below the 90th percentile. As shown in the results section below, this process leaves only patchy connections that strongly resemble experimental plots of patchy long-range connections in layer 2/3 of macaque and tree shrew, extending up to 4 hypercolumns from the neuron.

Sparsifying the weights in this way has the potential to alter the activity but since the connections mostly have a modulatory influence this does not significantly

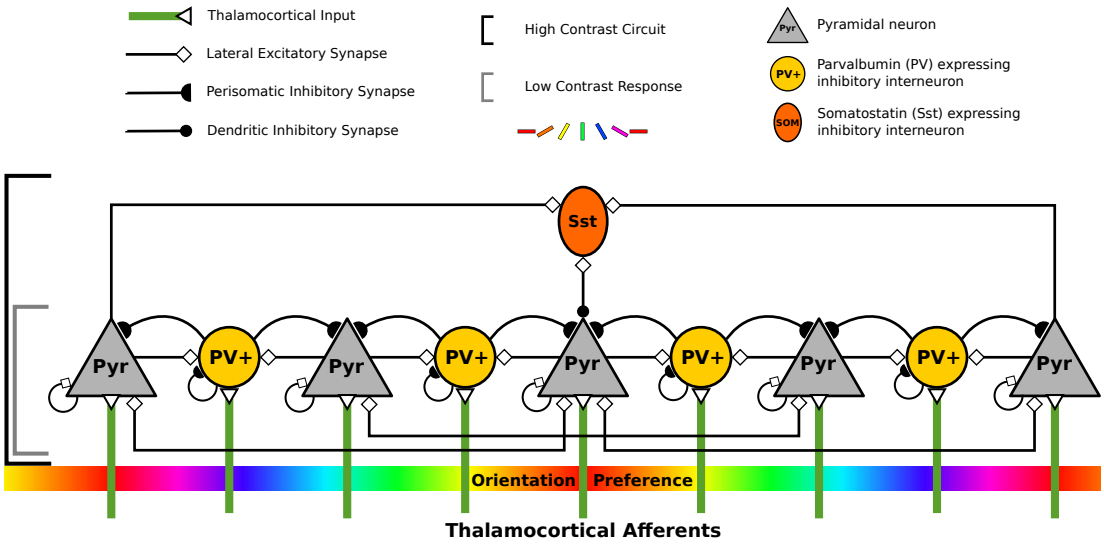


Figure 6.2: High-level circuit diagram of the LESPI model. The schematic represents multiple hypercolumns varying smoothly in their orientation preference. PV neurons provide feedforward and recurrent inhibition to neurons in the local region, regardless of their orientation tuning. Long-range excitatory connections link pyramidal cells with similar orientation preference, while the Sst population integrates both local and long-range input to provide tuned orientation suppression under high-contrast input.

change network responses especially in the initial response. However influences on the overall tuning and therefore the map structure cannot be ruled out. To eliminate these issues future work should therefore pursue an approach that employs connection sparsity from the start.

6.1.2 *Visual statistics*

In the models simulated in previous chapters, we have largely ignored the effect of visual statistics, instead using only simple artificial training patterns like 2D Gaussians that are always brighter than the background. These patterns were used because the models develop simple cells, strongly tuned for spatial phase. Patterns like natural images have a variety of spatial phases, such as light areas surrounded by dark and dark areas surrounded by light. When given such patterns, SCAL or SEPI simple cells with opposite phase preferences (swapping ON for OFF excitation) will be anti-correlated and thus far from each other on the cortical surface, disrupting the maps for orientation and retinotopy (Miikkulainen et al., 2005). In contrast, animal maps in the superficial layers used in imaging studies typically contain complex cells not highly selective for phase, which modeling suggests will lead them to develop

smooth representations for orientation and retinotopy regardless of the training patterns (Antolík, 2010).

To study contextual modulation, in this chapter we will need to use input patterns that have long-range correlations, instead of localized 2D Gaussians confined to the width of a receptive field. Otherwise, no long-range connectivity will develop by Hebbian learning. We will test the effects of a variety of natural image datasets as well as synthetic stimuli that include long-range correlations, while bearing in mind that the map patterns when trained on natural images should not be expected to closely follow results from animals, due to the phase dependence. Future versions of the model can include complex cells to remove this restriction, at a cost of greatly increased computational and memory requirements and additional parameters (Antolík, 2010).

Natural stimuli

The visual statistics in different natural stimuli can vary considerably. For instance, man-made objects often have very different statistics from natural objects. Visual scenes in nature typically contain smoothly varying somewhat circular contours, while man-made objects exhibit more straight and perpendicular lines (Perrinet and Bednar, 2015). Experiments suggest that humans can rapidly classify images of animals and non-animals based solely on the co-occurence statistics of visual contour elements in the image (Serre et al., 2007a; Perrinet and Bednar, 2015). Here we will make use of the various image datasets used in these studies as training data for the model, allowing us to compare their effects on contextual modulation.

The image datasets used for these purposes include the target dataset used by Serre et al. (2007a), consisting of animal images, as well as multiple datasets obtained by James Bednar in 2010 meant to replicate the visual experience of a developing animal, both in a laboratory environment, dominated by high-contrast bars of the animal cages, and a natural environment dominated mostly by grass, leaves, and trees.

6.1.3 *Surround modulation*

To reveal the effects of long-range connections, we will replicate a series of surround-modulation measurement protocols. In addition to the simple area-summation-curve measurements described above, we test two further protocols. In particular, we are interested in how stimuli outside the receptive field of a set of neurons can affect their response, and how those effects vary with contrast. For that purpose we replicate two measurement protocols, a simple annulus-based contrast-suppression measurement

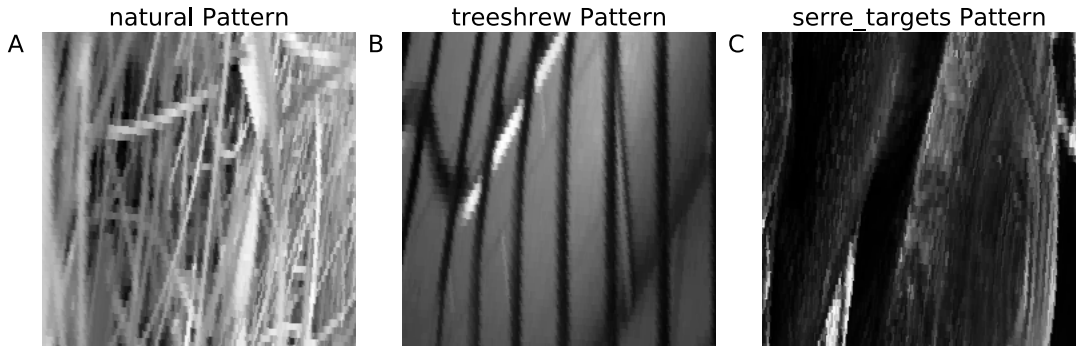


Figure 6.3: Example image patterns sampled from different image datasets and then randomly positioned and rotated. A) Forest and grass from the outdoor dataset taken in Gif-sur-Yvette. B) Inside of a tree-shrew cage taken in David Fitzpatrick's animal rearing facilities at Duke University. C) Serre et al. (2007a) target patterns of animals and natural landscapes.

(Jones et al., 2002), and a protocol using target and flanker stimuli (Kapadia et al., 1995).

Orientation-contrast suppression Orientation-contrast suppression is perhaps one of the most well-studied surround-modulation paradigms. The measurement involves presenting a central masked sine grating, optimized to the preferred size, frequency, and orientation of the neuron being measured. Once the baseline activity has been measured, an additional surround annulus with the same frequency is added and varied in contrast, size, and orientation to measure orientation-dependent interactions between center and surround (as shown in Figure 6.4). In in-vivo experiments, these measurements are performed using drifting gratings covering all spatial phases. Here we are only simulating simple cells, and so the gratings are optimized for the preferred phase of each neuron instead.

The center stimulus and the annulus were separated by 0.25°, which is smaller than the size of the separation employed by Jones et al. (2002), largely because the modeled cortical area is limited. However the size of the center and surround stimuli were roughly the same scale with a 1° and 3.5° diameter respectively.

In order to quantitatively assess the orientation-contrast suppression we used the orientation-contrast suppression index defined as:

$$\mathrm{OCSI} = \frac{R_0 - R_{\frac{\pi}{2}}}{R_0} \tag{6.5}$$

where R_0 represents the response at the preferred orientation and $R_{\frac{\pi}{2}}$ at the orthogonal orientation.

Figure 6.4: Orientation-contrast stimulus measuring modulation by a sine grating annulus on the response of a central neuron responding to a central sine grating disk of the same spatial frequency. Stimulus is varied by center and surround contrast, surround orientation and the offset between the central disk and the surround annulus.

Pop-out Another well-studied effect in the surround-modulation literature is "pop-out", where a visual stimulus embedded within a number of dissimilar features strongly stands out. It is thought this effect relies on similar features suppressing each other, leaving the dissimilar unaffected and therefore "popping-out" of the background (Kastner et al., 1997). In order to test this surround-modulation effect we generate visual patterns containing a target bar stimulus surrounded by a circle of flanker stimuli, which are either co-aligned with the target (iso-condition) or orthogonal to the target (cross-condition). The response to the targets and the flankers is then computed as the mean activity of the non-zero responding units with two annuli, one covering the central 1 degrees and the other covering the remaining simulated area.

6.2 RESULTS

As previously discussed, the surround-modulation effects in the model are highly dependent on the statistics of the visual input used for training. The following analyses all use either the "natural" or the "treeshrew" (cage) dataset shown in Figure 6.3. This is because the Gaussian patterns are drawn randomly and do not contain any long-range correlations the model could learn and encode in its lateral connectivity.

Before characterizing the responses to more complex patterns, we will once again show how the response to a simple Gaussian pattern evolves over time before and after development, which is shown in Figure 6.5. When comparing this response to

the SCAL model (see Figure 4.9) it is immediately noticeable that the model exhibits significantly more suppression, which can be attributed to inhibition from the additional inhibitory population (having kept the same strength for the first population when the second was added). Later on we will see that the timecourse of responses is highly stimulus dependent, resulting from the richer set of interactions when compared to the SCAL and SEPI models.

Secondly it is important to get an understanding of the map organization and how that may influences the results. In Figure 6.6 an orientation map and the afferent weight patterns for the simple Gaussians and two natural image datasets is shown. Visual comparison immediately highlights the disrupted organization of the orientation map when the model is trained on natural images. The afferent connection patterns highlight what is causing this disruption. In the natural image trained models the afferent connections form a phase map, flipping between ON-center and OFF-center receptive fields. This is because these patterns are locally anti-correlated and therefore segregate into separate regions. Additionally lower spatial frequency preference in the treeshrew case can be observed, made evident by the two-lobed, rather than three-lobed receptive fields.

6.2.1 *Orientation contrast suppression*

Orientation-contrast suppression is one of the most well-studied paradigms in the surround-modulation literature. One of the most thorough analyses was performed by Jones et al. (2002), varying the relative size and offset of the center stimulus and surround annulus. The measurements were performed in the model trained on the "natural" dataset, which does not have as highly biased statistics as the laboratory datasets.

As a first step, the orientation-contrast curve was measured for a number of randomly chosen neurons under high-contrast conditions. Non-responding units were rejected, making up roughly 29% of the neurons. This was likely due to imperfect fitting of the optimal parameters. By normalizing the response and averaging the curves and computing the standard error, an average contrast suppression curve was computed and can be seen next to the orientation-tuning suppression curve from a single neuron measured by Jones et al. (2002) in Figure 6.7. The result shows that the model neurons have realistically strong iso-orientation suppression with a well defined shape, which does not vary considerably in width.

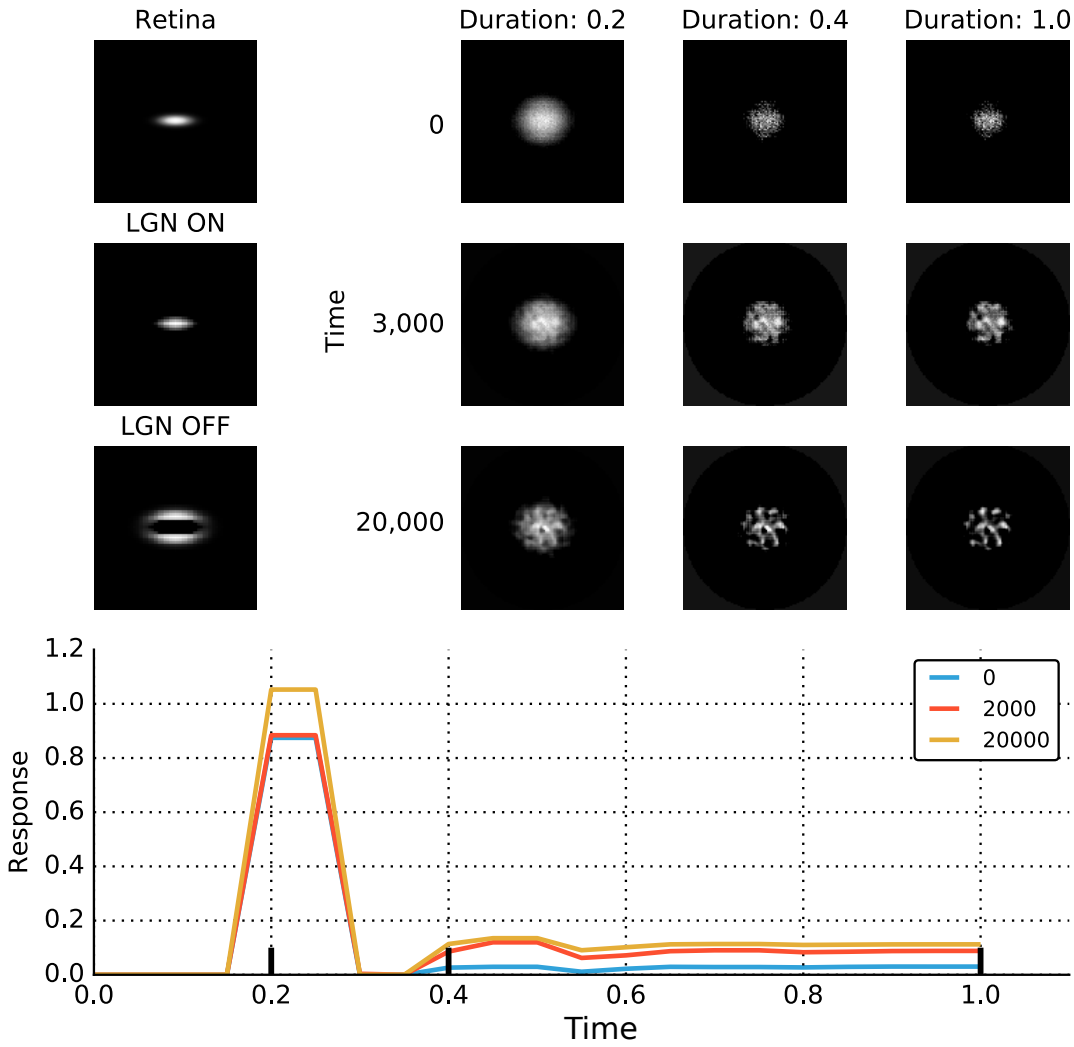


Figure 6.5: LESPI responses to a simple Gaussian pattern before and after self-organization on the natural dataset. Top left shows the input pattern and the LGN ON/OFF response to the pattern. Top right shows the evolving patterns of activity before, during, and after development. The response of the central neuron at the different stages of development is shown at the bottom. Neurons exhibit strong suppression after the initial onset, depending on the stimulus being shown.

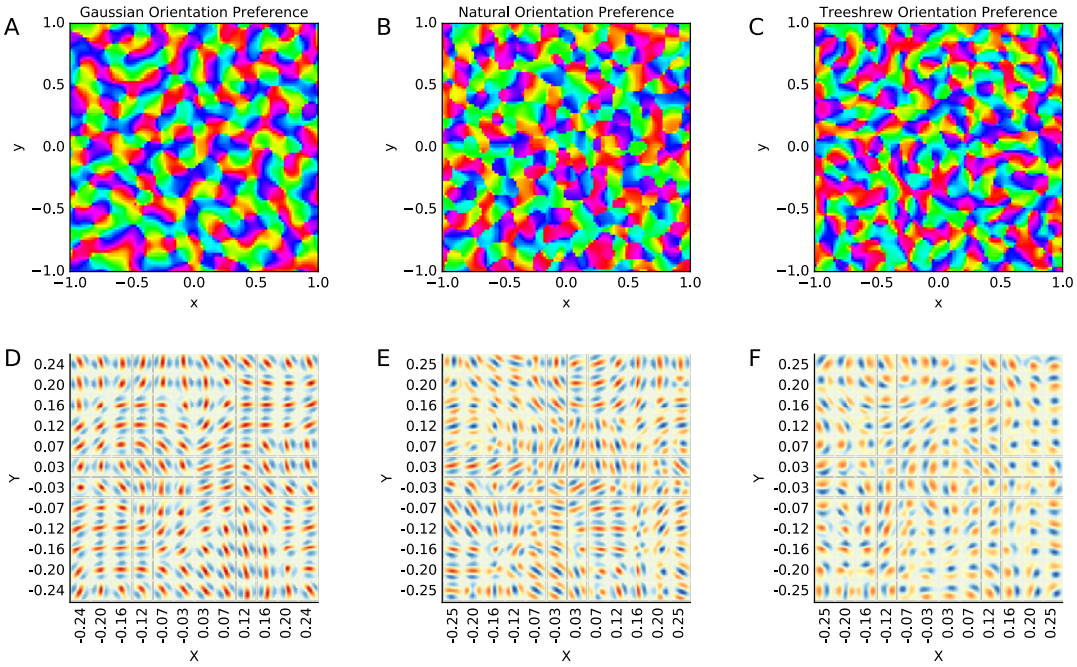


Figure 6.6: LESPI model orientation map when trained on Gaussian patterns and treeshrew and natural image datasets. A, B, C) Orientation maps showing good organization in the Gaussian case, and phase disrupted organization in the treeshrew and natural conditions. D, E, F) Afferent connections for the three conditions, highlighting the phase organization in the natural condition and spatial frequency differences in the treeshrew condition.

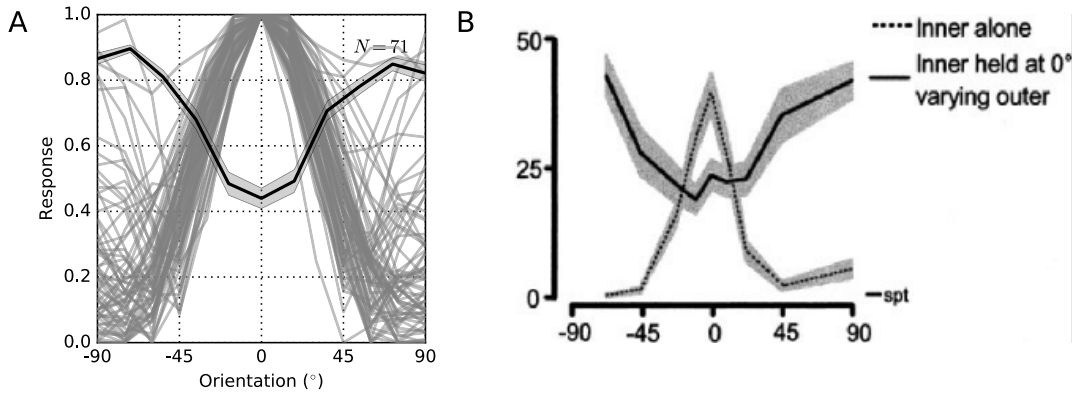


Figure 6.7: Averaged orientation-contrast suppression curve compared against Jones et al. (2002) example curve. A) Shows the normalized and averaged orientation contrast suppression curve from 71 neurons against the individual orientation tuning curves of those neurons (spread represents standard error). B) A sample orientation-contrast suppression curve reproduced from Jones et al. (2002). Results from model trained on natural dataset.

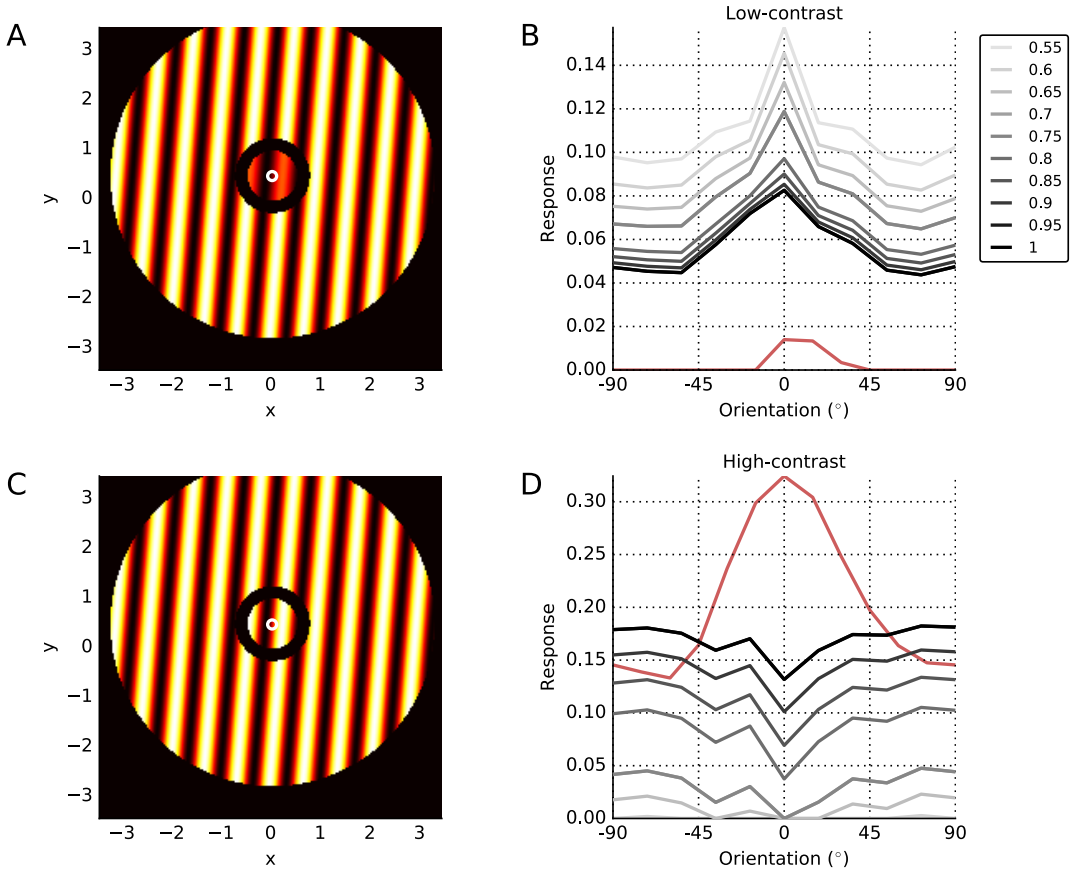


Figure 6.8: Dependence of orientation-contrast suppression on local and global contrast. A, C) Low- and high-contrast orientation-contrast suppression patterns. B, D) Corresponding orientation-contrast suppression curves showing facilitation at low contrast and suppression at high contrast, respectively. Results from model trained on natural dataset.

Contrast Dependence

In addition to measuring the classical high-contrast suppression curve, the equivalent measurement was performed for an individual neuron under both low and high-contrast conditions. The orientation contrast tuning curves were measured at each settling step of the model and are shown in Figure 6.8. As these results clearly show, the neuron exhibits very different effects under the two conditions. When the local and global contrast is low, the response of the neuron is enhanced, while high-contrast stimulation demonstrates the same suppressive effect that is evident in the averaged suppression curve in Figure 6.7.

In order to find the cross-over point at which facilitation turns to suppression, the contrast of both the center and surround pattern was varied and the OCSI was calculated for different durations. The corresponding plot is shown in Figure 6.9,

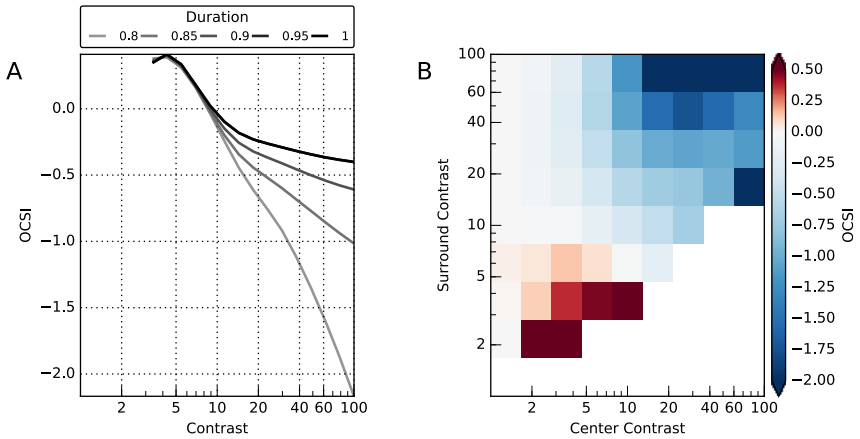


Figure 6.9: The relationship between contrast and orientation contrast suppression. A) Effect of varying center and surround contrast on the OCSI, demonstrating a shift from facilitation (positive OCSI) to suppression (negative OCSI). B) Effect of varying center and surround contrast independently, demonstrating that both local and surround context influence contextual modulation. Results from model trained on natural dataset.

highlighting again that suppression weakens as the response is settling, but more importantly clearly shows that the inflection point between excitation and suppression for this particular neuron lies at about 9% contrast.

This analysis also demonstrates how the suppression varies over the time course of the response, beginning shortly after onset (which occurs after a duration of about 0.25 and is excluded from analysis) and peaking at around 0.7 before weakening again. Since this model has not been temporally calibrated, it is hard to relate these times to experiments. Investigating the precise time-course of the different cell classes will make it clearer how the contrast dependence of the surround modulation emerges from the circuit.

Timecourse

In order to get a better understanding of how the different cell classes and synaptic projections interact to give rise to the contextual-modulation effects seen above, the timecourse of the activity can be visualized. The timecourse of the responses under four different conditions is shown in Figure 6.10. The top-row shows the response under low-contrast conditions for the iso-orientation and cross-orientation condition and highlights how the iso-orientation facilitation emerges. Comparing A and B it is clear how the lateral excitation increases the excitatory response in the iso-orientation condition but activates more weakly in the cross-orientation condition. Similarly in the high-contrast case, the Sst population activates much more strongly in the iso-

orientation condition (C) than in the cross-orientation condition, as expected due to the feature selectivity of the Sst neurons.

The model only has a rough temporal calibration, which makes it difficult to compare. However one iteration in the model corresponds to roughly 250 ms of response, and the lateral delay is calibrated to a maximum propagation speed of about 0.3 m/s, which has been confirmed in experiments (Hirsch and Gilbert, 1991). The time course of the iso-orientation surround suppression and facilitation arrive roughly 40-60 ms after the onset of the excitatory response, which is also roughly in line with experiments (Hupé et al., 2001; Bair et al., 2003). This indicates that lateral connections can realistically exert iso-oriented surround suppression within the spatial ranges explored here over the time-course that has been observed in experiments. However, over larger spatial scales it is likely that feedback connections would have to be involved as Bair et al. (2003) have argued previously.

