Circuit mechanisms underlying the encoding of ethologically relevant visual stimuli in the retina

Maxwell H. Turner

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2017

Reading Committee:

Fred Rieke, Chair

Eric Shea-Brown

Wyeth Bair

Program Authorized to Offer Degree:
Neuroscience

University of Washington

Abstract

Circuit mechanisms underlying the encoding of ethologically relevant visual stimuli in the retina

Maxwell H. Turner

Chair of the Supervisory Committee:
Dr. Fred Rieke
Department of Physiology & Biophysics

In the human retina, the axons of roughly 10^6 retinal ganglion cells (RGCs) carry all of the information underlying visual perception and visually guided behavior. RGC computation is the last processing step before the imposition of this sensory bottleneck. Because of the importance of RGC computation for visual function and because of the accessibility of the retina to physiological investigation, RGCs are among the best studied class of early sensory neurons. Many decades of investigation using artificial visual stimuli (e.g. spots, gratings, white noise etc.) has revealed a great deal about the physiology of RGCs, the computations they can perform, and the circuitry underlying these computations. What is lacking, however, is an understanding of RGC computation and encoding during natural visual stimulation. A complete understanding of RGC function requires extending what we have learned using artificial, mostly static visual stimuli to the dynamic and spatially structured conditions that characterize natural vision. A major focus of the thesis presented here is to begin to bridge this gap.

In Chapter 2 I will review some of what is known about how neural circuits in the early visual system encode naturalistic visual inputs. In Chapters 3 & 4 I will present work that connects a ubiquitous concept in early visual processing - the receptive field - to the encoding of natural visual stimuli in the retina. Classical models of the retinal ganglion cell (RGC)

receptive field assume linear integration across visual space, and this assumption guides most modern day models that aim to predict RGC responses to visual stimulation. I show that spatial nonlinearities can be important for encoding spatial contrast within a natural scene. Furthermore, we can use what we know about nonlinear receptive field structure to improve models of RGCs. Finally, I show that spatial contrast encoding by RGCs can be modulated by visual context.

Regardless of the stimulus being encoded, a fundamental limit to the fidelity with which sensory information can be encoded is the variability of neural responses. Repeated presentations of the same stimulus will elicit variable responses from a single neuron. Noise is a feature of neural population responses as well, and in practice it is often found that this noise is correlated within a population. In Chapter 5, I will present a strategy used by retinal circuits to minimize the effect of noise on the encoding of a behaviorally-relevant feature of the visual world - namely an object's direction of motion.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Good science and good scientific training can't be done in isolation. I am extremely fortunate to have had the opportunity to do this work within so many amazing support systems. The Rieke lab is a fabulous environment in which to work and grow as a scientist. Establishing this environment takes thoughtful cultivation, and Fred has achieved this wonderfully.

Fred is sort of a one-man support system himself. I have heard more than one person remark that Fred seems capable of doing the work of several people, and I have to agree. He is also one of the most thoughtful and brilliant scientists I have ever met. As a mentor, Fred patiently leads by example. Talk to any person who was trained by Fred and they will all tell you how amazing he is. I am extremely grateful to have had the opportunity to work with him and learn from him.

There are many wonderful people who have been a part of the lab during my time here and without them nothing presented in this thesis would have been possible. Greg Schwartz first sparked my interest in nonlinear spatial integration and receptive field structure. Will Grimes has been a fixture in the lab since I joined and his enthusiasm and helpful advice, especially when it comes to experimental techniques, have been invaluable. Sid Kuo, Raunak Sinha, and Braden Brinkman have similarly provided great feedback and moral support throughout most or all of my graduate career. Zhou Yu and Gabrielle Gutierrez, the newest members of the lab, are also amazing colleagues with whom I've very much enjoyed bouncing ideas off of.

My graduate-student "big brother" is Juan Angueyra. I remember Juan vigorously nodding in support at my first year rotation talk, and ever since then he has been an extremely valuable presence, even after he left the lab. As the senior graduate student in the lab I hoped to be half as supportive as Juan. The other graduate students in the lab - Ali Weber, Phil Mardoum, and Jacob Baudin - are extremely talented scientists and fun to be around both in the lab and outside of it.

The Rieke lab also benefits from very capable technical support. As an electrophysiologist I recognize how lucky I was that I never worried about the quality of my internal solution, or

whether the mouse I was using for an experiment was the right genotype. Paul Newman and Shellee Cunnington have been responsible for experimental solutions and generally making sure the lab runs smoothly. Mike Ahlquist skillfully maintains our mouse colonies (a truly heroic effort). Mark Cafaro is extremely talented and helped with everything from day-to-day computer and rig maintenance to designing and developing the acquisition and visual stimulation software we use in the lab. His acquisition software made many of the experiments in this thesis (especially in Chapters 3 & 4) possible by allowing us to easily do data analysis on the fly, during an experiment. This allowed me to tailor stimuli to a particular cell's receptive field and ask questions I would not otherwise have been able to ask.

All of the macaque monkey retina that was used for the studies presented in Chapters 3 & 4 was obtained through the Tissue Distribution Program at the Washington National Primate Research Center. This has been a truly unique resource for us, and I will always be grateful for it. I want to especially highlight the fantastic support we received from Chris English and Drew May, who went above and beyond to make sure that retinal tissue was made available to us whenever an opportunity arose.

I'd also like to thank some of my collaborators over the years including Adam Bleckert, Gautam Awatramani and Tim Sexton, all of whom let me help with their projects as a sort of hired gun for patch physiology. I would like to highlight the work of my collaborators on the study presented in Chapter 5 - Joel Zylberberg, Jon Cafaro, Fred Rieke and Eric Shea-Brown. This project was extremely collaborative and generated countless illuminating discussions. I am grateful for their diligence and insight, and I am very proud of the work we did together.

The members of my thesis advisory committee - Rachel Wong, Wyeth Bair, Eric Shea-Brown, and Maureen Neitz - have provided invaluable feedback all throughout my graduate career. Their advice and guidance was hugely helpful in finishing projects and planning for the next step.

Finally, I would like to thank the friends and family who have helped me get to this point, and never doubted that I could be a scientist. I am especially grateful to my soon-to-be wife, Sarah Pickett, a talented scientist in her own right. She has been an incredible source of support over the years.

Chapter 1
INTRODUCTION & ORGANIZATION OF THE THESIS

The work presented in this thesis is motivated by the question of how retinal circuits encode behaviorally-relevant visual stimuli, and how this is done efficiently. The following chapter serves as a reader's guide to introduce some basic background on the topics that will be discussed in this thesis. It also provides a road map for what is to come in subsequent chapters.

A major focus of the work presented here is the retinal representation of naturalistic visual stimuli, including those subjected to eye movements. Chapter 2 will present an in-depth review of what is known about the encoding of naturalistic stimuli in the early visual system. Chapters 3 & 4 will relate retinal receptive field structure to natural scene encoding.

A major impediment to the efficient representation of any visual stimulus is the presence of noise in neural responses. In Chapter 5 I will present work that explores a strategy for mitigating some of the deleterious effects of neural response variability in a case where the visual encoding task is clear - retinal direction selectivity.

1.1 Circuitry of the vertebrate retina

For most of the history of neuroscience, the vertebrate retina has attracted the attention of both anatomists and physiologists. Because of this focus, we now benefit from a quite detailed understanding of the anatomical organization of the retina, its basic physiological properties, and many of the visual computations it can support. Chapters 3 & 4 of this thesis will present work in the peripheral retina of the macaque monkey, while Chapter 5 uses the mouse retina as a model system. While these two systems have several important differences, their basic neural circuit organization is the same (see Fig. 1.1).

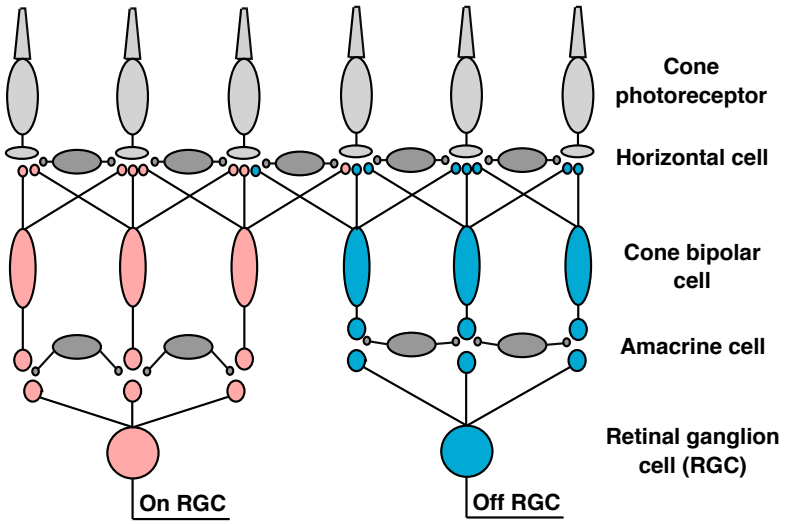

Figure 1.1: **Circuitry of the vertebrate retina** The five major neural cell classes in the vertebrate retina. On pathways (left, pink cells) and Off pathways (right, blue cells) are defined by their preference for increments and decrements in light intensity, respectively. For the purposes of this thesis, only the major cone pathways are included. Separate pathways (not shown) are responsible for transmitting rod-derived visual signals to the retinal ganglion cell layer.

Light is first transduced into an electrical signal in the photoreceptors (rods and cones) of the outer retina. Photoreceptors make glutamatergic synapses onto excitatory interneurons called bipolar cells. Broadly speaking, there are two classes of cone bipolar cells (those that receive input from cone photoreceptors): On cone bipolar cells depolarize in response to increments in light intensity, while Off cone bipolar cells depolarize to light decrements. Within this broad clustering, there are approximately 13 different types of cone bipolar cells [70]. Bipolar cells provide excitatory, glutamatergic synaptic input to retinal ganglion cells (RGCs), which are the output cells of the retina. At each synaptic layer (called "plexiform layers") there are interneurons that further shape the visual signal. In the outer plexiform layer, horizontal cells mediate lateral inhibitory interactions between signals from photoreceptors. In the inner retina, a very diverse class of interneurons called amacrine cells mediate mostly lateral, mostly inhibitory, interactions [141].

The diversity of photoreceptors, horizontal cells, bipolar cells and amacrine cells ulti-

mately produces a wide diversity of signals present in the retinal output. The axons of RGCs make up the optic nerve, which projects to downstream visual areas. In the mouse retina, there are upwards of 40 different classes of RGC [15]. It is unclear how many RGC classes exist in the macaque retina, but 17 different classes have been identified and described [76, 52]. Each RGC class forms an independent mosaic in the retina such that a given point in visual space is encoded by each RGC class. There is a general consensus in the field that each RGC class constitutes a separate channel for visual information, and that each class extracts different features of the visual world [142].

Which visual features are extracted from a scene by different RGC classes is very much an ongoing area of investigation. Very little is known about this sort of feature extraction in the context of natural visual stimuli. Chapters 3 & 4 will reveal some of the naturalistic feature sensitivities of two types of macaque monkey RGCs.

1.2 The retinal ganglion cell receptive field

In early studies of the electrical responses of visual neurons, a neuron's receptive field was defined as the region of visual space in which a change in illumination causes a change in the cell's response [97, 96]. A broader definition is that a receptive field describes the manner in which a cell integrates visual inputs over space and/or time to determine its response. Indeed when modern visual neuroscientists refer to a cell's receptive field, they are often referring to a model of the cell, or at least a part of such a model [208].

The first descriptions of the spatial receptive field of RGCs highlighted an antagonistic structure in which there was an excitatory "center" region and a neighboring suppressive "surround" region [125]. Observations like these led to an enduring model of the RGC receptive field called the "difference-of-Gaussians" model (Fig. 1.2A). In this model, the receptive field center is described by a two-dimensional Gaussian integrating function in visual space. Similarly, the surround is also composed of a Gaussian integrating function, but it is somewhat larger than the center and, importantly, has the opposite polarity to account for its suppressive nature.

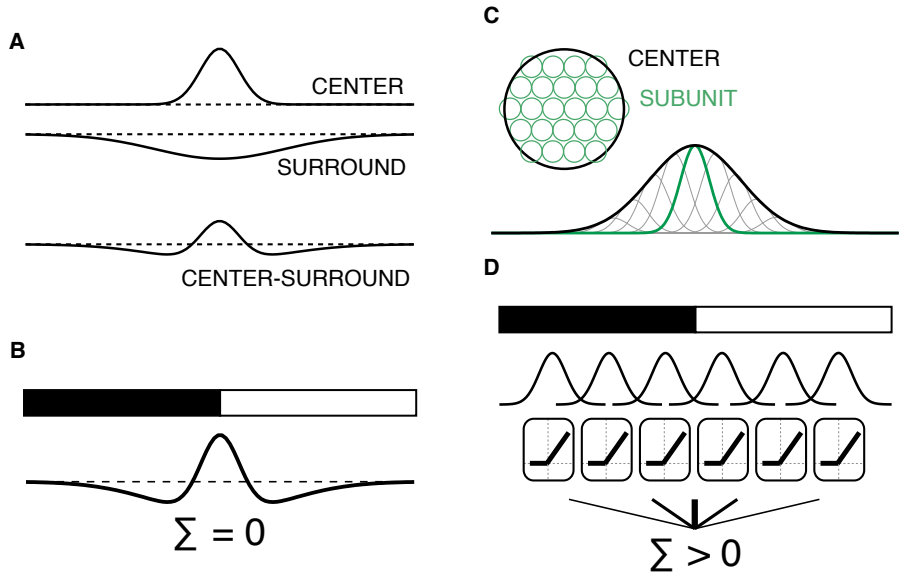

Figure 1.2: **Models of retinal ganglion cell spatial receptive fields** (A) The "Difference-of-Gaussians" spatial receptive field model consists of an excitatory center and a larger, suppressive surround. Each is described by a two-dimensional Gaussian integrating function in visual space. The linear combination of these two receptive field regions gives rise to the full, center-surround receptive field. (B) A spatially-linear receptive field model, such as the difference-of-Gaussians model, predicts no response to visual stimuli that contain equal and opposite positive and negative contrasts, which cancel when summed linearly. (C) The receptive field center is composed of many nonlinear "subunits" (green) that tile visual space (inset). (D) Small subunits that are treated with private, rectifying nonlinearities before spatial summation can account for nonlinear responses to grating stimuli. The rectification of subunit output prevents cancelation of light and dark regions of the visual input.

The difference-of-Gaussians receptive field model can account for RGC responses to many visual stimuli like spots and annuli [20]. When presented with some visual stimuli, however, the spatially linear difference-of-Gaussians model fails to account for the RGC response. This was originally shown using sine- or square-wave grating stimuli [69]. For a spatially linear cell, one can find a position of the grating over the receptive field that drives no neural response. This is because the positive and negative regions of the grating cancel when summed linearly over visual space. For other RGCs no such grating position can be found, and the cell responds robustly at all grating positions. This is termed "nonlinear spatial integration" and it conflicts with the spatially-linear difference-of-Gaussians model.

Nonlinear spatial integration can be explained using a model of the receptive field that contains small, nonlinear subunits [100] (Fig. 1.2C,D). These subunits correspond to the bipolar cells presynaptic to the RGC [61, 192]. A subunit has its own, small receptive field and the output of each subunit is treated with a rectifying nonlinear transfer function before summation over visual space, which takes place in the RGC. This rectification of subunit output ensures that a subunit will not transmit a suppressive signal to the RGC, and ultimately this prevents cancelation of light and dark regions in a visual scene. Chapters 3 & 4 will explore the impact of this receptive field organization on the encoding of naturalistic visual stimuli.

It is important to note that many receptive field models are spatially linear, not just the classical difference-of-Gaussians model. The assumption of spatial linearity has informed the architecture of most modern day predictive models of RGCs [38, 164, 208, 99], which aim to predict responses to arbitrary visual stimuli. In order to construct successful predictive models, we need to understand the conditions under which linear spatial integration is an appropriate abstraction of neural processing, and when we might need to re-evaluate the architecture of predictive models.

1.3 Direction selectivity in the retina

Spatial receptive field models like those illustrated in Fig. 1.2 suggest that visual encoding by RGCs is relatively simple, and that these neurons are simply sensitive to features of a certain size. However the use of more dynamic, hypothesis-driven stimuli has revealed that RGCs often perform complex computations on visual inputs. One of the most well known examples of a complex retinal computation is direction selectivity. First described by Barlow & Levick [21], direction-selective (DS) RGCs respond robustly to motion in one direction (the cell's "preferred" direction) but do not respond to motion in the opposite direction (the "null" direction).

DS RGCs have been observed in several vertebrate species including rabbit [21, 157] and mouse [239, 68]. There are several classes of DS RGCs that have been described. On DS RGCs are tuned for light increments [216], while Off DS RGCs are tuned for light decrements [119]. The most numerous and well-studied are On-Off DS RGCs [230], which receive inputs from both On and Off cone bipolar cells and correspondingly respond to both light increments and decrements [218]. There are four sub-types of On-Off DS RGCs, each of which prefers motion in a different cardinal direction in visual space [157].

Retinal direction selectivity relies on several overlapping and partially redundant mechanisms (for review see [59, 230]). The most important seems to be spatially offset inhibitory input provided by a class of amacrine cells called starburst amacrine cells, which release the inhibitory neurotransmitter GABA. The result is that preferred-direction motion elicits relatively small inhibitory synaptic inputs to a DS RGC, while null-direction motion drives very strong GABAergic inhibitory input to the cell [218, 83]. This strong null direction inhibition is evident in voltage clamp recordings from DS RGCs (e.g. see Chapter 5).

Retinal direction selectivity is an attractive system in which to examine the fidelity of encoding of complex visual features because the feature being encoded is clear. This allows for precise measurement of coding efficiency in the context of a behaviorally-relevant aspect of the visual environment. We can, for instance, ask how well one can decode the direction of

motion based on responses of direction-selective RGCs. We perform exactly such an analysis in Chapter 5.

Chapter 2

NEURAL COMPUTATION AND CODING DURING NATURAL VISION

2.1 Summary

The neural circuits that process sensory inputs have been shaped through evolution by the properties of the stimuli they encounter as well as the behavioral demands of the animal. Because of this, a deep understanding of sensory circuits and the computations they support requires connecting what we know about sensory systems to naturalistic stimulus conditions. The visual system is one of the best-understood sensory systems in neuroscience, but much more is known about how visual neurons encode simple, parameterized stimuli compared to more complex stimuli, including natural scenes. Translating what we learn using simple stimuli to the more complex conditions that characterize natural vision has not been straightforward. This is reflected in the failure of predictive models of visual neurons to generalize to naturalistic stimuli. Here we review some of the progress made towards understanding neural processing and representation of naturalistic stimuli in the early visual system. We will present a case for the use of naturalistic visual stimuli in the study of vision, and explore some of the considerations that may help make this approach successful.

2.2 Introduction

Sensory circuits evolved to extract behaviorally relevant information from stimuli encountered in the real world. To fully understand the function of a sensory circuit, we need to place it in the context of the constraints imposed by a realistic sensory environment. Studying sensory systems using natural stimuli can help provide a functional context for sensory processing. From a practical perspective, replacing a damaged or diseased sensory circuit

with a prosthetic device requires understanding what processing is being done to stimuli encountered in the real world as well as how that processing is achieved by the biological circuit under investigation.

Despite the appeal of an natural approach to sensory neuroscience [72], the field has progressed quite successfully using parametric artificial stimuli while avoiding stimuli that mimic the inputs encountered during natural conditions. This focus on simpler, more controlled stimuli has allowed for a rapid expansion of our understanding of how neurons process certain types of sensory inputs. This is probably best illustrated in the early vertebrate visual system, where neural computations and the circuits that underlie them are relatively well understood [36].

Efforts to translate the insights gained from studies that use simple, artificial visual stimuli to more natural conditions have been stymied by a number of complicating factors, chief among which is the sheer complexity of natural visual inputs [186]. Fortunately, recent advances in our understanding of a few areas - including the properties of natural visual inputs, computational methods for describing visual neurons, and circuit mechanisms of neural computation - has placed modern neuroscientists in a unique position to put vision in its natural context.

First, we'll discuss some of the evidence that early visual processing is tuned to the statistical properties of natural stimuli, as might be expected from evolutionary principles. Second, we'll review some of what is known about the neural representation of natural visual stimuli, highlighting the fact that the use of natural stimuli has provided key insights into the nature of the neural code. Third, we'll explore the possibility that because what we infer about the receptive field relies critically on the stimuli used to measure it, the likelihood of "solving" vision using exclusively artificial stimuli is low. Finally, we'll review progress on constructing computational models of visual neurons that can predict responses to arbitrary visual inputs, including natural stimuli.

2.3 Neural processing strategies reflect features of natural scenes

2.3.1 Efficient coding of natural visual inputs

An early and still influential hypothesis that undergirds much of the study of natural scene processing is that processing in early sensory systems produces an efficient representation of stimuli encountered in the natural world. This "efficient coding hypothesis," first proposed by Barlow [18, 23], was influenced quite heavily by Shannon's earlier work on information theory [203].

An efficient representation in this context means one that is maximally informative about the stimulus being encoded, subject to some constraint on the capacity of the sensory channel to transmit information. On behaviorally relevant timescales, there are a finite number of discernible messages (e.g. spike patterns) that a single sensory neuron can send to downstream neural circuits. Even analog signals in single cells (synaptic currents, membrane potentials, etc.) have a limited range over which they can operate, and a finite number of discernible signals due to noise. These physical constraints place a limit on the amount of information that a single cell or neural population can transmit about any stimulus. The efficient coding hypothesis is essentially the recognition of this problem and the prediction that the goal of early sensory processing is to use this limited capacity as efficiently as possible.

Barlow argued that efficient coding hinges on the reduction of redundancy among neural representations [18]. Two representations are redundant to the extent that knowledge of one gives you information about the other. Several features identify an efficient neural code. First, on a population level, neural responses should be uncorrelated, as correlations imply redundant representations of the same visual feature. Second, single cells should distribute their responses uniformly, such that each possible neural response occurs with equal frequency, since to do otherwise would mean that a cell is not making full use of its dynamic range.

The first example of efficient coding through redundancy reduction in a neural circuit comes from Laughlin [128]. Laughlin found that the distribution of contrasts in natural

scenes, when transformed by the contrast-response function of a blowfly visual interneuron produces a uniform distribution of membrane potentials in the cell. This sort of evenly-distributed representation of natural contrasts is in accordance with the efficient coding hypothesis as laid out by Barlow.

Through the lens of efficient encoding via redundancy reduction, natural stimuli present a particular challenge to visual neurons. This is because natural visual inputs contain strong statistical regularities across time and space (For an excellent review, see [211]). Put simply, knowing something about a particular region of a visual scene tends to tell you quite a lot about nearby regions. This is a problem because if visual neurons simply faithfully represented natural scenes, the neural representation of natural scenes would contain correlations that are inherited from the strong correlations present in the stimuli. This, the argument goes, would make for a terrible neural code, with neurons wasting precious spikes by conveying highly redundant information.

Field analyzed the statistical regularities present in natural images and found that the spatial correlations obeyed a power law scaling: the spatial power spectrum (in the Fourier domain) tended to fall as the inverse of the square of the spatial frequency [74]. This is sometimes referred to as the "scale invariant" spatial structure of natural scenes and it was the first step towards making concrete the statistical structure of natural images. Here, scale invariance means that the statistics of natural images do not change as one changes the viewing angle - e.g. if we compare a collection of natural images to a collection of zoomed-in natural images, we would find that the statistics of the two ensembles are very similar. It is often suggested that this results from the fact that objects in the real world can appear at any distance away from the observer [184]. Natural movies also show power law scaling in space and time, indicating the presence of correlations across time [62].

Atick & Redlich examined the suitability of "retinal filters" for efficiently encoding stimuli with the second-order redundancy characteristic of natural images [8]. Psychophysically measured spatial frequency sensitivity curves are band-pass, with both high and low spatial frequencies attenuated [117, 58]. Atick & Redlich found that the suppression of low spatial

frequencies was sufficient to flatten the amplitude spectrum of natural scenes, especially at these low spatial frequencies, effectively removing a large portion of the redundancy present in natural images. This flattening of the amplitude spectrum is referred to as "whitening" and removes correlations at the spatial frequencies that have been flattened. Similar decorrelation in both space and time was subsequently shown for LGN responses to natural movies [53], with spatial decorrelation the result of center-surround receptive field organization in retina and LGN and decorrelation in the temporal domain owing to the biphasic temporal filters of these neurons.

While these early examples had justifiably been cited as successes of the redundancy reduction hypothesis [7], several lines of evidence support the view that redundancy reduction is at best a partial explanation for center-surround receptive field organization seen in the retina and LGN.

First, a better understanding of the statistics of natural visual inputs under the influence of self-generated movement may call into question the original formulation of the problem: that natural visual stimuli contain a particular spatial correlation structure that obeys power law scaling. This is because, when subjected to some types of human eye movements, the spatial frequency spectra of natural images is roughly flat at low spatial frequencies [124]. This whitening effect is the result of fixational eye movements, which take place between the saccades that rapidly change the fixation point of the observer [30].

This means that, owing to the fixational eye movements that are constantly and unconsciously made by primates, the retinal image may already be whitened, prior to any neural processing at all. The decorrelating effect of fixational eye movements was later shown experimentally by measuring correlations in a population of retinal ganglion cells (RGCs) [196]. More generally, the effect of eye movements on the retinal image will be an important consideration in future studies that aim to learn something about neural processing of natural inputs [183]. Fortunately, the statistics of eye movements (especially those of the primate) are quite well studied [138, 139, 140, 182], and resources exist that may be of use to experimentalists who would like to incorporate eye movements into their visual stimulus toolbox

[227, 244].

Second, despite RGCs generally having a center-surround receptive field organization, natural movies produce highly correlated RGC population responses [168]. So at minimum, the actual receptive field surrounds in the retina tend to incompletely decorrelate the population response. Moreover, nonlinearities in retinal circuitry can contribute much more to RGC decorrelation than does center-surround receptive field organization [165].

Despite its limitations (see also [89]), redundancy reduction is still a useful framework for thinking generally about the efficient coding of natural images. Of course, natural images contain statistical regularities beyond those described by the power law scaling of the spatial amplitude spectrum [184, 185], and efficient coding predicts that neural processing should reduce redundancies that arise due to the higher-order structure of natural scenes as well. For instance, a model of V1 neurons that uses divisive normalization to decorrelate responses relative to neighboring units is able to produce behaviors that resemble measured V1 neurons [195], showing that gain control in V1 may be well suited to remove higher order correlations present in natural scenes.

In a pair of insightful studies, Coen-Cagli and colleagues showed that V1 receptive field surrounds gate their modulatory influence based on the statistical homogeneity of visual inputs, including natural images [41, 42]. Thus, when a V1 cell encounters a region of a natural scene with relatively homogeneous surroundings, its response is suppressed. If, however, a statistically heterogeneous region of a scene falls in a cell's receptive field, it will be able to respond robustly. The authors suggest this might be well suited to highlight regions of perceptual salience.

2.3.2 Beyond redundancy reduction: naturalistic feature sensitivity

Much of the classical work on efficient encoding of natural images considers only second-order statistics - i.e. the variance or power spectrum (e.g. [8]). There is, however, much more to natural images than their spatial frequency spectra. This becomes obvious when one generates artificial stimuli with the "natural" 1/f amplitude spectrum, as such images

look very different than natural images [211]. Moreover, natural images that have been whitened (and thus lack any pairwise correlations) retain much of the structure of the original image. Randomizing the phase structure of natural images, on the other hand, renders them unrecognizable and relatively featureless [220].

Given the importance of these higher order statistics in making natural scenes appear "natural," we might expect visual neurons to be sensitive to the complex features present in natural scenes. In the retina, for instance, an ethological approach to ganglion cell encoding has helped reveal the complexity of retinal computation.

In a well-known early study of retinal feature sensitivity, Lettvin and colleagues interpreted RGC types in explicit ethological terms, famously going so far as to speculate that one class of ganglion cell in the frog retina may be a "bug perceiver" [130]. The idea that the earliest neurons in the visual system are tuned to highly specific features of the visual world was a bit ahead of its time, as a different view of early visual processing seems to have dominated for several decades thereafter. In what has been called the "standard" or "textbook" view [143], the retina implements basic processing including lateral inhibition (via a center-surround spatial receptive field) and simple forms of luminance adaptation. Different ganglion cell classes do have different receptive field sizes, adaptation and gain control properties, and some niche behaviors like nonlinear spatial integration [69, 100, 205, 206]. But ultimately downstream visual areas, including V1, receive a minimally processed image and it is here that all of the computational heavy lifting is done.

