ABSTRACT

Title of Document: HIERARCHICAL NEURAL COMPUTATION IN THE MAMMALIAN VISUAL SYSTEM

Yuwei Cui,
Doctor of Philosophy, 2015

Directed By: Prof. Daniel A Butts
Department of Biology

Our visual system can efficiently extract behaviorally relevant information from ambiguous and noisy luminance patterns. Although we know much about the anatomy and physiology of the visual system, it remains obscure how the computation performed by individual visual neurons is constructed from the neural circuits. In this dissertation, I designed novel statistical modeling approaches to study hierarchical neural computation, using electrophysiological recordings from several stages of the mammalian visual system.

In Chapter 2, I describe a two-stage nonlinear model that characterized both synaptic current and spike response of retinal ganglion cells with unprecedented accuracy. I found that excitatory synaptic currents to ganglion cells are well described by excitatory inputs multiplied by divisive suppression, and that spike responses can be explained with the addition of a second stage of spiking nonlinearity and refractoriness. The structure of the model was inspired by known elements of the

retinal circuit, and implies that presynaptic inhibition from amacrine cells is an important mechanism underlying ganglion cell computation.

In Chapter 3, I describe a hierarchical stimulus-processing model of MT neurons in the context of a naturalistic optic flow stimulus. The model incorporates relevant nonlinear properties of upstream V1 processing and explained MT neuron responses to complex motion stimuli. MT neuron responses are shown to be best predicted from distinct excitatory and suppressive components. The direction-selective suppression can impart selectivity of MT neurons to complex velocity fields, and contribute to improved estimation of the three-dimensional velocity of moving objects.

In Chapter 4, I present an extended model of MT neurons that includes both the stimulus-processing component and network activity reflected in local field potentials (LFPs). A significant fraction of the trial-to-trial variability of MT neuron responses is predictable from the LFPs in both passive fixation and a motion discrimination task. Moreover, the choice-related variability of MT neuron responses can be explained by their phase preferences in low-frequency band LFPs. These results suggest an important role of network activity in cortical function.

Together, these results demonstrated that it is possible to infer the nature of neural computation from physiological recordings using statistical modeling approaches.

HIERARCHICAL NEURAL COMPUTATION IN THE MAMMALIAN VISUAL SYSTEM

By

Yuwei Cui

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Prof. Daniel A. Butts, Chair
Prof. David W. Jacobs
Prof. Patrick O. Kanold
Prof. Joshua H. Singer
Prof. Jonathan Z. Simon

Preface

The work presented in this dissertation reflects collaborative efforts between my advisor Daniel Butts, numerous experimental collaborators and myself. I wrote he introduction and conclusion chapters. The three main chapters (Chapter 2-4) were adapted from manuscripts that are published or under preparation for publication. In all cases, DAB helped design the experiment, suggested analyses, contributed to the interpretation of the results and supervised the production of the manuscript.

Chapter 2 arose from collaboration with Jonathan Demb's lab at Yale University. JDB and WYV designed and conducted the experiments. DAB and I developed the model, analyzed the data and wrote the manuscript. The manuscript is under preparation for publication. Part of the work has been presented at peer-reviewed conferences (Cui et al., 2013c).

Chapter 3 and 4 arose from collaboration with Chris Pack's lab at McGill University. DAB, CCP, LDL and myself designed the experiments. LDL performed the experiments. I analyzed the data, developed the model, and wrote the manuscript. Chapter 3 has been published (Cui et al., 2013b). Chapter 4 has been presented at conferences (Cui et al., 2014; Cui et al., 2013a, 2015) and is under preparation for publication.

Acknowledgements

The work presented in this dissertation results from tremendous support from colleagues, collaborators, friends and family. First and foremost, I am grateful to my advisor Daniel Butts. I have benefited immensely from his guidance through countless research projects during my graduate study. I am grateful to Jeff Hawkins, whose immense knowledge and insightful thoughts on neuroscience problems have been a constant source of motivation for me. I also would like to thank Subutai and Donna for giving me the flexibility to work on this dissertation while doing a research internship at Numenta.

The work presented here would have been impossible without the help of my collaborators: Chris Pack's lab at McGill University, Jonathan Demb's lab at Yale University, and many others. I would especially thank Dave Liu and Farhan Khawaja, who tirelessly did neurophysiological recordings that support many of the research projects, and Yanbin Wang, who generously shared valuable dataset of Retina recordings with me. I would also like to thank my current and past lab mates who helped me along the way: James McFarland, Oliver Rourke, Ninad Kothari, and Timm Lochmann.

I would like to thank all my friends who have supported and encouraged me in innumerable ways: Kevin, Franklin, Keith, Haosu, Pengfei, Ruiliang. I am also grateful to have the love and support of my parents, Jiahong and Haihua, and my girlfriend Asia, who is always there for me.

Finally, this dissertation study was made possible by a collaborative research in computational neuroscience (CRCNS) grant from the National Science Foundation

(IIS-0904430), by the Graduate School Summer Research Fellowships from the University of Maryland, and by the William Hodos dissertation assistantship.

Table of Contents

List of Figures

Chapter 1 Introduction

"*What does it mean to see? The plain man's answer would be, to know what is where by looking. In other words, vision is the process of discovering from images what is present in the world, and where it is.*"

-- David Marr

The problem of vision

Human beings are predominately visual creatures. Vision makes it possible for us to move though the world, catch or avoid moving objects, and base immediate decisions on a detailed understanding of the world around us. In our nervous system, there are more fibers dedicated to transmitting the output from the eyes to the brain than from any other peripheral sense organs. Roughly one third of human cortex and half of the cortex in macaques is related to processing of visual information, consisting of more than thirty visual and visual-association areas (Felleman and Van Essen, 1991).

There is a reason for the large amount of brain machinery dedicated to vision (Kaas, 1989). In natural environment, visual systems faced the daunting task of extracting relevant information from the ambiguous, noisy, and constantly changing patterns of light falling through a tiny aperture at the back of our eye. The physicist Richard Feynman once illustrated the problem of vision with an extraordinary analogy of a water bug floating on the surface of a swimming pool, which is confronted with the

task of reconstructing the position and motion of all the people swimming in the pool using only information comes from the movements of its body caused by the water waves (McIlwain, 1996). Although the water bug analogy sounds a bit of an exaggeration because we humans are so good at vision, scientists started to appreciate the difficulty of vision as they tried to develop computer vision system in the 1960-70s. The common feeling of despair shared amongst the early investigators of computer vision was that "practically anything could happen in an image and further more that practically everything did." (Marr, 1982).

Despite the difficulty of natural vision, biological visual systems have been highly successful through evolution. As we look at the world around us, we have immediate access to the composition of the visual scene into objects, as well as our relationship in space to those objects. The entire processing must take place in a fraction of second to be behaviorally relevant. Biological vision thus represents an amazing feat of natural intelligence. Understanding the computational principles underlying biological vision will not only guide development of machine vision systems, but also shed light onto the nature of intelligence.

Vision as an information-processing task: computation vs. representation

From a theoretical perspective, vision can be regarded as an information-processing task. There are two aspects of this task: (1) extracting behaviorally relevant information through a cascade of neuronal processing (i.e., computation); and (2)

forming internal representations of visual information, thus making it readily available for decisions about our actions (representation). Our brain processes information much differently than computers do. A modern computer can easily store the hue and luminance of every pixel of thousands of images at high resolution, but even with the best available computer vision algorithm, it cannot segment an arbitrary natural image into its underlying components. In other words, the ability to capture and accurately store visual scenes does not mean understanding content of the image.

Many studies in visual neuroscience equate visual computation with the ability to represent visual scenes through spikes of neurons in the visual cortex. For example, a number of studies showed that visual scenes could be reconstructed from responses of visual neurons (Bair et al., 1997; Kelly and Lee, 2004; Stanley et al., 1999) or brain activities by fMRI (Miyawaki et al., 2008; Nishimoto et al., 2011). While these studies are no doubt interesting from an engineering perspective (e.g., for development of brain-reading devices), they do not reveal the underlying computational principles in the visual system. Moreover, several theoretical frameworks stated the goal of visual information processing as maximizing information transmission (Kastner et al., 2015) or minimizing stimulus reconstruction error (Boerlin et al., 2013). Although these theoretical frameworks provide insightful predictions and explanations of neuron response properties of early visual system, it is hard to extend them to higher visual areas. In fact, except for processing in the early stages of the visual pathway, the goals of visual computation is likely quite different from image representation. To illustrate this difference, consider a thought

experiment where subjects are asked to judge whether a dot is present on an empty screen with a "yes" or "no" answer. The answer to this question relies on information extracted from the image, but may not be useful for reconstructing the image: placing a dot in the wrong location is worse than placing not dot at all. We can of course try to memorize the exact size, color and location of the dot if it is necessary for the completion of the task. But in real life, such detailed information is often not important. Instead, visual computation extracts very particular information from an image that is crucial for segmenting natural images into objects or for the completion of certain behavior task, while ignoring information that is not relevant to the answer. Visual neuroscientist Edward Adelson describing this distinct aspect of biological vision as: “Our world contains both things and stuff, but things tend to get the attention” (Adelson, 2001).

For decades, visual neuroscientists have tried to understand vision in physiological terms by tracing visual signals throughout the visual system: from the retina, to LGN, to primary visual cortex, and to many other downstream visual areas. Neural responses in early visual systems can be reasonably well characterized with phenomenological models. However, it is extremely challenging to develop functional models for later stages of visual processing. One reason for this difficulty is that throughout most, if not all, of the visual system, it is not clear what the goal of computation is. In a recent interview, Nobel laureates David Hubel and Torsten Wiesel identified it as “the next big question in the field” (Hubel and Wiesel, 2012): “We have almost no examples of neural structures in which we know the difference

between the information coming in and what is going out--what the structure is for. We have some idea of the answer for the retina, the lateral geniculate body, and the primary visual cortex, but that's about it. It is one thing to know that Broca's area has to do with language, but that is far from having any idea of the transformations of information taking place there." In fact, even for early visual systems, detailed analysis often reveal that the standard phenomenological models miss important elements of the response properties of sensory neurons (Butts et al., 2011; Carandini et al., 2005; Olshausen and Field, 2005), suggesting inadequate understanding of the underlying neural computation.

In this dissertation, I tried to tackle these difficult problems with a goal of understanding the nature of computation along the visual hierarchy. By using a combination of the state-of-the-art statistical modeling techniques and electrophysiological recordings from various stages of the mammalian visual system, I tried to explain the detailed response properties of visual neurons and understand how they are constructed from elements of the neural circuits. In Chapter 2 and 3 of the dissertation, I developed nonlinear functional models that provide detailed characterizations of how stimulus selectivity of visual neurons in the retina and in area MT is constructed from the neural circuits. These functional models addressed Hubel's concern and described "the difference between the information coming in and what is going out". By extending these models to a population level, important insights may be gained regarding what information is being represented in a visual area, or "what the structure is for".

Modeling stimulus processing of visual neurons

Neural encoding and receptive field models

A fundamental problem in sensory neuroscience is to characterize the relationship between sensory stimulus and the electrical activity of sensory neurons. This is known as the problem of neural encoding (Dayan and Abbott, 2005) (Fig. 1-1). Understanding the neural encoding process represents an important first step to revealing the underlying neural computation in the visual system. Over the past few decades, neuroscientists and statisticians have devoted much effort to the development of accurate neural encoding models. Lying at the heart of many of the encoding models is the concept of receptive field (RF). A broad definition of receptive field is "a portion of sensory space that can elicit neuronal responses when stimulated" (Alonso and Yao, 2009; Sherrington, 1906). In general, this sensory space can only be accurately defined with multiple dimensions, such as space, time, and luminance patterns, etc. It would require an improbable large amount of experimental data to fully capture the receptive field of single visual neurons. Instead, to simplify the problem, many investigators assumed a neuron's response only depend on a low-dimensional projection of the high-dimensional stimulus. The templates for this low-dimensional projection are known as linear receptive field. In this section, I review previous studies on neural encoding and receptive field models.

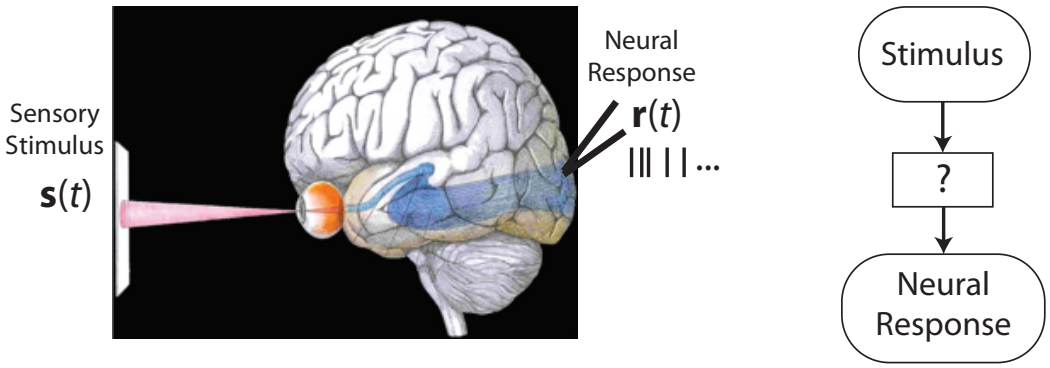

Figure 1-1 The problem of neural encoding.
In typical studies of the neural encoding, experimenters present sensory stimulus **s**(*t*) to the subject while recording visual neuron response **r**(*t*) from a neuron in the brain. The problem of neural encoding is to establish the relationship between stimulus and neural response.

Linear receptive field model

The simplest possible neural encoding model is a linear one. Mathematically, the stimulus input **s**, was assumed to be linearly related to the observed response r via projection onto a linear receptive field **k**,

$$r = \mathbf{k} \cdot \mathbf{s} \tag{1.1}$$

The linear receptive field models have been widely used through out multiple sensory systems (DiCarlo et al., 1998; Eggermont et al., 1983; Ramirez et al., 2014; Wilson, 2001; Wu et al., 2006). In the visual system, such models are known as spatiotemporal receptive field. Specifically, the linear receptive field $\mathbf{k}(\mathbf{x}, \tau)$ is a function of space **x** and time lag τ. The linear prediction of neural response at time t given a stimulus is given by:

$$g(t) = \sum_{\mathbf{x}} \sum_{\tau} k(\mathbf{x}, \tau) s(\mathbf{x}, t - \tau) \tag{1.2}$$

The output of the linear model can be mapped to a firing rate prediction through a static nonlinearity F[] to account for nonlinear properties of spike generation, such as spike threshold and response saturation. The predicted response is then given as,

$$r(t) = \mathrm{F}[g(t)] \tag{1.3}$$

This structure yields the canonical LN (Linear-Nonlinear) model (Chichilnisky, 2001; Hunter and Korenberg, 1986; Paninski, 2004). The linear receptive field can be efficiently estimated with reverse correlation techniques (Ringach and Shapley, 2004), such as spike-triggered averaging (Chichilnisky, 2001; Pillow and Simoncelli, 2006).

Linear RF models have been successfully applied to various stages in the visual system, from the retina (Chichilnisky, 2001; Gollisch and Meister, 2010), to the LGN (Cai et al., 1997; Reid et al., 1997), and to simple cells in V1 (DeAngelis et al., 1993; Ringach et al., 2002a). Despite the multiple levels of neural processing, the early visual system seems surprisingly linear at first glance, presumably due to specialized ribbon synapse in the retina (Shapley, 2009) and interplay between excitation and inhibitory in the primary visual cortex (Isaacson and Scanziani, 2011; Okun and Lampl, 2009; Priebe and Ferster, 2005; Wehr and Zador, 2003). However, detailed analysis revealed that even in the early retinal pathway, several important aspects of neural response properties were not captured by such linear models, most notably contrast adaptation (Kim and Rieke, 2001; Meister and Berry, 1999; Ozuysal and Baccus, 2012; Shapley and Victor, 1978) and the high temporal precision (Passaglia

and Troy, 2004; Uzzell and Chichilnisky, 2004). The inadequacy of linear models is because that stimulus selectivity of single neuron is constructed from distinct excitatory and suppressive elements. The linear model can not describe neurons that are selective to more than one visual feature in principle: a linear model that is "selective" to two features $\mathbf{k}_1$ and $\mathbf{k}_2$ is mathematically the same as being selective to their average:

$$r = \mathbf{k}_1 \cdot \mathbf{s} + \mathbf{k}_2 \cdot \mathbf{s} = (\mathbf{k}_1 + \mathbf{k}_2) \cdot \mathbf{s} \tag{1.4}$$

As a result, the linear model would not identify either feature properly.

Nonlinear neural encoding models

Due to the fundamental limitation of linear models, much effort has been devoted to the development of nonlinear encoding models. Unfortunately, the space of possible nonlinear models is unbounded. On the one hand, applying advanced machine learning techniques (e.g., multi-layer artificial neural network models) to neural data could yield good predictions with sufficient data, but the resulting model fits can be very hard to interpret and map to the elements of neural circuits, therefore not providing insight into how computation is constructed from the neural circuit. On the other hand, one might be inclined to include much detailed anatomical and circuitry information into the model. However, more complicated models generally require more data for parameter estimation, and often involve poorly behaved or intractable optimization problems. As a result, there is a tradeoff between model complexity and biological plausibility when developing such nonlinear encoding models.

To address the issue of model complexity, one common approach is to assume that neurons are selective to a small number of stimulus features. One can then identify the low dimensional "feature subspace" and predict neural responses with a nonlinear function applied only to these stimulus features (Fig. 1-2 A). In this framework, the predicted neural response is,

$$r = \mathrm{F}[\mathbf{s}^{\mathrm{T}}\mathbf{k}_1, \mathbf{s}^{\mathrm{T}}\mathbf{k}_2, \ldots.], \tag{1.5}$$

where ($\mathbf{k}_1$, $\mathbf{k}_2$, …) collectively define the feature subspace. Examples of this approach include dimensionality reduction approach such as spike-triggered covariance (STC) analysis (de Ruyter van Steveninck and Bialek, 1988; Schwartz et al., 2006) and information-theory-based approaches such as maximally informative dimensions (MID) analysis (Sharpee et al., 2004).

Another approach is to impose assumptions on the form of nonlinearities in the system. Examples include application of neural network models (Lau et al., 2002; Prenger et al., 2004) and using second-order and higher-order approximation of the nonlinear stimulus-response relationship, such as the Wiener-Volterra expansion (Deweese, 1996; Eggermont, 1993; Marmarelis and Marmarelis, 1978):

$$r = \sum_i k_i^{(1)} s_i + \sum_{i,j} k_{ij}^{(2)} s_i s_j + \sum_{i,j,k} k_{ijk}^{(3)} s_i s_j s_k + \cdots \tag{1.6}$$

Theoretically, the Wiener-Volterra expansion can capture arbitrarily nonlinear computations with the addition of higher order kernels. However, the number of parameters in this type of nonlinear model increases exponentially with the degree of the kernel, thus suffering from "the curse of dimensionality". In practice, the degree of the kernel considered is generally limited to be one or two (Wu et al., 2006). A

common problem of this approach and the subspace approach mentioned earlier is the difficulty to link resulting model components to physiological terms that present in the neural circuit.

Finally, some investigators have directly incorporated relevant nonlinearities into the linear model to capture specific response properties, such as using spike-history terms to explain refractoriness and precision (Berry and Meister, 1998; Keat et al., 2001; Paninski, 2004; Truccolo et al., 2005), using feedback terms to explain contrast adaptation (Mante et al., 2008; Meister and Berry, 1999; Shapley and Victor, 1978), and other type of nonlinearities to capture response properties such as sensitivity to local context (Ahrens et al., 2008a) and stimulus intensity (Carandini and Heeger, 2012; Heeger, 1992). While such approaches are often successful at describing the particular nonlinear effects targeted, they necessarily lump all nonlinear sources into the single nonlinear description, and usually can only explain the specific responses properties targeted.

Hierarchical models of sensory processing and the nonlinear input model

It is well known that visual system has a hierarchical organization (Felleman and Van Essen, 1991). Throughout the visual hierarchy, neurons in higher cortical areas receive inputs from neurons in lower areas. The receptive field size, as well as the complexity of the receptive field structure increases at successive stages of the visual pathway (Alonso and Yao, 2009). As a result, neural response in earlier stages of the

visual system is generally better understood and may be used to infer functions of later stages.

Hierarchical models has a long history since Hubel and Wiesel's seminal studies of simple and complex cells in the primary visual cortex (Hubel and Wiesel, 1962). In Hubel and Wiesel's model, RF of a simple cell is constructed by pooling over LGN inputs with center-surround RFs along a preferred axis of orientation. At the next stage, complex cells pool over afferent simple cells with the same preferred orientation to build translation-invariance of the preferred stimuli. Since then, many studies use hierarchical models to explain the organization and physiological properties of the visual system (Poggio and Serre, 2013), such as object and face recognition in the ventral visual pathway (Axelrod and Yovel, 2012; Riesenhuber and Poggio, 1999; Serre et al., 2007) and motion processing in the dorsal visual pathway (Mineault et al., 2012; Rust et al., 2006; Simoncelli and Heeger, 1998). One of the key principles in such hierarchical processing models is the RFs of neurons at one level of the hierarchy are constructed by first pooling a large number of inputs from neurons at a lower level, and then processing the combined signal with a set of canonical nonlinear functions, such as exponentiation (a form of thresholding) and normalization (Carandini and Heeger, 2012). Similar computational principles have been developed somewhat independently in machine learning field, such as the Neocognitron (Fukushima, 1980), the convolutional neural network (Lecun, 1998) and other deep learning architectures (Schmidhuber, 2014). The impressive

performance of such models on image recognition tasks also demonstrate the power of hierarchical computation.

In this dissertation, I use a statistical model for sensory processing, the "Nonlinear Input Model" (NIM) (McFarland et al., 2013), which was inspired by the hierarchical nature of sensory processing. In the NIM framework, the complex computations performed by sensory neurons are assumed to arise in large part from the accumulation of relatively simple nonlinear operations across the sensory processing hierarchy. At each level of the hierarchy, the NIM describes neuronal processing as a sum over excitatory and inhibitory inputs (Fig. 1-2 B). The space of nonlinear functions are limited with the assumption that nonlinearities in sensory processing are dominated by spike generation, which causes both rectification of the inputs to the neuron, as well as rectification of the neuron's output. Because the form of NIM is consistent with the basic description of neural computation in terms of the integrate-and-fire neuron, model selection and parameter estimation can be guided by knowledge about the inputs to a given neuron, and the elements of the resulting model can often be related to specific physiological terms.

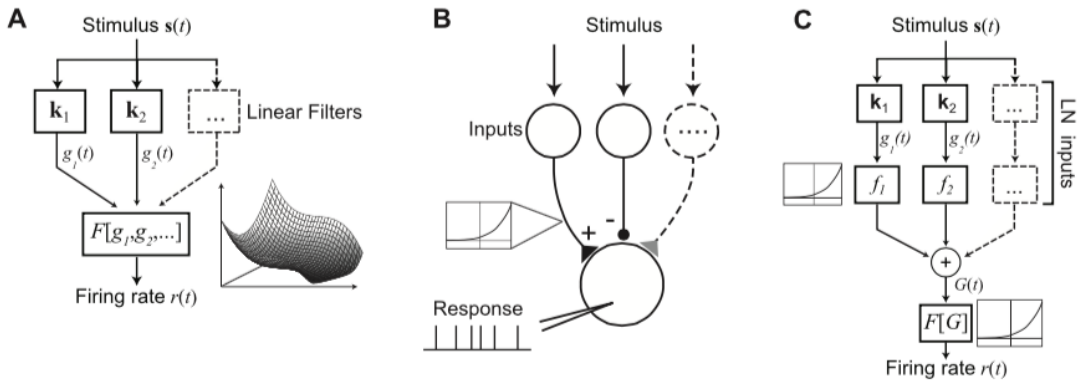

Figure 1-2 Schematics of the nonlinear input model (NIM).
A. Schematic diagram of a general LN model, with multiple filters ($\mathbf{k}_1$, $\mathbf{k}_2$, ...) that define the linear stimulus subspace. The outputs of these linear filters (g_1, g_2, ...) are then transformed into a firing rate prediction r(t) by the static nonlinear function F[g_1,g_2,...], depicted at right for a two-dimensional subspace. Note that while the general LN model thus allows for a nonlinear dependence on multiple stimulus dimensions, estimation of the function F[.] is typically only feasible for low (one- or two-) dimensional subspaces. **B**. Schematic illustration of a generic neuron that receives input from a set of 'upstream' neurons that are themselves driven by the stimulus **s**. Each of the upstream neurons provides input to the model neuron that is rectified due to spike generation (inset at left), and thus is either excitatory or inhibitory. The model neuron then integrates its inputs and produces a spiking output. **C**. Block diagram illustrating the structure of the NIM, based on (B). The set of inputs are represented as (one-dimensional) LN models, with a corresponding stimulus filter $\mathbf{k}_i$, and ''upstream nonlinearity'' f_i(.). These inputs are then linearly combined, with weights w_i, and fed into the spiking nonlinearity F[.], resulting in the predicted firing rate $r(t)$. The NIM thus has a 'second-order LN' structure (or LNLN), with the neuron's own nonlinear processing shaped by the LN nature of its inputs. Adapted from (McFarland et al., 2013).

In Chapter 2 and 3 of the dissertation, I describe the applications the NIM framework to the retina, and the motion processing area MT, respectively. In each case, I modified the original NIM model based on physiological and anatomical knowledge of the specific parts of the visual system. For the application to the retina, I utilized the fact that retinal ganglion cells (RGCs) receive synaptic inputs from bipolar cells, which are under the influence of presynaptic modulation mediated by amacrine cell (Euler et al., 2014; Wässle, 2004). The excitatory synaptic inputs to ganglion cells

were captured by a divisive suppression model, which provided an accurate description of both the intracellular and extracellular response of RGCs across contrasts. For the MT application, I incorporated the knowledge that MT receives inputs from direction-selective cells in V1 (Movshon and Newsome, 1996), which exhibit nonlinear response properties that are well characterized in literature. The resulting models for RGC and MT computation not only explain response properties not captured by the linear model, but also shed light onto the underlying computation along the visual pathway.

Variability of cortical neuron responses

Predictive power of feedforward models is limited by cortical variability

We have focused on the feedforward stimulus processing models so far. Indeed, in the standard framework of systems neuroscience, the activity of neurons in sensory cortex is assumed to be a fixed function of the stimulus, which is typically characterized by their "receptive field". This dominant view is manifested in the claim of "we know what the visual system does once we can predict neural responses to arbitrary stimuli" (Carandini et al., 2005). Such "receptive-field" based stimulus processing models are very successful in predicting neural responses in early visual system such as the retina (Pillow et al., 2005) (Chapter 2) and the LGN (Butts et al., 2011; Mante et al., 2008). Nevertheless, the ability of these models to explain neural responses is severely limited in visual cortex, even in principle. This is because neurons through out the cortical visual systems exhibit a remarkably high degree of

variability in response to repeated presentations of the same stimulus (Shadlen and Newsome, 1998; Softky and Koch, 1993; Tolhurst et al., 1983), which by definition cannot be captured in any type stimulus-processing models.

Intrinsic noise does not explain cortical variability

The high variability of cortical response was once thought to be intrinsic noise due to the stochastic nature of spike generation and neural transmission. Nevertheless, *in vitro* studies showed that spike responses to intracellular current injections occur with high fidelity (Mainen and Sejnowski, 1995). Although unreliability of synaptic transmission may have a bigger impact (Faisal et al., 2008), it is possible overcome this source of variability by integrating from many of them (Masquelier, 2013), and under some experimental conditions neural response does exhibit high degree of reliability (Haider et al., 2010; Herikstad et al., 2011; Kayser et al., 2010; Panzeri and Diamond, 2010). Therefore, it is improbable that the high variability of cortical neuron response is due to the intrinsic inaccuracy of biological systems. Instead, "randomness is only a measure of our ignorance of the different causes involved in the production of events" (Laplace, 1825), the reported high degree of cortical variability likely reflects our limited, incomplete understanding of the function of visual systems (Masquelier, 2013).

Non-stimulus-driven inputs to sensory neurons

Despite the focus on studying feedforward processing, the visual system is not an automatic machine that passively responds to the stimuli. Instead, visual processing must constantly adapt to changing in behavioral context and task demands. Indeed, as scientists start to study vision in more natural, behaviorally relevant conditions, visual system appears to be much more complex than the passive visual stimulus processing system that was originally imagined. Recent studies showed that responses of neurons in primary visual cortex are modulated by many factors beyond sensory stimulation, such as attention (Harris and Thiele, 2011; Reynolds and Heeger, 2009; Roberts et al., 2007), reward timing (Shuler and Bear, 2006), neuromodulatory signals (Avery et al., 2014; Goard and Dan, 2009), behavioral states (Niell and Stryker, 2010), locomotion signals (Ayaz et al., 2013; Keller et al., 2012; Saleem et al., 2013) and perceptual decisions (Britten et al., 1996; Nienborg et al., 2012). Even without any visual stimulation, visual cortex exhibits a rich set of spontaneous activities (Arieli et al., 1996; Buzsaki and Draguhn, 2004; Fiser et al., 2004), which is related with prior expectations (Berkes et al., 2011), changes in behavioral and cognitive state (Engel et al., 2001; Fries, 2005; Lakatos et al., 2005) and visual perception (Busch et al., 2009; Chakravarthi and VanRullen, 2012; Dugue et al., 2011).

Understanding the properties of the various non-stimulus-related inputs to visual cortical neurons presents a major challenge to the study of visual processing, both experimentally and computationally. This is because, during typical visual

experiments, investigators can only record from a very small fraction of neurons in a given area and often have no experimental control on the non-stimulus related input. However, several recent studies have demonstrated the possibility of inferring cortical network states from the local field potential (Arieli et al., 1996; Haslinger et al., 2012; Kelly et al., 2010; Kisley and Gerstein, 1999; Lakatos et al., 2005), a common element of extracellular recordings (Buzsaki and Draguhn, 2004). In Chapter 4 of the dissertation, I describe the problem of cortical variability in awake, behaving macaques under both passive and active conditions. Using sophisticated statistical models, I estimated the amount of cortical variability and quantified to what extent cortical variability can be accounted for by ongoing cortical network dynamics inferred from the LFPs (Chapter 4). By studying variability in a motion discrimination task, I found that the observed correlation between single neuron responses and perceptual decisions could be explained by slow oscillations in the cortex. These results provide unique insights into the source and functional significance of neural variability.

Linking visual neuron responses with perception.

Neural correlates of perceptual decision

The information carried by sensory neurons is ultimately used to create perception, our ability to interpret and understand the surrounding environment from sensory inputs (Schacter, 2011). The relationship between sensory neuron activity and perception remains one of the most important and intriguing problems in

neuroscience (Parker and Newsome, 1998), and understanding this relationship is essential for establishing a scientific account of the neural mechanisms underlying our mental life. In the past few decades, much effort has been devoted to search for perceptually relevant signals within the cerebral cortex. In this section, I review several landmark studies in this field and discuss how the work presented in this dissertation (Chapter 4) contribute to the understanding of the link between visual neuron responses and perceptual behaviors.

In 1989, Newsome et al. carried out a landmark study where neural responses and perceptual behaviors were simultaneously measured: single neuron responses in area MT were recorded while the subject performed a visual motion discrimination task (Newsome et al., 1989). The stimulus consists a field of dots, a fraction of which moves coherently in one direction while the rest moves randomly (Fig. 1-3 A). The subject has to report the motion direction with a saccade at the end of the stimulus presentation. By changing the fraction of dots that moved coherently, the investigators manipulated the difficulty of the task. The behavioral performance of the subject, quantified with a psychometric curve (Fig. 1-3 C), was found to be, on average, comparable to that of a hypothetical observer who performs the same task with decision based on responses of single MT neurons (i.e., the neurometric curve). This result demonstrated that single MT neuron responses are sufficient to explain the behavioral performance of the subject. Many subsequent studies have reported that when measured in appropriate perceptual decision tasks, behavioral performance can be explained by single or small number of neurons in sensory cortices, including V1

(Geisler and Albrecht, 1997; Nienborg and Cumming, 2006; Palmer et al., 2007), V2 (Nienborg and Cumming, 2006), S1 (Romo et al., 1998) S2 (Romo et al., 2004), MT (Andersen and Bradley, 1998; Croner and Albright, 1999; Uka and DeAngelis, 2004), MST (Celebrini and Newsome, 1994) and LIP (Law and I., 2008). In subsequent studies, investigators manipulated responses of relevant neurons with microstimulation (Britten and van Wezel, 1998; Cohen and Newsome, 2004; Romo et al., 1998; Salzman et al., 1990; Tehovnik et al., 2004) or chemical inactivation (Dürsteler and Wurtz, 1988; Newsome and Pare, 1988; Yamasaki and Wurtz, 1991), to show that activity of certain groups of sensory neurons are necessary for perceptual decisions.

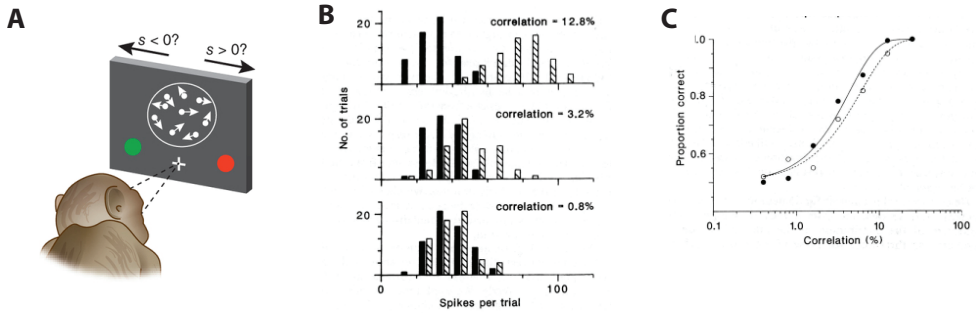

Figure 1-3 Neural correlates of perceptual decision.
A. The original motion discrimination task used in Newsome et al., 1989. The strength of the motion stimulus is controlled by varying the proportion of the dots moving coherently: at 0% coherence, the dot motion is completely random, at 100% coherence, all the dots are moving in the same direction. The monkey has to report the overall motion direction with a saccade at the end of the trial to get a liquid reward. **B**. The responses of a directionally selective MT neuron at three different motion strength levels. The hatched bars indicate responses to the preferred motion and the solid bars indicate responses to the null motion. **C**. Comparison of the psychophysical performance of the monkey (solid circle) with that of a hypothetical ideal observer who based its decision on the single neuron response shown in (B). (A): adapted from (Wohrer and Machens, 2013) (B-C): adapted from Newsome et al. (1989).

Choice probability for sensory neurons

Because the tight relationship between sensory neuron activity and perceptual decisions, it is natural to ask whether the high degree of response variability in sensory cortices is manifested in the behavior. Britten et al. first investigated this problem using data from the same motion discrimination task described above (Britten et al., 1996). This time the investigators turned their attention to the 0% coherence trials where the stimulus was completely ambiguous and reward was given randomly. Even though the behavioral performance was at chance level, the subject was not aware of the lack of motion signal and still tried to make the best guess of the motion direction on those trials. Because the stimulus-driven component of the response was the same regardless of the decision, this condition was ideal to isolate the decision relevant component of neural response. The researchers thus quantified how well responses of single MT neurons predicted perceptual decisions on a trial-by-trial basis using a novel non-parametric metric called "choice probability" (CP). CP was defined as the area under the receiver operating characteristic (ROC) curve calculated using the choice-conditioned distributions of neural responses. CP ranges from zero to one, and takes a value of 0.5 if the two choice-conditioned distributions of neural responses are completely overlapping. Conceptually, CP represents how well a hypothetical ideal observer can predict the subject's decision based on responses of single MT neurons. The critical finding of Britten et al., 1996 was that at the population level, CP significantly (though just slightly) above the chance level of 0.5 (mean=0.555), meaning that MT neuron activity is correlated with decision on a

trial-by-trial basis. Significant CP have been reported in a variety of cortical areas, including visual areas V1 (Nienborg and Cumming, 2014; Palmer et al., 2007), V2 (Nienborg and Cumming, 2006), MT (Bradley et al., 1998; Cohen and Maunsell, 2009; Purushothaman and Bradley, 2005), MST (Gu et al., 2007), IT (Uka et al., 2005), LIP (Roitman and Shadlen, 2002; Shadlen and Newsome, 2001), somatosensory cortex (Romo et al., 2002), motor and premotor cortex (de Lafuente and Romo, 2006; Romo et al., 2004), auditory cortex (Niwa et al., 2013) and subcortical vestibular neurons (Liu et al., 2013).

Relationship between choice probability and network activity

CPs for sensory neurons are typically small (<0.6). However, it is remarkable that they are detectable at all. A sensory stimulus typically activates many thousands of cortical neurons. If all of them contribute to the decision independently, the chance of finding one neuron that shows significant correlation with the behavior is vanishingly small. This apparent paradox can be resolved if activity of nearby cortical neurons fluctuate in a correlated manner (Haefner et al., 2013; Nienborg et al., 2012; Shadlen et al., 1996). This is indeed the case in sensory cortex. The correlation between the trial-to-trial fluctuations in activities of a pair of neurons, which is often termed as the noise correlation (Bair et al., 2001), is positive (but weak) for nearby neurons with similar stimulus tuning properties (Cohen and Kohn, 2011). Recent experiment evidences and theoretical analysis further showed that CP does not simply depend on the overall level of correlation; instead, it depends on the structure of these correlations (Bondy and Cummings, 2013; Haefner et al., 2014). Specifically, CP will

arise when correlations within a pool supporting one decision are higher than those between pools (Haefner et al., 2013; Nienborg et al., 2012).