6.2.2 *Size tuning*

The effects of both distance and contrast can best be explained when looking at the size-tuning curves of all three populations. A representative example of size-tuning curves measured from all three populations at a single location is shown in Figure 6.11. The excitatory neuron shows a significant shift towards larger size preferences at lower contrast ($> 3\times$). This shift is largely driven by the fact that both of the inhibitory populations demonstrate even more significant shifts in size tuning. The most striking feature of these results is the very gradual increase of Sst activation as the stimulus grows at low contrasts. This means that for very small stimuli the Sst neurons provide very little inhibition, allowing facilitatory effects to occur.

6.2.3 *Flanker Modulation*

A second often-used paradigm to measure surround-modulation effects is presenting a target stimulus optimized for a particular neuron, and then adding flanker stimuli, which may vary in orientation from the central neuron. These patterns are much more sparse and therefore do not engage the circuitry in the same way a natural image or even a sine grating would. We are particularly interested whether the surround-suppression effects that were seen above are sufficient to account for various saliency-based computations that are thought to originate in the early visual cortex. Specifically, can iso-orientation suppression drive a pop-out effect where visual elements with similar orientations suppress each other while leaving any heterogeneous elements unaffected?

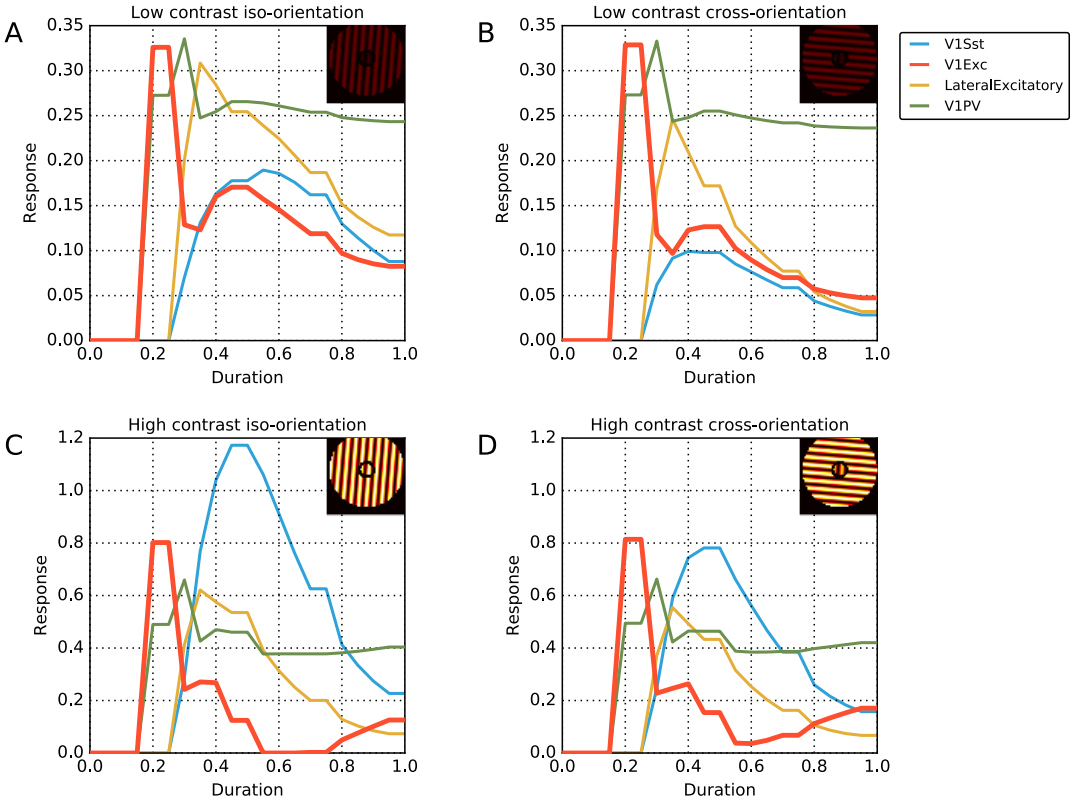


Figure 6.10: Time-course of neural and synaptic projection responses under four conditions, demonstrating how a contrast-dependent switch between facilitation and suppression occurs in the model. Visualizes how interneuron activity and lateral excitatory projection affect the response of the excitatory projection (in red). A) In the low-contrast iso-orientation condition, lateral excitation exceeds the activation of the V1 Sst population, resulting in facilitation. B) In the low-contrast cross-orientation condition, very little surround modulation occurs. C) In the high-contrast iso-orientation condition, the Sst population activates strongly causing strong suppression. D) Under the high-contrast cross-orientation condition, Sst neurons are only weakly recruited, resulting in a moderate amount of modulation. Results from model trained on natural dataset.

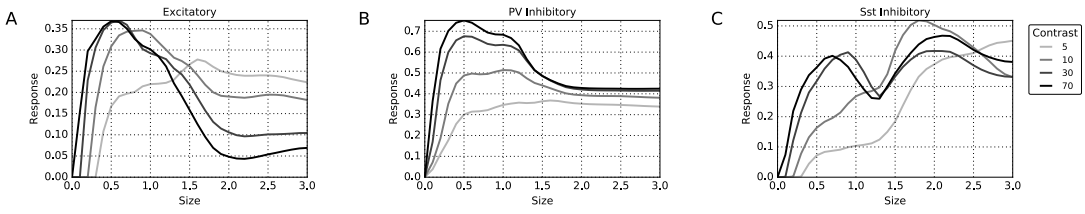


Figure 6.11: size-tuning curves measured in the LESPI model trained on natural stimuli for the A) excitatory, B) PV inhibitory, and C) Sst inhibitory population. The curves show clear modulation by stimulus contrast shifting toward higher size preferences, driven primarily by the shifts towards larger maximal suppression at low contrasts in the inhibitory populations. Results from model trained on natural dataset.

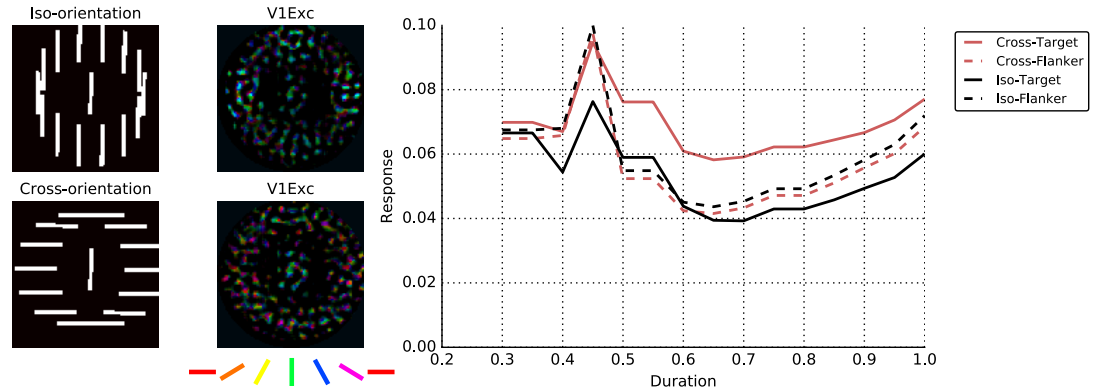


Figure 6.12: Demonstration of pop-out when presenting a target along with either iso-oriented or cross-oriented flankers, then averaging the response in two annuli corresponding to the target and flanker region. The V1Exc sheet responses are colored by the orientation preference of the neurons, showing that the center and surround regions can activate different orientations, and that the central response is suppressed by iso-oriented flankers but pops out when surrounded by cross-oriented flankers. The average response in the target region in both conditions is represented by the solid lines, while the annulus region is shown as dotted lines. The averaged responses highlight pop-out of the target stimulus from the background in the cross-orientation condition, but not the iso-orientation condition. Results from model trained on treeshrew dataset.

The simplest means of testing this effect is to embed a target bar stimulus in a ring of iso- or cross-oriented flankers, then compute the average activity in the region corresponding to the flanker against the region corresponding to the target. The effect of this can be seen in Figure 6.12, showing the two stimulus conditions and a plot of the average non-zero response in the central target region and in an annulus around it corresponding to the flanker region. The target activity is elevated above both the flanker activation in the same stimulus condition but more importantly it is also elevated when compared to the iso-orientation condition where all flankers have the same orientation.

6.3 DISCUSSION

In this chapter we introduced the LESPI model as an extension of the models presented in previous chapters, demonstrating how the model will exhibit many of the contextual-modulation effects that are known to occur in the primary visual cortex. Unlike other surround-modulation models, this model explains mechanistically how the long-range patchy excitatory connections emerge in the model, and how a specific class of inhibitory neurons can explain contrast- and context-dependent changes in surround modulation.

Using standard experimental paradigms for assessing surround modulation in the primary visual cortex, we show that long-range lateral connectivity targeting both excitatory and inhibitory neurons can provide a good match to electrophysiological and psychophysical results. The model allows inspecting the responses of every single neuron, providing comprehensive predictions for how neural responses are modulated.

6.3.1 *Feature-dependent modulation*

There are numerous models that have been used to explain feature-specific surround-modulation effects in the visual cortex. These models usually start by hardwiring a set of feature detectors, which are then connected according to their similarity in the feature space (Li, 2002; Schwabe et al., 2006). When modeling V1, the models thus connect neurons with similar orientation preference, to replicate the orientation-specific patchy connectivity that has been observed Bosking et al. (1997). However, while such models are useful to study specific phenomena, they provide an unsatisfyingly narrow explanation of cortical function, which does not generalize across different brain regions and sensory modalities.

Developmental models such as the model presented here explain the emergence of surround-modulation effects as a consequence of the way the circuit self-organizes and learns the statistics of the natural world. In other words, unlike other models, none of the surround modulation effects are hard-coded. Instead, the effects emerge naturally due to simple and general mechanisms: Hebbian learning, divisive gain control, homeostasis, and local microcircuits composed of various cell types. The emergence of iso-orientation facilitation and suppression is therefore the result of the model extracting co-occurrences in the training patterns. Indeed, by comparing models trained on simple independent Gaussian patterns, with the same model trained on natural or laboratory images, it is clear that the long-range modulatory effects only emerge when the training patterns contain consistent long-range correlations.

These results add to the growing amount of literature suggesting that the cortex uses the statistical dependencies between sensory features to optimize neural coding (Vinje and Gallant, 2000; Simoncelli and Olshausen, 2001) and improve performance in specific tasks (Geisler et al., 2001). In the LESPI model the long-range connections learn a statistical distribution of the inputs, and can then modulate the response to give rise to both facilitatory effects and suppressive effects. These effects are mediated through a divisive normalization mechanism similar to those used in models using higher-order correlations to optimize neural coding by reducing redundancy (Spratling, 2011; Coen-Cagli et al., 2015). This work suggests that the patchy lateral

connections in primary visual cortex could be the first stage for feature-specific redundancy reduction, without requiring specific feedback from higher-level areas. How much the lateral connections actually contribute to decorrelating activity and influencing development is an empirical question that will require further investigation, but the results here show that these connections could account for quite a range of contextual effects.

At the same time, the spatial scales over which the lateral connections can effectively act are very limited, dropping off almost completely after about 1° in visual angle in the model. This suggests that the lateral connections could mediate strong visual effects in relatively limited circumstances, such as at the borders of different textures or for more peripheral stimuli where the lateral connections would cover a large visual area compared to central vision.

6.3.2 *Contrast-dependent changes in suppression*

The contrast dependence of surround-modulation effects has been of great interest for many years, and a number of models have been proposed to explain it. Some form of asymmetry between excitatory and inhibitory neurons would be sufficient, as Somers et al. (1998) suggested. Their original proposal relied on an asymmetry in thresholds between excitatory and inhibitory neurons, which has little experimental support. However they also suggested that excitatory-inhibitory connections which increase in efficacy as contrast increases while excitatory-excitatory connections decrease (e.g. due to synaptic depression; Abbott et al. 1997; Tsodyks and Markram 1997), would be another possible explanation.

Since then a large number of studies have focused on different classes of inhibitory neurons, and the Sst suppressing interneurons fit precisely this description. Not only does the LESPI model demonstrate the contrast-dependent switch from facilitation to suppression that has been so widely described (Levitt and Lund, 1997; Polat et al., 1998; Dragoi and Sur, 2000; Wang and Fitzgibbon, 2009) as can be seen in Figure 6.9 and 6.8, we can now directly observe the emergence of an asymmetry in the efficacy of excitatory-excitatory and excitatory-inhibitory connections, targeting the excitatory and Sst populations respectively. The asymmetry is evident both in the timecourse of responses under various stimulus conditions as shown in Figure 6.10, and in the spatial integration properties (see Figure 6.11). Very long-range spatial integration by the Sst population has so far been described only in the mouse (Adesnik et al., 2012), but we predict that a cell class with a similar function should also exist in other mammals, for the various reasons outlined above.

In the model, the long-range projections integrate over both space and time due to their large spatial scale and hysteresis. These projections effectively compute, based on the local stimulus drive and the contextual activations, whether to suppress or enhance the neural response. While the contrast interactions have been characterized well in the organized model and correspond well with experimental results, further analysis should focus on how modulating the strength affects the sparsity of responses during self-organization, thereby influencing the developmental processes in the model.

6.3.3 *Spatial dependence*

One of the major features of this model over previous models is the calibration of spatial profiles of connectivity. Rigorous calibration is of great importance to understand the contributions of lateral connectivity to surround-modulation effects in the cortex. Much previous work has focused on teasing apart the contributions of lateral connectivity and feedback connections to surround suppression (Angelucci et al., 2002b; Bair et al., 2003; Schwabe et al., 2006). As many other studies suggest, the LESPI model demonstrates that while the lateral connections can mediate significant surround-modulation effects, they do so over relatively small spatial scales, covering up 6-7 mm in total diameter, translating to a visuotopic extent of about 2.5°. This range is not sufficient to account for the longest-range surround suppression reported, which can subtend more than 7° in visual angle (Bair et al., 2003; Levitt and Lund, 2002).

Thus the lateral connections modeled here cannot account for the full extent of surround modulation. At the same time, it is likely that they do provide some integration to take into the account the local context of the receptive field, perhaps aiding in computing the continuation of a contour or increasing the saliency of borders between different textures. Further experimental work should focus on considering the relationship between the neuron's receptive field size and the extent of its lateral connections, as that would provide a clearer indication whether the lateral connections are involved purely in local processing in the classical RF or also mediate non-classical receptive field effects. Indeed, since this relationship varies considerably between species, the role of lateral connections could also be to a large part species dependent. In species with a smaller cortex, these connections can span a considerable fraction of the visual field and could therefore mediate even long range surround modulation effects, while in species with larger brains such as macaques their role would be limited to near-surround suppression and facilitation, at least in central vision.

6.3.4 *Temporal scales*

In the current model the temporal properties of the neurons are only modeled on a very coarse level. Nonetheless, the classical shape of PSTHs emerges naturally from the model, with initial peaks and a smaller sustained response. Combining this model with temporally calibrated models as developed in Stevens (2016) would shed more light on the propagation speeds of the lateral connections. Based on a rough calculation, the long-range projections propagate at a speed of up to 0.2 m/s in the model, which is well within the reported physiological range of 0.1-0.3 m/s (Bair et al., 2003). With this speed, surround modulation arriving from lateral connections will act fast enough to have an effect, but any longer-range effects would have to be mediated by longer-range connections. On the other hand, since the model behavior already so closely replicates the observed contextual phenomena, it may be a relative simple extension to include feedback connections that mathematically act just like the lateral connections, but covering a wider spatial area made possible by the wide spread of feedback connections.

6.3.5 *Simple and Complex cells*

One major limitation of the current model is the lack of complex cells, which causes disruptions to the map organization when trained on natural images. This limitation also means that neurons respond optimally at one particular phase, so all surround modulations are dependent on the phase of the stimuli. In future, the model could be extended to incorporate layer-specific circuits that allow the emergence of complex cells, which has already been shown to improving the organization of the orientation map (Antolík, 2010). This possibility is of particular interest, because it is unclear to what extent lateral connections will capture the relative phase between different edges if they are primarily connecting complex cells in layer 2/3, which could have strong implications for their ability to represent specific spatial arrangements and mediate contour integration effects. The relationship of these connections to spatial phase is a very important question, and more experimental work should focus on the differences in lateral connections in simple and complex cells.

6.3.6 *Development*

In this chapter we mainly examined the role of lateral connectivity in surround-modulation effects. In the current model the lateral connections are simulated as

initially diffuse connections, which organize a patchy architecture connecting iso-orientation regions after development and finally they are thresholded so as to reflect the highly sparse organization that is observed in tracing studies. This is approach is not a good approximation of their actual process of development, as neurons slowly develop patchy connections in concert with orientation map development. The existence of long-range excitation also partially disrupts the organization of the model, as the initially diffuse and isotropic connections introduce destructive long-range correlation (Miikkulainen et al., 2005).

As part of the thesis research, a sparse-matrix implementation of the core algorithms was implemented, along with a variety of sprouting and retraction algorithms. In future, a more realistic approach to matching the timecourse of lateral connectivity development, which begins after the emergence of an orientation map (Ruthazer and Stryker, 1996), should be developed based on these early explorations. Initial results suggested that several different sprouting and retraction approaches could produce equivalent orientation maps and often resulted in more selective and stable tuning properties, while being faster to simulate.

It is also interesting to consider how reducing the correlations between distant neurons using surround suppression affects learning in the model. While decorrelating the activity of feedforward components will help in maintaining an unbiased distribution and therefore good coverage of features, it also reduces the precise correlations the network is trying to learn. In other words, if the long-range suppression reduces correlations in the input, it is eliminating the exact thing it is trying to capture. Indeed, through initial investigation (not shown), we were able to show that increasing the strength of lateral excitatory connections targeting the Sst population reduces the strength of both iso-orientation biases and isotropy biases in the lateral connections. This suggests that the lateral connections might learn differently from afferent connections, i.e. that afferent connections should learn on the final settled response, while the lateral connections should strengthen depending on the coherence of the signals arriving from afferent and lateral connections. Indeed this provides an intriguing connection to predictive coding models, where the sign of the lateral modulation represents the sign of the error (Rao and Ballard, 1999).

6.3.7 *More complex surround modulation effects*

In this chapter we mostly focused on fairly simple contextual modulation phenomena such as pop-out and orientation contrast suppression and facilitation. Generally these simple phenomena were modeled well but further investigation will have to be applied to determine whether this extends to phenomena such as contour integration

and more complex pop-out effects. In particular the model is highly sensitive to the total amount of excitation so it is unclear whether the results here extend to pop-out effects in crowded arrangements such as those tested in (Li, 2002). However from a conceptual perspective the mechanisms underlying simple and more complex phenomena are largely the same so the model is likely to be able to model these effects too.

There are however also some experimental observations that are not as easily explained such as the observation by Sillito et al. (1995) that when a stimulus with an orientation orthogonal to the preferred orientation is presented in the center and with the preferred orientation in the surrounding annulus, some neurons responded strongly even though presenting the center and surround independently elicited no response. This may be explained by the fact that an orthogonal stimulus in the center does not recruit much excitation on its own but also elicits little suppression, such that when a surround annulus with the preferred orientation is added the additional excitation arriving through long-range excitatory connections is sufficient to get the neuron to respond. In future these more subtle phenomena can be investigated using these models too, allowing us to make predictions on how the interaction between various cell classes can give rise to effects that are usually extremely difficult to disentangle.

6.3.8 *Effects of natural image statistics on the model*

In this chapter we switched from simple Gaussian patterns for training the model to samples from natural images, to allow the model to learn co-occurrence statistics in the lateral connections, which give rise to the surround-modulation effects observed here. However, the statistics of various natural image datasets can differ considerably, and relating the statistics of the sensory environment an animal is reared in and the surround-modulation effects that develop would be a huge step forward in understanding how the cortex learns higher-order correlations in the natural world.

Before such an analysis can be performed, we have to robustly quantify the statistics of the natural image patterns and what the lateral connections have actually learned, before relating these observations back to specific surround modulation results in the psychophysics and electrophysiology literature. In the next chapter we will present various analyses to do precisely that, which will let us make concrete predictions of how the visual rearing environment affects the processing of simple and complex stimuli.

7

MODELING THE EFFECTS OF VISUAL STATISTICS ON LONG-RANGE LATERAL CONNECTIVITY IN VISUAL CORTEX

One of the major problems in computational neuroscience is in understanding how the brain can robustly capture information about its environment to improve how new information is encoded and processed. One of the major strengths of developmental models such as those developed in the previous chapters is that the developed synaptic connections reflect the visual statistics of the input. Therefore we can make predictions about how the statistics embedded in the visual inputs will be reflected in the organization of the model and could be used to aid cortical computations.

A number of studies have investigated the role visual statistics play in shaping the organization of cortex. In particular it has been shown that the distribution of orientations in the orientation maps can be strongly affected by altering the visual experience of an animal through manipulations like goggle rearing (Tanaka et al., 2006). However, the evidence for the encoding of second-order statistics in lateral connections has been much harder to study, and even coarse approaches such as measuring the isotropy of lateral connections along the axis of preferred orientation has not yielded uniform results. While a number of studies have found that lateral connections are elongated along the axis of preferred orientation in tree shrew (Bosking et al., 1997), cat (Schmidt et al., 1997), and owl monkey (Sincich and Blasdel, 2001), in macaque this anisotropy may simply be explained by the anisotropy in cortical magnification such that the connections are not actually elongated in visual space (Angelucci et al., 2002b). Whether the different results reported per species are truly species differences or simply differences between the various laboratories and rearing environments involved is not fully clear. Unfortunately, the tracer injections required to reconstruct the lateral connections can be performed at most on a few cells in a single animal, making systematic collection of comprehensive data infeasible.

Although there is a clear lack of data in this area, a few attempts have been made to go further and establish whether the lateral projections connect co-circular orientation domains, reflecting the co-circularity in natural images (Hunt et al., 2011). These studies have again been inconclusive due to the sparsity of clear data.

In this chapter, we will employ the model introduced in the previous chapter to analyze to what extent the detailed statistics of the visual input shape the long-range excitatory connections. In particular, we will attempt to reconstruct those statistics

based purely on synaptic weights and the orientation map. In doing so we will determine to what extent the network has captured the statistics of the training dataset, and establish whether considerable biases in the inputs could affect the isotropy of connections.

7.1 METHODS

We will first introduce the methods for computing comparable statistics for natural images and for the LESPI network, which we can use to analyze how well the network has extracted the underlying patterns in the training data.

7.1.1 *Co-occurence statistics*

Natural image statistics have been well characterized in a number of papers. For instance, Perrinet and Bednar (2015) analyzed the edge co-occurence statistics in natural images by labeling images with edges at varying frequencies, scales, orientations, and phases through a greedy algorithm, and then computed both the relative orientation between each pair of edges, and the normalized azimuthal angle between them. An example image with labeled edges is shown in Figure 7.1, along with a diagram showing how the co-occurrences were computed. Similar analysis approaches have been used to visualize the co-occurrence statistics in natural images, such as those produced by Geisler et al. (2001) to predict the performance of human subjects in contour-detection tasks.

The afferent receptive fields in the visual cortex are usually assumed to act as feature-selective units. In primary visual cortex, connections between these neurons could therefore represent the co-occurence of the features, i.e. second- or higher-order correlations between simple oriented edges. In order to test this idea, the models were trained on various synthetic and natural image datasets, and the statistics embedded in the lateral connections were decoded.

The geometrical relationship between any pair of neurons each representing an edge can be defined by three parameters. The difference in orientation between the two edges θ, computed by taking the difference in orientation preference between each pair of neurons. Secondly the distance between the two edges in visual degrees d, computed by calculating the distance between the center-of-gravity of the weights of each neuron pair. Finally the angles φ and ψ, representing the relative and normalized azimuth between the two edges. The diagram in Figure 7.1C, demonstrates how the angles θ, φ and ψ were computed between the edges.

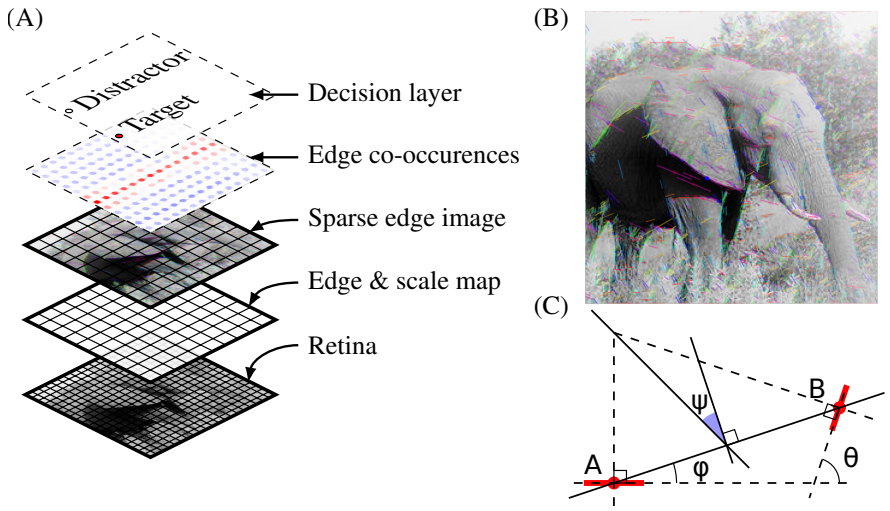


Figure 7.1: Diagrammatic representation of how a classifier is trained to distinguish between outdoor images and inanimate objects. A) The different layers used to train the classifier. The outdoor images are first fed through a model retina, all the edges are labeled with positions and scales. Using a greedy algorithm a set of edges accounting for the largest amount of luminance variance within the original image are selected. Using this set of edges the edge co-occurence statistics were computed and finally the classifier was trained based on these statistics. B) The sparse set of labeled edges extracted from a single image. C) Diagram showing how the angular difference θ and azimuth angle φ and relative azimuth ψ are computed from two edges. Reproduced from Perrinet and Bednar (2015).

As part of this analysis only post-synaptic neurons with a local homogeneity index above the 50th percentile, to mirror the procedure of Bosking et al. (1997) to select only neurons in iso-orientation domains. As in the previously described analyses the weights were first thresholded, leaving only the strongest 10% to ensure that the model roughly matches the known sparse and patchy pattern observed in anatomical tracing studies. To compute the preference of the network for particular arrangements the synaptic weights are combined with the orientation selectivity of the pre- and post-synaptic neurons and binned by the three geometric variables described above.

The results of this statistical analysis represents a three dimensional probability density function based on the synaptic weights between each arrangement of edges. This data can be binned and collapsed in a number of ways to evaluate uni-, bi- and multi-variate relationships that are encoded in the lateral connections.

Univariate statistics

The simplest way to analyze these statistics is to collapse across the dimensions to visualize the univariate distributions of the θ, φ, and distance of connections. This procedure provides (1) an easily understood analysis to demonstrate how strongly the connections are biased for similar orientations and co-linear directions, (2) an

indication of how isotropic the lateral connections are in the model, and finally (3) the distribution of weights by distance.

Multi-variate statistics

Using the computed histogram, various plots can be generated to compare against the plots produced by directly extracting the edge co-occurrences from images in the Perrinet and Bednar (2015) and Geisler et al. (2001) studies.