There is now a consensus in the retina community that the retina produces many parallel outputs, each carrying information about different features of the image [142, 238]. The particular output channels will vary from species to species, but in each case the information extracted from a visual scene will be shaped, through evolution, by the visual conditions and behavioral demands associated with the animal.

A great deal of evidence has accumulated that retinal computation is more complex than was once appreciated (for an excellent review see [87]). A wide variety of "non-standard" RGC computations have now been described (and often explained at the circuit and synaptic

level). An incomplete but representative sample includes: direction-selectivity [22, 59, 83, 230], orientation selectivity [150], sensitivity to differential foreground/background motion [12, 155], and detection of image blur [136].

Complex retinal computations are typically explored using low dimensional, artificial stimuli. But the use of natural stimuli to probe retinal computation can give insight into how retinal feature extraction might occur in the real world. In particular, we'd like to know how robust a particular feature extraction may be to the complex spatial or temporal structure characteristic of natural visual inputs.

For example, [123] discovered a very unusual response characteristic in so-called image-recurrence-sensitive RGCs in the mouse retina. When presented with a series of flashed gratings, these cells are typically suppressed after a grating stimulus is removed and a different grating appears. However, if the same grating re-appears, no such suppression occurs. The same effect can be seen using natural images, but the characteristic response was more variable and less reliably signaled image-recurrence. Any image recurrence signal in these cells' responses is likely present alongside signals about other features of the scene, like spatial structure within the image.

In the visual systems of insects one can also find evidence that neurons are sensitive to features present in natural stimuli. The fly vision community has a long history of examining visual processing as it relates to behaviors like flying [98]. The dissection of the circuits and computations underlying elementary motion detection has been a particularly successful endeavor (for reviews see [33, 67, 210]). Comparison of several insect species revealed that motion processing pathways in these animals appears tuned to the particular flight behaviors exhibited by each species [151]. Moreover, consideration of the visual motion induced by optic flow has led to a re-classification of the motion sensitivities of some visual neurons in the fly. For instance, horizontal system neurons in the blowfly seem to encode information about yaw rotation during flight [122].

Accounting for the sorts of visual input encountered during real world flying behavior has led to several insights about what motion sensitive neurons are encoding about the visual

world as well as how this encoding is achieved. For example, when flies are presented with naturalistic optic flow stimuli that are generated using real flight paths, motion-sensitive cells were found to be sensitive not only to elementary motion signals induced by rotation (as would be expected based on experiments using artificial stimuli) but also motion due to translation [29, 228, 113].

Heterogeneous spatial structure can complicate the extraction of motion signals from natural scenes, and higher order correlations can be used to improve motion estimation from these stimuli [40]. Clark, Fitzgerald & colleagues showed that flies are sensitive to higher order correlations and can use them to detect motion [40]. Sensitivity to higher order correlations relies on separate pathways sensitive to increments and decrements in intensity, and interestingly, they found evidence for this strategy in humans as well [40, 78].

2.4 Sparse representation of natural visual stimuli

The task in understanding neural encoding is to provide an account of the mapping from external stimulus to neural response [177], and justifiably much effort has been put into questions of this sort in sensory neuroscience. To account for the encoding of naturalistic stimuli one must know how the nature of the neural code depends on the statistics of natural and artificial stimuli.

Mounting experimental observations, sometimes motivated by theoretical considerations, suggest that visual information is represented by small numbers of neurons at any time. This idea is known as "sparse coding" [154]. A sparse neural code is one in which, for a given visual stimulus, only a small proportion of neurons in a population is active [75]. A related (but distinct) concept to this "population sparseness" is the idea of "lifetime sparseness," in which a single cell is only rarely active in the course of natural visual stimulation [174]. Both of these features of neural responses can arise due to specificity of neural feature selectivity, and Barlow pointed out that such specificity tends to increase as one progresses along the visual hierarchy [19]. For example, a retinal photoreceptor will hyperpolarize whenever it passes over a brighter region of a visual scene, which for natural scenes happens a bit less

than half of the time. A neuron in V1, however, responds more infrequently when presented with natural scenes [234]. Again the concept of redundancy reduction is a useful framework to keep in mind: a neural population could represent an image with relatively few active units by exploiting the statistical regularities present in natural scenes [19].

Field observed that the known structure of simple cell receptive fields in V1 promotes a sparse representation of natural images [74]. Several subsequent studies used learning algorithms to derive basis sets that represent natural images according to some coding constraint like explicitly maximizing sparseness of representation [152] or enforcing statistical independence of the bases [26, 229]. Such analyses revealed model receptive fields that were oriented, band-pass gabor filters, much like receptive fields in V1. Results like these have been interpreted to mean that sparse representation of natural scenes is a central feature of the neural code, at least in V1 [153]. If this is true, we should be able to observe sparse responses to natural stimuli in V1 neurons, and this sparsity should depend on features or statistics of natural stimuli.

There is substantial direct experimental evidence for lifetime sparseness in the representation of natural scenes in V1 [13]. A single cell's lifetime sparseness is straightforward to see by examining repeated responses to some visual stimulus, and can be quantified using metrics related to the tails of the cell's response distribution [180]. Sparse representation in V1 is enhanced by stimulation of the extra-classical receptive field (i.e. the region beyond the classical, oriented and band-pass receptive field) [94, 160, 234, 235].

It is important to note that lifetime sparseness is not necessarily indicative of population sparseness. Correlations among cells' feature sensitivities or heterogeneity in tuning properties can obscure the relationship between sparseness at the single cell level and sparseness in a population code [243].

More recent advances including the development of fluorescent neural activity reporters and wide-scale cortical imaging [45, 111] have enabled researchers to directly observe population sparseness in V1 [84]. Phase-scrambling natural movies decreases observed population response sparsity in mouse visual cortex [84], suggesting that the higher order structure

present in natural scenes (and not just the amplitude spectrum) promotes a sparse representation. It will be important going forward to further probe the dependence of population sparseness on different features of natural stimuli.

RGC responses also show high lifetime sparseness in some species, including in Salamander retina [11, 27]. This is not just a feature of responses to natural stimuli, but appears to be a general feature of the responses of some ganglion cell types to many visual stimuli. Sparse RGC responses result in part from a rectifying nonlinearity in the retinal circuit or intrinsic to ganglion cells themselves [165]. Indeed in the retina there is some evidence that the sparseness of neural responses is more a function of the ganglion cell type than the particular stimulus used [120].

2.5 Natural stimuli and the receptive field

Since its earliest description, the visual receptive field has been a hugely useful concept in describing sensory processing. Initially, the receptive field was described as the region of visual space over which light is integrated to determine a neuron's response.

The earliest descriptions of the physiology of visual neurons in retina [96, 97, 125], lateral geniculate nucleus (LGN) [106], and primary visual cortex (V1) [105] made use of simple stimuli like spots and bars of light to map a neuron's receptive field. Later, the use of more complex visual stimuli, including natural stimuli, revealed that the receptive field depends on the statistics of the stimulus used to probe it [16]. We will return to this important point shortly.

Early receptive fields were little more than hypotheses that were consistent with the observations gleaned using a few visual stimuli like spots, bars of light, or gratings. A push to make the receptive field more concrete led to methods to systematically probe some region of stimulus space and "measure" a cell's receptive field. Here we'll touch on the progress that has been made in describing a receptive field, some of the features of the receptive field that have been revealed through the use of natural stimuli, and explore some of the limitations of the receptive field.

2.5.1 Measuring receptive fields - the legacy of linear systems analysis

The trend towards a precise, systematic measurement of a cell's receptive field birthed a host of techniques for describing visual receptive fields. Chief among them is a family of analyses that were inspired by linear systems analysis. The idea is to treat a receptive field as a linear filter in time and/or space (or, more generally, a combination of multiple linear filters). To measure these linear filters, simple parameterized stimuli are used to search stimulus space in an unbiased way and ask which regions of stimulus space tend to (de)activate the neuron. Examples of these stimuli include sinusoidal gratings [204, 233] or spatiotemporal white noise. We will omit many details of these sorts of analyses here, as several very nice treatments of them already exist [38, 194, 208].

Gaussian noise has practical advantages for measuring receptive fields. By presenting random stimuli to a visual neuron for long periods of time and examining the statistics of the stimuli that tend to modulate neural activity, one can compute the neuron's receptive field, which can be thought of as the stimulus that maximally activates the cell. If the receptive field is an accurate description of visual processing, we should be able to predict how a neuron will respond to a given stimulus by comparing that stimulus to the measured receptive field.

Techniques have been developed that attempt to correct for the correlations and asymmetries present in natural scenes in order to use them to measure receptive fields via reverse correlation techniques. In some cases these correction methods, when applied to V1 neurons, seem to work in the sense that the first order structure of receptive fields measured using natural and synthetic stimuli are qualitatively similar [178, 212, 219, 221]. However, even in these early studies there were signs that there could be systematic differences between receptive fields measured using natural and artificial stimuli.

It is important to keep in mind that this sort of comparison between linear receptive fields measured under artificial and natural conditions is not necessarily the right way to test the hypothesis that neural computation is similar under these conditions. This is because in both

cases we are assuming a model of the neuron that relies on linear spatiotemporal filtering of the stimulus at the first stage of processing. If nonlinearities substantially contribute to a neuron's response and if natural stimuli engage such nonlinear mechanisms differently than do artificial stimuli, a comparison of the linear receptive fields might not reveal these differences. A better approach to this question would be to test whether the inferred receptive field can be used to predict responses to both artificial and natural stimuli.

2.5.2 The receptive field is stimulus dependent

Mounting evidence suggests that nonlinearities in neural circuits can substantially shape neural responses - indeed the presence of even a static, post-filtering nonlinearity can distort the receptive field estimate from natural scene stimulation [209]. For example, David and colleagues fit a nonlinear RF model to V1 simple and complex cells using either a series of gratings or a natural movie and found that the recovered receptive field differed substantially under these two conditions [55]. Such differences between artificial and natural stimulus conditions seem to depend on nonlinearities in the circuit, with greater differences seen for complex cells in V1 compared to simple cells [72]. This phenomenon is not a special feature of natural images, as two different artificial stimuli reveal different receptive field structures in superficial cortical neurons but not neurons in layer 4, where presumably fewer chances for nonlinear interactions have arisen [246]. Findings like these have highlighted the malleability of the receptive field [16].

Even in the retina, which is considered much simpler and more linear than visual cortex [204], receptive field structure can show substantial stimulus dependencies. For example, magnocellular-projecting On-center parasol RGCs show strong spatial nonlinearities when presented with gratings stimuli, but are well approximated with a spatially linear receptive field in the context of natural scenes. Off-center parasol RGCs, on the other hand, show nonlinear responses to both grating and natural stimuli [226] (see Chapter 3). RGC temporal receptive fields also differ when measured using a naturalistic stimulus compared to a Gaussian noise stimulus [129].

These results show that the concept of a receptive field is somewhat fluid and the choice of stimuli used to probe the receptive field can dramatically influence what is inferred about its structure. The stimulus dependence of a receptive field limits its utility as the basis for a neural model - since we would like to be able to construct a model of a cell that is flexible enough to account for responses to any arbitrary stimulus. Going forward it is important to keep in mind that except in the rare case of a perfectly linear neuron, one cannot measure a receptive field that will be versatile enough to describe neural computation under any stimulus condition. Thus it will be important to explore more complex and flexible neural models, and to fit them using many stimuli, including natural scenes.

2.6 Predicting neural responses to natural visual stimuli

Constructing a computational model that can predict neural responses to natural visual stimuli is important for a number of reasons. First, the practical: in order to design and build prosthetic devices that can replace a damaged or diseased visual circuit, we need to know how that circuit processes the sorts of visual stimuli encountered in the real world. The most commonly discussed such device is a retinal prosthetic for use in patients with some form of retinal degeneration [110, 132]. To build an effective retinal prosthetic, we must first understand what information higher visual areas expect to receive from the retina and how that information is encoded.

A useful analogy here is the cochlear implant, which has in many respects been hugely successful development in the treatment of hearing impairments. In that case the processing strategy to implement is relatively straightforward: the amplitude across frequency bands is converted to electrical stimulation along a linear array of electrodes that stimulate the spiral ganglion neurons that form the auditory nerve. While cochlear implants cannot match the performance of a healthy cochlea [225], they work remarkably well. This success is in large part due to our firm grasp of the simple tonotopic coding scheme that exists in the cochlea. Unfortunately no such understanding of retinal processing exists. Part of this stems from the massively parallelized nature of retinal output, with upwards of 40-50 RGC classes each

transmitting different features of the visual world [15]. But even within a single cell class, we are unable to predict neural responses to natural scenes as well as we can predict responses to some synthetic stimuli.

The second reason a predictive model of visual neurons is important is because such a model is a compact representation of what we know about visual computation. The fact that current predictive models cannot generalize to complex visual inputs, including natural scenes, indicates that we do not understand something fundamental about how visual circuits work.

While not common, some researchers have directly tested the ability of a model to predict responses to full, time-varying natural stimuli. For example, Heitman and colleagues fit a generalized linear model (GLM), which is a descendant of earlier linear nonlinear (LN) models, to primate RGC responses [99]. They presented both spatiotemporal white noise stimuli as well as natural images under the influence of eye-movement like translations. GLMs consistently performed better when tasked with predicting responses to white noise compared to natural scenes. Interestingly, the GLM predictions did slightly improve when the models were fit to natural scenes. This study touches on two important issues. The first is that the assumption of a linear spatiotemporal front-end may be fundamentally flawed, and that the limitations of this linear assumption are especially problematic in the context of naturalistic stimulation (see also [226]). Second, it might generally be important to fit computational models to a range of stimuli, including natural scenes, in order to explore the relevant stimulus space efficiently.

The lasting influence of models with a linear spatiotemporal front end is largely due to the computational advantages of fitting such models. Recently, however, techniques have been developed to fit models to neural responses that don't make such strong assumptions about neural processing. For example, procedures used for fitting convolutional neural network models have recently been applied to retinal data. A convolutional neural network model of salamander RGCs impressively outperformed LN and GLM models in predicting responses to natural scenes [144], and a stacked LN-LN model could outperform linear front end models

with less fitting data [134]. Still, the overall accuracy with which these models predicted natural stimulus responses was limited. These models likely performed better than linear front-end models because they entail much less strict assumptions about retinal computation (e.g. they do not assume linear spatiotemporal filtering at the front end). Moreover the model architecture in both of these studies is reminiscent of the nonlinear subunit structure of RGCs [101].

We have an abundance of information about retinal circuits and computation, but one would not know this from looking at state of the art predictive models. Incorporating known circuit nonlinearities and adaptive mechanisms into predictive models may vastly improve them, but this has very rarely been tested. Fortunately more techniques for constraining models with known circuit organization are emerging [82, 172]. Moreover, incorporating even basic features of neural processing like localized receptive fields can make model fitting much more efficient [159]. We suspect that, compared to discovering new exotic mechanisms that can complicate neural responses, incorporating known circuit architecture, nonlinear processing, and adaptational mechanisms will be much more important in the development of truly predictive models.

Chapter 3

SYNAPTIC RECTIFICATION CONTROLS NONLINEAR SPATIAL INTEGRATION OF NATURAL VISUAL INPUTS

The study presented in this chapter has been published as:
Turner, M.H., and Rieke, F. (2016). Synaptic rectification controls nonlinear spatial integration of natural visual inputs. Neuron. 90, 1257-1271.

3.1 Summary

A central goal in the study of any sensory system is to predict neural responses to complex inputs, especially those encountered during natural stimulation. Nowhere is the transformation from stimulus to response better understood than the vertebrate retina. Nevertheless, descriptions of retinal computation are largely based on stimulation using artificial visual stimuli, and it is unclear how these descriptions map onto the encoding of natural stimuli. We demonstrate that nonlinear spatial integration, a common feature of retinal ganglion cell (RGC) processing, shapes neural responses to natural visual stimuli in primate Off parasol RGCs, whereas On parasol RGCs exhibit surprisingly linear spatial integration. Despite this asymmetry, both cell types show strong nonlinear integration when presented with artificial stimuli. We show that nonlinear integration of natural stimuli is a consequence of rectified excitatory synaptic input, and that accounting for nonlinear spatial integration substantially improves models that predict RGC responses to natural images.

3.2 Introduction

Our understanding of the computations performed by sensory neurons has benefitted from many decades of study using artificial stimuli designed to probe particular features of a

sensory receptive field (RF). For example, the RF structure of retinal ganglion cells (RGCs) has been well characterized using artificial visual stimuli, beginning with spots and annuli to reveal center-surround organization [125, 20]. Other studies have used moving stimuli to measure direction selectivity [22], spatiotemporal white noise to characterize linear feature selectivity and nonlinear processing [146, 38, 164], and fine gratings to describe nonlinear spatial integration [69, 101, 100]. These studies reveal a variety of non-classical RF features that can be exercised using appropriately designed stimuli (for review see [115, 87]), but understanding the roles of these RF features in visual function requires knowing how they shape the encoding of natural inputs.

A well-known failure of classical RF models is the inability to capture nonlinear integration of signals across space [69, 215, 60, 162, 31]. Nonlinear integration is mediated by subunits in the RF that provide rectified input to the RGC [100, 232, 60, 48, 192]. In short, this means that each subunit can effectively transmit preferred contrast inputs but that the suppressive impact of non preferred contrast inputs will be minimal. Such subunit structure limits cancelation of light and dark regions in a scene and thus shapes RGC responses to spatially structured stimuli - e.g. causing the RGC response to depend on the spatial distribution of light, not just the total amount of light, within the RF center. At least in some cases, bipolar cells presynaptic to the RGC are the physiological basis of the rectified RF subunits [61, 192, 82, 126].

Thus, an impressive body of work demonstrates the ubiquity of nonlinear spatial integration, the circuit and synaptic mechanisms that give rise to a nonlinear RF, and some of the computations that a nonlinear RF can support [155, 86, 192, 126] (for review see [87, 193]). Beyond the retina, the classical energy model of complex cells in V1 also relies on nonlinear transformation of multiple linear filter outputs before integration [2], and more recent evidence suggests that a nonlinear subunit RF structure is appropriate for both simple and complex cells in primary visual cortex [187, 237, 236].

Nonlinear spatial integration is absent from most predictive models for RGC responses. Current models can accurately predict responses to the same sort of visual stimuli on which

they were trained [116, 163]. Yet such models are much less successful in generalizing to predict responses to natural stimuli [99] (for review, see [36]). Most predictive RGC models incorporate an initial linear filtering step that accounts for feature selectivity in time and/or space (i.e. a linear RF). But the many studies demonstrating nonlinear spatial integration suggest that a linear spatial RF may not be an appropriate description. It is unclear, however, whether nonlinear responses are primarily elicited by artificial stimuli designed to test spatial linearity or are instead elicited by natural stimuli. The aim of this study is to help bridge the gap between this classical understanding of nonlinear spatial integration and modern computational models.

Our central goals were to answer three related questions about the importance of nonlinear spatial integration in RGC encoding of natural inputs: (1) Does nonlinear spatial integration affect RGC responses to natural stimuli? (2) What circuit or synaptic mechanisms control spatial integration of natural scenes? And, (3) does incorporation of nonlinear RF structure substantially improve models that predict responses to natural stimuli? We will answer these questions for responses of On and Off-center parasol (magnocellular-projecting) RGCs in the monkey retina. Parasol RGCs provide key signals for visually-guided behavior and exhibit pronounced spatial nonlinearities [162, 48, 35]. This makes them an ideal class of RGC in which to answer these questions.

3.3 Results

We used a combination of in vitro recording and RF modeling to explore whether and how nonlinear spatial integration in the RF center shapes parasol RGC responses to natural visual stimuli. We focused on spatial integration within the RF center for two reasons. First, the RF center of parasol RGCs can show strong spatial nonlinearity when probed with artificial stimuli [162, 48, 35] (see also Fig. 3.5), but the impact on coding of natural stimuli is unknown. Second, understanding coding in the RF center is a tractable but essential step towards the broader question of natural stimulus encoding.

We start by showing that Off, but not On, parasol RGCs nonlinearly integrate natural

visual inputs, although both cell types exhibit strong spatial nonlinearities when probed with artificial stimuli. This On/Off asymmetry persists across a range of light levels. We then provide evidence that the asymmetry in natural stimulus integration results in part from differences in rectification of the cells' excitatory synaptic inputs. Finally, we show that including rectified subunits in a RF model can substantially improve predictions of Off parasol RGC responses to natural images.

3.3.1 Spatial integration of natural retinal inputs

To begin, we designed a stimulus to test whether a linear model of spatial integration could account for responses of parasol RGCs to natural stimuli. This stimulus was based on an online database of natural scenes and measured eye movements called the Database Of Visual Eye movementS (DOVES, [227], live.ece.utexas.edu/research/doves). The DOVES database contains eye tracking data measured from human observers freely viewing grayscale images from the van Hateren natural image database [229]. Using these data, we reconstructed the image impinging on the retina during these periods of natural viewing (Fig. 3.1A,B). Our DOVES-based movie stimuli capture several important features of natural retinal input, including the statistics of natural images, fixation lengths, and fixational eye movements (see Experimental Procedures for details).

To probe integration in the RF center with minimal activation of the surround, we presented DOVES movies with an aperture that restricted the image to the RF center. We scaled the DOVES movies to a mean luminance level of ∼4,000 cone opsin isomerizations (R*)/M or L cone/s), which corresponds to low-photopic light levels. During presentation of DOVES stimuli, we set the mean monitor intensity (i.e. the mean gray level of the aperture around the image and of the blank screen between trials) to 4000 R*/cone/s.

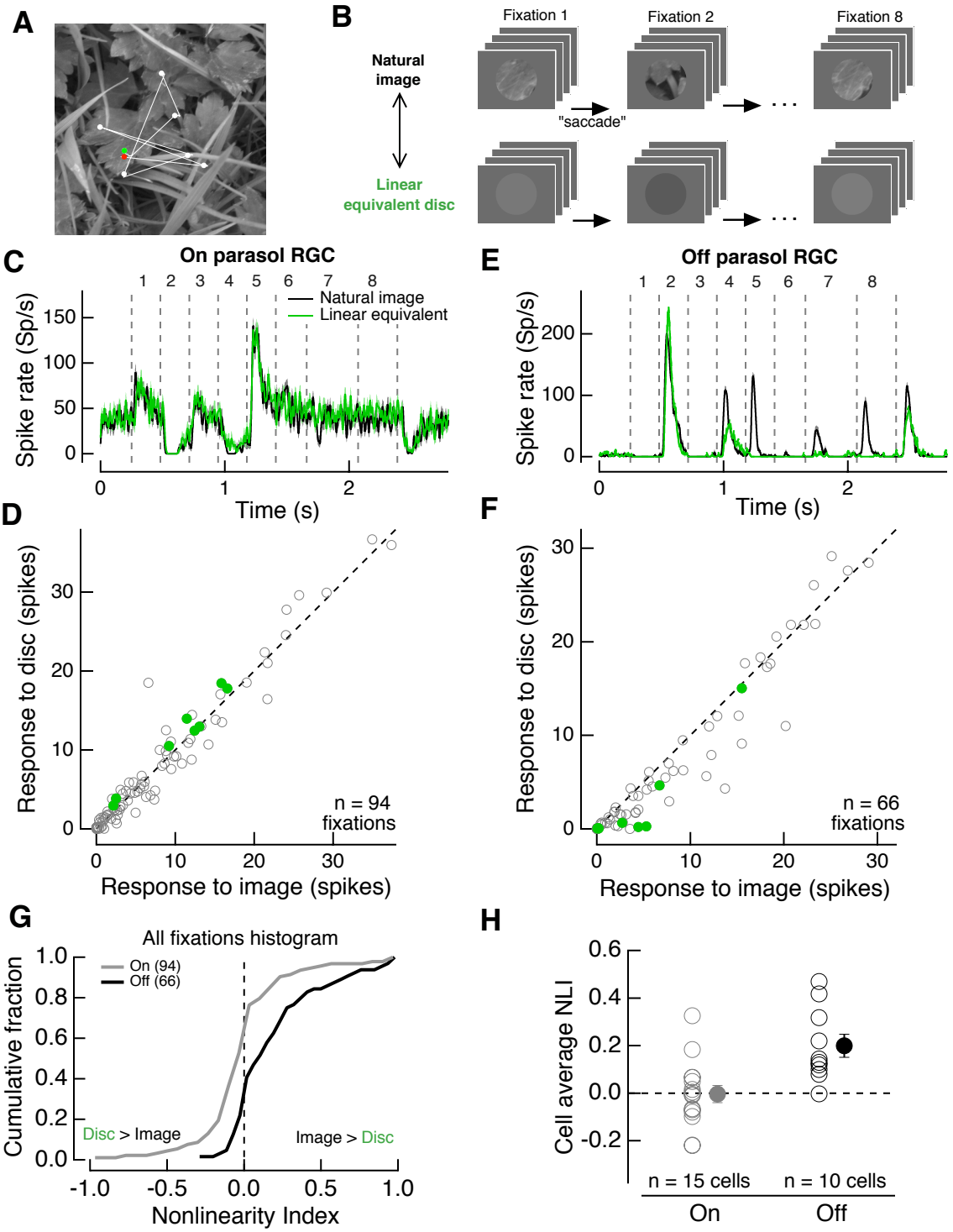
A
B
Natural image
Linear equivalent disc
Fixation 1
Fixation 2
Fixation 8
"saccade"
C
On parasol RGC
Natural image
Linear equivalent
Spike rate (Sp/s)
Time (s)
E
Off parasol RGC
Spike rate (Sp/s)
Time (s)
D
Response to disc (spikes)
Response to image (spikes)
n = 94 fixations
F
Response to disc (spikes)
Response to image (spikes)
n = 66 fixations
G
All fixations histogram
On (94)
Off (66)
Cumulative fraction
Disc > Image
Image > Disc
Nonlinearity Index
H
Cell average NLI
n = 15 cells
n = 10 cells
On
Off

Figure 3.1: (Previous page) **Natural retinal inputs drive nonlinear responses in Off but not On parasol RGCs.** (A) A natural image and associated eye movement trajectory from the DOVES database. White dots indicate fixation points, green and red dots indicate the first and last fixations. (B) Top: Schematic illustrating the DOVES movie, restricted to the RF center. Bottom: A linear equivalent disc stimulus movie was generated using a circular Gaussian RF sampling each frame in the natural movie. (C) Example On parasol RGC spike responses to these stimuli. Vertical dashed lines indicate frames at which a change of fixation ("saccade") occurred. Numbers above indicate fixations. (D) On parasol RGC population data showing mean spike count during each fixation (94 total fixations from 15 RGCs) in response to DOVES and associated linear equivalent disc movies. Open circles correspond to one fixation measured in one cell, green circles indicate the eight fixations from the example cell in (C). (E-F) Same as (C-D) for an example Off parasol RGC and population data (66 fixations from 10 RGCs). (G) Cumulative histogram of the nonlinearity index (NLI, see text) for every fixation response measured in On (gray trace) and Off (black trace) parasol RGCs. (H) For each cell, we averaged the NLI over all presented fixations. Open circles indicate cell-average NLIs for individual cells, closed circles show population mean ± S.E.M.

We designed a second stimulus to explicitly test linear spatial integration of the DOVES movies. This stimulus was a uniform disc whose intensity was equal to the average intensity of the corresponding frame in the original movie weighted by the cell's RF center; we refer to this stimulus as the linear equivalent disc movie. To construct the linear equivalent disc movie, we sampled each frame of the original movie with a circular Gaussian function and linearly integrated the (weighted) pixel values to yield a single intensity value for each frame (Fig. 3.1B, bottom). The size of the Gaussian RF center was determined using the area-summation curve generated by presenting spots, centered over the RF, of various sizes (e.g. see Fig. 3.5). The key property of this manipulation is that a RGC that integrates spatial inputs according to a linear Gaussian RF will produce the same response to the original movie ("Natural Image") and the sequence of linearly integrated intensity values ("Linear Equivalent Disc").