The origin of noise correlation that contributes to CP is highly debatable. Originally, it is thought to reflect shared, feed-forward sensory noise (Shadlen et al., 1996), such as noise of sensory receptors, synaptic transmission, and action potential generation. However, recent studies found that noise correlation in V1 can present over much larger distance in the V1 (Smith and Kohn, 2008) than in the retina (Greschner et al., 2011), suggesting that feedforward sensory noise cannot be the only source of noise correlation. Another potential source of shared inputs is the intrinsic connections within a cortical area. A number of recent studies demonstrated a preference of connections between neurons with similar stimulus tuning properties (Ahmed et al., 2012; Ko et al., 2011). Although intrinsic connection is likely a source of noise correlation, it cannot explain how the correlation structure in the cortex can flexibly change depending on the task context (Bondy and Cummings, 2013; Cohen and Newsome, 2008) and spatial attention (Cohen and Maunsell, 2009; Mitchell et al., 2009). Instead, these results are consistent with a third possibility that noise correlations and CPs reflect a top-down signal (Nienborg et al., 2012). In this top-down explanation of CP, a feature-selective feedback signal occurs after the decision is formed (Nienborg and Cumming, 2009), which may contribute to perceptual stability when the sensory signals are weak and ambiguous.

In Chapter 4 of the dissertation, I tested whether the choice related shared inputs to local population of sensory neurons that underlies noise correlation in perceptual discrimination tasks could be inferred from elements of the local field potentials (LFPs). I found that CP of single MT neuron could be well predicted by the phase of ongoing delta oscillation (<4 Hz). Moreover, over a wide range of stimulus strength, slow (delta-band) oscillations in the LFPs contain more information of the perceptual decisions than the stimulus. This result not only makes delta band LFP signals a potential candidate for perceptually relevant signal, but also has profound implications on the correlation structure between neurons across the cortex.

Chapter 2 : Divisive suppression describes retinal ganglion cell responses with high precision

Introduction

Neural computations in the retina are generated by complex circuits specific to each ganglion cells type (Masland, 2001). Despite the multiple levels of processing within the retina shaping the output layer of retinal ganglion cells, retinal processing appears to be surprisingly linear, and as a result, ganglion cells have typically been modeled using linear-nonlinear (LN) cascade model, where a linear spatiotemporal receptive field is combined with a static nonlinearity that captures the spike threshold and response saturation (Baccus and Meister, 2002; Chichilnisky, 2001). However, many aspects of ganglion cell spike trains deviate from predictions of the LN model – especially when studying responses at time resolution higher than the stimulus itself (Butts et al., 2007), most notably contrast adaptation (Kim and Rieke, 2001; Meister and Berry, 1999; Shapley and Victor, 1978) and the high temporal precision (Passaglia and Troy, 2004; Uzzell and Chichilnisky, 2004).

To explain the various nonlinear aspects of ganglion cell responses, a multitude of nonlinear descriptions of retinal computation have been proposed, and in general have taken two approaches. The first describes the nonlinear function between stimulus and response as a mathematical expansion, expanding from the linear receptive field (Chichilnisky, 2001) to second-order quadratic terms, using spike-triggered covariance (Fairhall et al., 2006; Gollisch, 2013; Vaingankar et al., 2012) or

maximally informative dimension analyses (Sharpee et al., 2004). While such "expansion terms" can lead to better predictions of the spike train, they are challenging to interpret both functionally and with respect to the underlying circuit (Butts et al., 2011; McFarland et al., 2013). A second approach targets specific aspects of the response, such as spike-refractoriness (Berry and Meister, 1998; Keat et al., 2001; Pillow et al., 2005), gain changes associated with contrast adaptation (Bonin et al., 2005; Mante et al., 2008; Meister and Berry, 1999; Shapley and Victor, 1978) and the interplay of excitation and inhibition (Butts et al., 2011). While such approaches are often successful at describing the particular nonlinear effects targeted, they necessarily lump all nonlinear sources into the single nonlinear description, and usually can only explain the specific responses properties targeted.

Here we derive a novel nonlinear modeling framework inspired by a physiological description of retinal circuitry, constrained by recordings at two stages of retinal processing: synaptic input currents (using voltage clamp) and spike output (through cell-attached loose patch) of ON-Alpha ganglion cells. These recordings reveal that temporal precision and some elements of contrast adaptation are already present in the synaptic current input. We thus devise a tractable model of excitatory currents that incorporates a nonlinear structure based on the relevant circuit elements, In particular, by including a divisive suppression mechanism meant to capture presynaptic inhibition from amacrine cells, the model captures the detailed structure of synaptic current response and explained and contrast adaptation in the synaptic current. We then extend this model to explain spike outputs of the ganglion cell with the inclusion

of a spike history term and the spiking nonlinearity. The multiple nonlinear elements in the model accurately captured the precision of the spike outputs and also predicted the observed contrast adaptation of the spike output. Therefore, our study established a unified model of nonlinear processing within ganglion cells that accurately captured both the generation of temporal precision and contrast adaptation, providing a foundation for understanding in more complex circumstances.

Methods

Neural recordings

Data were recorded from identified types of retinal ganglion cells from mouse retinas using whole cell recording techniques that have been described previously (Wang et al., 2011). Spikes were recorded in the loose-patch configuration using a patch pipette filled with external Ames medium and synaptic currents were recorded using a second pipette filled with intracellular solution. The targeted cell was voltage clamped at E_{Cl} (-67 mV) to record excitatory currents after correcting for the liquid junction potential. Cells in the ganglion cell layers with large somas (20-25 μm diameter) were recorded to target the Y/alpha-type ganglion cells. Recorded cells were confirmed to be On-Alpha cells based on previously established criteria (Borghuis et al., 2013): (1) each cell had a relatively wide dendritic tree (300-400 μm diameter) and (2) each cell stratified on the vitreal side of the nearby ON cholinergic (starburst) amacrine cell processes. All procedures were conducted in accordance

with National Institutes of Health guidelines under protocols approved by the Yale University Animal Care and Use Committee.

Visual Stimulation

The temporally modulated spot stimulus were described in (Wang et al., 2011). The retina was stimulated by UV LEDs (peak, 370 nm; NSHU-550B; Nichia America) to drive cone vision. UV LEDs were diffused and windowed by an aperture in the microscope's fluorescence port, with intensity controlled by pClamp 9 software via a custom noninverting voltage-to-current converter using operational amplifiers (TCA0372; ON Semiconductor). The stimulus was a flickering spot (1000 μm diameter), with intensity generated from low pass Gaussian noise with a 30 Hz cutoff frequency. We used a contrast-switching paradigm (Baccus and Meister, 2002; Kim and Rieke, 2001; Zaghloul et al., 2005), where the temporal contrasts alternate between high and low every 10 sec. The contrast of the stimulus is defined by the standard deviation of the Gaussian noise. The stimulus comprised 10 cycles of 10 sec for each contrast. The first 7 sec were unique in each cycle, and the last 3 sec were repeated across cycles. The stimuli were programmed in *Matlab* (Mathworks, Natick) using the *Psychophysics Toolbox* (Brainard, 1997).

The center-surround stimuli (Fig. 2-4B) were generated using custom-written language (Stim-Demo, C language), and presented with video projector (M109s DLP; Dell, or identical HP Notebook Companion; HP), modified to project UV light (single LED NC4U134A, peak wavelength 385 nm; Nichia)(Borghuis et al., 2013). In this stimulus, a spot covered the receptive field center (e.g., 0.3 mm), and an annulus

extended into the surround (e.g., inner/outer diameters of 0.35/1.0 mm). The center and surround stimuli were independently generated Gaussian noise with 30 Hz cutoff frequency. For all methods of stimulation, the gamma curve was corrected to linearize output, and stimuli were centered on the cell body and focused on the photoreceptors.

Statistical modeling for synaptic current response

In the traditional linear-nonlinear (LN) cascade model (Paninski, 2004; Truccolo et al., 2005), the transformation of the stimulus $\mathbf{s}(t)$ to the synaptic current response $c(t)$ was described as,

$$c(t)=f_1[\mathbf{k}_{lin}\cdot\mathbf{s}(t)] + c_0, \tag{2.1}$$

where $\mathbf{k}_{lin}$ denoted the linear filter, $f_1(\cdot)$ denoted a static nonlinearity and c_0 is a baseline offset.

To incorporate a second synaptic pathway into the model, we first considered a two-dimensional nonlinearity $F[\cdot,\cdot]$, with

$$c(t)=F[\mathbf{k}_e\cdot\mathbf{s}(t), \mathbf{k}_s\cdot\mathbf{s}(t)], \tag{2.2}$$

where $\mathbf{k}_e$ and $\mathbf{k}_s$ denoted the excitatory and suppressive filters respectively. To limit the number of parameters, we fit the filters $\mathbf{k}_{lin}$, $\mathbf{k}_e$, $\mathbf{k}_s$ using a family of orthonormalized basis functions suggested by (Keat et al., 2001):

$$\zeta(t) = \sin[\pi n(2t/t_F - (t/t_F)^2)], \tag{2.3}$$

where t_F=250 ms. The filters can be alternatively optimized for a given choice of nonlinearity using gradient-descent based method to minimize the mean-squared error,

$$\mathrm{MSE} = \sum_t [c(t) - c_{obs}(t)]^2 \tag{2.4}$$

The 2-D nonlinearity was represented using piece-wise planar surfaces and can be estimated non-parametrically for a given choice of filters. Specifically, we divided the 2D space into a set of uniform squares, and applied the same triangulation scheme to all squares to subdivide each square into two triangles. Each basis function was defined as a hexagonal pyramid function centered at one of the vertices. The 2-D nonlinearity function was expressed as a combination of the basis,

$$F[x,y] = \sum_{i,j} w_{ij} f_{ij}(x,y), \tag{2.5}$$

where $f_{ij}(x,y)$ is the basis centered at the ij*th* grid vertex, and w_{ij} is the weight coefficient, which can be optimized by minimizing MSE for a given choice of filters.

There is degeneracy in the combined optimization of stimulus filters and 2D nonlinearity. For example, one can choose a linear combination of the two stimulus filters and achieve the same model performance by refitting the 2D nonlinearity. To alleviate this problem, we constrained the 2D nonlinearity to be monotonically increasing along the first dimension, i.e.,

$$\text{If } x>x', \text{ then } F[x,y] \geq F[x',y], \ \forall\, y \tag{2.6}$$

Therefore, the first dimensions in the nonlinearity can be interpreted as excitatory dimension. No constraint was applied on the second dimension, as suppression was found to exhibit selectivity to both ON and OFF stimuli.

While this formulation of nonlinear model is general, the structure of the 2-D nonlinearity is generally not straightforward to interpret. To determine if there is simpler computational structure underlying $F[\cdot,\cdot]$, we considered a low-rank

decomposition of the 2-D nonlinearity as the sum of one-dimensional divisive interactions

$$c(t)=f_e[\mathbf{k}_e\cdot\mathbf{s}(t)] \times f_s[\mathbf{k}_s\cdot\mathbf{s}(t)] + c_0. \quad (2.7)$$

Each divisive term is in the form of a 1-D LN model. Although in principle a large number of divisive terms were required to fully capture the 2-D nonlinearity, we found in practice that only one term was sufficient. The resulting divisive suppression model (DS model) often has comparable performance as the 2-D model with much less parameters. To enhance the interpretability of the model, we constrained the excitatory nonlinearity f_e to be monotonically increasing, and the suppressive nonlinearity f_s be one when the input is zero, and less than or equal to one everywhere else. By alternatively optimizing the filters and the nonlinearities until convergence, and sampling different filter initializations, we can robustly find filters that minimize the MSE.

Statistical modeling for spike response

We have applied several statistical models to describe the spike response of ganglion cells. We first considered the generalized linear modeling (GLM) framework (Paninski, 2004; Simoncelli et al., 2004; Truccolo et al., 2005). The GLM makes prediction of the instantaneous firing rate of the neuron r(t) based on both the stimulus $\mathbf{s}(t)$ and the recent history of observed spike train $\mathbf{R}(t)$:

$$r(t)= F_{spk}[\mathbf{k}_{lin}\cdot\mathbf{s}(t) + \mathbf{h}_{spk}\cdot\ \mathbf{R}(t) - \theta], \quad (2.8)$$

where $\mathbf{k}_{lin}$ is the linear receptive field, $\mathbf{h}_{spk}$ is the spike history term and θ is the spiking threshold. Here, the parameters of the model are all linear functions inside the spiking nonlinearity F_{spk}. The linear-nonlinear (LN) model consists of only the linear receptive field and the spiking threshold; the full GLM further includes the spike history term (denoted as LN+RP in figures). The spiking nonlinearity has a fixed functional form $F_{spk}[g] = \log[1+\exp(g)]$, satisfying conditions for efficient optimization (Paninski, 2004). The model parameters are estimated using maximal-likelihood optimization. The log-likelihood (LL) of the model parameters given the observed neural response is (Paninski, 2004):

$$LL=\sum_t[r_{obs}(t) \log r(t) - r(t)] \tag{2.9}$$

Given a set of model parameters, the LL and its gradient can be directly calculated. The optimal model parameters can be found using gradient-descent based optimization of LL.

To capture nonlinear properties of the spike response, we modified the Nonlinear Input Model (NIM) (McFarland et al., 2013). The NIM framework incorporates nonlinear processing underlying a neuron's inputs. The predicted firing rate is given as,

$$r(t)= F_{spk}[\sum_i f_i[\mathbf{s}(t)] + \mathbf{h}_{spk} \cdot \mathbf{R}(t) - \theta], \tag{2.10}$$

where f[] represents a set of nonlinear subunits reflecting upstream processing. In this case, base on knowledge of nonlinear processing in synaptic current response, we assumed the nonlinear subunit takes the form of a DS model,

$$f[\mathbf{s}(t)] = f_e[\mathbf{k}_e \cdot \mathbf{s}(t)] \times f_s[\mathbf{k}_s \cdot \mathbf{s}(t)]. \tag{2.11}$$

Similar to parameter estimation of the DS models of synaptic current response, we alternatively estimated the filters and nonlinearities until converge. The same set of constraints was applied to the excitatory and suppressive nonlinearities.

Quantification of contrast adaptation with LN analysis

We performed LN model analysis to ganglion cell responses across contrasts and subsequently quantified the change of linear filter to gauge the extent of contrast adaptation. We used a shared nonlinearity and separate linear filters for the low and high contrast condition such that contrast adaptation effects are attributable entirely to changes in linear filter. The two linear filters can be simultaneously fit to the data using the shared nonlinearity, and the nonlinearity can then be estimated with the filters fixed. This approach is similar to earlier attempts to "align" nonlinearities across contrasts and applied the associated scaling factor to the linear filters (Chander and Chichilnisky, 2001). For the synaptic current data, we further introduced an offset factor to compensate for the change in the overall mean of the response (Baccus and Meister, 2002).

Once the linear filters at both contrasts were obtained, we identified the global peak and valley of the filters denoted as k_{peak} and k_{valley} respectively, and then calculated the percentage change of peak-to-valley amplitude (k_{peak}- k_{valley}) and biphasic index ($|k_{valley}/k_{peak}|$). These indices reflect the change of filter magnitude and shape across contrasts.

Evaluation of model performance

Because complex models, such as the DS model, are more prone to overfitting than simpler models. It is important to evaluate model performance on datasets not used for model fitting (i.e., cross-validation). We used cross-validation throughout the paper and evaluated model performance on the 3-seconds repeat trials that are set aside for model fitting. Several complementary metrics were used for performance evaluation. First, we used predictive power, or percent of explainable variance (David and Gallant, 2005; Sahani and Linden, 2003), to quantify how well the model captured the trial-averaged response for both intracellular and extracellular recordings. This metric corrects for noise-related bias due to limited number of trials. Second, we used cross-validated likelihood for the spike models. This metric concerns the timing of spikes on a trial-by-trial basis. For example, LN model with and without the spike refractoriness term have similar predictive power but very different cross-validated likelihood.

Coherence analysis of synaptic current response

The general model performance metrics such as predictive power and cross-validated likelihood do not reflect which aspect of the response does the model not capture. We thus devised a new coherence-based metric to quantify how well the model performs across frequencies. The coherence between the model predicted current response c(*t*) and the recorded current response on the *i*th trial $c_{\mathrm{obs}}^{i}(t)$ is (Butts et al., 2007):

$$\gamma_i^2(\omega) = \frac{\langle |\, C_{\mathrm{obs}}^{i}(\omega)\overline{C(\omega)}|^2\rangle}{\langle |\, C_{\mathrm{obs}}^{i}(\omega)|^2\rangle\langle |C(\omega)|^2\rangle} \tag{2.12}$$

where $\mathrm{C}(\omega)$ and $\mathrm{C}_{\mathrm{obs}}^{i}(\omega)$ are the Fourier transforms of $\mathrm{c}(t)$ and $\mathrm{c}_{\mathrm{obs}}^{i}(t)$ respectively, and the bar denotes complex conjugate. We used angular frequency $\omega = 2\pi f$ instead of f to be consistent with common conventions. Once we get the coherence measure on individual trials, we average it across repeat trials for each cell.

Because the observed response on each trial contains noise, a coherence of one throughout the frequency is not a realistic target. To correct for this bias, we calculated the coherence between the trial-averaged current response (i.e., the ideal predictor of response) and the recorded current on each trial. This noise corrected coherence metric represents an upper bound of coherence that can be achieved by any stimulus-processing model. It also reflects the consistency of current response at each frequency range. For example, in the low contrast condition, the response contained little high frequency component (Fig. 2-7A-B), the measured coherence was then close to zero above 30 Hz.

Event analysis of spike trains

We modified a previously established method to identify short spiking episodes "events" in the spike train (Butts et al., 2010). Analyses in the same spirit have been previously applied by (Berry et al., 1997; Kumbhani et al., 2007). Specifically, events were first defined in the peristimulus time histogram (PSTH) as times of firing interspersed with periods of silence lasting ≥ 8 ms. Each resulting event was further analysis by fitting the PSTH with a two-component Gaussian mixture models. An event was broken into two events if the differences of means of the two Gaussian components exceed two times the sum of standard deviations. Event boundaries were

defined as the midpoint between neighboring event centers and were used when assigning event labels to simulated spikes. Events were excluded from further analysis if no spike was observed on more than 50% of the trials during the event window. This criterion excluded spontaneous spikes that occur only on a few of the trials. Event analysis was first performed on responses at high contrast condition. Once the events in high contrast were determined, events at low contrast were then defined using the event boundaries obtained from high contrast data. These particular methods were chosen because they gave the most reasonable results with regards to visual inspection, but the results presented here do not qualitatively depend on the precise methods.

Once event were parsed, we measured a variety of properties associated with the spikes that comprised them (Fig. 2-6). The first-spike SD measured the standard deviation of the first spike of the event on each trial, which reflects the precision of event. The event time scale is estimated as the standard deviation of all spike times in each event, which is related to the duration of each event. The event Fano factor measures the ratio between the variance of spike count and the mean spike count in each event.

Results

We recorded spikes from identified ON-Alpha ganglion cells in the mouse retina with a loose-patch configuration, during the presentation a temporally modulated 1-mm spot that was centered on the neuron's receptive field (Fig. 2-1A, *top*). Contrast was changed between two levels every 10 seconds. In high contrast (HC), the stimulus

yielded spike responses that were precisely timed from trial to trial (Fig. 2-1A, *left*), as observed in numerous previous studies (Berry and Meister, 1998; Butts et al., 2007; Passaglia and Troy, 2004; Reinagel and Reid, 2000; Uzzell and Chichilnisky, 2004). Typically, ganglion cell responses are predicted using a cascade model based on the neuron's linear receptive field (RF) (Chichilnisky, 2001; Hunter and Korenberg, 1986). For the linear-nonlinear (LN) model, the neural response is predicted by the degree that the stimulus $\mathbf{s}(t)$ matches the linear receptive field $\mathbf{k}$ (Fig. 2-1B, *left*), and then scaled to a firing rate through the spiking nonlinearity (*right*). Although the LN model predicts the coarse dynamics of firing rate fluctuations, it does not capture the fine temporal features of the response (Fig. 2-1A). Specifically, using an established "event" analysis that divides the spike train into separate episodes of firing separated by silence (Butts et al., 2010), we compared both the average first-spike jitter and overall event duration between LN model and observed data (see Methods), demonstrating the failure of the LN model to capture either (Fig. 2-1C).

In the low contrast (LC) condition, the overall firing rate was comparable to the HC condition, consistent with many previous studies demonstrating contrast adaptation in the retina (Baccus and Meister, 2002; Rieke, 2001; Zaghloul et al., 2005). This is reflected in the LN model fits as the change of the temporal filter magnitude across contrasts (Fig. 2-1B, *left*). On the other hand, the response precision degraded significantly in LC condition (Fig. 2-1A, *right*): individual events were much broader than the HC condition, and the first spikes in each event were less precise. As a result,

there was no significant difference in the precision of the measured ganglion cell spike trains and the LN model predictions for low contrast, suggesting the adequacy of the linear model in LC, and thus a change in the amount of nonlinearity across contrast.

Nonlinear processing distributed across two stages of retinal processing

To understand whether temporal precision and contrast adaptation are generated within ganglion cells – as suggested in models using spike-refractoriness (Berry and Meister, 1998; Gaudry and Reinagel, 2007a; Keat et al., 2001) – or arise earlier in the retinal circuit and inherited by ganglion cells, we recorded excitatory synaptic current inputs from the same ganglion cells as those recorded with spikes (Fig. 2-1D). The resulting excitatory current could also be fit with an LN model (Baccus and Meister, 2002; Zaghloul et al., 2005), which for the HC condition resembled that of the spike-based model (Fig. 2-1E, *left*), except that its nonlinear stage was not rectifying and relatively linear (*right*), as expected.

Like the comparison between the spike response and LN model prediction, the excitatory currents also had more precise features than the LN model prediction. At the same time – like the spike responses also – the LC excitatory current was well described by the LN model. The precision in excitatory currents relative to the LN model suggests the currents have higher frequency content than the stimulus. To test this, we measured the coherence between trial-averaged response and responses on individual trials (see Methods), which reflects the consistency of the response at

different frequency bands across repeats. Indeed, there was appreciable coherence at frequencies higher than the stimulus in HC, while the LN model did not have frequency content any higher than that of the stimulus (Fig. 2-1F). In contrast, in LC the coherence of the excitatory current matched that of the LN model prediction. Together, these observations suggest that precision in ON-Alpha ganglion cell spike responses is generated (at least in large part) by retinal circuitry upstream of ganglion cells.

While the temporal structure of excitatory currents ad spike trains appears similar, they had markedly different adaptation to contrast, as gauged by comparing the linear receptive fields of the LN model (Fig. 2-1B). While the contrast was decreased by a factor of three between HC and LC conditions, the magnitude of the spike response did not undergo a similar decrease, reflecting contrast gain (Shapley and Victor, 1978). This gain is reflected in the magnitude of the linear receptive field (Fig. 2-1B) (Chander and Chichilnisky, 2001). In contrast, the decrease in stimulus amplitude was directly reflected in a similar decrease in the magnitude of excitatory current (Fig. 2-1D), reflecting very little gain change in the LN model (Fig. 2-1E). Therefore, much of the contrast adaptation observed in the spike outputs was not inherited from upstream processing, and is likely due additional processing and intrinsic mechanisms at the level of the ganglion cell (Kim and Rieke, 2001).

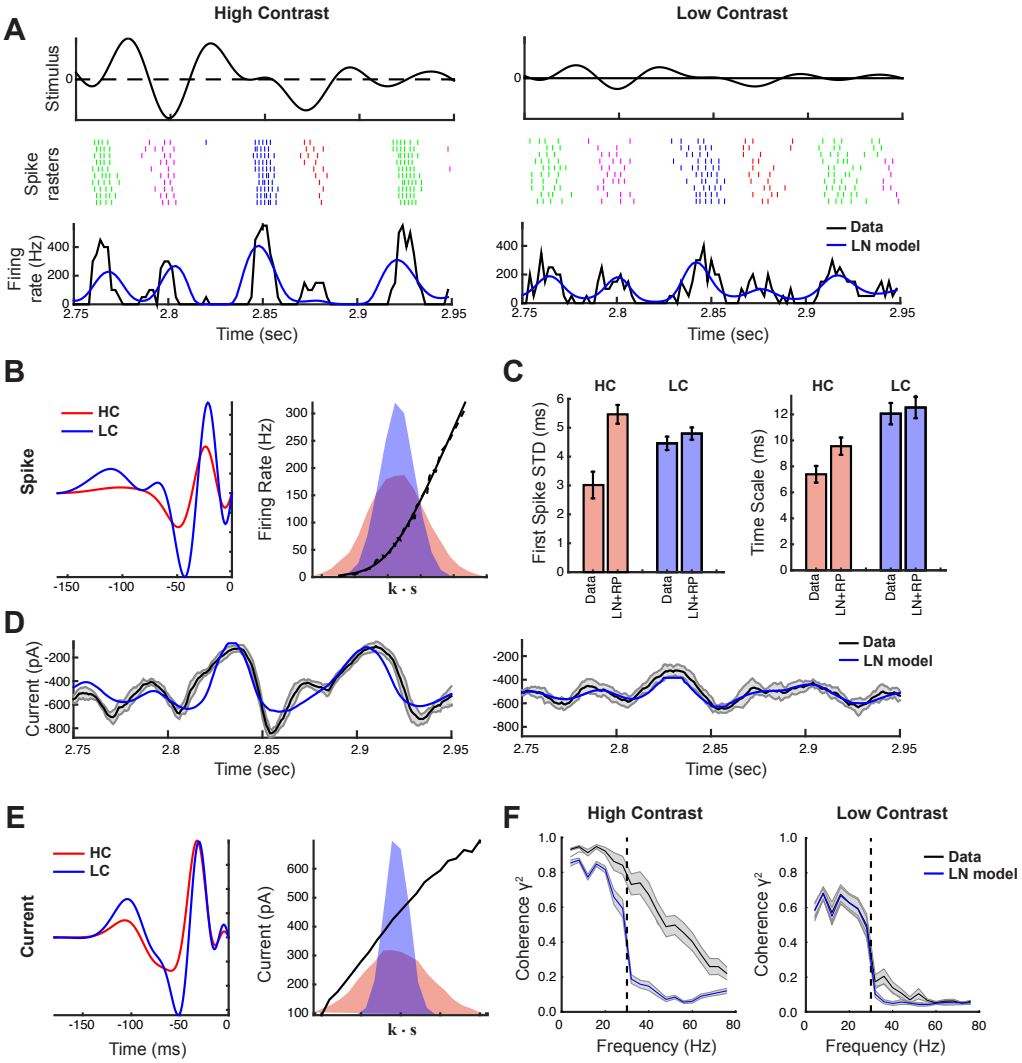

Figure 2-1 Nonlinear processing distributed across two stages of retinal processing.

A Spike response of an ON-Alpha cell to 10 repeated presentations of a temporally modulated noise stimulus (top) across two contrast levels (left vs. right). Colors in the spike raster label separate spike "events" using established analyses (middle). The firing rate predictions of the LN model (blue) are compared with the PSTH of the neuron's response (black) at the bottom. **B**. The linear filters of the LN model exhibit contrast adaptation (left). A shared spiking nonlinearity function is used across contrast (right) so that contrast adaptation effect is solely reflected in changes of temporal filters. We used a parametric mathematical from of the spiking nonlinearity (solid black line, see Methods), which agrees well with the non-parametric estimation (dashed line). The shaded areas are distribution of the filtered stimulus at high (red) and low (blue) contrasts. **C**. Temporal properties of real and LN-model simulated spike events. Left: SD of first spike in each event. Right: Event time scale as estimated by the SD of all spikes in the event. The LN model has much worse

precision than observed at high contrast. This difference is diminished at low contrast. **D**. Excitatory synaptic current response from the same cell with the same stimulus (black) compared with the LN model predictions (blue). The gray shaded area indicates SD of the response across trials, reflecting minimal variability in the recording. **E**. LN model fits to the current data. The temporal filters (left) show less contrast adaptation effect than the spike data (Fig. 2-1B). The nonlinearity in the LN model is close to a linear function (Fig. 2-1E). **F**. The coherence between response on individual trials and trial-averaged current response (black) or LN model predictions (LN). This data measures the consistency of responses across frequencies, and the model predicted coherence (blue) reflects how well the model captured each frequency components. The LN model fails to capture high frequency component in the current response at high contrast (left), but agrees well with the data at low contrast (right).

The nonlinear computation underlying synaptic inputs to ganglion cells

Knowledge of retinal circuitry greatly informs the ability to understand the sources of nonlinear computation in ganglion cells (Fig. 2-2A). Ganglion cells receive excitatory inputs solely from bipolar cells. Because voltage responses of bipolar cells are well described by an LN model (Baccus and Meister, 2002), the nonlinear mechanisms observed in the excitatory current to ganglion cells are likely localized at the bipolar-ganglion cell synapse (Fig. 2-2A). In particular, many studies showed that glutamate release from the bipolar cell terminal is subject to presynaptic inhibition provided by amacrine cell (Euler et al., 2014; Schubert et al., 2008; Wässle, 2004), representing a likely source of nonlinear computation. Note that an alternative mechanism, considered later, is synaptic depression at the bipolar cell terminal (Jarsky et al., 2011; Ozuysal and Baccus, 2012).

We thus constructed a "divisive suppression" (DS) model to capture the computations potentially performed by presynaptic inhibition mediated by amacrine cells (Fig. 2-2A, *bottom left*). We modeled both the bipolar cell excitation and amacrine cell

inhibition each as an LN model, with a multiplicative interaction of the amacrine LN output on the bipolar LN output (Fig. 2-2A, *bottom right*). In this formalization, the amacrine model has an output that drops below one when the stimulus matches the suppressive filter, causing a proportional decrease in excitatory current. Such a model could be tractably fit to current data using recent advances in statistical modeling (Ahrens et al., 2008b; McFarland et al., 2013). Note that the DS model reduces to the LN model if the suppressive nonlinearity is estimated to be one over the entire range of suppressive filter output.

The DS model fits were highly consistent across the population, with an excitatory filter and closely matching, but delayed, suppressive filter (Fig. 2-2B). The suppressive filters were delayed relative to the excitatory filters for every cell (mean delay = 8.9±1.4 ms, p=0.002, Fig. 2-2C). The excitatory nonlinearity was linear over the range tested (Fig. 2-2D, *left*), whereas the suppressive nonlinearity decreased below one when the stimulus either matched or was opposite to the suppressive kernel (Fig. 2-2D, *right*), resulting in both ON and OFF selectivity.

The DS model assumed a particular mathematical form of interaction between excitation and suppression (*i.e.*, separable and multiplicative). We also tested more general forms of nonlinear interactions by directly estimating the two-dimensional function, which maps each combination of the outputs of the excitatory and suppressive filters to a predicted current (Fig. 2-2E, *left*; see Methods). Although this 2D model contains many more parameters than the DS model, it did not have

significantly better model performance (Fig. 2-2F), and the estimated 2D nonlinearities were well approximated by the separable mathematical form of the divisive suppression model (R-squared for 2D nonlinearity reconstruction = 0.94±0.02; Fig. 2-2G). We also tested an additive suppression (AS) model, where suppression interacts with excitation additively (see Methods), but this had significantly worse predictive performance (Wilcoxon signed rank test, p=0.002; Fig. 2-2F) and less resemblance to the 2D nonlinearities than the DS model (Wilcoxon signed rank test, p=0.002; Fig. 2-2G). Therefore, the DS model gives a parsimonious description of the nonlinear computation at the bipolar-ganglion cell synapse and yields more interpretable model components.

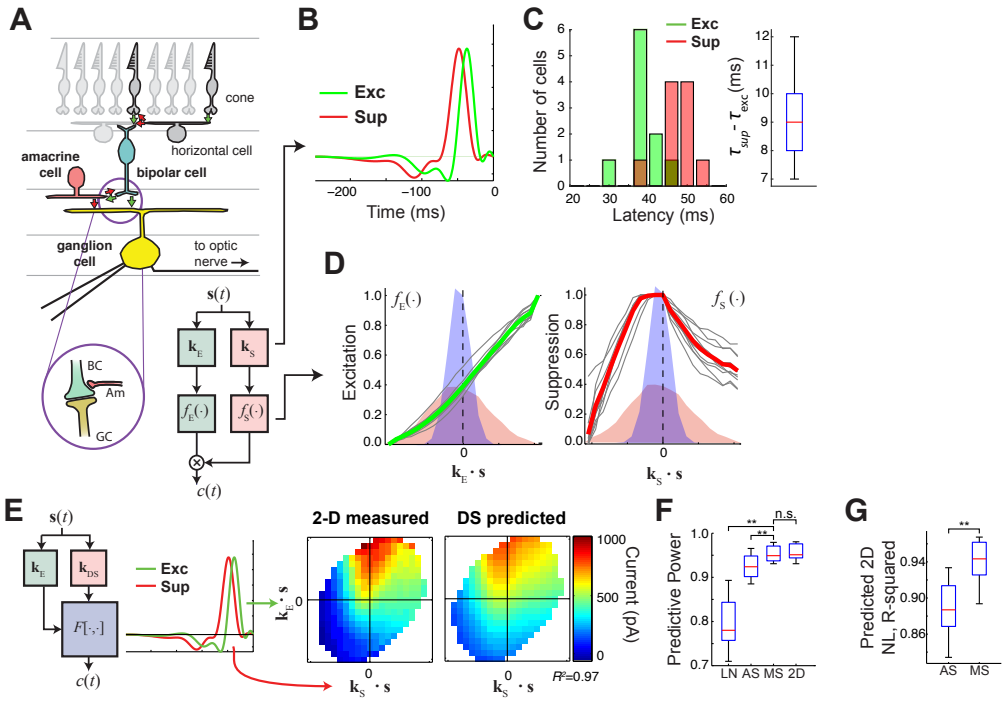

Figure 2-2 The divisive suppression (DS) model of synaptic currents.
A. Diagram of retinal circuitry: light is detected by cone photoreceptors (top), which excite bipolar cells (cyan), which in turn excite ganglion cells (yellow). Amacrine cells (red) are interneurons that receive excitatory input from bipolar cells and can

inhibit presynaptic terminals of ganglion cell. The presynaptic inhibition from amacrine cells is modeled as a divisive suppression model (bottom inset), where amacrine cell outputs multiplicatively modulates excitatory inputs from the bipolar cell. **B**. The excitatory (green) and suppressive (red) temporal filters of an example cell. **C**. Distribution of latencies for the excitatory and suppressive filter (left). Across the population, suppression is consistently delayed relative to excitation (right), with a mean delay = 8.9±1.4 ms (Wilcoxon signed rank test, p=0.002). **D**. Excitatory (left) and suppressive nonlinearities (right) for the DS model. The solid line indicates model fits for the example cell, and the gray lines are from other cells in the population. For each nonlinearity, the distribution of filtered stimuli is also shown as the shaded area for HC (red) and LC (blue). Note that suppression occurs for both stimuli that match the kernel or are opposite, implying that it is ON-OFF. **E**. To validate the divisive assumption of the DS model, a more general model is fit using a 2-D nonlinearity, which makes no assumptions about the interaction between the filters (see Methods). For the example ON Alpha cell, the filters are identical (left), and the measured 2-D nonlinearity (middle) matches that predicted (right) by multiplying the two 1-D nonlinearities shown in (B). **F**. The performance of the different models of the ON Alpha cell is tested on the three-second repeats that were not used for model fitting. The DS model significantly outperforms the LN model and an additive suppression model (AS), and the 2-D model does no better. ** Wilcoxon signed rank test, p=0.002. **G**. Accuracy of 2-D nonlinearity reconstruction with DS model and AS model. The multiplicative interaction assumed in the DS model performs significantly better than the additive assumption (** Wilcoxon signed rank test, p=0.002).

Divisive suppression explains temporal precision and contrast adaptation in the synaptic current input

In addition to the nearly perfect performance in predicting the excitatory synaptic current in the HC condition (Fig. 2-2), the DS model also generalized across contrast. A single set of parameters was fit to HC and LC conditions, and the resulting model performance in both conditions was similarly accurate (Fig. 2-3A). In contrast, the LN model with a single set of parameters fit to both contrasts ("LN-HL") performed much worse. Indeed, contrast adaptation is often measured by comparing the LN model filter in HC versus LC conditions (*e.g.*, Fig. 2-1E), with the change in their temporal filter implying the presence of contrast adaptation (Chander and

Chichilnisky, 2001). Likewise, separately fit models ("LN-H" and "LN-L") performed better than LN-HL, but were still outperformed by the DS model in both conditions. As this implies, the DS model predicted the fine time scales of the excitatory current in HC (Fig. 2-3B, *left*), and also predicted the lack of fine time scales – past what was predicted by the LN model – in LC (*right*). This demonstrates that the DS model implicitly adapts to contrast with no associated changes in parameters.

How does divisive suppression explain the fine temporal features of the excitatory current in HC, and yet predicts their absence in LC? To understand this, we consider a typical cross-validated response and how it is predicted by the DS model in both contrasts (Fig. 2-3C, top). The stimulus is first processed by the excitatory and (delayed) suppressive filter. The output of the filters represents how closely each matches the stimulus as a function of time (second row). Note that because the stimulus contrast is three times smaller in the LC condition, the filter output is proportionately smaller, but otherwise identical. Because the nonlinearity that is processing the excitatory term is roughly linear (Fig. 2-2D, left), the excitatory output is proportional to the excitatory filter output (bottom row, green).

Maximal suppression occurs at both positive (orange) and negative (purple) peaks of the suppressive filter output, corresponding to ON and OFF suppression, respectively. In HC, the presence of these suppressive peaks results in large deviations of the DS prediction from the excitatory filter output both before and after peak excitation

(*bottom row*). However, in LC, the much smaller values of the suppressive filter output scale down the amount of suppression (gray), and thus the output of the suppressive term (*third row*) barely deviates from unity. As a result, the DS model in LC barely deviates from an LN model that matches the output of the excitatory current (*bottom row*, *right*).

The close match between observed data and DS predictions across contrasts suggest that the DS model should "adapt" to contrast similarly to the recorded currents on a cell-by-cell basis, as measured by LN analysis (*e.g.*, Fig. 2-1E). Indeed, the model predicted filters are very similar to the observed ones in each case (Fig. 2-3D). In the LC condition, the linear filters generally are more biphasic (increase of biphasic index=15.2%±16.7%, p=0.034, Wilcoxon signed rank test) and have larger peak-to-valley amplitude (increase of peak-to-valley amplitude=9.9%±16.7%, p=0.034, Wilcoxon signed rank test). These effects are accurately captured by the DS model predictions (Fig. 2-3E,F).