Chevron Maps The first of these plots (Figure 7.4), called a "chevron map", presents the bivariate distribution of θ and ψ, highlighting which spatial arrangement of edges is most likely to co-occur. In these plots, the central 1° diameter region of the lateral field was ignored so the plot would reflect co-occurrence statistics of connections *outside* the receptive field of the neuron. The color in each bin represents the strength of weights for each arrangement of edges relative to the mean. Therefore dark blue bins represent edge arrangements that the model learned to be less likely to co-occur while dark red bins represent arrangements more likely to co-occur. This procedure also allowed comparing the co-occurrences between datasets by computing the ratio of normalized weights in each of the bins, highlighting how the edge co-occurrences differ between datasets.

Geisler co-occurence plots The Geisler et al. (2001) type of co-occurence plots, on the other hand, reflect three dimensions, the angles θ and φ and the distance d. The plot represents the probability of an edge of a particular orientation, co-occurring at a particular distance and azimuth relative to the reference edge. It is then possible to plot the data to ask two related questions corresponding to the first and second co-occurrence properties: (1) what is the most likely orientation difference given a distance and azimuth (i.e., $\mathrm{argmax}_\theta P(\theta \mid d, \varphi)$), and (2) where is the most likely position or azimuth of an edge of a specific orientation at a specific distance (i.e., $\mathrm{argmax}_\varphi P(\varphi \mid d, \theta)$).

While the normalized weights cannot be directly compared to the relative probabilities obtained by Geisler et al. (2001) through analysis of various image datasets, they should provide a reasonable proxy for the co-occurrence probabilities as encoded by the network. Since each bin represents a circular annulus of different area the weights in each bin were additionally normalized by their area. In addition to obtaining the magnitude of the vector sum, a selectivity was also obtained by dividing the magnitude by the sum of weights. In the final plots the magnitude is indicated by the color, while the selectivity at least for the first occurrence property plots is indicated by size.

von Mises Model

Finally, the lateral weights were again fit using the von Mises model with both Gaussian and orientation-preference–dependent components, which was first introduced in Section 4.1.3. This procedure allows us to provide a quantitative assessment of how well the lateral connections can be approximated with a simple model that is made up of orientation, direction, and spatially dependent components. Most importantly, the direction-dependent component allows us to quantify how strongly the lateral connection fields are biased along the axis of preferred orientation, a key issue from the literature.

7.1.2 *Training stimuli*

Because natural images introduce a variety of RF shapes and phases, not just long-range correlations, we will first look at the results from synthetic stimuli constructed to have long-range correlations. These stimuli are simple extensions of the elongated 2D Gaussian stimuli used to train the SCAL and SEPI models. The model will draw 1 pattern per unit area of simulated model, which consists of a chain of three Gaussian patterns (see Figure 7.2). The middle Gaussian determines the overall orientation (θ_M) of the pattern, and the outer Gaussians are offset in orientation by a value drawn from a von Mises distribution with a mean of θ_M and a κ of either 0.5 or 8. The patterns are then spatially offset so that they form a chain forming an 'S' shape. By varying the κ we can vary how the distribution of co-occurence statistics affects the preference for simple elongated bars.

We will also train the model on two natural-image datasets that had previously been analyzed for their co-occurence statistics (Perrinet and Bednar 2015; see Figure 7.1), plus an additional image dataset recorded in 2010 in juvenile tree-shrew cages in David Fitzpatrick's lab at Duke University. The tree-shrew cage images feature great numbers of highly extended, high-contrast bars (shown in figure 6.3), and are examples of the rearing environment for the animals in the Bosking et al. (1997) study. Thus the statistics of these patterns could be what the Bosking et al. (1997) tree shrews encoded in their visual cortex, if the hypotheses from this chapter are valid.

7.2 RESULTS

We will start with the synthetic-stimulus–trained model, to ensure the approaches work well in the simple case, which we will then extend to the more complex natural-image–trained models.

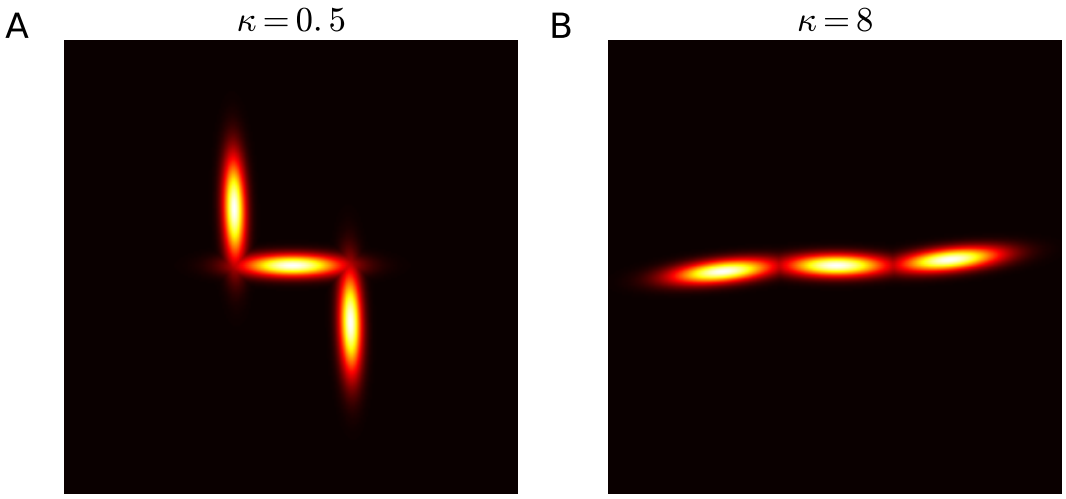


Figure 7.2: Gaussian patterns with co-occurence statistics where the orientation offset is drawn from a von Mises distribution with different κ values changing the distribution of orientation offsets.

7.2.1 *Synthetic Stimuli*

The stimuli described in Section 7.1.2 vary in both relative orientation and azimuth, which means that if the lateral connections capture the correlations in the inputs, then these correlations should also be captured and vary depending on the width of the von Mises distributions the orientation offsets are drawn from.

By analyzing and binning the lateral weights by orientation difference, relative azimuth and distance we can get begin to understand what the lateral connections are actually capturing. The difference in lateral fields between a model trained on the synthetic stimuli with a very wide distribution ($\kappa = 0.5$) and a much tighter distribution ($\kappa = 8$) is shown in Figure 7.3. These univariate distributions already demonstrate a clear preference for similar orientations and co-linear stimuli under both conditions. The histogram of relative azimuths highlights a strong isotropy along the axis of preferred orientation. This can indeed be seen even when just looking at the sample lateral connection fields directly (Figure 7.3A, B). The analysis also highlights that the model trained on patterns drawn from the tighter distribution ($\kappa = 8$) has a slightly stronger preference for similar orientations, and more weights at larger distances. This reflects the stronger bias for long iso-oriented and co-linear contours in the input patterns.

The chevron map of edge co-occurrences in Figure 7.4 highlights very similar results. In particular it demonstrates a greater spread in co-occurring orientations as the distribution of orientation offsets in the input pattern widens. Both the azimuth distribution and the co-occurrence histogram also emphasize an increased preference

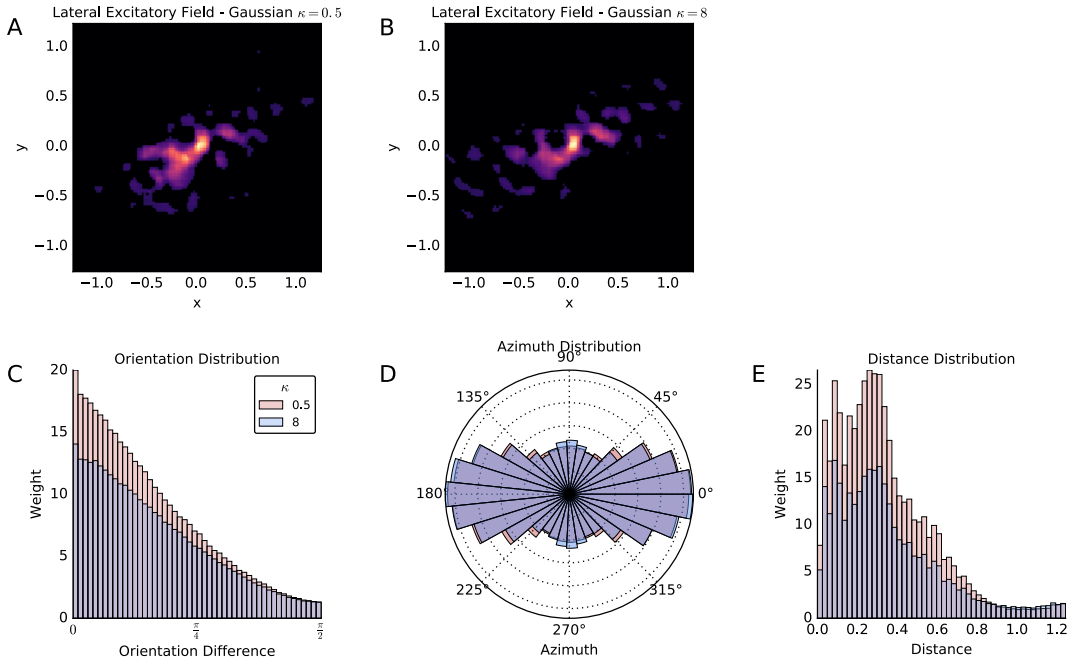


Figure 7.3: Lateral connections and histograms describing their orientation, azimuth and distance dependent distributions for narrowly and widely distributed synthetic stimuli. A, B) Example lateral excitatory weights after thresholding for the widely distributed ($\kappa = 0.5$) and narrowly distributed ($\kappa = 8$) condition. C) Orientation distribution of lateral connections, showing stronger bias for iso-orientations in the narrowly distributed condition. D) Azimuth distribution, showing strong co-linear bias for both conditions. E) Distance distribution, showing more weight at distant locations in the narrowly distributed condition.

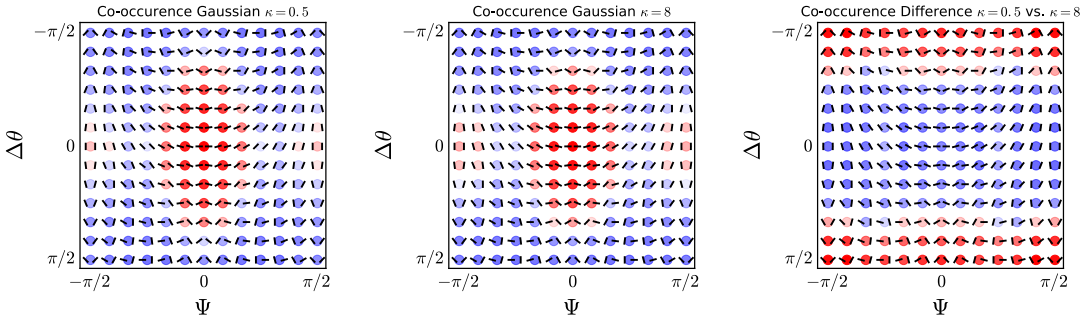


Figure 7.4: Chevron map showing the distribution of orientation and azimuth differences between pre- and post-synaptic neurons weighted by connection strength for the wide distribution (left), narrow distribution (middle), and the difference between them (right). Modeled after the Perrinet and Bednar (2015) co-occurrence analysis of natural images. Each 'chevron' represents one particular configuration of two edges, varying by ψ (relative azimuth) along the x-axis and θ (orientation difference) along the y-axis. Increased probability is shown in red, and lower probabilities shown in blue. These results also highlight that the bias toward similar orientations is significantly stronger than for azimuths.

of parallel alignment for the tighter distribution. It is not quite clear what drives these correlations, but through parameter exploration not shown here it was determined that they could be reduced by presenting fewer overlapping patterns, suggesting that they were due to increased likelihood of chance alignments.

The most striking results are obtained by the Geisler-style co-occurrence histograms in Figure **??** highlighting how well the network has captured the statistics of the input and additionally demonstrating how distribution of input patterns differs with a varying κ value. The line segments shown in the top row of the figure, show, at each of 6 distances and 72 azimuths, the most probable orientation difference. Expressed differently, for each distance and direction we plotted the preferred line orientation. The line segments in the bottom row of the figure on the other hand show, at each distance, the most probable direction for each possible orientation difference.

In both types of plot we can observe that the same orientation is encountered along the co-linear axis and as we move away from this direction the preferred angle slowly shifts away from a parallel to a co-circular arrangement, which is precisely how the input pattern is defined. The strength of the co-occurrence probabilities naturally falls off with distance, and this effect is even stronger in the $\kappa = 0.5$ condition, likely because the distribution is wider and extended, iso-orientation contours are less likely to occur. This effect is evident both in the orientation and azimuth co-occurrence plots.

The top row showing the preferred orientation differences particularly highlights the increasing preference for extended, co-linear contours with increasing κ. The bottom row showing the preferred direction (or azimuth) on the other hand highlights a slightly greater spread in azimuths when the distribution of azimuths is wider, demonstrating that the model can capture differences between the co-occurence distributions.

To make the differences even clearer we computed the ratio in weights for each orientation difference at each distance and direction between these two conditions and then plotted relative over- and under-representation of the weights for the $\kappa = 8$ compared to the $\kappa = 0.5$ condition. Figure 7.6 highlights just how strongly co-linear arrangements are overrepresented in the $\kappa = 8$ condition, with a 20-fold increase in the weight assigned to this arrangement. The underrepresentation of oblique and orthogonal arrangements in the $\kappa = 8$ condition also reflects the actual differences between input pattern distribution since a higher κ will result in a narrower distribution in the orientation difference between patterns. This suggest that the network does not only learn a coarse representation of co-occurrences but also captures more subtle differences in the input statistics. Having confirmed that the model can cap-

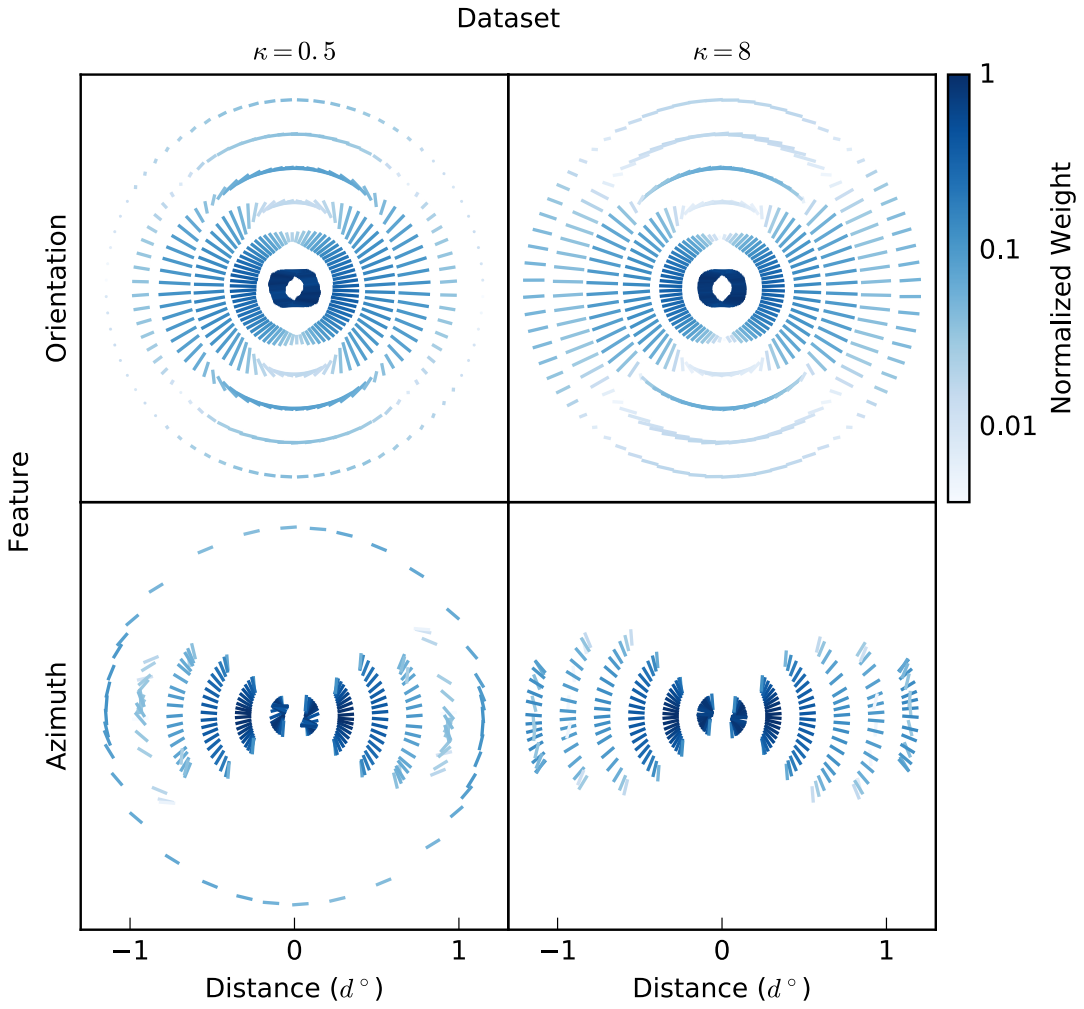


Figure 7.5: Co-occurrence statistics extracted from lateral connections to highlight the most likely orientation at each distance and azimuth for the two conditions (top row) and the most likely azimuth for each orientation and distance (bottom row). The colormap represents the relative probability of each configuration, while the size reflects the selectivity for that particular arrangement. The more narrow distribution (κ) results in more weight at the furthest distances and greater confidence in co-circular arrangements, however in both conditions co-linearity and co-circularity are decoded almost perfectly.

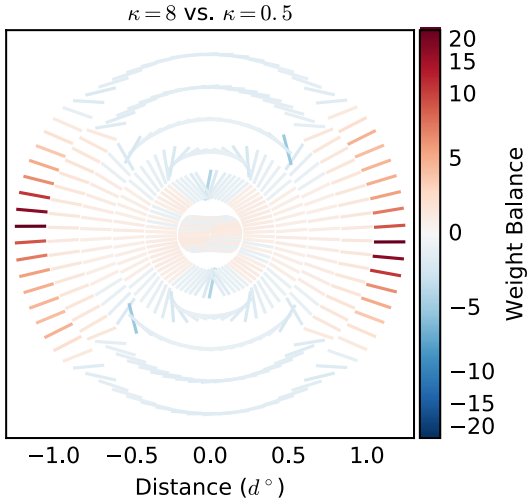


Figure 7.6: The difference in the first co-occurrence property between the models trained on synthetic stimuli with flankers drawn from a von Mises distributions with $\kappa = 8$ and $\kappa = 0.5$. Shows the ratio between their respective weights for each azimuth and weight bin. Underrepresentation is represented by negative (blue) values, while over-representation is represented by positive (red) values. This analysis highlights the strong over-representation of co-linear contours at the furthest distances and the slight underrepresentation of co-circular edges non-axis aligned azimuths in the $\kappa = 8$ condition when compared to the $\kappa = 0.5$ condition.

ture such differences for simple, synthetic patterns, we can extend the analysis to the far greater complexity that is present in natural images.

7.2.2 *Natural Images*

Natural and man-made environments have very different statistics, which may be reflected in the lateral connections in the primary visual cortex. So far we have seen that the LESPI model can indeed capture the statistics of a simplified input pattern and distinguish specific differences in the input statistics. Now we will train the model on a number of different image datasets, either from outdoor environments or artificial structures such as laboratory environments in which animals are raised. Using detailed labeling and analysis of these datasets, which include those used by Perrinet and Bednar (2015) and Serre et al. (2007a), we will demonstrate that the lateral connections in V1 can already capture the differences between different datasets. We will suggest that the rearing environment of an animal can have strong effects on the organization of the visual system, which should be considered when using animal models.

Man-made environments are characterized by co-linear, parallel, and orthogonal arrangements of edges, while true natural images, i.e., images of outdoor environ-

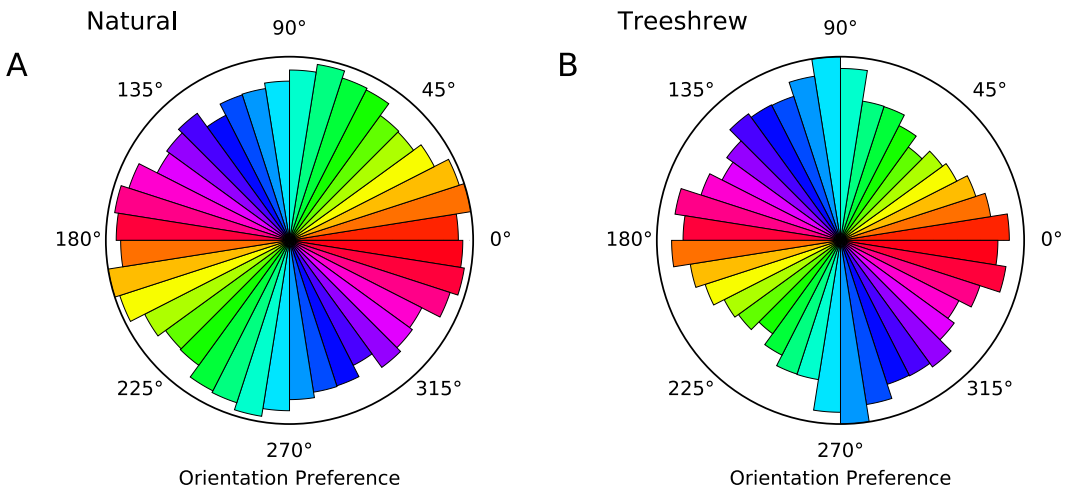


Figure 7.7: Distribution of orientations in the orientation map of models trained on (A) outdoor and (B) laboratory images. Although all images were rotated to eliminate any inherent biases in the original dataset, long-range interactions and border effects still cause small biases in the model.

ments, exhibit more curvature and textured patterns (Perrinet and Bednar, 2015). In order to test whether the model would capture these differences, three natural-image datasets were used: (1) the "outdoor" dataset containing a lot of grass textures, (2) the "treeshrew" laboratory cage containing long, high contrast bars, and (3) the Serre et al. (2007a) target dataset of natural scenes and animals. These represent three highly distinct visual environments, which should be reflected in long range connections.

First-order statistics

Before investigating more complex second-order statistics, we analyzed whether particular orientations were hugely overrepresented in the orientation map, which could introduce systematic biases to the results. As can be seen in Figure 7.7, both datasets show some bias for specific orientations, with the tree shrew condition exhibiting a stronger bias along the horizontal and vertical axes. These biases are likely an artifact arising from the artificial border at the edge of the simulated sheet, which can affect the whole map through long-range interactions.

Second-order histograms

Once again we compare the isotropy and the orientation and distance histograms between the two conditions (see Figure 7.8), which highlights significant differences between the datasets. Specifically, the orientation histogram is biased much more strongly towards iso-orientations in the outdoor condition. The outdoor condition is also considerably more orientation selective on average, which is what drives the

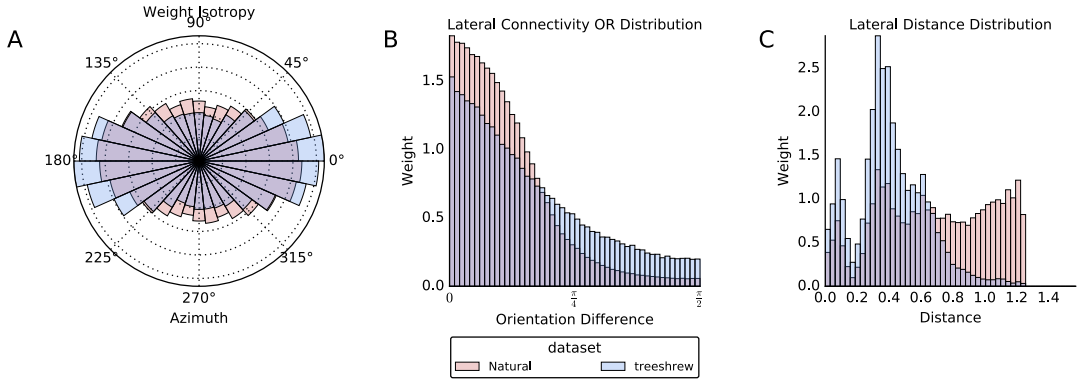


Figure 7.8: Distributions of lateral weights broken down by (A) azimuth, (B) orientation, and (C) distance for the outdoor and treeshrew datasets. The treeshrew dataset statistics are reflected via much stronger co-linear bias in the azimuth histogram even though they are more weakly selective and do not extend as far, emphasizing just how strong the co-linear bias in the data is.

patchiness of the lateral connections. However, even though the treeshrew neurons are much less selective, they actually exhibit a stronger bias along the axis of preferred orientation, with highly anisotropic lateral connection fields. Finally, we can see there is much more weight at distant locations, suggesting the outdoor patterns exhibit more long-range correlations overall.

Chevron Maps

The chevron maps offer a different view of these effects (Figure 7.9), as we can now see the relative co-occurrence along both the relative azimuth and orientation dimensions, providing an overview of the likelihood of various geometric arrangements. The views analyzing each dataset individually highlight once again just how much more biased the connections are along the θ dimension, i.e. that the connections are more biased for similar orientations rather than specific azimuths. However, the results also demonstrate that in the outdoor condition co-linear lines are not much more likely than any other azimuth, which reaffirms the more circular azimuth histogram that can be seen in Figure 7.8.

Comparing the distributions to ask whether certain configurations are more likely in one dataset than the other highlights some more interesting differences. Specifically, we can clearly see a stronger bias for parallel lines in the outdoor dataset and one for orthogonal lines in the treeshrew dataset. This may reflect the difference between textured grass patterns with numerous parallel arrangements, and man-made structures with far more right-angles than would usually be seen in nature. At the same time, certain curvatures are seen more strongly in the tree-shrew condition, which is surprising but may merely reflect the lower orientation selectivity that is

evident when training on that dataset. An analysis that considers the relative orientation independently from the azimuth may therefore shed more light on the actual differences.

Co-occurrence properties

Instead of considering both orientation and azimuth co-occurrences at the same time, we can treat each separately to ensure that one does not drown out the other. The first and second co-occurrence property plots, shown in Figure 7.10 alongside the results from Geisler et al. (2001) and Perrinet and Bednar (2015), make these differences very clear. Note that while the image analysis performed in these two studies can directly compute the probability, the model results use the normalized and weighted synaptic weights as a proxy for the probability. The magnitude and confidence obtained by computing the vector sum of the orientation and azimuth co-occurrences are indicated through color and size respectively.

The right half of Figure 7.10 shows the co-occurrence of orientation differences obtained from the model and previous analyses. The results obtained by Geisler et al. (2001) bear out strong preference for iso-oriented edges, close to and placed along the axis aligned with the orientation of the reference edge. The results from the model also show a preference for this arrangement, however there seems to some distortion in the azimuth for close distances, which may be caused by imperfect decoding or as a result of the local interactions that drive column formation. In comparing the direct image analysis to the model results for the Serre animal dataset and the tree-shrew–cage dataset, we can make several additional observations. The tree-shrew data has a stronger bias for parallel and co-linear edges at almost all azimuths and distances. The animal dataset on the other hand exhibits more co-circularity in the orientation differences but has very little preference and selectivity for arrangements that are orthogonal to the co-linear axis. This result presumably reflects the fact the treeshrew cages largely consist of long, parallel lines in form of the bars, while the animal dataset will contain more smooth, circular contours.