For each cell, we presented one of ten reconstructed DOVES stimuli (each movie was generated using a different natural image and/or measured eye movement trajectory). For many fixations, On parasol RGCs generated near-identical responses to the original DOVES movie and the linear equivalent disc movie (Fig. 3.1C,D). Spike responses to these two stimuli

did not differ systematically across all measured fixations (Fig. 3.1D; $p = 0.65$). However, the same stimuli often produced large differences in the spike responses of Off parasol RGCs (Fig. 3.1E,F). The linear equivalent disc and the natural movie produced approximately equivalent responses for some fixations (points along unity in Fig. 3.1F), but responses to other fixations differed substantially (points below unity in Fig. 3.1F). On average, Off parasol RGCs responded more strongly to the DOVES movie than to the linear equivalent disc movie (Fig. 3.1F; $p = 8.7 \times 10^{-6}$).

To quantify the deviations from linear integration, we defined a nonlinearity index (NLI) as

$$NLI = \frac{r_{image} - r_{disc}}{|r_{image}| + |r_{disc}|} \quad (1)$$

where, for each fixation, r_{image} and r_{disc} are the average (over repeated trials) spike count in response to the original DOVES movie and to the linear equivalent disc movie. Thus a NLI near 0 indicates near-linear spatial integration, negative NLIs indicate stronger responses to the linear equivalent disc movie and positive NLIs indicate stronger responses to the DOVES movie. The absolute value operators in the denominator ensure that this sign relationship holds for responses that can take negative values (e.g. synaptic current responses in Fig. 3.2).

Figure 3.1G shows the cumulative distributions of NLI values for all fixations and all recorded On and Off parasol RGCs. NLI values for both cell types varied considerably across fixations; NLI values in Off parasol RGCs were shifted to positive values relative to On parasol RGCs (median NLI of 0 in On versus 0.12 in Off parasol RGCs, $p = 6 \times 10^{-6}$, Wilcoxon rank sum test on the distributions in Fig. 3.1G). We averaged NLIs across fixations for each cell to yield the cell-average NLI (Fig. 3.1H). On parasol RGCs had an average NLI that was not significantly different than zero ($p = 0.54$), whereas the average NLI for Off parasol RGCs was positive and significantly larger than that for On parasol RGCs ($p = 0.002$). Thus, the linear equivalent disc effectively mimicked natural stimuli for On parasol RGCs, but the same linear prediction systematically underestimated responses of Off parasol RGCs to natural stimuli.

Off parasol RGC spike responses were typically more transient than On parasol RGC responses to the same stimuli (e.g. compare Fig. 3.1C and E). This was also true for artificial stimuli like gratings (Fig. 3.5) and spots (Fig. 3.6). Despite this difference, we saw no obvious history effects in responses to DOVES movies that could account for the On/Off asymmetry in spatial integration. Moreover, as we show below, the asymmetry in spatial integration is present in responses to flashed natural images from a mean gray background (Fig. 3.3).

Each cell was presented with a small number of fixations (typically 5-7 in a single DOVES movie), whereas the variability in spatial structure within and across natural images is quite large. This means that it is difficult to identify features of the DOVES movies that drive deviations from linearity in Off parasol RGC responses. Another consequence of this limited sampling is that much of the variability in cell-average NLIs (Fig. 3.1H) is due to variability across natural image stimuli rather than biological variability in the parasol RGC population. This is illustrated in Fig. 3.9, which shows that NLI values for a single cell can vary considerably across movies. We will return later to the question of what features of an image produce deviations from linear spatial integration.

Could a failure to accurately capture the linear RF account for the differences between the responses to DOVES and linear equivalent disc movies? Three observations indicate that this is not the case. First, the systematic undershoot of the linear model across fixations is characteristic of a failure of (linear) cancelation of positive and negative contrasts and is not predicted by errors in linear RF estimation. Second, there is no reason to suspect that our linear RF estimation is systematically wrong for Off but not On parasol RGCs. Third, we can account for most of the failures of linear integration in the context of natural scenes using a simple nonlinear subunit model (Figs. 3.3, 3.4, 3.7).

A spatially linear RF allows responses to light and dark regions in a scene to cancel. For a spatially nonlinear RF, however, such cancelation is incomplete. This difference is a consequence of rectified subunits within the RF: because of rectification, subunits that are suppressed by non-preferred contrast do not effectively transmit a suppressive signal to the

RGC, while subunits that are activated transmit a strong positive signal. Thus the smaller Off parasol RGC responses to the linear equivalent disc compared to the structured image patches are consistent with an impact of rectified subunits in the RF center.

If the nonlinear responses are due to presynaptic cells acting as rectified subunits, we should see nonlinear responses in the synaptic inputs to parasol RGCs and such nonlinearities should differ for On and Off RGCs. To test these hypotheses, we presented DOVES movies and corresponding linear equivalent disc stimuli while measuring synaptic currents in whole-cell voltage-clamp recordings (Fig. 3.2). We isolated inhibitory and excitatory synaptic currents using holding potentials near the cation and chloride reversal potentials (see Experimental Procedures). We then compared the charge transfer (integrated current) in response to DOVES and linear equivalent disc movies. The excitatory synaptic inputs to On parasol RGCs show relatively small, but significant, deviations from linear integration (Fig. 3.2A,C; $p = 0.02$), whereas the inhibitory inputs to the same cells show strong nonlinear integration (Fig. 3.2A,C; $p = 1.8 \times 10^{-5}$). For Off parasol RGCs, both excitatory and inhibitory inputs are larger in response to the original natural movie (Fig. 3.2B,D; excitatory charge transfer: $p = 8 \times 10^{-5}$, inhibitory charge transfer: $p = 3 \times 10^{-5}$). Fig. 3.2E shows an asymmetry in the degree of nonlinear spatial integration in these different inputs: NLIs for Off parasol RGC excitatory inputs were strongly skewed, with 88% of fixations producing positive NLIs, whereas NLIs for On parasol RGC excitatory inputs were more symmetrical, with positive NLIs for 64% of fixations. Cell-average NLIs were significantly more positive for Off than On RGC excitatory inputs (Fig. 3.2F; on average $\sim$5 times greater for Off than On RGCs, $p = 0.01$).

These data show that spatial integration in parasol RGC spike outputs qualitatively resembles that in the cells' excitatory synaptic inputs [47]. While there is an asymmetry in the strength of nonlinear integration in the excitatory inputs to On and Off parasol RGCs, the excitatory inputs to both cell types deviate substantially from linear spatial integration. Indeed, neither the excitatory nor inhibitory synaptic inputs to On parasol RGCs are as linear as their spike outputs. Moreover, the inhibitory input to On and Off parasol RGCs

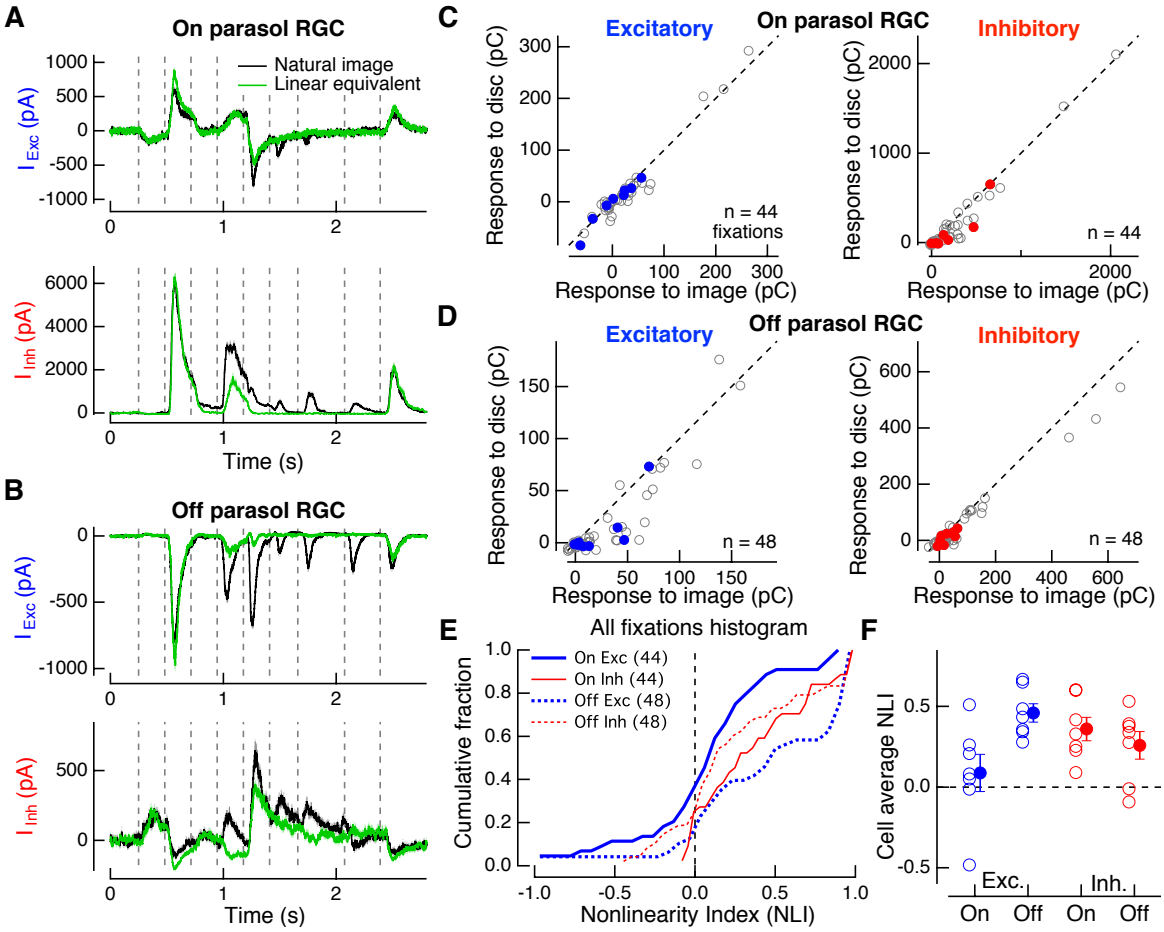

Figure 3.2: **Nonlinear synaptic current responses to natural visual stimuli.** (A) Mean synaptic current responses - in the same example On parasol RGC as in Fig. 3.1 - to DOVES and linear equivalent disc movies. (B) Same as A, for an example Off parasol RGC's synaptic inputs. (C) Analogous to Fig. 3.1D. Average excitatory charge transfer (integrated current) in response to the DOVES movie and associated linear equivalent disc movie; each point corresponds to one fixation (44 fixations from 7 RGCs). Colored points correspond to the example On parasol RGC in (A). Left: excitatory charge transfer, right: inhibitory charge transfer. (D) Same as (C) for Off parasol RGC synaptic inputs (48 fixations from 7 RGCs). Colored points indicate example Off parasol RGC in (B). (E) Cumulative histograms of NLI values for all fixations measured in On and Off parasol RGC excitatory and inhibitory currents. (F) Cell-average NLI values, 7 cells in each case.

is differentially driven by these natural stimuli. In particular, the inhibitory input to On parasol RGCs is stronger than inhibition to Off parasol RGCs (compare the example traces in Fig. 3.2A and B as well as the population data in Fig. 3.2C and D). These observations suggest that synaptic integration and intrinsic cell properties shape spatial encoding by On parasol RGCs. We will return to this issue in the Discussion.

3.3.2 A nonlinear subunit model predicts failures of linear integration

Our experiments with DOVES stimuli (Fig. 3.1 & 3.2) show that a spatially linear RF model fails to predict responses of Off parasol RGCs to natural inputs. To test the ability of simple models to capture these failures of linear integration, we turned to a stimulus paradigm that does not include eye movements. These stimuli were stationary image patches selected from a larger natural image [229, 227](Fig. 3.3A). This simpler stimulus allowed us to focus on spatial integration in isolation, without history dependence created by interactions between eye movements and the spatial structure of natural scenes [124, 183].

We compared how well two models captured failures of linear integration: a linear-nonlinear (LN) RF model and a nonlinear subunit RF model (Fig. 3.3B; see Experimental Procedures for details). Image patches were passed through a square grid of RF subunits (Fig. 3.3A, inset), each modeled as a circular Gaussian. The subunit models shown in Fig. 3.3B are illustrated with On (positive contrast-preferring) subunits, but models with Off (negative contrast-preferring) subunits produced similar results.

In the LN model, the output of each subunit was weighted by a larger Gaussian representing sampling in the RF center (Fig. 3.3A, inset). The subunit outputs were linearly summed and passed through a rectifying output nonlinearity. In the nonlinear subunit model, the output of each subunit was passed through the same rectifying nonlinearity before being weighted and summed with the other (rectified, weighted) subunit outputs. These models were purposefully designed to be simplistic and minimal, with the key difference being the site of nonlinear transformation: before (nonlinear subunit model) or after (LN model) integration by the RGC. Models were not fit directly to RGC responses. Instead, we set each

subunit diameter to be 1/6 of the RF center diameter, consistent with the subunits sizes measured in Fig. 3.5 below. The behavior of the model was similar as long as the subunits were $\sim$1/2 the width of the RF center or smaller (See Fig. 3.10).

Outputs of each RF model were computed for 10,000 randomly-selected patches from a natural image. The nonlinear subunit model always produced greater or equal responses than the LN model, with some image patches producing much stronger responses in the nonlinear subunit model than the LN model (Fig. 3.3C). This can also be seen in a histogram of model response differences (nonlinear subunit minus LN, Fig. 3.3D, gray trace). Relatively homogeneous patches drive the models similarly while patches with both positive and negative contrast regions produce stronger responses in the subunit model than the LN model (Fig. 3.3C, insets). This is because rectification of subunit outputs prevents cancelation of positive and negative contrast regions.

If nonlinear subunit RF organization underlies failures of linear integration, we should be able to predict which image patches produce nonlinear responses in parasol RGCs. To test this, we selected patches from natural images and presented them to On and Off parasol RGCs while measuring spike responses. Image patches were not randomly drawn for these experiments (unlike Fig. 3.3C and the gray trace in Fig. 3.3D); instead one patch was drawn from each bin of the full histogram such that the patches chosen uniformly span the observed range of differences in model outputs (black trace in Fig. 3.3D, note that the displayed bin size has been increased for clarity). This allowed us to explore the entire range of nonlinear response strengths within a single experiment.

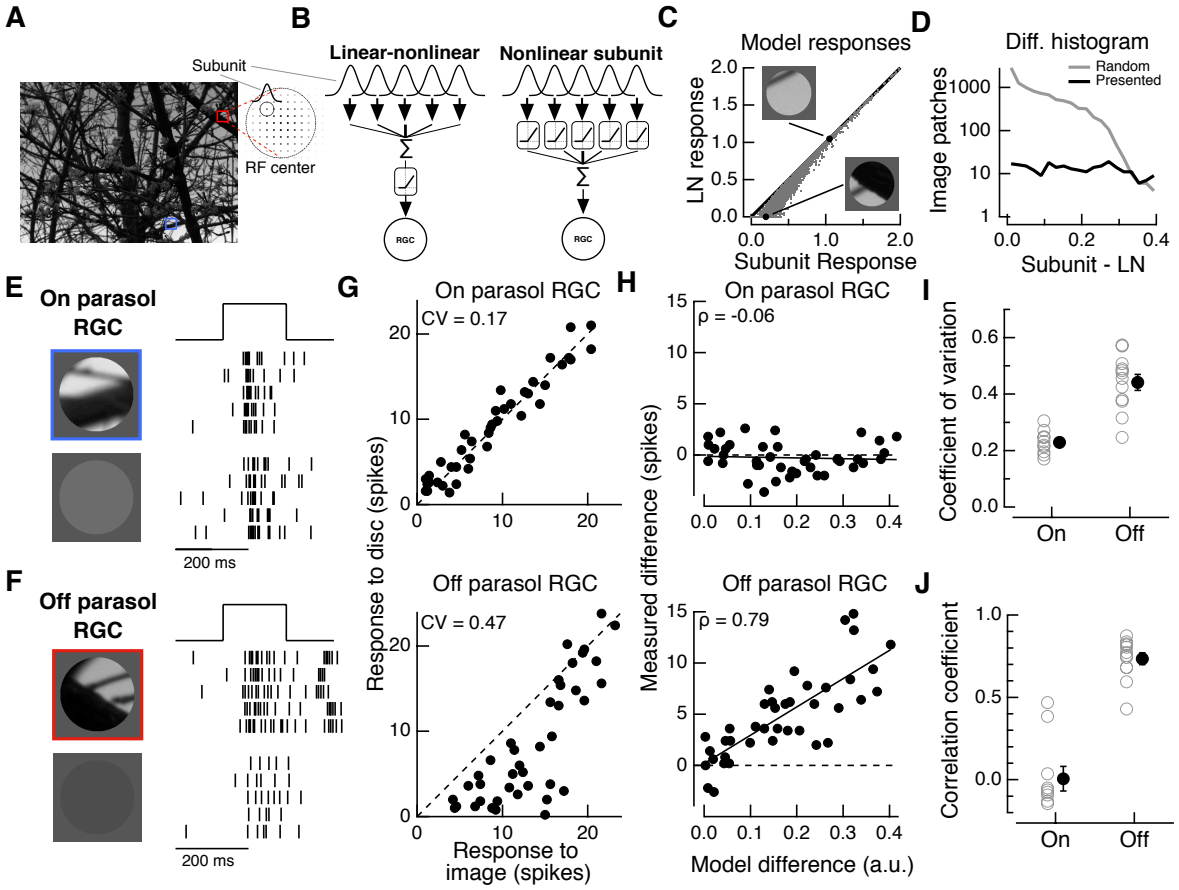
A
Subunit
RF center
B
Linear-nonlinear
Nonlinear subunit
RGC
RGC
C
Model responses
LN response
Subunit Response
D
Diff. histogram
Random
Presented
Image patches
Subunit - LN
E
On parasol RGC
200 ms
F
Off parasol RGC
200 ms
G
On parasol RGC
CV = 0.17
Off parasol RGC
CV = 0.47
Response to disc (spikes)
Response to image (spikes)
H
On parasol RGC
ρ = -0.06
Off parasol RGC
ρ = 0.79
Measured difference (spikes)
Model difference (a.u.)
I
Coefficient of variation
On
Off
J
Correlation coefficient
On
Off

Figure 3.3: (Previous page) **Failures of linear integration in Off parasol RGCs is due to nonlinear subunit RF structure.** (A) Example of a natural image used in the modeling and experiments below. Diagram shows the spatial arrangement of subunits used in the RF models outlined in (B). (B) In both models, the RF center is composed of subunits. The linear-nonlinear (LN) RF model (left) takes the linear sum of subunit outputs and passes that sum through a rectified-linear output nonlinearity to yield the response. The nonlinear subunit RF model (right, "subunit") applies the rectified-linear nonlinearity at the output of each subunit before summation at the ganglion cell. (C) Model outputs for 10,000 randomly-selected patches (gray points) from the image in (A). Black line indicates unity. (D) Histogram of the model response differences (subunit model response minus LN model response) for randomly-selected image patches (gray line). For the following experiments, we sampled these image patches to uniformly span the range of model differences (black line). (E) Spike rasters from an example On parasol RGC in response to a flashed stationary natural image patch (top) and its linear equivalent disc (bottom). Colored outlines indicate the region of the scene in (A) from which this patch was drawn. (F) Same as (E) for an example Off parasol RGC. (G) Mean spike counts (over five presentations of each patch) in response to 40 different patches from a natural image, for the example On parasol RGC (top) and Off parasol RGC (bottom). Dashed line indicates unity. (H) We subtracted the response to each linear equivalent disc from the response to its associated image, giving us a measure of the strength of nonlinear integration ("measured difference") and compared that to the difference in model outputs for each image patch presented. Solid lines show linear fits, with linear correlation coefficient indicated at top left. (I) Population data showing the coefficients of variation (standard deviation between image and disc responses divided by average image response) for On and Off parasol RGCs. Open symbols correspond to individual cells, closed symbols denote mean ± S.E.M. (10 On parasol RGCs, 12 Off parasol RGCs). (J) Population data for On and Off parasol RGCs showing linear correlation coefficients between experimentally measured differences and model response differences.

We presented each patch, restricted to the RF center, for 200 ms. We then presented the corresponding linear equivalent disc for each image patch (Fig. 3.3E,F). As in Fig. 3.1, On parasol RGC responses to the linear equivalent disc were similar to responses to the corresponding natural image (Fig. 3.3G, top). However, for many patches, Off parasol RGCs responded more weakly to the linear equivalent disc than to the image (Fig. 3.3G, bottom). Responses that deviated from linearity tended to be those of weaker or intermediate response strengths. This behavior is seen in the responses to DOVES movies (Fig. 3.1F) and is predicted by the model (Fig. 3.3C). The stronger deviations for intermediate response strengths are likely because the patches that strongly drive the cell are often uniformly the preferred contrast, and thus there is little opportunity for cancelation of light and dark

regions of the scene. On average the difference between the responses to the flashed image patches and linear equivalent disc stimuli was greater in Off than On parasol RGCs (Fig. 3.3I, $p = 2 \times 10^{-4}$).

For each image patch, we calculated the difference between the spike response to the image and to the linear equivalent disc, and compared this to the difference in subunit and LN model outputs to the same patch. For On parasol RGCs, the measured differences were near zero, and there was little correlation between measured response differences and model response differences (Fig. 3.3H, top). Off parasol RGCs, however, showed a strong correlation between measured and modeled response differences (Fig. 3.3H, bottom, Fig. 3.3J). Thus, the same image patches that produced nonlinear responses in Off parasol RGCs produced stronger responses in the nonlinear subunit model than in the LN model.

The experiments in Fig. 3.3 were performed at low-photopic light levels of ∼4,000 R*/cone/sec. Previous work has shown that mean luminance can change the spatial processing properties of RGCs [20], including nonlinear spatial integration [93]. These observations suggest two questions: (1) Does the observed On/Off asymmetry persist when we change the mean light level? (2) Do changes in mean light level change the spatial RF structure in ways that impact natural image encoding? To answer these questions, we repeated the experiments and analysis shown in Fig. 3.3 at "low" (∼500 R*/cone/s), "medium" (∼5,000 R*/cone/s, similar to the data in Fig. 3.3), and "high" (∼50,000 R*/cone/s) mean luminance levels. The correlation between measured response differences and model response differences was larger for Off compared to On parasol RGCs at the two highest light levels ($p < 0.005$ for both medium & high), but not at the lowest level (Fig. 3.4F). For Off parasol RGCs, this correlation value was smaller at the lowest light level compared to the higher light levels ($p < 0.005$ in each case). Thus, nonlinear spatial integration contributes less to Off parasol RGC responses to natural scenes at lower light levels. These experiments revealed an additional change in spatial processing in the On parasol RGCs. In particular, at the highest light levels, there was a tendency for the (relatively small) deviations from linear integration to negatively correlate with our subunit model predictions (Fig. 3.4F).

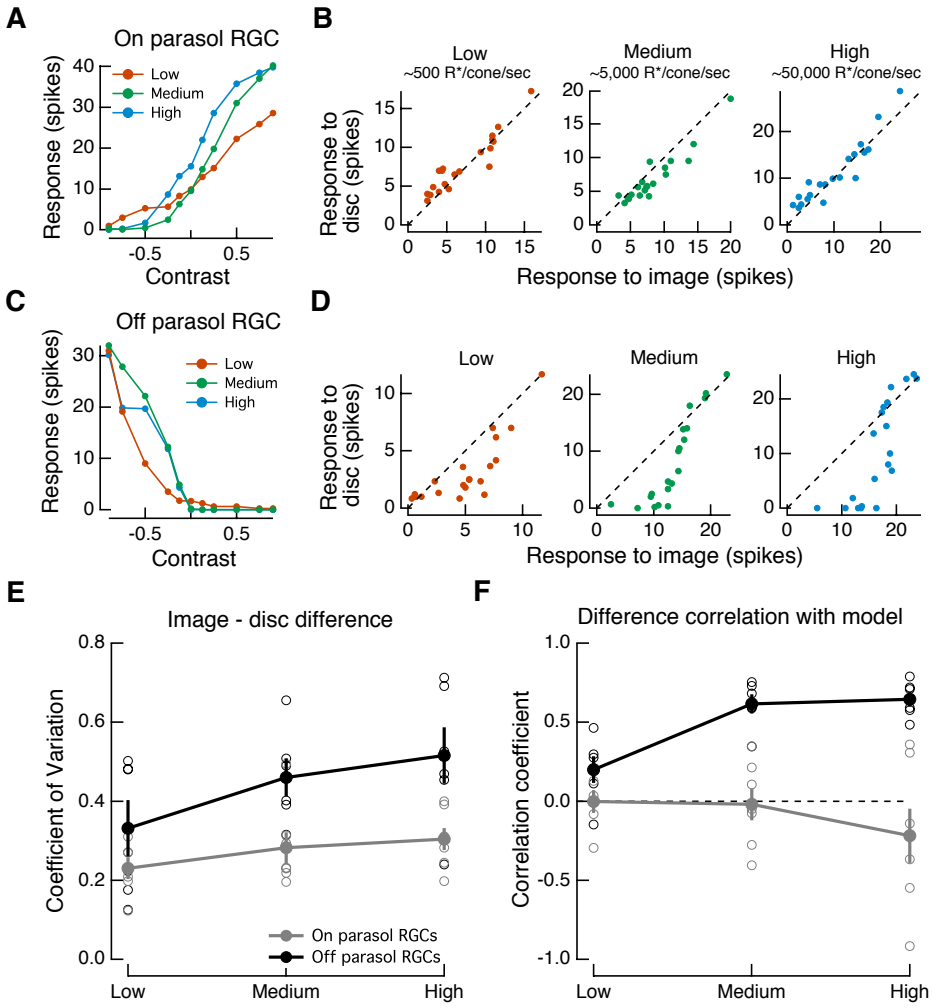

Figure 3.4: **Mean light-level dependence of nonlinear spatial integration of natural images.** (A) Contrast response functions measured in an example On parasol RGC at three luminance levels: high (∼50,000 R*/cone/s), medium (∼5,000 R*/cone/s) and low (∼500 R*/cone/s). (B) At the same three light levels, we measured the spike responses to flashed natural image patches and their corresponding linear equivalent discs, as in Fig. 3.3. (C,D) Same as (A,B) for an example Off parasol RGC. (E) Population data showing the coefficient of variation (CV) between image and disc responses (as in Fig. 3.3I) across these three luminance levels. (F) The linear correlation coefficient between measured differences and modeled differences (as in Fig. 3.3J) across luminance levels. For On parasol RGCs: n = 6, 7, 7 (low, medium, high); for Off parasol RGCs: n = 6, 6, 6 (low, medium, high).

This suggests a light-level dependent contribution from an additional nonlinear mechanism, operating at a small spatial scale - for example, adaptation to luminance or contrast in the photoreceptors.

3.3.3 On and Off parasol RGC receptive fields contain small nonlinear subunits

The asymmetry in the impact of nonlinear spatial integration on On and Off parasol RGC responses to natural images is puzzling given that both cell types exhibit nonlinear responses to high spatial frequency grating stimuli [48, 162]. Could a difference in nonlinear RF structure between On and Off parasol RGCs underlie the difference in spatial integration revealed using natural stimuli? To answer this question, we used grating stimuli to probe the spatial scale and contrast dependence of nonlinear RF subunits.

As in our experiments with DOVES movies and natural image stimuli, we started each recording by finding the center of the RF and measuring its size (Fig. 3.5A,B). To test for the presence of small nonlinear subunits in the RF, we centered a contrast reversing grating stimulus over the RF. A spatially linear RF will produce no response to the temporal modulation of this stimulus (Fig. 3.5C, left), but a RF composed of subunits with rectified output will produce a response at every half cycle of the temporal modulation (Fig. 3.5C, right). This frequency-doubled (F2) response is a classical signature of nonlinear spatial integration [69, 101].