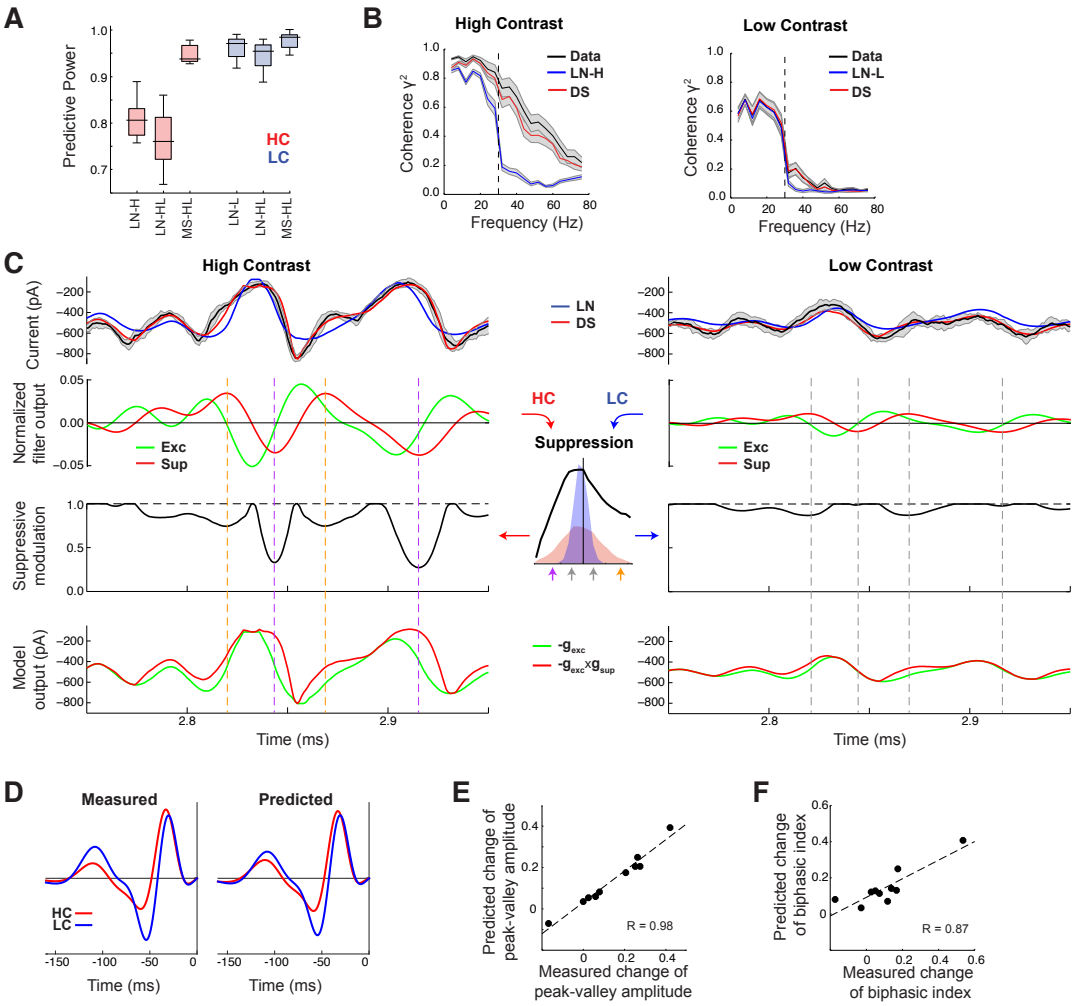

Figure 2-3 DS model explains temporal precision and contrast adaptation in the synaptic currents.

A. The predictive power of models across contrasts. A single DS model is fit to both contrasts, and outperforms both the LN model fit separately to each contrast (LN-H and LN-L). The LN model fit to both contrasts (LN-HL) performs even worse, as the LN model of course would ideally change with contrast. **B**. Average coherence between model predictions and recorded synaptic currents on individual trials (n=10). The DS model prediction (red) is almost identical with using the trial-averaged response (Data), suggesting that it captured all the current response across all frequency ranges. **C**. DS model explains precision and contrast adaptation through the interplay of excitation and suppression. Top: comparison of predictions of synaptic current response of the LN model (blue) and the DS model (red) for the cell shown in Fig. 2-1. 2nd row: normalized output of the excitatory (green) and delayed suppressive (red) filter. 3rd row: suppressive modulation is obtained by passing the filtered output through the suppressive nonlinearity (middle inset). Bottom: excitatory output of the DS model before (green) and after (red) the suppressive modulation. **D**. Comparison of the measured (left) and DS model predicted (right) filters across contrasts. **E-F**.

The DS model captured changes of both peak-to-valley amplitude (E) and biphasic index (F) of the temporal filters across contrasts.

Divisive suppression arises largely from the surround of the receptive field

We have suggested the delayed divisive suppression is due to presynaptic inhibition mediated by amacrine cells. However, another mechanistic explanation of the divisive suppression is synaptic depression (Jarsky et al., 2011; Ozuysal and Baccus, 2012), because depletion of a signaling mechanism can lead to a multiplicative decline of response amplitude (Zucker and Regehr, 2002). In this sense, the DS term that we fit could be capturing either the effects of presynaptic inhibition or synaptic depression, as both would lead to a multiplicative suppression on the synapse. Indeed, we evaluated this alternative explanation by applying the linear-nonlinear-kinetic model (LNK model), which fits ganglion cell intracellular recordings using a LN model with additional terms that simulate synaptic depression (Ozuysal and Baccus, 2012). The LNK model also outperforms the LN model and was able to capture fine time scale features of the observed currents (data not shown), although not good as the DS model (Fig. 2-4A).

These two mechanistic explanations might be distinguished by their spatial profiles: because synaptic depression acts on the same inputs that drive excitation (*i.e.*, it is a reduction in the amount of excitation), suppression should have the same spatial footprint as excitation. In contrast, presynaptic inhibition reflects modulation through a separate pathway and may have a different spatial profile. We thus presented stimuli

with two independently varying spatial components – a center and an annulus – to measure the spatial profiles of excitation and suppression (Fig. 2-4B, left). We found that the DS term indeed has a very distinct spatial profile than excitation, and it is more driven by the surround annulus, while excitation is mostly from the center (Fig. 2-4B). Moreover, the LNK model failed to capture the nonlinear response in this scenario, with no better model performance than the LN model (Fig. 2-4C). These experiments thus support presynaptic inhibition as a mechanistic explanation of DS for ON-Alpha ganglion cells, although does not preclude the presence or role of synaptic depression elsewhere in the retinal circuit.

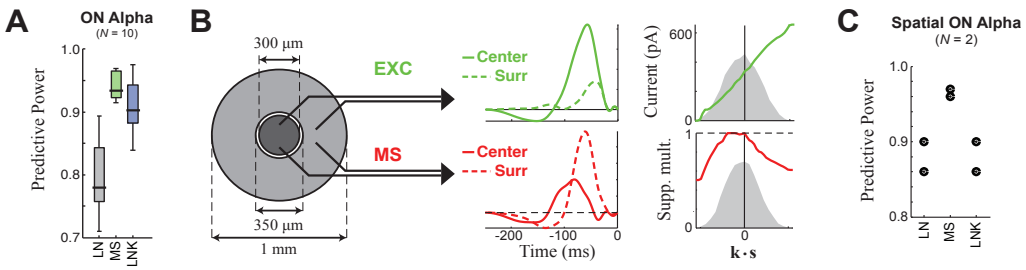

Figure 2-4 Probing mechanism with center-surround stimuli.
A. Cross-validated performance of the LN, the DS, and the Linear-Nonlinear-Kinetic model (LNK) applied during temporal only stimuli (note that DS models match form in Fig. 2-2). **B**. Separate modulation of center and surround (left) allows for the two different mechanistic explanations for divisive suppression to be distinguished. The spatiotemporal DS model finds different spatial footprints for excitation (green) and suppression (red), with excitation largely driven by the center (solid) and suppression by the surround (dashed), which is inconsistent with synaptic depression as an explanation for this suppression. **C**. Consistent with this, the performance of the spatiotemporal LNK model does no better than the LN model, and the DS model shows large improvement ($n = 2$).

Nonlinear mechanisms underlying spike outputs of ganglion cell

We next applied the nonlinear computation governing the inputs to the ganglion cell to a full model of the observed ganglion cell spike trains. Following previous likelihood-based models of ganglion cell spikes (Butts et al., 2011; McFarland et al., 2013; Paninski, 2004; Pillow et al., 2005) the output of a spike-refractoriness term was added to our previous model of the current, and the final result passed through a spiking nonlinearity (Fig. 2-5A), to yield the final predicted firing rate. Using this likelihood-based framework, all terms of the model – including the excitatory and suppressive LN models that comprised the current prediction – could be tractably fit using spike data alone.

The first stage of the resulting spike DS model fits closely resembled the DS model of the excitatory currents made from the same neurons (e.g., Fig. 2-2B,D). Suppression was consistently delayed relative to excitation (Fig. 2-5B), and exhibited both ON and OFF selectivity (Fig. 2-5C). The spike history term was mostly suppressive and had two distinct components, a strong "absolute refractory period" that last 1-2 ms and a second "relative refractory period" lasting more than 15 ms, similar in scale to that found by previous studies of neural refractoriness (Berry and Meister, 1998; Keat et al., 2001; Paninski, 2004; Pillow et al., 2005).

The resulting DS model successfully captured nearly all the predictable variance in the firing rate (Fig. 2-5G, predictive power = 95%), representing the best model of ganglion cell spike trains considered at millisecond resolution. In contrast, the

standard LN model has a predictive power of 64%, and modestly increases to 69% upon inclusion of a spike-history term. This suggests that the exceptional performance of the DS model is mostly due to the presence of the divisive suppression term.

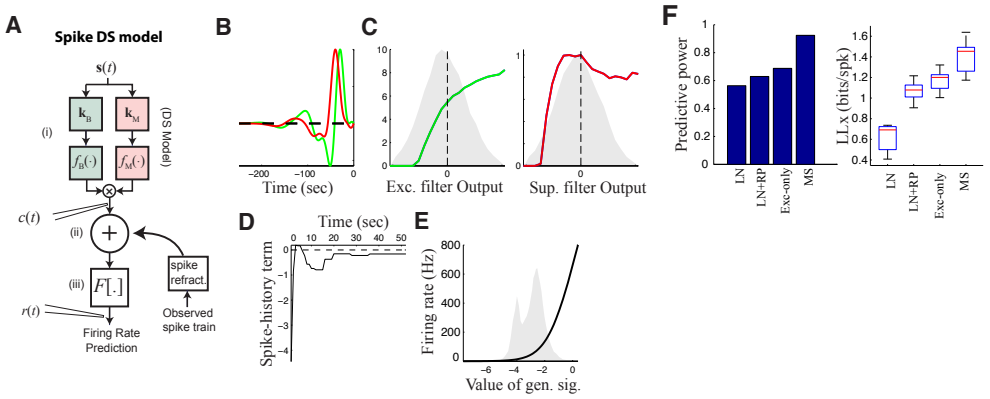

Figure 2-5 Divisive suppression model for spiking response.
A. Model schematics for divisive suppression model of spike trains. In addition to the DS model for the current data (i), the spike DS model further included spike history term and a rectifying spiking nonlinearity. (B-E) The model components for the same example neuron considered in Fig. 2-1-3. **B**. The excitatory and suppressive filters. **C**. The excitatory and suppressive nonlinearities. The filters and nonlinearities are similar to the DS model fit from current data (Fig. 2-2B). **D**. The spike history term, demonstrating an absolute and relative refractory period. **E**. The spiking nonlinearities, with shaded area indicating the distribution of generating signals. **F**. Left: predictive power of different models applied to the spike data. The DS model has much better performance than other models tested, including the LN model, the LN model with spike history term (LN+RP), and the excitation only model (Exc-only). Right: cross-validated likelihood (LL_x) of different models reflects model performance on a trial-by-trial basis. There is a big improvement in cross-validated likelihood (LL_x) with the inclusion of the spike history term, which is not reflected by the predictive power metric.

Precision of spike trains arises from complementary mechanisms of divisive suppression and spike refractoriness

Precision of spike trains depends on the interplay of multiple nonlinear mechanisms, including spike refractoriness and divisive suppression. To evaluate the contribution of the different model terms in capturing the detailed structure in the ganglion cell spike trains, we simulated spike trains using different combinations of the model components, and performed event analysis on both real and simulated spikes (Fig. 2-6A). We first tested what the effect of the spike history term was, by comparing models fit to the data with and without it. The Fano factor for each event, which is a measure of response variability, is much higher for models without a spike history term (Fig. 2-6B), and the inclusion of the spike history term also shapes the interspike interval (ISI) distribution to closely resemble the measured distribution (Fig. 2-6C).

We next compared models with and without the divisive suppression term to evaluate its effect. The spike history term alone (LN+RP) is clearly not sufficient to explain precision in the spike train. The LN model predicts firing rate varies at a slower time scale than observed even with the inclusion of the history term (Fig. 2-6A, red and green)(Butts et al., 2011). As a result, the LN model predicted duration of each response event is much longer than observed (Fig. 2-6B), and the precision of the first spike in each event is significantly worse than the real data (Fig. 2-6C). In contrast, the full DS model predicts sharp change in firing rate that matches the observed response through the interplay between excitation and suppression in the DS model,

which creates brief time windows for the neuron to respond (Fig. 2-6A). As a result, the DS model generated spikes resembles the real spikes more closely (Fig. 2-6A-C).

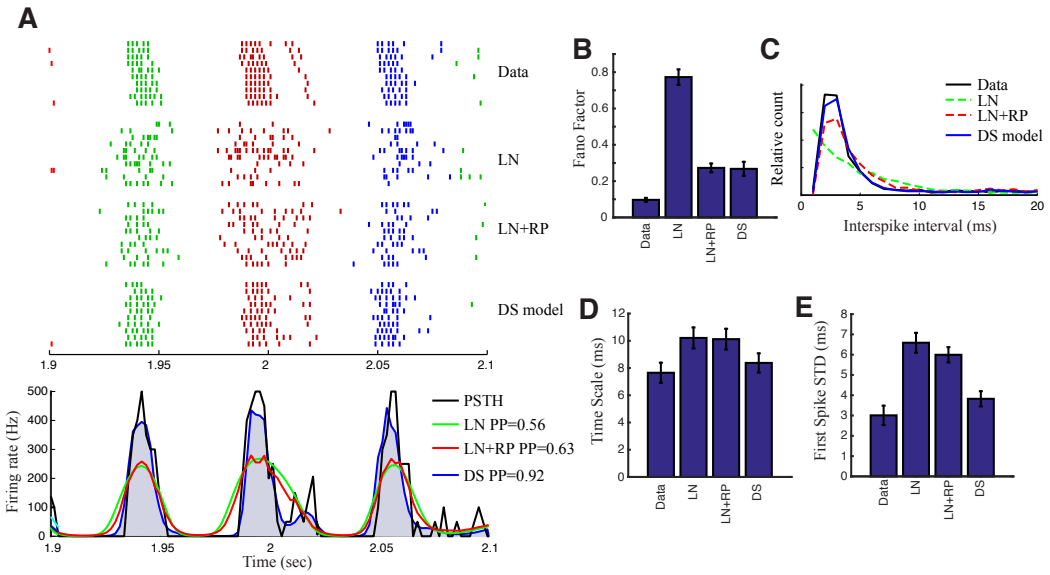

Figure 2-6 Spike patterning is shaped by a combination of nonlinear mechanisms.
A. Top: Spike rasters recorded over ten repeats for an example cell ("Data") compared with simulated spikes from three models: LN, LN model with spike history (refractory period) term (LN+RP), and the DS model with spike history ("Full"). Colors in the raster label separate spike "events" consistent across multiple trials in the data (see Methods). Bottom: The PSTH for each model, demonstrating that suppressive terms are important in shaping the overall envelope of the firing rate (models with suppression: gray shaded area). (**B-E**). Using event labels, spike statistics across repeats for the data and each model were compiled to gauge the impact of different model components. **B**. The Fano factor for each event is a measure of reliability, which is increased (Fano factor decreased) for models with a spike history. **C**. The spike history term has the clearest effect on spike patterns, as gauged by the interspike interval distribution: models with a spike history term (blue, red) most closely resemble the measured distribution (black). **D**: The event scale is the standard deviation of spikes in each event, with the median and 25-75% range across events shown by the boxes. **E**. First spike STD measures the precision of each event. The suppressive terms have the largest effect on the time scale of the event, whereas the refractory period has little effect.

Contrast adaptation is enhanced via spike refractoriness in ganglion cell output

In addition to accurate reproduction of precise spike outputs of ganglion cells, the spike DS model also captured the effects of contrast adaptation observed in the ganglion cell spike trains. In both contrast conditions, the simulated spike trains from the DS model are almost indistinguishable from the real data (Fig. 2-7A), which the DS model accomplishes with a single set of parameters fit to both contrasts. Like with the fits of excitatory current (Fig. 2-3), it outperforms LN models that are separated fit for each contrast level (Fig. 2-7B). The change of temporal precision across contrast is largely inherited from the synaptic current inputs and thus arises from the DS term of the model (Fig. 2-3C). Similar to the currents, the resulting spikes in the LC condition do not drive the suppression mechanism strongly and the DS model predicts the same time scales as the LN model (Fig. 2-7C).

However, contrast adaptation of spikes is different in that there is a much larger effect of contrast gain (Shapley and Victor, 1978), meaning the firing rate is proportionally higher in LC than expected from the reduction in stimulus contrast. The full spike DS model predicts the filter change across contrasts observed in the real data (Fig. 2-7D). We found that the contrast gain control effect is a result of the spike history term. as the gain control effect is largely diminished if we eliminated the spike history term from the DS model (Fig. 2-7D, *right*). This is because at HC, individual response event has shorter time scale and the impact of spike history term has much stronger effect on subsequent spikes within the event (Fig. 2-7E). Indeed, although the average number of spikes is comparable across the contrasts, the effect of spike history term

in LC is significantly smaller when the cell received excitatory current because spikes are dispersed over longer period (Fig. 2-7F).

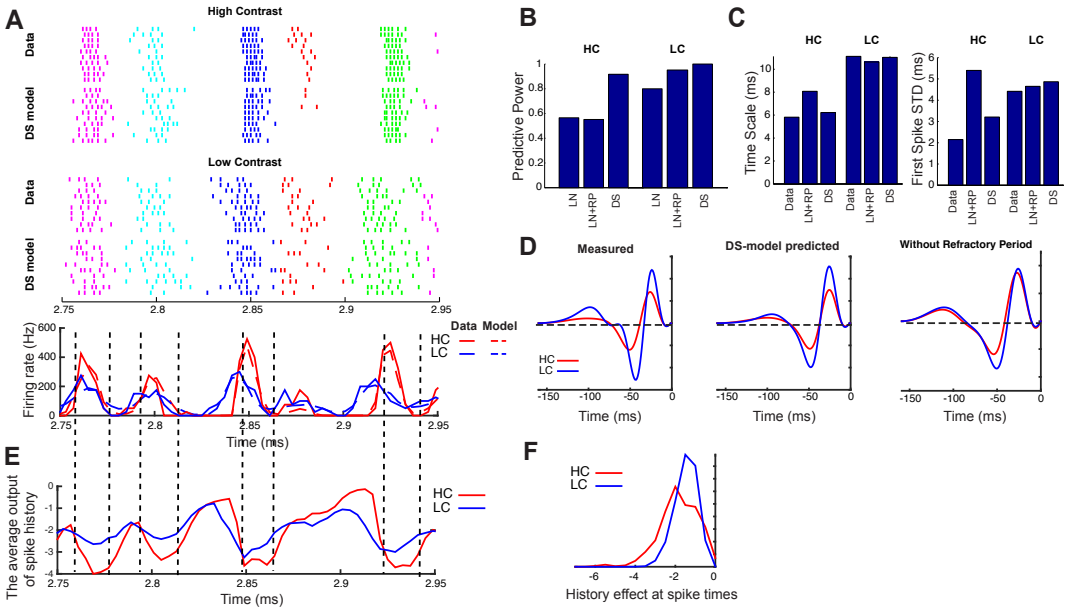

Figure 2-7 Contrast adaptation in the spike depends on both divisive suppression and spike refractoriness.

A. The full DS model accurately captured ganglion cell outputs across contrasts. Top: observed ("Data") and simulated ("DS model") spikes across at both high and low contrasts. Bottom: the model predicted firing rate (dashed line) agrees well with the PSTH of the neuron's response (solid line). **B**. The predictive power of models across contrasts. A single DS model is fit to both contrasts, and outperforms both the LN and LN+RP models that are fit separately to each contrast. **C**. The DS model captured temporal precision of the spike outputs, as measured by event time scale (left) and first spike STD (right), across both contrasts. **D**. The full DS model predicted change of temporal filter shape and magnitude across contrasts as observed in the data (left and middle). This prediction relies on spike refractoriness. There is little predicted change in temporal filter across contrasts if the spike history is removed from the DS model (right). **E**. Output of the spike history term during the same time window as in (A). Spike history term has stronger suppressive effects during firing events (dashed line). **F**. The suppressive effect from the spike history term is much stronger at high contrast condition (red) when the ganglion cell responds to the stimuli.

Discussion

In this study we derived a retina-circuit-inspired model for ganglion cells computation using recordings of both the synaptic inputs and spike outputs of the ON-Alpha ganglion cell, coupled with nonlinear statistical modeling whose structure was inspired by details of the retinal circuitry. The physiological data were used to fit model parameters and evaluate different hypotheses of how the retinal circuit processed visual stimuli. The result model explains both high temporal precision and contrast adaptation properties of their responses with unprecedented accuracy. We provide evidence that precise timing is already present in the excitatory synaptic current inputs, and can be explained by divisive suppression, which is likely due to presynaptic inhibition from amacrine cells. The interplay between nonlinear mechanisms, including divisive suppression, spike refractoriness and spiking nonlinearity, accurately captured detailed structures in both the synaptic current and spike response across contrast levels, and thus explained a variety of nonlinear properties of retinal processing.

Generation of temporal precision in the retina

One important nonlinear response property of early sensory neurons is high temporal precision. Temporal precision of spike responses has been observed in the retinal pathway with both noise stimuli (Berry et al., 1997; Reinagel and Reid, 2000) and natural movies (Butts et al., 2007). Presence of precise spike timing has been suggested to be evidence for temporal coding in the nervous system (Berry et al., 1997), or simply reflects that analog processing in the retina must be “oversampled”

in order to preserve information about the stimulus (Butts et al., 2007). Temporal precision also has been shown to play an important role in downstream processing of information provided by ganglion cells (Stanley et al., 2012; Usrey et al., 2000).

The generation of temporal precision involves nonlinear mechanisms within the retina; which may include both spike-refractoriness within ganglion cells (Berry and Meister, 1998; Keat et al., 2001; Pillow et al., 2005) and the interplay of excitation and inhibition (Baccus, 2007; Butts et al., 2011). Such distinct mechanisms contributing to ganglion cell computation are difficult to distinguish using recordings of the spike outputs alone, which naturally reflect the total effects of these various factors. By recording at two stages of the ganglion cell processing, we were thus able to demonstrate that high temporal precision already present in the synaptic current inputs in the HC condition, and temporal precision of both current inputs and spike outputs can be accurately explained by the divisive suppression model.

For models of both current inputs and spike outputs, divisive suppression is consistently delayed relative to excitation and exhibits selectivity to both ON and OFF stimuli. The same suppression mechanism is also likely underlying high temporal precision of LGN responses, which can be captured by a model with delayed suppression (Butts et al., 2011). Indeed, there is evidence that precision of LGN responses is inherited from the retina and enhanced across the retinogeniculate synapse (Carandini et al., 2007; Casti et al., 2008; Rathbun et al., 2010; Wang et al., 2010). Therefore, our results demonstrate that the temporal precision in the early

visual system likely originates from nonlinear processing in the inputs to retinal ganglion cells.

Our results shows that the contribution of spike history term to precision – as measured by the time scale of events and first-spike jitter – seems minor, consistent with an earlier study in the LGN (Butts et al., 2011). Nevertheless, spike history term does play an important role in spike patterning within the event (Pillow et al., 2005) and the resulting neuronal reliability (Berry and Meister, 1998). In fact, we could not fit the divisive suppression term robustly without the spike history term in place, suggesting that both nonlinear mechanisms are important to explain ganglion cell responses.

Contrast adaptation relies on both divisive suppression and spike refractoriness

The visual system continuously adjusts its sensitivity to efficiently encode the visual stimuli. It has long been known that retinal ganglion cell adapts its response to contrast, the range of intensity fluctuations around the mean. Ganglion cells respond with proportionately lower firing rates in a high contrast environment than in low contrast environment (Sakai et al., 1995; Shapley and Victor, 1978). Many experimental studies have shown that contrast adaptation in ganglion cells arises through multiple cellular mechanisms (Demb, 2008), including presynaptic inhibition of the bipolar cell terminal by amacrine cell (Zaghloul et al., 2007), synaptic depression in bipolar cell output (Jarsky et al., 2011; Manookin and Demb, 2006), and intrinsic cellular properties of ganglion cells (Kim and Rieke, 2003). Previous

models of contrast adaptation suggests different underlying explanations, such as feedback gain control (Shapley and Victor, 1978), spike refractoriness (Gaudry and Reinagel, 2007a, b), presynaptic inhibition (Zhang et al., 2015) and synaptic depression (Ozuysal and Baccus, 2012).

Here we modeled contrast adaptation at the level of synaptic currents and spikes from the same ganglion cell. We found there is contrast adaptation in the synaptic current inputs of ganglion cells, consistent with previous studies (Zaghloul, Boahen et al. 2005, Beaudoin, Borghuis et al. 2007), which could be explained by divisive suppression, which takes a mathematical form similar to previously proposed gain control models (Shapley and Victor 1979, Heeger 1992). Because the suppressive nonlinearity has very different shape than the excitatory nonlinearity, divisive suppression has much stronger effect in HC and result in a decrease of filter magnitude and a change in filter shape. Moreover, the same divisive suppression mechanism may also explain nonlinear spatial summation properties of ganglion cells (Shapley and Victor 1979) because suppression generally has broader spatial profiles than excitation.

We found that divisive suppression alone is not sufficient to explain contrast adaptation in the spike outputs. Instead, the effect of contrast adaptation is amplified in the spike outputs mostly due to spike refractoriness and changes of neural precision across contrast. At high contrast, neural response has higher precision and occurs within shorter event windows (Butts et al., 2010). As a result, the accumulated effect of spike refractoriness is stronger within each response events. Note that the effect of

the spike history term is highly dependent on the ability of the model to predict high temporal precision in the HC condition, which is largely originated from the divisive suppression term as discussed earlier. Therefore, the two nonlinear properties of retinal processing, contrast adaptation and temporal precision, are tightly related mechanistically and can be simultaneously explained by the DS model.

Circuits and mechanisms underlying the divisive suppression

Divisive suppression has been observed in many places in both peripheral and central nervous system, such as the invertebrate olfactory system (Olsen and Wilson, 2008), the lateral geniculate nucleus (Bonin et al., 2005), the primary visual cortex (Heeger, 1992), and higher visual areas like area MT (Carandini and Heeger, 2012; Simoncelli and Heeger, 1998). A number of biophysical and cellular mechanisms that could give rise to divisive suppression have been proposed in the literature, including shunting inhibition (Carandini et al., 1997; Hao et al., 2009), synaptic depression (Abbott et al., 1997), presynaptic inhibition (Olsen and Wilson, 2008) and fluctuation in membrane potential due to ongoing activity (Finn et al., 2007).

We evaluated different mechanistic explanations of the divisive suppression identified in this study. Divisive suppression underlying synaptic inputs to ganglion cells cannot be attributable to fluctuations in membrane potential or shunting inhibition since we recorded synaptic current responses under voltage-clamp conditions that minimize inhibitory inputs. Although synaptic depression could also explain fast transient responses and contrast adaptation (Ozuysal and Baccus, 2012),

it would predict excitation and suppression has the same spatial profile, which is not the case as we found suppression largely arise form the surround while excitation is mostly from the center of the receptive field. Therefore, the divisive suppression in our model is most likely due to presynaptic inhibition from amacrine cells, which is consistent with the retinal circuitry (Euler et al., 2014; Schubert et al., 2008; Wässle, 2004). Indeed, the response properties of the divisive suppression, including delayed response, larger spatial summation area and selectivity to ON-OFF stimuli matches properties of some classes of amacrine cells in the retina (Kaneko, 1973; Lin and Masland, 2006; Miller et al., 2006) and would be directly verifiable through pharmacological blockage of inhibitory synapses from amacrine cells (Beaudoin et al., 2007).

Notably, we specifically focused on one ganglion cell type, ON-Alpha cells, and these results might be specific to this cell type. Indeed, different manifestations of contrast adaptation (Kastner and Baccus, 2013) as well as underlying mechanisms, likely exist in different cell types. Our models were very consistent for ON-Alpha cells, but will show different nonlinear structure for other cell types (data not shown).

Chapter 3 : Diverse suppressive influences in area MT and selectivity to complex motion features

3.1 Introduction

Our ability to perceive natural scenes relies on efficient extraction of higher-order structure, which permits the decomposition of the visual scene into objects and surfaces. The extrastriate cortex of primates is devoted to such higher-order processing of visual signals (Maunsell and Newsome, 1987; Orban, 2008). Consequently, characterizing the relationship between extrastriate cortical activity and complex visual scenes can reveal a great deal about the neuronal computations underlying sensory processing. However, such stimuli often have very high dimensionality and necessarily involve complicated spatiotemporal correlations, making it difficult to isolate those potentially complex aspects of the stimulus that are driving neuronal responses.

One of the most thoroughly studied regions of the extrastriate cortex is the middle temporal (MT) area. MT is somewhat unusual among extrastriate regions in that the vast majority of its neurons are highly selective within a low-dimensional stimulus feature space. Specifically, MT neurons are selective for motion, which in natural vision occurs in the context of optic flow, comprising the spatiotemporal stimuli observed during translations and rotations of objects in the environment relative to the observer. However, most previous explanations of MT stimulus selectivity have focused on a subset of optic flow patterns that corresponds to the *two-dimensional*

velocities of objects translating within a single depth plane (Lisberger and Movshon, 1999; Nishimoto and Gallant, 2011; Perrone and Thiele, 2001; Simoncelli and Heeger, 1998). While this is a very important aspect of MT responses, such studies implicitly neglect the effect of different velocities at different positions within the visual field, which is an important component of optic flow.

There has been much evidence that MT is also selective for complex motion patterns, given the observation of the powerful suppressive surrounds (Allman et al., 1985; Born, 2000; Rust et al., 2006; Tsui and Pack, 2011; Xiao et al., 1997). Previous studies have isolated various aspects of suppression, including its contrast sensitivity (Pack et al., 2005; Hunter and Born, 2011), direction tuning (Allman et al., 1985), and spatial structure (Xiao et al., 1995, 1997), but these experiments have focused on a subset of properties of suppression, often using highly tailored stimuli. Thus, it has been left unclear how the multiple forms of suppression combine with excitation in more natural contexts – where stimuli driving each element are related due to the statistics of optic flow – in order to potentially result in higher-order selectivity.

Here, we use a continuously varying optic flow stimulus to measure the combination of excitation and multiple forms of suppressive tuning, using a nonlinear modeling approach that can be fit to the recorded neuronal spike trains. Our analysis reveal the full spatial and temporal structure of excitatory and suppressive influences for MT neurons, and demonstrates the diversity of computation in MT from this perspective. Using simulations, we show that these nonlinear properties of MT receptive fields are

functionally useful for extracting information about the *three-dimensional* velocities of moving objects, and hence directly facilitate further motion processing in higher cortical areas (Mineault et al., 2012; Zemel and Sejnowski, 1998).

3.2 Methods

3.2.1 Electrophysiology Recordings and Behavioral Task

Data were recorded from two adult rhesus macaque monkey (one female, one male, referred to as M1 and M2 hereafter), prepared using standard surgical techniques that have been described previously (Mineault et al., 2012). Animals were trained to fixate within 2° of a small fixation point on a computer monitor in return for a liquid reward. Eye movements were monitored at 500 Hz by an infrared eye tracker (EyeLink II; SR Research). Extracellular recordings were performed on 102 well-isolated single units in area MT, which was located using exterior cranial landmarks, anatomical images from magnetic resonance imaging (MRI), and/or physiological properties. Data were recorded using either single electrodes (n=63, M1; n=18, M2) or a multi-site linear electrode array (n=21, M1). Signals were amplified, bandpass filtered, sorted on-line and resorted off-line, using spike-sorting software (Plexon) to identify single units. All aspects of the experiments were approved by the Animal Care Committee of the Montreal Neurological Institute and were conducted in compliance with regulations established by the Canadian Council on Animal Care.

3.2.2 Visual Stimuli

Upon isolation of a single MT unit, we measured the neuron's direction tuning, speed tuning, and size tuning using random-dot motion stimuli. We then presented a continuously varying optic flow stimulus composed of moving dots whose velocity varied over space as well as time (Mineault et al., 2012). The velocity field was generated as a random combination of six optic flow components: horizontal/vertical translation, expansion, rotation and horizontal/vertical shears. The magnitude of each optic flow components varied independently based on low-pass filtered Gaussian noise with a cutoff of 2 Hz or 5 Hz. For the majority of cells (n=84), the stimulus was displayed in a slowly moving aperture with a diameter ranging from 8° to 20° (depending on size of the receptive field), the position of which was determined by another pair of low-pass filtered Gaussian noise with a cutoff of 0.05-0.10 Hz, with its mean at the center of the receptive field, and its standard deviation ranging between 3-10°, depending on the size of the receptive field. The other neurons (n =18, recorded from monkey M2) used the same optic flow stimulus, although it remained centered on the neurons' receptive field (location estimated from hand-mapping), and there was no moving aperture; instead, stimuli were displayed either on the full screen or in a very large static aperture with 30° diameter. The stimuli were presented on a LCD monitor (Dell 2707WFP) with a display resolution of 1,600×1,000 pixels (49°×36° of visual field at a distance of 50 cm) and a refresh frame rate of 60 Hz (n=73, M1) or 75 Hz (n=11, M1; n=18, M2). There were no notable differences for cells recorded with different refresh rate. In all cases, the dots were 0.1° in diameter

against a black background. The luminance of white dots was 194 cd/m^2 and the luminance of black background was 0.2 cd/m^2.

During the experiment, the stimulus was displayed in 6 min or 8 min blocks until the animal stopped behaving or the unit was lost. The stimulus presented in each block was different, and the data were thus combined to form a longer continuous stimulus with a median length of about 18 min per unit. The stimuli would revisit a given spatial position an average of 36 times. For a subset of recordings (n=20), we showed a repeated short segment of the stimulus (5 sec) that had its aperture centered on the cell's receptive field, in order to measure the response reliability and to calculate the explained variance (R-squared) of the model.

3.2.3 Data Analysis

The measured spike trains were binned at 25 ms temporal resolution to obtain the observed response rate $r_{obs}(t)$. We excluded data from 100 ms before fixation breaks (when the animal's gaze location deviated by more than 1.5° from the fixation point) to 500 ms after the recovery of fixation. To avoid any saccade-related effects or transients, only periods with fixations that were longer than 1 sec were used. The total recording of each neuron was broken into 10 sec segments, which were randomly divided into two groups: 80% of these segments were used to estimate model parameters ("training set"), and model performance was evaluated on the remaining 20% of the data ("cross-validation set"). The use of cross-validation ensures that the improved model performance is truly due to an increased ability to capture the

relationship between the stimulus and response, rather than a tendency to characterize random fluctuations (noise).

Of the 102 units recorded, we excluded 8 units for the following reasons. From monkey M1, two units were excluded because they did not have a consistent stimulus-dependent response, which meant that a stimulus-independent "null model" (that only predicted the average firing rate) outperformed all stimulus-dependent models tested. We also excluded six units recorded from monkey M2, where the measured receptive field did not align with the hand-mapped receptive field center (less than 30% of the excitatory weights were within 7.5° of the receptive field center). Although we did obtain significant model fits for these neurons, the spatial elements of the model were hard to interpret because the stimuli were not centered on the receptive field. The remaining 94 units were included for this study.

Finally, because units recorded from the same electrode often had very similar relationships between excitation and suppression, we only included a single neuron from each multi-site electrode experiment (n=5/21) in the figures relating to the overall distribution of MT neuron properties (Figs. 5F,G), in order to avoid sampling bias.

We also carefully controlled for the effects of fixational eye movements on the 53 recordings for which there were eye signals of sufficient quality, including both fixational drift and microsaccades (Martinez-Conde et al., 2013). The speed of

fixational drift was small during fixation (median = 1.515±0.090 °/sec) compared to the typical speed of the stimulus (~20 °/sec), which had an undetectably small effect on the neuronal response (data not shown). Microsaccades were detected using the algorithm proposed by (Engbert and Mergenthaler, 2006) and were observed to occur with a frequency of 1.77±0.10 Hz during the experiment. The impact of microsaccades on the neural response was direction dependent as previously reported (Bair and O'Keefe, 1998). However, we found this effect was largely unrelated to the stimulus-dependent terms of the model and had no impact on the results presented. As a result, we did not include these further analyses here.

3.2.4 Modeling of MT neurons

To understand how MT units respond to the complex motion stimuli, we developed a hierarchical modeling framework. We assumed that MT neuron responses are generated by an inhomogeneous Poisson process with an instantaneous rate $r(t)$. The log-likelihood of the model is then given by (up to an additive constant):

$$LL[r_{obs}(t), r(t)] = \sum_t [r_{obs}(t) \log r(t) - r(t)], \tag{3.1}$$

where $r_{obs}(t)$ is the measured neuronal response, and $r(t)$ is the model predicted firing rate (Paninski, 2004).