Next we can look at the azimuth co-occurrence preferences (shown on the left of Figure 7.10. The results from Geisler et al. (2001) and Perrinet and Bednar (2015) indicate some general patterns, parallel or similarly oriented edges are most likely to occur along the co-linear and co-circular directions, while oblique and orthogonal edges mostly occur at angles orthogonal to the reference edge. The model results also show a similar preference for the co-occurrence of parallel and similarly oriented edges, but only the model trained on outdoor images places a number of orthogonal edges along the axis orthogonal to the reference edge. This suggests the models is more selective for the azimuth of the co-occurring edge than its orientation, perhaps

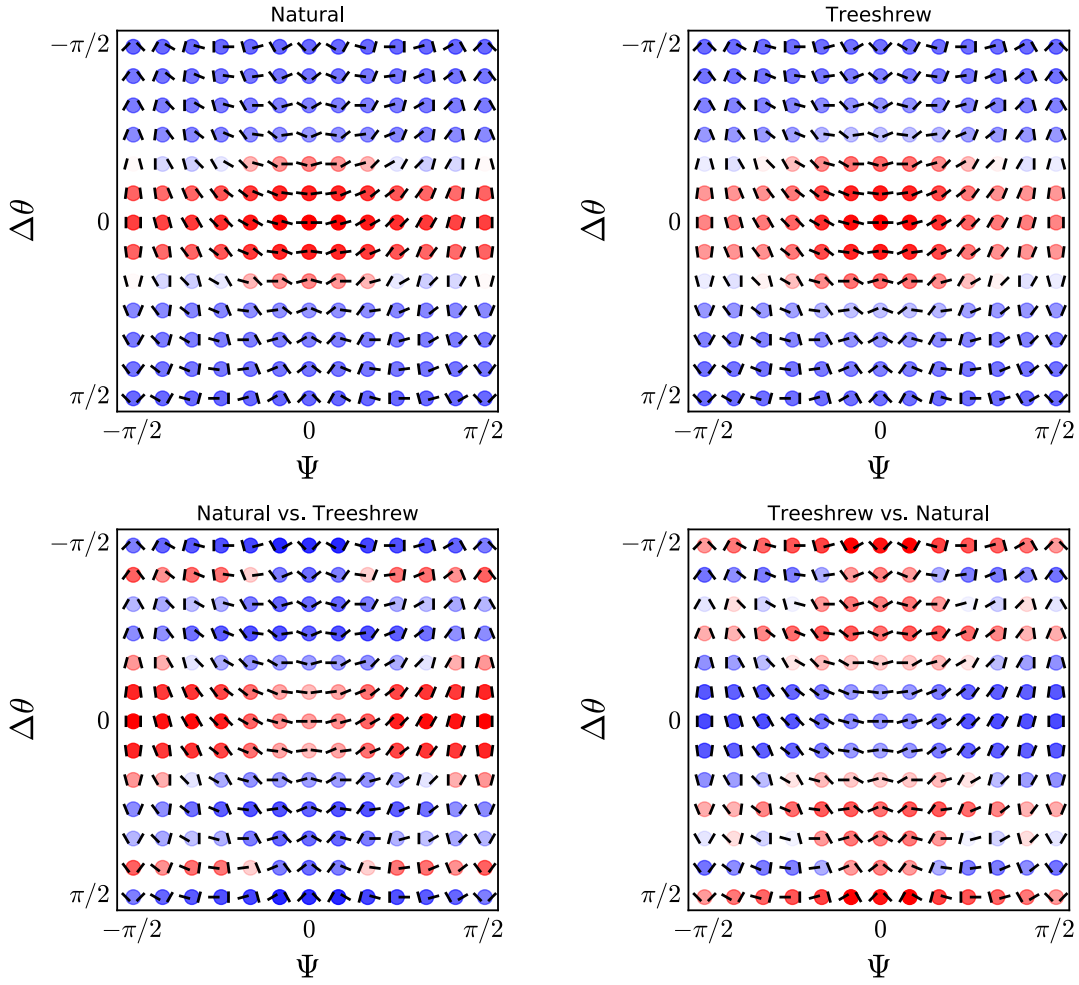


Figure 7.9: Chevron map highlighting co-occurrence statistics of geometrical arrangements in natural images. Red indicates higher probability of co-occurrence, while blue indicates lower probability. Chevron maps for outdoor and tree-shrew dataset shown on top and probability ratios comparing one dataset against the other below. Highlights strong bias for co-linear and fewer parallel arrangements in the treeshrew dataset, in addition to a greater bias towards perpendicular arrangements, which are rare in natural images.

reflecting low orientation selectivity. Looking more specifically at the model trained on the treeshrew dataset highlights a considerably tighter distribution of azimuth co-occurrences centered around the co-aligned axis. This suggests that the treeshrew dataset has more extended, co-linear contours and fewer co-circular or orthogonal edges when compared to the Serre animal dataset. While the model results are somewhat noisy this result is also strongly borne out by the image analysis performed by Perrinet and Bednar (2015), and makes intuitive sense, since the cages are made up of parallel bars while animals and their environments will exhibit more co-circularity and criss-crossing textures.

Quantifying the anisotropy

The analyses so far have allowed us to make qualitative assessments of how well the lateral connections in the model match the statistics in the input patterns. In order to test whether the lateral connections in the model are comparable to the long-range patchy connections in layer 2/3 of V1 we will also confirm how well the lateral connections fields are fit by the Buzas et al. (2006) model of lateral connectivity that was introduced in 4.1.3. By adding the aspect ratio of the long-range Gaussian as an additional free parameter we can also quantitatively assess the isotropy along the axis of preferred orientation to test our hypothesis that the anisotropy of lateral connections observed in experiments (Bosking et al., 1997) could be explained by differences in rearing environments.

After fitting the model to all the neurons we selected the best fits for the tree shrew and outdoor conditions and the error between the fit and the actual weight pattern, allowing us to see features the model did not capture very well. The comparison between the two fits is shown in 7.11. It is immediately obvious that the tree shrew weights are considerably more elongated along the axis of preferred orientation. The patterns of the error also highlight several issues, however. First of all, the fitting procedure seems to assign too little weight to the nearby iso-orientation patches, which are strongly connected in the actual weights. There are also patches that have been assigned weight, yet do not have any in the actual weight patterns. These are likely associated with phase-inversions, since the model consists entirely of simple cells, which means that locally iso-orientation patches with opposite phase are anti-correlated, a feature the Buzas model does not capture.

In order to confirm that the model indeed captures the orientation dependent component well we also let it fit the orientation directly and confirmed the fitted orientation was well correlated with the orientation preference in the orientation map. Overall this analysis showed very high correlation between the estimated and measured orientation, as can be seen in Figure 7.12A We also plotted the aspect ratio of

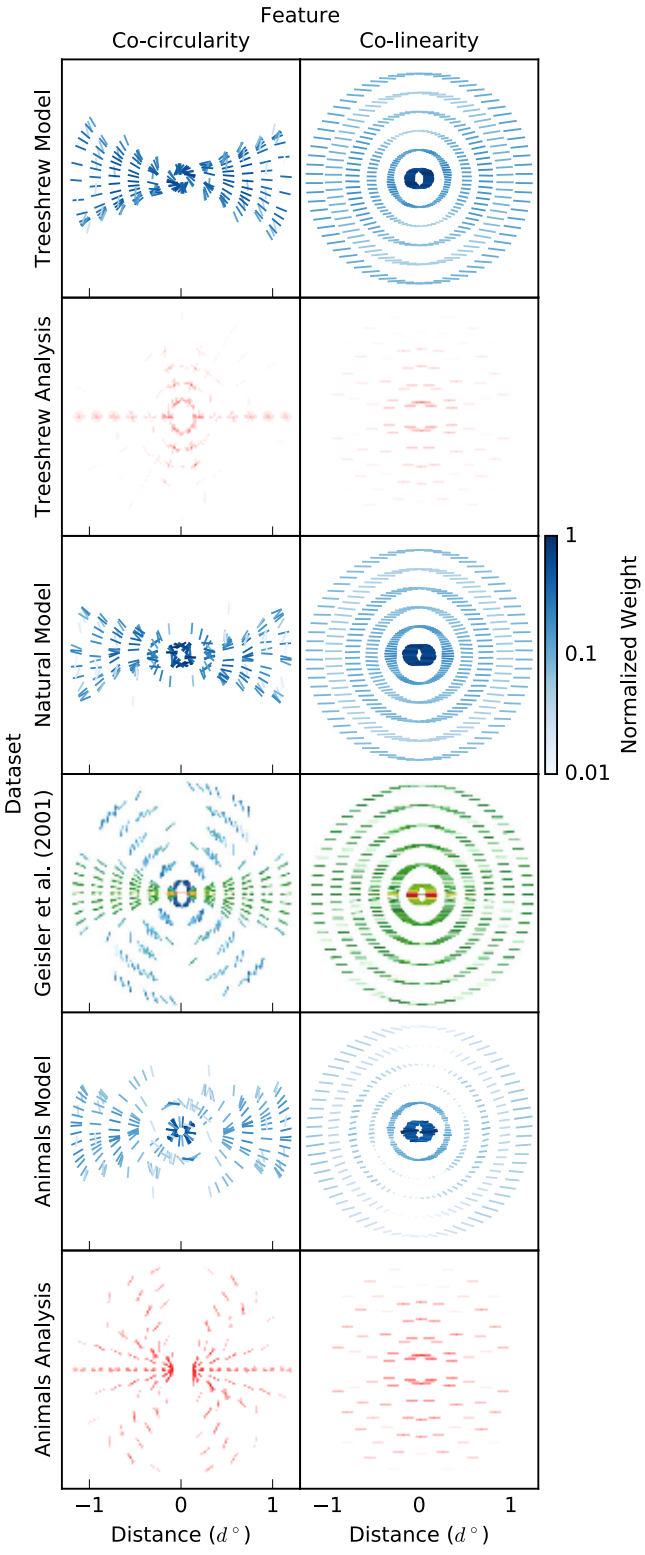


Figure 7.10: Comparison between co-linearity and co-circularity statistics extracted from the model and measured directly from the image datasets. Animals and tree-shrew analyses were provided by Laurent Perrinet by sparsely labeling edges in the image datasets and computing the statistics directly and should be directly comparable to the corresponding model results. Probability is indicated through the alpha level in the animals/tree-shrew analysis, and through color in the Geisler and model results. Model results additionally indicate confidence through the size of the edge. A number of clear conclusions can be drawn. First of all co-linear arrangements and extended contours are recognizable in all datasets. In the treeshrew condition the lateral decoding is very confident about these arrangements, while in the outdoor conditions similarly oriented edges occur at almost all azimuths. Cross-oriented edges co-occur with low probability (i.e. strength) but at least in the outdoor condition it is confident those do not generally occur on the co-aligned axis. Overall this provides strong evidence the lateral connections have captured the statistics of the input well and although less precise show strong parallels to the statistics extracted directly from the images by Geisler et al. (2001) and Perrinet and Bednar (2015) can be seen.

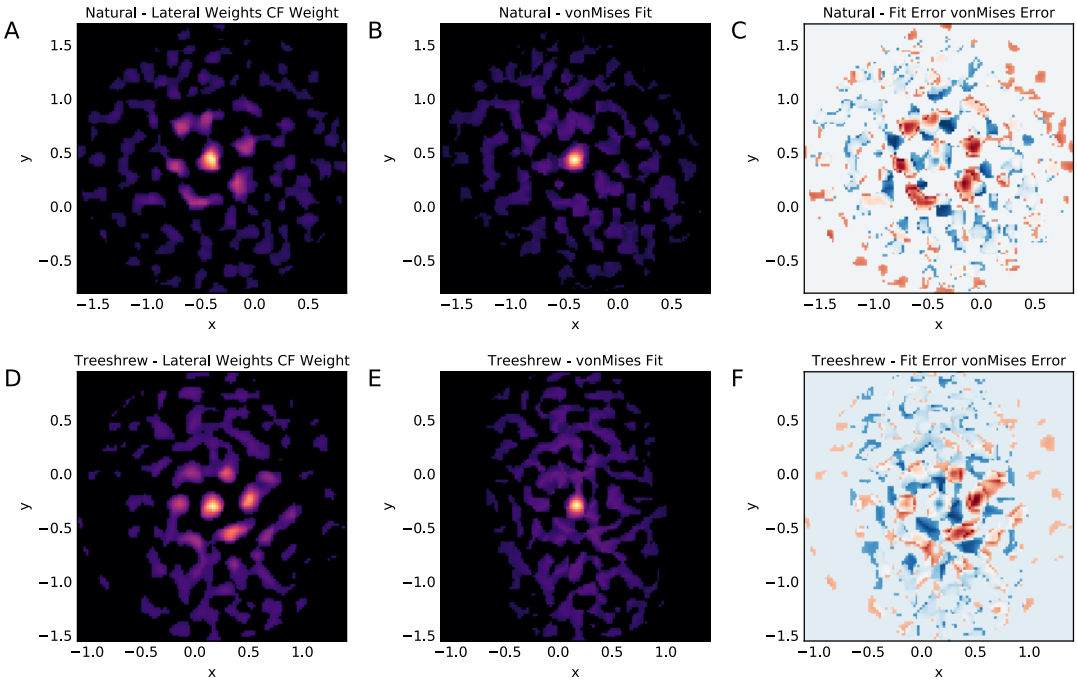


Figure 7.11: Comparison of Buzas et al. (2006) von Mises model fits between the outdoor and tree-shrew–trained models. A, D) Thresholded lateral fields in the outdoor and tree-shrew condition. B, E) Model-fitting result for the two conditions. C, F) Error between fitted and actual lateral weight fields. Overall the patchy organization is captured well but the large error at one hypercolumn distance suggests the model is not capturing the spatial weight distribution perfectly.

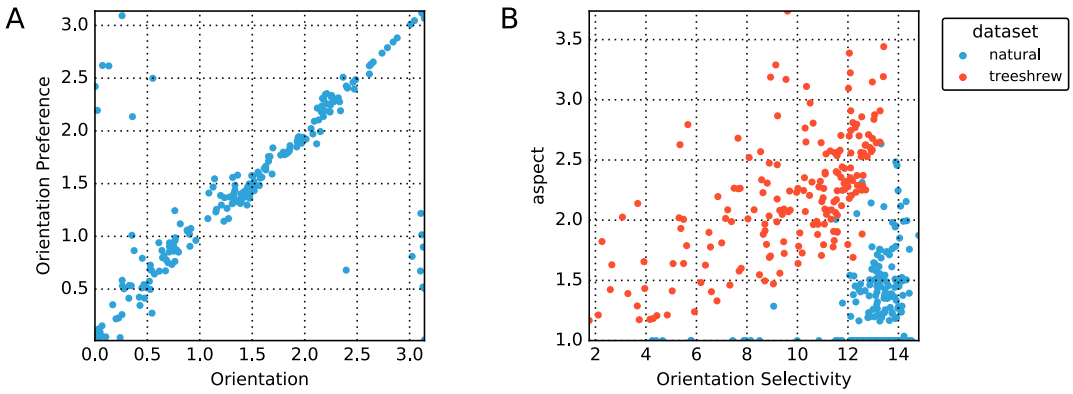


Figure 7.12: Results from Buzas model fitting. A) Correspondence between neurons' preferred orientations and the orientation estimated based on the lateral connection field. B) Dependence between orientation selectivity and aspect ratio for the two conditions. Clearly highlights the anisotropy in lateral connections trained on the treeshrew dataset.

the long range Gaussian pattern against the orientation selectivity of the neurons. In the treeshrew model, the selectivity was highly correlated with the aspect, while in the outdoor condition this correlation was much weaker. Overall, the mean aspect ratio for the outdoor condition was 1.29, while the treeshrew trained model exhibited an aspect ratio of 2.2, suggesting a much higher anisotropy along the axis of preferred orientation.

7.3 DISCUSSION

In this chapter we explored the effect of changing the co-occurrence statistics of the visual training on the long-range lateral connections in a developmental model of V1, in order to test whether various results concerning the co-circularity (Hunt et al., 2011) and anisotropy of lateral connections (Bosking et al., 1997) could be explained by differences rearing environments. We are also interested to what extent the early visual cortex is involved in computations concerning the co-occurrence statistics, to determine whether they could be involved in computing second or higher-order properties such as the difference between animal and non-animal objects, which has been shown to be computed very rapidly in human psychophysics experiments (Serre et al., 2007b).

In performing the analysis we demonstrated that the model could capture the statistics of simplified stimuli almost perfectly (see Figure 7.5) and demonstrated that it could even extract the statistics of far more complex natural-image stimuli to a reasonable extent, including clear differences between various datasets. This provides

the first demonstration that a developmental of V1 can indeed capture the statistics of natural images and can learn various Gestalt rules for edge co-linearity and co-circularity without hard-coding them. The results suggest that encoding second- and possibly higher-order statistics in the lateral connections is a general principle even in the early sensory cortex and predict that early sensory regions of all modalities should capture these correlations in some way. The model also makes it clear that the common assumption that patchy connections in the primary visual cortex specifically link columns with similar orientation preference is highly simplistic, with the more interesting story that such connectivity may simply reflect the co-occurrence probabilities in the natural world, whatever they may be for a given modality.

Furthermore, we quantify the effect of training the model in a laboratory environment with very long, high-contrast cage bars, and suggest that this highly biased rearing environment will have large implications for the organization of long-range lateral connections. Specifically, such images exhibit a considerably stronger anisotropy along the axis of preferred orientation than do outdoor images of nature. The tree shrew lateral-connection patterns have an average anisotropy ratio of 2.2, which is considerably higher than the anisotropy ratio of 1.29 in the model trained on outdoor images. On this basis we predict that the large anisotropy ratios observed by (Bosking et al., 1997) may be at least partially explained by the rearing environment of these animals, which is very different from outdoor images of nature.

The novel analyses described in this chapter provide a framework to answer questions about how second-order correlations are captured in the model. In future, these analyses should be extended to link the statistics encoded in the lateral connections back to the surround-modulation effects we previously showed can be mediated by them. Before such an analysis can be performed, a number of issues should be addressed.

7.3.1 *Spatial Frequency*

The spatial frequency distribution of natural images has been well described in the literature as having a $\frac{1}{f}$ distribution. Perrinet and Bednar (2015) have shown that the co-occurrence statistics are generally independent of the spatial frequency. However, the models in this thesis currently only uses a single spatial frequency filter at the level of the LGN, meaning that the other spatial frequencies will be filtered out. The image patterns in the tree shrew dataset appear to diverge considerably from the $\frac{1}{f}$ distribution that has been found empirically. Indeed, by investigating the selectivity of the model trained on tree shrew images, we could show that they generally have lower spatial frequency preference than the model trained on the outdoor dataset.

This in turn affects the selectivity of the model, since broader spatial frequency tuning also results in lower orientation selectivity for a given CF size. A future analysis should either employ a wider range of spatial frequency filters at the LGN level or ensure that the distribution of spatial frequency is approximately equal across the tested datasets.

Such calibration may be particularly important because the orientation selectivity between the models trained on the treeshrew and outdoor dataset differs quite considerably, which makes comparing between them very difficult. The chevron maps in particular are dominated by the difference in orientation selectivity, which partially obscures the the differences in azimuths.

7.3.2 *Complex cells and phase preference*

Another major question regarding the current analysis concerns the fact that all the modeled neurons are inherently simple cells. As it is known that long-range patchy connections emerge in layer 2/3, where there is a mix of simple and complex cells, having only simple cells in the model makes drawing concrete conclusions very difficult. In particular, complex cells pool over phase, which means they discard some of the positional information, complicating the encoding of relative azimuth between two edges. Extending the model to incorporate complex cells as outlined in Antolík (2010) may address some of these questions and would allow us to make concrete predictions about the differences of lateral connections linking simple and complex cells, which could be tested in experiments.

Indeed, the large difference in anisotropy that are observed in tree-shrew lateral connections compared to other species could reflect the fact that layer 2/3 in tree shrews has a greater proportion of simple cells, since orientation selectivity is thought to emerge from the connections between layer 4 and 2/3 rather than thalamocortical projections as in other species (Van Hooser et al., 2013).

7.3.3 *Local isotropy and suppression*

So far we have mainly discussed to what extent the connections do reflect the statistics of the natural images the model was trained on. There are however also clear and systematic differences which likely reflect properties of the underlying circuit. Since the cortex has to map high-dimensional visual features onto the 2D surface of the cortex, there are some tradeoffs in representing all the features perfectly. In particular, since orientation is mapped onto discrete columns with other orientations intervening, there is some distortion in the way position is represented locally. These

distortions can be seen in some of the co-circularity plots, which show maximal co-linear enhancement slightly offset from the center. Additionally, the local isotropic suppression provided by the PV population will strongly suppress cross-oriented stimuli locally, which means they do not generally show up in the co-circularity plots as is particularly evident in the outdoor condition, where they do appear at longer distances.

From a functional perspective, these do not seem like major issues, if the lateral connections are primarily concerned with capturing co-occurrences rather than simply reflecting the preference of the neuron in the classical receptive field. This issue once again highlights the importance of considering the visuotopic extent of lateral connections when compared to the size of the receptive field. Depending on the species and eccentricity, the neurons could shift from being primarily devoted to mediating effects within the classical receptive field to playing a modulatory role in transmitting information from the extra-classical receptive field. Systematically characterizing both the receptive field size and the extent of lateral connections may therefore shed some light on their primary function.

As we concluded in the surround-modulation chapter, the lateral connections in parafoveal regions of the macaque are just big enough to mediate interactions at the borders of different textures or between contour elements and beyond the cRF of a neuron, while more peripheral regions could integrate larger stimuli.

7.3.4 *Implications for surround modulation and perception*

Having confirmed that lateral connections can capture the co-occurrence statistics of the input, it is important to ask what purpose this may serve. One of the guiding hypothesis of this thesis and the developmental models on which the thesis is built is that the mammalian brain self-organizes based on activity-dependent processes in order to best represent the statistics of the input. This approach means that identical processes can give rise to the development of orientation maps in visual cortex and tonotopic maps in the auditory cortex. Rather than encoding the precise function of each brain region, the brain robustly yet adaptively organizes in such a way that it can robustly and accurately encode whatever input it is given. This idea is supported by experiments performed by the Sur lab, where the retinal projections to the LGN were rewired to the auditory thalamus. Sur et al. (1990) found that animals would develop visual receptive fields in auditory cortex, and were even able to perform various visual tasks with no primary visual cortex other than the rewired area (Von Melchner et al., 2000).

Based on this evidence and the results presented here, we argue that the organization of the visual cortex is to a large extent an experience dependent process and even surround-modulation effects such as contour integration and iso-orientation surround suppression are emergent phenomena, arising because the cortex learns to represent the visual statistics of the inputs. The role of lateral connections then is to learn co-occurrences of the inputs, expressing the model predictions either as suppressive or facilitatory effects depending of the geometric arrangement of the stimuli, the contrast and other contextual information. By learning that edges in the visual environment are generally co-linearly and co-circularly arranged, even the very earliest stages of visual processing can contribute towards complex computations. For instance, such a system could detect visually salient features through the learned Gestalt grouping laws, and fill in missing in information based on past experience when the input is weak or incomplete.

Indeed one of the major findings by Geisler et al. (2001) was the observation that the local grouping function for contour integration derived from absolute and Bayesian statistics are very similar. Since local grouping functions based upon Bayesian statistics are near optimal, this suggests that learning the absolute co-occurence statistics may be sufficient to achieve human level performance in contour grouping. In the analysis of the co-occurrence statistics we showed that the model was very successful at learning the preferences for parallel, co-linear and co-circular edge arrangements that are present in natural images. Future analysis should further explore whether decoding the neural responses could signal contour groupings, which could then be compared to the psychophysical results obtained by Geisler et al. (2001).

In order to predict to what extent the visual statistics of the rearing environment affect surround modulation, future work should focus on generating image patterns from the input statistics, in order to see how strongly the effects can be modulated when the statistics either match or clash with the statistics of the rearing environment. If the statistics play a significant role, then matching the statistics should maximize the surround-modulation effects and result in a sparser representation. This property alone may indeed be sufficient to explain why the responses to natural images are sparser than those for artificial stimuli.

8

GENERAL DISCUSSION

In this thesis we presented a series of models of the primary visual cortex to integrate our understanding of how different cell classes and their synaptic connections can give rise to a robust, yet dynamic circuit to extract and encode visual features and statistics of the world and replicate a wide range of contextual modulation phenomena. Additionally we presented a workflow for analyzing complex, high-dimensional datasets easily and reproducibly.

One of the major goals set out in this thesis was to begin integrating across the various levels of description that David Marr laid out in his seminal work on vision (Marr, 1982). Individual models describe individual phenomena very well and we have gathered a huge collection of data about the cortex from increasingly complex and precise experimental studies. However, the theoretical frameworks for understanding computation in the cortex have for the large part been unchanged since the early days of neuroscience, when concepts like the receptive and association field were first developed. This lack of progress stands in stark contrast to the field of machine learning, which recently has made huge advances in the area of deep neural networks, based loosely on the architecture of the cortex. The models developed as part of this model closely resemble the architecture of these networks, but also model the complexities of the cortical circuit to far greater detail. Integrative models are essential to begin to make sense of the large amounts of data that are being generated by modern experimental techniques, placing them into a overall framework that explains mechanistically how low-level network interactions give rise to behavior through computation and learning.

On this basis, we have proposed a series of models, which let us relate the process of learning to the actual computations performed by the cortex. Unlike many other models, integrative models such as those proposed here can mechanistically explain how different cell classes and synaptic connections interact at an algorithmic and implementational level to give rise to a developed circuit, which exhibits many of the features observed in the visual cortex.

At the same time, it is important to note the limitations of the modeling approach employed in this thesis. The models used here are particularly suited towards studying network-level phenomena, while lower-level phenomena at the level of individual neurons or even synapses are largely ignored. This approach allows simulating large

cortical networks and studying the response of the network to naturalistic inputs, but also means that the precise details of the molecular interactions giving rise to them are not accounted. This approach should thus be seen as complementary to both lower and higher level modeling frameworks, but it is my conviction that it is the appropriate level at which the computational principles of the cortex can be studied, ignoring details at the level of individual neurons but at the same time not completely ignoring inter-neuronal interactions. As such, this type of modeling is meant as a first order approximation of cortical function, and it is highly probable that the more detailed biophysical interactions at the level of individual synapses and spiking neurons will eventually be required to fully explain cortical computation.

The main aim of these models then is to identify a number of canonical computations by the cortex that explain a wide range of phenomena. In the past, both divisive gain control (Carandini and Heeger, 2012) and homeostasis (Marder and Goaillard, 2006) have been identified as core computations, and previous developmental models have shown that these mechanisms are sufficient to explain the robust, yet adaptive organization of the cortex into feature maps (Stevens et al., 2013b). In this thesis, we have made the case for an additional canonical computation, one that allows the brain to extract the statistics of the visual inputs in order to optimize coding and aid the organism in making sense of its environment by extracting the important information about the world by appropriately using feedforward information and statistical inference.