We used stationary, contrast-reversing gratings modulated at 4 Hz to characterize the spatial properties of nonlinear RF subunits of On and Off parasol RGCs. For both cell types, spike outputs and excitatory and inhibitory synaptic inputs showed strong F2 responses to this stimulus (Fig. 3.5D-G). Varying the width of the bars composing the grating revealed that the F2 response in all cases was present at a spatial scale much smaller than the RF center size (Fig. 3.5D-G). This indicates that nonlinear RF subunits are several times smaller than the RF center. F2 responses of excitatory and inhibitory inputs peaked for bar widths $\sim$1/3rd of the RF center size and declined for larger bar widths. This characteristic dependence on bar width can be explained by a center-surround organization of the subunits.

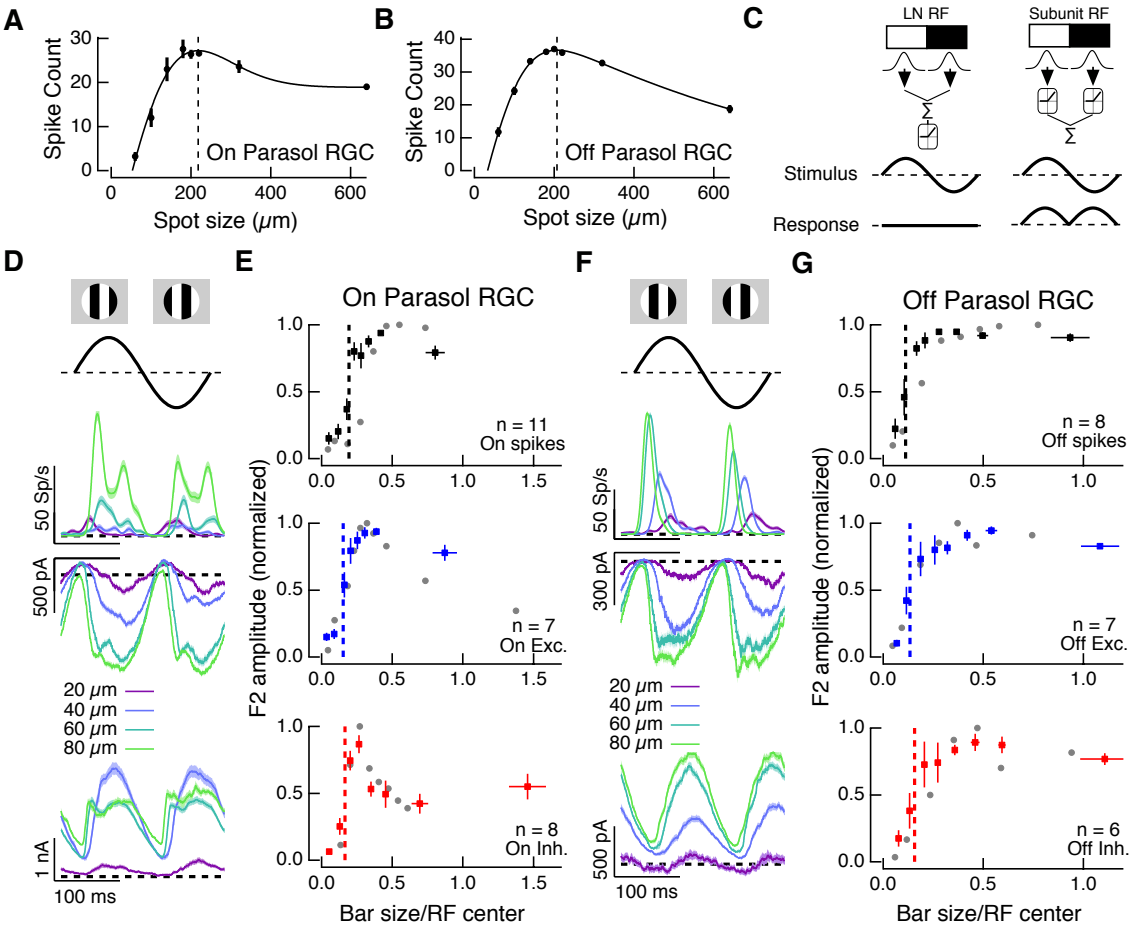

Figure 3.5: **On & Off parasol RGC receptive fields contain small nonlinear subunits.** (A) Area summation curve from an example On parasol RGC. Points indicate mean S.E.M., curve indicates the best fit difference-of-Gaussians model (see Experimental Procedures). Dashed vertical line indicates the RF center size of this cell. (B) Same as (A) for an example Off parasol cell. (C) For a grating stimulus centered over the RF, a LN RF produces no response to contrast modulation, while a nonlinear subunit RF produces a frequency-doubled (F2) response. (D) Cycle average responses in an example On parasol RGC. A contrast reversing grating stimulus, centered over the RF, produces robust F2 responses in spike output (top), excitatory input (middle) and inhibitory input (bottom) across a range of spatial frequencies. Different colored traces correspond to different bar widths of the grating stimulus. (E) On parasol RGC population summary data showing the nonlinear F2 response amplitude as a function of bar width. For each sampled bar width, the F2 response amplitude was calculated from 20-60 cycles of the stimulus. Within each cell, the bar width was normalized by the size of the RF center (measured as in A), to control for cell-type and eccentricity based differences in RF center size. Gray points indicate individual measurements from the example cell in (D), and colored squares indicate mean S.E.M. for equally populated bins. Vertical dashed lines indicate the (normalized) bar width that drove half-maximal F2 responses, on average. Top: spike responses, 11 cells. Middle: Excitatory synaptic currents, 7 cells. Bottom: Inhibitory currents, 8 cells. (F-G) Same as (D-E) for Off parasol RGCs. These cells also show small-scale nonlinear RF structure in their spike output (top, 8 cells) as well as their excitatory (middle, 7 cells) and inhibitory synaptic inputs (bottom, 6 cells).

These experiments show that both On and Off parasol RGCs exhibit nonlinear responses at a spatial scale substantially smaller than the RF center.

To test the contrast sensitivity of the nonlinear response, we presented sinusoidally-modulated uniform or split-field grating stimuli over the RF center at a range of contrasts (Fig. 3.11). Spike and synaptic current responses in both cell types showed nonlinear F2 responses across a range of contrasts. Off parasol RGCs showed 1.5- to 2-fold stronger F2 responses in spike output and excitatory synaptic input than On parasol RGCs. The reverse was true for inhibitory inputs to these cells. It is not clear, however, how these differences between On and Off parasol cell responses can explain the large asymmetry in spatial integration of natural images.

3.3.4 Asymmetric synaptic rectification in On and Off parasol RGC pathways

Figure 3.5 shows that grating stimuli reveal small nonlinear subunits in RFs of both On and Off parasol RGCs. If deviations from linear integration are the result of nonlinear subunit RF structure (Fig. 3.3 & 3.4), what then accounts for the On/Off asymmetry in spatial integration of natural images? A key property of nonlinear subunit RF models - including the one we present above - is rectification of non-preferred contrasts. Previous work has identified subunit rectification as an important regulator of integration of spatially structured visual inputs in mouse RGCs [93, 192].

To test for differences in rectification between On and Off parasol RGCs, we measured proxies for the subunit nonlinearities in each of these cells. Because nonlinear subunit RF structure is present at the level of synaptic currents (Fig. 3.5), the relevant nonlinearity will be reflected in these inputs. We used voltage clamp recordings to measure the contrast-response functions of synaptic currents in On and Off parasol RGCs for uniform discs. In the simplified case where a population of homogeneous subunits conveys the synaptic input, the shape of the contrast response function is identical to that of the subunit output nonlinearity. More generally, the contrast response function constrains the shape of subunit nonlinearities, and the degree of rectification correlates with the degree of subunit output rectification.

Excitatory currents in On parasol RGCs were approximately linear for low contrasts and became rectified for stronger negative contrasts (Fig. 3.6A). Inhibitory input to the same cells was very sharply rectified at non-preferred (positive) contrasts. On parasol RGCs received a small amount of feedforward (i.e. On-derived) inhibitory input (Fig. 3.6B), indicated by the increase in inhibitory input at high positive contrasts.

Excitatory input to Off parasol RGCs was sharply rectified even for low contrasts, and correspondingly the cells received very little tonic excitatory input (Fig. 3.6C). Off parasol RGCs receive a sizable feedforward (Off-derived) inhibitory input at high negative contrasts (Fig. 3.6D). The asymmetry in rectification of the excitatory inputs to On and Off parasol RGCs is consistent with previous reports of stronger rectification of Off parasol RGC spike responses [39]. Rectification of responses to white noise stimuli showed a similar asymmetry (Fig. 3.12; see also [224]).

We modified our LN and nonlinear subunit models to use a smooth function fit to the average excitatory contrast-response functions for either On or Off parasol RGCs (Fig. 3.6E). We presented these models with randomly-selected patches from a natural image (Fig. 3.6F). The differences in model responses were stronger with the Off parasol nonlinearity than with the On parasol nonlinearity (Fig. 3.6F). For each of the 101 natural images in the DOVES database, we computed the mean NLI for subunit and LN model outputs across all patches drawn from that image (Fig. 3.6G). For both the On and Off parasol RGC nonlinearity, the mean NLI was positive for every image, indicating that the subunit model tended to produce stronger responses than the LN model. For every image, the Off nonlinearity produced a more positive NLI value than did the On nonlinearity. We saw similar results when we assigned the more rectified Off nonlinearity to the On cell model and vice-versa (by reflecting the nonlinearities across the vertical axis, not shown), indicating that the shape of the nonlinearity produced these differences rather than some feature of negative contrast regions of natural images.

The stronger nonlinear behavior seen with the Off nonlinearity is consistent with the stronger nonlinearities seen in Off parasol RGC excitatory responses to DOVES stimuli and

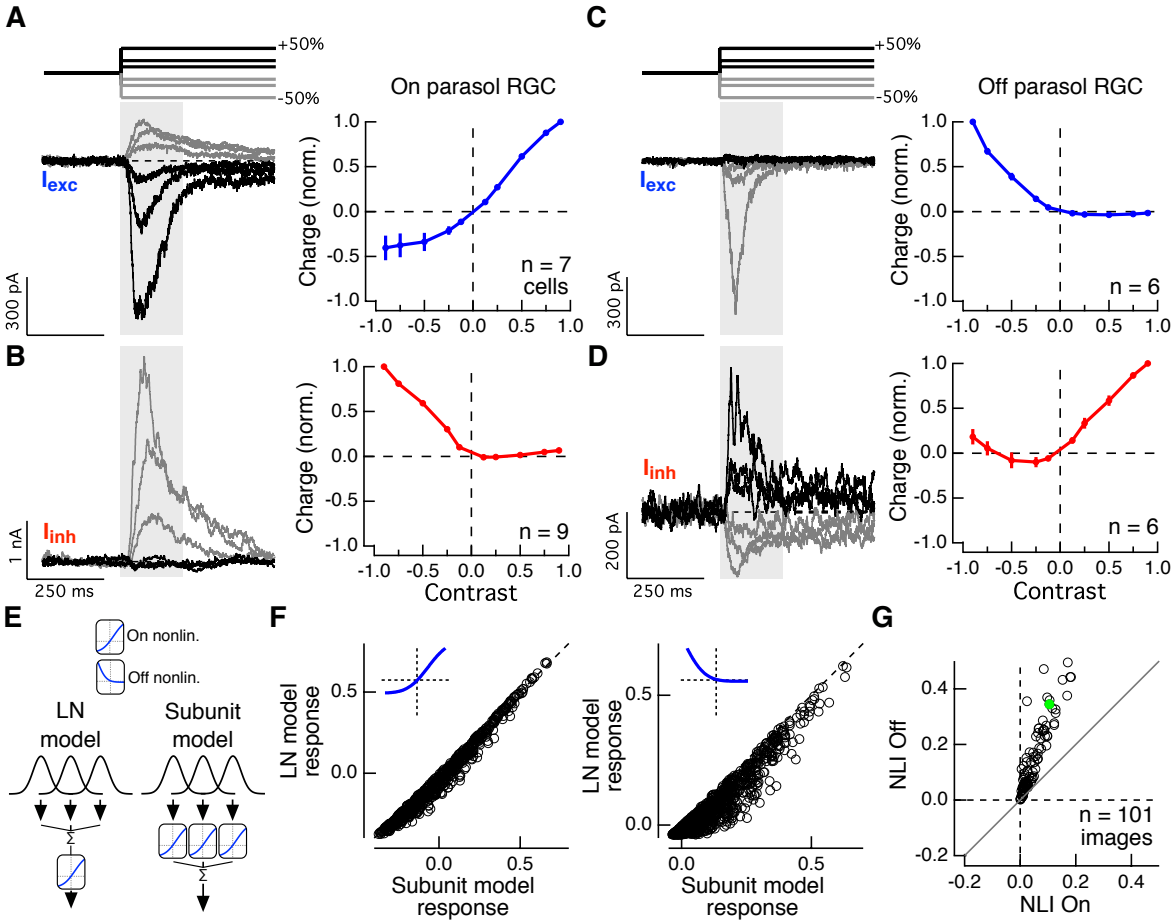

Figure 3.6: **Sharp Off pathway rectification drives strongly nonlinear responses.** (A) Left: Mean excitatory current responses in a representative On parasol RGC presented with spots of varying contrasts. Gray traces denote negative contrast steps and black traces indicate positive contrasts (responses to ± 12.5%, 25%, 50% contrast are shown). Shaded region indicates time over which charge was integrated to yield the contrast response functions. Right: Mean ± S.E.M. contrast response function for On parasol RGC excitatory currents, normalized within each cell (7 cells). (B) Contrast step responses (left) and the mean contrast response function (right) for On parasol RGC inhibitory input (9 cells). (C-D) Same as (A-B) but for Off parasol cell excitatory (C, 6 cells) and inhibitory (D, 6 cells) synaptic currents. (E) We modified our LN and nonlinear subunit models to use the excitatory contrast response function from either On or Off parasol cells in place of the simple piecewise-linear nonlinearity. (F) We presented each model with 1,000 randomly selected patches from a natural image and compared the two model outputs. Left: responses when the models include the On parasol RGC excitatory nonlinearity (inset); right: responses when the models instead use the Off parasol RGC excitatory nonlinearity. (G) We repeated the analysis in (F) for all 101 natural images in the DOVES database and measured the median nonlinearity index across all sampled patches. Each point corresponds to one image. Green point indicates the example image used in (F).

image patches. The decrease in Off parasol rectification and NLI with decreasing light level (Fig. 3.4) further supports the importance of rectification in nonlinear spatial integration. These models of the excitatory inputs to On and Off parasol RGCs can also account for the cells responses to grating stimuli (Fig. 3.13). The NLIs predicted by the On parasol nonlinearity are stronger than those observed in the spike output of these cells (see Fig. 3.1). This suggests that other mechanisms shape spatial (non)linearity in On parasol RGCs, for example the large, nonlinear inhibitory inputs to these cells (Fig. 3.2).

These models show that the degree of subunit rectification determines the degree to which a nonlinear RF can be approximated as linear in the context of natural visual stimulation. A similar effect of subunit rectification can be seen using a simplified piecewise-linear nonlinearity (Fig. 3.14), which shows that this is a general consequence of rectified subunit nonlinearities and does not depend on the precise shape of the nonlinearity.

3.3.5 A nonlinear subunit model improves predictions of responses to natural images

To determine the extent to which incorporating a nonlinear RF structure can improve predictive models of Off parasol RGC responses, we used our excitatory subunit and LN models from above (Fig. 3.6E-G) to build simple predictive models of Off parasol RGC excitatory synaptic current and spike output responses (Fig. 3.7A). We start by predicting excitatory synaptic inputs in response to natural image patches, and then expand the model to predict spike responses.

For each cell, we used the measured RF center size (based on area-summation curves, e.g. see Fig. 3.5A,B), as the RF size in the model, and set the subunit size to be 1/6 of this measured RF center size. Models incorporated the measured nonlinear contrast-response function for excitatory synaptic inputs to Off parasol RGCs (from Fig. 3.6C). As with the modeling above, performance was similar for a range of subunit and RF center sizes, as long as the RF center was several times larger than a subunit (Fig. 3.10).

To test how well these models predict excitatory current responses to natural stimuli, we voltage clamped Off parasol RGCs and measured the excitatory charge transfer in response

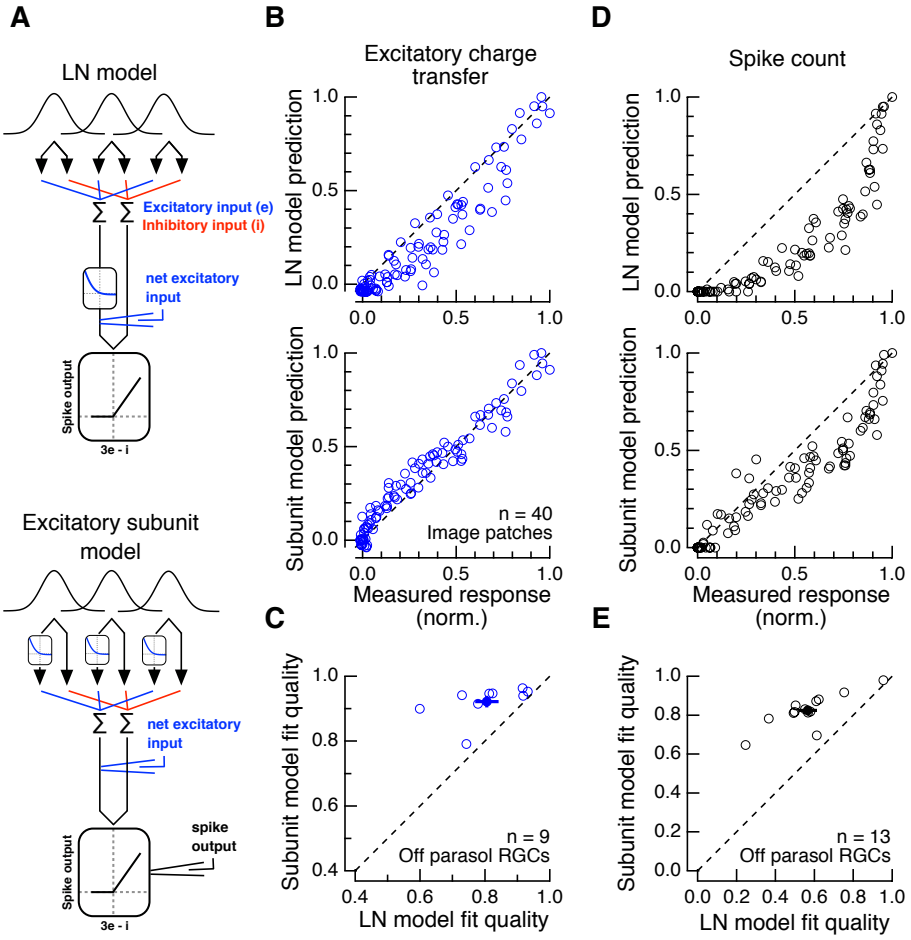

Figure 3.7: **Inclusion of nonlinear excitatory subunits improves a predictive model of natural image responses.** (A) We modified our nonlinear subunit and LN RF models (Fig. 3.3B) to include the measured contrast response function for Off parasol RGC excitatory inputs as well as a linear inhibitory channel. The nonlinear subunit model applies the measured nonlinearity before spatial integration (bottom), while the LN model applies this nonlinear transform to excitatory inputs after spatial integration (top). In both models inhibitory input is purely cross-over (on-tuned) and linear. Excitatory and inhibitory inputs are linearly combined and passed through a threshold-linear input-output function to yield spike output. (B) We measured an Off parasol RGC's mean excitatory charge transfer in response to presentation of 40 randomly selected patches from a natural image. We then passed those image patches through each model and measured the net excitatory input for each stimulus. The subunit model (bottom) provided a better prediction of the excitatory input than the LN model (top), which generally underestimated the excitatory input. (C) Population data for the experiments in (B), showing that the fraction of explained variance is higher for the nonlinear subunit model than the LN model (9 cells). (D) We compared the spike output of each of our models to measured Off parasol RGC spike responses. As with the excitatory inputs to these cells, the nonlinear subunit model more accurately predicted Off parasol RGC spike outputs. (E) Population data showing that the nonlinear subunit model outperforms the LN model (13 cells).

to presentation of 40 randomly-selected patches from a natural image. We normalized the excitatory responses to control for across-cell differences in response strength. We then compared the measured excitatory responses to those predicted by each of the two models (Fig. 3.7A). The LN model underestimated the excitatory current response for many image patches, and the nonlinear subunit model gave a better prediction (Fig. 3.7B,C, $p = 3.9 \times 10^{-3}$). On average, the subunit model explained 92% of the measured variance, while the LN model accounted for 81% (Fig. 3.7C).

To test whether a nonlinear excitatory RF structure could also improve predictive models of spike output, we expanded the excitatory models above to include simplified forms of inhibitory input and spike generation. In both models the inhibitory input is purely linear and exclusively cross-over (On-tuned). The excitatory and inhibitory inputs are combined linearly (with a scaling of excitatory inputs to account for the approximately 3-fold difference in driving force acting on excitatory conductances at spike threshold) and passed through a rectifying nonlinear function to yield the spike output of the cell. Even with these simplified forms of inhibition and spike generation, the subunit model improves the spike count prediction (Fig. 3.7D,E; $p = 2.4 \times 10^{-4}$). On average, the subunit model accounted for 82% of the response variance while the LN model captured only 57%, showing that most of the errors made by an LN model in predicting spike responses to natural images can be corrected by incorporating nonlinear excitatory subunits into the RF. This improvement in prediction with the excitatory subunit model is robust to changes in the form of inhibitory input. For example, using a rectified (rather than linear) cross-over inhibition or even removing inhibition altogether did not qualitatively change the results shown here. This suggests that the improvement in prediction of Off parasol RGC spike responses stems from the improved predictions of the cell's excitatory inputs. There are many aspects of retinal circuitry and intrinsic RGC computation that these models exclude. Notable examples include local adaptational mechanisms [63, 114], heterogeneous/non-Gaussian linear RF structure [77, 192], and realistic inhibitory inputs and spike generation mechanisms. These and other features will likely be important to include in models that predict parasol cell

responses, especially when presented with time-varying stimuli (like the DOVES stimuli in Fig. 3.1 & 3.2). Indeed, the observation that an On parasol RGC's synaptic inputs can be spatially nonlinear, while its spike output is linear (when probed with natural images), implies that additional mechanisms, beyond rectification of excitatory subunits, shape spatial integration at the level of spike outputs in On parasol RGCs. Nevertheless, the substantial improvement in predictions of Off parasol RGC responses to natural images suggests that excitatory nonlinear subunits will be an essential part - indeed perhaps the most important of the mechanisms listed above - of full predictive models for some RGC types.

3.4 Discussion

We investigated the extent to which nonlinear spatial integration affects RGC responses to natural visual stimuli. We found that responses of On but not Off parasol RGCs were well approximated using a spatially linear RF, despite both cell types showing nonlinear responses to grating stimuli designed to probe spatial nonlinearity. Sharply rectified excitatory subunits in the RF center mediate the nonlinear responses of Off parasol RGCs. Finally, we found that a simple RF model that includes rectified subunits accurately predicts responses to natural images, capturing approximately 80-90% of the response variance; this represents a substantial improvement over models with linear spatial integration.

3.4.1 Synaptic rectification controls nonlinear integration

The sharply rectified excitatory subunits in the Off pathway mediate the nonlinear spatial integration of Off parasol RGCs; excitatory inputs to On parasol RGCs exhibited substantially milder rectification. Excitatory inputs to brisk transient RGCs in guinea pig show a similar On/Off asymmetry of rectification [61]. In addition to differences among cell types, synaptic rectification can differ within the same cell type as lighting conditions change [93]. The degree of synaptic rectification may be a common control point for retinal circuits that can change with cell type and stimulus conditions. Indeed, Off parasol RGCs behaved more linearly at lower light levels (Fig. 3.4).

On parasol RGC responses to natural images were more linear than expected given the degree of excitatory subunit rectification and the nonlinear responses to contrast-reversing grating stimuli. Thus other mechanisms appear to linearize On parasol RGC responses. Two lines of evidence support this conclusion.

First, the excitatory synaptic inputs to On parasol RGCs in response to natural inputs showed only modest deviations from linear spatial integration (Fig. 3.2). This is surprising considering the clear nonlinear responses elicited by low-contrast grating stimuli (Fig. 3.11). The grating stimuli used here, and in past work, were designed to elicit nonlinear (F2) responses while minimizing linear (F1) response components. This was achieved by carefully centering the stimulus over the RF center and/or using high spatial frequency gratings. When presented with natural images, however, very rarely is the linear response component so effectively nulled. Most natural image responses contain both linear and nonlinear response components, and it is not known whether these response components interact in a straightforward fashion or if, for instance, the linear response component can suppress nonlinear responses. Consistent with this hypothesis, the On parasol RGC excitatory subunit model (Fig. 3.6) could reproduce observed grating responses (Fig. 3.13), but over-estimated the NLI in response to natural images.

Second, the small deviations from linearity in the On excitatory inputs (Fig. 3.2) are absent at the level of spike output. Nonlinear cross-over inhibitory input to On parasol RGCs could further "linearize" responses to images [35, 240]. The images that drive strong nonlinear responses in the excitatory input would also be expected to drive nonlinear responses in inhibitory input, resulting in weaker nonlinear responses in spike output compared to excitatory synaptic inputs.

3.4.2 A nonlinear receptive field enhances responses to spatial structure in natural images

Nonlinear responses to natural images depended strongly on the particular region of the image presented. This was true for the DOVES movies (Fig. 3.1), the stationary natural images (Fig. 3.3), and the nonlinear RF model responses (Fig. 3.3, 3.4, 3.6, 3.7). What

is it about these images that drive strong deviations from linear integration? The mechanism of nonlinear integration - rectification of non-preferred contrasts and thus incomplete cancelation of light and dark regions of a scene - suggests an answer. In particular, homogeneous images that drive most of the subunits in the same direction (e.g. a population of On subunits all subjected to strong positive contrast) will produce the same output in a LN or a comparable nonlinear subunit model. However, images containing a wide distribution of contrasts strongly activate some subunits and deactivate others, causing a linear and a nonlinear RF to produce very different responses.

To explore this intuition, we compared the difference in model outputs for randomly sampled natural image patches (as in Fig. 3.3) with the normalized variance of each image patch (variance in pixel intensities divided by the average intensity of the patch). These two values were positively correlated (0.43 ± 0.18 mean $\pm$ standard deviation, linear correlation coefficient) for 10,000 patches from each of 101 natural images. Thus heterogeneous natural images more strongly drive nonlinear spatial integration than homogeneous images.

Where in natural scenes are the relatively "nonlinear" patches located? To illustrate this, we passed each RF model over every location within several natural images and computed the model output difference (nonlinear subunit minus LN model). Figure 3.8 shows example heat maps of model differences and the original images from which they were generated. For some images, it is clear that spatial inhomogeneity drives nonlinear responses (Fig. 3.8A,B), whereas regions of the scene that are relatively homogeneous at the scale of the RF center (e.g. the tree trunk or sky in Fig. 3.8C) activate both models approximately equally. Some images show strong nonlinear responses clustered along boundaries or edges within the scene (Fig. 3.8A,C).

Thus, a nonlinear RF composed of rectified subunits enhances responses to regions of scene with a large degree of spatial inhomogeneity, including edge structure. Whether and how downstream circuits use this information is unknown. Psychophysical evidence suggests that during free-viewing, new fixation locations are not randomly dispersed through the visual field, but that saccades are typically directed towards regions of the scene with edges

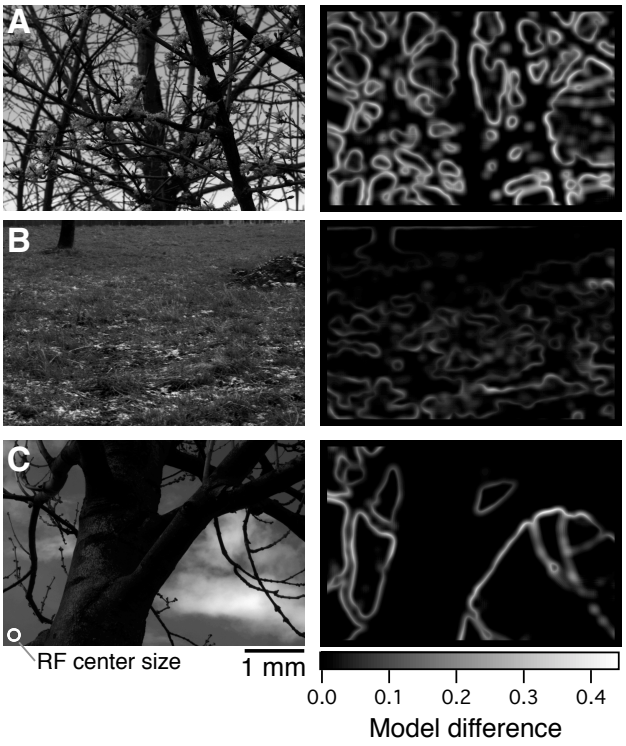

Figure 3.8: **Rectified subunit RF structure enhances responses to spatial structure in natural images.** (A-C) Left: example natural images were passed through each of the models in Fig. 3.4, with each point in the image sampled. Right: for each image, we constructed a heat map of model differences (subunit minus LN model) to illustrate the locations within the scene where the two models differ. For scale, the RF center size is indicated in panel (C). Color scale in (C, right) applies to all heat maps.