All models we consider use a fixed spiking nonlinearity $F[\cdot]$ that acts on the stimulus-dependent terms of the model, which we refer to as the generating signal $g(t)$,

$$r(t) = F[g(t) - b], \tag{3.2}$$

where b is the spiking threshold, and we choose the spiking nonlinearity function $F[\cdot]$ to be of the form $\log[1+\exp(\cdot)]$. This functional form resembles a familiar rectified-linear function, and additionally facilitates well behaved model optimization (McFarland et al., 2013; Paninski, 2004). To validate the use of this parametric form of $F[\cdot]$, we also measure the spiking nonlinearity using non-parametric histogram method (Chichilnisky 2001, Paninski 2004). In all cases, we find the chosen form of nonlinear function gives good description of the measured spiking nonlinearity. Other nonlinear functions, such as the power law transformation (Ghose and Bearl 2010, Nykamp and Ringach 2002) can fit the measured spiking nonlinearity equally well (data not shown), but have additional parameters, and are not as well behaved for parameter optimization (Paninski, 2004).

The model acts on the continuously varying optic flow stimulus, which is described by a local motion speed $\rho(t, x, y)$ and direction $\theta(t, x, y)$, sampled at a spatial resolution of 2°. This local motion signal is first processed by a set of subunits, which are described by a speed tuning function $f_v[\cdot]$ and a direction tuning function $f_\theta[\cdot]$. The subunit output is thus given by,

$$\text{subunit}(t, x, y) = f_v[\rho(t, x, y)]f_\theta[\theta(t, x, y)]. \quad (3.3)$$

In order to test the idea that V1 neurons are detectors of one-dimensional velocity, we also implemented an alternative formulation for the subunit, in which velocity was first projected onto the preferred direction of the subunit before being processed by the subunit nonlinearity,

$$\text{subunit}(t, x, y) = f[\rho(t, x, y)\cos(\theta(t, x, y) - \theta^{pref})]. \quad (3.4)$$

This formulation is consistent with the assumptions of various models of MT (e.g., Simoncelli and Heeger, 1998). However, because direction tuning and speed tuning are entangled in this formulation, the resulting models were more difficult to fit and interpret. Thus, because we did not find any significant difference in model prediction for these two formulations, we used the first subunit model (Eq. 3) for the majority of this work.

The generating signal of the model is computed by integrating the subunit outputs over space and time. In the motion-opponency model (MO model), this is given as,

$$g^{MO}(t) = \sum_{\tau} k_T(\tau)[\sum_{x,y} w(\mathrm{x},\mathrm{y})\text{subunit}(t, x, y)] \quad (3.5)$$

where k_T is the temporal kernel and *w* is the spatial weighting function. With a fixed speed tuning function, the temporal kernel and the spatial kernel can be efficiently estimated with the GLM framework (Paninski, 2004). The nonlinear speed tuning function can also be efficiently optimized by expressing it as a linear combination of basis functions $f_v(x) = \sum_n \alpha_n \xi_n(x)$, which were chosen to be overlapping tent-basis functions (Ahrens et al., 2008b). We chose the center of the tent-basis to be equally spaced on a logarithmic scale.

The nonlinear direction-tuning functions can be similarly optimized by expressing them using the tent basis functions. However, in practice we found it was more reliable to assume a parametric form for the direction tuning functions. To mimic the local motion opponency mechanism (Qian and Andersen, 1994), we used the von Mises function as the direction tuning function,

$$f_\theta(\phi) = \exp[b\cos(\phi - \phi_p)], \tag{3.6}$$

where ϕ_p denotes the preferred direction and *b* controls the direction tuning width. The cosine function implements local opponency, because a non-preferred stimuli will lead to suppressed output. These functions are always rectified (positive) because of the use of the exponential function. Speed tuning, direction tuning, spatial weights and temporal kernels can be alternatively optimized with appropriate choice of initial guesses.

For the excitation-suppression model (ES model), the stimulus processing is performed separately by excitatory and suppressive components, each using their own direction tuning functions and rectified spatial weighting functions. Here, the superscripts 'E' and 'S' denote the excitatory and the direction selective suppressive (DS-Sup) components, respectively. The generating signal is then given as:

$$
\begin{aligned}
g^{ES}(t) &= g^{E}(t) + g^{S}(t) \\
&= \sum_{\tau} {k_T}^{E}(\tau) \sum_{x,y} w^{E}(x,y) {f_{\mathrm{v}}}^{E}[\rho(\mathrm{t}-\tau,\mathrm{x},\mathrm{y})] {f_{\theta}}^{E}[\theta(\mathrm{t}-\tau,\mathrm{x},\mathrm{y})] \\
&\quad + \sum_{\tau} {k_T}^{S}(\tau) \sum_{x,y} w^{S}(x,y) {f_{\mathrm{v}}}^{S}[\rho(\mathrm{t}-\tau,\mathrm{x},\mathrm{y})] {f_{\theta}}^{S}[\theta(\mathrm{t}-\tau,\mathrm{x},\mathrm{y})],
\end{aligned}
\tag{3.7}
$$

where ${k_T}^{E}$ and ${k_T}^{S}$ are the temporal kernels, w^{E} and w^{S} the spatial weighting functions, ${f_v}^{E}[\cdot]$ and ${f_v}^{S}[\cdot]$ the speed tuning functions, ${f_\theta}^{E}[\cdot]$ and ${f_\theta}^{S}[\cdot]$ the direction tuning functions. For some models, an additional non-direction selective suppression (NS-Sup) is included, which takes the form of:

$$
g^{NS}(t) = \sum_{\tau} {k_T}^{NS}(\tau) \sum_{x,y} w^{NS}(x,y) {f_{\mathrm{v}}}^{NS}[\rho(\mathrm{t}-\tau,\mathrm{x},\mathrm{y})]. \tag{3.8}
$$

Because this component lacks a direction tuning function, its output is independent of stimulus direction.

This model is thus structured as a multilinear model of functions of the stimulus (Ahrens et al., 2008b), and each set of parameters can be efficiently optimized while holding the others constant. Because the overall multilinear optimization procedure can get stuck in local optima, it was important to choose an appropriate initialization of these parameters for the optimization procedure. The excitatory component was first optimized as the only model component, and then different types of suppressive components were added to the model to improve its performance. For the excitatory component, the temporal kernel was initialized to unity over the range of 50 ms to 150 ms, and was zero for all other time lags; based on the typical response latency of MT cells. The direction tuning parameter b (Eq. 6) was initialized to be 1, corresponding to a standard deviation of 42.6° for the direction tuning function (similar to measured tuning width using random dot patterns, Snowden et al., 1992: 46.5°), and the speed tuning function was initialized to be linear over a logarithmic scale. The direction preference was initialized to be one of 8 possible directions, equally spaced from 0° to 360°. Spatial weighting functions were optimized for each direction, and the best one was selected for further refinement. We then alternatively optimized the direction tuning parameters, the temporal kernel, the speed tuning functions, and the spatial weights.

Similar procedures were carried out when a DS-Sup component was added to the model: the suppressive direction that could most improve the model was selected, and then other model components were refined alternatively. For models with NS-Sup components, NS-Sup was simultaneously optimized with the other model

components. To determine which components to include in the model of a given neuron, four separate models were fit and compared for each cell: models with (1) only excitation; (2) excitatory and DS-Sup; (3) excitation and NS-Sup and (4) excitatory and both type of suppression. We selected the model with the best cross-validated performance.

Each component of the model has 225 parameters (15×15) for the spatial weighting function, 40 parameters for the temporal kernel, 10 parameters for the speed tuning function, and 2 parameters for the direction tuning function (preferred direction and tuning width). There is an additional parameter for the spiking nonlinearity for each model. The MO model thus contains 307 parameters, an NS-Sup component adds an additional 306 parameters, and an NS-Sup component adds an additional 304 parameters. A model with both DS-Sup and NS-Sup thus has 917 parameters. Note that such models are well constrained by the median stimulus duration of 18 mins, which corresponds to 43,200 different stimuli at a sampling rate of 25 ms. Note that the evaluation of model performance using a cross validation dataset avoids bias towards models with more parameters.

Due to the complexity of the model, regularization techniques were used to prevent over-fitting. A penalty term was added to the log-likelihood proportional to the second derivative or the Laplacian of the temporal kernel and the speed tuning function, e.g., $-\lambda \Sigma_t \partial^2 k_T / \partial \tau^2$. For the two-dimensional spatial weighting function, we used both "smoothness" and "sparseness" penalty. The smoothness penalty was

proportional to the sum of the squared slope relative to four nearest neighbors (up, down, left right), and the sparseness penalty was proportional to the sum of absolute value of the weighting function. The use of regularization techniques imposes certain prior distributions on model parameters and reduces the "effective" number of free parameters. For example, although we used 225 parameters to describe each spatial weighting function to make it flexible enough, the usage of smoothness and sparseness regularization ensure that only a small fraction of the parameters will be non-zero, and these non-zero weights usually vary smoothly across space.

Because the temporal kernels and the speed tuning functions usually had a stereotyped shape, we used the same regularization parameters for these functions across all cells. In contrast, the smoothness regularization parameters were adjusted for each cell and each model component individually using a nested cross-validation scheme. Specifically, 20% of the fitting data were randomly selected and reserved during the optimization of the spatial weights (note that this was different from the cross-validation data, which was never used during model estimation). The regularization parameters that gave best performance on the reserved data were used for the final fits to the data.

3.2.5 Measurement of the properties of suppressive components

Properties of the DS-Sup components were analyzed when detectable (n=62/94). The spatial profiles, direction tuning strengths and temporal dynamics of DS-Sup and NS-Sup components were also compared when both were detectable (n=33/94).

Following the fitting procedure described above, we investigated our assumptions about the direction tuning of each component by refitting the direction tuning curves using tent-basis functions (Ahrens et al., 2008b; Butts et al., 2011), which yields a non-parametric estimate of each direction tuning function. The direction tuning strength was then measured using circular variance (Ringach et al., 2002b), defined as Var=1-|R|, where R is given by

$$R = \frac{\sum_k f_\theta(\rho_k) e^{i\rho_k}}{\sum_k f_\theta(\rho_k)}, \tag{3.9}$$

where ρ_k are the centers of the tent basis, and angles are expressed in radians. The value of circular variances ranges from 0 to 1, with lower values indicating tighter clustering around a single mean. A circular variance close to 1 indicates no direction tuning.

Three indexes were calculated to describe spatial profiles of suppression. First, we calculated the center of mass of the excitatory weights and the suppressive weights. The weighted distance between suppression and excitation was then calculated as:

$$D_{ES} = \sum_{x,y} w_S(x,y) \sqrt{(x - {x_c}^E)^2 + (y - {y_c}^E)^2}, \tag{3.10}$$

where $({x_c}^E, {y_c}^E)$ is the center of excitation. This distance reflects how far suppression is from the receptive field center. A second index was calculated to reflect the dispersion of suppression:

$$Dispersion = \sum_{x,y} w_S(x,y)\sqrt{(x-{x_c}^S)^2+(y-{y_c}^S)^2} \tag{3.11}$$

where $({x_c}^S, {y_c}^S)$ is the center of suppression. Both indexes were separately calculated for DS-Sup and NS-Sup components. Finally, an overlap index was calculated as

$$Overlap = -\frac{\sum_{x,y} w_E(x,y)w_S(x,y)}{\|w_E\|\|w_S\|}, \tag{3.12}$$

where the minus sign is introduced because the suppressive weights $w_S(x,y)$ are negative.

The temporal kernel was specified at 25 ms resolution. To more accurately measure the latency of a given model component, we used cubic spline interpolation with five points around the peak of the temporal kernel, and used the peak of the cubic spline function as the component's latency.

3.2.6 Measurement of the selectivity of MT neurons for non-translational optic flow.

To gauge the selectivity of MT neurons to non-translational optic flow, we calculated the correlation coefficient ρ between each optic flow component and the neural response. For a given optic flow component, we computed the difference between the measured correlation coefficient and that predicted by the model, which reveals how well the models predicted the actual optic flow selectivity of the neuron.

We also compared the responses of the MO and ES models directly to optic flow stimuli, where each model was presented with a randomly selected combination of optic flow components centered on its receptive field, and the fraction of non-translational stimuli was systematically varied from 0 to 1, with 200 random stimulus samples drawn at each level. For each simulation, we calculated the correlation coefficient between model outputs, and also calculated the correlation coefficient between model response and the translational optic flow component.

3.2.7 Population Decoding Simulations

To simulate an MT population response to different stimuli, we "cloned" each model fit by replicating it across space (translating its position) and direction (rotating its spatial features and direction selectivity). For each of the 33 cells for which both types of suppression were detected, the MO model and the ES model fits were shifted on a 5×5 grid with spacing of 5° and rotated to include the anti-preferred direction and both perpendicular directions. Therefore, we generated 25×4=100 virtual models

for each cell. We pooled all virtual models together to create an entire virtual population of 3,300 cells.

We simulated the response of this population to velocity field induced by 3D motion in different directions. For these simulations, an object covering the central 20° of the visual field was simulated undergoing 3D motion in 200 directions sampled from a spherical distribution randomly. For the purposes of the simulation, we assumed the distance from the observer to the object was 5 m, and the speed of the object was uniformly distributed in the range of 0.87 m/sec to 2.62 m/sec, matching the speed range we explored with the optic flow stimuli (10 °/sec~30 °/sec). The spiking threshold of different model types was re-adjusted to give the same average firing rate (20 Hz) across all stimulus patterns.

We compared the capacity of an optimal linear decoder to extract information relevant to behavior, which were the physical parameters of the stimulus (3D velocity (v_x, v_y, v_z)), given the output of a population of model MT cells. Specifically, we computed a weight vector **w** to minimize the mean squared error between the parameter to decode (e.g. v_x) and the estimated parameter of the decoder **Xw**, where X is the matrix with one row for each stimulus and one column for each simulated MT response. At each repeat, we randomly selected 200 virtual models from the entire population or from a subset of the population and evaluated the reconstruction error as the ratio between the root mean squared error (rMSE) and the range of the

physical parameter. The procedure was repeated 20 times to estimate the error of our evaluation.

3.3 Results

To explore the selectivity of MT neurons to complex motion stimuli, we recorded from single units in area MT during the presentation of a continuous optic flow stimulus composed of a random dot field with the velocity field specified by a random combination of six optic flow dimensions (Fig. 3-1A) (Mineault et al., 2012). MT neurons generally responded very reliably to this stimulus, as demonstrated by the reproducible patterns of spikes in response to multiple repeats of the same stimulus sequence (Fig. 3-1B).

MT neurons are thought to primarily be selective for the direction of motion of stimuli in their receptive field (Albright, 1984; Maunsell and Van Essen, 1983; Mikami et al., 1986), as characterized by their average firing rate as a function of motion direction (Fig. 3-1C). Such "first-order" tuning is also reflected in more complex stimulus contexts such as during the continuous optic flow stimulus, which can be demonstrated by fitting a linear model to explain responses in this context (Fig. 3-1D) (Richert, 2008; Weber et al., 2010). Measurements of tuning in this stimulus context have the added advantage that the optic flow stimulus can reveal more complex aspects of MT tuning, because the stimulus incorporates different combinations of velocities across space and time. Indeed, as we show, more accurate models of stimulus processing are necessary to capture these details.

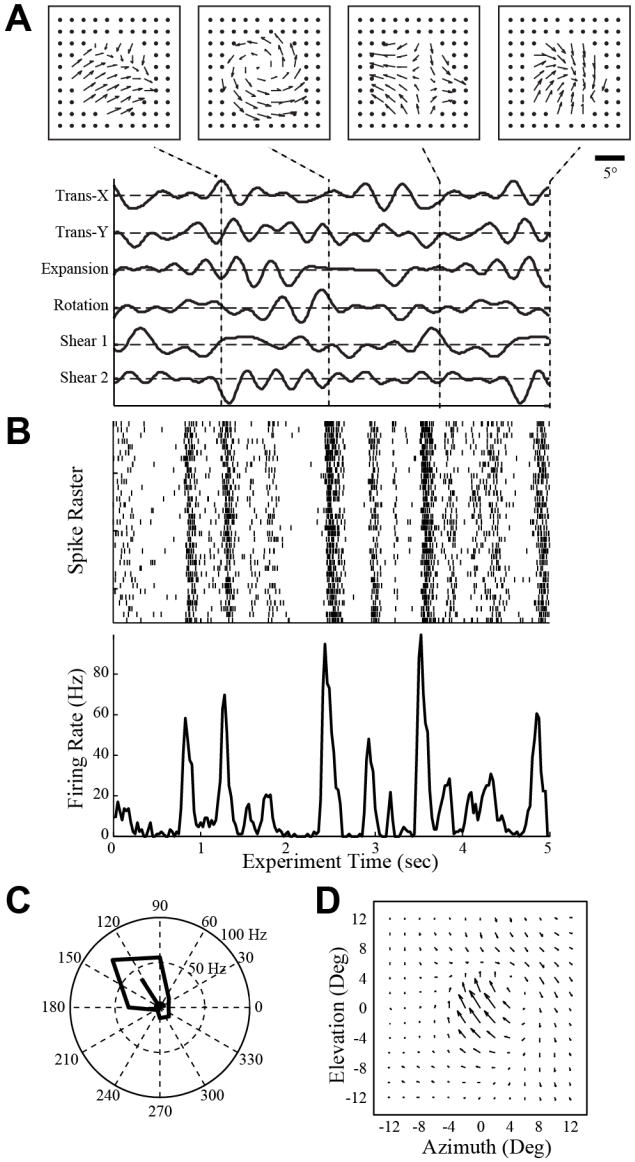

Figure 3-1 Response of MT neurons to continuous optic flow stimuli.
A. The naturalistic optic flow stimuli used in this study is composed of six independently varying optic flow components: horizontal and vertical translation (Trans-X and –Y), expansion rotation, and shear along both axes. Each optic flow component is independently specified by low-passed Gaussian noise (*bottom*), and displayed over a circular aperture is moving around slowly to explore the spatial profile of MT receptive field. The resulting velocity fields explore different types of flow, as shown by four example velocity fields that occur over the 5 second period shown. While stimulus is displayed at high spatial resolution, the models use the velocity field sampled over a uniform grid at a resolution of 2°. **B.** Response of an

example MT neuron to the optic flow stimulus shown. The repeated spike responses are represented in a raster plot (*top*), with each vertical bar indicating a spike, and the peristimulus time histogram (PSTH) is shown at the bottom. **C.** The direction tuning curve of this example MT neuron, measured as the average firing rate in response to random dot motion in the receptive field. **D.** Spatial kernel of the linear model fit to the continuous optic flow stimuli for the same MT neuron.

3.3.1 Hierarchical modeling framework for MT neurons

To interpret the neuronal response in this complex motion context, we adopt a hierarchical modeling framework (Fig. 3-2A). Receptive fields of MT cells are much larger than those at earlier stages of the visual hierarchy, and thus likely reflect aggregated responses of a large number of V1 neurons. We therefore choose an analysis resolution that gives fine detail for MT spatial receptive fields, by dividing them into smaller subunits, each presumably representing pooled responses of V1 neurons from that location. We assume that speed and direction are separately processed within each subunit (Hammond and Reck, 1980; Movshon et al., 1986; Rodman and Albright, 1987), such that subunit output is given by its direction tuning acting on the velocity (magnitude and direction) at each spatial location (Fig. 3-2B), multiplied by its speed tuning function (Fig. 3-2C).

The output of each subunit is then integrated across space with a spatial weighting function (Fig. 3-2D), and integrated over time with a temporal kernel (Fig. 3-2E). Our model assumes the same direction and speed processing at all relevant positions across space, in part because we found that relaxing these constraints does not improve model predictions (data not shown, although see Discussion). The final

signal after integrating over space is converted into a spike rate through a spiking nonlinearity (Fig. 3-2F).

Note that the structure of this model follows earlier models of MT responses (Qian et al., 1994), which assumes that the response of each subunit is enhanced by the preferred stimulus and suppressed by the non-preferred stimulus, followed by a rectified nonlinearity (Fig. 3-2B, *top*). This is often called "motion opponency", since the direction preference of suppression is always opposite from that of excitation (Adelson and Bergen, 1985; Simoncelli and Heeger, 1998); earlier studies showed that this opponent-direction suppression is local (Pack et al., 2006; Qian and Andersen, 1994). We thus refer to this model as the "motion opponency" (MO) model.

The parameters of this model structure can be tractably optimized using maximum likelihood estimation methods, in which parameters representing the stimulus selectivity of individual subunits, the spatial weights, and the temporal kernels can be fit to the observed neuronal response (see Methods). Alternative formulations of MT processing, such as ones in which V1 cells only respond to the velocity component projected onto their preferred velocities (Movshon et al., 1986), do not substantively change the model performance (Wilcoxon rank sum test on the cross-validated likelihood across the population between models of each type, $p>0.1$), but the resulting model is more difficult to fit and interpret because speed tuning and direction tuning become entangled.

In general the model fitted from data recorded in the context of complex optic flow stimuli has direction preferences that correlate with the neuron's preferred direction measured using the standard approaches (see Methods). (Fig. 3-2H, R=0.97).

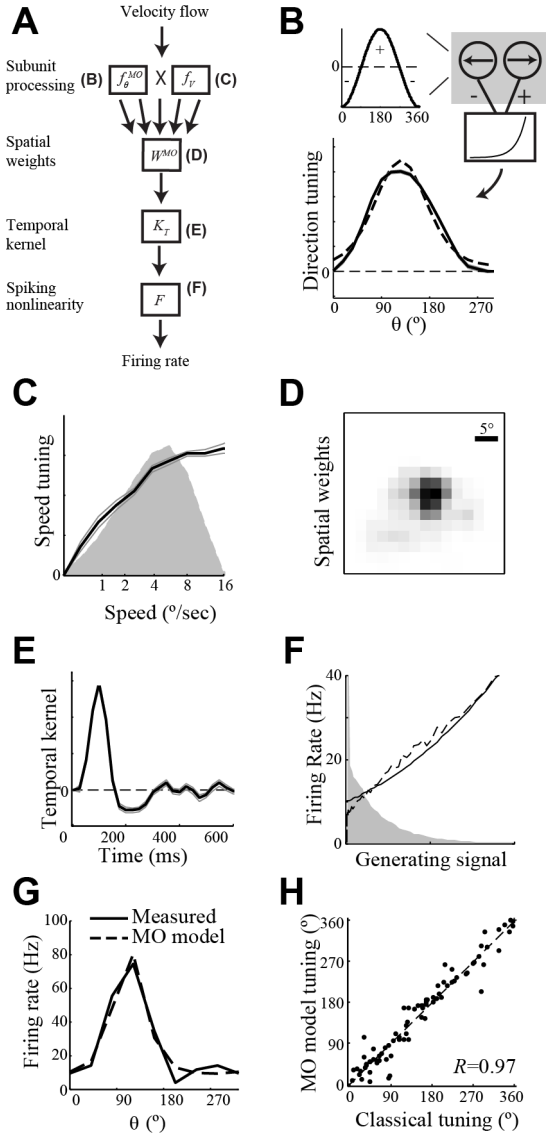

Figure 3-2 Analysis of MT neurons with the motion-opponency (MO) model.
A. MO model schematic: motion is first processed by local subunits with direction and speed tuning functions (*top*), and then is pooled across space (*2nd row*), with the resulting output integrated over time (*3rd row*). Finally, a spiking nonlinearity is applied on this signal to transform it into a firing rate prediction (*bottom*). **B-F**: The

model components for the example MT neuron with standard deviation of the fits (indicated as gray shaded area) calculated using bootstrapping techniques (100 repeated resampling with replacement). **B.** The direction tuning function for the subunits. The model is fit by assuming a von Mises function (dashed), which contains a motion opponency stage followed by a rectified nonlinearity (top). This can be validated by non-parametrically fitting the direction tuning directly (solid), which is in close agreement. **C.** The speed tuning function for the subunits (solid), compared with the distribution of stimulus speed (shaded gray). **D.** Spatial weighting function. **E.** Temporal kernel. **F.** The measured spiking nonlinearity (dashed) is fit parametrically (solid), with the distribution of the generating signal indicated in shaded gray. **G.** The model has nearly identical direction tuning curve (generated through simulation; dashed) as that of the neuron (solid). **H.** The model predicted direction preferences are highly correlated with the measured ones across the population of neurons in the study (R=0.97, $p<10^{-49}$, N = 84).

3.3.2 Extension of the Hierarchical Modeling framework with Suppression

While the MO model provides a good description of MT neurons' direction tuning (Fig. 3-2G), it only incorporates a single form of suppression, which is always spatially localized with excitation, and with opposite direction tuning. In contrast, numerous studies that probed stimulus selectivity with more than one spatially localized motion component have observed suppression outside the classical receptive field – in the "surround" – in a majority of MT neurons (Allman et al., 1985; Born, 2000; Tsui and Pack, 2011; Xiao et al., 1997). This suggests an extension of the MO model in order to include suppression that is not simply co-localized with excitation.

We thus extend the hierarchical framework to model a spatially distinct suppressive influence (Fig. 3-3A) in addition to the excitation described above, comprising what we call the "excitation-suppression model" (ES model). Motivated by detailed studies of suppression in V1, where suppression can have both an orientation-selective

component (Ringach et al., 2002) and a non-selective component (Sengpiel et al., 1997), we incorporate both forms of suppression into the ES model: a direction selective component (DS-Sup, *blue*) and a non-selective component (NS-Sup, *green*). The DS-Sup component has a direction tuning function, which is potentially distinct from the excitatory direction tuning, while the NS-Sup responds equally to all directions. The excitatory, DS-Sup, and NS-Sup components are pooled across space and integrated over time with their own spatial weighting functions and temporal kernels. The excitatory weights are constrained to be positive and the suppressive weights (both DS and NS) to be negative. As a result, the contribution of each is distinct, and simultaneous optimization of these components is tractable, resulting in a single optimal description of the excitatory and suppressive substructure of the MT receptive field of this form.

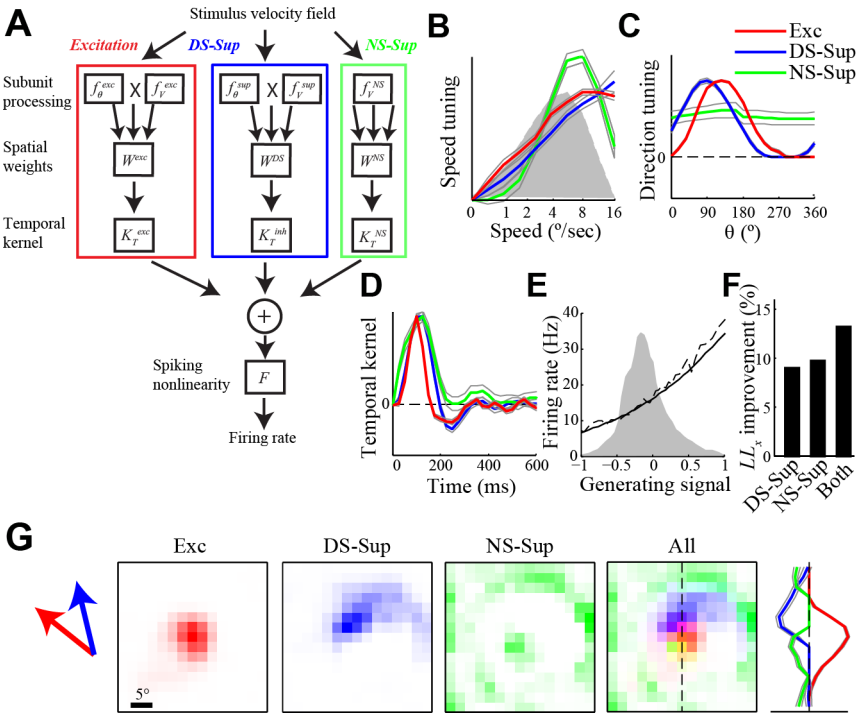

Figure 3-3 Incorporation of suppressive components into the description of MT processing.

A. Model schematic for the Excitation-Suppression (ES) model, which inputs excitation (Exc, red), as well as direction selective suppression (DS-Sup, blue) and non-selective suppression (NS-Sup, green). The DS-Sup component has the same computational structure as the excitatory component, but with the spatial weights constrained to be negative. The NS-Sup component is like DS-sup, but does not have a direction tuning function, and thus responds equally to all directions. **B.** Each component has its own speed tuning function, but they are all very similar in this case. Note gray lines around each curve show standard deviation estimated with bootstrapping techniques. **C.** Subsequent unconstrained fits of the direction tuning functions for each model component validates the forms used in fitting them, with Exc (red) and DS-Sup (blue) with very similar (antagonistic) tuning, and NS-Sup component tuning (green) not selective for direction. **D.** Temporal kernels of Exc (red), DS-Sup (blue) and NS-Sup (green), demonstrating a slight delay of suppressive components. **E.** The measured spiking nonlinearity (*dashed*) and the corresponding parametric fit (*solid*), relative to the distribution of the generating signal (shaded gray). **F.** The improvement of cross-validated likelihood of ES models with each (or both) suppressive components added over the MO model, expressed as a fraction of the performance of the model with both DS-Sup and NS-Sup. **G.** Spatial footprints of Exc (red), DS-Sup (blue) and NS-Sup (green). The arrows on the left indicate direction preference of excitation and DS-Sup. A vertical slice of the weighting function is shown on the right, with the gray lines indicating standard error of each function over multiple model fits.

Across the population of MT neurons, we find that excitation, DS-Sup and NS-Sup typically have different spatial profiles and direction preferences, and in some cases, different temporal dynamics. For example, the previously considered neuron (Fig. 3-2) has a DS-Sup component with similar direction preference as excitation (Fig. 3-3C), but forms an asymmetric surround structure that is largely non-overlapping with the excitatory region (Fig. 3-3F). Furthermore, there is also a distinct NS-Sup component, which has spatial weights in two areas: a central area that is largely overlapping with the excitatory component, and a surround component that is further away from the receptive field center than the DS-Sup. The temporal kernels of both suppressive components are slightly delayed relative to excitation (Fig. 3-3D). Although the ES model does not predict a different excitatory tuning of the neuron, it has a much better cross-validated performance than the MO model (Fig. 3-3E).

To gain additional insight into the degree of model performance improvement, for a subset of the recorded neurons, we also presented repeats of a short segment of the stimulus (e.g., Fig. 3-1B), allowing us to evaluate model performance using both cross-validated log-likelihood (LL_x) and more traditional peristimulus histogram (PSTH)-based methods such as explained variance (R^2). This analysis reveals that the models can explain 34.5±3.1% of the variance of the response, comparable with the performance of other models of MT processing (Nishmoto et al. 2011). Furthermore, the improvement of model performance after inclusion of suppression is significant using both metrics, and corresponds to a median of 25.3% of explained variance and 26.2% of LL_x (Figs. 4A,B). Moreover, there is a correlation between log-likelihood

and R^2 (Fig. 3-4C), indicating a consistency of the two metrics. Because it provides a much more reliable metric of model performance in this case and does not require repeated stimulus presentations, we will only use cross-validated log-likelihood to measure accuracy of model predictions in the rest of the study.

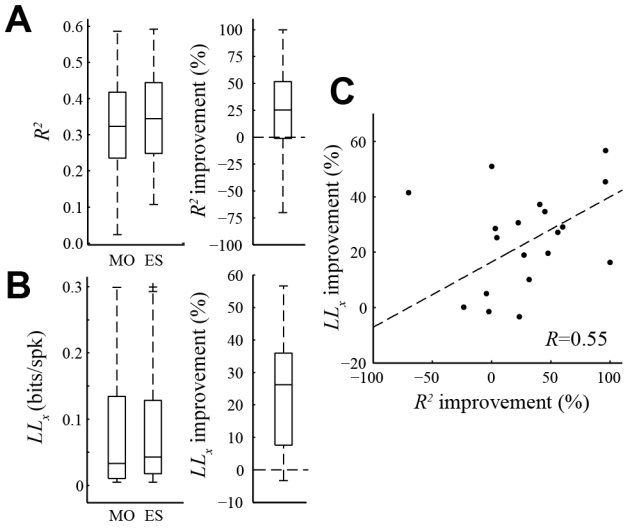

Figure 3-4 Model prediction accuracy.
A. Fraction of explained variance (R^2) for the MO and ES models (left). Paired comparisons (*right*) demonstrate significant improvement: 22.0%±11.7% (p<0.05, Wilcoxon signed rank test). **B.** Cross-validated likelihood (LL_x) for the MO and ES models, showing a similar trend as (A). Percentage improvement is 21.7%±5.0% (p<0.05, Wilcoxon signed rank test). **C.** A comparison between the two metrics shows the correlation between them (R=0.55, p<0.05).

3.3.3 Properties of suppressions in MT

Across the population of cells we recorded from, the addition of one or both types of suppression improves model prediction for most of the neurons (81/94). For one third of the neurons, the best model has both DS-Sup and NS-Sup (33/94, LL_x improvement=14.6±3.2%). For the rest of the neurons in the population, the best

model either only has DS-Sup (29/94, LL_x improvement=11.9±5.3%), or only has NS-Sup (19/94, LL_x improvement=15.8±9.1%). Because a model with more components requires more data to fit and is more susceptible to overfitting, the fraction of cells that have both suppressive components might be underestimated using our criteria of requiring better cross-validated likelihood. For the neurons with both DS-Sup and NS-Sup components, we verified that there was a significant improvement in model performance over models with only one suppressive component (Fig. 3-5A), suggesting that DS-Sup and NS-Sup are describing different aspects of neuronal selectivity.

The two forms of suppression thus appear to be distinct contributions to the computation performed by the neuron. To explore this further, we first tested whether we correctly assumed the forms of DS-Sup and NS-sup tuning by relaxing constraints on the tuning functions and fitting these functions using non-parametric methods directly (see Methods). The resulting direction tuning of the NS-Sup components are indeed flat in almost every case, reflected as a circular variance very close to unity (Fig. 3-5B). In contrast, the direction tuning of DS-Sup fit in the same way has a circular variance comparable with that of the excitatory component (Fig. 3-5B).

The two components also typically have different spatial profiles (*e.g.*, Fig. 3-3F), with the spatial weights of NS-Sup usually located further from the excitatory center (Fig. 3-5D) and with more dispersion than that of the DS-Sup (Fig. 3-5E). Both DS-Sup and NS-Sup temporal integration is significantly delayed relative to the

excitatory components (relative latency of DS-Sup=12.6±4.0 ms, relative latency of NS-Sup=25.5±6.9 ms). The latency of DS-Sup is consistent with previous observations of MT surrounds (Perge et al., 2005),

In summary, the NS-Sup shares many characteristics of the previously described MT surround (Born and Bradley, 2005; Hunter and Born, 2011), which is further from the center, covers a larger area, and is very broadly tuned and delayed. In contrast, DS-Sup is more likely to influence stimulus selectivity directly within or close to the receptive field. The different direction selectivity and spatial profiles of DS-Sup and excitation, however, can lead to local tuning differences in subregions of the RF, seen for example in (Richert, 2008). In the rest of the paper, we will thus focus on the properties of MT neurons imparted by DS-Sup.

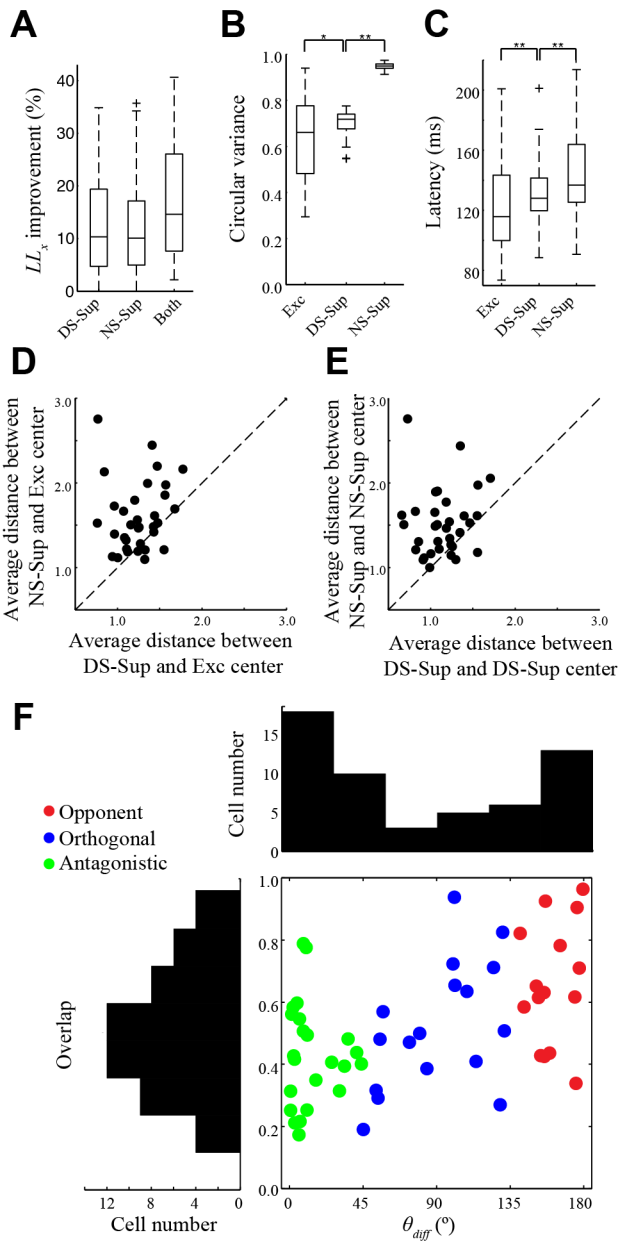

Figure 3-5 Properties of the suppressive components.
A. Box plots showing the improvement of cross-validated likelihood (LL_x) over the MO model, applied to MT neurons with detectable DS-Sup and NS-Sup components (n=33). Including only DS-Sup on average increases the LL_x by 12.4% (*left*); the improvement is 12.1% for a model with only NS-Sup (*middle*), and 20.7% for a model with both DS-Sup and NS-Sup (*right*). **B.** Circular variance, which measures direction-tuning width, of the direction-tuning curve for excitation (*left*), DS-Sup (*middle*) and NS-Sup (r*ight*). DS-Sup has comparable tuning variance as excitation, while NS-Sup is almost completely non-selective to motion direction (circular variance close to 1). **C.** Both types of suppression are delayed relative to excitation, with latency relative to excitation for DS-Sup (l*eft,* 12.6±4.0 ms, p<0.001) and for NS-Sup (*right,* 25.5±6.9 ms, p<0.001). NS-Sup is further delayed relative to DS-Sup

(p<0.05) **D.** Average distance from the receptive field center for DS-Sup (x-axis) and NS-Sup (y-axis), in units of receptive field size. NS-Sup is further away from the center than DS-Sup (p<0.001). **E.** NS-Sup is more dispersed than DS-Sup, as demonstrated by measuring the distance from the centroid of DS-Sup (x-axis) and NS-Sup (y-axis) (p<0.001). **F.** The population of MT neurons demonstrates a diversity of relationship between excitation and suppression, as shown by plotting each neuron's direction difference between excitation and DS-Sup (horizontal) and amount of spatial overlap between them (vertical) (n=55). Each dot shows an individual neuron; while the distribution is continuous, we draw color-coded distinctions in order to analyze different regimes of tuning (green: antagonistic suppression, blue: orthogonal suppression, red: overlapping opponent suppression). Marginal distributions over direction difference and overlapping extent are shown at the bottom and the left respectively.