8.1 A REPRODUCIBLE WORKFLOW FOR THE ANALYSIS OF COMPLEX DATASETS

As part of this thesis we have described a general workflow for the analysis of complex datasets using the new HoloViews and Lancet Python libraries. Every single non-diagrammatic plot in this thesis has been generated using the HoloViews library, and is available with a completely reproducible record of how the figures were generated available at `thesis.philippjfr.com`. These tools in themselves provide a major contribution to greatly improve the productivity of researchers while ensuring, almost as a byproduct, that each stage of the research is completely reproducible. This process starts with launching instances of a model either locally or on a cluster, monitoring and collecting those results and finally generating complex figures and analyses with minimal amounts of code. The HoloViews library in particular is now a popular open-source project, and is in use in a wide range of scientific disciplines, has been used in numerous publications and for teaching materials. As a core designer and developer of the library, this work represents one of the most fundamental contributions

of this thesis and will improve productivity and reproducibility far beyond the field of neuroscience.

8.2 SPATIAL CALIBRATION

One of the first steps in approaching the hugely complex problem of modeling the neuronal interactions that give rise to development of a cortical circuit was to adapt existing developmental models in such a way that they would more closely reflect the real interactions between cell classes and synaptic connections in the cortex. Thus the SCAL model was introduced with spatially calibrated extents, allowing us to relate known anatomical measurements directly to the model and replicate various measurement protocols, which are highly dependent on the spatial profiles of connections relative to the distances in the model. The model was also switched from subtractive to divisive inhibition, as divisive contrast-gain control has long been identified as a canonical computation in the cortex (Carandini and Heeger, 2012). These small modification were sufficient to account for contrast-dependent size-tuning shifts, suggesting that the shift toward larger size preferences at low contrasts could almost completely be accounted for by changes in the gain of the center and surround mechanisms, as suggested by Cavanaugh et al. (2002b). Additionally, the spatially calibrated model provides a foundation to study effects that depend on the precise spatial configuration of the stimuli.

Furthermore through careful evaluation of the literature we have shown that the extent of a certain classes of inhibitory neurons is sufficient to account for the effective Mexican-hat connectivity in the model, which is required for robust self-organization into an orientation map, while even longer-range excitatory projections provide the true long-range interactions in the circuit.

8.3 CANONICAL COMPUTATIONS

The spatial calibration was but a first step, making it possible to relate anatomy to function in the cortex and begin integrating the vast amount of information that has been gathered in the pursuit of understanding the organization of the cortical circuit. Our main interest was in establishing core mechanisms in the interactions of excitatory and inhibitory neurons in the cortex to determine their role in cortical computations and if possible relate them back to the neural subtypes and connections that mediate them. Here we will identify these mechanisms, summarize the findings, and make the case that the mechanisms perform a number of canonical computations

in the cortex that are valid for multiple sensory modalities and perhaps even for non-sensory areas.

8.3.1 *Feedforward inhibition*

Specifically, we discovered that the known properties of PV neurons make this class of large, feedforward and recurrently driven interneurons particularly well suited towards driving the organization of V1 into a smooth orientation map. By changing the properties of the inhibition, we were also able to show that sub- or supra-linear integration by this cell class would result in disruptions to the development, which may reflect the changes in visual processing that have been observed in pathologies as diverse as autism and schizophrenia.

Our analysis also suggested that the feedforward-driven PV neurons are perfectly positioned to provide feedforward inhibition gating how much afferent drive arrives in the cortex, thereby controlling the effective contrast gain of the circuit. While a model cannot directly confirm that PV+ neurons serve the same functions in macaque visual cortex as has been suggested in mouse cortex, the analysis does demonstrate that feedforward inhibition, implemented as divisive gain-control, is an essential property to account for robust and stable organization and contrast dependent size tuning shifts. On that basis we predict that PV interneurons in higher mammals mediate global, feedforward driven contrast gain control, as suggested in a number of studies in mouse V1 (Ma et al., 2010; Atallah et al., 2012; Wilson et al., 2012; Nienborg et al., 2013).

At the same time, the class of PV-immunoreactive interneurons is very broad and may exhibit considerable heterogeneity. Specifically, clutch cells, which are primarily located in layer 4, likely have very different properties than the large basket cells that cover a larger area of layer 2/3. Indeed, there is some evidence that clutch cells are involved in providing the "pull" component in a push-pull model of orientation selectivity. This issue may also explain the ongoing debate about the extent to which this cell class is tuned for orientation, with studies in different layers and species coming to very different conclusions (Cardin et al., 2007; Ma et al., 2011; Hofer et al., 2011).

The major contribution of this model in the context of the thesis is in establishing a robust model of orientation map development that is spatially calibrated and has distinct excitatory and inhibitory populations. The model produces orientation maps that match biological maps very closely (Kaschube et al., 2010; Stevens et al., 2013b), with high stability and selectivity. The strong divisive gain control afforded by the inhibitory population also accounts for contrast dependent size tuning shifts, which

matches experiments both in diversity and magnitude. As such it provides a solid basis to build further models on, including models for surround modulation.

In a wider context, we suggest that feedforward gain control is a fundamentally important mechanism for computation in the cortex. It allows gating the amount of feedforward drive that is received by an area in the cortex, providing a powerful mechanism to control the balance between feedforward and recurrent processing that can be modulated in a task-dependent manner, controls development and fundamentally shapes the organization of the cortex. The PV interneuron population provides an excellent candidate to mediate this suppression, and experimentalists should work towards precisely characterizing their properties across species to establish whether their function is evolutionarily conserved, which would provide a good indicator of how central they are to cortical computations.

8.3.2 *Feedback inhibition*

While feedforward inhibition responds very quickly to afferent input, surround modulation operates over longer spatial and temporal scales. The other major contribution of this thesis is in proposing a general cortical circuit to mediate long-range interactions via slowly integrating long-range connections, which mediate both direct facilitation and di-synaptic suppression through a secondary interneuron population. After a detailed review of the literature we concluded that the Sst population, which receive extensive recurrent inputs from lateral and feedback connections in layer 2/3 integrate over long spatial scales (Xu and Callaway, 2009; Adesnik et al., 2012; Nienborg et al., 2013), provide an excellent candidate to mediate long-range and feature specific surround suppression. Additionally we have shown that an accelerating response function on the lateral connections between the excitatory cells and this inhibitory population is sufficient to account for the contrast-dependent changes from surround suppression to facilitation that have been well characterized in the brain (Levitt and Lund, 1997; Polat et al., 1998; Dragoi and Sur, 2000; Wang and Fitzgibbon, 2009).

The full proposed circuit therefore accounts for a huge range of response properties, developmental processes, and surround modulation effects using a small, general set of mechanisms that can be used to inform future experiments. In particular, the modeling suggests a strong role for a secondary laterally driven inhibitory interneuron population to provide long-range surround suppression, one that is distinct from the much faster, feedforward driven inhibition we have attributed to the PV population.

Whether feedforward and feedback inhibition can be directly mapped onto the PV- and Sst- expressing inhibitory cell classes will have to be experimentally validated.

In any case, we suggest that feedback inhibition is another core computation performed by the cortex. Specifically we suggest that it allows placing the information arriving via the feedforward pathways into context, allowing the cortex to compute visual saliency and fill in missing information. Additionally feedback inhibition can decorrelate neurons across wider spatial scales allowing the feedforward responses to extract a maximally independent and therefore sparse representation of the inputs. However, unlike many previous proposals, we suggest that instead of being hard-coded this suppression emerges as a consequence of the fact that the cortex captures the statistics of its sensory inputs automatically in the patterns of lateral connections.

8.4 COMPARISON WITH OTHER MODELS OF SURROUND SUPPRESSION

Over the past decade a number of studies have found evidence that V1 may operate as a so called inhibition-stabilized network (ISN). The first experimental findings that pointed to this mode of circuit operation was that the activity of inhibitory neurons and the inhibition received by excitatory neurons actually decrease during surround suppression (Ozeki et al., 2009b). Models based on the principles of ISNs such as the stabilized supra-linear network suggested by Rubin et al. (2015), suggest that surround suppression and normalization can be emerge naturally from a number of circuit properties, a) supra-linear neuronal input/output functions, b) sufficiently strong recurrent excitation, c) feedback inhibition and d) simple spatial properties of intracortical connections.

An ISN is characterized by a network that is unstable by itself but is stabilized through feedback inhibition. Without formal verification it is difficult to say whether the models presented in this thesis operate as ISNs. They meet many of the requisites including supra-linear I/O functions, strong recurrent excitation and feedback inhibition. Indeed without recurrent inhibition the network would quickly destabilize as we could observe when varying the exponent of the PV neuron output function in Chapter 4. At the same time however surround suppression in the model is in general accompanied by an increase in inhibitory responses particularly in the somatostatin-expressing inhibitory neurons.

This leaves the possibility that the network, just as the real cortex, moves in and out of an ISN regime. Indeed experimental studies on SOM-expressing interneurons in mouse V1 suggest that this class of neurons is surround facilitated (Adesnik et al., 2012), while PV and excitatory neurons are suppressed, which is what drives feature dependent surround suppression in the LESPI model. Even then net inhibition received by all neuron classes may still be reduced because SOM neurons make up a fairly small fraction of all inhibitory neurons in macaque cortex.

Without formal verification its not clear whether the models presented as part of this thesis ever fall into an ISN regime, but they do share many of the properties of an ISN and increasing the strength of recurrent, excitatory interactions relative to afferent input may push it further into that direction. Further investigation will be needed analyze whether operating under such a regime would improve the function of the models, particularly in regard to amplification of weak signals at low contrast conditions. This may be achieved by increasing the exponent in the excitatory response function.

It is also interesting to consider how ISN and non-ISN regimes relate to the balance between PV- and SOM-interneurons, which recurrently inhibit each other in the cortex. Independent of the operating regime of the network, many conclusions that can be drawn are the same. Rubin et al. (2015) for instance suggest that the role of attentional enhancement and modulatory suppression may simply be shifting the network between states dominated by excitation or inhibition, effectively turning the knobs of balanced amplification.

8.5 ROLES OF LATERAL CONNECTIVITY

In the model, the inhibitory populations themselves do not exhibit any long-range or feature-specific connectivity patterns. Long-range interactions are predominantly driven by long-range excitatory connections spanning several visual degrees and across multiple hypercolumns, as has been extensively established in the literature (Weliky et al., 1995; Bair et al., 2003). Furthermore, detailed analysis of the connectivity patterns of these long-range connections established not only that they link iso-orientation columns, but also that they are able to capture the co-occurrence statistics of both simple synthetic stimuli and more complex natural images. On that basis we suggest that the long-range patchy connections in layer 2/3 of the cortex reflect patterns of co-linearity by connecting iso-orientation regions, but that they also capture more subtle properties of natural image statistics such as co-circularity and other Gestalt rules of visual perception and are therefore directly responsible for mediating contextual effects such as contour facilitation, iso-orientation suppression, and pop-out.

In order to understand what the lateral connections were actually representing, this thesis contributes a number of analyses based on existing methods for analyzing the co-occurrence statistics in natural images put forward by Geisler et al. (2001) and Perrinet and Bednar (2015). By adapting these analyses to extract the co-occurrence statistics from the weight matrices in the model, we were able to conclusively demonstrate that lateral connections could capture the co-occurrence statistics of the visual envi-

ronment. On this basis we predict that lateral connections across the cortex capture higher-order statistics of the inputs, providing a statistical model of the relationship between lower level feature-selective units.

This work also predicts that the rearing environment of an animal strongly affect the organization of lateral connections in the cortex, which could explain species differences in the anisotropy of lateral connectivity. To test this possibility directly, the model was trained on both image datasets drawn from nature and from treeshrew cages, resulting in highly significant differences in the anisotropy of the connectivity. This suggests that the differences in lateral connection isotropy between treeshrews (Bosking et al., 1997) and macaques (Angelucci et al., 2002b) could at least partially be explained by the differences in rearing environments of the animals. Experimentalists should test for this possibility explicitly because it could significantly influence the properties of surround modulation.

Establishing to what extent these connections are involved in higher-order processing is still under heavy debate. However, based on the results from the model we can draw several specific conclusions. First of all, their spatial scale is sufficient to capture interactions beyond the size of the receptive field of most neurons, suggesting they are at least involved in processing contextual influences from the extra-classical surround. The extent to which they are involved in these processes probably depends on the species and cortical eccentricity, as receptive field sizes vary considerably across species, but at least in some species they can cover a significant portion of the visual space (Bosking et al., 1997). Even so, they are neither fast nor extensive enough to account for the full range of surround effects.

While these specific results are very specific to the field of vision, lateral connections are found extensively across the cortex, suggesting they play a crucial role in cortical computation. Indeed there is a growing amount of evidence suggesting that statistically optimal surround suppression can improve figure/ground separation and computation of visual salience (Coen-Cagli and Schwartz, 2013; Coen-Cagli et al., 2015). One of the core hypothesis put forward in this thesis is that lateral connections provide a general mechanism for the cortex to generate a statistical model of the inputs to a particular cortical area, which the cortex uses dynamically for a number of crucial computations, depending on the sensory and behavioral context. Specifically, through interaction with the inhibitory interneurons they can aid in discrimination through suppression of similar stimuli and detection of surprising stimuli, which are not well captured by the statistical model and are therefore unsuppressed. At the same time the connections provide a substrate for the cortex to make statistical inferences when feedforward input is not sufficient, resulting in effects like contour completion.

8.6 LATERAL VS. FEEDBACK CONNECTIONS

A major component of this thesis was to determine whether lateral connections could, at least theory mediate many of the surround modulation phenomena in the cortex. A wide range of experimental data argues that the involvement of lateral connections in these phenomena is at the very least limited (Levitt and Lund, 2002; Bair et al., 2003; Schwabe et al., 2006), with the majority of the modulation originating from higher cortical areas. Given the spatial scales over which surround modulation can operate in visual cortex and the speed with which it propagates makes this conclusion all the more likely. However lateral connections form a major source of input for neurons in the visual cortex and encode the appropriate information to detect local discontinuities in textures and contours.

It is also worth noting that lateral connections are spatially far more precise than feedback connections (Angelucci and Bressloff, 2006), which would make them well suited both to reduce local spatial correlations, and serve as a mechanism by which local texture discontinuities can be detected and continuous contours can be grouped together. The less precise feedback connections can then control the balance between recurrent excitation and inhibition to provide task- and context-dependent modulation of the incoming activity, as well as driving interactions that operate over larger spatial scales.

8.7 FUTURE WORK

The models, analyses, and tools presented in this thesis cover a huge range and individually provide major contributions to our understanding of cortical circuits and computation. However, perhaps more importantly, they lay the groundwork to study a wide range of phenomena in the cortex. In this section we will lay out how the models could be extended to study phenomena reaching well beyond the primary visual cortex and suggest specific avenues for future studies that were outside of the scope of this project.

8.7.1 *Linking statistics to perception*

A major component of the work presented here was in demonstrating how lateral connections could mediate a number of surround modulation effects that have been observed in the primary visual cortex. Secondly, through analyzing the connectivity between neurons with different feature preferences, we also showed that they were

able to accurately capture the statistics of the developmental training patterns. As an obvious next step, future experiments should provide detailed analyses to what extent differences in the visual statistics during rearing can affect surround modulation in the adult. Such experiments will link up all the ideas proposed in this thesis, connecting low-level mechanistic circuit interactions to learning and learning to behavior. As far as we are aware, this approach would lead to the first fully integrative model able to link low-level processes to high level behavior with an overarching theory of the computations performed by the cortex.

One particular avenue for future experiments would be to generate stimuli that either match or do not match the statistical dependencies encoded in the lateral connections, demonstrating significant differences in surround modulation. This procedure would provide concrete predictions, which could be tested by rearing animals in environments with significantly different visual statistics and confirming that surround modulation is indeed affected.

8.7.2 *Interactions between cell classes*

In the current model interactions between the inhibitory cell classes were largely ignored. However, based on experiments it is thought that PV and Sst neurons strongly inhibit each other. Through some exploratory work it was observed that Sst suppression of PV neurons could lead to interesting dynamics where they would essentially disinhibit the local excitatory population, reducing the amount of contrast gain-control and therefore reducing the effective integration area. This is a very interesting avenue to pursue, as this could effectively model some of the features of visual attention, where feedforward drive is increased relative to the lateral processing, effectively mirroring the effects of increasing the contrast.

8.7.3 *Sparsity of connections*

One major avenue of work that was developed for this thesis but ultimately was not used was developing a sparse-matrix implementation of the core algorithms used to simulate the model. This work has since been extended to work on GPUs, enabling significant speedups in simulating large models. This work will not only allow for the implementation of much larger models, but will also make the existing models more realistic. Specifically, it is known that V1 neurons in layer 4 receive afferent input from just a small number of afferent projections, and lateral connections are highly sparse in their pattern of innervation.

Simulating sparsity could therefore improve the realism of the models and at the same time improve a variety of properties in the model. Specifically, initial exploration of sprouting and retraction algorithms for controlling the sparsity of synaptic connections demonstrated that orientation map development could proceed with just a small fraction of the connections that are simulated in the dense version. The resulting models were also considerably more stable because the afferent weights could not easily move around. This approach would provides more realism for the simulation of orientation map development.

More interestingly, however, letting the lateral connections develop sparsely could match the known time-course of the emergence of patchy connections in the cortex very well (Ruthazer and Stryker, 1996), developing in concert with the orientation map and then slowly clustering in increasingly distant iso-orientation patches. This approach would also remove the need to threshold the weights after development, achieving the sparse distribution of connections required for strong surround modulation without an artificial mechanism.

8.7.4 *Complex cells*

One major limitation of the models presented in this thesis is that they have purely simple cells, even though long-range connections and surround modulation effects are generally ascribed to layer 2/3, which has both simple and complex cells. This may have considerable implications for the results offered here, since complex cells trade off precise positional information for invariance. If lateral connections primarily link complex cells it is possible that there is not enough spatial precision to accurately capture the co-circular properties of visual scenes over the spatial scales that they are known to project over. Additionally the current set of models suffers from disruptions in the orientation map caused by phase reversals, which would be eliminated by the introduction of complex cells (Antolík, 2010).

An extension of the models presented here, which explicitly models layer 4 and layer 2/3 and therefore develops orientation invariance would therefore be an excellent avenue of further study and make concrete predictions about the differences in lateral connectivity, linking simple and complex cells while improving our understanding of what purposes they serve.

8.7.5 *Tuning diversity*

For reasons of computational and conceptual tractability, the models introduced as part of this thesis have approximated the incredible diversity in LGN receptive fields

by focusing on a single average response profile. However, there are complex interactions between spatial frequency preferences and orientation selectivity that could therefore not be sufficiently explored. Modeling the full diversity of spatial frequency filters would provide a better account of the diversity of spatial and orientation tuning diversity in primary visual cortex. Additionally, very little work has been done to begin to understand how spatial frequency maps and relate to other feature maps, which would go some way towards providing an understanding of how the cortex is able to map features varying in multiple dimensions onto the 2D surface of the cortex.

Having such a model would also make it easier to relate the statistics across datasets with significantly different spatial frequency distributions, which has proved problematic when extracting the natural image statistics from the lateral connectivity.

8.7.6 *Temporal responses*

While significant effort was put into spatially calibrating the models presented in this thesis, it only has a very coarse model of the temporal properties of neuronal responses. Even though these characteristics were not explicitly considered, the stereotypical response of the model neurons bears close resemblance to the peri-stimulus time histogram (PSTH) profiles that have been measured in experiments, with an initial larger response followed by strong suppression and a period of settling reflecting the modulatory inputs to the neuron. Incorporating a more realistic model of the time course of neuronal responses by applying some of the proposals set out in Stevens (2016), such as a finer temporal resolution, latency scatter, and distance dependent delays, could be used to study a wide range of phenomena.

Specifically, combining models of different cell classes and the large spatial spread of lateral connections could be used to explain the spatio-temporal properties of the VSD and optical imaging techniques. Additionally, precisely characterizing the speed of propagation could help in disentangling the role of lateral and feedback connections in contextual modulation.

8.7.7 *Modeling different sensory areas*

As we have emphasized throughout this thesis, the mechanisms that we have identified in the models presented in this thesis underlie fully canonical computations, which are not in any way specific to V1 or even the visual pathway in general. The organization of both receptive fields in the afferent connections and association fields in the lateral connections is an emergent phenomenon driven by the statistics of the

inputs. It would therefore be of great interest to move beyond just primary visual cortex and begin modeling both later stages in the visual processing hierarchy and other cortical areas such as those devoted to auditory or somatosensory cortex or even motor areas, to understand to what extent these mechanisms are truly general.

Modeling other sensory cortices in this way may shed light on the organization and computational principles behind other sensory areas. This is particularly interesting because while visual neuroscience has been extensively characterized in some important aspects, other sensory areas such as auditory or somatosensory cortex are not nearly as well understood. Much could therefore be learned about the role of lateral connections in learning co-occurrence statistics between auditory frequencies, or between whisker activations in rodent barrel cortex. If learning feedforward and feedback inhibition and learning of co-occurrence statistics are indeed general phenomena preserved across the cortex, the model should be able to predict the organization of lateral connections in all these areas.

Furthermore, modeling higher level sensory areas such as V2, V4 or MT could shed light on the kinds of computations they are involved in, by decoding the higher-order correlations embedded in their connectivity. This work would be particularly important because many hierarchical models of vision ignore recurrent and feedback processing entirely in favor of a simpler feedforward arrangement, which ignores a lot of the complexity and models the cortical hierarchy as a static circuit, ignoring context or task dependent modulation of neuronal responses. This could have wide ranging consequences even for modern machine learning approaches, which have generally avoided these complexities.

8.7.8 *Cortical feedback*

Modeling higher-level sensory areas such as V2 also opens up the possibility to start modeling feedback projections to V1, which will be central to building a comprehensive model of primary visual cortex and cortical computation in general. Both lateral connections and feedback projections have the potential to capture, represent, and transmit higher-order statistical relationships in the world and at least according to some theories can represent both prediction and error signals (Rao and Ballard, 1999; Spratling, 2011). Understanding how higher-level areas communicate contextual information to lower-level areas to perform statistical inference and rapidly identify surprising or novel stimuli is one of the major theoretical problems facing neuroscience. The integrative approach taken by the family of models presented here provides one of the best chances to begin understanding at a mechanistic level how different cortical areas interact.

To begin modeling these interactions, the architecture laid out here could simply be stacked, with connections between the higher and lower levels allowed to develop based on the usual Hebbian learning rules. Once such a model has been set up, comparing the feedback connections to the known feedback connections could be used to validate the models, and the analyses developed as part of this thesis could be used to analyze the statistical relationships in the feedback connections, which would go a long way towards understanding what these connections are doing from a computational perspective.

8.7.9 *Neuromodulation of visuo-cortical information processing*

One of the original aims when developing the models that were presented as part of this thesis was to develop a circuit that would allow for realistic modulation of specific neural subtypes and connections. In particular, the models are well suited to investigate how the circuit can be dynamically modulated for task-specific demands, e.g. to model attentional-modulation effects.

Specifically, the existence of individual populations of inhibitory neurons and a robust long-range excitatory network could be used to investigate how different neuromodulatory inputs to the cortex affect the balance between feedforward and recurrent processing. In particular, it is hypothesized that cholinergic and adrenergic neuromodulators are associated with attentional modulation.

One particular target of interest for further investigation is the role of cholinergic inputs from the basal forebrain in mediating attentional effects. Currently there are two main competing theories of how cholinergic inputs to the cortex alters circuit dynamics, outlined in the review by Thiele (2013). The first has been around for at least two decades, and suggests that cholinergic release increases the feedforward drive and indiscriminately disrupts intracortical processing. Recent evidence, however, seems to contradict this hypothesis as a gross simplification, pointing out that different intracortical circuits are affected heterogeneously. I.e., while feedforward drive may be increased, layer 4 spiny cell firing is decreased, leading to weak inputs being filtered out, and Sst-ir cell firing is also up-regulated, indicating some similarity between cholinergic modulation and the regular processing taking place within the V1 circuit under high contrast.

In order to test these hypotheses they could first be modeled using specific and well defined manipulations in the LESPI model. The most basic manipulation would involve increasing thalamocortical drive and reducing the strength of lateral excitatory connections in the cortex, shifting the balance of the circuit towards feedforward processing. Additionally, more specific modulations could be applied by suppressing

the Pv-ir neurons and by steepening the Sst-ir neuron activation function. Overall, these changes should shift the circuit towards a regime akin to the processing that occurs under high contrast, reducing spatial and contextual integration in favor of more accurately representing feedforward input (Roberts et al., 2005; Roberts and Thiele, 2008a,b).

9

CONCLUSION

The overarching goal in computational and theoretical neuroscience has always been providing a comprehensive theory of cortical computation. As we have discussed throughout this thesis, this goal has seemingly not moved much closer over the past decades, despite the tremendous amount of data that has been collected over this time and more recently the huge success of deep learning models in achieving performance that rivals humans in very narrowly defined domains. The main contribution of this thesis is in providing a mechanistic account of how even early sensory areas such as the primary visual cortex can accurately extract and encode higher order statistical relationships, and can then make use of these statistics to dynamically modulate neural responses depending on the context and task-specific demands. We suggest that the ability of cortical areas to encode the dependencies between its inputs is a fundamental property that is shared across cortical areas, allowing the brain to dynamically adapt to contextual and task-specific demands.

Part I

APPENDIX

A

APPENDIX

A.1 HOLOVIEWS: BUILDING COMPLEX VISUALIZATIONS EASILY

HoloViews: Building Complex Visualizations Easily for Reproducible Science

Jean-Luc R. Stevens‡†*, Philipp Rudiger‡†, James A. Bednar‡

https://www.youtube.com/watch?v=hNsR2H7Lrg0

Abstract—Scientific visualization typically requires large amounts of custom coding that obscures the underlying principles of the work and makes it difficult to reproduce the results. Here we describe how the new HoloViews Python package, when combined with the IPython Notebook and a plotting library, provides a rich, interactive interface for flexible and nearly code-free visualization of your results while storing a full record of the process for later reproduction.

HoloViews provides a set of general-purpose data structures that allow you to pair your data with a small amount of metadata. These data structures are then used by a separate plotting system to render your data interactively, e.g. within the IPython Notebook environment, revealing even complex data in publication-quality form without requiring custom plotting code for each figure.

HoloViews also provides powerful containers that allow you to organize this data for analysis, embedding it whatever multidimensional continuous or discrete space best characterizes it. The resulting workflow allows you to focus on exploring, analyzing, and understanding your data and results, while leading directly to an exportable recipe for reproducible research.

Index Terms—reproducible, interactive, visualization, notebook

Introduction

Scientific research alternates between stretches of speculative, exploratory investigation and periods where crucial findings are distilled and disseminated as publications or reports. The exploratory phase typically involves running many different analyses with interactive plotting tools before the important aspects of the data are determined. The final results are then typically prepared as static figures for dissemination, often putting together many subfigures into a complicated figure that reveals multiple interrelated aspects of the results.