[14, 175]. The highly contrast sensitive and nonlinearly integrating peripheral Off parasol RGCs would be well-suited to detect these regions of a natural scene.

Another consequence of the enhanced sensitivity to spatial contrast in Off parasol RGCs is that some natural visual stimuli will activate both On and Off parasol RGCs simultaneously. In particular, this will occur when a saccade is made to an image patch that contains high spatial contrast (e.g. an edge) but is, on average, brighter than the previous fixation. For example, see the spike responses to fixation number 5 in Fig. 3.1C,E. This suggests that during natural vision, the standard division of labor in which On cells signal increases in luminance and Off cells signal decreases in luminance is too simplistic. In reality, RGCs encode multiple features of the scene (here, for example, mean luminance and spatial contrast), and RGC responses are shaped by a combination of sensitivities to these features.

3.4.3 Implications for constructing generalizable models

A linear RF is a common feature of predictive models in the visual system, but its ability to capture responses to natural visual stimuli has been largely untested (see [99] for an exception). Our observations suggest that to predict responses to natural visual inputs, models of some RGC types will need to account for nonlinear RF structure, while for other RGC types a linear approximation suffices. In the case of Off parasol RGCs, we found that the precise subunit organization was not as important as simply having several rectified subunits that are smaller than the RF center. The same model may apply quite directly to other classes of RGC that contain rectified subunits.

Which RGCs can be approximated using a linear RF is not clear from responses to typical artificial stimuli such as gratings. The use of natural stimuli to test models of neural computation engages a variety of mechanisms in ways that are not easily predicted from responses to artificial stimuli. Artificial stimuli probe only a very small region of stimulus space, which makes them ideal for dissecting a particular mechanism of interest (e.g. grating stimuli to probe nonlinear spatial integration), but poorly suited when attempting to place circuit mechanisms in their broader functional context. A similar approach may

enhance understanding of the coding impact of several other mechanisms, including interactions between the center and surround, photoreceptor adaptation in the context of spatially structured stimuli, and integration of spatially nonlinear synaptic inputs.

While the results presented here suggest that spatially nonlinear models will be required to capture natural scene responses for some classes of visual neurons, further work is needed to develop methods of constructing such models and fitting them to data. Already some important progress has been made on this front (e.g. see [82, 237, 236]). We hope that the work presented here will guide the development of predictive models by highlighting specific mechanisms that substantially shape neural responses to natural stimuli.

3.5 Experimental Procedures

3.5.1 Tissue preparation

Retinal tissue was obtained from terminally anesthetized Macaque monkeys (*M. nemestrina*, *M. mulatta*, or *M. fascicularis*) of either sex via the tissue distribution program at the Washington National Primate Research Center. All procedures were approved by the Institutional Animal Care and Use Committee at the University of Washington. After enucleation, the eye was hemisected and the vitreous humor was removed mechanically. In some cases, the eye cup was treated for ∼15 minutes with human plasmin (∼50 μg/mL, Sigma or Haematologic Technologies Inc.) to aid in removal of vitreous. We observed no differences in retinal health or sensitivity after plasmin treatment. The retina was dark adapted for ∼1 hr, and all subsequent procedures were performed under infrared light using night-vision goggles. The retina and pigment epithelium were separated from the sclera and stored in oxygenated (95%O2/5%CO2) Ames bicarbonate solution (Sigma) in a light-tight container. Retinal mounts (∼2-3 mm on a side) were removed from the pigment epithelium and laid flat, photoreceptor-side down, onto a poly-D-lysine coated coverslip (BD biosciences) before being placed in a recording dish that was continuously perfused at 7-9 mL/min with Ames solution at 31-35°C. During recording, the tissue was visualized using infrared illumination

and DIC optics.

3.5.2 Patch recordings

Electrophysiological recordings were performed using a Multiclamp 700B amplifier (Molecular Devices). Spike responses were measured using extracellular or loose-patch recordings with an Ames-filled pipette (tip resistance ∼2-6 MΩ). For voltage clamp recordings, we used low-resistance pipettes (tip resistance ∼1.5-4 MΩ) filled with a Cs-based internal solution (containing, in mM: 105 CsCH3SO3, 10 TEA-Cl, 20 HEPES, 10 EGTA, 5 Mg-ATP, 0.5 Tris-GTP, and 2 QX-314, pH 7.3, ∼280 mOsm). We compensated for access resistance (∼4-12 MΩ) online by 50-75%. Reported voltages have been corrected for an approximately -10 mV liquid junction potential.

We isolated excitatory (inhibitory) synaptic inputs by holding at the reversal potential for inhibitory (excitatory) synaptic currents. Reversal potentials were estimated for each cell by delivering light steps near the expected reversal potentials (approximately 0 mV for excitatory inputs and -60 mV for inhibitory currents) and adjusting the holding potential to minimize the appropriate current.

3.5.3 Cell identification and selection

On and Off parasol RGCs were first identified under DIC optics by the size and morphology of their soma, which was generally a very reliable indicator of cell type. We confirmed cell type by delivering light stimuli over the RF center to observe the transient spike responses and high contrast sensitivity characteristic of parasol RGCs. The overall health and sensitivity of the retina was confirmed by delivering a uniform, 5% contrast, 4 Hz modulated stimulus, which produces a robust spike response in On parasol RGCs in sufficiently sensitive tissue. Sensitivity was continuously monitored (typically before each recording) in this way. Mean spike rates in healthy, sensitive tissue at a mean light level of ∼4,000 R*/cone/s were 12.8 ± 1.5 Hz for On parasol RGCs and 2.2 ± 0.6 Hz (mean ± S.E.M.) for Off parasol RGCs.

3.5.4 Visual stimulation

Visual stimuli were presented on an OLED microdisplay monitor (eMagin, Bellevue, WA) focused, via a microscope condenser, onto the photoreceptors. The 800×600 monitor display was presented with a resolution of 1.2 μm/pixel at the retina. Stimuli were presented and data acquired using custom written stimulation and acquisition software packages Stage (stage-vss.github.io) and Symphony (symphony-das.github.io). Monitor outputs were linearized by gamma correction. Stimuli were calibrated using monitor power outputs, the spectral content of the monitor, macaque photoreceptor spectral sensitivity [24], and a collecting area of 0.37 μm^2 for cones [190] and 1 μm^2 for rods. Unless otherwise noted (i.e. Fig. 3.4), mean light levels produced ∼4,000 isomerizations (R*)/M or L-cone/s, ∼1,000 R*/S-cone/s and ∼8,000 R*/rod/s.

Before every parasol RGC recording, we found the center of the cell's RF using a split-field contrast reversing grating stimulus at 4 Hz and 90% contrast. To do this, we shifted the grating until the two F2 response cycles were balanced (i.e. we minimized the F1 while maximizing the F2 component of the response). We performed this search in both the horizontal and vertical dimensions. This method of mapping typically allowed us to find the RF center to within 10 μm. With the exception of the contrast reversing grating stimuli (Fig. 3.5) and the contrast response flashes (Fig. 3.6), which covered 600 μm on the retina, all stimuli were restricted to the RF center. For each cell, the RF center size was based on the area summation curve, given by the spike count during presentation of various sized spots, centered over the RF center (e.g. see Fig. 3.5A,B). The extent of the RF center was estimated as the point in the area summation curve where the response begins to roll off due to the strength of the surround overtaking that of the center.

Naturalistic retinal input movies (Fig. 3.1 & 3.2) were generated using the Database Of Visual Eye movementS (DOVES, [227], live.ece.utexas.edu/research/doves). In the DOVES database, eye movements for each subject were measured for a variety of natural images. We selected a subset of the trajectories from the database to use as visual stimuli. Because of the

frame rates used in this study (60 Hz), we were unable to capture the entire eye movement trajectory without jumping large retinal distances between frames during saccades. Because of this, we selected stable fixations of at least 200 ms from an eye movement trajectory and concatenated them such that, instead of the full saccade trajectory between fixations, the image simply jumped from one fixation to the next between frames. Fixational eye movements, while smaller and slower than saccades, can move the image on the retina tens of microns during a fixation and are included in these stimuli. The stimuli we constructed had an average length of 1.9 seconds (range 0.6 - 3.8 seconds) and contained 5 fixations on average (range 2-9 fixations). The average fixation was 0.38 seconds in duration (range 0.23 - 1.25 seconds).

For both DOVES movies (Fig. 3.1 & 3.2) and stationary natural image patches (Fig. 3.3, 3.4 & 3.7), we scaled the stimuli such that the brightest point in the entire image (from which the stimuli were derived) was set as the highest monitor intensity (i.e. 255 on our 8-bit display). The eye movement data in the DOVES database were collected at a scale of about 1 arcmin/pixel, which is 4.8 μm on the human retina. We presented DOVES movies and natural image patches at the same scale on the retina (4.8 μm / pixel).

To construct the linear equivalent disc stimulus (Fig. 3.1-3.4), we took a weighted average pixel intensity of each frame of a given DOVES movie with a circular Gaussian function. The two-standard deviation width of this Gaussian function was given by the diameter of the aperture placed over the image (which was equal to the measured RF center size). Each frame of the original movie thus yielded a single intensity value, which we assigned to the disc on that frame of the linear equivalent disc stimulus. We confirmed the accuracy of the linear prediction, and of our monitor's gamma correction, by sampling with a uniform (instead of Gaussian) function and comparing the power output of a DOVES movie and its corresponding linear equivalent disc using an optometer.

3.5.5 Analysis and modeling

Data analysis was performed using custom written scripts in MATLAB (Mathworks). Spike detection was performed using high-pass filtering and a k-means clustering algorithm on positive and negative peaks in the spike waveform. The near perfect accuracy of spike detection was confirmed by visual inspection. Spike PSTH traces were generated using a Gaussian kernel ($\sigma = 5$ msec) smoothing operation on binary spike trains. Current recordings were baseline-subtracted based on time points before onset of the stimulus. Charge transfer was calculated by numerical integration of baseline-subtracted, averaged current traces. For cycle-averaged responses (Fig. 3.5), the first cycle of the response was excluded. F1 and F2 amplitudes were calculated using the built in MATLAB FFT function.

For RF modeling, subunits were located on a square grid. The RF center weighting function was a circular Gaussian, 2-st. dev. width of 230 μm; the subunits were circular Gaussians 1/6 the size of the RF center. The spacing between subunit centers on the grid was defined as two standard deviations of the subunit filter. These sizes were chosen to be consistent with our observations in parasol RGCs (Fig. 3.5). Variation of RF center and subunit sizes within the range of sizes observed in Fig. 3.5 did not substantially affect the results presented here (see Fig. 3.10). For the predictive models in Fig. 3.7, we used each cell's measured RF center size and subunits that were 1/6 the size of the RF center.

Natural image patches were selected from a natural image and each pixel was converted to a Weber contrast value before being run through the RF models. Contrast, C was defined in the Weber sense, i.e. $C = (I - B)/B$, where I is the intensity of each pixel, and B is the mean intensity value for the entire image (not just that patch). The background intensity when presenting natural image patches to cells (Fig. 3.3, 3.4 & 3.7) was similarly defined as the mean intensity across the entire image from which the patches were drawn.

To pass an image through the models, we computed the activation of each subunit (by sampling the image with each spatially-offset subunit or, equivalently, convolving the image with the subunit filter and sampling the filtered image at each of the subunit locations),

passed each subunit activation through a (non)linear transfer function, took a weighted sum of the subunit outputs, and in the case of the LN model, applied a final output nonlinearity. See equations 2 and 3, below, which give the responses to an image patch of the nonlinear subunit model, r_{sub}, and the linear-nonlinear model, r_{LN}

$$r_{sub} = \sum_{i=1}^{n} w_i \times f(a_i) \quad (2)$$

$$r_{LN} = f(\sum_{i=1}^{n} w_i \times a_i) \quad (3)$$

where a_i is the activation of subunit i, w_i is the weighting (by the RF center) applied to subunit i, and f is a nonlinear input-output function, the variants of which are described below. For the models in Fig. 3.3, 3.4 & 3.8, we used a threshold-linear rectifying nonlinearity. For the nonlinearities in Fig. 3.6 & 3.7 we fit a modified cumulative Gaussian function (equation 4) to the across-cell average of the normalized excitatory contrast-response functions. For these smooth curves, the (normalized) response as a function of contrast was,

$$R(C) = \epsilon + \frac{\alpha}{\sqrt{2\pi}} \int_{-\infty}^{C} e^{\frac{-(\beta t+\gamma)^2}{2}} dt \quad (4)$$

where C is contrast, on -1 to 1. The curve has four free parameters: ϵ is an offset in the vertical direction, γ is an offset along the horizontal (contrast) axis, β determines the sensitivity or slope of the contrast response function, and α is a scale factor which determines the maximal response. For the excitatory contrast response functions used in the modeling in Fig. 2.6 and 2.13 (appendix), the values of these fitted parameters were, for the On parasol RGC model: $[\alpha = 1.64, \beta = 1.99, \gamma = -0.705, \epsilon = -0.407]$; and for the Off parasol model: $[\alpha = 2.70, \beta = -1.88, \gamma = -2.00, \epsilon = -0.038]$.

The contrast response functions shown in Fig. 3.6 were measured by flashing 600 μm diameter spots for 500 ms. We measured the charge transfer (integrated current) during the first 200 ms of the response to yield the contrast response functions in Fig. 3.6 and 3.7.

To compute the RF center size (as in Fig. 3.5), we fit a difference of Gaussians model to the area summation curve generated by presenting spots of various sizes over the RF center. The difference of Gaussians function is shown in equation 5 below. The response R, for a spot size of diameter, D is given by

$$R(D) = R_0 + \int_{-D/2}^{D/2} k_c e^{-(\frac{t}{2\sigma_c})^2} - k_s e^{-(\frac{t}{2\sigma_s})^2} dt \quad (5)$$

There are five free parameters for this fit: R_0 is a vertical offset, σ_c and σ_s determine the sizes of the center and surround, respectively, and k_c and k_s determine the weights of the center and surround, respectively.

Throughout the paper, reported p-values are from either a Wilcoxon signed-rank test (for paired samples) or a Wilcoxon rank-sum test (Mann-Whitney U test, for unpaired samples).

3.6 Acknowledgments

We thank Shellee Cunnington, Paul Newman, and Mark Cafaro for excellent technical support. Mark Cafaro also designed the acquisition and stimulation software packages (Symphony & Stage) which made this work possible. Tissue was provided by the Tissue Distribution Program at the Washington National Primate Research Center (WaNPRC), and we are grateful for assistance from the WaNPRC staff, especially Chris English. Raunak Sinha, Mike Manookin, and Juan Angueyra assisted in tissue preparation. We thank Wyeth Bair, Nora Brackbill, David Brainard, E.J. Chichilnisky, Phil Mardoum, Dean Pospisil, and Ali Weber for helpful discussion and feedback on a previous version of this manuscript. This work was supported in part by NIH grants T32-EY07031 and F31-EY026288 (to MHT) and EY11850 (FR), and the Howard Hughes Medical Institute (FR).

3.7 Supporting Information

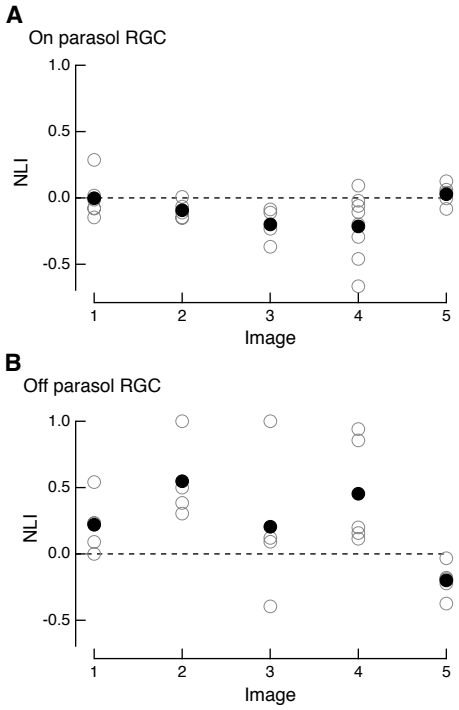

Figure 3.9: **Variability in nonlinear integration of natural stimuli due to variability in structure of natural stimuli.** (A) We presented five DOVES movies (using five different natural images) to an On parasol RGC and measured the nonlinearity indices using the cell's spike responses (as in Fig. 3.1). Open circles denote the NLI for individual fixations, while the filled circle for each stimulus image indicates the movie-average mean NLI (analogous to the cell average NLI in Figs. 3.1 & 3.2, where each cell was only presented with one DOVES movie). (B) Same as (A) but for an Off parasol RGC. The variability of NLIs (across fixations within a single DOVES movie and between different DOVES movies) indicates that a large part of the observed variability in nonlinear responses is due to the variable statistics of the natural image stimuli.

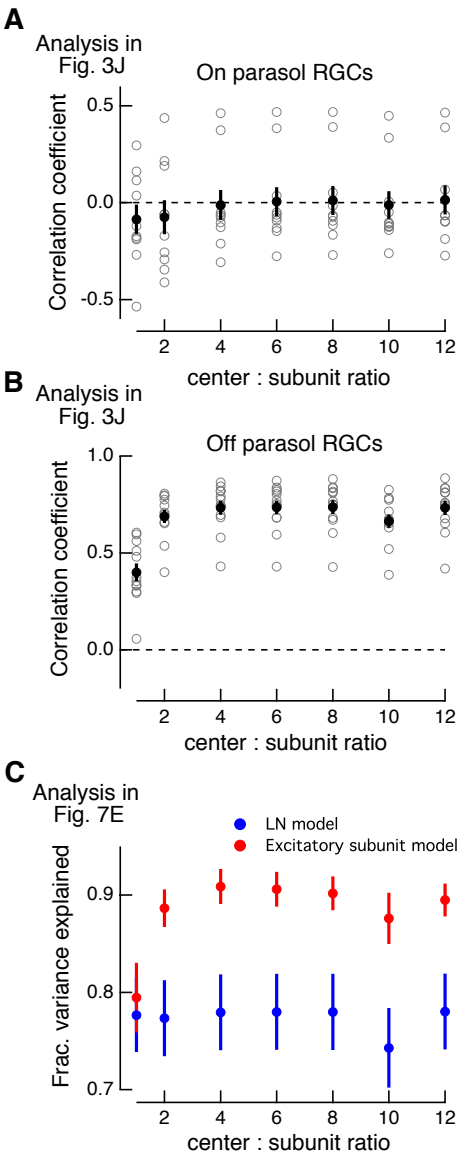

Figure 3.10: **Sensitivity of subunit RF models to subunit size.** (A) We repeated the analysis in Fig. 3.3J using a variety of subunit sizes in the model. We kept the RF center size constant and varied the ratio of subunit to RF center diameter. As in Fig. 3.3J, open circles denote individual cells and filled circles denote mean ± S.E.M. across the measured population (n = 10 On parasol RGCs). (B) Same as (A) for the population of Off parasol RGCs (n = 12). (C) We varied the relative size of the subunits and repeated the predictive modeling analysis of Fig. 3.7E. Shown are the population mean ± S.E.M. fraction of explained variance for the LN (blue) and excitatory subunit (red) predictive models. For all of these analyses, as long as the subunits are ∼1/2-1/4 the size of the RF center (or smaller), the model's predictive power is relatively stable. This indicates that capturing the exact details of the subunit structure does not matter as much as simply having several sub-RF scale nonlinear subunits.

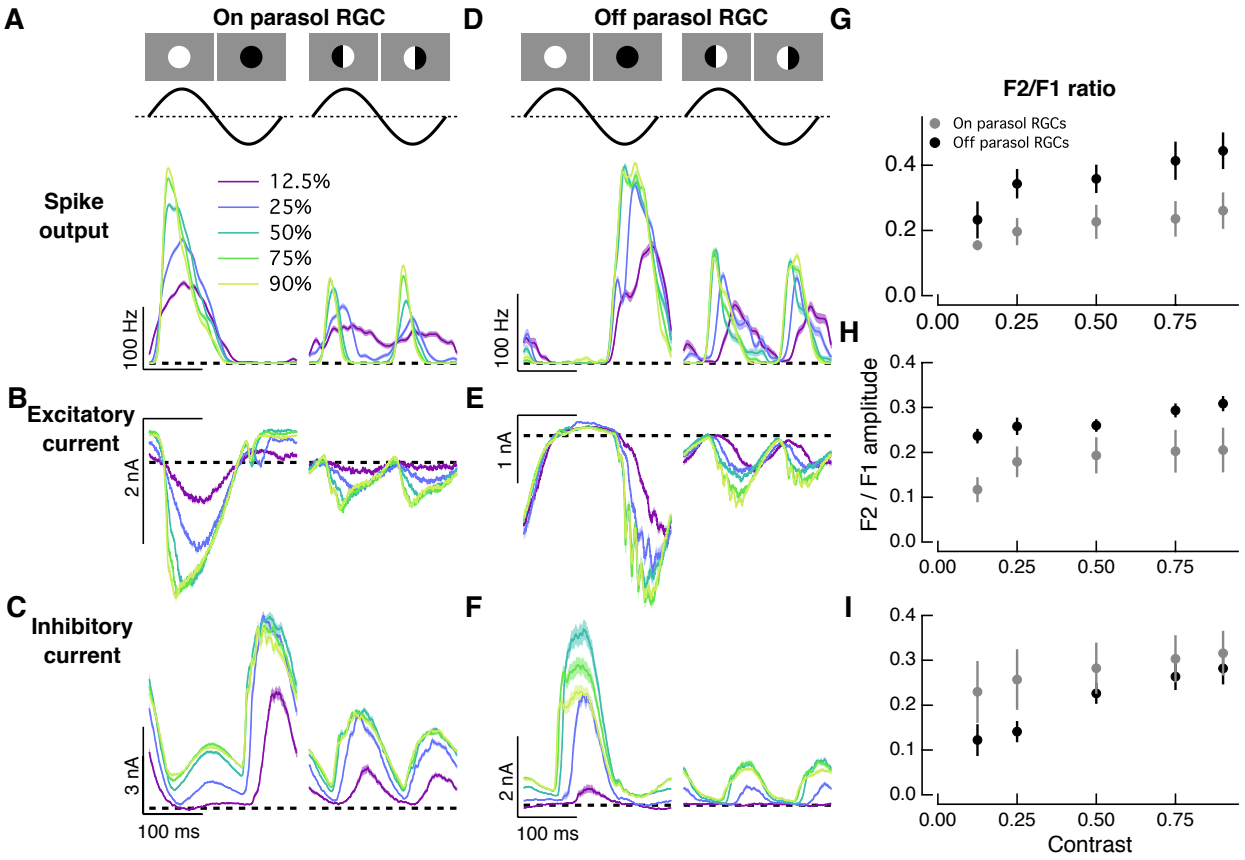

Figure 3.11: **Contrast dependence of nonlinear responses.** (A) Example cycle-average On parasol RGC spike responses to uniform (left) and split-field (right) contrast-reversing stimuli restricted to the RF center. We measured F1 (4 Hz) responses to the uniform stimuli and F2 (8 Hz) responses to the split-field stimuli at a variety of contrasts, indicated by legend. Dashed line indicates zero spikes/sec. (B-C) Same as (A) for excitatory and inhibitory synaptic currents, measured in whole-cell voltage clamp recordings. Dashed line indicates baseline current. (D-F) Same as (A-C) for an example Off parasol RGC. (G) Population summary showing F2 response amplitude, normalized by F1 amplitude, for On (gray) and Off (black) parasol RGC spike responses. (n = 7 On parasol RGCs and 5 Off parasol RGCs). (H-I) Same as (G) for excitatory (H, n = 4 On parasol RGCs and 6 Off parasol RGCs) and inhibitory (I, n = 3 On parasol RGCs and 5 Off parasol RGCs).

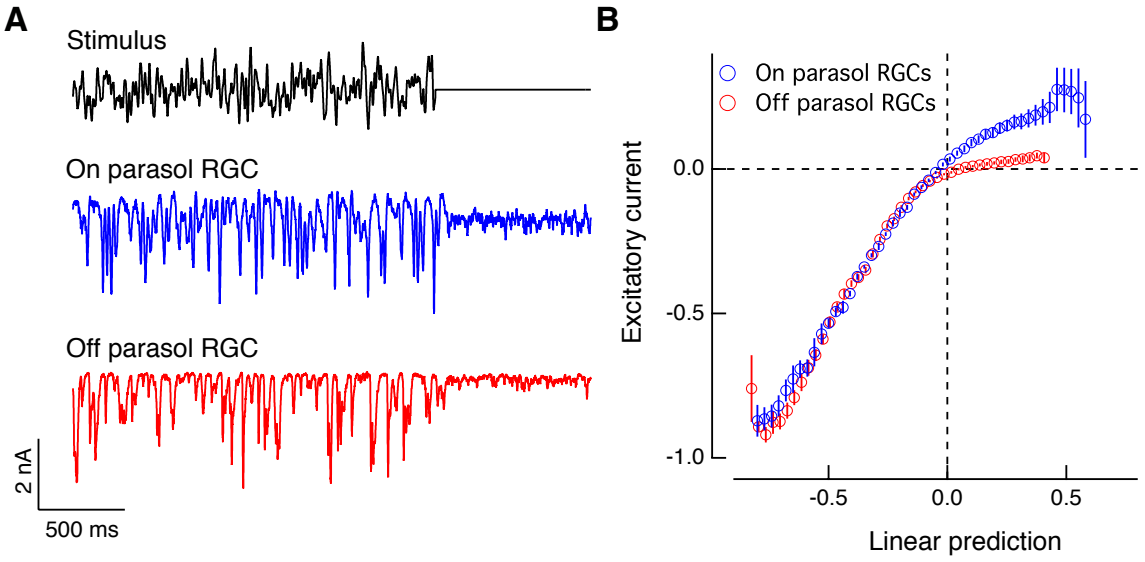

Figure 3.12: **Asymmetric rectification in On and Off parasol RGC excitatory inputs measured using an LN model fit to white noise stimulation.** We presented On and Off parasol RGCs with random, full-field white noise stimulation while measuring excitatory currents in voltage clamp recordings. (A) Example traces showing a portion of the stimulus (top), which was spatially uniform white noise and single trials of the excitatory currents, measured in whole-cell voltage clamp recordings, in On (middle, blue) and Off (bottom, red) parasol RGCs. Note the larger, noisier tonic current in the On parasol RGC compared to the Off parasol RGC. These currents were recorded simultaneously in paired recordings (see [224]). (B) Using responses to the white noise portion of the stimulus, we computed linear filters using reverse correlation. Shown are the linear predicted currents (linear filters convolved with the stimulus) compared to the measured excitatory current. Excitatory current responses have been normalized within each cell, and inward (excitatory) currents are negative. Off parasol RGC excitatory inputs show sharp rectification, while On parasol RGC excitation is relatively more linear near zero, which agrees with estimates of the excitatory nonlinearity using the contrast-response function. Shown for each bin is mean $\pm$ S.E.M. n = 8 On parasol RGCs, n = 8 Off parasol RGCs.