3.3.4 Diversity of MT selectivity reflected in direction-selective suppression

The combination of their excitatory and suppressive influences can potentially give MT neurons selectivity to complex motion. To understand such selectivity, we first characterize the relationship between excitation and DS-Sup with two parameters: the differences of direction preferences and the overlap extent of spatial weights (Fig. 3-5F). Note that here we only consider cells with DS-Sup detected, and only include a single neuron from each multi-site electrode experiment (see Method), resulting in a population size of 55. It is most common for MT neurons to have suppression and excitation with matching tuning (antagonistic suppression, *green*), as well as with opposite preference (opponent suppression, *red*). Nevertheless, this distribution is continuous – with no apparent holes – and there are also many cells with suppression orthogonal to excitation (orthogonal suppression, *blue*). However, the distribution also shows a certain degree of diversity, especially for cells with orthogonal suppression. In order to highlight examples from different parts of this distribution and further investigate the diversity of neuron computation pictured, we divide cells

into three groups based on their position within this *two-dimensional* distribution (Fig. 3-5F). Example model fits for cells in each group are shown in (Fig. 3-6).

Cells with antagonistic suppression represent the majority of our population (23/55). Spatially, DS-Sup has an asymmetric surround structure around excitation (Fig. 3-6A). Note that the lack of observed neurons in the upper-left corner of Fig. 3-5F, corresponding to overlapping antagonistic suppression, is likely due to the model's inability to detect overlapping suppression that has similar selectivity as excitation, because that adding excitation and suppression with identical tuning at the same position will have no effect on the model output in the stimulus contexts we studied. In contrast, other studies have observed overlapping, antagonistic suppression using tailored stimuli (DeAngelis et al., 1992; Cavanaugh et al., 2002). In most cases, NS-Sup can also be detected at a further distance from the center, suggesting that MT surrounds can be decomposed into a direction-selective and a more distant non-selective component. In principle, both suppressive components could contribute to classically measured "size tuning" (Allman et al., 1985). However, we find that size tuning is not limited to cells with antagonistic suppression, and only the strength of NS-Sup is significantly correlated with the extent of size tuning (R=0.50, p<0.05). This is also consistent with earlier reports of broader direction tuning of the surround than that of the center (Born and Bradley, 2005; Hunter and Born, 2011). On the other hand, the spatial footprint of DS-Sup resembles the asymmetric antagonistic surround seen in other studies (Orban, 2008; Raiguel et al., 1995; Xiao et al., 1995), and may contribute to selectivity for complex motion features such as speed gradients and

surface orientation (Gautama and Van Hulle, 2001; Sanada et al., 2012; Xiao et al., 1997).

Orthogonal suppression is found in a proportion of MT neurons (17/55) (Fig. 3-6B). This type of suppression is not as well documented as the other type of suppressions, presumably because previous studies of MT surround usually restrict motion in the same or the opposite direction as the preferred direction. The spatial footprint of orthogonal DS-Sup is generally very different than excitation, and exhibits a large degree of diversity across neurons with orthogonal suppression. Usually this suppression is confined to one side of the receptive field. In addition to image discontinuities, this arrangement may provide selectivity to curvature in the motion field, which is an important aspect of natural motion.

Finally, cells with opponent suppression comprise the rest of MT neurons in the population where we detected DS-Sup (15/55). DS-Sup is typically co-localized with excitation for these cells (Fig. 3-6C), resembling the organization of the motion opponent subunit we used in the MO model (Qian et al., 1994) (Fig. 3-2A). However, the ES model still represents a different computation from the MO model due to the rectified nonlinearity in the subunit. As a result, the subunits will not give any response for non-preferred stimuli, and the MT neuron will respond at its spontaneous level, whereas for models with opponent suppression, the neuronal response will be truly suppressed below the baseline for non-preferred stimuli. Indeed, we can observe this in the measured direction-tuning curve, where we calculated the amount of

suppression relative to the baseline firing rate in the non-preferred direction. We find that the amount of suppression below baseline is significantly larger for cells with opponent suppression compared the rest of the cells in the population ($p<0.05$, Mann-Whitney U test).

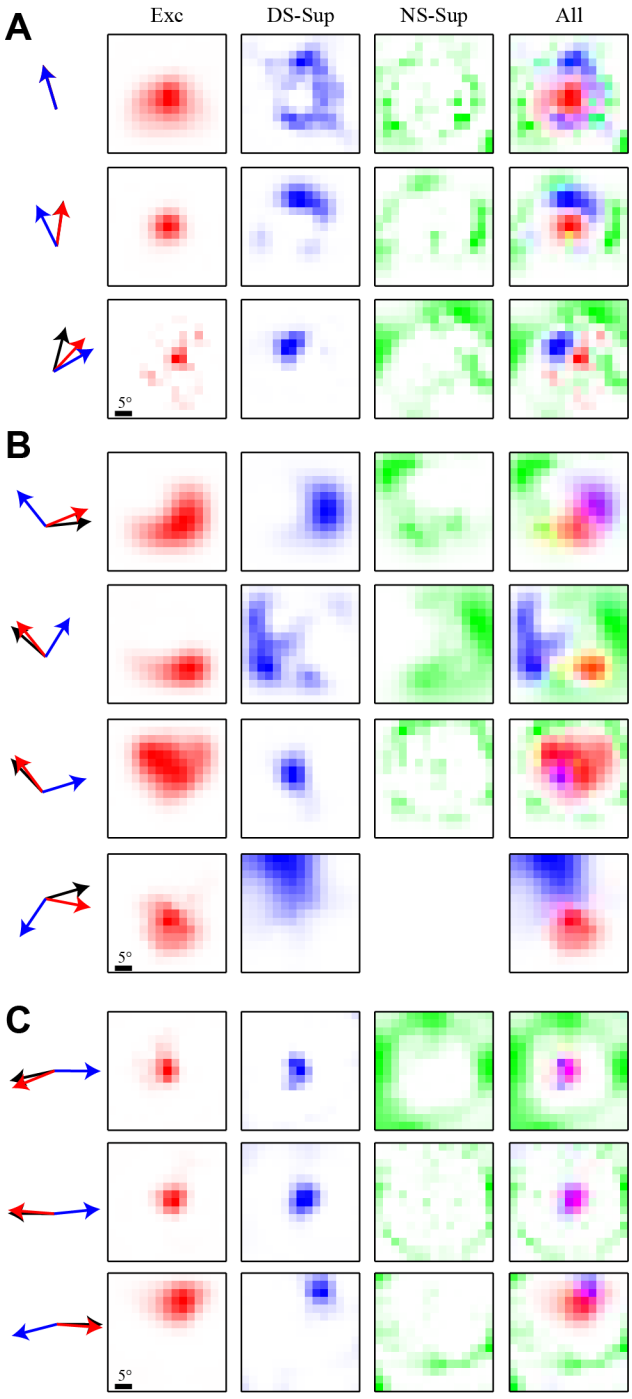

Figure 3-6 Example model fits for different types of tuning.
A-C: Example fits are shown for different types of tuning: (**A**) antagonistic suppression, (**B**) orthogonal suppression, and (**C**) overlapping opponent suppression. Each row shows (from *left* to *right*) direction preferences of excitation (red), DS-suppression (blue) and the measured preferred direction (black), spatial weights for excitation, DS-Sup, NS-Sup and overlapping of them.

3.3.4 Selectivity and coding of complex optic flow

The diversity of both excitatory and suppressive tuning direction and their spatial footprints suggest that MT neurons might be specifically tuned to velocity fields with different directions of motion at different spatial locations. Such visual stimuli are quite common in more natural settings where velocity fields are due to motion of the observer and with objects at different depths. A simple way to reveal optic flow selectivity is to ask to what extent the neuron's firing rate is modulated by a given flow component, which we gauge using the correlation coefficient ρ between each of the six optic flow components and 1) the actual neural response; 2) the predicted response of the MO model 3) the predicted responses of the ES model (i.e., with both types of suppression). Because MT responses are significantly correlated with non-translational flow components in many cases, the models with suppressive components predicted their correlation with each optic flow component better than the MO model (Fig. 3-7A). Therefore, optic flow selectivity in MT relies at least in part on the suppressive influences we characterized before.

We can also gauge the effect of adding suppression by comparing the responses of MO and ES models fit to the same neuron, as different amounts of non-translational optic flow are presented. First, for stimuli that consist of only translation, the outputs of the MO and ES models are highly correlated (R=0.91±0.06), which is expected because both models can predict the standard direction tuning properties of the cell. However, the correlation between the two models decreases as more non-translational optic flow is introduced into the stimulus (Fig. 3-7B). When the stimulus is composed

of only non-translational components, the average correlation coefficient is only 0.67, with a standard error of 0.06. This implies that, as expected, direction-selective suppression results in different responses to more complex motion stimuli. Such selectivity to complex optic flow appears to arise from the DS-Sup component rather than the NS-sup, because models with excitation and DS-Sup alone (Fig. 3-7B, blue) are similar to the full ES models, while models with only excitation and NS-Sup (red) are generally more correlated to the MO models.

Next we examine which components of the stimulus are most correlated with the predicted response for models with and without suppression. In general, the MO model outputs are much more correlated with the translational components than the ES models (Fig. 3-7C). The difference is most significant when a moderate amount of complex optic flow is introduced (10-40%). Consistent with Fig. 3-7B, we find DS-Sup alone is enough to explain this difference, while models with only NS-Sup mostly resemble behaviors of the MO models. This suggests that the difference between model predictions is due to the fact that MO models are only selective to translational motion, while the ES models exhibit additional selectivity to more complex optic flow stimuli. When more complex optic flow is introduced, both types of models are driven by the non-translational flow components, but the responses of the two models generally depend on the stimulus in different ways, as reflected by the decreased correlation in (Fig. 3-7B).

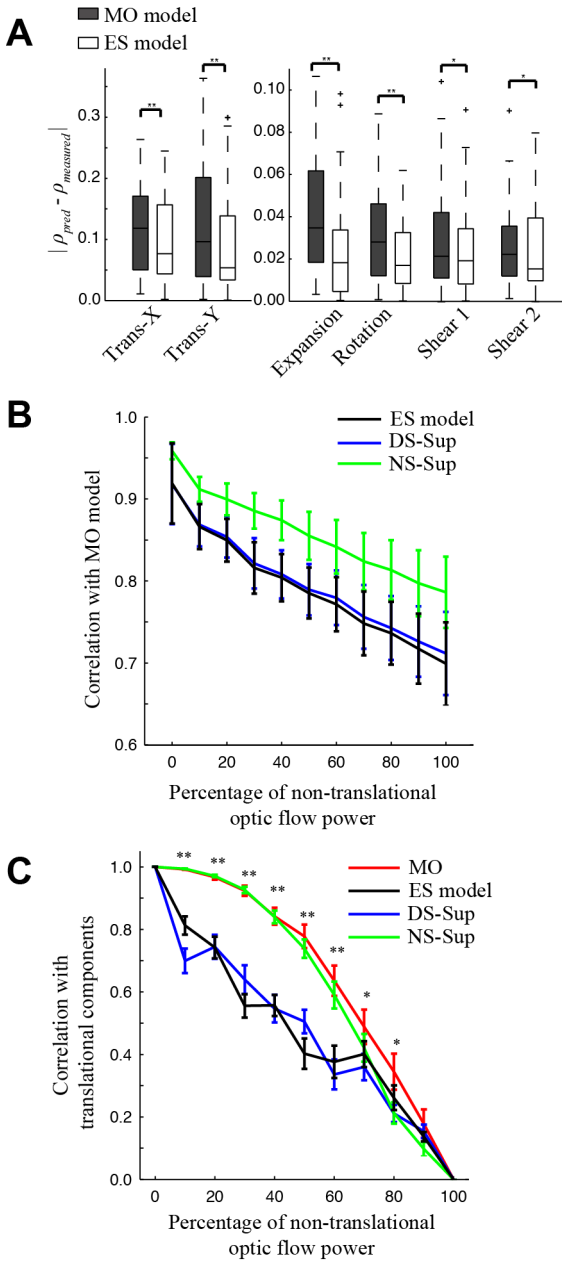

Figure 3-7 Suppression enhances selectivity to complex optic flow.
A. To gauge the ability of the MO and ES models to capture selectivity of MT neurons to different components of optic flow, we calculate the correlation coefficient ρ between each flow component and the neuron response, and compare that predicted by each model across the population of neurons with measured suppression (n=33). There are significant improvements for all six optic flow components with incorporation of suppression into the model (** p<0.001, * p<0.05 Wilcoxon signed rank test). **B.** As a second way to gauge this selectivity, we simulate the response of MO and ES models of the same neuron to different combinations of optic flow, and measure how correlated the responses are, as a function of the amount of non-translational optic flow present in the stimulus. Each data point shows the average

correlation coefficient between the responses of the MO and ES models over the 33 MT neurons that have both types of suppression. Responses of the two models, with both components included (black) are highly correlated for translational stimuli (R=0.92±0.05), but decrease progressively for stimuli that contain more non-translational optic flow components (black). This trend is highly significant (Spearman's rank correlation coefficient ρ=-0.38, $p<10^{-11}$). DS-Sup appears to contribute to this selectivity to non-translational optic flow, as demonstrated by including only the DS-Sup term with excitation (blue) or NS-Sup term with excitation (green). Models with only NS-Sup are more correlated with the MO model in general ($p<0.05$ when percentage of non-translational optic flow is not zero, *t-test*), while models with only DS-Sup are similar to the full ES model (Spearman's rank correlation coefficient ρ=-0.36, $p<10^{-11}$). **C.** To measure to what extent model output is determined by translational component of the stimulus, we also report the correlation coefficients between model outputs and the translational components of the optic flow stimulus in the same context as (B), as a function percentage of non-translational optic flow. Significant differences between MO model (red) and ES model (black) are revealed when the stimuli contains moderate amount of complex optic flow (** $p<0.001$ * $p<0.05$, t-test). This difference is also observed for models with only DS-Sup (blue), but is absent for models with only NS-Sup (green).

Thus, the suppressive components of the ES model contribute to selectivity to non-translation components of optic flow stimuli. Does such selectivity allow for a better representation of such stimuli by MT neurons? Previous work has shown that individual MT neurons exhibit fairly weak tuning for optic flow patterns such as expansion and rotation, and that such tuning is highly dependent on stimulus position (Lagae et al., 1994). Here we ask whether a population of MT neurons could encode three-dimensional motion patterns and to what extent this encoding depended on the presence of nonlinear surround suppression.

To address these points we applied an optimal linear decoding framework (DiCarlo and Cox, 2007) to recover the velocity of a simulated object moving with different 3D velocities at a specific position relative to the observer (Fig. 3-8A). This task is also related to decoding egomotion, where one typically assumes the visual scene is

static and the observer is moving. In this case the "object" is the entire visual for egomotion decoding, as we are not examining the more complex case of recovering egomotion during simultaneous rotations of the eyes, head, or body. The results we report here are not sensitive to the size of the object.

For each simulation we calculate the velocity pattern generated by 3D object motion in different directions (Fig. 3-8A-B), and then simulate a population of MT neuron responses to these stimuli (see Methods). An optimal linear decoder is trained to infer the 3D motion direction based on the simulated neuronal response either using the MO models, or the ES models. The reconstruction performance of the decoder reflects a lower bound on the information that would be available for a downstream brain area (DiCarlo and Cox, 2007; Mineault et al., 2012).

The results of this simulation show that decoding performance using the ES model responses is significantly greater (i.e., less decoding error; Fig. 3-8C) compared to that using the MO model, while further incorporation of NS-Sup does not further improve the performance. Moreover, the decoding performance depends on properties of suppression: although significant improvement is observed if we only use cells with antagonistic suppression or orthogonal suppression for the decoding task, such difference is not observed for cells with opponent suppression (Fig. 3-8D). This is likely because spatial offsets between excitatory and suppressive components may be instrumental in creating non-translational optic flow sensitivity, but opponent suppression tends to have high degree of overlap with excitation and thus little spatial

offset (Fig. 3-5F). Note that the performance of both models significantly degrades when stimuli are not presented in the center of the receptive field, in marked contrast to the same decoding task applied in area MST (Mineault et al., 2012), suggesting that computations described in MT represent one element of what is likely a hierarchical computation. That is, the initial selectivity developed in MT is further refined and generalized across spatial positions in higher-level areas such as MST, as has been suggested for analogous computations in other areas (Riesenhuber and Poggio, 2000).

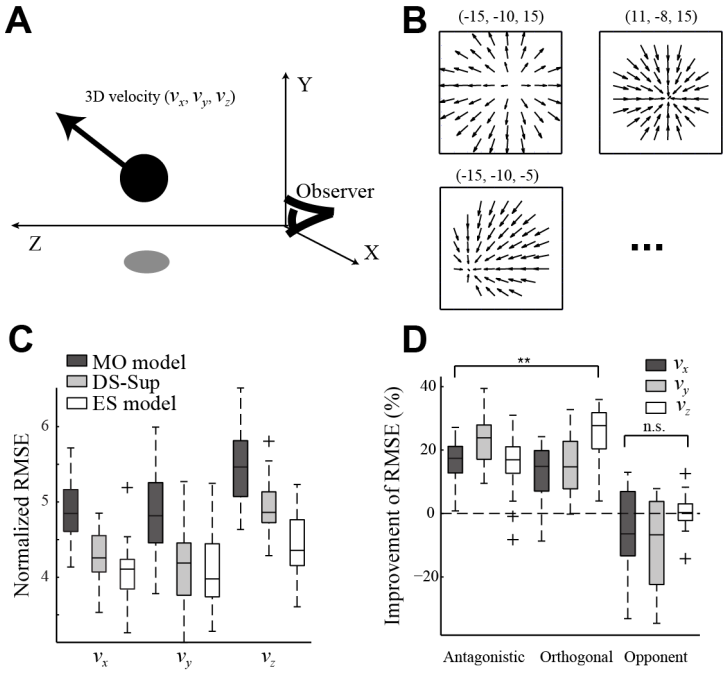

Figure 3-8 Role of suppressive surround revealed by population decoding of 3D velocity.
A. Schematic of the 3D motion population decoding task. We calculate the optic flow pattern generated by object motion in a 3D space, with the motion direction and speed randomly selected (see Methods). The resulting velocity field is processed by a population of MT models, and a linear decoder is fitted to reconstruct the 3D motion based outputs of 200 randomly selected models. **B.** Example velocity patterns used in the decoding task, with the 3D motion direction labeled above each pattern. **C.** Performance of the decoder based on input from MO model (black) and ES model

with only DS-Sup (gray) and ES model with both type of suppression (white). Results are quantified as the ratio of root mean squared error to the range of the parameter; smaller values indicate better performance. Using outputs of ES models gives better reconstruction performance along all three dimensions than using the MO models population. (p<0.05, t-test). **D.** Percentage improvement of reconstruction performance between MO model and the full ES model for model cells with antagonistic suppression (*left*), orthogonal suppression (*center*) and opponent suppression (*right*). Significant improvements of reconstruction are observed for cells with antagonistic suppression and orthogonal suppression, but not for cells with opponent suppression (** p<0.01, n.s. p>0.05, t-test).

3.4 Discussion

In this study, we have recorded the responses of MT cells in the context of naturalistic optic flow stimuli. To interpret the neuronal responses in this rich stimulus context, we have constructed a hierarchical modeling framework, which describes MT processing as integration over excitatory and both direction-selective and non-selective suppressive inputs. Most previous studies of MT have focused on isolating either the excitatory tuning of MT neurons in the receptive field center or individual suppressive influences, such as the surround suppression (Xiao et al., 1995; Born 2000) or motion opponency (Snowden et al., 1991; Qian et al. 1994). For example, experimental studies of suppression usually use the preferred stimulus to drive the center and gauge the effects of suppression in this simplified context (Snowden et al., 1991; Xiao et al., 1995; Xiao et al., 1997). At the same time, previous modeling studies have largely focused on explaining how MT computes velocity from local motion signals (Qian et al., 1994; Rust et al., 2006; Simoncelli and Heeger, 1998) within its receptive field.

Here, by fitting a hierarchical excitation-suppression (ES) model with both excitatory and suppressive components, our results provide the most complete picture of how different types of suppressive influences interact with excitation to impart selectivity to higher-order motion stimuli. In particular, we find suppression can be divided into a direction-selective component, which exhibits diverse structure and imparts functionally useful higher-order selectivity, and a non-selective component, which seems to play the role of surround suppression and normalization.

3.4.1 The use of complex motion stimuli to probe MT

The majority of work on MT has explored its role in estimating the velocity of a rigidly translating object (Lisberger and Movshon, 1999; Rust et al., 2006; Simoncelli and Heeger, 1998). The idea dates back to the discovery that response of MT neurons is primarily dependent on the motion direction within a two-dimensional plane (Zeki, 1974). However, when more complex motion fields are used, selectivity to higher-order features, such as speed gradients (Treue and Andersen, 1996; Xiao et al., 1997) and surface orientation (Nguyenkim and DeAngelis, 2003; Sanada et al., 2012) has been observed, raising the question of what features of natural vision are represented by MT neurons. Our results show that although excitatory contributions dictate direction-selectivity, they are not sufficient to explain responses to motion stimuli with different velocities across visual space: suppressive contributions with different spatial profiles than excitation also significantly modulate MT responses and thus impart selectivity to complex motion features.

Here, we focus on understanding the role of spatial heterogeneity in natural motion fields, and thus use a stimulus – a random dot field – that produces an unambiguous velocity signal as a function of space. This allows us to map the excitatory and suppressive influences on an MT neuron across space, and purposefully avoids the complexities associated with extracting velocity from texture patterns, which is another known aspect of MT processing (Jazayeri et al., 2012; Rust et al., 2006). In particular, one recent modeling study (Nishimoto and Gallant, 2011) extended such texture-based processing to explain complex motion stimuli, although as a result did not focus on the specific roles of spatially distributed suppression in processing such stimuli. We thus regard this approach as being orthogonal to understanding motion estimation for textures patterns, and expect the models investigated here might be consistent with, and/or ultimately combined with models that address those complexities (Bradley and Goyal, 2008; Nishimoto and Gallant, 2011).

3.4.2 Different forms of suppression in MT

Several forms of suppression have been documented in the literature. For example, neurons in MT are often suppressed by motion in the anti-preferred direction (Mikami et al., 1986). This is often termed "motion opponent suppression", which is likely related to similar phenomena that have been reported in psychophysical studies (Levinson and Sekuler, 1975; Qian et al., 1994), single unit recordings (Mikami et al., 1986; Rodman and Albright, 1987) and functional magnetic resonance imaging (Heeger et al., 1999).

Surround suppression is another well-studied property of MT receptive fields. Most neurons in MT have receptive fields with antagonistic surrounds (Allman et al., 1985; Born, 2000; Tanaka et al., 1986; Tsui and Pack, 2011), and the classical view is that maximal suppression occurs when the surround stimulus moves in the same direction as that in the center (Born and Bradley, 2005; Bradley and Andersen, 1998). Other studies have shown that the suppressive surrounds could be quite complex relative to the center, exhibiting such properties as asymmetric spatial organization (Raiguel et al., 1995; Xiao et al., 1997), different contrast sensitivity (Pack et al., 2005), and less direction selectivity (Hunter and Born, 2011).

Although different suppressive mechanisms are often separately studied using targeted stimuli, our modeling approach provides a unified framework for characterizing suppression. We incorporate the motion opponency assumption at the local unit level and allow for both direction-selective suppression and non-selective suppression. Interestingly, for the majority of the cells we studied, the best model has both selective and non-selective suppression, but with different spatial profiles. The direction-selective suppression is closer to the receptive field center and appears to directly modulate motion selectivity, while the non-selective suppression is more like to be the previously described surround suppression.

Although our model assumes a feed-forward structure, the source of different forms of suppression remains unclear. Given the longer latency of suppression, it may be that these suppressive contributions actually reflect horizontal connections within MT

or feedback from higher areas such as MST. The contribution of different types of suppression may also depend on the stimulus (Huang et al. 2007). While the goal of this study is to reveal types of suppression that are functionally relevant to processing of naturalistic optic flow, other types suppression, such as those that is spatially and directionally aligned with excitation (DeAngelis et al., 1992; Cavanaugh et al., 2002), may not be revealed by our method.

3.4.3 Spatial heterogeneity of MT processing

Although the classical view of MT receptive fields is that preferred directions in the center are largely homogeneous, and the surround is antagonistic and circularly symmetric (Tanaka et al. 1986), a handful of studies have focused on spatial heterogeneity of processing, using stimuli that either separately drove center and surround regions (Orban, 2008; Raiguel et al., 1995; Xiao et al., 1995), or separately mapped direction preferences (Richert et al., 2013) or sensitivity (Britten and Heuer, 1999) within the classical receptive field.

In our model, the spatial heterogeneity is reflected by the different spatial profiles and direction preferences of excitation and suppression. Although we assume the subunit selectivity is the same for each model component, different direction selectivity and sensitivity could emerge within the receptive field due to the different combinations of excitation and suppression across space due to their different spatial footprints. This effect is most prominent for cells with orthogonal oriented suppression. The complex surround of MT has been shown to play important roles in 3D shape

estimation and motion segmentation (Buracas and Albright, 1996; Gautama and Van Hulle, 2001). The additional heterogeneity in direction preferences may further contribute to selectivity to curvature in the motion field, which is an important aspect of natural motion. Indeed, the improvement of accuracy in 3D velocity estimation is most significant when orthogonal suppression is introduced to the model (Fig. 3-8D).

3.4.4 The role of MT in visual motion processing

From a computational perspective, the goal of visual motion processing is much broader than estimating 2D velocity. For a behaving animal, motion stimulus not only depends on object motion and visual depth, but also results from optic flow patterns imparted by ego-motion and eye-movements. Multiple problems are thus involved in motion processing, such as detection of independently moving objects, ego-motion estimation, 3D velocity estimation, and structure from motion (Beauchemin and Barron, 1995; Bradley et al., 1998; Pauwels et al., 2010; Sanada et al., 2012). While other cues (e.g., disparity; DeAngelis et al., 1998) are also important, some of these problems involve recognition of complex motion patterns (Fermüller and Aloimonos, 1995) and discontinuities of the motion field (e.g., motion segmentation), which likely require selectivity to higher-order motion features. Although motion processing is certainly not complete at the stage of MT, our results suggest that such higher-order selectivity is already present in feed-forward MT processing and can support behaviorally relevant tasks, such as estimation of the 3D velocity of a moving object.

MT neurons project to the medial superior temporal area (MST) (Tanaka et al., 1993; Ungerleider and Desimone, 1986), which is thought to calculate the heading direction of the observer (Perrone and Stone, 1994, 1998) and to estimate 3D velocity (Zemel and Sejnowski, 1998). MST neurons are more selective to optic flow components than MT neurons (Lagae et al., 1994) and are more invariant to the stimulus shape (Geesaman et al., 1997) and position (Duffy and Wurtz, 1995). This selectivity can be partly explained in a hierarchical framework using model MT neurons as inputs (Mineault et al., 2012), and our results here suggest processing in MT neurons might more directly facilitate these ultimate goals of motion processing.

Chapter 4 : Cortical variability arising from spontaneous and decision-related network activity

4.1 Introduction

Sensory neuron responses to identically presented stimuli can be highly variable (Faisal et al., 2008; Goris et al., 2014; Masquelier, 2013; Softky and Koch, 1993). While one source of variability may simply be sensory noise that originates at early stages of processing (Faisal et al., 2008; Zohary et al., 1994), another possible source is ongoing network activity that is simply not explicitly timed with respect to the stimulus (Arieli et al., 1996; Masquelier, 2013). Indeed, cortical neurons receive only a fraction of their inputs from upstream sensory areas (Felleman and Van Essen, 1991), and many areas of cortex have been observed to undergo spontaneous fluctuations in activity, particularly in anesthetized animals (Ecker et al., 2014; Kelly et al., 2010; Steriade et al., 1993). In awake animals, similar dynamics have been observed in conjunction with changes in cognitive and behavioral state (Cohen and Newsome, 2008; Engel et al., 2001; Morishima et al., 2009) related to, for example, working memory (Mendoza-Halliday et al., 2014) and the allocation of attention (Saalmann et al., 2007).

Neuronal variability has repeatedly been found to exhibit substantial correlation across a range of spatial and temporal scales (Cohen and Maunsell, 2009; Herrero et

al., 2013; Zohary et al., 1994). One way to detect such correlated network activity is through the local field potentials (LFPs), which are low-frequency fluctuations in the extracellular voltage that are often recorded in combination with single-unit responses (Buzsaki and Draguhn, 2004; Einevoll et al., 2013). Indeed, the timing of cortical neuron spikes is often biased by the phase of LFP fluctuations in certain frequency bands in a task-dependent manner (Fries et al., 2001; Lee et al., 2005). These considerations suggest that the LFP signals themselves may contain relevant information from a range of sources that is useful for predicting – and understanding – variability observable in single neurons.

Here, we relate response variability in the middle temporal area (MT) of macaque visual cortex to ongoing network activity, inferred using LFPs and the activity of other simultaneously recorded neurons. The stimulus selectivity of neurons in area MT has been characterized very thoroughly, and computational models can successfully predict their responses to a variety of motion stimuli (Cui et al., 2013b; Nishimoto and Gallant, 2011; Rust et al., 2006). Furthermore, MT neuron activity can reflect both processes linked with perception and behavior, such as attention (Treue and Martinez-Trujillo, 1999) and motion perception (Britten et al., 1992). Indeed, the prominent role of area MT linking representation of motion stimuli to perception is suggested by correlation between the variability of MT neuron responses (in the context of identical or ambiguous stimulus presentations) and perceptual decisions (Britten et al., 1996; Nienborg et al., 2012).

We use a model-based approach, in which the spiking responses of individual MT neurons are predicted using the combinations of the stimulus, LFPs, and activity of nearby MT neurons. We find that the inclusion of LFP-related inputs can yield dramatic improvements in the ability to predict spiking responses measured during the passive viewing of naturalistic optic flow stimuli. This improvement relates specifically to their ability to predict response variability; *i.e.*, activity that is not locked to the stimulus. Furthermore, we demonstrate that, during the performance of a motion discrimination task, the locking of neuronal responses to low-frequency bands of the LFP can be used to predict the relationship between neural responses and perceptual decisions. These results thus suggest a central role of network inputs in linking sensory neuron activity to perceptual choices.

4.2 Methods

4.2.1 Electrophysiology Recordings

Data were recorded from two adult rhesus macaque monkeys, prepared using standard surgical techniques that have been described previously (Mineault et al., 2012). Eye movements were monitored at 500 Hz by an infrared eye tracker (EyeLink II; *SR Research*). Extracellular recordings were performed on 134 well-isolated single units during a passive fixation task, of which 63 units were recorded with single electrodes, and 71 units were recorded using a multi-site linear electrode array with 16 recording sites with 100 μm spacing. An additional 79 units were recorded during performance of a two-alternative motion discrimination task. Signals were amplified, bandpass filtered, sorted on-line and resorted off-line, using spike-sorting software

(*Plexon*) to isolate single units. This system includes a dedicated low-pass filter for LFP signals (four-pole high cut at 170 Hz) followed by digitization at 1 kHz. As described in an earlier study of our lab (Zanos et al., 2011), this type of filtering can introduce spurious correlations into spike-LFP coherence. We therefore obtained custom hardware modifications to the acquisition system that included wideband analog filters (two-pole high cut at 2.5 kHz) and a higher digitization rate (10 kHz). Line noise at 60 Hz and spike waveforms were removed from LFPs using a Bayesian spike removal algorithm (Zanos et al., 2011). All aspects of the experiments were approved by the Animal Care Committee of the Montreal Neurological Institute and were conducted in compliance with regulations established by the Canadian Council on Animal Care.

Data for the LGN neurons were recorded extracellularly from anesthetized and paralyzed cat performed by the Alonso Lab for other studies (Butts et al., 2010). The stimulus consists of a 32x32 pixel natural movie, refreshed at 60 Hz, which was recorded from a camera mounted on top of a cat's head (Kayser et al., 2003). We included 52 neurons recorded with repeated presentation of a 10 seconds segment of the stimulus, and the stimulus was typically repeated more than 50 times. Responses to the first 10 repeats of the stimulus were omitted to avoid non-stationarity resulting from stimulus onset (Butts et al., 2011).

4.2.2 Passive Fixation Task

In the passive fixation task, animals were trained to fixate within 2° of a small fixation point on a computer monitor, which displayed a continuous optic flow

stimulus. The animals were given a liquid reward periodically for maintaining fixation. All stimuli were presented on a CRT monitor with a display resolution of 1,600×1,000 pixels (49°×36° of visual field at a distance of 50 cm) and a refresh frame rate of 60 Hz (*n*=123/134) or 75 Hz (*n*=11/134). Stimuli consisted of a continuously varying optic flow stimulus (Fig. 3-1A), as described in detail in previous studies (Cui et al., 2013b; Mineault et al., 2012). Briefly, the stimulus was composed of moving dots whose velocity field was generated as a random combination of six optic flow components: horizontal/vertical translation, expansion, rotation and horizontal/vertical shears. For the majority of the cells (*n*=117/134), the stimulus was displayed in a slowly moving aperture with a diameter ranging from 8° to 20°, the position of which was drawn from a Gaussian noise distribution, low-pass filtered in time with a cutoff of 0.05-0.10 Hz, with its mean at the center of the receptive field. The other neurons (*n*=17/134) used the same optic flow stimulus, although it remained centered on the neurons' receptive field. All neurons were presented with a long continuous trial lasting between 10 and 36 min (mean=18 min). For a subset of recordings (*n*=86/134), we showed a short segment of the stimulus (5 sec) that was centered on the cell's receptive field and repeated between 60 and 240 times (mean=100 times). We used these repeated presentations to measure the response reliability and to calculate the predictive power of the model, as well as distinguish between stimulus-locked and trial-variable elements of the response. The repeats were presented continuously (with no gap in between successive trials) to minimize transients due to onset responses.

For both long continuous trials and repeated trials, we excluded all data associated with fixation breaks (when the animal's gaze location deviated by more than 1.5° from the fixation point), from 100 ms before the break to 500 ms after the recovery of fixation. Only periods with fixations that were longer than 1 sec were used for model fitting and evaluation, resulting in an average of 10.1±3.2 min of usable data for each unit. The usable data was then broken into 10 sec segments, which were randomly divided into two groups: 80% of these segments were used to estimate model parameters, and model performance was evaluated on the remaining 20% of the data (*i.e.*, for cross-validation).

4.2.3 Motion Discrimination Task

Animals were trained to perform coarse motion discrimination tasks with Gabor patches, illustrated in Fig. 3-8. During this task, Gabor patches with 8-9 contrasts were presented in a block of randomly interleaved trials, while the animal maintained fixation within a 2°×2° fixation area. The range and spacing of contrasts were chosen to sample the linear range of each neuron's contrast response function. The monkey performed between 20-80 repetitions of each distinct stimulus. All incomplete trials were excluded from analysis. Model fitting and performance evaluation (Fig. 3-S9) was performed also using an 80/20 data partition on all non-zero contrast trials. Predictions of CP (described below) were performed on zero-contrast trials, which were not used for model fitting.

4.2.4 Measuring response variability

Variability of the neural response was estimated as the ratio between the power of the trial-averaged response and the total response power, after correcting for the effect of sampling noise and spike refractoriness (see Fig. 3-4-2 for more details).

Any firing rate recorded over multiple repeated presentations $\lambda^{(n)}(t)$ can be decomposed into a stimulus-locked rate $\lambda_{\text{stim}}(t)$ (*i.e.*, the PSTH) and a trial-variable component $\lambda_{\text{var}}^{(n)}(t)$,

$$\lambda^{(n)}(t) = \lambda_{\text{stim}}(t) + \lambda_{\text{var}}^{(n)}(t) \tag{4.1}$$

The variance (or "power") of the stimulus-locked rate $P[\lambda_{\text{stim}}(t)]$ is estimated following Sahani and Linden 2003 (Sahani and Linden, 2003) at 25 ms resolution. This time resolution was chosen to sufficiently capture the time scale of the observed firing rate fluctuations (Fig. 3-4-1). The power of the trial-variable rate variation $P[\boldsymbol{\lambda}_{\text{var}}]$ was estimated as the difference of the total response variance between real spikes, binned at 25 ms resolution in time, and simulated spikes from a Generalized Linear Model that accurately captured the observed stimulus-locked rate and spike history dependence (see Fig. 3-4-2). We report the ratio $P(\boldsymbol{\lambda}_{\text{stim}})/[P(\boldsymbol{\lambda}_{\text{stim}})+ P(\boldsymbol{\lambda}_{\text{var}})]$ as the "stimulus-locked rate variation".

For comparison, we also used a dataset recorded in anesthetized cat LGN in response to natural movies in the Alonso lab, described in detail in (Butts et al., 2010). This

dataset was comprised of 52 neurons recorded with repeated presentation of a 10 seconds segment of a natural movie stimulus. The stimulus was typically repeated more than 50 times.