Current software tools provide relatively poor support for this dual exploring/reporting nature of scientific research, severely limiting scientific progress. On the one hand, developing new exploratory visualizations typically requires large amounts of custom software coding, which is slow, error-prone, and distracts from the actual scientific analysis. Moreover, this process typically involves a large amount of trial and error, generating transitory code and analyses that make it difficult to later reproduce the steps that led to any particular result [Cro13]. Switching to different tools for final, non-interactive, publication-quality figures exacerbates this problem, further disconnecting the reported results from the process by which they were created. This lack of reproducibility is a serious handicap both for progress within a single lab and for the community as a whole, making it nearly impossible for researchers to build on each others' work even for purely computational projects [Cro13].

Here we will describe a new Python software package built to address these problems directly, by providing simple tools for gradually building elaborate visualizations and analyses interactively yet reproducibly. HoloViews supports immediate exploration of data as it is obtained, without requiring custom coding, and then supports incrementally revealing more complex relationships between datasets, culminating in the final publication of fully reproducible scientific results.

In this paper we will focus on the high-level design principles that allow HoloViews to achieve these goals and we encourage the reader to visit holoviews.org for concrete examples. As detailed below, we show how this is achieved by enforcing a strict separation in the declaration of the semantic properties of the data and the specification of plotting options, allowing the user to declaratively specify their intent and let HoloViews handle the visualization.

The interactive interpreter

To understand this approach, we need to consider the history of how we interact with computational data. The idea of an interactive programming session originated with the earliest LISP interpreters in the late 1950s and remains a popular way to interact with dynamic languages such as Python.

However, like most such command prompts, the standard Python prompt is a text-only environment. Commands are entered by the user, parsed, and executed, with results displayed as text. This offers immediate feedback and works well for data that is naturally expressed in a concise textual form. Unfortunately, this approach begins to fail when the data cannot be usefully visualized as text, as is typical for the large datasets now commonplace. In such instances, a separate plotting package offering a rich graphical display would normally be used to present the results outside the environment of the interpreter, via a graphical user interface.

† *These authors contributed equally.*
∗ *Corresponding author: jlstevens@ed.ac.uk*
‡ *Institute for Adaptive and Neural Computation, University of Edinburgh*

This disjointed approach reflects history: text-only environments, where interactive interpreters were first employed, appeared long before any graphical interfaces. To this day, text-only interpreters are standard due to the relative simplicity of working with text. Proprietary attempts to overcome these limitations, such as the Mathematica Notebook [Wol03], have remained constrained by limited interoperability and a lack of standardized open formats. Other approaches focusing explicitly on reproducibility involve building a recipe for reproducing results only at the end of the scientific project [knitr], when it is often too late to capture the important steps involved. Here we consider how graphical output can be integrated fully into an interactive workflow, addressing both exploration and reproducibility simultaneously.

Fixing the disconnect between data and representation

At the same time as text-based interpreters have failed to overcome the inherent limitations of working with rich data, the web browser has emerged as a ubiquitous means of interactively working with rich media documents. In addition to being universally available, web browsers have the benefit of being based on open standards that remain supported almost indefinitely. Although early versions of the HTML standard only allowed passive page viewing, the widespread adoption of HTML5 has made it possible for anyone to interact with complex, dynamic documents in a bi-directional manner.

The emergence of the web browser as a platform has been exploited by the Python community and the scientific community at large with tools such as the IPython Notebook [Per07] and SAGE MathCloud [Ste05]. These projects offer interactive computation sessions in a notebook format instead of a traditional text prompt. Although similar in design to the traditional text-only interpreters, these notebooks allow embedded graphics or other media (such as video) while maintaining a record of useful commands in a rich document that supports the gradual development of a document with interleaved code, results, and exposition.

Yet despite the greatly improved interactive capabilities of these tools, the spirit of the original interpreter has not yet been restored: there is still an ongoing disconnect between data and its representation. This artificial distinction is a lingering consequence of text-only displays, forcing a strict split between how we conceptualize "simple" and "complex" data. Although the IPython notebook now offers the means to give objects rich media representations, few packages have so far embraced this and none have supported easy composition of related figures. As a result the most common way to visualize complex data remains for the user to specify a detailed list of steps to get subfigures using an external plotting package such as Matplotlib [Hun07], then often combining subfigures using a GUI-based image editor.

Here we introduce HoloViews, a library of simple classes designed to provide an immediately available representation for even complex data in notebooks, analogous to the way simple datatypes are displayed in interactive sessions. HoloViews is not a plotting package; instead, it offers a set of useful data structures paired with rich, customizable visual representations that display effortlessly in the IPython Notebook environment. The result is research that is more interactive, concise, declarative, and reproducible. Figure 1 shows a self-contained example of building a complex visualization showing the declaration of an `Image` object followed by an example of how to compose HoloViews objects together.

Design principles

The core design principle of HoloViews is to *automatically* and *transparently* return and display declarative data structures to the user for immediate feedback without requiring additional code. Although this concept is familiar and intuitive when interactively working with simple data types, it is worth reviewing explicitly what is going on so that the appropriate graphical extension of these ideas is clear.

When executing an addition operation like `1 + 2.5` at a Python prompt, the expression is parsed, converted into bytecode, and then executed, resulting in the float value `3.5`. This floating-point value is immediately returned to the user in the appropriate displayable representation, giving the user immediate feedback. Of course, this representation is not the float itself, but the string `"3.5"`. Such strings are automatically generated by the interpreter, via the displayed object's `__repr__` method.

The Python interpreter also provides such automatic, immediate feedback for more complex data types like large NumPy arrays, but for such data the displayed string has very little utility because it is either incomplete or impractical. In a terminal, this restriction is a result of the `__repr__` method only supporting a text-based display value. Using HoloViews in the IPython Notebook, you can give your array a more useful, interpretable default visual representation as an image, curve, or similar plot according to the following principles:

- It must be easy to assign a useful and understandable default representation to your data. The goal is to keep the initial barrier to productivity as low as possible -- data should simply reveal itself.
- These atomic data objects (elements) should be almost trivially simple wrappers around your data, acting as proxies for the contained arrays along with a small amount of semantic metadata (such as whether the user thinks of some particular set of data as a continuous curve or as a discrete set of points).
- Any metadata included in the element must address issues of *content* and not be concerned with *display* issues -- elements should hold essential information only.
- There are always numerous aesthetic alternatives associated with rich visual representations, but such option settings should be stored and implemented entirely separately from the content elements, so that elements can be generated, archived, and distributed without any dependencies on the visualization code.
- As the principles above force the atomic elements to be simple, they must then be *compositional* in order to build complex data structures that reflect the interrelated plots typical of publication figures.

The outcome of these principles is a set of compositional data structures that contain only the essential information

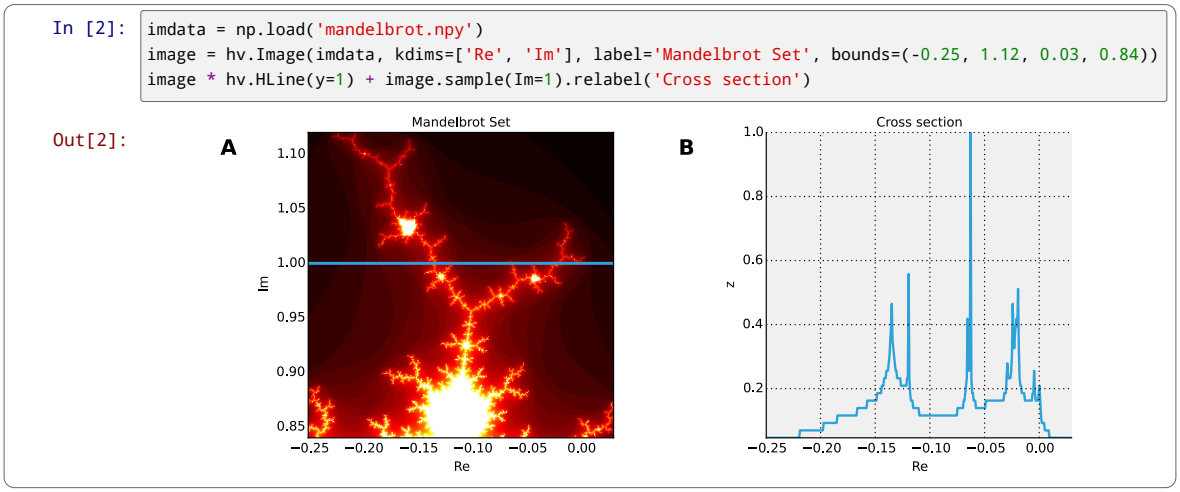


Fig. 1: *Example of a composite HoloViews data structure and how it is displayed in an IPython Notebook session. The* `imdata` *array loaded using Numpy corresponds to the displayed portion of the Mandelbrot set.* ***A.*** *The* `Image` *element displays* `imdata` *overlaid via the* `*` *operator with a horizontal line element (*`HLine`*).* ***B.*** *A* `Curve` *element generated via the* `.sample()` *method of the image, showing a cross-section of the fractal along the indicated blue horizontal line. The curve is concatenated with the* `Overlay` *in* ***A*** *via the* `+` *operation.*

underlying potentially complex, publication-quality figures. These data structures have an understandable, default visualization that transparently reveals their contents, making them a useful proxy for the data itself, just as the text `3.5` is a proxy for the underlying floating-point value. This default visualization may then be customized declaratively to achieve the desired aesthetics, without complicating the objects themselves.

In the next section we will discuss the data structures that hold the important content. Starting with the simple primitive elements, we examine how they can be composed into complex figures and embedded in high-dimensional spaces for exploration. Along the way we will discover how our implementation realizes the design principles outlined and manages to keep the state of the data separate from its visual representation.

Data Structures

In this section we discuss the data structures that hold the raw data and the essential semantic content of interest. The Elements section introduces each of the primitives, and the Collections section explains how they can be combined. Finally, we will discuss working with Elements embedded in high-dimensional continuous or discrete spaces.

Elements

The atomic classes that wrap raw data are the `Element` primitives. These classes are named by the natural representation they suggest for the supplied data, with `Image`, `Curve`, and `Scatter` being some simple examples. These elements are easily constructed as they only require the raw data (such as a NumPy array) to display.

In Figure 1, we have some examples of the Element primitives. On the left, in subfigure **A**, we see the `Image` primitive containing a two-dimensional NumPy array. This `Image` is declared by supplying the NumPy array `imdata` along with the optional metadata, including a suitable label and a declaration of the bounding region in the complex plane. The visual output is automatically generated and shows that the array is a part of the Mandelbrot set. Our object merely holds the supplied NumPy array, which remains easily accessed via the `.data` attribute. In part **B** of Figure 1 we have an example of a `Curve` containing a horizontal cross section of the image, as computed by the `sample` method.

Although the names of the `Elements` suggest that these objects are about visualization, they are primarily concerned with content and *not* display. The visually meaningful class names offer a convenient way to intuitively understand the dimensionality of the data in terms of an appropriate visual representation. For instance, in Figure 1 **A**, the name `Image` conveys the notion that the contained data is in the form of a two-dimensional NumPy array that can be meaningfully displayed as an image.

The particular `Image` shown in Figure 1 **A** was constructed as a visualization of the Mandelbrot Set, defined in the complex plane. In particular, the `kdims` argument declares that the x-axis is along the real axis and that the y-axis is along the imaginary axis. This information is then reflected in the visual output by assigning the appropriate axis labels. This semantic information is also passed to the `Curve` object generated by sampling the image using `image.sample(Im=1)`.

This `Curve` object is also able to pass on this semantic information to other Elements with different visual representations so that they faithfully reflect the space in which the Mandelbrot Set is defined. For instance, you can pass the curve directly to the constructor of the `Scatter` or `Histogram` elements and a new visual representation of the resulting object will retain the original semantic dimension labels. This type of operation merely changes the representation associated with the supplied data.

Note that in the declarations of `Image`, the dimensions of the axes are declared as key dimensions (`kdims`). Key

dimensions correspond to the independent dimensions used to index or slice the element, with the remaining dimensions called value dimensions (`vdims`). In the case of this image, there is a single value dimension, for the values in the supplied NumPy array, which are then visualized using the default colormap of the `Image` elements (the 'hot' color map).

As key dimensions are indexable and sliceable, we can slice the `Image` to select a different subregion of the Mandelbrot Set. Continuous values are supported when slicing an `Image` and the result is then a new `Image` containing the portion of the original NumPy array appropriate to the specified slice. The mapping between continuous space and the discrete array samples is specified by the bounds, allowing us to apply the slice `[-0.2:0,0.85:1.05]` to select the corresponding part of the complex plane. The first component of this slice selects the first key dimension (the real axis `'Re'`) from *-0.2* to *0.0* while the second component of the slice selects the second key dimension (the imaginary axis `'Im'`) from *0.85* to *1.05*. You can apply a similar slice along the real axis to select a portion of the curve object shown in Figure 1 **B**.

There are many additional element classes, one for each of the common visual representations for data. These elements form an extensible library of primitives that allow the composition of data structures with complex, meaningful visualizations. Within the set of all elements, you can cast your data between representations so long as the number of key and value dimensions is consistent. You can then index and slice your elements along their respective key dimensions to get new elements holding the appropriately sliced data of interest.

Collections

The elements are simple wrappers that hold the supplied data and allow a rich, meaningful default representation. An individual element is therefore a data structure holding the semantic contents corresponding to a simple visual element of the sort you may see in a publication. Although the elements are sufficient to cover simple cases such as individual graphs, raster images, or histogram, they are not sufficient to represent more complex figures.

A typical published figure does not present data using a single representation, but allows comparison between related data items in order to illustrate similarities or differences. In other words, a typical figure is an object composed of many visual representations combined together. HoloViews makes it trivial to compose elements in the two most common ways: concatenating representations into a single figure, or overlaying visual elements within the same set of axes.

These types of composition are so common that both have already been used in Figure 1 as our very first example. The + operation implements concatenation, and * implements overlaying elements together. When you compose an object using the + operator, a default four-column layout is used but you can specify the desired number of columns using the `.cols` method. Layouts are easily specified but also support multiple options for customizing the position and sizing of elements.

When we refer to subfigures 1 **A** and 1 **B**, we are making use of labels generated by HoloViews for representing a composite data structure called a `Layout`. Similarly, subfigure 1 **A** is itself a composite data structure called an `Overlay` which, in this particular case, consists of an `Image` element overlaid by the `HLine` element.

The overall data structure that corresponds to Figure 1 is therefore a `Layout` which itself contains another composite collection in the form of an `Overlay`. The object in Figure 1 is in fact a highly flexible, compositional tree-based data structure: intermediate nodes correspond either to `Layout` nodes (+) or `Overlay` nodes (*), with element primitives at the leaf nodes. Even in this potentially complex tree, all the raw data corresponding to every visual element is conveniently accessible via key or attribute access by selecting a leaf element using its path through the tree, and then inspecting the `.data` attribute, making it simple to declare which part of a complex dataset you want to work with at a given time.

As any element may be a leaf of such a tree, there needs to be an easy way to select subtrees or leaf elements. This is achieved with a semantic, two-level labeling system using "group" and "label" strings supported throughout HoloViews. We have seen an example of a label string in Figure 1, where it was used to title the image "Mandelbrot Set". The textual representation of the layout in Figure 1 (see Out[6] of Figure 4) shows how the supplied label is used in the attribute-based indexing scheme of the layout. The strings "Image", "Overlay", "HLine" and "Curve" are default group names, but you can supply your own names to define semantic groupings for your data. To illustrate this system, you can access the sampled data (a NumPy array) in Figure 4 using `content.Curve.Cross_Section.data`.

With the ability to overlay or concatenate any element with any other, there is great flexibility to declare complex relationships between elements. Whereas a single element primitive holds semantic information about a particular piece of data, trees encode semantic information between elements. The composition of visual elements into a single visual representation expresses some underlying semantic value in grouping these particular chunks of data together. This is what composite trees capture; they represent the overall *semantic content* of a figure in a highly composable and flexible way that always preserves both the raw data and associated metadata for further interactive analysis and reproduction.

Spaces

A single plot can represent at most a few dimensions before it becomes visually cluttered. Since real-world datasets often have higher dimensionality, we face a tradeoff between representing the full dimensionality of our data, and keeping the visual representation intelligible and therefore effective. In practice we are limited to two or at most three spatial axes, in addition to attributes such as the color, angle, and size of the visual elements. To effectively explore higher dimensional spaces we therefore have to find other solutions.

One way of dealing with this problem is to lay out multiple plots spatially. Plotting packages like ggplot [Wic09] and seaborn [Was14] have shown how this can be done easily using

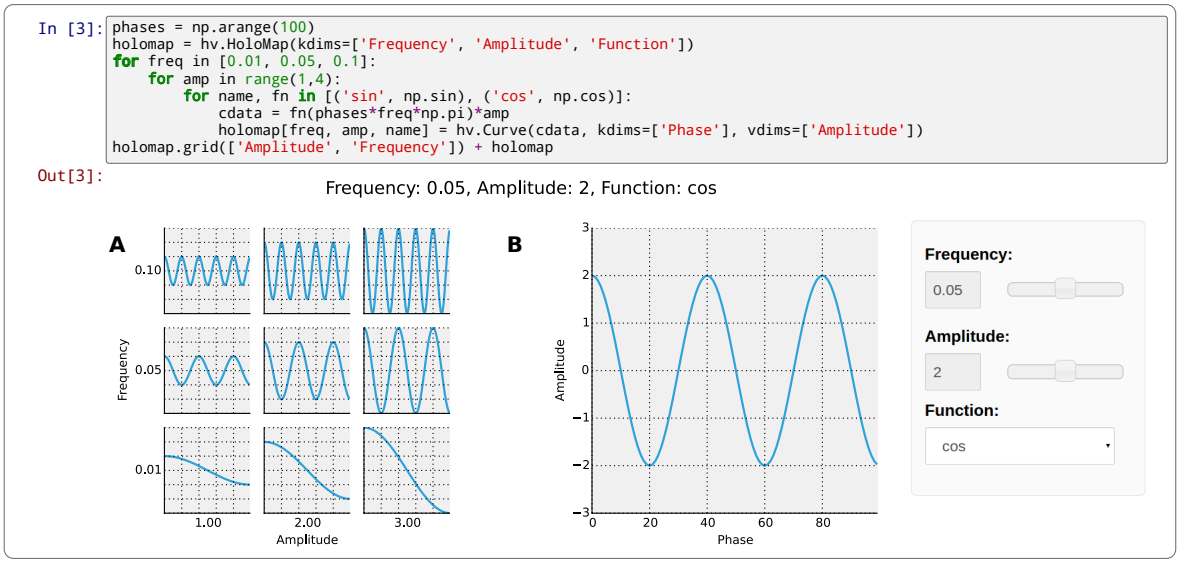


Fig. 2: *Example of a Layout object containing two different representations of a multi-dimensional space. Both representations contain* `Curve` *objects embedded in three dimensions (*`Frequency`*,* `Amplitude`*,* `Function`*), but not all of these dimensions can be visualized at once. In* ***A****, two of the dimensions are mapped onto the rows and columns of a grid, and the remaining* `Function` *dimension can be selected using the widget at the right. In* ***B****, only a single curve is shown, with the three sliders at the right together selecting the appropriate curve from the 3D HoloMap space. When two* `HoloMaps` *are joined in a* `Layout` *like this, it will automatically find the joint set of dimensions the HoloMaps can be varied over. In this way HoloMaps allow users to explore data naturally and conveniently even when its dimensionality exceeds what can be sensibly displayed on the screen at once.*

various grid-based layouts. Another solution is to present the data sequentially over time as an animation. A third solution is to provide interactive control, allowing the user to reveal further dimensionality by interacting with the plots using various widgets.

HoloViews provides support for all three of these approaches, via composable data structures that embed collections of `Element` objects in any arbitrarily dimensioned space. Fundamentally, this set of data structures (subclasses of `NdMapping`) are multi-dimensional dictionaries that allow the user to declare the dimensionality of the space via a list of key dimensions (`kdims`).

The list of supported `NdMapping` classes includes:

- `HoloMaps`: The most flexible high-dimensional data structure in HoloViews, allowing `Element` instances to be embedded in an arbitrarily high-dimensional space, to be rendered either as a video animation or as an interactive plot that allows exploration via a set of widgets.
- `GridSpaces`: A data structure for generating spatial layouts with either a single row (1D) or a two-dimensional grid. Each overall grid axis corresponds to a key dimension.
- `NdLayouts`/`NdOverlays`: Similar to `Layout` or `Overlay` objects, where the contained objects vary over one or more dimensions.

To explore a high-dimensional space of height as a function of age across different countries and years, you could declare `space=HoloMap(kdims=['Country', 'Year'])`. Now we can treat `space` as a dictionary and insert instances of classes such as `Curve` or `Scatter` with the appropriate `(country, year)` keys. For instance, the age and height `Curve` for the USA in 1988 (`usa`) can be inserted using `space['USA', 1988] = usa`. Note that the order of the indexing corresponds to the order of the declared key dimensions.

All of the above classes are simply different ways to package and view a high-dimensional dataset. Just as with `Elements`, it is possible to cast between these different spaces via the constructor. In addition, they can all be tabularized into a HoloViews `Table` element or a pandas `DataFrame` [McK10], a feature that is also supported by the `Element` primitives.

To get a sense of how composing data and generating complex figures works within this framework, we explore some artificial data in Figure 2. Here we vary the frequency and amplitude of sine and cosine waves, demonstrating how we can quickly embed this data into a multi-dimensional space. First, we declare the dimensions of the space we want to explore as the key dimensions (`kdims`) of the HoloMap. Next, we populate the space iterating over the frequencies, amplitudes, and the two trigonometric functions, generating each `Curve` element individually and assigning to the HoloMap at the correct position in the space.

We can immediately go ahead and display this HoloMap either as an animation or using the default widgets, as in Figure 2 **B**. Visualizing individual curves in isolation is not very useful, of course; instead we probably want to see how the curves vary across `Frequency` and `Amplitude` in a single plot. A `GridSpace` provides such a representation and by using the space conversion method `.grid()` we can easily transform our three-dimensional HoloMap into a two-dimensional GridSpace (which then allows the remaining dimension, the choice of trigonometric function, to be varied

via the drop-down menu). Finally, after composing a `Layout` together with the original `HoloMap`, we let the display system handle the plotting and rendering.

If we decide that a different representation of the data would be more appropriate, it is trivial to rearrange the dimensions without needing to write new plotting code. Even very high-dimensional spaces can be condensed into an individual plot or expressed as an interactive plot or animation, by simply specifying which part of the data we are interested in rather than writing new brittle and error-prone custom plotting code.

Customizing the visual representation

In this section we show how HoloViews achieves a total separation of concerns, keeping the composable data structures introduced above completely separate from both customization options and the plotting code. This design is much like the separation of content and presentation in HTML and CSS, and provides the same benefits of making the content easily maintainable while the presentation is easily controllable.

The only required connection between the above data structures and the custom display options is a single, automatically managed integer. Using this integer attribute we can make the data structures behave as if they were rich, stateful, and individually customizable objects, without actually storing anything to do with visualization on the objects. We will show how this separation is useful and extensible so that the user can quickly and easily customize almost every aspect of their plot. For instance, it is easy to change the font size of text, change the subfigure label format, change the output format (e.g. switch from PNG to SVG) and even alter the plotting backend (currently defaulting to Matplotlib) without changing any part of the underlying object being rendered.

Figure 3 provides an overall summary of how the different components in the display system interact. The declarative data structures define what will be plotted, specifying the arrangements of the plots, via Layouts, Overlays, and spaces. The connection between the data structure and the rendered representation is made according to the object type, the aforementioned integer attribute, and optionally specified group and label strings. By collecting the display options together and associating them with particular objects via these attributes, the visual representation of the content may be easily customized, e.g. to tweak aesthetic details such as tick marks, colors and normalization options. Once the user has specified both content and optionally customized the display the rendering system looks up the appropriate plot type for the object in a global registry, which then processes the object and looks up the specified options in order to display it appropriately. This happens transparently without any input from the user. Once the plotting backend has rendered the plot in the appropriate format, it will be wrapped in HTML for display in the notebook.

The default display options are held on a global tree structure similar in structure to the composite trees described in the previous section, but with nodes holding custom display options in the form of arbitrary keywords. In fact, these option trees also use labels and groups the same way as composite

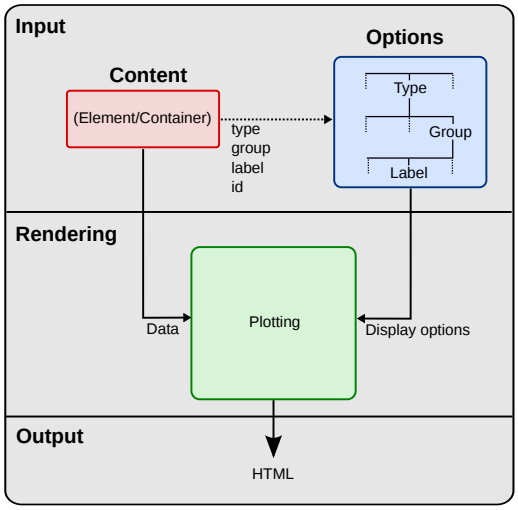


Fig. 3: *This view of the HoloViews display and customization systems illustrates the complete separation between the content (data) to be displayed, the display options, and the rendering/plotting system. The display options are stored entirely separately from the content as a tree structure, with the appropriate options being selected with user-controllable levels of specificity: general options for all objects of a given type, more specific options controlled by user-definable* `group` *and* `label` *strings, or arbitrarily specific options based on the integer* `id` *assigned to each content object. Plotting and rendering happens automatically through the use of IPython display formatters. These combine the content with the specified display options, call an external plotting library, which returns an HTML representation that can then be rendered in the notebook.*

trees except they additionally support type-specific customization. For instance, you may specify colormap options on the `Image` node of the tree that will then be applied to all `Images`. If this chosen colormap is not always suitable, you can declare that all `Image` elements belonging to a group (e.g. `group='Fractal'`) should use a different colormap by overriding it on the `Image.Fractal` node of the tree. This form of inheritance allow you to specify complex yet succinct style specifications, applying to all objects of a particular type or just to specific subsets of them.

To explore how option setting works in practice, Figure 4 shows an example of customizing Figure 1 with some basic display options. Here we use an optional but highly succinct method for setting the options, an IPython cell magic `%%opts`, to specify aspect ratios, line widths, colormaps, and sublabel formats. By printing the string representation of the content (`Out[6]`) and the options (`Out[7]`), we can see immediately that each entry in the options tree matches a corresponding object type. Finally, in the actual rendered output, we can see that all these display options have taken effect, even though the actual data structure differs from the object rendered in Figure 1 only by a single integer attribute.