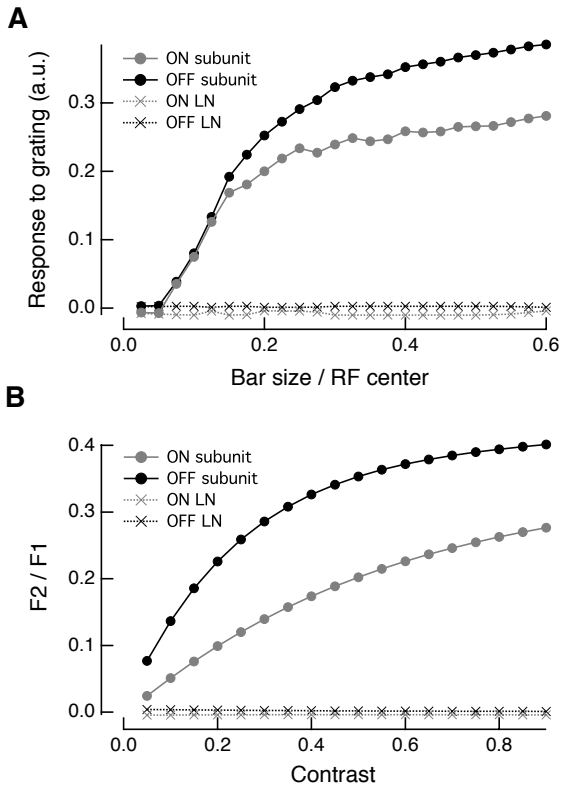

Figure 3.13: **Subunit models can account for excitatory current responses to gratings stimuli.** (A) We presented gratings stimuli to the models in Fig. 3.6 which use measured excitatory subunit nonlinearities. These models have no temporal dynamics, so the modeled F2 response is simply the model's output to a grating centered over the RF center. For both On and Off parasol RGCs, the LN models fail to respond to the gratings, as expected. The nonlinear subunit models show nonlinear responses that rise at a similar spatial scale as that observed in the experimental results (Fig. 3.5). Note that the subunits in this model do not contain surrounds, which is why the nonlinear response does not fall as observed in the data. (B) When we presented a split-field grating stimulus at various contrasts, the On cell model produced weaker nonlinear responses compared to the Off cell model. Compare this to the excitatory current responses shown in Fig. 3.11).

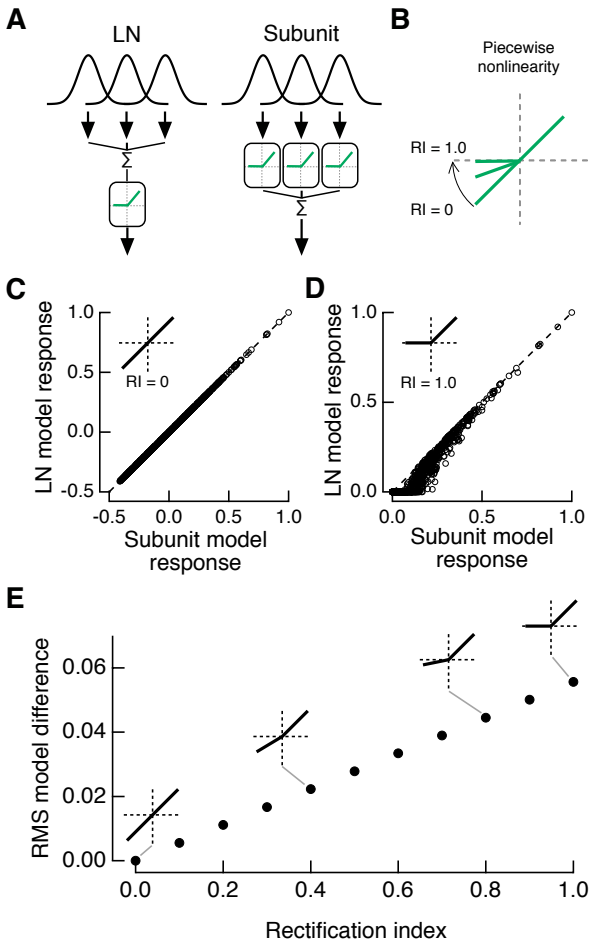

Figure 3.14: **Rectification of subunits controls nonlinear integration of natural images.** (A) We modified our LN and nonlinear subunit models (Fig. 3.3) to include a piecewise-linear input-output function with a variable degree of rectification. (B) The Rectification Index (RI) determines the strength of rectification (see text). (C) We passed 1,000 random patches from a natural image through each of our models with a range of input-output functions, defined by their RIs. For an RI of 0 (inset), the linear model is equivalent to the subunit model, as expected. (D) For a perfectly rectified function ($RI = 1.0$, inset), the subunit model generates a larger response than the LN model for many image patches, as in Fig. 3.3. (E) We measured the root mean squared difference between the two model outputs with a variety of transfer functions. As the RI increases (i.e. more strongly rectified nonlinearity), the difference between the two models grows.

To illustrate the impact of subunit rectification on natural image integration, we modified our LN and nonlinear subunit models to include a piecewise-linear nonlinearity with a variable degree of rectification (Fig. 3.14A,B). This function is linear with slope 1.0 for preferred contrasts, and responses to non-preferred contrasts are controlled by a single parameter - the Rectification Index (RI):

$$RI = 1 - \frac{|r_n|}{r_p} \quad (6)$$

where r_n and r_p are outputs to equal and opposite non-preferred and preferred contrasts. Thus a RI of 0 corresponds to a linear function and a RI of 1.0 corresponds to a completely rectifying function (the same nonlinearity used in the original models of Fig. 3.3). We then presented natural image patches to each of these models and compared their outputs, as in Fig. 3.3C,D, while varying the RI of the nonlinear transfer function in both models. For a RI of 0, the two models are equivalent and are both purely linear models. Accordingly, model outputs were the same in response to every image patch (Fig. 3.14C). For a RI of 1.0, as we have seen (Fig. 3.3C), many image patches more strongly activate the nonlinear subunit model than the LN model (Fig. 3.14D). As the degree of rectification increases, the differences between the model outputs grow (Fig. 3.14E).

Chapter 4

RECEPTIVE FIELD CENTER-SURROUND INTERACTIONS MEDIATE CONTEXT-DEPENDENT SPATIAL CONTRAST ENCODING IN THE RETINA

4.1 Summary

Almost every neuron in the early visual system has some form of an antagonistic receptive field surround. But the function of the surround is poorly understood, especially under naturalistic stimulus conditions. We use both natural and synthetic stimuli to show that the surround of retinal ganglion cells interacts nonlinearly with the receptive field center. This nonlinear center-surround interaction is especially prominent for stimuli that contain intensity correlations across visual space, like natural images. The surround enters the feed-forward retinal pathway before a spatially-localized nonlinearity that shapes signaling in the retinal ganglion cell center, and one consequence of this is that inputs to the surround can modulate the sensitivity of the center to spatial contrast in a scene, including in natural images. This work shows that one unexpected function of the surround in encoding natural scenes may be to modulate spatial contrast sensitivity based on visual context.

4.2 Introduction

The receptive field (RF) surround is a ubiquitous feature of early visual computation. Classically its function in the vertebrate retina has been understood to enhance sensitivity to edges in a visual scene via lateral inhibition [137]. The RF surround has also been suggested to play a role in promoting efficient representation of naturalistic visual stimuli. Natural scenes contain strong correlations across visual space [62, 74, 211], and the RF surround can decorrelate responses across a population of visual neurons [8, 7, 53], thereby increasing

encoding efficiency via a reduction of redundancy in the population code [18, 214].

Recent findings have challenged this simple conception of the function of the RF surround in the context of natural scenes. For example, other mechanisms can play a larger role in decorrelation (or "whitening") of natural images compared to the surround, which seems to provide incomplete decorrelation [168]. For instance, nonlinear processing in the retina can be approximately twice as important for decorrelating retinal ganglion cell (RGC) responses to natural scenes compared to the RF surround ([165], see also [81, 134]).

A better understanding of the impact of eye movements on retinal inputs has also complicated the redundancy reduction interpretation of the RF surround. While the original hypothesis posited that the RF surround is responsible for whitening the spatial frequency content of natural images, recent work has shown that human fixational eye movements themselves whiten the retinal image, before any neural processing takes place [124, 183, 196, 30]. In light of this, the function of the RF surround, especially as it relates to natural scene encoding, is unclear.

Anatomical and functional characterization of the RF surround has shown that the surround may be well poised to impact the spatially nonlinear organization of the RF center. A rectifying nonlinearity applied to the output of cone bipolar cells has been shown to endow a RGC with sensitivity to spatial structure much finer than the RF center [93, 192], and this is true in the context of naturalistic visual inputs as well [226]. It is this nonlinearity which mediates the nonlinear subunit structure of the RGC RF center [60, 126, 192].

The surround is generated in either or both of the synaptic layers in the retina, with some contribution from horizontal cells in the outer retina [46, 54, 135, 145, 231], and some from amacrine cells in the inner retina [44, 71, 79, 103, 217]. The relative contributions of inner and outer retinal mechanisms to the RF surround may depend on species, cell-type [135] or light level [107]. Whether generated in the outer or inner retina, the surround is present at the level of the bipolar cells presynaptic to a RGC, this can be seen in direct recordings from bipolar cells [51], measurements of bipolar cell output [32, 81] or in recordings of RGC excitatory input (see Fig. 4.1D,E). Importantly, then, inputs from the surround enter the

retinal circuit upstream of the crucial nonlinear transformation that takes place at the bipolar cell to ganglion cell synapse.

With this circuit organization in mind, and given the poor understanding of the function of the RF surround, here we set out to ask two distinct, but related questions about the role of the surround in natural vision. (1) How does the surround combine with the center to jointly determine a cell's response? And (2) how does the surround affect the impact of the nonlinear subunits within the RF center?

We explore these issues in On and Off-center parasol (magnocellular-projecting) RGCs in the macaque monkey retina. The nonlinear subunit structure of these cells has been well characterized [35, 48, 162], and the impact of this RF organization has been shown to be important for encoding of natural visual stimuli [226].

4.3 Results

Here we use single cell patch electrophysiology in an *in vitro* macaque retinal preparation in conjunction with computational modeling of the RGC RF to explore the impact of the surround on natural scene encoding. First, we use natural movies and natural luminance stimuli to show that inputs to the RF center and surround combine nonlinearly. We find that this nonlinear interaction is especially prominent under natural stimulus conditions because of strong intensity correlations present in natural images. Next we show that a simple model in which inputs to these two RF regions combine linearly before passing through a shared rectifying nonlinearity can capture the nonlinear interaction between center and surround. We hypothesized that the shared nonlinearity governing center-surround interactions is the nonlinear synaptic transfer from bipolar cell to RGC, and that because of this inputs to the surround could modulate nonlinear spatial integration (and thus spatial contrast sensitivity) in the RF center. Finally, we confirm this hypothesis using both artificial grating stimuli and natural images. These results suggest that RGC sensitivity to spatial contrast in natural scenes is modulated by context via a simple nonlinear interaction between the RF center and surround.

4.3.1 RF center and surround interact nonlinearly during naturalistic visual stimulation

To explore how inputs to the RF center and surround combine to determine a cell's response under naturalistic stimulus conditions, we used a visual stimulus designed to mimic natural primate viewing conditions. We masked this naturalistic visual stimulus to selectively activate either the RF center or surround. At the beginning of each experiment, we found the center of the cell's receptive field (see Experimental Procedures) and mapped the linear center-surround RF structure using expanding spots stimuli. Both On and Off parasol RGCs show classical center-surround RF organization in their spike responses (Fig. 4.1A,B shows an example Off parasol RGC). By holding the cell at the inhibitory reversal potential in whole-cell voltage clamp recordings, we isolated the excitatory synaptic input to the cell. Expanding spots stimuli show that the center-surround RF organization is present at the level of excitatory input as well (Fig. 4.1C,D), in agreement with previous observations in parasol RGCs [47, 226] as well as other primate RGC types [169].

Our naturalistic visual stimuli were based on the Database Of Visual Eye movementS (DOVES, [227]), which combine grayscale natural images [229] and measured human eye movements. An example image and a corresponding eye movement trajectory is shown in Fig. 4.1E. Natural movies were presented at a mean luminance level that caused approximately 5,000 cone opsin isomerizations (R*) per M or L cone per second (see Experimental Procedures for details on visual stimulation).

Using measured linear RF sizes as in Fig. 4.1A-D, we presented our natural movie stimuli while isolating regions of the RF using gray masks and/or apertures over each frame (Fig. 4.1E, right). Fig. 4.1F shows an example Off parasol RGC's spike responses to each of the three movie stimuli: stimulation of the center alone (Fig. 4.1F, purple), stimulation of the surround alone (Fig. 4.1F, blue), or simultaneous stimulation of both the center and surround (Fig. 4.1F, black). Responses to isolated center or surround stimuli are shown in Fig. 4.1G for the spike output of the cell (Fig. 4.1G, top) and the excitatory synaptic input to the cell (Fig. 4.1G, bottom).

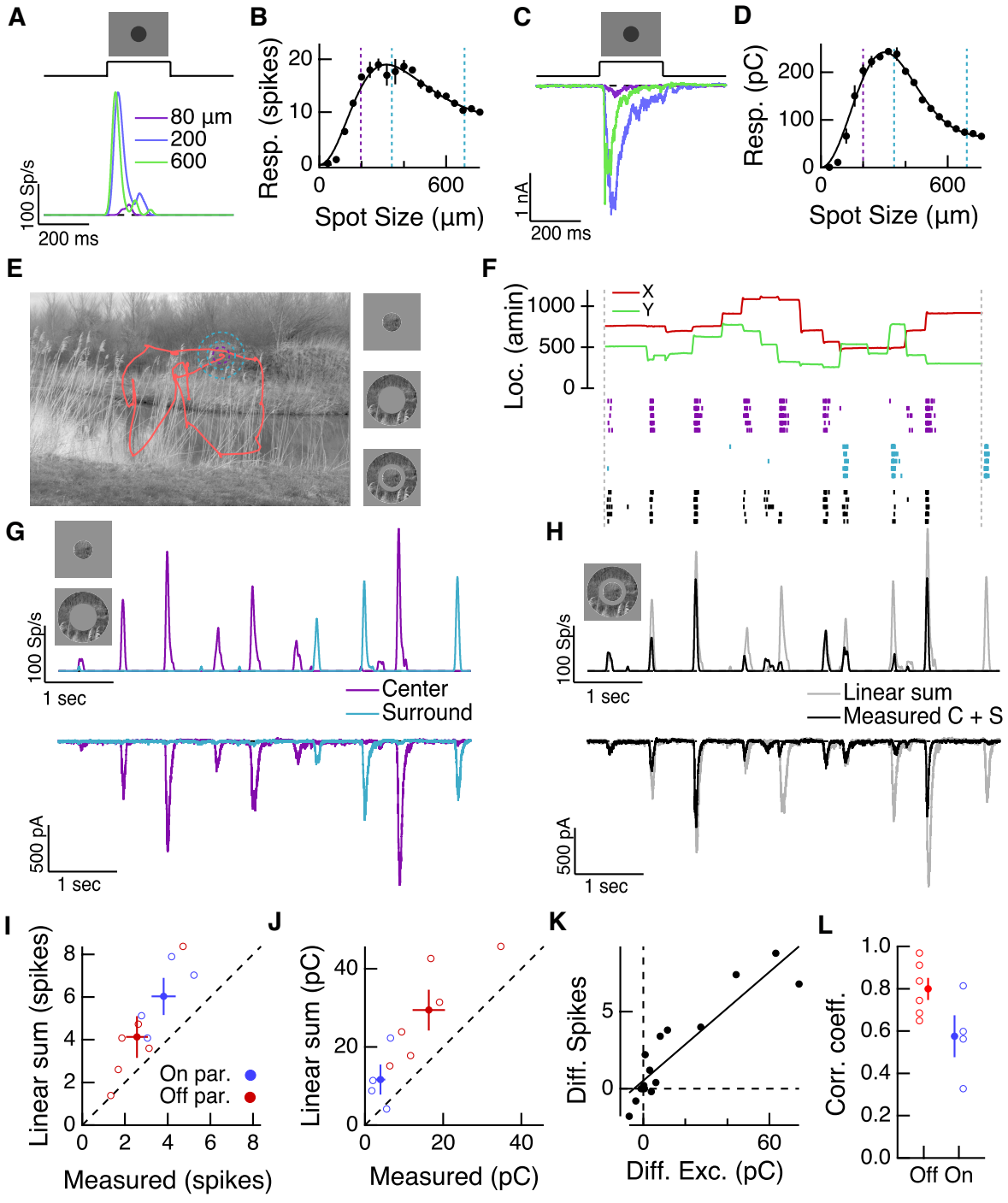
A
80 µm
200
600
100 Sp/s
200 ms
B
Resp. (spikes)
20
10
0
0
600
Spot Size (µm)
C
1 nA
200 ms
D
Resp. (pC)
200
100
0
0
600
Spot Size (µm)
E
F
Loc. (amin)
1000
500
0
X
Y
G
100 Sp/s
1 sec
Center
Surround
500 pA
1 sec
H
100 Sp/s
1 sec
Linear sum
Measured C + S
500 pA
1 sec
I
Linear sum (spikes)
Measured (spikes)
On par.
Off par.
J
Linear sum (pC)
Measured (pC)
K
Diff. Spikes
Diff. Exc. (pC)
L
Corr. coeff.
Off On

Figure 4.1: (Previous page) **Natural movie stimuli elicit nonlinear interactions between the RF center and surround.** (A) Off parasol RGC spike responses to expanding spots stimuli. (B) Area-summation curve from the cell in (A), points show mean ± S.E.M., smooth curve is a difference-of-Gaussians RF model fit (see Experimental Procedures). (C-D) Same as (A-B) for excitatory current responses showing that the surround is present in the excitatory inputs to these cells. (E) Natural image and associated eye movement trajectory from [227]. Right: example movie frames showing isolated center (top), surround (middle), and center-surround stimuli (bottom). (F) Rasters show example Off parasol RGC spike responses to these three movie stimuli. Top shows eye movement position. (G) Spike output (top) and excitatory synaptic input (bottom) responses to isolated center and surround stimuli. (H) Spike and excitatory synaptic input responses to the center-surround stimulus. Gray trace shows the linear sum of isolated center and surround responses. The measured center-surround response is generally weaker than the linear sum of center and surround responses. (I-J) We isolated natural movie fixation responses (see Experimental Procedures) and compared the mean response magnitude (spike count in I, excitatory charge transfer in J) in response to the center-surround stimulus to the linear sum of isolated center and surround response magnitudes. Each point is a different natural movie. The tendency of the points to lie above the line of unity indicates that center and surround, on average, sum sub-linearly under these stimulus conditions (Spike responses (I), $p = 0.02$; excitatory current responses (J), $p = 0.005$). (K) For the example in (E-H), the difference between measured and linearly-summed spike responses was correlated with differences in excitatory synaptic inputs ($r = 0.91$). (J) Population data for the analysis in (K), which is consistent with the interpretation that sub-linear center-surround additivity arises before nonlinearities in synaptic integration or spike generation.

We asked whether inputs to the center and surround interacted linearly by comparing the linear sum of center and surround responses (Fig. 4.1H, gray traces) to the measured response to simultaneous stimulation of both the center and surround (Fig. 4.1H, black traces). For both spike and excitatory current responses, the measured center-surround response was weaker than the linear sum of the two responses measured independently. This sub-linear additivity indicates a nonlinear interaction between the RF center and surround. Moreover, because this interaction is present at the level of excitatory synaptic input to the cell, it is not due to nonlinearities in synaptic integration or spike generation.

We quantified responses to these stimuli by detecting fixations within the stimulus using a threshold on the instantaneous velocity of the eye trajectory and integrating a cell's response within each fixation. The mean fixation response across several different natural movies and cells shows that for both spike output (Fig. 4.1I) and excitatory synaptic input (Fig. 4.1J),

center and surround inputs sum sub-linearly. Moreover, the difference between the linear sum of center and surround inputs and the measured simultaneous response is correlated between excitatory inputs and spike outputs within each cell (Fig. 4.1K,L). This is consistent with the interpretation that the nonlinear interaction seen at the level of spike output is present in the excitatory synaptic inputs as well.

4.3.2 Natural spatial correlations promote nonlinear center-surround interactions

Does the nonlinear center-surround interaction depend upon some feature of natural scenes that may not be prominent in artificial stimuli? The independent center and surround stimuli tended to drive the cell at non-overlapping periods of the stimulus. This is consistent with the antagonistic nature of the surround (e.g. an Off-center parasol RGC has an On-surround) as well as the spatial correlations in intensity that characterize natural images [211]. This means that periods of the stimulus that strongly activate the center tend to deactivate the surround and vice versa. We tested the effect of spatial correlations on the nonlinear center-surround interaction using a synthetic visual stimulus inspired by our natural movie stimuli.

For this stimulus, we randomly sampled intensities from a natural image (Fig. 4.2A,B) and presented those to either the center, surround, or center and surround simultaneously. To compute the center and surround intensities, we passed a circle (for the center) and annulus (for the surround) over a given image and computed the mean intensity within each region. The intensity correlations characteristic of natural scenes were evident when we plotted the center intensity against the corresponding surround intensity (Fig. 4.2C, left, "Control surround"). When we shuffled the surround intensities relative to those of the centers, those spatial correlations were eliminated but the stimulus still maintained the same marginal distributions (Fig. 4.2C, right, "Shuffled surround").

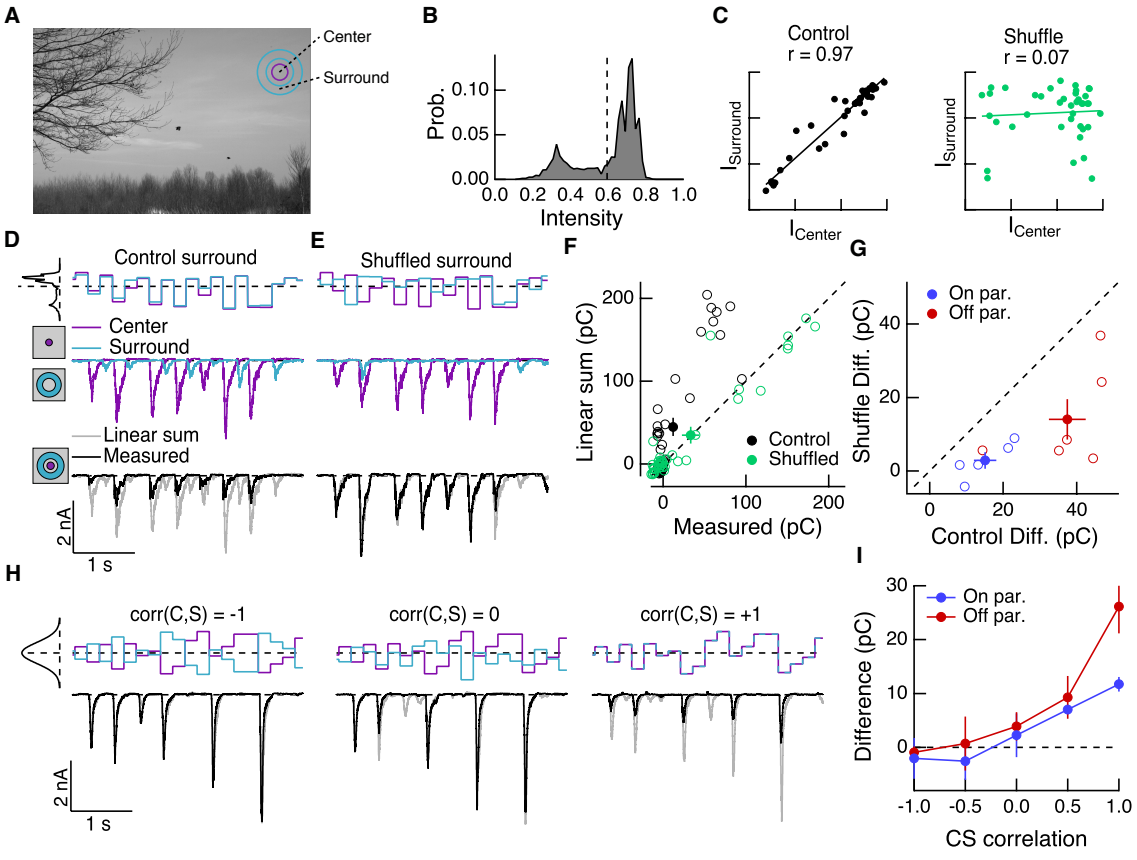
A
Center
Surround
B
Prob.
0.10
0.05
0.00
0.0 0.2 0.4 0.6 0.8 1.0
Intensity
C
Control
r = 0.97
I Surround
I Center
Shuffle
r = 0.07
I Surround
I Center
D
Control surround
Center
Surround
Linear sum
Measured
2 nA
1 s
E
Shuffled surround
F
Linear sum (pC)
200
100
0
0 100 200
Measured (pC)
Control
Shuffled
G
Shuffle Diff. (pC)
40
20
0
0 20 40
Control Diff. (pC)
On par.
Off par.
H
corr(C,S) = -1
corr(C,S) = 0
corr(C,S) = +1
2 nA
1 s
I
Difference (pC)
30
20
10
0
-1.0 -0.5 0.0 0.5 1.0
CS correlation
On par.
Off par.

Figure 4.2: (Previous page) **Spatial correlations in natural scenes promote nonlinear center-surround interactions.**(A) Example natural image [229] used to construct natural intensity stimuli. Inset shows center and surround regions used in local mean intensity measurements. (B) Intensity histogram from the image in (A). Dashed vertical line indicates the mean intensity, which was used as the mean gray level in experiments that follow. (C) Center and surround intensity values for 40 image patches from the image in (A). When surround measurements were associated with their respective center measurements, strong intensity correlations were present (left). Shuffling surround values relative to the center breaks these correlations (right). (D) Example stimuli (top) and Off parasol RGC excitatory current responses to isolated center and surround (middle) and center-surround (bottom) stimulation. Gray trace in bottom shows linear sum of isolated center and surround responses. (E) Same as (D) for shuffled surround intensities. (F) The response magnitude (charge transfer) of each fixation is plotted for measured center-surround and linearly-summed center and surround responses. Open points show mean responses for each fixation, filled points show mean ± S.E.M. across all fixation responses in this example cell. Control responses (black) generally show sub-linear center-surround additivity (as in Fig. 4.1), but shuffling surround intensities relative to the center decreases these nonlinear interactions. (G) Population data showing the mean difference between responses to the center-surround stimulus and the linearly-summed response. Open circles show average differences for each cell tested, and filled points show population mean ± S.E.M. For all cells tested, the control stimulus generated stronger sub-linear additivity than did the shuffled stimulus ($n = 5$ On parasol RGCs, $p = 0.001$; $n = 6$ Off parasol RGCs, $p = 0.006$). (H) We generated white noise center-surround stimuli that had variable center-surround correlations but constant marginal distributions. Shown are example excitatory current responses in an Off parasol RGC. Black traces show the measured center-surround stimulus response and gray traces show the linearly-summed center and surround responses. When center and surround inputs were anticorrelated (left), center and surround interacted approximately linearly. As center-surround correlations increased (center and right), sub-linear center-surround additivity became more pronounced. (I) Population data from the experiments in (H) showing that nonlinear center-surround interactions depend on the correlation between center and surround inputs ($n = 3$ On parasol RGCs; $n = 2$ Off parasol RGCs).

We presented a modulated disc to the center and/or annulus to the surround, updating the intensity of each region every 200 msec, which is consistent with typical human fixation periods [227]. When spatial correlations were intact, inputs to the center and surround combined sub-linearly in the excitatory synaptic input to the cell (Fig. 4.2D), as with the full natural movie responses (Fig. 4.1). When we shuffled the surround intensities relative to the center, thus eliminating the spatial intensity correlations seen in natural images, nonlinear center-surround interactions were much weaker (Fig. 4.2E,F). This was true for

excitatory synaptic inputs to both On and Off parasol RGCs (Fig. 4.2G).

To further probe the impact of intensity correlations on center-surround interactions, we generated Gaussian random noise stimuli with a tunable degree of correlation between center and surround inputs, ranging from -1 (perfectly anti-correlated noise) to 0 (uncorrelated noise) to +1 (perfectly correlated noise that is modulated in unison). When noise stimuli in the center and surround were negatively correlated, inputs to the center and surround summed linearly (Fig. 4.2H, left). As the center-surround intensity correlations increased, sublinear additivity became more obvious (Fig. 4.2H, middle and right), with strongly positively correlated noise stimuli inducing center-surround interactions that resembled those seen using naturally correlated luminances (Fig. 4.2H, right). This dependence of nonlinear center-surround interactions on center-surround intensity correlations was robust and present in both On and Off parasol RGCs (Fig. 4.2I).