4.2.5 Statistical modeling framework for describing MT neuron responses

We used a maximum-likelihood based framework that included both stimulus-dependent terms and network-activity-dependent terms to model MT neuron responses. We assumed that neuron responses are generated by an inhomogeneous Poisson process with an instantaneous rate $r(t)$. Measured spike trains were binned in time at 5 ms resolution to obtain the observed response rate $r_{\text{obs}}(t)$. The log-likelihood of the model is then given by (up to an additive constant):

$$LL[r_{obs}(t), r(t)] = \sum_t [r_{obs}(t) \log r(t) - r(t)], \tag{4.2}$$

where $r_{\mathbf{obs}}(t)$ is the measured neuronal response, and $r(t)$ is the model predicted firing rate (Paninski, 2004).

All models used a fixed spiking nonlinearity $F[\cdot]$ that acts on both the stimulus-dependent terms and network activity dependent terms; we refer to these as $g_{stim}(t)$ and $g_{network}(t)$ respectively. The model response is given by

$$r(t) = F[g_{stim}(t) + g_{network}(t) - b], \tag{4.3}$$

where b is the spiking threshold and the spiking nonlinearity function $F[\cdot]$ was chosen to be of the form $\log[1+\exp(\cdot)]$. Such a model can be optimized within the maximum-likelihood framework efficiently (Cui et al., 2013b; McFarland et al., 2013).

4.2.5.1 The stimulus-processing component of MT neuron model

For neurons recorded during the continuous optic flow stimuli, we based the stimulus-processing model component on a previously established model fit to this stimulus condition (Cui et al., 2013b). Briefly, the stimulus selectivity of MT neurons is based on nonlinear combinations of localized processing by V1-like inputs. There were three stimulus-processing terms (Fig. 3-1B): direction-selective excitation (Exc), direction selective suppression (DS-Sup), and non-direction selective suppression (NS-Sup), each with their unique direction selectivity (except NS-Sup), velocity sensitivity, and spatial and temporal integration. All components of the stimulus model were fit to the data, as previously described (Cui et al., 2013b).

For neurons recorded during the motion discrimination task, we modeled the time-varying firing rate of the neuron to the brief motion stimulus directly, resulting in the following generating signal for trials with contrast level c and motion direction θ:

$$g_{stim}^{c,\theta}(t) = w^{c,\theta} \sum_i k_{\mathrm{i}}^{\theta} B_{i,2}(t - t_s), \tag{4.4}$$

where t_{stim} is the stimulus onset time of the current trial, $B_{i,2}(t)$'s are a set of piece-wise linear (*i.e.*, second-order) B-spline basis functions with 10 ms spacing. The coefficients k_{i}^{θ} capture the dynamics of the stimulus-triggered response. Because the temporal response at different contrasts to a given motion direction was very similar up to an overall scaling, we introduce a gain factor $w^{c,\theta}$ that scales the temporal kernel for each stimulus conditions. The temporal kernel k^{θ} and gain factor $w^{c,\theta}$ were alternately fit until converged.

4.2.5.2 The network components of the MT neuron model

Our model inferred the relevant network inputs from both the LFPs and MUA recorded from the electrode array. To process the LFPs, a continuous wavelet transformation was performed either on the LFP signal itself or, in the case of multi-electrode array recordings, on the second spatial derivative of the LFP signals, using a complex Morlet mother wavelet at 16 logarithmic scales from 0.5 Hz to 70 Hz. Calculation of the spatial derivatives of the LFPs closely related to current source density analysis (Nicholson and Freeman, 1975), and yielded a more localized distribution of model weights compared to those without this spatial transform.

The instantaneous phase ϕ_{db}(t) and amplitude ψ_{db}(t) for the (transformed) LFPs at a given depth *d* and frequency band *b* were given by the phase and amplitude. We used amplitude-modulated sine and cosine terms as linear predictors, with model coefficients α_{db} and β_{db}:

$$g_{LFP}(t) = \sum_{d,b} \psi_{db}(t)[\alpha_{db} \cos[\phi_{db}(t)] + \beta_{db} \sin[\phi_{db}(t)]] = \mathbf{k}_{\mathrm{LFP}} \cdot \mathbf{X}_{\mathrm{LFP}}(t), \quad (4.5)$$

where $\mathbf{X}_{\mathrm{LFP}}(t)$ is a design matrix with components $\psi_{db}(t)\cos[\varphi_{db}(t)]$ and $\psi_{db}(t)\sin[\varphi_{db}(t)]$, and $\mathbf{k}_{\mathrm{LFP}}$ is the LFP-receptive field vector that contains the corresponding model parameters α_{db} and β_{db}. Note that by fitting different linear weights on the cosine and sine phases of each signal, this implicitly yields the reported model weights sqrt(α_{db}^2+ β_{db}^2) and preferred phases arctan(β_{db}/α_{db}), for each frequency band and depth.

We also used simultaneously recorded MUA to predict neuronal responses with a causal coupling filter. The generating signal based on MUA is,

$$g_{spk}(t) = \sum_{d,\tau} k_{d\tau} r_d(t-\tau) = \mathbf{k}_{\mathrm{spk}} \cdot \mathbf{X}_{\mathrm{spk}}(t), \quad (4.6)$$

where $r_d(t)$ is the MUA from depth d, and $k_{d\tau}$ reflects the dependence of the spike response from depth d at time τ before. The MUA from the same depth was always excluded to avoid contamination due to imperfect spike sorting, and for channels where other single units were isolated, the single unit replaced the MUA signal at those depths.

The network components of the model were fit simultaneously with the stimulus-processing components (described above), and thus the model weights implicitly were not redundant with the stimulus-model predictions.

4.2.5.3 Regularization of the LFP model

Since we sampled the LFP signal at 16 frequencies, the LFP model has 32 parameters for each depth and 512 parameters across 16 depths. The MUA filters include 15 depths and 15 lags for each unit. To avoid overfitting, we applied L2-regularization to the model coefficients (McFarland et al., 2013), using a penalty term added to the *LL* that penalizes non-smoothness of the filter. For the LFP model, this term can be expressed as,

$$LL_{smooth}^{LFP} = \eta_d^{LFP}(\|\mathbf{L}_\mathrm{d}\alpha_{db}\|_2 + \|\mathbf{L}_\mathrm{d}\beta_{db}\|_2) + \eta_b^{LFP}(\|\mathbf{L}_\mathbf{b}\alpha_{db}\|_2 + \|\mathbf{L}_\mathrm{b}\beta_{db}\|_2), \quad (4.7)$$

where $\mathbf{L}_\mathrm{d}$ and $\mathbf{L}_\mathrm{b}$ are discrete Laplacian operators with respect to the depth and frequency dimensions. The hyper-parameters η_d^{LFP} and η_b^{LFP} control smoothness across frequency bands and across cortical depth respectively. For the MUA filter, the penalty term is similarly expressed as,

$$LL_{smooth}^{spk} = \eta_d^{spk}\|\mathbf{L}_\mathrm{d}k_{d\tau}\|_2 + \eta_\tau^{spk}\|\mathbf{L}_\tau k_{d\tau}\|_2 \quad (4.8)$$

The hyper-parameters η_d^{LFP}, η_b^{LFP}, η_d^{spk}, and η_τ^{spk} were adjusted using a nested cross-validation scheme, where 20% of the fitting data were randomly selected and reserved during optimization of the LFP filter. The regularization parameters that gave the best performance on the reserved data were used to evaluate performance in

the cross-validation data (which thus was neither used to fit model coefficients, nor the hyperparameters).

4.2.5.4 Model Evaluation on Repeat Trials

Data recorded with repeated stimuli to were used to calculate predictions of the stimulus-driven response and the non-stimulus locked response separately. The generating signal of the model was decomposed into a stimulus-locked component and a (zero-mean) trial-variable component, $g^{(i)} = g_{stim} + g_{var}^{(i)}$. We then used either the stimulus-locked component g_{stim} or the residual component $g_{var}^{(i)}$ to predict responses on the repeat trials. Because overall gain and offset of the spiking nonlinearity often drifts over long periods, we refit these two parameters from the overall model (fit using long continuous stimuli) using the odd repeat trials. We thus excluded the odd repeats from cross-validation, and used only the even repeated trials to measure model performance.

4.2.6 Predicting Choice Probability of MT neurons

Choice probabilities (CPs) of single MT neurons were calculated on trials with no stimulus presented (0% contrast condition), using standard procedures(Britten et al., 1996). The Receiver Operating Characteristic (ROC) curve was computed using spike counts over 100-200 ms after (expected) stimulus onset throughout, except in Figure 6B where we measured CP over time with a 100 ms sliding window. The area under the ROC curve is the choice probability of the neuron. We only used neurons in this analysis that showed clear direction-selective motion response (n=79).

Predicted CPs were calculated based on the model-predicted response on every zero-contrast trial. [Note that models were not fit to the zero-contrast trials, and thus these represent cross-validated tests.]

4.2.7 Statistical Analyses

We used robust statistics when comparing performances of different models as the log-likelihood were often not normally distributed. Paired comparisons of group medians were performed with two-sided Wilcoxon signed rank tests. The variability of medians was estimated using bootstrapping techniques (random sampling with replacement with 1000 repetitions). Normality of the measured and predicted choice probability was confirmed with the One-sample Kolmogorov-Smirnov test, and correlations between measured and predicted choice probability were computed using Pearson's product-moment correlation coefficient. Significance of correlation is computed by transforming the correlation to create a t statistic having n-2 degrees of freedom. Significance of the difference between two correlation coefficients was computed using the Fisher r-to-z transformation.

4.2.8 Analysis of the impact of microsaccade on choice probability

The effect of microsaccades on stimulus processing models during the continuous optic flow stimulus was minor, as previously analyzed and reported (Cui et al., 2013b). However, previous work has demonstrated a link between microsaccades and choice probability (Herrington et al., 2009), which we further investigated. In the task of the previous study, choice probability was measured in the context of a "zero-

coherence" stimulus (Britten et al., 1996), and microsaccade induced different effects on neural firing rates depending on their direction due to the presence of a stimulus. Such effects were not expected in our task, due to the absence of a stimulus in the context of computing choice probabilities, but verified this through the following analyses.

Microsaccades were detected using an established algorithm (Engbert and Mergenthaler, 2006). Across the 22 behavior experiment sessions, the average microsaccade frequency was 1.71±0.57 Hz for the zero-contrast conditions (calculated from 100 ms prior to the [expected] motion stimulus onset time, to 250 ms after the stimulus onset). The frequency of microsaccades was comparable to that seen in other studies (Martinez-Conde et al., 2013). We measured the microsaccade-triggered average neural response as a function of direction in the zero-contrast trials, and – as expected – did not observe any direction-selective bias. Moreover, we did not observe a difference in the microsaccade rate between the preferred- and null-choice trials in any of the experiments sessions ($p>0.05$, χ^2 *test*), meaning that microsaccade frequency also could not explain choice probability. Therefore, we concluded that microsaccades do not contribute to choice probability in this study.

4.3 Results

4.3.1 Response variability of MT neurons

We presented a naturalistic optic flow stimulus (Fig. 4-1A, *top*), while recording from neurons in area MT in macaques during a passive fixation task. This stimulus robustly

modulates MT neuron firing rates (*bottom*), permitting a thorough characterization of neuronal selectivity to visual motion (Cui et al., 2013b). As a result, a sophisticated stimulus-processing model (Fig. 4-1B) captured a sizeable fraction of the neuron's firing rate response (Fig. 4-1A), as reflected by the predictive power (*PP*), a measure of model performance corresponding to the fraction of stimulus-locked (*i.e.*, firing rate) response predicted by the model (Sahani and Linden, 2003). We found a mean predictive power of 29.6%±2.7% (n=86), comparable to the performance of other models of stimulus processing in visual cortex (David and Gallant, 2005; Nishimoto and Gallant, 2011).

However, the fraction of the *total* response explained by this and previous models is in fact much lower, because, by construction, stimulus-processing models cannot capture elements of the response that are not stimulus-locked. Across the population, we observed a range of trial-to-trial variability, with some neurons showing a great deal of stimulus-locked activity, as evidenced by the highly repeatable responses to identical stimuli (Fig. 4-1C, *right*), and other neurons showing much more variability (*left*).

To estimate what fraction of the total response could not be explained from the stimulus alone, we partitioned the neuron's response to repeated presentations of the same stimulus into a *stimulus-locked* component, and a *trial-variable* component. The fraction of MT neuron response power that was stimulus-locked was calculated as the ratio of the variance in the PSTH to the total response power, after correcting both for

sampling noise and spike history effects (Czanner et al., 2008). We found that there was a broad range in the amount of stimulus-locked modulation (Fig. 4-1D), with a mean of 51.0%±27.8%. By comparison, a similar analysis of data recorded in the LGN of anesthetized cats to natural movies showed that 87.0%±13.7% (n=52) of the firing rate modulation was stimulus locked. This is consistent with previous observations that the amount of variability grows through successive levels of visual processing (Kara et al., 2000), and suggests that stimulus-processing models cannot describe more than half of MT neuron activity.

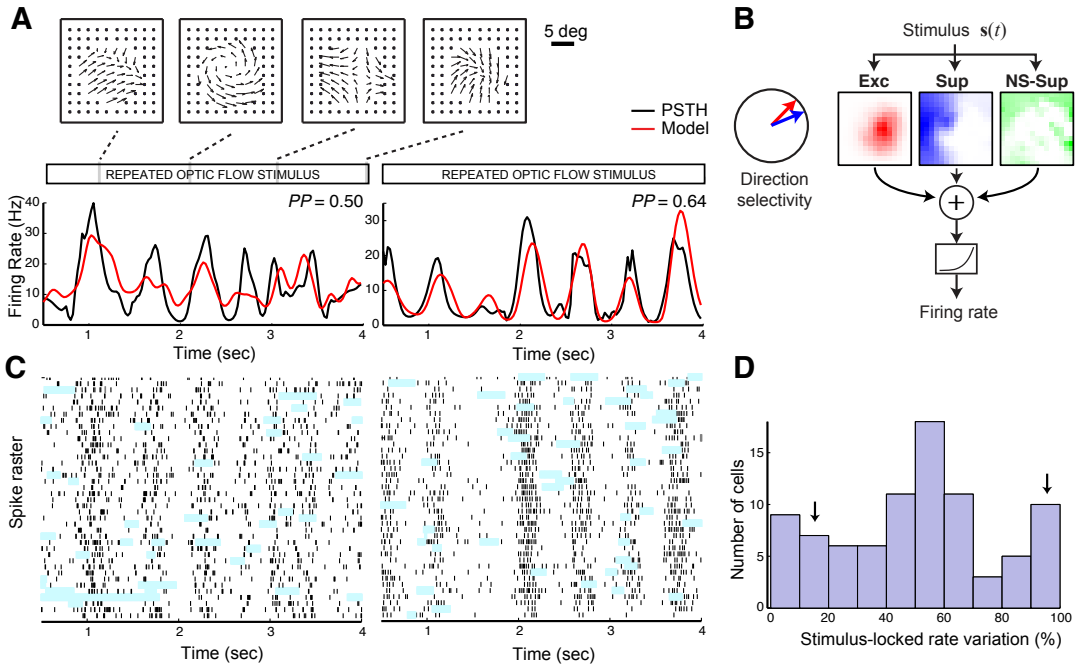

Figure 4-1 Response variability of MT neurons.
A. The responses of two MT neurons to repeated presentations of a continuously varying optic flow stimulus, with four example frames from a section of the repeated stimulus segment (*top*). The peristimulus time histograms (PSTHs, *bottom*, black) are compared to the firing rate prediction of the stimulus-processing model (red), labeled with the fraction of stimulus-locked response explained by the model (predictive power, *PP*). **B.** Model schematic illustrating the main components of the stimulus processing model: the stimulus composed is first processed locally by direction- and speed selective local subunits, and then gets pooled across space separately by

excitatory (Exc, *left*), direction-selective suppressive (DS-Sup, *middle*) and non-selective suppressive (NS-Sup, *right*) components. Red and blue arrows (*left*) indicate direction selectivity for the Exc and DS-Sup terms. Finally, a spiking nonlinearity is applied on this signal to transform it into a firing rate prediction (*bottom*). **C.** Spike rasters of the same neurons as in A, from which the PSTH is generated. The shaded areas mark where data excluded from analysis due to periods when eye position was outside the 2° fixation window. **D.** The distribution of stimulus-locked rate variation across MT neurons (n=86, mean=51.0%±3.0%) The black arrows indicate the location of the two example neurons in (A) (stimulus-locked variance = 16.5% and 96.4%).

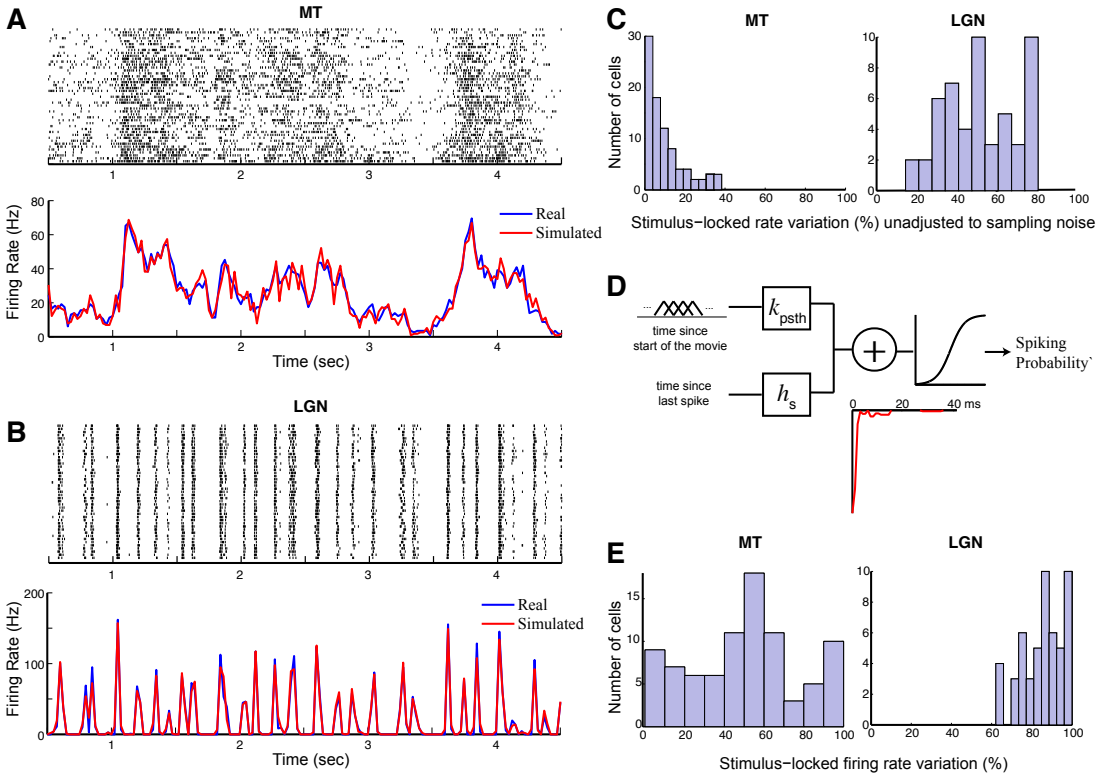

Figure 4-2 The amount of stimulus-locked firing rate modulation.
A. Spike raster (*top*) showing the responses of an example MT neuron to repeated presentations of a continuous optic flow stimulus, with the resulting PSTH (*bottom*, blue) compared with the firing rate predicted by the model described here (red). **B.** Spike raster and PSTHs of an example LGN neuron responding to a natural movie stimulus, displayed as in (A). We analyze an LGN dataset recorded from anesthetized cat – where little trial-variable firing rates are expected, in order to validate the methods described here (and for comparison). **C.** To address how much of each neuron's response is stimulus-locked, we calculated the ratio of the variance of the PSTH to the total variance of the spike-count within each trial, using 25 ms bins. The resulting distributions of stimulus-locked variance across neurons for each data set

are shown for MT neurons (*left*, 10.2%±1.0%, *n*=86) and for LGN neurons (*right*, mean=52.4%±2.5%, *n*=52). Note that although LGN neuron response are highly reliable across repeats, there is still a large fraction of trial-variable spike count variance, which is likely due to sampling noise. **D.** To estimate the amount of this sampling noise, we construct a model that explicitly fits the firing rate over time in the context of a "spike-history term", both of which are fit in the context of a generalized linear model (model schematic shown). The spike-history term models the effect of a spike on subsequent firing rate (*lower inset*, red). The model is used to generate spike counts in each bin that are consistent with the observed stimulus-locked firing rate – which is the same on every trial – and effect of spike history. Because the variances of independent processes are additive (*i.e.*, variance of stimulus-locked and trial-variable responses), the additional variance in the observed response is what we call the trial-variable firing rate modulation. **E.** The fraction of stimulus-locked rate variation is then given by the ratio of the adjusted stimulus-locked variance to the total firing rate modulation (that includes both stimulus-locked and trial-variable components), and is shown for the population of neurons in MT (*left*, mean=52.7%±3.1%, *n*=86) and LGN (*right*, mean=87.0%±1.9%, *n*=52).

4.3.2 Using network activity to predict MT neuron responses

In considering such a magnitude of response variability, a fundamental question is whether it can be related to any other experimental, or otherwise observable, variables. One possibility is that response variability might reflect ongoing activity in the cortical network that is coordinated within and/or between cortical areas (Arieli et al., 1996; Masquelier, 2013). Such coordinated network activity can often be detected in the LFPs (Banerjee et al., 2012). Different frequency bands of the LFP are thought to represent unique aspects of network dynamics (Buzsaki and Draguhn, 2004), and can be further resolved using LFP profiles across depth (de Cheveigne et al., 2013; Schroeder and Lakatos, 2009) (Fig. 4-3). Furthermore, relevant components of network activity might also be detected in the multi-unit activity (MUA) recorded with multi-electrode recordings (Bair et al., 2001; Kelly et al., 2010).

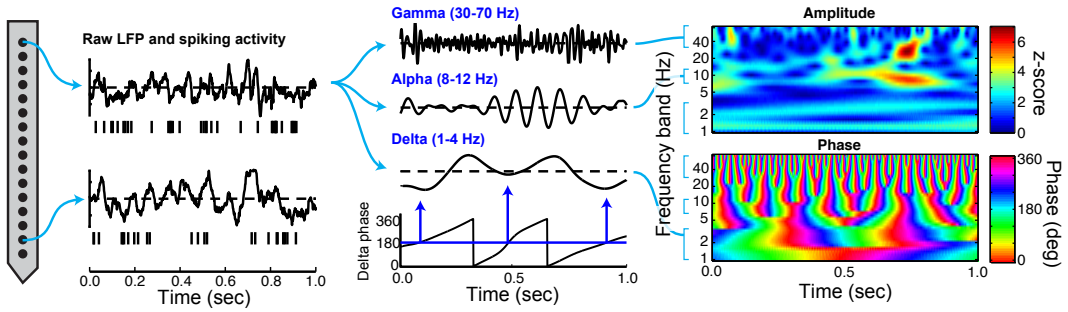

Figure 4-3 Time-frequency analysis of the LFPs.
Left: LFP signals and spiking activity were recorded with a 16-channel multielectrode array, and used as inputs to a model that predicts a given well-isolated single unit on the array (described below). *Middle*: To illustrate how each LFP was processed for use in the model, we show an LFP segment band-passed in the gamma (30-70 Hz), alpha (8-12 Hz) and delta (1-4 Hz) bands. *Right*: To process each LFP into input to our model, we use a continuous wavelet transform on the LFP signal to dissociate the LFP signal into amplitude and phase in each band. To demonstrate how phase relates to a given band-passed signal, the delta phase is shown below the delta-band-filtered LFP (*bottom middle*), with times that phase is 180° (blue arrows) corresponding to dips in the band-passed signal.

In order to examine the relationship between response variability and ongoing network activity, we extended our previously developed stimulus-processing model to incorporate network components derived from LFPs and MUA (Fig. 4-4A), recorded across cortical depth with linear multielectrode arrays (Fig. 4-3). The LFP component of the model was composed of a relative weight and preferred phase (Fig. 4-4B) for each frequency band and channel (Fig. 4-3), and the MUA component of the model used a temporal filter that weighted past MUA activity from each electrode (Fig. 4-4C). Model parameters of all three components were estimated using a maximum-likelihood framework (see Methods). By simultaneously fitting all three types of model components, this approach automatically accounts for the strong correlations between the various model inputs, assigning weights to those elements that best explain the MT neuron response.

We found that inclusion of the network inputs into the model dramatically improved the model performance in explaining the observed MT spike trains, compared to the stimulus-processing model alone (LL_x improvement = 64.0%±14.0%, $p<10^{-11}$, Wilcoxon signed rank test), which is more than double the performance improvement achieved by the sophisticated stimulus processing model in our previous study (Cui et al., 2013b). Much of this model improvement was due to inclusion of the LFP recorded on the same channel as the recorded unit (LL_x improvement = 30.8%±4.5%, n=71, $p<10^{-11}$, *Wilcoxon signed rank test*), which we verified using separate set of recordings with single electrodes (LL_x improvement = 37.8%±10.3%, n=63, $p<10^{-10}$, *Wilcoxon signed rank test*). The full model performance was achieved by including LFPs from other channels (LL_x improvement = 12.5%±4.8%), and furthermore using MUA (LL_x improvement = 4.2%±1.6%) (Fig. 4-4D). In order to ensure that the improvement in model performance was not due to contamination of the LFPs or the MUAs by spike waveforms, we used a robust spike-removal algorithm to subtract spikes from the LFP signal (see Methods) and excluded MUA on the same electrode. We also validated the LFP components of our models using the LFP signal from a different recording channel (Zanos et al., 2012) (Fig. 4-4-5).

To examine which frequency bands in the LFP were important in predicting spiking responses, we measured the impact of removing individual frequency bands on model performance. This analysis revealed that a broad range of frequency bands were predictive of spiking responses (Fig. 4-4E), with large contributions coming from the

delta band (1-4 Hz), the beta band (12-30 Hz) and the gamma band (30-70 Hz range). Across the population, MT neurons had very consistent preferred phases of the local LFP signal (Fig. 4-4F).

While there were a variety of weight profiles across the range of depths spanned by the electrode array, LFP weights in a given band were often concentrated at a particular depth relative to the channel from which the spikes were obtained (Fig. 4-4B). The gamma band modulation usually originated at the same depth as the recorded neuron relative to other frequency bands while delta band modulation tended to originate from more superficial depths. In contrast, the model weights on the MUA did not exhibit this rich structure across depth, and weighted nearby neurons most strongly (Fig. 4-4C).

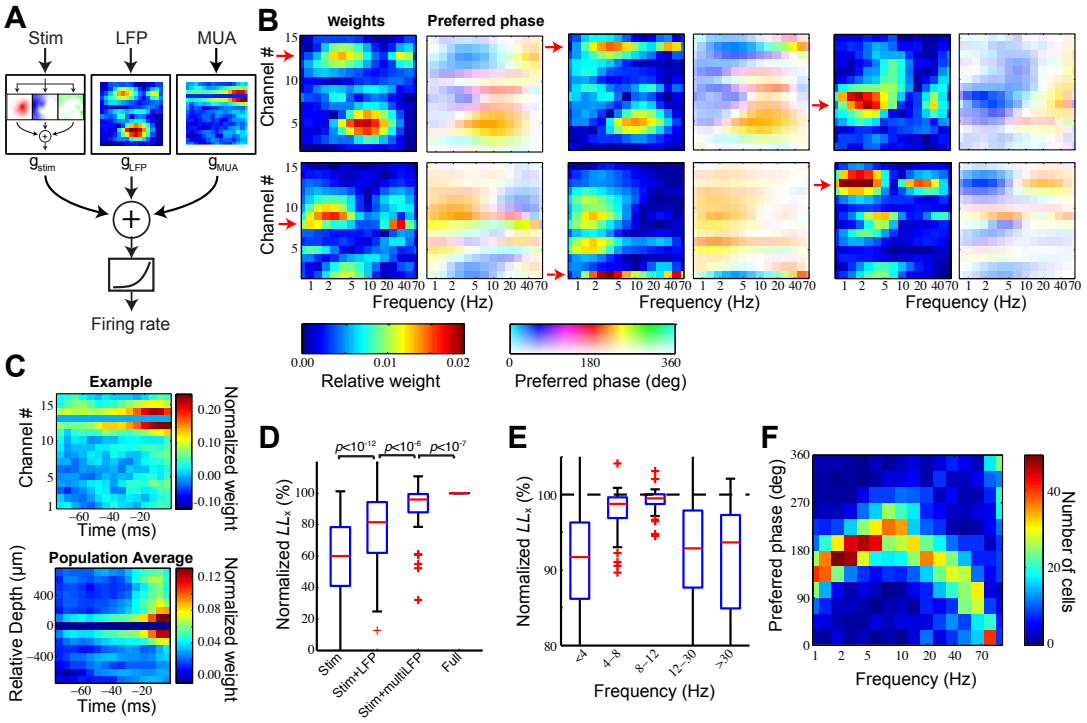

Figure 4-4 Predicting MT responses from the stimulus and inferred network activity.

A. Schematic of the full MT model, which includes a stimulus-processing component (from Fig. 4-1), and two network-based components: using information from LFPs and multi-unit activity (MUA). **B**. There is a diverse array of LFP components, shown for six examples, with the red arrow designating the depth of the unit being modeled. Model weights (*left*) and preferred phases (*right*) are displayed as a function of both frequency bands (*horizontal axis*) and cortical depth (*vertical axis*). Phase values are depicted by hue, with the color brightness representing the corresponding weights in order to emphasize bands and depths that are most important. **C**. *Top*: example MUA model component, showing weights on MUA as a function of time lag (*horizontal axis*) and depth (*vertical axis*), for the same neuron as the upper-left example in (B). *Bottom*: Average of all MUA model components across population of recorded neurons. **D.** Contribution to model performance of adding the network components, using the log-likelihood relative to the full model based on cross-validation data (normalized LL_x). **E**. To gauge what frequency bands were important for the LFP, we measured the decrease of model performance relative to the full model after removal of different frequency bands across neurons (n=71). **F.** The consistency across neurons of phase preferences in each LFP band is demonstrated by two-dimensional histograms of each neuron's preferred phase in each frequency band (n=134; note this includes recordings with both single-electrode and multi-electrode array).

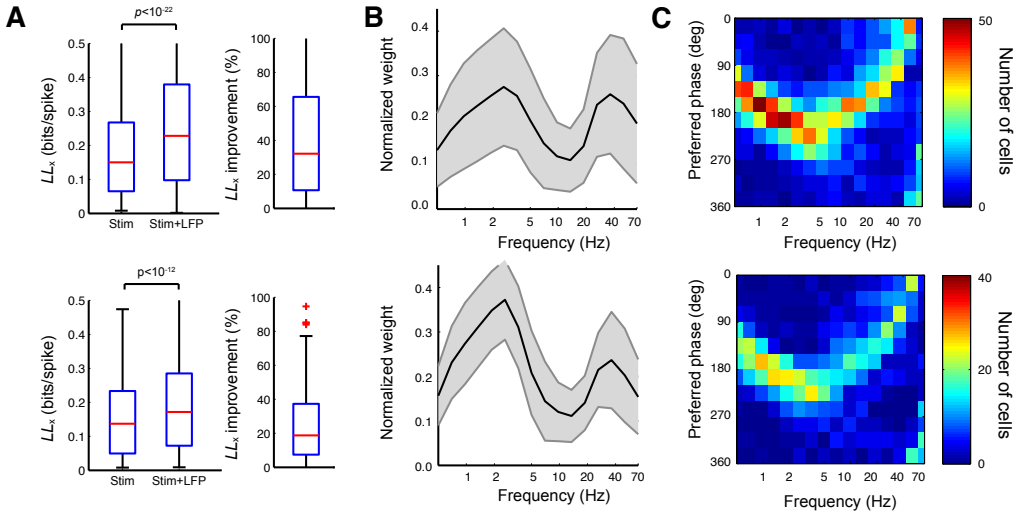

Figure 4-5 Predictions of MT neuron responses are similar between same and neighboring electrodes.
To provide additional evidence that the LFP model components were not simply using residual elements of the neuron's own spiking response, for multi-electrode array recordings (n=71), we compared models fit using the same electrode that is used for recording spikes (*top*) versus a neighboring electrode 100 μm away (*bottom*). Note that we expect they will have relatively similar LFPs, but any residual artifacts from the single unit spikes will be either gone or at least radically diminished. Nevertheless, LFP model weights, preferred phases, and overall performance are comparable. **A.** Cross-validated model performance for stimulus processing models and models with LFP signals. There is a highly significant improvement in model performance in both cases ($p<10^{-22}$ for the same electrode, $p<10^{-12}$ for using a different electrode, Wilcoxon signed rank test, n=71). **B.** Normalized LFP model component weights across frequency bands. These are comparable weights overall, although the weights in the gamma band were slightly smaller for the model based on a different electrode. **C**. Distribution of the preferred phase across frequency. There was no noticeable difference in the preferred distribution across the two conditions.

4.3.3 Trial-to-trial variability predicted using signals within the LFP

Did LFP-based improvements in the model arise from the better ability to predict the response to the stimulus, or trial-to-trial variability? To answer this question, we analyzed the model output over repeated presentations of the stimulus to distinguish between stimulus-locked and trial-variable contributions to model predictions (Fig. 4-

6A). By construction, the stimulus-dependent component of the model had the same output on every trial. In contrast, the output of the LFP component had very few stimulus-locked features, suggesting that it mainly contributed to trial-variable components of the spiking response. Finally, the MUA model component typically had both stimulus-locked and trial-variable components, with clear similarities to both the stimulus-dependent and LFP component outputs.

To gauge the contribution of each component to model performance, we separated the stimulus-locked and trial-variable components of the full model prediction, and computed how much each contributed to the likelihood improvement over the stimulus-processing model alone (Fig. 4-6B, *left*). This revealed that the bulk of the model improvement gained from adding the network components was attributable to ability to better predict trial-variable firing rate, with only a small contribution to improvement in predicting the stimulus-locked response. The performance improvement of the full model was largely a result of the LFP component, since the omission of the MUA term only reduced this contribution slightly (Fig. 4-6B, *middle*). In contrast, the absence of the LFP model component lead to a large decrease in the ability to predict trial-variable responses, even with the MUA term present (Fig. 4-6B, *right*). This suggests that – although some information about both the stimulus-locked and trial variable response is present in the activity of neighboring neurons – such information does not contribute significantly in the presence of effective stimulus-processing and LFP model components.

Because the contributions of the LFP component were almost entirely trial-variable, we expected that the LFP component of the model would increase model performance even when the stimulus model was perfect. We tested this using a subset of recordings that had enough repeated stimulus presentations to directly fits the stimulus-locked rate (see Methods), and thus perfectly predicts the entire stimulus-locked response (Haslinger et al., 2012). Indeed, the addition of an LFP component to this "perfect" stimulus-processing model lead to a comparable level of improvement over the stimulus model (Fig. 4-4-7), suggesting that the LFP model was indeed capturing the trial-variable fluctuations in firing rate not accessible to even the best possible predictions of the stimulus response.

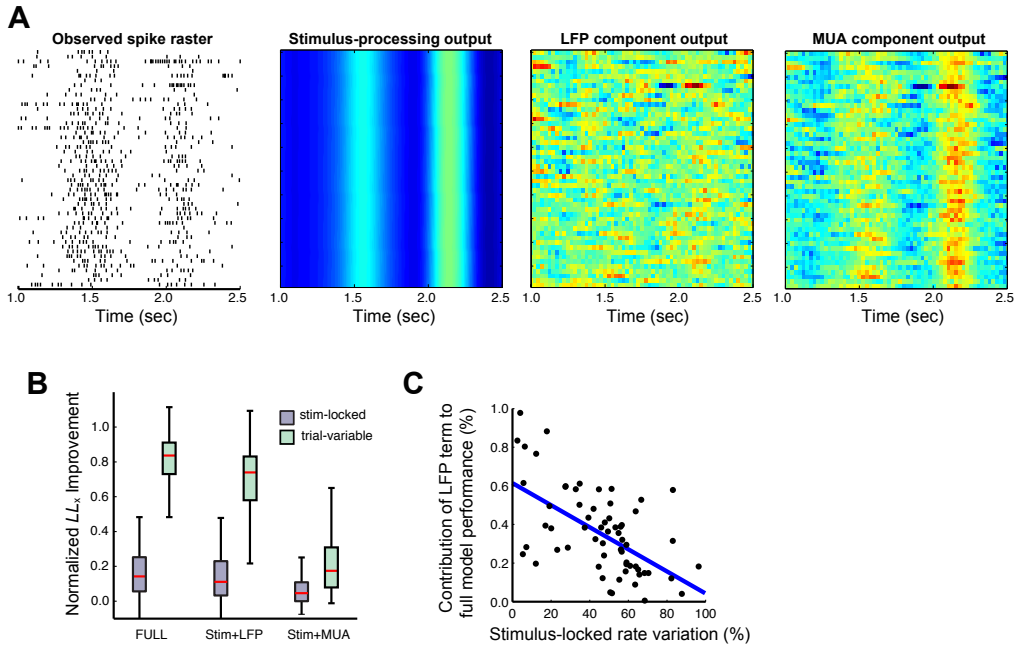

Figure 4-6 Network inputs predict trial-to-trial variability.
A. Spike raster across repeated trials (*left*), and the output of the three model components across repeated trials of the same motion stimulus (*right*). The stimulus component output is identical across trials; the output of the LFP-dependent term is

highly variable across trials, and the MUA component has both stimulus-locked and trial-variable elements. **B.** The fractional contribution to model performance of each network input (*i.e.*, fraction of LL_x improvement between the stimulus-processing model to the full model), divided into the stimulus-locked and trial-variable parts of the response predicted by the model. The trial-variable component accounts for most of the model improvement of the full model (*left*). Removal of the MUA component from the full model only reduced this contribution slightly (*middle*), whereas the absence of the LFP signal leads to a large decrease of the trial-variable component's contribution (*right*). **C.** Neurons that had more trial-to-trial variability had larger model improvements (R=-0.57, $p<10^{-6}$, n=71).