A major benefit of separating data and customization options in this way is that all the options can be gathered in one place. There is no longer any need to dig deep into the documentation of a particular plotting package for a particular option, as all the options are easily accessible via a tab-completable IPython magic and are documented via the `help`

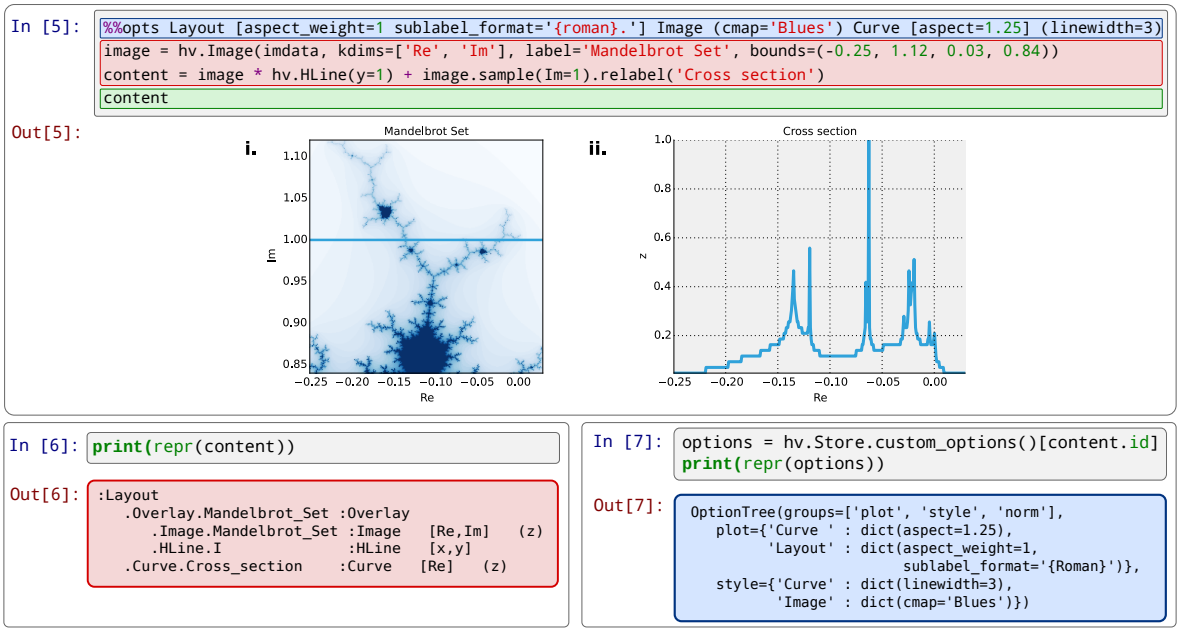


Fig. 4: *An example of customizing the display of Figure 1's data using the default Matplotlib backend.* `In[5]` *is color coded according to the components in Figure 3, where red is the content, blue is the display options (using an optional IPython-specific succinct syntax), and green is what triggers the the rendering.* `Out[5]` *shows how the supplied options have affected the final plots, compared to Figure 1. Finally,* `Out[6]` *and* `Out[7]` *show the textual representations of the content and the style specification respectively, demonstrating how the two are separate yet linked.*

function. This ease of discovery enables a workflow where the visualization details of a plot can be easily and quickly iteratively refined once the user has found data of interest.

The options system is also inherently extendable. New options may be added at any time, and will immediately become available for tab-completion. In fact, the plotting code for each element and container type may be switched out completely and independently, and the options system will automatically reflect the changes in the available customization options. This approach lets the user work with a variety of plotting backends at the same time, without even having to worry about the different plotting APIs.

The separation between content, options and plotting explicitly supports the workflows that are common in science, repeatedly switching between phases of exploration and periods of writing up. Interesting data can be collected and curated over time, where each step is instantly and transparently visualizable without any custom code cluttering up the notebook. Visualizations of data that are worth keeping can be customized through an interactive and iterative process, and the final set of plotting options can then be expressed as a single data structure separate from the actual displayed data, ready to be applied to the next batch of data from a subsequent measurement or experiment. Throughout, the scientist curates the data of interest, as revealed in associated visual representations, along with the visualization options and a separate codebase of general-purpose plots (mostly included in HoloViews, but potentially extended locally for specific domains). Each of these three aspects of the process (data, options, and code) can be developed, maintained, archived, and improved independently, providing comprehensive support for the natural process of exploration and dissemination common to all scientific disciplines.

Discussion

This paper demonstrates a succinct, flexible, and interactive approach for data exploration, analysis, and visualization. HoloViews restores the immediate feedback cycle that is characteristic of working with simple data in an interpreter. This is achieved by having declarative objects display themselves with good defaults allowing the user to immediately understand their data. In the majority of cases this eliminates the need to write plotting code and allows the user to keep a concise and reproducible recipe of their work, from exploration to the final publication. HoloViews thus allows scientists to capture the entire workflow involved in a research project.

Without a strictly enforced separation of concerns, workflow stages often end up mixing both data processing and visualization. Although a displayed representation is always necessary for understanding, it has been a dead end for further data processing. Because HoloViews objects represent themselves visually but also contain the raw data, the ability to continue processing is never terminated and exploration can continue. Furthermore, the chosen representation can easily be changed, turning what used to be a highly disjointed workflow into a open-ended process concerned with the semantics of the data. Only once results worth disseminating are attained does it become necessary to consider the details of visualization.

The compositionality of HoloViews is superficially reminiscent of systems such as the Grammar of Graphics [Wil05] for the R language, but the aim of HoloViews is quite different. Instead of expressing all the complexities of graphics, the

declarative data structures in HoloViews define a language for the semantics of the actual data. This language focuses on how the researcher conceptualizes it, *independent* of the exact details of plotting. The need for an automatic and useful visual representation is driven by the need to immediately present the data in a meaningful format.

HoloViews is one of many packages designed for working with large, multidimensional datasets, but it differs from each of these in important ways. For instance, Python's `seaborn` [Was14] and R's `ggplot2` [Wic09] library support laying out high-dimensional data into subplots and grids, while Python's Bokeh library and R's shiny [shiny] web application framework provide widgets for interactive data exploration. While each of these packages can provide extremely polished interactive graphics, getting them set up for specific sets of data requires significant additional effort and custom code, placing a barrier to their primary use case, the interactive exploration of data. HoloViews instead tries to avoid custom coding altogether as far as possible, with users instead supplying metadata to declare the properties of the data and option settings to control its visual appearance.

Although HoloViews is a general purpose library for working with data at every stage, it actually represents a significant advance over previous approaches focused only on achieving reproducibility of the final result. Simply by keeping specifications for figures succinct, HoloViews allows the entire recipe to be preserved in the notebook, not scattered over separately imported plotting code files. Secondly, because HoloViews can directly express the complex relationships between different bits of data as subfigures, it can capture entire figures within notebooks that would previously have required unreproducible work in external drawing programs. Lastly, HoloViews exports the actual data alongside published figures, allowing it to be tested automatically (as is done for the project web site) without conflating it with arbitrary display choices. HoloViews makes it possible to reproduce results from every step of the project, up to and including the final published figures, in a way that has not previously been practical.

Although HoloViews aims to provide good default behavior, scientific work often requires highly specialized visualizations. For that reason we have made it easy to extend the defaults and integrate new visualizations. Firstly, as many plotting and styling options as possible are exposed in an easily accessible manner, while providing a powerful, inheritance-based system for changing these options when required. Secondly, the options system has been designed to work well with the compositional data structures provided by HoloViews. Thirdly, HoloViews makes it trivial to add completely novel types of Elements with corresponding plots (or to override specific code in existing plots) using custom code when needed, and these custom plots will then combine seamlessly with other objects to make composite figures. Finally, not only is it possibly to implement new plot classes but entire plotting backends may be added and exposed to the user, such as the prototype Bokeh backend, which is well suited to live interaction and large datasets. Thus default plots are simple and straightforward, but even complex figures are easily achievable. Many such examples, ranging from simple to complex, can be found in the Tutorials and Examples sections of holoviews.org.

In this paper, we have focused on how a user can quickly build data structures for their content of interest. An even more powerful approach is for a developer to integrate HoloViews directly into a library, analysis tool, or simulator. By returning HoloViews objects (which do not depend on any plotting library), any Python package can immediately have access to flexible, compositional data structures that automatically double as a visualization system. This is exactly the approach taken by the ImaGen image generation library and the Topographica neural simulator, two very different projects that both output data wrapped in HoloViews data structures.

Conclusion

Based on the key principles of: (1) making data immediately and transparently visualizable, (2) associating data directly with its semantic description, (3) keeping display option settings separate from the data, (4) keeping display code separate from both data and display options, (5) explicitly expressing the relationships between data elements compositionally, and (6) keeping the original data accessible even in complex visualizations, Holoviews supports the entire life cycle of scientific research, from initial exploration, to dissemination and publication, to eventual reproduction of the work and new extensions. Existing approaches for achieving some of these goals individually have been very limiting and only partially successful, each adding significant new costs along with the benefits they offer. HoloViews instead addresses the underlying problems fundamental to current methods for scientific research, solving seemingly intractable issues like reproducibility almost as a side effect of properly supporting the basic process of doing science.

Acknowledgments

This work was funded in part by grant 1R01-MH66991 to the University of Texas at Austin from the USA National Institute of Mental Health, by grant EP/F500385/1 from the UK EPSRC and MRC research councils, and by the Institute for Adaptive and Neural Computation at the University of Edinburgh.

References

[Cro13] Crook et al., "Learning from the Past: Approaches for Reproducibility in Computational Neuroscience", *20 Years of Computational Neuroscience*, J.M. Bower, ed., Springer, 9:73-102, 2013.

[Wol03] Stephen Wolfram, *The Mathematica Book*, Fifth Edition, Wolfram Media/Cambridge University Press, 2003.

[knitr] Foundation for Open Access Statistics, *knitr*, http://yihui.name/knitr, 2015.

[Per07] Fernando Perez and Brian E. Granger, IPython: a System for Interactive Scientific Computing, *Computing in Science and Engineering*, 9:21-19, 2007.

[Ste05] William Stein and David Joyner. SAGE: System for Algebra and Geometry Experimentation. *ACM SIGSAM Bulletin*, 39:61-64, 2005.

[Hun07] John D. Hunter, *Matplotlib: A 2D graphics environment*, Computing In Science & Engineering, 9(3):90-95, 2007.

[Wic09] Hadley Wickham, *ggplot2: elegant graphics for data analysis*, Springer New York, 2009.

[Was14] Michael Waskom et al.. *seaborn: v0.5.0*, Zenodo. 10.5281/zenodo.12710, November 2014.

[McK10] Wes McKinney, *Data Structures for Statistical Computing in Python*, Proceedings of the 9th Python in Science Conference, 51-56, 2010.

[Wil05] Leland Wilkinson, *The Grammar of Graphics*, Springer-Verlag New York, 2005.

[shiny] RStudio, Inc, *shiny: Easy web applications in R.*, http://shiny.rstudio.com, 2014.

A.2 MODEL PARAMETERS

Symbol	Description	Value
	RGC/LGN	
γ_A	Strength of the retinogeniculate projection	14
γ_L	Strength of the lateral gain control	0.6
σ_C	Size of the center component	0.2
σ_S	Size of the surround component	0.3
σ_L	Size of the lateral gain control component	0.25
c	Constant offset of the lateral gain control function	0.11
κ_A	Cut-off radius of the afferent projection	0.3
κ_L	Cut-off radius of the lateral gain control projection	0.5
	SCAL V1	
γ_A	Strength of the geniculocortical projections	2.4
γ_E	Strength of the local excitatory projection	1.6
γ_I	Strength of the divisive inhibitory projection	1.8
γ_{LE}	Strength of lateral excitatory projection	0
σ_E	Size of local excitatory projection	0.06
σ_I	Size of inhibitory projection	0.12
σ_L	Size of lateral excitatory projection	1.25
κ_A	Cut-off radius of afferent projection	0.5
κ_E	Cut-off radius of local excitatory projection	0.14
κ_I	Cut-off radius of inhibitory projection	0.18
κ_L	Cut-off radius of lateral excitatory projection	1.25
λ_A	Learning-rate of the afferent projections	0.1
λ_E	Learning-rate of the local excitatory projection	0
λ_I	Learning-rate of the inhibitory projection	0.3
λ_L	Learning-rate of the lateral excitatory projection	1
τ_L	Time constant of lateral projection hysteresis function	0.2
	SEPI V1 Excitatory (Only changed)	
γ_A	Strength of the afferent projection	3
γ_E	Strength of local excitatory projection	2.5
γ_L	Strength of lateral excitatory projection	0
λ_A	Learning-rate of the afferent projections	0.25
λ_E	Learning-rate of the local excitatory projection	0
λ_L	Learning-rate of the lateral excitatory projection	1
	SEPI V1 Parvalbumin (Only changed)	
γ_P	Strength of the divisive inhibitory projection	2.5
γ_{RP}	Strength of the recurrent divisive inhibitory projection	0.8
λ_P	Learning rate of lateral inhibitory projection	0.25
	LESPI V1 Excitatory (Only changed)	
γ_L	Strength of lateral excitatory projection	3
γ_{ES}	Strength of local excitatory projection to Sst cells	0.6
γ_{LS}	Strength of lateral excitatory projection to Sst cells	6
λ_{LS}	Learning rate of lateral excitatory projection to Sst cells	5
β_{LS}	Exponent of the lateral excitatory projection to Sst cells	3
	LESPI V1 Somatostatin	
γ_S	Strength of Sst inhibition projection	-1
τ_S	Time constant of Sst neuron hysteresis function	0.2
λ_S	Learning rate of Sst inhibition projection	0.1

LIST OF FIGURES

BIBLIOGRAPHY

Abbott, L. F., Varela, J. a., Sen, K., and Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science (New York, N.Y.)*, 275(5297):220–4.

Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z. J., and Scanziani, M. (2012). A neural circuit for spatial summation in visual cortex. *Nature*, 490(7419):226–31.

Ahissar, M. and Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8(10):457–64.

Albus, K. and Wahle, P. (1994). The Topography of Tangential Inhibitory Connections in the Postnatally Developing and Mature Striate Cortex of the Cat. *European Journal of Neuroscience*, 6:779–792.

Angelucci, A. and Bressloff, P. C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Progress in Brain Research*, 154(6):93–120.

Angelucci, A. and Bullier, J. (2003). Reaching beyond the classical receptive field of V1 neurons: horizontal or feedback axons? *Journal of Physiology*, 97(2-3):141–54.

Angelucci, A., Levitt, J. B., and Lund, J. S. (2002a). Anatomical Origins of the Classical Receptive Field of Single Neurons in Macaque Visual Cortical Area V1. *Progress in Brain Research*, 136:373–88.

Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupe, J.-M., Bullier, J., and Lund, J. S. (2002b). Circuits for local and global signal integration in primary visual cortex. *The Journal of Neuroscience*, 22(19):8633–46.

Angelucci, A. and Sainsbury, K. (2006). Contribution of Feedforward Thalamic Afferents and Corticogeniculate Feedback to the Spatial Summation Area of Macaque V1 and LGN. *The Journal of Comparative Neurology*, 498(April):330–351.

Antolík, J. (2010). *A unified developmental model of maps, complex cells and surround modulation in the primary visual cortex*. Doctoral thesis, The University of Edinburgh.

Antolík, J. and Bednar, J. A. (2011). Development of maps of simple and complex cells in the primary visual cortex. *Frontiers in Computational Neuroscience*, 5(April):17.

Atallah, B. V., Bruns, W., Carandini, M., and Scanziani, M. (2012). Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron*, 73(1):159–70.

Atick, J. J. and Redlich, N. A. (1992). What does the Retina Know about Natural Scenes? *Neural Computation*, 4:196–210.

Bair, W., Cavanaugh, J. R., and Movshon, J. A. (2003). Time course and time-distance relationships for surround suppression in macaque V1 neurons. *The Journal of Neuroscience*, 23(20):7690–701.

Barlow, H. B. and Földiák, P. (1989). Adaptation and decorrelation in the cortex. In Durbin, R., Miall, C., and Mitchison, G., editors, *The Computing Neuron*. Addison-Wesley.

Bartley, A. F., Huang, Z. J., Huber, K. M., and Gibson, J. R. (2008). Differential activity-dependent, homeostatic plasticity of two neocortical inhibitory circuits. *Journal of Neurophysiology*, 100(4):1983–94.

Bear, M. F., Connors, B., and Paradiso, M. (2006). *Neuroscience: Exploring the Brain*. Lippincott Williams and Wilkins, 3rd revise edition.

Bednar, J. a. (2012). Building a mechanistic model of the development and function of the primary visual cortex. *Journal of Physiology*.

Bednar, J. a. and Miikkulainen, R. (2003). Self-organization of spatiotemporal receptive fields and laterally connected direction and orientation maps. *Neurocomputing*, 52-54(02):473–480.

Beierlein, M., Gibson, J. R., and Connors, B. W. (2003). Two dynamically distinct inhibitory networks in layer 4 of the neocortex. *Journal of Neurophysiology*, 90(5):2987–3000.

Bezrukov, S. and Vodyanoy, I. (1997). Stochastic resonance in non-dynamical systems without response thresholds. *Nature*, 385:319–321.

Bienenstock, E. L., Cooper, L. N., and Munro, P. W. (1982). Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *The Journal of Neuroscience*, 2(1):32–48.

Binzegger, T., Douglas, R. J., and Martin, K. a. C. (2004). A quantitative map of the circuit of cat primary visual cortex. *The Journal of Neuroscience*, 24(39):8441–53.

Blakemore, C., Carpenter, R., and Georgeson, M. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature*, 228(228):37–39.

Blasdel, G. G. (1992). Orientation Selectivity, Striate Cortex Preference, and Continuity in Monkey Striate Cortex. *The Journal of Neuroscience*, 12(8):313903161.

Blasdel, G. G. and Fitzpatrick, D. (1984). Physiological organization of layer 4 in macaque striate cortex. *The Journal of Neuroscience*, 4(3):880–895.

Bonin, V., Mante, V., and Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus. *The Journal of Neuroscience*, 25(47):10844–56.

Bosking, W. H., Zhang, Y., Schofield, B., and Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, 17(6):2112–27.

Budd, J. M. and Kisvárday, Z. F. (2001). Local lateral connectivity of inhibitory clutch cells in layer 4 of cat visual cortex (area 17). *Experimental Brain Research*, 140(2):245–250.

Bullier, J., Hupe, J.-M., James, A. C., and Girard, P. (2001). The role of feedback connections in shaping the responses of visual cortical neurons. *Progress in Brain research*, 134:193–204.

Burkhalter, A. (2008). Many specialists for suppressing cortical excitation. *Frontiers in neuroscience*, 2(2):155–67.

Buzás, P., Eysel, U. T., Adorján, P., and Kisvárday, Z. F. (2001). Axonal topography of cortical basket cells in relation to orientation, direction, and ocular dominance maps. *The Journal of comparative neurology*, 437(3):259–85.

Buzas, P., Kovacs, K., Ferecskó, A. S., Budd, J. M., Eysel, U. T., and Kisvárday, Z. F. (2006). Model-Based Analysis of Excitatory Lateral Connections in the Visual Cor-

tex. *The Journal of comparative neurology*, 881(May):861–881.

Cang, J., Rentería, R. C., Kaneko, M., Liu, X., Copenhagen, D. R., and Stryker, M. P. (2005). Development of precise maps in visual cortex requires patterned spontaneous activity in the retina. *Neuron*, 48(5):797–809.

Carandini, M. and Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1):51–62.

Carandini, M., Heeger, D. J., and Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(22):10053–65.

Cardin, J. A., Palmer, L. A., and Contreras, D. (2007). Stimulus feature selectivity in excitatory and inhibitory neurons in primary visual cortex. *The Journal of Neuroscience*, 27(39):10333–10344.

Cavanaugh, J. R., Bair, W., and Movshon, J. A. (2002a). Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. *Journal of Neurophysiology*, 88(5):2530–46.

Cavanaugh, J. R., Bair, W., and Movshon, J. A. (2002b). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology*, 88(5):2547–56.

Chapman, B., Stryker, M. P., and Bonhoeffer, T. (1996). Development of orientation preference maps in ferret primary visual cortex. *The Journal of Neuroscience*, 16(20):6443–53.

Coen-Cagli, R., Kohn, A., and Schwartz, O. (2015). Flexible gating of contextual influences in natural vision. *Nature neuroscience*.

Coen-Cagli, R. and Schwartz, O. (2013). The impact on midlevel vision of statistically optimal divisive normalization in V1. *Journal of Vision*, 13:1–20.

Connolly, M. and Van Essen, D. (1984). The representation of the visual field in parvicellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. *The Journal of Comparative Neurology*, 226(4):544–64.

Cruikshank, S. J., Lewis, T. J., and Connors, B. W. (2007). Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex. *Nature Neuroscience*, 10(4):462–8.

Davison, A., Brüderle, D., Kremkow, J., Muller, E., Pecevski, D., Perrinet, L., and Yger, P. (2008). Pynn: a common interface for neuronal network simulators. *Frontiers in Neuroinformatics*, 2:11.

DeAngelis, G. and Robson, J. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex. *Journal of Neurophysiology*, 68(1):144–163.

Derrington, A. M. and Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *The Journal of Physiology*, 357:219–40.

Di Cristo, G., Wu, C., Chattopadhyaya, B., Ango, F., Knott, G., Welker, E., Svoboda, K., and Huang, Z. J. (2004). Subcellular domain-restricted GABAergic innervation in primary visual cortex in the absence of sensory and thalamic inputs. *Nature Neuroscience*, 7(11):1184–6.

Dosher, A. B. and Lu, Z.-l. (2000). Noise Exclusion in Spatial Attention. *Psychological Science*, 11(2):139–146.

Douglas, R. and Martin, K. (1991). A functional microcircuit for cat visual cortex. *The Journal of Physiology*, 440:735–769.

Dragoi, V. and Sur, M. (2000). Dynamic properties of recurrent inhibition in primary visual cortex: contrast and orientation dependence of contextual effects. *Journal of Neurophysiology*, 83(2):1019–1030.

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., and Rasmussen, D. (2012). A large-scale model of the functioning brain. *science*, 338(6111):1202–1205.

Fagiolini, M., Fritschy, J.-M., Löw, K., Möhler, H., Rudolph, U., and Hensch, T. K. (2004). Specific GABAA circuits for visual cortical plasticity. *Science (New York, N.Y.)*, 303(5664):1681–3.

Fagiolini, M. and Hensch, T. K. (2000). Inhibitory threshold for critical-period activation in primary visual cortex. *Nature*, 404(6774):183–6.

Fanselow, E., Richardson, K. A., and Connors, B. W. (2008). Selective, state-dependent activation of somatostatin-expressing inhibitory interneurons in mouse neocortex. *Journal of Neurophysiology*, 100:2640–2652.

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

Field, D. J., Hayes, A., and Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local "association field". *Vision research*, 33(2):173–193.

Firth, S. I., Wang, C.-T., and Feller, M. B. (2005). Retinal waves: mechanisms and function in visual system development. *Cell Calcium*, 37(5):425–32.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current opinion in neurobiology*, 10(4):438–43.

Fregnac, Y. and Imbert, M. (1978). Early development of visual cortical cells in normal and dark-reared kittens: relationship between orientation selectivity and ocular dominance. *The Journal of physiology*, 278:27–44.

Freund, T. F. and Katona, I. (2007). Perisomatic inhibition. *Neuron*, 56(1):33–42.

Fu, Y., Tucciarone, J. M., Espinosa, J. S., Sheng, N., Darcy, D. P., Nicoll, R. a., Huang, Z. J., and Stryker, M. P. (2014). A cortical circuit for gain control by behavioral state. *Cell*, 156:1139–1152.

Gabernet, L., Jadhav, S. P., Feldman, D. E., Carandini, M., and Scanziani, M. (2005). Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. *Neuron*, 48(2):315–27.

Geisler, W. S., Perry, J. S., Super, B. J., and Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision research*, 41(6):711–24.

Gentet, L. J., Kremer, Y., Taniguchi, H., Huang, Z. J., Staiger, J. F., and Petersen, C. C. H. (2012). Unique functional properties of somatostatin-expressing GABAergic neurons in mouse barrel cortex. *Nature Neuroscience*, 15(4):607–12.

Gewaltig, M.-O. and Diesmann, M. (2007). Nest (neural simulation tool). *Scholarpedia*, 2(4):1430.

Gilbert, C. and Wiesel, T. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision research*, 30(Il).

Gilbert, C. C. and Wiesel, T. T. N. (1983). Clustered Intrinsic Connections in Cat Visual Cortex. *The Journal of Neuroscience*, 3(5):1116.

Gonchar, Y., Wang, Q., and Burkhalter, A. (2007). Multiple distinct subtypes of GABAergic neurons in mouse visual cortex identified by triple immunostaining. *Frontiers in neuroanatomy*, 1(March):3.

Graham, N. V. (2011). Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. *Vision Research*, 51(13):1397–1430.

Grinvald, A., Lieke, E. E., Frostig, R. D., and Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *The Journal of Neuroscience*, 14(5):2545–68.

Hawken, M. J., Shapley, R. M., and Grosof, D. H. (2009). Temporal-frequency selectivity in monkey visual cortex. *Visual Neuroscience*, 13(03):477.

Hennig, M. H., Grady, J., van Coppenhagen, J., and Sernagor, E. (2011). Age-dependent homeostatic plasticity of gabaergic signaling in developing retinal networks. *The Journal of Neuroscience*, 31(34):12159–12164.

Hensch, T. K. and Stryker, M. P. (2004). Columnar architecture sculpted by GABA circuits in developing cat visual cortex. *Science (New York, N.Y.)*, 303(5664):1678–81.

Higley, M. (2014). Localized GABAergic inhibition of dendritic Ca2+ signalling. *Nature Reviews Neuroscience*, 15(9).

Hines, M. (1994). The neuron simulation program. In *Neural network simulation environments*, pages 147–163. Springer.

Hirsch, J. a. and Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *The Journal of Neuroscience*, 11(6):1800–9.

Hirsch, J. A. and Martinez, L. M. (2006). Circuits that build visual cortical receptive fields. *Trends in neurosciences*, 29(1):30–39.

Hirsch, J. a., Martinez, L. M., Pillai, C., Alonso, J.-M., Wang, Q., and Sommer, F. T. (2003). Functionally distinct inhibitory neurons at the first stage of visual cortical processing. *Nature Neuroscience*, 6(12):1300–8.

Hofer, S. B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., Lein, E., Lesica, N. a., and Mrsic-Flogel, T. D. (2011). Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nature Neuroscience*, 14(8):1045–52.

Hogan, D., Terwilleger, E., and Berman, N. (1992). Development of subpopulations of GABAergic neurons in cat visual cortical areas. *NeuroReport*, 3:1069–1072.

Hubel, D. and Wiesel, T. N. (1961). Integrative Action in the Cat's Lateral Geniculate Body. *Journal of Physiology*, 155:385–398.

Huberman, A. D., Feller, M. B., and Chapman, B. (2008). Mechanisms underlying development of visual maps and receptive fields. *Annual Review of Neuroscience*, 31:479–509.

Hunt, J. J., Bosking, W. H., and Goodhill, G. J. (2011). Statistical structure of lateral connections in the primary visual cortex. *Neural systems & circuits*, 1(1):3.

Hupé, J.-M., James, A. C., Girard, P., Bullier, J., Ponce, C. R., Hunter, J. N., Pack, C. C., Lomber, S. G., Richard, T., Anderson, J. C., Martin, K. A. C., Allen, H. A., Humphreys, G. W., Colin, J., and Neumann, H. (2001). Response Modulations by Static Texture Surround in Area V1 of the Macaque Monkey Do Not Depend on Feedback Connections From V2. *Journal of Neurophysiology*, 85:146–163.

Hupé, J.-M., James, A. C. A., Payne, B. R. B., Lomber, S. S. G., Girard, P., and Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2 and V3 neurons. *Nature*, 394(6695):784–787.

Huxlin, K. and Pasternak, T. (2001). Long-term neurochemical changes after visual cortical lesions in the adult cat. *Journal of Comparative Neurology*, 429:221–241.