4.3.3 Inputs to the center and surround combine linearly and pass through a shared nonlinearity

What is the source of the observed nonlinear center-surround interaction, and what sort of circuit architecture is needed to account for this behavior? To answer this question we used a linear-nonlinear cascade modeling approach. We presented Gaussian-distributed noise to either the center, surround or both while measuring excitatory synaptic inputs to On and Off parasol cells (Fig. 4.3A). Note that here the inputs to the center and surround are uncorrelated, which allows us to explore the space of center and surround activation in an unbiased manner. Using trials in which only the center or surround were stimulated in isolation (Fig. 4.3A, left and middle), we computed linear filters for each RF region using reverse correlation. For this analysis we present the excitatory conductance by dividing the measured excitatory currents by the driving force. We also presented trials where the center and surround were driven simultaneously (Fig. 4.3A, right).

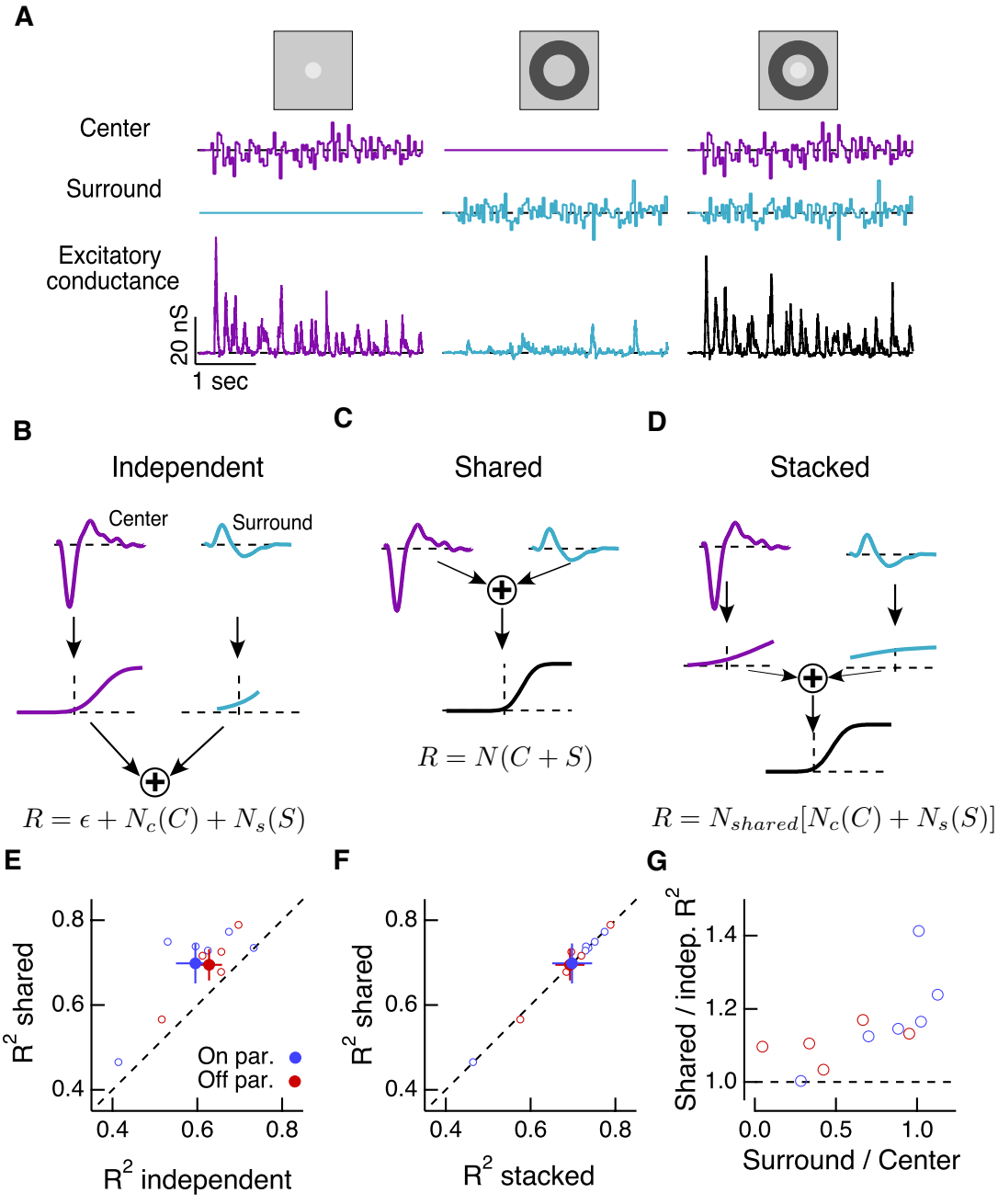
A
Center
Surround
Excitatory conductance
20 nS
1 sec
B
Independent
Center
Surround
$R = \epsilon + N_c(C) + N_s(S)$
C
Shared
$R = N(C + S)$
D
Stacked
$R = N_{shared}[N_c(C) + N_s(S)]$
E
R^2 shared
R^2 independent
On par.
Off par.
F
R^2 shared
R^2 stacked
G
Shared / indep. R^2
Surround / Center

Figure 4.3: **Linear-nonlinear cascade modeling supports an architecture where center and surround combine linearly before passing through a shared nonlinearity.** (A) We presented Gaussian white noise (uncorrelated between center and surround) to either the center (left), surround (middle) or center and surround simultaneously (right) while measuring excitatory synaptic current responses. Note that we converted measured excitatory currents to an estimate of the excitatory conductance (in nS) by dividing measured currents by an estimate of the relevant driving force. Example traces are from a representative Off parasol RGC. From isolated center and surround trials we computed the linear filters for each RF region via reverse correlation. We used these measured linear filters and the trials in which center and surround were stimulated simultaneously to generate three models of center and surround interactions. (B) The independent model treats the filtered center and surround inputs with private nonlinear functions, and the outputs of these two nonlinearities are then summed to produce the excitatory conductance response. (C) The shared model integrates filtered center and surround inputs linearly, and this summed input is then passed through a single, shared nonlinearity. (D) The stacked model combines the independent and shared models by treating center and surround with independent nonlinearities before summation and treatment with a third, shared nonlinearity. (E) We tested the ability of each of these models to predict held-out responses to center-surround stimulation. The shared nonlinearity model better accounted for excitatory current responses than did the independent nonlinearity model for both On and Off parasol RGCs ($n = 6$ On cells, $p = 0.02$; $n = 5$ Off cells, $p = 0.01$). (F) The shared model performed no worse than the stacked model ($p > 0.75$ for both On and Off cells), despite the former being a special case of the latter with many fewer free parameters. (G) The relative improvement of the shared over the independent model is positively correlated with each cell's relative surround strength ($r = 0.67$, $p = 0.02$).

Using the linear filters measured from independent trials, we constructed three models of how center and surround inputs combine to determine the cell's excitatory conductance response: 1) In the "independent" model (Fig. 4.3B), inputs to the center and surround are filtered using their respective linear filters and then passed through separate nonlinear functions. The outputs of the two nonlinearities are then summed to give the response of the cell. 2) In the "shared" model (Fig. 4.3C), filtered center and surround inputs are summed linearly before passing through a single, shared nonlinear function. The output of this nonlinearity is the cell's response. 3) The "stacked" model (Fig. 4.3D) combines models (1) and (2) by treating center and surround with private nonlinearities before summation and treatment with a third, shared nonlinearity. See the Experimental Procedures section for details of the modeling.

We fit each model using a subset of simultaneous center-surround trials (Fig. 4.3C, right) and used the remaining simultaneous trials to test how well each model could predict the cell's response. For all cells tested, we found that the shared model outperformed the independent model (Fig. 4.3E). The shared model performed as well as the more complicated stacked model (Fig. 4.3F), despite the latter having many more free parameters (10 free parameters) than the shared model (5 free parameters).

Note that the shared model is a special case of the stacked model. This is strong support for the notion that the dominant nonlinear interaction is characterized by a shared nonlinearity, and that upstream of this nonlinearity center and surround interact approximately linearly. This is also consistent with the observation that the private nonlinearities fit in the stacked model tended to be quite shallow and much nearer to linear than the sharply rectified downstream, shared nonlinearity (e.g. see Fig. 4.3D).

Finally, there is variability across cells in the relative improvement of the shared over the independent model (Fig. 4.3E). We found that some of this variability is the result of the strength of the surround. Cells that had a stronger surround (indicated by the amplitude of their surround filter relative to their center filter) tended to show a greater improvement with the shared model (Fig. 4.3G).

4.3.4 The RF surround regulates effective rectification of synaptic inputs to the center

The experiments used in the modeling above also allowed us to directly examine the relationship between center and surround activation and the cell's excitatory synaptic response. An example of such a joint response distribution for an Off parasol RGC is shown in Fig. 4.4A. These response distributions show that the nonlinear relationship between center activation and excitatory conductance response depends on the degree of surround activation: when the surround is strongly suppressive, this nonlinearity is sharply rectified (Fig. 4.4A,B, blue trace), but when the surround is facilitatory, this nonlinearity becomes less rectified and nearer to linear (Fig. 4.4A,B, red trace).

We reasoned that this rectifying nonlinearity is located at the synaptic output of the

bipolar cells presynaptic to the ganglion cell. This is supported by the observation that the surround is present at the level of the bipolar cell, and therefore at least some portion of the surround enters the circuit before the nonlinearity at the bipolar cell synapse. A rectified bipolar cell synaptic nonlinearity has been well described in the retina [32, 93, 108, 192], including at the output synapse of diffuse cone bipolar cells in primate retina [92, 226].

This circuit interpretation accounts for the dependence of center rectification on surround activation by a dynamic change in the set point of the synaptic nonlinearity (Fig. 4.4C). When the surround is weakly activated or suppressive, inputs to the center explore a relatively rectified local region of the synaptic nonlinearity; a facilitatory surround, on the other hand, pushes the synapse into a more linear state (Fig. 4.4C, right).

For the sake of concreteness we have assumed that the nonlinear relationship between bipolar cell input and excitatory input to the RGC arises at the presynaptic terminal, but we note that some contribution from nonlinearities earlier in the bipolar cell is likely as well. For example, active conductances in the axons of diffuse bipolar cells would also be expected to shape the nonlinear relationship between bipolar cell input and synaptic release [170]. Active conductances in the bipolar cell axon would be expected to more heavily shape a surround generated in the outer retina compared to an inner-retinal surround.

Modulation of the bipolar cell synaptic set point can result from changes in mean luminance [93] as well as genetic or pharmacological manipulations that change the resting membrane potential of AII amacrine cells [149]. The principle here is similar except in this case the synaptic set point is being dynamically modulated with surround activation. We hypothesized that one consequence of dynamic modulation of synaptic rectification would be to adjust spatial contrast sensitivity in the RF center. This is because for many classes of RGC, including parasol cells, bipolar cells with a rectifying synaptic nonlinearity act as nonlinear subunits within the RF center [61, 126, 192]. Nonlinear spatial integration can confer sensitivity to spatial contrast at a sub-RF center spatial scale. This manifests itself in, for example, sensitivity to drifting or contrast-reversing grating stimuli [35, 48, 69, 100, 162] that would otherwise not drive responses in spatially linear RGCs.

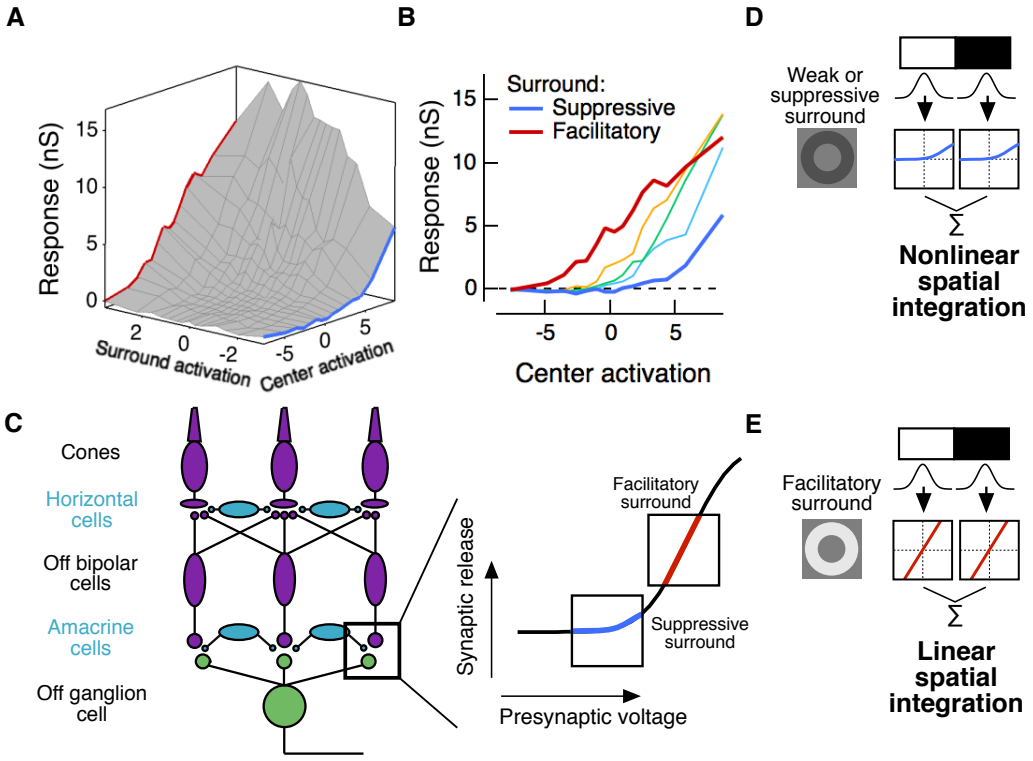

Figure 4.4: **The RF surround changes the apparent rectification of inputs to the center.** (A) From the white noise experiments in Fig. 4.3, we generated a response surface showing the mean excitatory conductance response from an Off parasol RGC as a two-dimensional function of filtered inputs to both the center and surround (center or surround "activation", i.e. their generator signals). (B) Sections through this surface at various levels of surround activation reveal that the shape of the nonlinear dependence of excitatory conductance on center activation changes dramatically as the surround is modulated. In particular, in the context of a suppressive surround, the center is sharply rectified (blue curve) while a facilitatory surround is associated with a more linear, less rectified center nonlinearity (red curve). (C) In light of the shared nonlinearity architecture revealed in Fig. 4.3 and what is known about the circuit basis of the RF surround, we hypothesized that the shared nonlinearity in this circuit lies at the synapse from presynaptic bipolar cell to postsynaptic RGC. A suppressive surround hyperpolarizes the bipolar cell terminal, pushing the synapse into a more rectified state (blue portion of synaptic nonlinearity). A facilitatory surround depolarizes the terminal and pushes the synapse into a more locally-linear state (red portion of curve). (D) If this circuit interpretation is correct, we should see predictable differences in the spatial integration properties of the RF center, since bipolar cells act as nonlinear subunits for these RGCs. A weak or suppressive surround will be associated with rectified subunit output nonlinearities, and thus nonlinear spatial integration and sensitivity to spatial contrast (e.g. a split-field grating stimulus). (E) A facilitatory surround will be associated with more linear subunit outputs and the RGC should integrate across visual space approximately linearly.

4.3.5 The RF surround modulates spatial contrast sensitivity in the center

Because nonlinear spatial integration relies so critically on rectification of subunit outputs (i.e. synaptic rectification at the bipolar cell axon terminal), we hypothesized that surround modulation could dynamically regulate nonlinear spatial integration in parasol RGCs. When a weak or suppressive surround is present, bipolar cell terminals presynaptic to the RGC are in a rectified state, and the RF center integrates nonlinearly over visual space (Fig. 4.4D). When a facilitatory surround is present, however, the bipolar cell synapse is relatively more linear, and the RF center integrates linearly over visual space (Fig. 4.4E).

We tested this hypothesis in Off parasol RGCs, because these cells show a stronger degree of rectification of subunit output compared to On parasol cells [39] as well as nonlinear spatial integration in the context of naturalistic visual stimuli [226]. We first measured nonlinear spatial integration by recording spike responses of Off parasol RGCs to a split-field grating restricted to the RF center. Because of the nonlinear subunit structure of the RF center, Off parasol cells respond robustly to such stimuli (Fig. 4.5A, left). We then presented the same grating within an annulus in the surround as well as, in separate trials, the surround annulus alone (Fig. 4.5A, right). When paired with a bright surround, the grating stimulus did not activate the cell much beyond its response to the surround alone.

We repeated this experiment for surround contrasts ranging from +0.9 to -0.9 (Fig. 4.5B). While the surround-free grating (surround contrast = 0) stimulus showed strong nonlinear integration (indicated by its distance away from the unity line in Fig. 4.5B), the presence of a surround diminished the response to the grating (indicated by the tendency of non-zero surround contrast points to lie closer to the line of unity). This was true for a range of central grating contrasts (Fig. 4.5C), indicating that this behavior is not the result of some post-integration saturation of the response.

We repeated these experiments while measuring excitatory synaptic inputs to Off parasol cells in voltage-clamp recordings (Fig. 4.5D). As predicted from our original hypothesis (Fig. 4.4), the surround modulation of nonlinear spatial integration was present in the excitatory

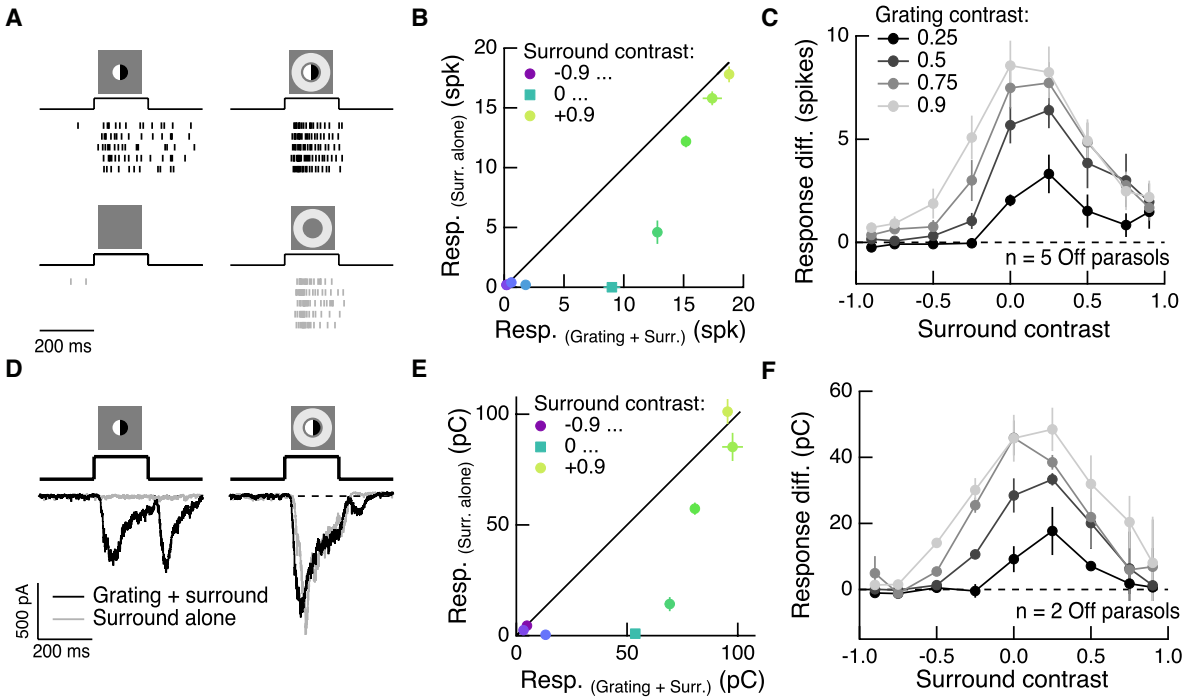

Figure 4.5: **The RF surround regulates nonlinear spatial integration in the RF center.** We set out to directly test the hypothesis outlined in Fig. 4.4D,E by presenting split field grating stimuli to Off parasol RGCs while modulating inputs to the RF surround. (A) Left column: Example Off parasol RGC spike response showing nonlinear response to an isolated split-field grating stimulus restricted to the RF surround. The linear equivalent stimulus (i.e. no stimulus) causes no response in the cell. Right column: when the center stimulus is paired with a bright surround, the grating and the linear equivalent stimulus produce very similar responses. (B) For the example cell in (A), we tested sensitivity to the center grating stimulus with a range of contrasts presented to the surround. Negative contrast surrounds (suppressive for these cells) decrease the spike response. Positive contrast surrounds (facilitatory) sum sub-linearly with the grating stimulus such that for the brightest surrounds, the addition of the grating only mildly enhances the cells response. Points show mean ± S.E.M. spike count response. (C) We measured the response difference between the grating stimulus and the surround-alone stimulus across a range of surround contrasts (horizontal axis) and for four different central grating contrasts (different lines). For each grating contrast, addition of either a bright or dark surround decreased sensitivity to the added grating. Points are population means ± S.E.M. ($n = 5$ Off parasol RGCs). (D-F) same as (A-C) for excitatory synaptic currents in an Off parasol RGC. Response magnitude is quantified as excitatory charge transfer and points represent mean ± S.E.M. response for the example cell in (E) and population mean ± S.E.M. ($n = 2$ Off parasol RGCs) in (F). The same change in spatial integration properties with surround modulation is seen in the bipolar cell input to the cell, which is consistent with the hypothesis outlined in Fig. 4.4D,E.

inputs to these cells and does not seem to be substantially shaped by inhibitory input, post-synaptic integration, or spike generation mechanisms (Fig. 4.5E,F).

For both the spike output and excitatory synaptic inputs to Off parasol RGCs, the peak spatial nonlinearity was observed with no surround or with slightly positive contrast surrounds (Fig. 4.5C,F). This indicates that (at least for the mean luminance levels used here), the bipolar cell synapse is almost ideally situated to promote nonlinear spatial integration in the absence of surround modulation. When the surround is strongly suppressive (i.e. for these cells, a dark surround), the response appears to be diminished as a result of the surround pulling the synapse into a highly quiescent state, and the grating stimulus cannot overcome the threshold for synaptic release. When the surround is facilitatory (i.e. a bright surround), the nonlinear sensitivity of the center is reduced.

Prior work has shown that nonlinear spatial integration by Off parasol RGCs endows these cells with sensitivity to spatial contrast present in natural images [226]. Because of this we next tested whether the surround can modulate spatial contrast sensitivity in the context of natural images as it can for artificial gratings stimuli. To do this, we measured Off parasol RGC spike responses to natural image patches that contain spatial contrast (see Experimental Procedures for details on images and patch selection). For each image patch we also presented a linear equivalent disc stimulus, which is a uniform disc with intensity equal to the weighted sum of the pixel intensities within the RF center. The weighting function used was a circular Gaussian function representing the linear RF center. The parameters of this linear RF center model were determined using a difference-of-Gaussians model fit to expanding spots data for each cell (as in Fig. 4.1B). A cell whose RF center is behaving according to this measured linear RF model will respond equally to a natural image and its associated linear equivalent disc.

As shown previously, Off parasol cells tend to respond much more strongly to natural images than a linear equivalent disc, especially when the natural image contains high spatial contrast (Fig. 4.6A, left, see also [226]). Similar to the experiments using grating stimuli, when a bright surround was presented simultaneously, the natural image and its linear

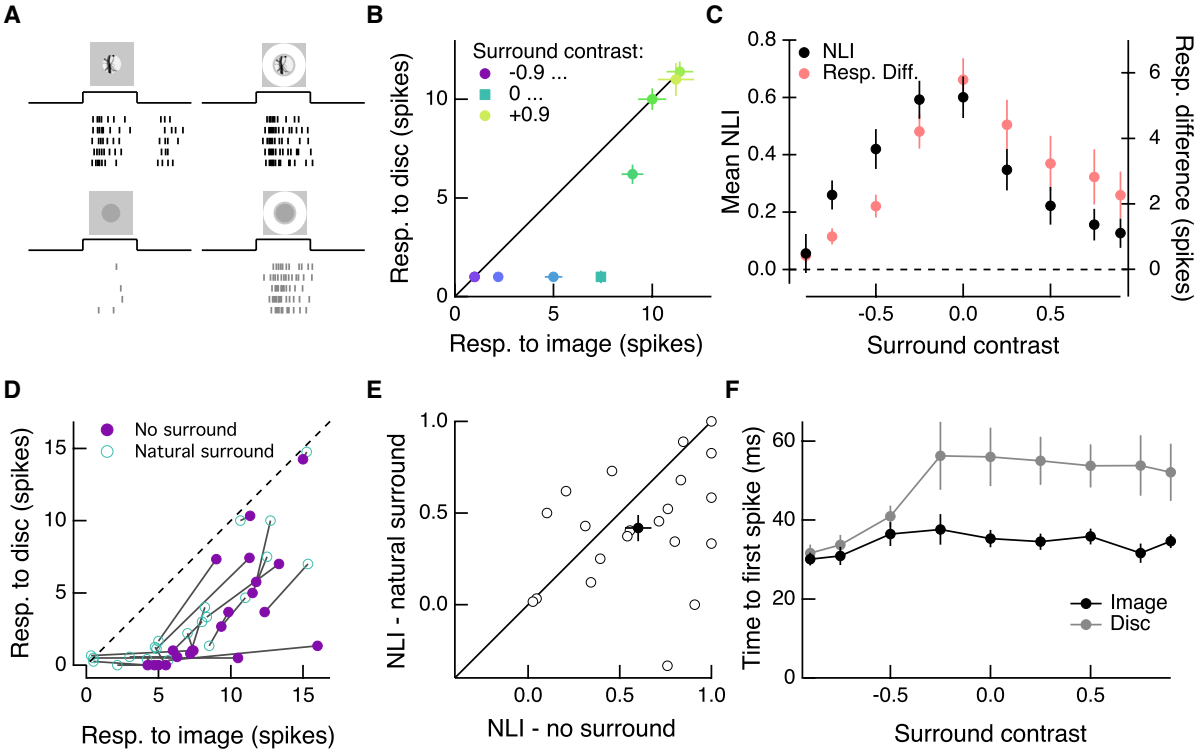

Figure 4.6: **Inputs to the RF surround modulate spatial contrast sensitivity of the center during natural scene stimulation.** (A) We presented a natural image patch and its linear-equivalent disc stimulus (see text and Experimental Procedures) to probe sensitivity to spatial contrast in natural scenes. Example Off parasol RGC spike responses are shown. The addition of a bright surround (right) makes responses to the natural image and its linear equivalent stimulus more similar. (B) Spike count responses to an example image patch and its linear equivalent disc across a range of surround contrasts. The addition of a sufficiently bright surround (top three points) eliminates sensitivity to spatial contrast in this image patch. (C) Population summary showing, as a function of surround contrast, the difference between image and disc responses (right axis, red) and the nonlinearity index (NLI, left axis, black, see text) ($n = 21$ image patch responses measured in 5 Off parasol RGCs). (D) We also presented the naturally-occurring mean intensity to the surround wile presenting each image patch and its associated linear equivalent stimulus. Connected points show the same patch measured without a surround (filled, purple points) and with the "natural" surround (open points). Most surrounds were suppressive (points pulled towards the origin), but many of those that were facilitatory decreased the response difference between the natural image and disc stimuli. (E) From the data in (D), the NLI in the presence of the "natural" surround compared to the center-only case. Each point is a different image patch. Filled point shows population mean ± S.E.M. The addition of a surround tended to decrease sensitivity to spatial contrast in the center (but not always). (F) We measured the time to the first spike during each response (for non-zero spike counts) to each image patch and its associated linear equivalent disc. With a weakly suppressive or absent surround, a response delay exists between the image and its equivalent disc. This delay persists, and stays relatively constant, with the addition of a facilitatory surround.

equivalent disc stimulus drove the cell approximately equally (Fig. 4.6A, right). The dependence of nonlinear integration of natural images on surround intensity was similar to that for gratings (Fig. 4.6B). This effect persisted across many image patches and cells tested (Fig. 4.6C), indicated by the change in response difference between image and disc stimuli as well as a drop in the nonlinearity index (NLI), which is simply the response difference normalized by the response sum (See equation 1).

$$NLI = \frac{r_{image} - r_{disc}}{r_{image} + r_{disc}} \quad (1)$$

The NLI was used as a measure of nonlinear integration because (unlike grating stimuli) different image patches drove responses that differed considerably in absolute scale, and the NLI normalizes each patch/surround pair response by the response amplitude.