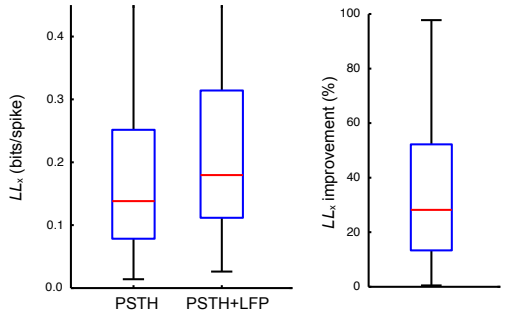

Figure 4-7 Additional evidence that the LFP component improves the predictions of trial-variable firing rate.
For neurons recorded with repeated presentations of the same stimulus (n=86), we could use a "perfect" prediction of the stimulus-locked term based on the observed PSTH, to verify that the improvements from the LFP component of the model indeed predicts trial-variable responses. In other words, if the stimulus-locked response is perfectly predicted, any additional model improvements must be in accurate prediction of trial-to-trial variability. The LFP model component for each neuron was fit as previously described, and then applied to the LFPs during each trial of the repeated stimulus, resulting in a the LFP model output $g_{LFP}(t)$. We used half of the repeat trials (odd-numbered trials) to fit an optimal gain (multiplier on g_{LFP}) and offset to augment the firing rate that matches the PSTH, and then cross-validated on the other half of the trials (even-numbered), We found that inclusion of the LFP signal yielded a significant improvement of the cross-validated likelihood (LL_x improvement = 30.9±4.8%, $p<10^{-12}$, *Wilcoxon signed rank test*, n=86), measured on trials not used for refitting gain and offsets (even trials). Because all the stimulus-locked response is captured by the stimulus-processing term, the inclusion of the LFP component must be contributing to the prediction of the trial-variable response.

Overall, these results suggest that much of the trial-variable response in MT neurons – which is typically considered "noise" in the context of stimulus processing – are the

result of network inputs. Indeed, the improvement in model performance gained by including the LFP component was more substantial for cells that had more trial-to-trial variability (*i.e.*, smaller percentages of stimulus-locked power) (Fig. 4-6C). This suggests that neurons with large trial-to-trial variations – rather than being unreliable – are simply more driven by network activity, and obtaining an accurate picture of their function requires taking such influences into account.

4.3.4 Predicting choice probability of MT neurons

Trial-to-trial fluctuations in MT neuron firing rates have also been studied in the context of motion discrimination tasks, and trial-variable MT neuron activity has famously been shown to correlate with perceptual decisions (Britten et al., 1996; Nienborg et al., 2012). We hypothesized that such choice-related variability can be inferred from the LFP signals in the same way as during the simple fixation task described above, leading to a network-level explanation of the link between MT neuron activity and behavior. To test this idea, we recorded from MT neurons during a simple two-alternative forced choice motion discrimination task.

In our task, the subject was required to maintain fixation during brief presentation of a moving Gabor patch stimulus, and subsequently to saccade to one of two targets designating the perceived motion direction (Fig. 4-8A). On each trial, stimulus motion was either in the neuron's preferred direction, or in the opposite "null" direction. We controlled task difficulty by adjusting the contrast of the Gabor patch, and in a subset of trials used zero contrast (*i.e.*, no stimulus present), with a reward

being given on 50% of trials randomly. Typical performance was near chance at contrast levels <4%, and increased to about 85% at the highest contrast (Fig. 4-8B).

As with the passive fixation task considered above, we used a model that used both the stimulus and LFP to predict neuron responses. For the stimulus-processing component of the model, we simply used time-varying rate fit to explain the observed PSTH (see Methods), which was fit simultaneously with the LFP component using all trials other than the zero contrast condition. The resulting LFP component of the models strongly resembled those fit with the continuous optic flow stimulus, and lead to a similarly large improvement in model performance over the model that used the stimulus-processing component alone (Fig. 4-4-9).

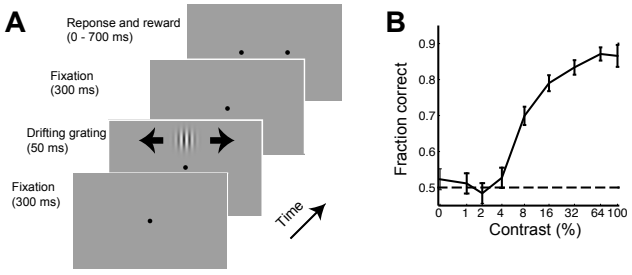

Figure 4-8 Motion discrimination task.
A. A given trial of the motion discrimination task began with the onset of the fixation dot. After 300 ms of continued fixation, a Gabor patch moving in one of two directions was presented for 50 ms at the location of the MT receptive fields. The two possible motion directions were either in the preferred or null direction of the recorded group of neurons on the array. After maintaining fixation for 200 ms from the stimulus presentation, the subject reported the perceived motion direction through a saccade to one of the two choice targets. The subject was required to make a saccade to the correct target within 700 ms after stimulus onset in order to receive a reward. **B.** Performance as a function of contrast of the Gabor patch on the motion discrimination task. The subjects' performance was at chance level when the contrast of the Gabor patch is low (<4%) and increased to about 85% at the highest contrast level.

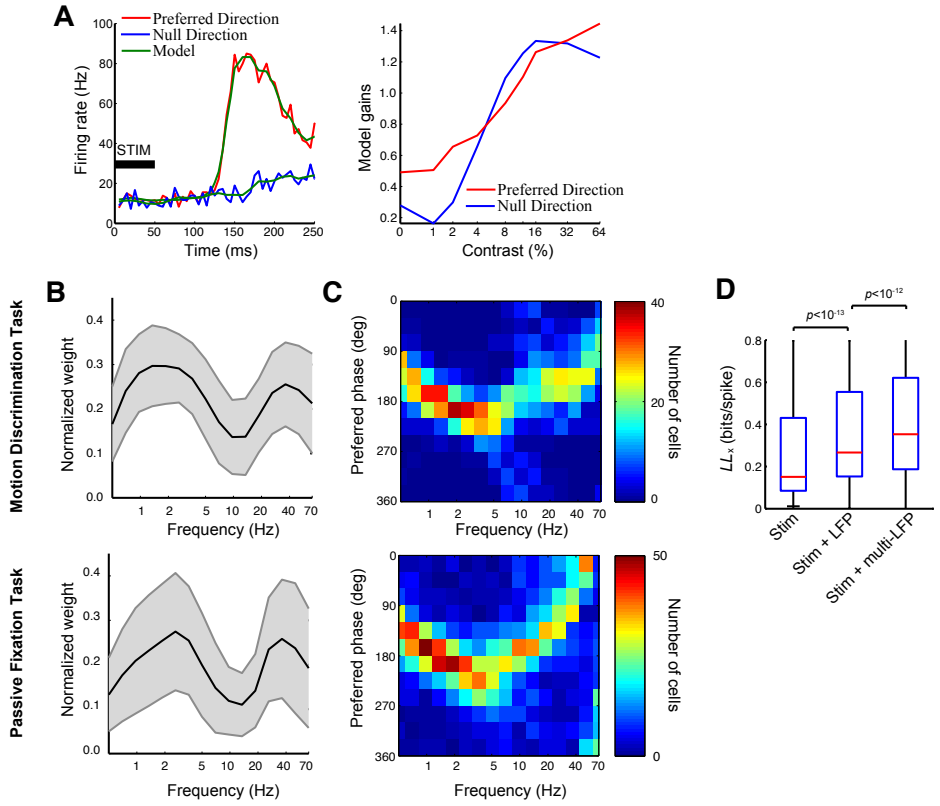

Figure 4-9 Models of MT neurons in the context of the motion discrimination task.

A. The stimulus-processing model component for the motion discrimination task is composed of firing rates for preferred (red) and null (blue) stimuli (*left*), and contrast dependent multipliers (*right*) that scale the firing rate to each contrast condition. All parameters (preferred and null firing rate at 10 ms resolution, and gains) are fit along with LFP model components. The model accurately captured average response to the motion stimuli in the preferred and null direction (*left*), and overall captured 0.93±0.01 of the stimulus-locked response. (**B,C**) LFP model components for the motion discrimination task (*top*) and passive fixation task (*bottom*) are very similar. **B.** Normalized LFP component weights across frequency for single-electrode LFP models (n=79). We considered single-electrode models here because they are less dependent on the relative depth of the electrode array (since they only use local signals), and thus yield a more fine-grained comparison across different recordings. **C.** Distribution of the preferred phases across frequency band for the same models. **D.** Cross-validated model performance for all models fit to the neurons in the motion discrimination task. The LL_x is that compared for models with stimulus processing alone (Stim), stimulus processing and single electrode LFP model components (Stim+LFP), and stimulus-processing and full LFP model components (Stim+multiLFP). As with the passive fixation task, the inclusion of the LFP model components significantly improved model performance (LL_x improvement =

29.8%±5.2%, $p<10^{-13}$, *Wilcoxon signed rank test*), and inclusion of LFP signal at other depths further improved model performance (LL_x improvement = 16.5%±3.7%, $p<10^{-12}$, *Wilcoxon signed rank test*). The magnitude of performance improvement is comparable to the passive conditions (Fig. 4-4D).

The correlation between neuron firing rates and decision in the presence of an ambiguous (or absent) stimulus is commonly measured using "choice probability" (CP), which reflects how well an ideal observer can predict the subject's choice using a given neuron's response(Nienborg et al., 2012). We calculated CPs with a 100 ms sliding window on trials in which no stimulus was present. While there was no change in firing rate during the expected stimulus onset time (Fig. 4-10B, *top*), we observed an increase in the mean choice probability over the typical response window (*bottom*), suggesting that trial-to-trial fluctuations in firing rate over this brief time window were correlated with the decision.

We thus used this window for subsequent analyses. Similar to previous studies(Britten et al., 1996), we found a range of choice probabilities exhibited by individual MT neurons (Fig. 4-10C, *top*). On a neuron-by-neuron basis, the LFP model gave a remarkably good prediction of each neuron's CP (*bottom*, R=0.53, $p<1\times10^{-5}$, n=79), meaning that the trial-to-trial differences in the LFP can predict fluctuations in each neuron's firing rate that are correlated with the perceptual decision. Thus, this shows that elements of the LFP signals are predictive of perceptual decisions.

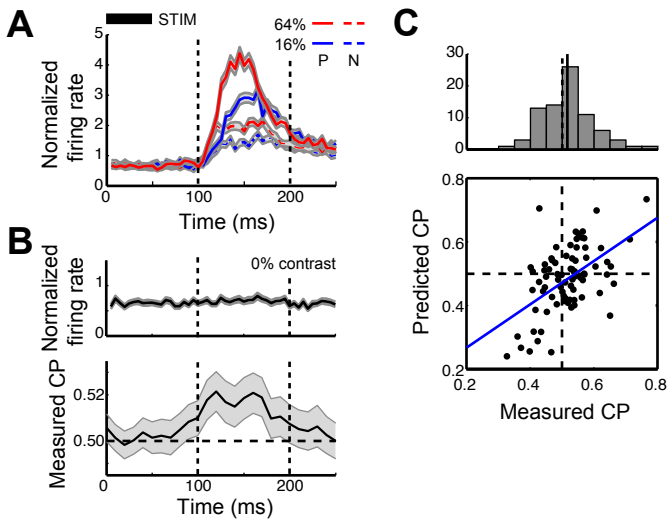

Figure 4-10 Prediction of choice probability (CP) of MT neurons
A. The average firing rate in response to the motion stimulus at the preferred ('P', solid) and null ('N', dashed) directions (n=79). Firing rate was normalized by the overall mean firing rate for each neuron. Shaded regions show the standard error of the mean across neurons. **B**. The average firing rate (*top*) and CP (*bottom*) on the zero-contrast trials (n=79). Significant CP was observed at the typical response window for trials with a stimulus (* p<0.05, ** p<0.01, *t-test*), despite the absence of a stimulus in the zero-contrast trials. **C.** The measured CP for each neuron (*horizontal axis*) compared to that predicted by the full model (*vertical axis*), demonstrating strong correlation CP (R=0.53, $p<5\times10^{-7}$, n=79). Each dot represents an individual cell. A histogram of measured CP is shown on the *top* (mean=0.52, n=79, p=0.032, *t-test*).

To determine which bands of the LFP contained choice-related information, we refitted the models with certain frequency bands omitted, and measured the impact on model performance. As with the LFP model components fit in the passive fixation task (Fig. 4-4E), removal of delta, beta and gamma bands had a large impact on cross-validated model performance (Fig. 4-11A). However, only the delta band was critical for predicting CP: removing the delta band significantly reduced the correlation coefficient between the measured and predicted CP (p<0.0013, two-tailed z-test, Fig. 4-11B).

To understand how the activity in the delta band could be used to predict CP, we compared the distributions of delta phase between trials at zero contrast in which the subject chose the preferred versus null motion direction (Fig. 4-11C). While these two distributions varied from experiment to experiment, in many cases they were clearly distinct from one another. Indeed, the delta band phases were more predictive of decisions than single neuron spiking activity (Fig. 4-11D, *bottom*, *paired t-test*, $p<10^{-11}$), and the choice-related signals in firing rates and delta phases were correlated across neurons (Fig. 4-11D, *top*, $R=0.39$, $p<0.001$).

How does delta band phase relate to CP in single neurons? For an example neuron with CP>0.5 (Fig. 4-11C, *top*), the delta phases on preferred trials were typically closer to the optimal delta phase of spiking as inferred from the LFP model fits, suggesting that output from the delta-band would result in elevated firing rate throughout the relevant time window for trials that the stimulus moved in the preferred direction. In contrast, for another neuron with CP<0.5 (Fig. 4-11C, *middle*), the delta phases on the null trials were closer to its optimal delta phase, and thus firing rates were lower for choices in the preferred direction. Finally, for a third example neuron with CP very close to 0.5, the distributions of delta phases were very similar on preferred trials and null trials (Fig. 4-11C, *bottom*). The link between delta band and CP in this case clearly relates to the time scale of the delta band itself (1-4 Hz). Delta band phase is slowly changing relative to the 100 ms counting window, meaning that the phase differences associated with preferred and null choices will

produce consistent firing rate modulation throughout the window, and thus a correlation with CP.

The observed CP can thus be predicted knowing each neuron's optimal phase in the delta band ϕ_0, and the locations of the means of the observed choice-related phase distributions for choices in the preferred and null directions (ϕ_{pref} and ϕ_{null}). The relative distances between ϕ_0 and ϕ_{pref} versus ϕ_{null} was indeed correlated with each neuron's observed CP (R = 0.37, p<0.001, Fig. 4-11E), with the quality of its prediction of CP nearly that derived from the full model. As expected, this predictor was valid using delta phase distributions for a range of times throughout the trial (Fig. 4-11F). Thus, this analysis demonstrates a clear link between the choice probability of MT neurons and network activity in the delta band, suggesting both salient network signatures of perception- and/or behavior-related modulation of neuron firing, and methods to uncover them.

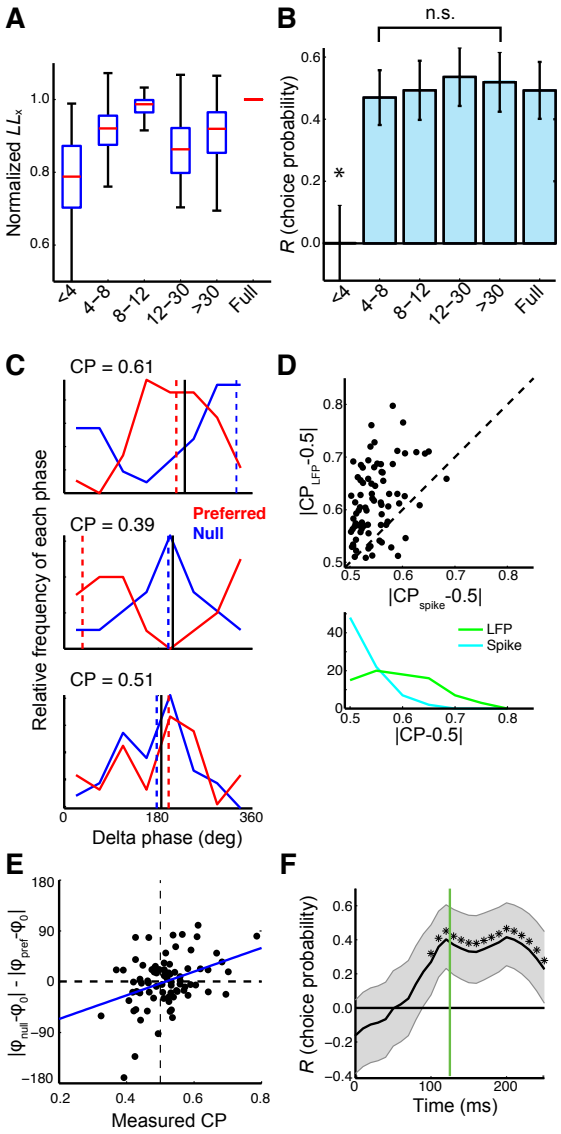

Figure 4-11 Choice probability (CP) of neurons predicted by delta band phase. **A.** The contribution of different frequency bands to model performance for the motion discrimination task (n = 79), computed in the same way as for the passive fixation task (Fig. 4-4E). **B.** Correlation between predicted and measured CP (*e.g.*, Fig. 4-6C) after removal of each frequency band from the model. Only removal of delta band affects prediction of CP (p<0.005, *two-tailed test*),. **C.** Histograms of the delta band phase measured 125 ms after the expected stimulus onset time for zero-contrast trials that the subject selected the preferred direction (red) versus the null direction (blue). The phase preference (dashed line) for three example neurons demonstrates why one neuron had a CP>0.5 (*top*), the other had a CP<0.5 (*middle*) and a third neuron had CP very close to 0.5 (*bottom*). **D.** The magnitudes of CP calculated using spike counts (*horizontal axis*) were compared with CP calculated from the delta band phase (*vertical axis*). Delta band phase is more predictive of choice than spikes (*paired t-test*, $p<10^{-11}$). **E.** Each neuron's CP can be predicted

using its preferred delta band phase ϕ_0 relative to the distributions of delta phases shown in (C). Comparison between the predicted CP using a simple metric (see *vertical axis*) and observed CP reveals a significant correlation (R = 0.37, p<0.005). **F.** Delta band phase reliably predicts CP over a range of times in the trial (*, p<0.05). Green line shows time of optimal prediction (125 ms), used for delta phase measurements in other parts of this Figure.

4.4 Discussion

While sensory neurons are typically studied with respect to their average responses to sensory stimuli, responses in the cortex can be highly variable across repeats of the same stimulus (Ermentrout et al., 2008; Goris et al., 2014; Masquelier, 2013). Variability can be indicative of noise that is disruptive to cortical function(Faisal et al., 2008), as well as unexplored aspects of cortical function that are not stimulus-locked (Reynolds and Chelazzi, 2004). Central to the question of function is how such variability is related to system-level activity, and whether it can be related to cognitive processes or behavior. We estimated that, on average, roughly half of the total rate modulation in neuronal responses in area MT is variable across trials, and showed that much of this variability can be predicted by incorporating network activity inferred from LFP recordings. Furthermore, we found that the relationship between MT neuron firing and network activity was relatively consistent across tasks, and could be used to predict the relationship between MT neuron activity and choice in a motion discrimination task.

In order to discover these relationships, we constructed a model that explained MT neuron responses using the stimulus, the LFP, and MUA. While all three signals are in principle correlated, the modeling framework only assigned weights to the signals

that could best predict MT neuron spikes. Thus, due to the effective stimulus-processing component of the model, the LFP component specifically identified contributions to MT neuron firing that were not stimulus-locked, and thus largely contributed to predicting variability. It also implicitly singled out particular frequency bands and depths that were most important at making these predictions. In contrast, given the LFP and stimulus-processing components, MUA from neighboring neurons did not contribute significantly to stimulus-locked or trial-variable predictions, even though there are raw correlations between all three signals(Burns et al., 2010). Thus, the modeling framework allowed us to isolate specific components of these signals that contributed to the trial-variable component of neural responses, ultimately allowing us to probe its relation to behavior during the motion discrimination task.

4.4.1 Relationships between LFPs, network activity, and cognitive processes

A fundamental barrier to the investigation of network influences on cortical activity is the difficulty in estimating a given cortical neuron's inputs, which will often come from a combination of nearby cortical neurons and those projecting from other brain areas. Here, we used recorded LFPs to infer the presence of network inputs relevant to recorded cortical neurons, following recent work (Ecker et al., 2014; Haslinger et al., 2012; Kelly et al., 2010). While LFPs have numerous sources, they are strongly influenced by synchronous synaptic inputs, and as a result can reflect inputs from other areas that are not necessarily reflected in local spiking activity (Buzsaki and Draguhn, 2004; Einevoll et al., 2013; Khawaja et al., 2009). We analyzed LFPs in the frequency domain, because different sources contributing to the LFP will often be

overlapping in time but exhibit dynamics on different time-scales (Buzsaki and Draguhn, 2004). Furthermore, particular oscillation frequencies have been attributed to the activation of different cortical microcircuits, which can have distinctive signatures as a function of depth within the cortical column (de Cheveigne et al., 2013; Schroeder and Lakatos, 2009; Xing et al., 2009). The LFP components of our model thus use separate cues available across frequency and depth to best infer the components of network activity relevant to a given cortical neuron.

The presence of power in different frequency bands in both LFP and EEG recordings has been associated with a variety of processes related to cognitive states, as well as behavioral and task-related variables (Buzsaki and Draguhn, 2004; Fries, 2005). In this work, we identify frequency bands of the LFP that are relevant for predicting neural activity that is not explained by stimulus processing. This analysis goes well beyond simple analyses of spike-LFP coupling such as coherence and phase-locking (Fries, 2005), in that it is not confounded by correlations such as those between the stimulus, other neuron activity, and LFPs, nor correlations across frequency bands within the LFP. As a result, our models typically focused on a small subset of frequency bands and recorded depths (relative to measures of spike-LFP coherence), and largely capture elements of the LFPs that are specifically predictive of the trial-variable component of the neurons' responses.

Furthermore, we found that coupling between network activity and MT neuron firing was consistent across the conditions tested, including passive fixation in the presence

of ongoing visual stimulation, a discrimination task with stimuli presented, and behavioral choice without visual stimulation (the zero contrast condition). This suggests that a given neuron's coupling with network activity may be relatively constant, while fluctuations in the amount and type of network activity, as well as of feedforward visual stimulation, may make it appear otherwise (Tan et al., 2014). Such observations suggest the exciting possibility that these results reflect mechanisms by which other brain areas can coordinate local populations of sensory neurons for task-specific reasons (Schroeder and Lakatos, 2009).

4.4.2 Choice probability and network activity

During tasks involving perceptual decisions, the responses of cortical neurons are correlated with the subject's choice, even if the stimulus is ambiguous or absent. This is commonly referred to as choice probability (CP), which measures how well an ideal observer can predict the subject's choice using the neuron response. The correlation between neuronal variability and perceptual decisions has been observed in a number of sensory areas (Nienborg et al., 2012). In area MT, choice-related variability has been found in direction, speed, and disparity discrimination tasks (Britten et al., 1996; Liu and Newsome, 2006; Uka and DeAngelis, 2004).

The origin of CP is highly debated (Nienborg et al., 2012). Originally, it was expected that CP arose because the subject based its decisions on the responses of sensory neurons that were selective for the relevant stimuli (*e.g.*, MT neurons for motion discrimination). As a result, noise in these neurons was expected to propagate up to

decision-making areas of the brain and is thus correlated with the behavioral outcome (Shadlen et al., 1996). However, other effects, such as microsaccades, can also contribute to the correlation between neural activity and behavior (Herrington et al., 2009), although this is unlikely the cause for CP in this study because we used the zero-contrast stimulus condition to measure CP and did not observe a difference in the microsaccade rate between the preferred trials and null trials in any of the experiments (p>0.05, n=22).

Recent work has called into question the directionality of causation, and suggests an alternative explanation where other factors could influence both the decision and the neural activity (Nienborg and Cumming, 2009). Our results are thus more consistent with this latter alternative, insofar as they link this variability related to choice to observable network-level activity. We cannot infer where such network influences arise based on this study, but by identifying their network signatures in the LFP, we hope to inform future work addressing the origin and functional role of such cortical modulation.

Chapter 5 : Conclusions

In this dissertation, I have developed a set of novel statistical models to explain sensory neuron response properties at various stages of the mammalian visual system. These models incorporate knowledge of the anatomy and physiology of the underlying neural circuits, with individual model components directly linked to physiological interpretations. As a result, these models not only provide unprecedented accuracy in predicting the neural response, but also suggest how neural computation is constructed from specific elements of the underlying circuit. The results provide experimentally testable hypotheses and shed light on the nature of neural computations across the visual system.

In this chapter, I summarize several important conceptual and methodological advances of the work presented in this dissertation.

5.1 Using nonlinear hierarchical models to characterize neuron response

A central hypothesis in my work is that neural computation in the sensory system emerges from nonlinear, hierarchical processing. Sensory processing involves nonlinear computations performed on the stimuli. However, characterizations of sensory neurons still typically rely on the assumption of linear stimulus processing due to the difficulty of construction and estimation of nonlinear models. Although the space of possible nonlinear models is unbounded, we may identify relevant types of

nonlinearity based on known physiological properties of neurons and anatomy of the neural circuits. For example, one dominant source of nonlinearity in the nervous system is spike generation, which results in both rectification of the inputs to the neuron (which is usually from other spiking neurons), as well as rectification of the neuron's output (McFarland et al., 2013). Other sources of nonlinearity include spike refractoriness (Berry and Meister, 1998), synaptic facilitation and depression (Ozuysal and Baccus, 2012; Tsodyks et al., 1998) , and presynaptic suppression (Euler et al., 2014). The models developed in this work incorporated these nonlinear properties when appropriate. By doing so, the models are able to successfully characterize complex neural response properties, such as temporal precision of spike and synaptic current response (Chapter 2), and selectivity to multiple stimulus features (Chapter 3), which cannot be explained in linear models (as discussed in Chapter 1).

Hierarchical processing is another important concept that is emphasized throughout this dissertation. The key principle of hierarchical processing is that the RFs of neurons at one level of the hierarchy are constructed by first pooling a large number of inputs from neurons at a lower level, and then processed the combined signal with a set of canonical nonlinear functions, most notably rectification due to spike generation. I incorporated the principle of hierarchical processing into the models in each project, utilizing knowledge of upstream processing. In Chapter 3, I assumed MT neurons receive inputs from a large number of V1 neurons with much smaller receptive fields. Each V1-like subunit is characterized with the product of a direction-

tuning function and speed tuning function based on physiology experiments. Similar hierarchical models have been developed for other visual cortical areas (Lochmann et al., 2013; Mineault et al., 2012; Rust et al., 2006; Willmore et al., 2010).

5.2 Tailored nonlinear models based on knowledge of specific neural circuits

In Chapter 2, I developed a two-stage nonlinear model to characterize the underlying computation represented by retinal ganglion cell responses, with both processing stages constrained by experimental data. The first stage represents upstream processing at the level of inputs from bipolar cells, and is fit to the observed excitatory synaptic currents recorded from retinal ganglion cells. The model structure reflects known anatomical and physiological properties of the retinal circuit, and suggests that presynaptic inhibition mediated by amacrine cells as an important component underlying ganglion cells computation. These results thus provide a set of testable predictions that guide future studies.

5.3 Identification of signatures of network activity related to cortical variability

The first half of the dissertation focused on feedforward stimulus processing in the visual system (Chapters 2-3). While analyzing cortical neural responses over repeated presentation of the same stimulus, we noticed a large amount of variability. This was consistent with many previous studies, which showed that there is an increase of

response variability along the visual hierarchy (Scholvinck et al., 2015; Vogel et al., 2005). Understanding the variability of cortical neuron response is an important problem in system neuroscience. In Chapter 4, I showed that a significant fraction of the cortical variability could be explained by distinct elements of the ongoing network activity reflected in local field potentials (LFPs) and multi-unit activity, thus establishing a direct link between cortical variability and ongoing network activity. This LFP-modeling framework can be readily applied to other cortical areas under different behavior conditions, thus representing a versatile tool to study the relationship between neural variability and cortical network activity.

5.4 Linking single neural response and network activity with perceptual behaviors

In the second half of Chapter 4, I studied the relationship between simultaneously recorded neuron responses, network activity, and perceptual decisions, while the subject actively participated in a motion discrimination task. In contrast to most studies of the relationship between neural response and perceptual decisions, I included into the model framework ongoing cortical network activity, which likely reflects non-stimulus driven inputs to cortical neurons. By doing so, I identified a novel choice-related component of the network activity, and showed that the widely studied correlation between single neuron response and perceptual decisions (i.e., choice probability) could be well predicted by network activity. This result has profound implications for further study of choice probability.

Glossary

CP	Choice Probability
DS	Direction-selective
fMRI	Functional Magnetic Resonance Imaging
LFP	Local Field Potential
LGN	Lateral Geniculate Nucleus
LL	Log-likelihood
MID	Maximally Informative Dimensions
MT	Middle Temporal Area
MST	Medial Superior Temporal Area
MUA	Multi-Unit Activity
NIM	Nonlinear Input Model
RF	Receptive Field
RGC	Retinal Ganglion Cell
ROC	Receiver Operating Characteristic
S1	Primary Somatosensory Cortex
S2	Second Somatosensory Cortex
STA	Spike Triggered Average
STC	Spike Triggered Covariance
V1	Primary Visual Area

Bibliography

Abbott, L.F., Varela, J.A., and Sen, K. (1997). Synaptic depression and cortical gain control. Science 275, 220-224.

Adelson, E.H. (2001). On seeing stuff: the perception of materials by humans and machines. In Proceedings SPIE human vision and electronic imaging VI, R. BE, and P. TN, eds., pp. 1-12.

Adelson, E.H., and Bergen, J.R. (1985). Spatiotemporal energy models for the perception of motion. J Opt Soc Am A 2, 284-299.

Ahmed, B., Cordery, P.M., McLelland, D., Bair, W., and Krug, K. (2012). Long-range clustered connections within extrastriate visual area V5/MT of the rhesus macaque. Cereb Cortex 22, 60--73.

Ahrens, M.B., Linden, J.F., and Sahani, M. (2008a). Nonlinearities and contextual influences in auditory cortical responses modeled with multilinear spectrotemporal methods. J Neurosci 28, 1929-1942.

Ahrens, M.B., Paninski, L., and Sahani, M. (2008b). Inferring input nonlinearities in neural encoding models. Network 19, 35-67.

Albright, T.D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. J Neurophysiol 52, 1106-1106.

Allman, J., Miezin, F., and McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. Annu Rev Neurosci 8, 407-430.

Alonso, J.-M., and Yao, C. (2009). Receptive field. Scholarpedia 4, 5393.

Andersen, R.A., and Bradley, D.C. (1998). Perception of three-dimensional structure from motion. Trends Cogn Sci 2, 222-228.

Arieli, A., Sterkin, A., Grinvald, A., and Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. Science 273, 1868-1871.

Avery, M.C., Dutt, N., and Krichmar, J.L. (2014). Mechanisms underlying the basal forebrain enhancement of top-down and bottom-up attention. Eur J Neurosci 39, 852--865.

Axelrod, V., and Yovel, G. (2012). Hierarchical processing of face viewpoint in human visual cortex. J Neurosci 32, 2442-2452.

Ayaz, A., Saleem, A.B., Schlvinck, M.L., and Carandini, M. (2013). Locomotion controls spatial integration in mouse visual cortex. Curr Biol 23, 890--894.

Baccus, S.A. (2007). Timing and computation in inner retinal circuitry. Annu Rev Physiol 69, 271-290.

Baccus, S.A., and Meister, M. (2002). Fast and slow contrast adaptation in retinal circuitry. Neuron 36, 909-919.

Bair, W., Cavanaugh, J.R., and Movshon, J.A. (1997). Reconstructing stimulus velocity from neuronal responses in area MT. In Advances in Neural Information Processing Systems, M. Mozer, M.I. Jordan, and C.C.H. Petersen, eds. (Cambridge: MA, MIT Press), pp. 34--40.

Bair, W., Zohary, E., and Newsome, W.T. (2001). Correlated firing in macaque visual area MT: time scales and relationship to behavior. J Neurosci 21, 1676-1697.

Banerjee, A., Dean, H.L., and Pesaran, B. (2012). Parametric models to relate spike train and LFP dynamics with neural information processing. Front Comput Neurosci 6, 51.

Beauchemin, S.S., and Barron, J.L. (1995). The computation of optical flow. ACM Computing Surveys 27, 433-466.

Beaudoin, D.L., Borghuis, B.G., and Demb, J.B. (2007). Cellular basis for contrast gain control over the receptive field center of mammalian retinal ganglion cells. Journal of Neuroscience 27, 2636-2645.

Berkes, P., Orbán, G., Lengyel, M., and Fiser, J. (2011). Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. Science 331, 83-87.

Berry, M.J., and Meister, M. (1998). Refractoriness and neural precision. J Neurosci 18, 2200-2211.

Berry, M.J., Warland, D.K., and Meister, M. (1997). The structure and precision of retinal spike trains. Proc Natl Acad Sci U S A 94, 5411-5416.

Boerlin, M., Machens, C.K., and Deneve, S. (2013). Predictive coding of dynamical variables in balanced spiking networks. PLoS Comput Biol 9, e1003258.

Bondy, A., and Cummings, B. (2013). Top down signals influence the distribution of noise correlations amongst sensory neurons. In Society for Neuroscience Annual Meeting (San Diego), p. 311.302.

Bonin, V., Mante, V., and Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus. J Neurosci 25, 10844-10856.

Borghuis, B.G., Marvin, J.S., Looger, L.L., and Demb, J.B. (2013). Two-photon imaging of nonlinear glutamate release dynamics at bipolar cell synapses in the mouse retina. J Neurosci 33, 10972-10985.

Born, R.T. (2000). Center-surround interactions in the middle temporal visual area of the owl monkey. J Neurophysiol 84, 2658-2669.

Born, R.T., and Bradley, D.C. (2005). Structure and function of visual area MT. Annu Rev Neurosci 28, 157-189.

Bradley, D.C., and Andersen, R.A. (1998). Center-surround antagonism based on disparity in primate area MT. J Neurosci 18.

Bradley, D.C., Chang, G.C., and Andersen, R.A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. Nature 392, 714-717.

Bradley, D.C., and Goyal, M.S. (2008). Velocity computation in the primate visual system. Nat Rev Neurosci 9, 686-695.

Brainard, D.H. (1997). The Psychophysics Toolbox. Spat Vis 10, 433-436.

Britten, K.H., and Heuer, H.W. (1999). Spatial summation in the receptive fields of MT neurons. J Neurosci 19, 5074-5084.

Britten, K.H., Newsome, W.T., Shadlen, M.N., Celebrini, S., and Movshon, J.A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. Vis Neurosci 13, 87-100.

Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. J Neurosci 12, 4745-4765.

Britten, K.H., and van Wezel, R.J. (1998). Electrical microstimulation of cortical area MST biases heading perception in monkeys. Nat Neurosci 1, 59-63.

Buracas, G.T., and Albright, T.D. (1996). Contribution of area MT to perception of three-dimensional shape: a computational study. Vision Res 36, 869-887.

Burns, S.P., Xing, D., and Shapley, R.M. (2010). Comparisons of the dynamics of local field potential and multiunit activity signals in macaque visual cortex. J Neurosci 30, 13739-13749.

Busch, N.A., Dubois, J., and VanRullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. J Neurosci 29, 7869--7876.

Butts, D.A., Desbordes, G., Weng, C., Jin, J., Alonso, J.M., and Stanley, G.B. (2010). The episodic nature of spike trains in the early visual pathway. J Neurophysiol 104, 3371-3387.

Butts, D.A., Weng, C., Jin, J., Alonso, J.M., and Paninski, L. (2011). Temporal precision in the visual pathway through the interplay of excitation and. J Neurosci 31, 11313-11327.

Butts, D.A., Weng, C., Jin, J., Yeh, C.I., Lesica, N.A., Alonso, J.M., and Stanley, G.B. (2007). Temporal precision in the neural code and the timescales of natural vision. Nature 449, 92-95.

Buzsaki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. Science 304, 1926-1929.

Cai, D., DeAngelis, G.C., and Freeman, R.D. (1997). Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kittens. J Neurophysiol 78, 1045-1061.

Carandini, M., Demb, J.B., Mante, V., Tolhurst, D.J., Dan, Y., Olshausen, B.A., Gallant, J.L., and Rust, N.C. (2005). Do we know what the early visual system does? J Neurosci 25, 10577-10597.

Carandini, M., and Heeger, D.J. (2012). Normalization as a canonical neural computation. Nat Rev Neurosci 13, 51-62.

Carandini, M., Heeger, D.J., and Movshon, J.A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. J Neurosci 17, 8621-8644.

Carandini, M., Horton, J.C., and Sincich, L.C. (2007). Thalamic filtering of retinal spike trains by postsynaptic summation. J Vis 7, 20 21-11.

Casti, A., Hayot, F., Xiao, Y., and Kaplan, E. (2008). A simple model of retina-LGN transmission. J Comput Neurosci 24, 235-252.

Celebrini, S., and Newsome, W.T. (1994). Neuronal and psychophysical sensitivity to motion signals in extrastriate area MST of the macaque monkey. J Neurosci 14, 4109--4124.

Chakravarthi, R., and VanRullen, R. (2012). Conscious updating is a rhythmic process. In Proc Natl Acad Sci U S A (United States), pp. 10599-10604.

Chander, D., and Chichilnisky, E.J. (2001). Adaptation to temporal contrast in primate and salamander retina. J Neurosci 21, 9904-9916.

Chichilnisky, E.J. (2001). A simple white noise analysis of neuronal light responses. Network 12, 199-213.

Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. In Nat Neurosci, pp. 811-819.

Cohen, M.R., and Maunsell, J.H. (2009). Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci 12, 1594-1600.

Cohen, M.R., and Newsome, W.T. (2004). What electrical microstimulation has revealed about the neural basis of cognition. In Curr Opin Neurobiol (England), pp. 169-177.

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

Croner, L.J., and Albright, T.D. (1999). Segmentation by color influences responses of motion-sensitive neurons in the cortical middle temporal visual area. J Neurosci 19, 3935--3951.