Ichida, J. M., Schwabe, L., Bressloff, P. C., and Angelucci, A. (2007). Response facilitation from the "suppressive" receptive field surround of macaque V1 neurons. *Journal of Neurophysiology*, 98(4):2168–81.

Isaacson, J. S. and Scanziani, M. (2011). How inhibition shapes cortical activity. *Neuron*, 72(2):231–243.

Jean-Luc R. Stevens, Philipp Rudiger, and James A. Bednar (2015). HoloViews: Building Complex Visualizations Easily for Reproducible Science. In Kathryn Huff and James Bergstra, editors, *Proceedings of the 14th Python in Science Conference*, pages 61 – 69.

Jin, J., Wang, Y., Swadlow, H. a., and Alonso, J. M. (2011). Population receptive fields of ON and OFF thalamic inputs to an orientation column in visual cortex. *Nature Neuroscience*, 14(2):232–238.

Jones, H. E., Wang, W., and Sillito, a. M. (2002). Spatial organization and magnitude of orientation contrast interactions in primate V1. *Journal of Neurophysiology*, 88(5):2796–808.

Jones, J. P. and Palmer, L. A. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of neurophysiology*, 58(6):1233–1258.

Kang, K., Shelley, M., and Sompolinsky, H. (2003). Mexican hats and pinwheels in visual cortex. *Proceedings of the National Academy of Sciences*, 100(5):2848–2853.

Kapadia, M. K., Ito, M., Gilbert, C. D., and Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron*, 15(4):843–56.

Kapadia, M. K., Westheimer, G., and Gilbert, C. D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proceedings of the National Academy of Sciences of the United States of America*, 96(21):12073–12078.

Kaschube, M., Schnabel, M., Löwel, S., Coppola, D., White, L., and Wolf, F. (2010). Universality in the evolution of orientation columns in the visual cortex. *Science*, 330(6007):1113–1116.

Kastner, S., Nothdurft, H.-C., and Pigarev, I. N. (1997). Neuronal correlates of pop-out in cat striate cortex. *Vision research*, 37(4):371–376.

Katz, L. C., Weliky, M., and Crowley, J. C. (2000). Activity and the Development of the Visual Cortex: New Perspectives. In *The New Cognitive Neurosciences*, chapter 13,

pages 199–221. MIT Press, Cambridge, MA.

Keemink, S. W. and van Rossum, M. C. (2015). A unified account of tilt illusions, association fields, and contour detection based on elastica. *Vision research*.

Keil, W., Kaschube, M., Schnabel, M., Kisvarday, Z. F., Löwel, S., Coppola, D. M., White, L. E., and Wolf, F. (2012). Response to comment on "universality in the evolution of orientation columns in the visual cortex ". *Science*, 336(6080):413–413.

Kepecs, A. and Fishell, G. (2014). Interneuron cell types are fit to function. *Nature*, 505(7483):318–26.

Khan, B. (2011). Modelling the pathway to auditory cortex. Master's thesis, The University of Edinburgh.

Kisvárday, Z., Ferecskó, A., and Kovács, K. (2002). One axon-multiple functions: specificity of lateral inhibitory connections by large basket cells. *Journal of Neurocytology*, 264(2002):255–264.

Kisvárday, Z. F., Beaulieu, C., and Eysel, U. T. (1993). Network of GABAergic large basket cells in cat visual cortex (area 18): implication for lateral disinhibition. *The Journal of comparative neurology*, 327(3):398–415.

Kisvárday, Z. F., Tóth, E., Rausch, M., and Eysel, U. T. (1997). Orientation-specific relationship between populations of excitatory and inhibitory lateral connections in the visual cortex of the cat. *Cerebral Cortex*, 7(7):605–18.

Klausberger, T., Roberts, J. D. B., and Somogyi, P. (2002). Cell type- and input-specific differences in the number and subtypes of synaptic GABA(A) receptors in the hippocampus. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 22(7):2513–21.

Kuhlman, S. J., Olivas, N. D., Tring, E., Ikrar, T., Xu, X., and Trachtenberg, J. T. (2013). A disinhibitory microcircuit initiates critical-period plasticity in the visual cortex. *Nature*, pages 1–7.

Le Roux, N., Cabezas, C., Böhm, U. L., and Poncer, J. C. (2013). Input-specific learning rules at excitatory synapses onto hippocampal parvalbumin-expressing interneurons. *The Journal of physiology*, 591(Pt 7):1809–22.

Lee, S., Kruglikov, I., Huang, Z. J., Fishell, G., and Rudy, B. (2013). A disinhibitory circuit mediates motor integration in the somatosensory cortex. *Nature neuroscience*, 16(11):1662–70.

Letzkus, J. J., Wolff, S. B. E., Meyer, E. M. M., Tovote, P., Courtin, J., Herry, C., and Lüthi, A. (2011). A disinhibitory microcircuit for associative fear learning in the auditory cortex. *Nature*, 480(7377):331–5.

Levitt, J. B. and Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature*, 387:73–76.

Levitt, J. B. and Lund, J. S. (2002). The spatial extent over which neurons in macaque striate cortex pool visual signals. *Visual Neuroscience*, 19(4):439–52.

Levitt, J. B., Schumer, R. a., Sherman, S. M., Spear, P. D., and Movshon, J. a. (2001). Visual response properties of neurons in the LGN of normally reared and visually deprived macaque monkeys. *Journal of Neurophysiology*, 85(5):2111–29.

Lewis, D. A., Curley, A. A., Glausier, J. R., and Volk, D. W. (2012). Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. *Trends in neuro-*

sciences, 35(1):57–67.

Li, L.-Y., Xiong, X. R., Ibrahim, L. a., Yuan, W., Tao, H. W., and Zhang, L. I. (2014). Differential Receptive Field Properties of Parvalbumin and Somatostatin Inhibitory Neurons in Mouse Auditory Cortex. *Cerebral Cortex*, 14:2–3.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences*, 6(1):9–16.

Löwel, S., Schmidt, K. E., Kim, D.-S., Wolf, F., Hoffsümmer, F., Singer, W., and Bonhoeffer, T. (1998). The layout of orientation and ocular dominance domains in area 17 of strabismic cats. *European Journal of Neuroscience*, 10(8):2629–2643.

Lund, J. and Yoshioka, T. (1991). Local circuit neurons of macaque monkey striate cortex: III. Neurons of laminae 4B, 4A, and 3B. *Journal of Comparative Neurology*, 311:234–258.

Lund, J. S. (1987). Local circuit neurons of macaque monkey striate cortex: IV. Neurons of laminae 4C and 5A. *The Journal of comparative neurology*, 257:60–92.

Ma, W.-p., Liu, B.-h., Li, Y.-t., Huang, Z. J., Zhang, L. I., and Tao, H. W. (2010). Visual representations by cortical somatostatin inhibitory neurons–selective but with weak and delayed responses. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(43):14371–14379.

Ma, W.-p., Liu, B.-h., Li, Y.-t., Huang, Z. J., Zhang, L. I., and Tao, H. W. (2011). Visual Representations by Cortical Somatostatin Inhibitory Neurons - Selective but with Weak and Delayed Responses. *Journal of Neuroscience*, 30(43):14371–14379.

Marder, E. and Goaillard, J.-M. (2006). Variability, compensation and homeostasis in neuron and network function. *Nature Reviews Neuroscience*, 7(7):563–574.

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., and Wu, C. (2004). Interneurons of the neocortical inhibitory system. *Nature Reviews Neuroscience*, 5(10):793–807.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* Freeman and Company, San Francisco.

Martin, K., Somogyi, P., and Whitteridge, D. (1983). Physiological and morphological properties of identified basket cells in the cat's visual cortex. *Experimental Brain Research*, 50:193–200.

McGuire, B. a., Gilbert, C. D., Rivlin, P. K., and Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *The Journal of comparative neurology*, 305(3):370–92.

McMains, S. a. and Somers, D. C. (2004). Multiple spotlights of attentional selection in human visual cortex. *Neuron*, 42(4):677–86.

Mehaffey, W. H., Doiron, B., Maler, L., and Turner, R. W. (2005). Deterministic multiplicative gain control with active dendrites. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(43):9968–77.

Miikkulainen, R., Bednar, J. A., Choe, Y., and Sirosh, J. (2005). Development of Maps and Connections. In *Computational Maps in the Visual Cortex*, chapter 5, pages 85–132. Springer, Berlin.

Miller, K., Keller, J., and Stryker, M. (1989). Ocular dominance column development: Analysis and simulation. *Science*, 245(4918):605–615.

Mizobe, K., Polat, U., Pettet, M. W., and Kasamatsu, T. (2001). Facilitation and suppression of single striate-cell activity by spatially discrete pattern stimuli presented beyond the receptive field. *Visual Neuroscience*, 18(3):377–91.

Nauhaus, I., Benucci, A., Carandini, M., and Ringach, D. L. (2008). Neuronal selectivity and local map structure in visual cortex. *Neuron*, 57(5):673–679.

Nelson, S., Toth, L., Sheth, B., and Sur, M. (1994). Orientation selectivity of cortical neurons during intracellular blockade of inhibition. *Science*, 265(5173):774–777.

Nienborg, H., Hasenstaub, a., Nauhaus, I., Taniguchi, H., Huang, Z. J., and Callaway, E. M. (2013). Contrast Dependence and Differential Contributions from Somatostatin- and Parvalbumin-Expressing Neurons to Spatial Integration in Mouse V1. *Journal of Neuroscience*, 33(27):11145–11154.

Nijholt, B. and Akhmerov, A. R. (2015). Orbital effect of magnetic field on the majorana phase diagram. *arXiv preprint arXiv:1509.02675*.

Nowak, L. G., Sanchez-Vives, M. V., and McCormick, D. a. (2008). Lack of orientation and direction selectivity in a subgroup of fast-spiking inhibitory interneurons: cellular and synaptic mechanisms and comparison with other electrophysiological cell types. *Cerebral Cortex*, 18(5):1058–78.

Obermayer, K., Ritter, H., and Schulten, K. (1990). A principle for the formation of the spatial structure of cortical feature maps. *Proceedings of the National Academy of Sciences of the United States of America*, 87(21):8345–9.

Ohki, K., Chung, S., Ch'ng, Y. H., Kara, P., and Reid, R. C. (2005). Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature*, 433(7026):597–603.

Olshausen, B. a. and Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images.

Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D., and Ferster, D. (2009a). Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron*, 62(4):578–92.

Ozeki, H., Finn, I. M., Schaffer, E. S., Miller, K. D., and Ferster, D. (2009b). Inhibitory stabilization of the cortical network underlies visual surround suppression. *Neuron*, 62(4):578–592.

Pérez, F. and Granger, B. E. (2007). IPython: a system for interactive scientific computing. *Computing in Science and Engineering*, 9(3):21–29.

Perrinet, L. U. and Bednar, J. A. (2015). Edge co-occurrences can account for rapid categorization of natural versus animal images. *Scientific reports*, 5.

Perry, V. H. and Cowey, A. (1985). The ganglion cell and cone distributions in the monkey's retina: Implications for central magnification factors. *Vision Research*, 25(12):1795–1810.

Pfeffer, C. K., Xue, M., He, M., Huang, Z. J., and Scanziani, M. (2013). Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. *Nature Neuroscience*, 16(8):1–12.

Poggio, T. (2012). The levels of understanding framework, revised. Technical report, MIT.

Polat, U., Mizobe, K., Pettet, M. W., Kasamatsu, T., and Norcia, A. M. (1998). Colinear stimuli regulate visual responses depeding on cell's constrast threshold. *Nature*, 391(February):580–584.

Poort, J., Raudies, F., Wannig, A., Lamme, V. a. F., Neumann, H., and Roelfsema, P. R. (2012). The role of attention in figure-ground segregation in areas v1 and v4 of the visual cortex. *Neuron*, 75(1):143–56.

Potjans, T. C. and Diesmann, M. (2014). The cell-type specific cortical microcircuit: relating structure and activity in a full-scale spiking network model. *Cerebral Cortex*, 24(3):785–806.

Ramsden, B. M., Hung, C. P., and Roe, a. W. (2001). Real and illusory contour processing in area V1 of the primate: a cortical balancing act. *Cerebral Cortex*, 11(7):648–65.

Rao, R. P. and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1):79–87.

Reynaud, A., Masson, G. S., and Chavane, F. (2012). Dynamics of local input normalization result from balanced short- and long-range intracortical interactions in area V1. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 32(36):12558–69.

Ringach, D. L. (2002). Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of neurophysiology*, 88:455–463.

Ringach, D. L. (2007). On the origin of the functional architecture of the cortex. *PLoS One*, 2(2):e251.

Roberts, M. J. and Thiele, A. (2008a). Attention and contrast differently affect contextual integration in an orientation discrimination task. *Experimental Brain Research*, 187(4):535–49.

Roberts, M. J. and Thiele, A. (2008b). Spatial integration and its moderation by attention and acetylcholine. *Frontiers in Bioscience*, 13:3742–3759.

Roberts, M. J., Zinke, W., Guo, K., Robertson, R., McDonald, J. S., and Thiele, A. (2005). Acetylcholine dynamically controls spatial integration in marmoset primary visual cortex. *Journal of Neurophysiology*, 93(4):2062–72.

Rochester, N., Holland, J. H., Haibt, L., and Duda, W. (1956). Tests on a cell assembly theory of the action of the brain, using a large digital computer. *Information Theory, IRE Transactions on*, 2(3):80–93.

Rodieck, R. W. (1965). Quantitative Analysis of Cat Retinal Ganglion Cell Response to Visual Stimuli. *Vision Research*, 5:583–601.

Rodieck, R. W. and Stone, J. (1965a). Analysis of Receptive Fields of Cat Retinal Ganglion Cells. *Journal of Physiology*, 28(5):833–849.

Rodieck, R. W. and Stone, J. (1965b). Response of Cat Retinal Ganglion Cells to Moving Visual Patterns. *Journal of Neurophysiology*, 28:819–832.

Rubin, D. B., Van Hooser, S. D., and Miller, K. D. (2015). The stabilized supralinear network: A unifying circuit motif underlying multi-input integration in sensory cortex. *Neuron*, 85(2):402–417.

Ruderman, D. L. (1997). Origins of scaling in natural images. *Vision research*, 37(23):3385–98.

Rudolph, U., Crestani, F., Benke, D., Brünig, I., Benson, J. a., Fritschy, J. M., Martin, J. R., Bluethmann, H., and Möhler, H. (1999). Benzodiazepine actions mediated by specific gamma-aminobutyric acid(A) receptor subtypes. *Nature*, 401(6755):796–800.

Rudy, B., Fishell, G., Lee, S., and Hjerling-Leffler, J. (2011). Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Developmental Neurobiology*, 71:45–61.

Runyan, C. a., Schummers, J., Van Wart, A., Kuhlman, S. J., Wilson, N. R., Huang, Z. J., and Sur, M. (2010). Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron*, 67(5):847–57.

Ruthazer, E. S. and Stryker, M. P. (1996). The role of activity in the development of long-range horizontal connections in area 17 of the ferret. *The Journal of Neuroscience*, 16(22):7253–69.

Sceniak, M. and Hawken, M. (2001). Visual Spatial Characterization of Macaque V1 Neurons. *Journal of Neurophysiology*, 85:1873–1887.

Sceniak, M., Ringach, D., and Hawken, M. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nature*, 2(8):733–739.

Sceniak, M. P., Chatterjee, S., and Callaway, E. M. (2006). Visual spatial summation in macaque geniculocortical afferents. *Journal of Neurophysiology*, 96(6):3474–84.

Schmidt, K. E., Goebel, R., Löwel, S., and Singer, W. (1997). The Perceptual Grouping Criterion of Collinearity Is Reflected by Anisotropies of Connections in the Primary Visual Cortex. *European Journal of Neuroscience*, 9(December 1996):1083–1089.

Schwabe, L., Obermayer, K., Angelucci, A., and Bressloff, P. C. (2006). The role of feedback in shaping the extra-classical receptive field of cortical neurons: a recurrent network model. *The Journal of Neuroscience*, 26(36):9117–29.

Schwartz, O. and Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature neuroscience*, 4:819–825.

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

Serre, T., Kreiman, G., Kouh, M., Cadieu, C., Knoblich, U., and Poggio, T. (2007a). A quantitative theory of immediate visual recognition. *Progress in brain research*, 165:33–56.

Serre, T., Oliva, A., and Poggio, T. (2007b). A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, 104(15):6424–6429.

Shatz, B. C. J., Strykert, M. P., Shatz, C., and Stryker, M. (1978). Ocular dominance in layer IV of the cat's visual cortex and the effects of monocular deprivation. *Journal of Physiology*, 281(1):267–283.

Shatz, C. J. (1996). Emergence of order in visual system development. *Proceedings of the National Academy of Sciences*, 93(2):602–608.

Shelley, M., McLaughlin, D., Shapley, R., and Wielaard, J. (2002). States of high conductance in a large-scale model of the visual cortex. *Journal of computational neuroscience*, 13(2):93–109.

Sherman, S. M. and Guillery, R. W. (2002). The role of the thalamus in the flow of information to the cortex. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 357(1428):1695–708.

Shushruth, S., Ichida, J. M., Levitt, J. B., and Angelucci, A. (2009). Comparison of spatial summation properties of neurons in macaque V1 and V2. *Journal of Neurophysiology*, 102(4):2069–83.

Sigman, M., Cecchi, G. a., Gilbert, C. D., and Magnasco, M. O. (2001). On a common circle: natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences of the United States of America*, 98(4):1935–1940.

Silberberg, G. and Markram, H. (2007). Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. *Neuron*, 53(5):735–46.

Sillito, A. (1979). Inhibitory mechanisms influencing complex cell orientation selectivity and their modification at high resting discharge levels. *The Journal of physiology*, 289:33–53.

Sillito, A. M., Cudeiro, J., and Jones, H. E. (2006). Always returning: feedback and sensory processing in visual cortex and thalamus. *Trends in Neurosciences*, 29(6):307–16.

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., and Davls, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities.

Simoncelli, E. and Olshausen, B. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24:1193–216.

Sincich, L. C. and Blasdel, G. G. (2001). Oriented axon projections in primary visual cortex of the monkey. *The Journal of Neuroscience*, 21(12):4416–26.

Skottun, B. C., De Valois, R. L., Grosof, D. H., Movshon, J. A., Albrecht, D. G., and Bonds, A. (1991). Classifying simple and complex cells on the basis of response modulation. *Vision research*, 31(7):1078–1086.

Solomon, S. G., Lee, B. B., and Sun, H. (2006). Suppressive surrounds and contrast gain in magnocellular-pathway retinal ganglion cells of macaque. *The Journal of Neuroscience*, 26(34):8715–8726.

Solomon, S. G. and Lennie, P. (2007). The machinery of colour vision. *Nature Reviews Neuroscience*, 8(4):276–86.

Solomon, S. G., White, A. J. R., and Martin, P. R. (2002). Extraclassical receptive field properties of parvocellular, magnocellular, and koniocellular cells in the primate lateral geniculate nucleus. *The Journal of Neuroscience*, 22(1):338–49.

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

Somers, D. C., Todorov, E. V., Siapas, a. G., Toth, L. J., Kim, D. S., and Sur, M. (1998). A local circuit approach to understanding integration of long-range inputs in primary visual cortex. *Cerebral Cortex*, 8(3):204–17.

Somogyi, P., Kisvárday, Z. F., Martin, K. a., and Whitteridge, D. (1983). Synaptic connections of morphologically identified and physiologically characterized large basket cells in the striate cortex of cat. *Neuroscience*, 10(2):261–94.

Spear, P. D., Moore, R. J., Kim, C. B., Xue, J. T., and Tumosa, N. (1994). Effects of aging on the primate visual system: spatial and temporal processing by lateral geniculate neurons in young adult and old rhesus monkeys. *Journal of Neurophysiology*, 72(1):402–20.

Spratling, M. (2011). A single functional model accounts for the distinct properties of suppression in cortical area v1. *Vision research*, 51(6):563–576.

Spratling, M. W. (2010). Predictive coding as a model of response properties in cortical area V1. *The Journal of Neuroscience*, 30(9):3531–43.

Srinivasan, M., Laughlin, S., and Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London*, 216(1205):427–459.

Stemmler, M., Usher, M., and Niebur, E. (1995). Lateral Interactions in Primary Visual Cortex: A Model Bridging Physiology and Psychophysics. *Science*, 269(5232):1877–1880.

Stettler, D. D., Das, A., Bennett, J., and Gilbert, C. D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron*, 36(4):739–50.

Stevens, J. L. (2011). A temporal model of neural activity and VSD response in the primary visual cortex. Master's thesis, The University of Edinburgh.

Stevens, J. L. (2016). *Modeling the spatiotemporal properties of evoked neural response in the mammalian primary visual cortex*. PhD thesis, The University of Edinburgh.

Stevens, J.-L., Elver, M., and Bednar, J. A. (2013a). An automated and reproducible workflow for running and analysing neural simulations using Lancet and IPython Notebook. *Frontiers in Neuroinformatics*, 7:44.

Stevens, J.-L. R., Law, J. S., Antolík, J., and Bednar, J. a. (2013b). Mechanisms for stable, robust, and adaptive development of orientation maps in the primary visual cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 33(40):15747–66.

Sullivan, T. J. and De Sa, V. R. (2006). Homeostatic synaptic scaling in self-organizing maps. *Neural Networks*, 19(6):734–743.

Sur, M., Pallas, S. L., and Roe, A. W. (1990). Cross-modal plasticity in cortical development: Differentiation and specification of sensory neocortex. *Trends in Neurosciences*, 13(6):227–233.

Swadlow, H. a. (2003). Fast-spike interneurons and feedforward inhibition in awake sensory neocortex. *Cerebral Cortex*, 13(1):25–32.

Tan, Z., Hu, H., Huang, Z. J., and Agmon, a. (2008). Robust but delayed thalamocortical activation of dendritic-targeting inhibitory interneurons. *Proceedings of the National Academy of Sciences*, 105(6):2187–2192.

Tanaka, S., Ribot, J., Imamura, K., and Tani, T. (2006). Orientation-restricted continuous visual exposure induces marked reorganization of orientation maps in early life. *NeuroImage*, 30(2):462–77.

Tenner, V. T., de Dood, M. J., and van Exter, M. P. (2016). Measurement of the phase and intensity profile of surface plasmon laser emission. *ACS Photonics*.

Thiele, A. (2013). Muscarinic signaling in the brain. *Annual Review of Neuroscience*, 36:271–94.

Thomson, A. and Deuchars, J. (1994). Temporal and spatial properties of local circuits in neocortex. *Trends in neurosciences*, 2:119–126.

Thomson, A., West, D., and Deuchars, J. (1995). Properties of single axon excitatory postsynaptic potentials elicited in spiny interneurons by action potentials in pyramidal neurons in slices of rat neocortex. *Neuroscience*, 69(3):727–738.

Treue, S. (2003). Visual attention: the where, what, how and why of saliency. *Current Opinion in Neurobiology*, 13(4):428–432.

Troyer, T. W., Krukowski, A. E., Priebe, N. J., and Miller, K. D. (1998). Contrast-Invariant Orientation Tuning in Cat Visual Cortex: Thalamocortical Input Tuning and Correlation-Based Intracortical Connectivity. *Journal of Neuroscience*, 18(15):5908–5927.

Tsodyks, M. V. and Markram, H. (1997). The neural code between neocortical pyramidal neurons depends. *Proceedings of the National Academy of Sciences*, 94:719–723.

Turrigiano, G. G. (1999). Homeostatic plasticity in neuronal networks: the more things change, the more they stay the same. *Trends in neurosciences*, 22(5):221–227.

Turrigiano, G. G. and Nelson, S. B. (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, 5(2):97–107.

Van Brederode, J. F., Mulligan, K. a., and Hendrickson, a. E. (1990). Calcium-binding proteins as markers for subpopulations of GABAergic neurons in monkey striate cortex. *The Journal of comparative neurology*, 298(1):1–22.

Van Essen, D. C., Newsome, W. T., and Maunsell, J. H. (1984). The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Research*, 24(5):429–48.

Van Hooser, S. D., Roy, A., Rhodes, H. J., Culp, J. H., and Fitzpatrick, D. (2013). Transformation of receptive field properties from lateral geniculate nucleus to superficial v1 in the tree shrew. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 33(28):11494–505.

Vinje, W. E. and Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287(5456):1273–1276.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14(2):85–100.

Von Melchner, L., Pallas, S. L., and Sur, M. (2000). Visual behaviour mediated by retinal projections directed to the auditory pathway. *Nature*, 404(6780):871–876.

Wang, C. and Fitzgibbon, T. (2009). Contrast dependence of center and surround integration in primary visual cortex of the cat. *Journal of Vision*, 9:1–15.

Webb, B. S., Dhruv, N. T., Solomon, S. G., Tailby, C., and Lennie, P. (2005). Early and late mechanisms of surround suppression in striate cortex of macaque. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(50):11666–75.

Weliky, M., Kandler, K., Fitzpatrick, D., and Katz, L. C. (1995). Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to orientation columns. *Neuron*, 15(3):541–52.

White, L. E. and Fitzpatrick, D. (2007). Vision and cortical map development. *Neuron*, 56(2):327–338.

Wickersham, I. R., Lyon, D. C., Barnard, R. J. O., Mori, T., Conzelmann, K.-k., Young, J. a. T., Callaway, E. M., Manuscript, A., and Neurons, G. T. (2007). Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons. *Neuron*, 53(5):639–647.

Wilson, N. R., Runyan, C. a., Wang, F. L., and Sur, M. (2012). Division and subtraction by distinct cortical inhibitory networks in vivo. *Nature*, 488(7411):343–8.

Wilson, S. P., Law, J. S., Mitchinson, B., Prescott, T. J., and Bednar, J. A. (2010). Modeling the emergence of whisker direction maps in rat barrel cortex. *PloS one*, 5(1):e8778.

Wöhr, M., Orduz, D., Gregory, P., Moreno, H., Khan, U., Vörckel, K., Wolfer, D., Welzl, H., Gall, D., Schiffmann, S. N., et al. (2015). Lack of parvalbumin in mice leads to behavioral deficits relevant to all human autism core symptoms and related neural morphofunctional abnormalities. *Translational psychiatry*, 5(3):e525.

Xing, D., Shapley, R. M., Hawken, M. J., and Ringach, D. L. (2005). Effect of Stimulus Size on the Dynamics of Orientation Selectivity in Macaque V1. *Journal of Neurophysiology*, 94:799–812.

Xu, H., Jeong, H.-Y., Tremblay, R., and Rudy, B. (2013). Neocortical somatostatin-expressing GABAergic interneurons disinhibit the thalamorecipient layer 4. *Neuron*, 77(1):155–67.

Xu, X. and Callaway, E. M. (2009). Laminar specificity of functional input to distinct types of inhibitory cortical neurons. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 29(1):70–85.

Xu, X., Roby, K. D., and Callaway, E. M. (2010). Immunochemical characterization of inhibitory mouse cortical neurons: three chemically distinct classes of inhibitory cells. *The Journal of comparative neurology*, 518(3):389–404.

Yousef, T., Tóth, É., Rausch, M., Eysel, U. T., and Kisvárday, Z. F. (2001). Topography of orientation centre connections in the primary visual cortex of the cat. *Neuroreport*, 12(8):1693–1699.