Cui, Y., Liu, D., McFarland, J., Pack, C., and Butts, D. (2014). The relationship between choice-related variability of MT neuron responses and ongoing network activity. In Society for Neuroscience Annual Meeting (Washington DC).

Cui, Y., Liu, D.L., McFarland, J.M., Pack, C.P., and Butts, D.A. (2013a). Modulation of stimulus processing by network activity across cortical depth in MT. In Society for Neuroscience Annual Meeting (San Diego).

Cui, Y., Liu, D.L., McFarland, J.M., Pack, C.P., and Butts, D.A. (2015). Predicting the relationship between sensory neuron variability and perceptual decisions. In Cosyne Abstract (Salt Lake City).

Cui, Y., Liu, L.D., Khawaja, F.A., Pack, C.C., and Butts, D.A. (2013b). Diverse suppressive influences in area MT and selectivity to complex motion features. J Neurosci 33, 16715-16728.

Cui, Y., Wang, Y.V., Demb, J.B., and Butts, D.A. (2013c). Evidence of presynaptic inhibition in shaping retinal processing. In Cosyne Abstract (Salt Lake City).

Czanner, G., Eden, U.T., Wirth, S., Yanike, M., Suzuki, W.A., and Brown, E.N. (2008). Analysis of between-trial and within-trial neural spiking dynamics. J Neurophysiol 99, 2672-2693.

David, S.V., and Gallant, J.L. (2005). Predicting neuronal responses during natural vision. Network 16, 239-260.

Dayan, P., and Abbott, L. (2005). Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems. (Boston, MA, The MIT press).

de Cheveigne, A., Edeline, J.M., Gaucher, Q., and Gourevitch, B. (2013). Component analysis reveals sharp tuning of the local field potential in the guinea pig auditory cortex. J Neurophysiol 109, 261-272.

de Lafuente, V., and Romo, R. (2006). Neural correlate of subjective sensory experience gradually builds up across cortical areas. In Proc Natl Acad Sci U S A (United States), pp. 14266-14271.

de Ruyter van Steveninck, R., and Bialek, W. (1988). Real-time performance of a movement-sensitive neuron in the blowfly visual system: coding and information transfer in short spike sequences. Proc R Soc Lond B 234, 379-414.

DeAngelis, G.C., Ohzawa, I., and Freeman, R.D. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. J Neurophysiol 69, 1118-1135.

Demb, J.B. (2008). Functional circuitry of visual adaptation in the retina. J Physiol 586, 4377-4384.

Deweese, M. (1996). Optimization principles for the neural code. Network 7, 325--331.

DiCarlo, J.J., and Cox, D.D. (2007). Untangling invariant object recognition. Trends Cogn Sci 11, 333-341.

DiCarlo, J.J., Johnson, K.O., and Hsiao, S.S. (1998). Structure of receptive fields in area 3b of primary somatosensory cortex in the alert monkey. J Neurosci 18, 2626--2645.

Duffy, C.J., and Wurtz, R.H. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. J Neurosci 15, 5192-5208.

Dugue, L., Marque, P., and VanRullen, R. (2011). The phase of ongoing oscillations mediates the causal relation between brain excitation and visual perception. In J Neurosci (United States), pp. 11889-11893.

Dürsteler, M.R., and Wurtz, R.H. (1988). Pursuit and optokinetic deficits following chemical lesions of cortical areas MT and MST. J Neurophysiol 60, 940--965.

Ecker, A.S., Berens, P., Cotton, R.J., Subramaniyan, M., Denfield, G.H., Cadwell, C.R., Smirnakis, S.M., Bethge, M., and Tolias, A.S. (2014). State dependence of noise correlations in macaque primary visual cortex. Neuron 82, 235-248.

Eggermont, J.J. (1993). Wiener and Volterra analyses applied to the auditory system. Hear Res 66, 177--201.

Eggermont, J.J., Johannesma, P.M., and Aertsen, A.M. (1983). Reverse-correlation methods in auditory research. Q Rev Biophys 16, 341--414.

Einevoll, G.T., Kayser, C., Logothetis, N.K., and Panzeri, S. (2013). Modelling and analysis of local field potentials for studying the function of cortical circuits. Nat Rev Neurosci 14, 770-785.

Engbert, R., and Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. Proc Natl Acad Sci U S A 103, 7192-7197.

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

Ermentrout, G.B., Galan, R.F., and Urban, N.N. (2008). Reliability, synchrony and noise. Trends Neurosci 31, 428-434.

Euler, T., Haverkamp, S., Schubert, T., and Baden, T. (2014). Retinal bipolar cells: elementary building blocks of vision. Nat Rev Neurosci 15, 507-519.

Fairhall, A.L., Burlingame, C.A., Narasimhan, R., Harris, R.A., Puchalla, J.L., and Berry, M.J., 2nd (2006). Selectivity for multiple stimulus features in retinal ganglion cells. J Neurophysiol 96, 2724-2738.

Faisal, A.A., Selen, L.P., and Wolpert, D.M. (2008). Noise in the nervous system. Nat Rev Neurosci 9, 292-303.

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

Fermüller, C., and Aloimonos, Y. (1995). Direct perception of three-dimensional motion from patterns of visual motion. Science 270, 1973-1976.

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

Fiser, J., Chiu, C., and Weliky, M. (2004). Small modulation of ongoing cortical dynamics by sensory input during natural vision. In Nature (England), pp. 573-578.

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

Fries, P., Reynolds, J.H., Rorie, A.E., and Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. Science 291, 1560-1563.

Fukushima, K. (1980). Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36, 193-202.

Gaudry, K.S., and Reinagel, P. (2007a). Benefits of contrast normalization demonstrated in neurons and model cells. J Neurosci 27, 8071-8079.

Gaudry, K.S., and Reinagel, P. (2007b). Contrast adaptation in a nonadapting LGN model. J Neurophysiol 98, 1287-1296.

Gautama, T., and Van Hulle, M.M. (2001). Function of center-surround antagonism for motion in visual area MT/V5: a modeling study. Vision Res 41, 3917-3930.

Geesaman, B.J., Born, R.T., Andersen, R.A., and Tootell, R.B. (1997). Maps of complex motion selectivity in the superior temporal cortex of the alert macaque monkey: a double-label 2-deoxyglucose study. Cereb Cortex 7, 749-757.

Geisler, W.S., and Albrecht, D.G. (1997). Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. In Vis Neurosci (England), pp. 897-919.

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

Gollisch, T. (2013). Features and functions of nonlinear spatial integration by retinal ganglion cells. J Physiol Paris 107, 338-348.

Gollisch, T., and Meister, M. (2010). Eye smarter than scientists believed: neural computations in circuits of the retina. Neuron 65, 150-164.

Goris, R.L., Movshon, J.A., and Simoncelli, E.P. (2014). Partitioning neuronal variability. Nat Neurosci 17, 858-865.

Greschner, M., Shlens, J., Bakolitsa, C., Field, G.D., Gauthier, J.L., Jepson, L.H., Sher, A., Litke, A.M., and Chichilnisky, E.J. (2011). Correlated firing among major ganglion cell types in primate retina. In J Physiol (England), pp. 75-86.

Gu, Y., DeAngelis, G.C., and Angelaki, D.E. (2007). A functional link between area MSTd and heading perception based on vestibular signals. In Nat Neurosci (United States), pp. 1038-1047.

Haefner, R.M., Berkes, P., and Fiser, J. (2014). Perceptual decision-making as probabilistic inference by neural sampling. arXiv 1409, 0257.

Haefner, R.M., Gerwinn, S., Macke, J.H., and Bethge, M. (2013). Inferring decoding strategies from choice probabilities in the presence of correlated variability. In Nat Neurosci (United States), pp. 235-242.

Haider, B., Krause, M.R., Duque, A., Yu, Y., Touryan, J., Mazer, J.A., and McCormick, D.A. (2010). Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. Neuron 65, 107--121.

Hammond, P., and Reck, J. (1980). Influence of velocity on directional tuning of complex cells in cat striate cortex for texture motion. Neurosci Lett 19, 309-314.

Hao, J., Wang, X.-d., Dan, Y., Poo, M.-m., and Zhang, X.-h. (2009). An arithmetic rule for spatial summation of excitatory and inhibitory inputs in pyramidal neurons. Proceedings of the National Academy of Sciences of the United States of America 106, 21906-21911.

Harris, K.D., and Thiele, A. (2011). Cortical state and attention. Nat Rev Neurosci 12, 509--523.

Haslinger, R., Pipa, G., Lima, B., Singer, W., Brown, E.N., and Neuenschwander, S. (2012). Context matters: the illusive simplicity of macaque V1 receptive fields. PLoS One 7, e39699.

Heeger, D.J. (1992). Normalization of cell responses in cat striate cortex. In Vis Neurosci (England), pp. 181-197.

Heeger, D.J., Boynton, G.M., Demb, J.B., Seidemann, E., and Newsome, W.T. (1999). Motion opponency in visual cortex. J Neurosci 19, 7162-7174.

Herikstad, R., Jonathan, B., Jean-Philippe, L., M., G.C., and Shih-Cheng, Y. (2011). Natural movies evoke spike trains with low spike time variability in cat primary visual cortex. J Neurosci 31, 15844--15860.

Herrero, J.L., Gieselmann, M.A., Sanayei, M., and Thiele, A. (2013). Attention-induced variance and noise correlation reduction in macaque V1 is mediated by NMDA receptors. Neuron 78, 729-739.

Herrington, T.M., Masse, N.Y., Hachmeh, K.J., Smith, J.E., Assad, J.A., and Cook, E.P. (2009). The effect of microsaccades on the correlation between neural activity and behavior in middle temporal, ventral intraparietal, and lateral intraparietal areas. J Neurosci 29, 5793-5805.

Hubel, D.H., and Wiesel, T.N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of physiology 160, 106-154.

Hubel, D.H., and Wiesel, T.N. (2012). David Hubel and Torsten Wiesel. Neuron 75, 182--184.

Hunter, I.W., and Korenberg, M.J. (1986). The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. Biol Cybern 55, 135-144.

Hunter, J.N., and Born, R.T. (2011). Stimulus-Dependent Modulation of Suppressive Influences in MT. J Neurosci 31, 678-678.

Isaacson, J.S., and Scanziani, M. (2011). How inhibition shapes cortical activity. Neuron 72, 231--243.

Jarsky, T., Cembrowski, M., Logan, S.M., Kath, W.L., Riecke, H., Demb, J.B., and Singer, J.H. (2011). A synaptic mechanism for retinal adaptation to luminance and contrast. J Neurosci 31, 11003-11015.

Jazayeri, M., Wallisch, P., and Movshon, J.A. (2012). Dynamics of Macaque MT Cell Responses to Grating Triplets. J Neurosci 32, 8242-8253.

Kaas, J.H. (1989). Why does the brain have so many visual areas? J Cogn Neurosci 1, 121-135.

Kaneko, A. (1973). Receptive field organization of bipolar and amacrine cells in the goldfish retina. J Physiol 235, 133--153.

Kara, P., Reinagel, P., and Reid, R.C. (2000). Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. Neuron 27, 635-646.

Kastner, D.B., and Baccus, S.A. (2013). Spatial segregation of adaptation and predictive sensitization in retinal ganglion cells. Neuron 79, 541-554.

Kastner, D.B., Baccus, S.A., and Sharpee, T.O. (2015). Critical and maximally informative encoding between neural populations in the retina. In Proc Natl Acad Sci U S A (United States), pp. 2533-2538.

Kayser, C., Logothetis, N.K., and Panzeri, S. (2010). Millisecond encoding precision of auditory cortex neurons. Proc Natl Acad Sci U S A 107, 16976--16981.

Kayser, C., Salazar, R.F., and Konig, P. (2003). Responses to natural scenes in cat V1. J Neurophysiol 90, 1910-1920.

Keat, J., Reinagel, P., Reid, R.C., and Meister, M. (2001). Predicting every spike: a model for the responses of visual neurons. Neuron 30, 803-817.

Keller, G.B., Tobias, B., and Mark, H. (2012). Sensorimotor mismatch signals in primary visual cortex of the behaving mouse. Neuron 74, 809--815.

Kelly, R.C., and Lee, T.S. (2004). Decoding V1 Neuronal Activity using Particle Filtering with Volterra Kernels. 15 1359--1366.

Kelly, R.C., Smith, M.A., Kass, R.E., and Lee, T.S. (2010). Local field potentials indicate network state and account for neuronal response variability. J Comput Neurosci 29, 567-579.

Khawaja, F.A., Tsui, J.M., and Pack, C.C. (2009). Pattern motion selectivity of spiking outputs and local field potentials in macaque visual cortex. J Neurosci 29, 13702-13709.

Kim, K.J., and Rieke, F. (2001). Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. J Neurosci 21, 287-299.

Kim, K.J., and Rieke, F. (2003). Slow Na+ inactivation and variance adaptation in salamander retinal ganglion cells. J Neurosci 23, 1506-1516.

Kisley, M.A., and Gerstein, G.L. (1999). Trial-to-trial variability and state-dependent modulation of auditory-evoked responses in cortex. J Neurosci 19, 10451-10460.

Ko, H., Hofer, S.B., Pichler, B., Buchanan, K.A., Sjöström, P.J., and Mrsic-Flogel, T.D. (2011). Functional specificity of local synaptic connections in neocortical networks. Nature 473, 87-91.

Kumbhani, R.D., Nolt, M.J., and Palmer, L.A. (2007). Precision, reliability, and information-theoretic analysis of visual thalamocortical neurons. In J Neurophysiol (United States), pp. 2647-2663.

Lagae, L., Maes, H., Raiguel, S., Xiao, D.K., and Orban, G.A. (1994). Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. J Neurophysiol 71, 1597-1626.

Lakatos, P., Shah, A.S., Knuth, K.H., Ulbert, I., Karmos, G., and Schroeder, C.E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. J Neurophysiol 94, 1904-1911.

Laplace, P. (1825). Essai Philosophique sur les Probabilités (Gauthier- Villars.).

Lau, B., Stanley, G.B., and Dan, Y. (2002). Computational subunits of visual cortical neurons revealed by artificial neural networks. Proc Natl Acad Sci U S A 99, 8974-8979.

Law, C.-T., and I., G.J. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. Nat Neurosci 11, 505--513.

Lecun, Y. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE 86, 2278-2324.

Lee, H., Simpson, G.V., Logothetis, N.K., and Rainer, G. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. Neuron 45, 147-156.

Levinson, E., and Sekuler, R. (1975). The independence of channels in human vision selective for direction of movement. J Physiol 250.

Lin, B., and Masland, R.H. (2006). Populations of wide-field amacrine cells in the mouse retina. J Comp Neurol 499, 797--809.

Lisberger, S.G., and Movshon, J.A. (1999). Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. J Neurosci 19, 2224-2246.

Liu, J., and Newsome, W.T. (2006). Local field potential in cortical area MT: stimulus tuning and behavioral correlations. J Neurosci 26, 7779-7790.

Liu, S., Gu, Y., DeAngelis, G.C., and Angelaki, D.E. (2013). Choice-related activity and correlated noise in subcortical vestibular neurons. Nat Neurosci 16, 89-97.

Lochmann, T., Blanche, T.J., and Butts, D.A. (2013). Construction of direction selectivity through local energy computations in primary visual cortex. PLoS One 8, e58666.

Mainen, Z.F., and Sejnowski, T.J. (1995). Reliability of spike timing in neocortical neurons. Science 268, 1503-1506.

Manookin, M.B., and Demb, J.B. (2006). Presynaptic mechanism for slow contrast adaptation in mammalian retinal ganglion cells. Neuron 50, 453-464.

Mante, V., Bonin, V., and Carandini, M. (2008). Functional mechanisms shaping lateral geniculate responses to artificial and natural stimuli. In Neuron (United States), pp. 625-638.

Marmarelis, P.Z., and Marmarelis, V.Z. (1978). Analysis of physiological systems: The white-noise approach (Plenum Press).

Marr, D. (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information (MIT Press).

Martinez-Conde, S., Otero-Millan, J., and Macknik, S.L. (2013). The impact of microsaccades on vision: towards a unified theory of saccadic function. Nat Rev Neurosci 14, 83-96.

Masland, R.H. (2001). The fundamental plan of the retina. Nature neuroscience 4, 877-886.

Masquelier, T. (2013). Neural variability, or lack thereof. Front Comput Neurosci 7, 7.

Maunsell, J.H., and Newsome, W.T. (1987). Visual processing in monkey extrastriate cortex. Annu Rev Neurosci 10, 363-401.

Maunsell, J.H., and Van Essen, D.C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. J Neurophysiol 49, 1148-1167.

McFarland, J.M., Cui, Y., and Butts, D.A. (2013). Inferring nonlinear neuronal computation based on physiologically plausible inputs. PLoS Comput Biol 9, e1003143.

McIlwain, J. (1996). An Introduction to the Biology of Vision (Cambridge, UK: Cambridge University Press).

Meister, M., and Berry, M.J. (1999). The neural code of the retina. Neuron 22, 435-450.

Mendoza-Halliday, D., Torres, S., and Martinez-Trujillo, J.C. (2014). Sharp emergence of feature-selective sustained activity along the dorsal visual pathway. Nat Neurosci 17, 1255-1262.

Mikami, A., Newsome, W.T., and Wurtz, R.H. (1986). Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. J Neurophysiol 55, 1308-1308.

Miller, R.F., Staff, N.P., and Velte, T.J. (2006). Form and function of ON-OFF amacrine cells in the amphibian retina. J Neurophysiol 95, 3171--3190.

Mineault, P.J., Khawaja, F.A., Butts, D.A., and Pack, C.C. (2012). Hierarchical processing of complex motion along the primate dorsal visual pathway. Proc Natl Acad Sci U S A 109, E972-980.

Mitchell, J.F., Sundberg, K.A., and Reynolds, J.H. (2009). Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. Neuron 63, 879--888.

Miyawaki, Y., Hajime, U., Okito, Y., Masa-aki, S., Yusuke, M., C., T.H., Norihiro, S., and Yukiyasu, K. (2008). Visual image reconstruction from human brain activity using a combination of multiscale local image decoders. Neuron 60, 915--929.

Morishima, Y., Akaishi, R., Yamada, Y., Okuda, J., Toma, K., and Sakai, K. (2009). Task-specific signal transmission from prefrontal cortex in visual selective attention. Nat Neurosci 12, 85-91.

Movshon, J.A., Adelson, E.H., Gizzi, M.S., and Newsome, W.T. (1986). The analysis of moving visual patterns. In *Experimental Brain Research Supplementum II: Pattern recognition mechanisms*, R.G. C. Chagas, C.Gross, ed. (Springer, New York), pp. 117-151.

Movshon, J.A., and Newsome, W.T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. J Neurosci 16, 7733-7741.

Newsome, W.T., Britten, K.H., and Movshon, J.A. (1989). Neuronal correlates of a perceptual decision. Nature 341, 52-54.

Newsome, W.T., and Pare, E.B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). J Neurosci 8, 2201-2211.

Nguyenkim, J.D., and DeAngelis, G.C. (2003). Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. J Neurosci 23.

Nicholson, C., and Freeman, J.A. (1975). Theory of current source-density analysis and determination of conductivity tensor for anuran cerebellum. J Neurophysiol 38, 356-368.

Niell, C.M., and Stryker, M.P. (2010). Modulation of visual responses by behavioral state in mouse visual cortex. Neuron 65, 472--479.

Nienborg, H., Cohen, M.R., and Cumming, B.G. (2012). Decision-related activity in sensory neurons: correlations among neurons and with behavior. Annu Rev Neurosci 35, 463-483.

Nienborg, H., and Cumming, B.G. (2006). Macaque V2 neurons, but not V1 neurons, show choice-related activity. In J Neurosci (United States), pp. 9567-9578.

Nienborg, H., and Cumming, B.G. (2009). Decision-related activity in sensory neurons reflects more than a neuron's causal effect. Nature 459, 89-92.

Nienborg, H., and Cumming, B.G. (2014). Decision-related activity in sensory neurons may depend on the columnar architecture of cerebral cortex. In J Neurosci (United States), pp. 3579-3585.

Nishimoto, S., and Gallant, J.L. (2011). A three-dimensional spatiotemporal receptive field model explains responses of area MT neurons to naturalistic movies. J Neurosci 31, 14551-14564.

Nishimoto, S., Vu, A.T., Naselaris, T., Benjamini, Y., Yu, B., and Gallant, J.L. (2011). Reconstructing visual experiences from brain activity evoked by natural movies. Curr Biol 21, 1641-1646.

Niwa, M., Johnson, J.S., O'Connor, K.N., and Sutter, M.L. (2013). Differences between primary auditory cortex and auditory belt related to encoding and choice for AM sounds. In J Neurosci (United States), pp. 8378-8395.

Okun, M., and Lampl, I. (2009). Balance of excitation and inhibition. Scholarpedia 4, 7467.

Olsen, S.R., and Wilson, R.I. (2008). Lateral presynaptic inhibition mediates gain control in an olfactory circuit. In Nature (England), pp. 956-960.

Olshausen, B.A., and Field, D.J. (2005). How close are we to understanding v1? Neural Comput 17, 1665-1699.

Orban, G.A. (2008). Higher order visual processing in macaque extrastriate cortex. Physiol Rev 88, 59-89.

Ozuysal, Y., and Baccus, S.A. (2012). Linking the computational structure of variance adaptation to biophysical mechanisms. In Neuron, pp. 1002-1015.

Pack, C.C., Conway, B.R., Born, R.T., and Livingstone, M.S. (2006). Spatiotemporal structure of nonlinear subunits in macaque visual cortex. J Neurosci 26, 893-907.

Pack, C.C., Hunter, J.N., and Born, R.T. (2005). Contrast dependence of suppressive influences in cortical area MT of alert macaque. J Neurophysiol 93, 1809-1815.

Palmer, C., Shao-Ying, C., and Eyal, S. (2007). Linking neuronal and behavioral performance in a reaction-time visual detection task. J Neurosci 27, 8122--8137.

Paninski, L. (2004). Maximum likelihood estimation of cascade point-process neural encoding models. Network 15, 243-262.

Panzeri, S., and Diamond, M.E. (2010). Information Carried by Population Spike Times in the Whisker Sensory Cortex can be Decoded Without Knowledge of Stimulus Time. Front Synaptic Neurosci 2, 17.

Parker, A.J., and Newsome, W.T. (1998). Sense and the single neuron: probing the physiology of perception. Annu Rev Neurosci 21, 227--277.

Passaglia, C.L., and Troy, J.B. (2004). Information transmission rates of cat retinal ganglion cells. J Neurophysiol 91, 1217-1229.

Pauwels, K., Kruger, N., Lappe, M., Woergoetter, F., and Van Hulle, M.M. (2010). A cortical architecture on parallel hardware for motion processing in real time. J Vis 10, 18.

Perge, J.A., Borghuis, B.G., Bours, R.J.E., Lankheet, M.J.M., and van Wezel, R.J.A. (2005). Dynamics of directional selectivity in MT receptive field centre and surround. Eur J Neurosci 22, 2049-2058.

Perrone, J.A., and Stone, L.S. (1994). A model of self-motion estimation within primate extrastriate visual cortex. Vision Res 34, 2917-2938.

Perrone, J.A., and Stone, L.S. (1998). Emulating the visual receptive-field properties of MST neurons with a template model of heading estimation. J Neurosci 18.

Perrone, J.A., and Thiele, A. (2001). Speed skills: measuring the visual speed analyzing properties of primate MT neurons. Nat Neurosci 4, 526-532.

Pillow, J.W., Paninski, L., Uzzell, V.J., Simoncelli, E.P., and Chichilnisky, E.J. (2005). Prediction and decoding of retinal ganglion cell responses with a probabilistic spiking model. J Neurosci 25, 11003-11013.

Pillow, J.W., and Simoncelli, E.P. (2006). Dimensionality reduction in neural models: an information-theoretic generalization of spike-triggered average and covariance analysis. J Vis 6, 414-428.

Poggio, T., and Serre, T. (2013). Models of visual cortex. Scholarpedia 8, 3516.

Prenger, R., Wu, M.C.K., David, S.V., and Gallant, J.L. (2004). Nonlinear V1 responses to natural scenes revealed by neural network analysis. Neural Netw 17, 663--679.

Priebe, N.J., and Ferster, D. (2005). Direction selectivity of excitation and inhibition in simple cells of the cat primary visual cortex. In Neuron (United States), pp. 133-145.

Purushothaman, G., and Bradley, D.C. (2005). Neural population code for fine perceptual decisions in area MT. Nat Neurosci 8, 99-106.

Qian, N., and Andersen, R.A. (1994). Transparent motion perception as detection of unbalanced motion signals. II. Physiology. J Neurosci 14, 7367-7380.

Qian, N., Andersen, R.A., and Adelson, E.H. (1994). Transparent motion perception as detection of unbalanced motion signals. I. Psychophysics. J Neurosci 14, 7357-7366.

Raiguel, S., Hulle, M.M.V., Xiao, D.K., Marcar, V.L., and Orban, G.A. (1995). Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. Eur J Neurosci 7, 2064-2082.

Ramirez, A., Pnevmatikakis, E.A., Merel, J., Paninski, L., Miller, K.D., and Bruno, R.M. (2014). Spatiotemporal receptive fields of barrel cortex revealed by reverse correlation of synaptic input. Nat Neurosci 17, 866--875.

Rathbun, D.L., Warland, D.K., and Usrey, W.M. (2010). Spike timing and information transmission at retinogeniculate synapses. J Neurosci 30, 13558-13566.

Reid, R.C., Victor, J.D., and Shapley, R.M. (1997). The use of m-sequences in the analysis of visual neurons: linear receptive field properties. Vis Neurosci 14, 1015-1027.

Reinagel, P., and Reid, R.C. (2000). Temporal coding of visual information in the thalamus. J Neurosci 20, 5392-5400.

Reynolds, J.H., and Chelazzi, L. (2004). Attentional modulation of visual processing. Annu Rev Neurosci 27, 611-647.

Reynolds, J.H., and Heeger, D.J. (2009). The normalization model of attention. Neuron 61, 168--185.

Richert, M. (2008). Classical and surround receptive field structure in cortical area MT as revealed by reverse correlation. (UC San Diego).

Richert, M., Albright, T.D., and Krekelberg, B. (2013). The complex structure of receptive fields in the middle temporal area. Front Syst Neurosci 7, 2.

Rieke, F. (2001). Temporal contrast adaptation in salamander bipolar cells. In J Neurosci (United States), pp. 9445-9454.

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

Riesenhuber, M., and Poggio, T. (2000). Models of object recognition. Nat Neurosci 3 Suppl, 1199-1204.

Ringach, D., and Shapley, R. (2004). Reverse correlation in neurophysiology. Cog Sci 28, 147--166.

Ringach, D.L., Hawken, M.J., and Shapley, R. (2002a). Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. J Vis 2, 12-24.

Ringach, D.L., Shapley, R.M., and Hawken, M.J. (2002b). Orientation selectivity in macaque V1: diversity and laminar dependence. J Neurosci 22, 5639-5651.

Roberts, M., S., D.L., Jose, H., A., G.M., and Alexander, T. (2007). Attention alters spatial integration in macaque V1 in an eccentricity-dependent manner. Nat Neurosci 10, 1483--1491.

Rodman, H.R., and Albright, T.D. (1987). Coding of visual stimulus velocity in area MT of the macaque. Vis Res 27, 2035-2048.

Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J Neurosci 22, 9475--9489.

Romo, R., A, H., and Antonio, Z. (2004). Neuronal correlates of a perceptual decision in ventral premotor cortex. Neuron 41, 165--173.

Romo, R., Hernández, A., Zainos, A., and Salinas, E. (1998). Somatosensory discrimination based on cortical microstimulation. Nature 392, 387--390.

Rust, N.C., Mante, V., Simoncelli, E.P., and Movshon, J.A. (2006). How MT cells analyze the motion of visual patterns. Nat Neurosci 9, 1421-1431.

Saalmann, Y.B., Pigarev, I.N., and Vidyasagar, T.R. (2007). Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. Science 316, 1612-1615.

Sahani, M., and Linden, J.F. (2003). How linear are auditory cortical responses? Advances in neural information processing systems, 125-132.

Sakai, H.M., Wang, J.L., and Naka, K. (1995). Contrast gain control in the lower vertebrate retinas. J Gen Physiol 105, 815-835.

Saleem, A.B., Ayaz, A., Jeffery, K.J., Harris, K.D., and Carandini, M. (2013). Integration of visual motion and locomotion in mouse visual cortex. In Nat Neurosci (United States), pp. 1864-1869.

Salzman, C.D., Britten, K.H., and Newsome, W.T. (1990). Cortical microstimulation influences perceptual judgements of motion direction. Nature 346, 174-177.

Sanada, T.M., Nguyenkim, J.D., and Deangelis, G.C. (2012). Representation of 3D surface orientation by velocity and disparity gradient cues in area MT. J Neurophysiol 107, 2109-2122.

Schacter, D. (2011). Pscyology (Worth Publishers).

Schmidhuber, J. (2014). Deep Learning in Neural Networks: An Overview. arXiv, 1004.7828.

Scholvinck, M.L., Saleem, A.B., Benucci, A., Harris, K.D., and Carandini, M. (2015). Cortical state determines global variability and correlations in visual cortex. J Neurosci 35, 170-178.

Schroeder, C.E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. Trends Neurosci 32, 9-18.

Schubert, T., Kerschensteiner, D., Eggers, E.D., Misgeld, T., Kerschensteiner, M., Lichtman, J.W., Lukasiewicz, P.D., and Wong, R.O. (2008). Development of presynaptic inhibition onto retinal bipolar cell axon terminals is subclass-specific. J Neurophysiol 100, 304-316.

Schwartz, O., Pillow, J.W., Rust, N.C., and Simoncelli, E.P. (2006). Spike-triggered neural characterization. J Vis 6, 484-507.

Sengpiel, F., Sen, A., and Blakemore, C. (1997). Characteristics of surround inhibition in cat area 17. Exp Brain Res 116, 216-228.

Serre, T., Oliva, A., and Poggio, T. (2007). A feedforward architecture accounts for rapid categorization. Proc Natl Acad Sci U S A 104, 6424-6429.

Shadlen, M.N., Britten, K.H., Newsome, W.T., and Movshon, J.A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. J Neurosci 16, 1486-1510.

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

Shadlen, M.N., and Newsome, W.T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. J Neurophysiol 86, 1916-1936.

Shapley, R. (2009). Linear and nonlinear systems analysis of the visual system: Why does it seem so linear? ☆: A review dedicated to the memory of Henk Spekreijse. 49, 907 – 921.

Shapley, R.M., and Victor, J.D. (1978). The effect of contrast on the transfer properties of cat retinal ganglion cells. J Physiol 285, 275-298.

Sharpee, T., Rust, N.C., and Bialek, W. (2004). Analyzing neural responses to natural signals: maximally informative dimensions. Neural Comput 16, 223-250.

Sherrington, C. (1906). The integrative action of the nervous system (New York: C Scribner and Sons).

Shuler, M.G., and Bear, M.F. (2006). Reward timing in the primary visual cortex. Science 311, 1606--1609.

Simoncelli, E., Paninski, L., Pillow, J., and Schwartz, O. (2004). Characterization of neural responses with stochastic stimuli. In The New Cognitive Neurosciences, M. Gazzaniga, ed. (MIT Press).

Simoncelli, E.P., and Heeger, D.J. (1998). A model of neuronal responses in visual area MT. Vis Res 38, 743-761.

Smith, M.A., and Kohn, A. (2008). Spatial and temporal scales of neuronal correlation in primary visual cortex. In J Neurosci (United States), pp. 12591-12603.

Softky, W.R., and Koch, C. (1993). The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. J Neurosci 13, 334-350.

Stanley, G.B., Jin, J., Wang, Y., Desbordes, G., Wang, Q., Black, M.J., and Alonso, J.M. (2012). Visual orientation and directional selectivity through thalamic synchrony. In J Neurosci (United States), pp. 9073-9088.

Stanley, G.B., Li, F.F., and Dan, Y. (1999). Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. J Neurosci 19, 8036-8042.

Steriade, M., Nunez, A., and Amzica, F. (1993). A novel slow (< 1 Hz) oscillation of neocortical neurons in vivo: depolarizing and hyperpolarizing components. J Neurosci 13, 3252-3265.

Tan, A.Y., Chen, Y., Scholl, B., Seidemann, E., and Priebe, N.J. (2014). Sensory stimulation shifts visual cortex from synchronous to asynchronous states. Nature 509, 226-229.

Tanaka, K., Hikosaka, K., Saito, H., Yukie, M., Fukada, Y., and Iwai, E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. J Neurosci 6, 134-144.

Tanaka, K., Sugita, Y., Moriya, M., and Saito, H.A. (1993). Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. J Neurophysiol 69.

Tehovnik, E.J., Slocum, W.M., and Schiller, P.H. (2004). Microstimulation of V1 delays the execution of visually guided saccades. Eur J Neurosci 20, 264--272.

Tolhurst, D.J., Movshon, J.A., and Dean, A.F. (1983). The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res 23, 775--785.

Treue, S., and Andersen, R.A. (1996). Neural responses to velocity gradients in macaque cortical area MT. Vis Neurosci 13, 797-804.

Treue, S., and Martinez-Trujillo, J.C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature 399, 575-579.

Truccolo, W., Eden, U.T., Fellows, M.R., Donoghue, J.P., and Brown, E.N. (2005). A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. In J Neurophysiol (United States), pp. 1074-1089.

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. Neural Computation 10, 821-835.

Tsui, J.M., and Pack, C.C. (2011). Contrast sensitivity of MT receptive field centers and surrounds. J Neurophysiol 106, 1888-1900.

Uka, T., and DeAngelis, G.C. (2004). Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. Neuron 42, 297-310.

Uka, T., Tanabe, S., Watanabe, M., and Fujita, I. (2005). Neural correlates of fine depth discrimination in monkey inferior temporal cortex. In J Neurosci (United States), pp. 10796-10802.

Ungerleider, L.G., and Desimone, R. (1986). Cortical connections of visual area MT in the macaque. J Comp Neurol 248.

Usrey, W.M., Alonso, J.M., and Reid, R.C. (2000). Synaptic interactions between thalamic inputs to simple cells in cat visual cortex. In J Neurosci (United States), pp. 5461-5467.

Uzzell, V.J., and Chichilnisky, E.J. (2004). Precision of spike trains in primate retinal ganglion cells. J Neurophysiol 92, 780-789.

Vaingankar, V., Soto-Sanchez, C., Wang, X., Sommer, F.T., and Hirsch, J.A. (2012). Neurons in the thalamic reticular nucleus are selective for diverse and complex visual features. Front Integr Neurosci 6, 118.

Vogel, A., Hennig, R.M., and Ronacher, B. (2005). Increase of neuronal response variability at higher processing levels as revealed by simultaneous recordings. J Neurophysiol 93, 3548-3559.

Wang, X., Hirsch, J.A., and Sommer, F.T. (2010). Recoding of sensory information across the retinothalamic synapse. J Neurosci 30, 13567-13577.

Wang, Y.V., Weick, M., and Demb, J.B. (2011). Spectral and temporal sensitivity of cone-mediated responses in mouse retinal. J Neurosci 31, 7670-7681.

Weber, F., Machens, C.K., and Borst, A. (2010). Spatiotemporal response properties of optic-flow processing neurons. Neuron 67, 629-642.

Wehr, M., and Zador, A.M. (2003). Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. Nature 426, 442--446.

Willmore, B.D.B., Prenger, R.J., and Gallant, J.L. (2010). Neural representation of natural images in visual area V2. J Neurosci 30, 2102-2114.

Wilson, D.A. (2001). Receptive Fields in the Rat Piriform Cortex. Chem Senses 26, 577--584.

Wohrer, A., and Machens, C.K. (2013). Percept and the single neuron. Nat Neurosci 16, 112-113.

Wu, M.C., David, S.V., and Gallant, J.L. (2006). Complete functional characterization of sensory neurons by system identification. Annu Rev Neurosci 29, 477-505.

Wässle, H. (2004). Parallel processing in the mammalian retina. Nat Rev Neurosci 5, 747-757.

Xiao, D.K., Raiguel, S., Marcar, V., Koenderink, J., and Orban, G.A. (1995). Spatial heterogeneity of inhibitory surrounds in the middle temporal visual area. Proc Natl Acad Sci U S A 92, 11303-11306.

Xiao, D.K., Raiguel, S., Marcar, V., and Orban, G.A. (1997). The spatial distribution of the antagonistic surround of MT/V5 neurons. Cereb Cortex 7, 662-677.

Xing, D., Yeh, C.I., and Shapley, R.M. (2009). Spatial spread of the local field potential and its laminar variation in visual cortex. J Neurosci 29, 11540-11549.

Yamasaki, D.S., and Wurtz, R.H. (1991). Recovery of function after lesions in the superior temporal sulcus in the monkey. J Neurophysiol 66, 651--673.

Zaghloul, K.A., Boahen, K., and Demb, J.B. (2005). Contrast adaptation in subthreshold and spiking responses of mammalian Y-type retinal ganglion cells. In J Neurosci (United States), pp. 860-868.

Zaghloul, K.A., Manookin, M.B., Borghuis, B.G., Boahen, K., and Demb, J.B. (2007). Functional circuitry for peripheral suppression in Mammalian Y-type retinal ganglion cells. J Neurophysiol 97, 4327-4340.

Zanos, S., Zanos, T.P., Marmarelis, V.Z., Ojemann, G.A., and Fetz, E.E. (2012). Relationships between spike-free local field potentials and spike timing in human temporal cortex. J Neurophysiol 107, 1808-1821.

Zanos, T.P., Mineault, P.J., and Pack, C.C. (2011). Removal of spurious correlations between spikes and local field potentials. J Neurophysiol 105, 474-486.

Zeki, S.M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. J Physiol 236.

Zemel, R.S., and Sejnowski, T.J. (1998). A model for encoding multiple object motions and self-motion in area MST of primate visual cortex. J Neurosci 18.

Zhang, D., Wu, S., and Rasch, M.J. (2015). Circuit motifs for contrast-adaptive differentiation in early sensory systems: the role of presynaptic inhibition and short-term plasticity. PLoS One 10, e0118125.

Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature 370, 140-143.

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