In addition to the parameterized surround contrasts presented around natural image patches, we also presented trials in which the intensity of the surround was meant to mimic the activation of the (spatially linear) surround that would occur in the full natural scene. We found this "natural" intensity by computing a weighted average of the pixels in the natural image surrounding each image patch using a circular Gaussian function representing linear spatial integration by the RF surround. The size of this circular RF surround was measured using a difference-of-Gaussians model fit to expanding spots data (as in Fig. 4.1B,D).

The inclusion of the naturally occurring surround luminance tended to decrease the NLI for an image patch (Fig. 4.6D,E), both because of a suppressive surround (points that move towards the origin in Fig 4.6D) and because of a facilitatory surround (points that move up towards the line of unity in Fig. 4.6D).

While the spike count responses between an image patch and its linear equivalent disc are more similar with the addition of a surround (Fig. 4.6C), we note that there seem to be differences in spike responses to these two stimuli that persist in the presence of a surround. For example, the surround drives a delayed response relative to the center. This can be seen most clearly in the excitatory responses to the grating stimulus in Fig. 5D. Indeed, we found that responses to the linear equivalent disc were delayed by about 20 msec relative to image responses, with or without the inclusion of a surround stimulus (Fig. 4.6F). This

timing delay is consistent with the delay seen between the center and surround linear filters measured in Fig. 4.3. Thus, although discriminating responses based on spike count becomes more difficult with the addition of a surround, significant spike timing cues seem to persist that could, in principle, be used to extract information about the spatial structure of an image.

4.4 Discussion

Here we have shown that inputs to the RF center and surround of parasol RGCs interact nonlinearly. Moreover, strong intensity correlations present in natural images enhance the impact of nonlinear center-surround interactions in shaping the ganglion cell response. These nonlinear center-surround interactions are consistent with a circuit architecture in which inputs to the center and surround combine linearly and then pass through the shared, rectifying nonlinearity at the output of bipolar cells. Because the surround enters the circuit before the formation of the full RF center, the surround has the ability to modulate the spatial integration properties of the RF center by changing subunit (non)linearity. Functionally, this means that inputs to the surround can change a cell's sensitivity to small-scale spatial contrast present in an image, including in natural scenes.

The circuit architecture revealed here is consistent with previous observations that the surround is present in diffuse bipolar cells in primate retina [51]. Past work in parasol RGCs as well as other RGC types has shown that the inclusion of nonlinear subunits in the RF center is important to capture responses to spatially structured stimuli [31, 48, 60, 82, 192], including natural scenes [226]. The present study suggests that to capture observed center-surround interactions and spatial contrast sensitivity, the nonlinear subunits in the RF center should themselves have a center-surround RF organization.

Our analysis of spatial contrast sensitivity using grating or natural image stimuli in the RF center (Figs. 4.5 & 4.6) used spike counts as a metric for RGC responses to argue that modulation of the surround tends to decrease spatial contrast sensitivity in the center. We note that spike timing differences persisted during presentation of the surround, and that

this information could, in principle, be used to extract information about spatial structure in the RF center even in the presence of surround modulation. Indeed, past work in salamander retina has shown that spike timing cues can be used to encode information about spatial structure in a scene [86].

The observation that the RF surround can change nonlinear spatial integration in the center highlights the importance of exploring how neural circuit mechanisms interact under diverse stimulus conditions, including during naturalistic stimulation. By probing neural computation using complex stimuli, including those that mimic natural conditions, we can effectively reveal interactions among common neural mechanisms. Synthetic stimuli can then be designed to further probe those interactions in a more systematic way.

4.5 Experimental Procedures

4.5.1 Tissue preparation

We obtained retinal tissue from Macaque monkeys (*M. nemestrina*, *M. mulatta*, or *M. fascicularis*) via the tissue distribution program at the Washington National Primate Research Center. All procedures were approved by the Institutional Animal Care and Use Committee at the University of Washington. Dissection procedures have been described previously [6, 226]. After enucleation, the eye was hemisected and the vitreous humor was removed mechanically, sometimes assisted by treatment with human plasmin ($\sim$50 μ g/mL, Sigma or Haematologic Technologies Inc.). Retina was dark adapted for $\sim$1 hr, and all subsequent procedures were performed under infrared light using night-vision goggles. The retina and pigment epithelium were separated from the sclera and stored in oxygenated (95%O2/5%CO2) Ames bicarbonate solution (Sigma) in a light-tight container. Retinal mounts were removed from the pigment epithelium and laid photoreceptor-side down onto a poly-D-lysine coated coverslip (BD biosciences). During experiments, retinal tissue was perfused at $7-9$ mL/min with Ames solution at $\sim$32°C.

4.5.2 Patch recordings

Electrophysiological recordings were performed using a Multiclamp 700B amplifier (Molecular Devices). Spike responses were measured using extracellular or loose-patch recordings with an Ames-filled pipette. For voltage clamp recordings, we used low-resistance pipettes (tip resistance ∼1.5-4 MΩ) filled with a Cs-based internal solution (containing, in mM: 105 CsCH3SO3, 10 TEA-Cl, 20 HEPES, 10 EGTA, 5 Mg-ATP, 0.5 Tris-GTP, and 2 QX-314, pH 7.3, ∼280 mOsm). We compensated for access resistance (∼4-8 MΩ) online by 75%. Reported voltages have been corrected for an approximately -10 mV liquid junction potential.

To measure excitatory synaptic inputs in voltage clamp recordings, we held the cell at the expected reversal potential for inhibitory inputs. This was typically around -60 mV, but was adjusted for each cell by delivering light steps at holding potentials near this value until the inhibitory response was eliminated.

4.5.3 Cell identification and selection

We identified On and Off parasol RGCs under infrared illumination based on soma size and morphology as well as characteristic spike responses to light steps centered on the cell. The overall health and sensitivity of the retina was confirmed by delivering a uniform, 5% contrast, 4 Hz modulated stimulus, which produces a robust spike response (at least a few spikes per cycle) in On parasol RGCs in sufficiently sensitive tissue. Sensitivity was continuously monitored (typically before each recording) in this way.

4.5.4 Visual stimulation

Stimuli were presented and data acquired using custom written stimulation and acquisition software packages Stage (stage-vss.github.io) and Symphony (symphony-das.github.io). Lab-wide acquisition packages can be found at https://github.com/Rieke-Lab/riekelab-package and protocols used in this study can be found at https://github.com/Rieke-Lab/turner-package.

Visual stimuli were presented with 60Hz frame rates on an OLED microdisplay monitor (eMagin, Bellevue, WA) focused onto the photoreceptors. Monitor outputs were linearized by gamma correction. Stimuli were calibrated using monitor power outputs, the spectral content of the monitor, macaque photoreceptor spectral sensitivity [24], and a collecting area of 0.37 μm^2 for cones [190] and 1 μm^2 for rods. Unless otherwise noted, mean light levels produced ∼9,000 isomerizations (R*)/M or L-cone/s, ∼2,000 R*/S-cone/s and ∼18,000 R*/rod/s.

Before every parasol RGC recording, we found the center of the cell's RF using a split-field contrast reversing grating stimulus at 4 Hz and 90% contrast. To do this, we translated the grating until the two F2 response cycles were balanced (i.e. we minimized the F1 while maximizing the F2 component of the response). We performed this search in both the horizontal and vertical dimensions. In experiments where we specifically targeted visual stimuli to the RF center or surround, we measured each cell's area-summation curve and fit, online, to these data a circular difference-of-Gaussians RF model in radial coordinates. This model is described in equation 2 below.

$$R = K_C \times (1 - exp(-\frac{r^2}{2\sigma_C^2})) - K_S \times (1 - exp(-\frac{r^2}{2\sigma_S^2})) \quad (2)$$

Where R is the response of the cell, r is the spot radius, and four free parameters describe the shape of the RF, two for each center and surround: K describes the amplitude scaling of each component and σ describes the size. We then chose a boundary for the center that would minimally activate the surround while still filling much of the center and likewise for the surround (e.g. see Fig. 4.1A-B). These boundaries were used to generate appropriate masks and apertures to isolate each RF subregion.

4.5.5 Natural visual stimuli

Naturalistic movie stimuli (Fig. 4.1) were generated using data from the DOVES database [227] (http://live.ece.utexas.edu/research/doves/). The images in the DOVES database were selected from the van Hateren natural image database [229], and we used the original van Hateren images instead of the images included in the DOVES database. This is because the DOVES images are cropped, and using the original, larger images allowed for a wider

selection of eye movement trajectories to be used. To select data from the DOVES database to use in experiments, we ensured that the eye trajectory never extended close enough to the boundaries of the image that our presented frames would extend beyond the boundaries of the image.

For natural movie (Fig. 4.1), luminance (Fig. 4.2), and image stimuli (Fig. 4.6), we scaled the image such that the brightest pixel in the image was assigned an intensity value of 1 (maximum monitor intensity). The mean gray level of the monitor was set to the mean pixel intensity over the entire image. This mean gray level was used in masked/apertured regions of the frame as well as the blank screen between trials. Natural movies were presented at a spatial scale of 1 arcmin/pixel, which is equal to 3.3 μm/pixel on the monkey retina. Natural images, from the van Hateren database, were presented at a scale of 6.6 μm/pixel on the retina.

In the experiments in Fig. 4.2, we updated the intensity of a disc (annulus) in the center (surround) every 200 msec, which is consistent with typical human fixation periods during free-viewing [227], but on the rapid side of the distribution (for efficient data collection). To compute natural intensity stimuli for the center and surround, we selected many random patches from a natural image and measured the mean intensity within a circle of diameter 200 microns (for the RF center) and within an annulus with inner and outer diameters 200 and 600 μm, respectively (for the RF surround).

For the natural image experiments in Fig. 4.6 we used a natural image patch selection method that has been described previously [226]. Briefly, we used a nonlinear subunit model developed in [226] to rank image patches based on their degree of response nonlinearity (i.e. how differently they drove a spatially linear compared to a spatially nonlinear RF model). We then sub-sampled the entire distribution of image patches in order to present an approximately uniform distribution of images that ranged from very nonlinear (i.e. lots of spatial contrast) to very linear (i.e. no spatial contrast). This ensured that we could efficiently explore the full range of spatial contrasts present in natural images within a single experiment.

4.5.6 Data analysis and modeling

Data analysis was performed using custom written scripts in MATLAB (Mathworks). The code used to analyze the data in this study, perform the computational modeling, and generate the figures presented can be found at https://github.com/mhturner/RFsurround. Throughout the study, reported p-values were computed using t-tests. The center-surround models in Fig. 4.3 were constructed as linear-nonlinear cascade models. Using trials where only the center or surround was stimulated in isolation, we used reverse correlation between the cell's response and the noise stimulus to compute the linear filter for both the center and the surround. Each of the three models presented in Fig. 4.3 took as inputs the filtered center and surround noise stimuli.

For each of the nonlinearities ("N" in fig. 4.3) in the models, we parameterized a smooth curve based on a cumulative Gaussian function to fit to the mean response as a function of generator signal (see [38]). Here, $C(\cdot)$ is the cumulative normal distribution (*normcdf* in MATLAB).

$$N(x) = \alpha \times C(\beta x + \gamma) \quad (3)$$

Equation 3 above shows the general form of this nonlinear function, which is determined by three free parameters: γ is an offset along the horizontal axis, β determines the sensitivity or slope of the contrast response function, and α is a scale factor which determines the maximal response. Equations for the three models are described in more detail below.

1. The **Independent model** is described by seven free parameters. Three each for each of the independent nonlinearities of the form described in equation (3) above, and one, ϵ, which sets the vertical offset of the output response. In this and the following equations, h_c (h_s) is the linear filter for the RF center (surround), s_c (s_s) is the stimulus in the center (surround) and the $*$ operator indicates a linear convolution.

$$R = \epsilon + N_c(h_c * s_c) + N_s(h_s * s_s) \quad (4)$$

2. The **Shared model** is described by five free parameters. Three for the single nonlinearity (equation 3), a vertical offset of the response (ϵ), and a scale factor, a, which

sets the relative weight of the center compared to the surround.

$$R = \epsilon + N(a \times (h_c * s_c) + h_s * s_s) \quad (5)$$

3. The **Stacked model** is described by ten free parameters. Three for each of the three nonlinearities (equation 3) and a single vertical offset, ϵ.

$$R = \epsilon + N_{shared}(N_c(h_c * s_c) + N_s(h_s * s_s)) \quad (6)$$

We fit each of the models on 90% of the simultaneous center-surround trials and held the remaining data out for cross-validation purposes. Models were fit with least-squares regression.

Chapter 5

DIRECTION-SELECTIVE CIRCUITS SHAPE NOISE TO ENSURE A PRECISE POPULATION CODE

The study presented in this chapter is the result of a collaborative effort among myself, Joel Zylberberg, Jon Cafaro, Fred Rieke, and Eric Shea-Brown. It has been published as:

∗Zylberberg, ∗J., Cafaro, J., ∗*Turner, M.H.*, Shea-brown, E., and Rieke, F. (2016). Direction-Selective Circuits Shape to Ensure a Precise Population Code. Neuron. 89, 369-383. ∗*Indicates equal contribution.*

I collected the paired extracellular recording data shown in Figs. 5.1 & 5.6 and performed the data analysis shown in Figs. 5.1, 5.3, 5.4, & 5.5. All authors contributed to the conception of the project and to the writing of the manuscript.

5.1 Summary

Neural responses are noisy, and circuit structure can correlate this noise across neurons. Theoretical studies show that noise correlations can have diverse effects on population coding, but these studies rarely explore stimulus dependence of noise correlations. Here, we show that noise correlations in responses of ON-OFF direction-selective retinal ganglion cells are strongly stimulus dependent and we uncover the circuit mechanisms producing this stimulus dependence. A population model based on these mechanistic studies shows that stimulus-dependent noise correlations improve the encoding of motion direction two-fold compared to independent noise. This work demonstrates a mechanism by which a neural circuit effectively shapes its signal and noise in concert, minimizing corruption of signal by noise. Finally, we generalize our findings beyond direction coding in the retina and show that stimulus-dependent correlations will generally enhance information coding in populations of diversely

tuned neurons.

5.2 Introduction

Basic biophysical considerations mean that sensory signals are inevitably corrupted with noise. Divergence of these noisy signals to multiple downstream neurons will cause those neurons' response to covary. The noise correlations that result from such common circuit mechanisms can have diverse effects on coding, ranging from redundant codes, in which groups of cells encode less information than would be predicted from studying the individual cells they contain, to synergistic codes, in which they encode more [9, 104, 109, 181, 191, 200, 201, 242, 245, 247]. Understanding the impact of noise correlations on coding is essential for understanding the fidelity with which neural circuits can compute and direct behavior.

Observed noise correlations are diverse in magnitude and structure. In cortex, average noise correlations are often positive, small, and depend on similarities between the cells' tuning to different stimuli [17, 43, 65, 64, 85, 173, 200]. The small amplitude of noise correlations has been attributed to circuits operating in a balanced state, in which correlated fluctuations in excitatory and inhibitory inputs cancel [91, 95, 176]. However, the balanced state does not always hold [34, 95], and noise correlations can be quite strong. Moreover, noise correlations can depend on neural firing rate [57], and on the stimulus presented [43, 121, 131]. Because of these issues, the extent of correlations between cells and how those correlations are constrained by the synaptic input cells receive is unclear.

Theoretical work provides guidelines for how noise correlations can affect sensory coding: noise that mimics the signals being conveyed by the population will be deleterious to the population code, whereas noise with different statistical structure than the signal is relatively benign. Most theoretical work considers the case where correlations are constant across stimuli and across neural firing rates [1, 9, 56, 156, 158, 200, 202, 247].

Other work suggests that stimulus-dependence can alter the impact of correlations on sensory coding [112, 147, 245]. The importance of this issue is highlighted by studies showing that correlations between cells can be strongly modulated by neural firing rates and stimuli

[28, 80, 90, 121, 57, 127, 131, 167, 189]. Previous theoretical work, however, did not isolate the impact of stimulus-dependence of the correlations in neural populations from other factors such as the diversity of correlation coefficients across the population [112, 245].

Thus, the extent, origins and coding impact of stimulus-dependent correlations remain unclear. This is largely because few experimental preparations permit direct investigation of the circuit mechanisms shaping signal and noise for physiologically relevant stimuli. A notable exception is the population of ON-OFF directionally-selective retinal ganglion cells (ooDS cells), in which the relevant stimulus space (direction of moving objects) is simple and well-described [22, 157]. Knowledge of the stimulus parameters that these cells encode and the accessibility of the retina to mechanistic investigations allowed us to answer several general questions about the role of noise correlations in neural population codes: What is the structure (including stimulus-dependence) of correlated noise in a neural population? How are these correlations generated by the circuitry? What impact do correlations have on the fidelity of sensory encoding? Computational modeling allowed us to generalize our findings beyond the ooDS system. This reveals that correlation structures that would be harmful to the population code in the absence of stimulus dependence lead to significant improvements in coding accuracy in the presence of stimulus dependence.

5.3 Results

Our aim is to understand the impact of noise correlations in ooDS cell populations. First, we show that the direction of a moving stimulus modulates noise correlations between pairs of ooDS cells. Second, we provide a simple example for why stimulus dependence might change how correlations affect population codes. Third, we use intracellular recordings to uncover the circuit mechanisms underlying stimulus-dependent noise correlations. Using this mechanistic description, we develop a computational model to extrapolate beyond our paired recordings and demonstrate that the observed stimulus-dependent correlations improve direction encoding by the ooDS cell population. Fourth, we perform theoretical calculations that suggest that stimulus-dependent noise correlations of the form that are observed in the

ooDS cell population will enhance population codes in other neural circuits.

5.3.1 Correlated variability in the spiking responses of ooDS cell pairs

To characterize noise correlations in ooDS cell spike outputs, we recorded simultaneously from cell pairs while projecting moving bars of light onto the mouse retina. There are four subtypes of ooDS cells based on direction tuning, each of which responds preferentially to motion in one of the cardinal directions [22, 157] (Fig. 5.1A). This means that there are three types of ooDS cell pairs as defined by the separation in their tuning curve peaks (0°, 90°, or 180°). We recorded from neighboring ooDS cell pairs with each possible tuning relation; these neighboring cells have partially overlapping receptive fields, and thus encode motion in the same region of visual space.

We presented moving bars in eight different directions and measured the spike count during the 2-3 second response to each stimulus (Fig. 5.1B). We computed the mean (neural tuning curves; Fig 5.1C) and covariability (noise correlations; Figs. 5.1D-G) of the spike count across many repeats of each bar direction. Both measures varied systematically with bar direction [5, 80]. Moreover, there was a clear relationship between the mean responses of the two cells and their correlation coefficient (Fig. 5.1E,F, Fig. 5.10): higher geometric mean responses (square root of the product of the tuning curves) occurred together with higher correlation coefficients [80]. This effect was independent of the type of ooDS pair recorded and all three types had mostly positive noise correlations (Fig. 5.1E-G). Thus, correlations were strongest when both cells were strongly spiking, regardless of their tuning.

Cell pairs with similar tuning curves (0° pairs) have many stimulus values where both cells fire strongly, and thus have high correlation coefficients when averaged across stimuli (Fig. 5.1G). The dependence of the (average) strength of noise correlations on tuning curve overlap is consistent with the notion of "limited range" correlations in the literature [1, 9, 43, 66, 200, 213, 247]. These "limited range" correlations are typically thought to be harmful to neural population codes [10]. However, most previous examinations of these correlations have only considered the case where they are independent of the stimulus (see [131] for

an exception). Below, we show that stimulus-dependent correlations can lead to different conclusions.

5.3.2 Why might stimulus dependence of noise correlations affect neural population codes?

Noise correlations can be strongly stimulus dependent [28, 80, 90, 121, 57, 127, 167, 189] (Fig. 5.1E,F). Nonetheless, reported noise correlations are often averaged across stimuli. Figure 5.2 below provides a simple example, based on cell pairs, to illustrate how stimulus dependence can change the impact of noise correlations (also see [80]).

The impact of noise correlations on neural encoding depends on the structure of the signal space [9], which describes how the mean population response changes as the relevant stimulus parameter changes. The black line in Fig. 5.2B illustrates the signal space for a hypothetical pair of neurons encoding motion direction with tuning curves shown in Fig. 5.2A (solid curves). Noise spreads responses out along the signal space, leading to ambiguity in the mapping between stimuli and neural responses. The impact of this noise is represented by the projection of the noise onto the signal space (Fig. 5.2B). Noise oriented along the signal space limits the precision with which the signal is encoded, while noise along axes orthogonal to the signal space does not [9, 158, 200].

Stimulus dependence of correlations alters coding by shaping the noise relative to the signal space. To illustrate this effect, we compare stimulus-dependent noise correlations resembling those in our data (Fig. 5.2A, blue line) with stimulus-independent noise correlations (Fig. 5.2A, red line); since the two types of correlations are equal when averaged across stimuli, we refer to the stimulus-independent correlations as "matched" to the stimulus-dependent ones. Stimulus dependence causes the projection of the noise along the signal space to be smaller than the stimulus-independent case for stimuli between the two cells' tuning curve peaks (Fig. 5.2B, "Stim. 1"); other stimuli yield similar projections for stimulus-dependent and matched stimulus-independent noise correlations ("Stim. 2" and "Stim. 3").

This example emphasizes that the responses of populations with stimulus-dependent and matched stimulus-independent correlations differ, and that this difference can affect

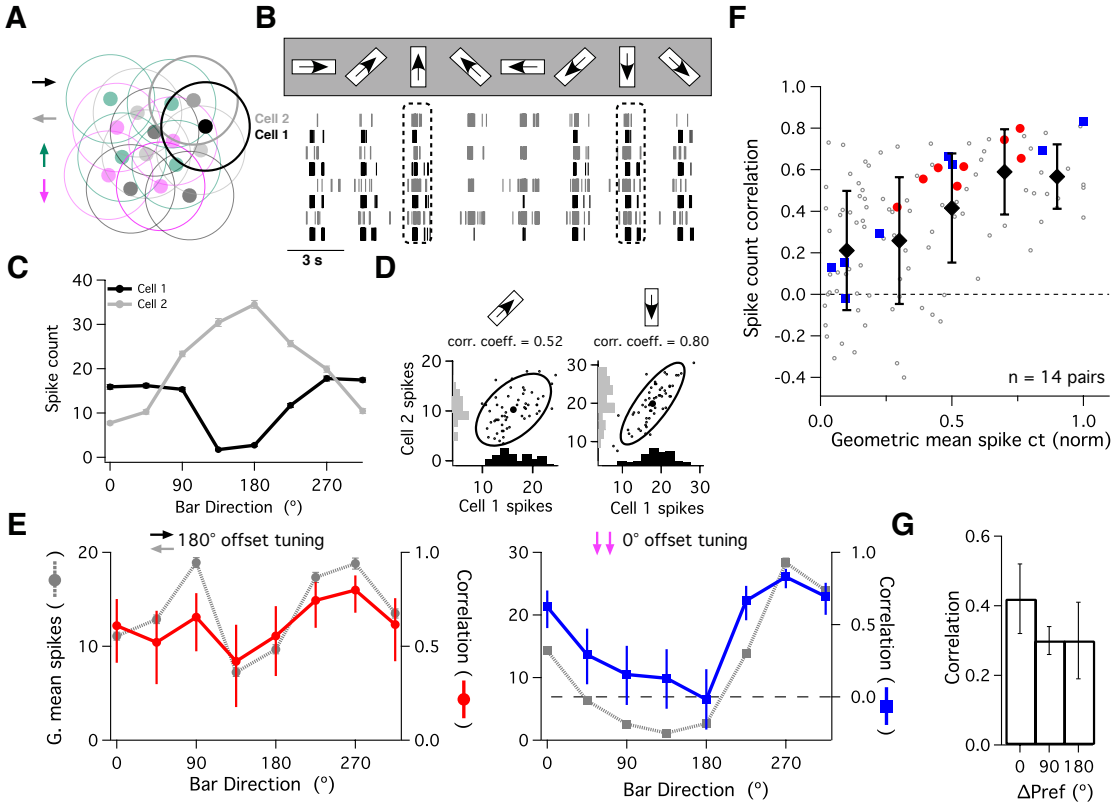

Figure 5.1: **Stimulus-dependence of ooDS cell pair responses.** (A) We performed simultaneous recordings from pairs of nearby ooDS cells with overlapping receptive fields. (B) Raster plot of spike responses from an ooDS cell pair in response to bars moving in eight different directions. See Experimental Procedures and Fig. 5.9 for details on ooDS cell identification. (C) Mean ± S.E.M. spike count as a function of bar direction ("tuning curve") for each cell in the pair - these cells have a 180° offset in their preferred directions. (D) Spike counts of the ooDS cell pair over many repeats of the same stimulus, shown for two directions of motion. Marginal distributions along axes show trial-to-trial variability in spike responses, while full response distributions show that this variability is correlated between cells (left distribution: linear correlation coefficient = 0.52, $p < 10^{-4}$; right distribution: linear correlation coefficient = 0.80, $p < 10^{-11}$). (E) We computed the geometric mean spike count for the pair by taking the square root of the product of the two tuning curves. For the cell pair in (B-D) (left panel), the noise correlation coefficient (red: error bars show 95% confidence intervals) varies with bar direction and follows the geometric mean ± S.E.M. spike count (grey, dotted line). The same is true for a pair of cells with the same preferred direction (right panel). Corresponding data for all 14 recorded cell pairs is shown in Fig. 5.10. (F) Summary of the relationship between the geometric mean responses and the noise correlation coefficient. Before computing the geometric mean responses, we first normalized the tuning curve of each cell. Each point corresponds to one stimulus presented to one cell pair (8 stimuli per pair, 14 recorded pairs). Filled symbols denote the mean ± std. dev. correlation in each bin. Note that the spike count correlation increases with geometric mean spike count (linear correlation coefficient = 0.49, $p = 5 \times 10^{-8}$). Filled, colored symbols represent the example pairs in (E). (G) Mean ± S.E.M. correlation coefficient over all eight directions for cell pairs with approximately 0, 90, and 180° offset in their preferred directions.

the fidelity of the population code. Sensory information is generally encoded by neural populations with more than two cells, and intuitions about the overall impact of correlations do not necessarily extrapolate from cell pairs to larger populations [104]. Nevertheless, the general rule still holds that fluctuations along axes orthogonal to the signal space do not limit coding fidelity.

Below, we use intracellular recordings and computational modeling to characterize the structure of both signal and noise in larger ooDS cell population responses. We then present theoretical calculations showing that, in general, stimulus-dependent correlations of the type seen in our data improve population coding in diversely-tuned neural populations.

5.3.3 Mechanisms underlying stimulus-dependent correlations

Biophysical origins of stimulus-dependent correlations

Several considerations suggest that noise correlations in the ooDS population result from common synaptic input rather than recurrent coupling. First, only one of the four ooDS cell types is gap-junction coupled to ooDS cells with the same direction tuning, while the other 3 sub-types have no such coupling [222, 223]. Second, there is no known coupling between ooDS cells of different sub-types. Hence neither 90° nor 180° pairs are coupled. Correlations for 0° pairs resembled those of 90° and 180° pairs (Fig. 5.1F), suggesting that recurrent coupling is not strongly influencing the observed correlations. Third, a feed-forward model of ooDS cell populations (Fig. 5.6) can account for the observed correlations, supporting the idea that feed-forward circuit architectures suffice for understanding them.

To elucidate the circuit mechanisms responsible for the noise correlations, we characterized the synaptic inputs to ooDS cells. We measured all four inputs to a cell pair (i.e. excitatory and inhibitory inputs to each cell) simultaneously using an alternating-voltage clamp technique [34](Fig. 5.3A-B; see Experimental Procedures). We switched the holding potential between the excitatory and inhibitory reversal potentials every 5 ms and measured the excitatory or inhibitory current at the end of each such period (Fig. 5.3A-B). Linear interpolation of these sampled currents yields a near-simultaneous estimate of both synaptic

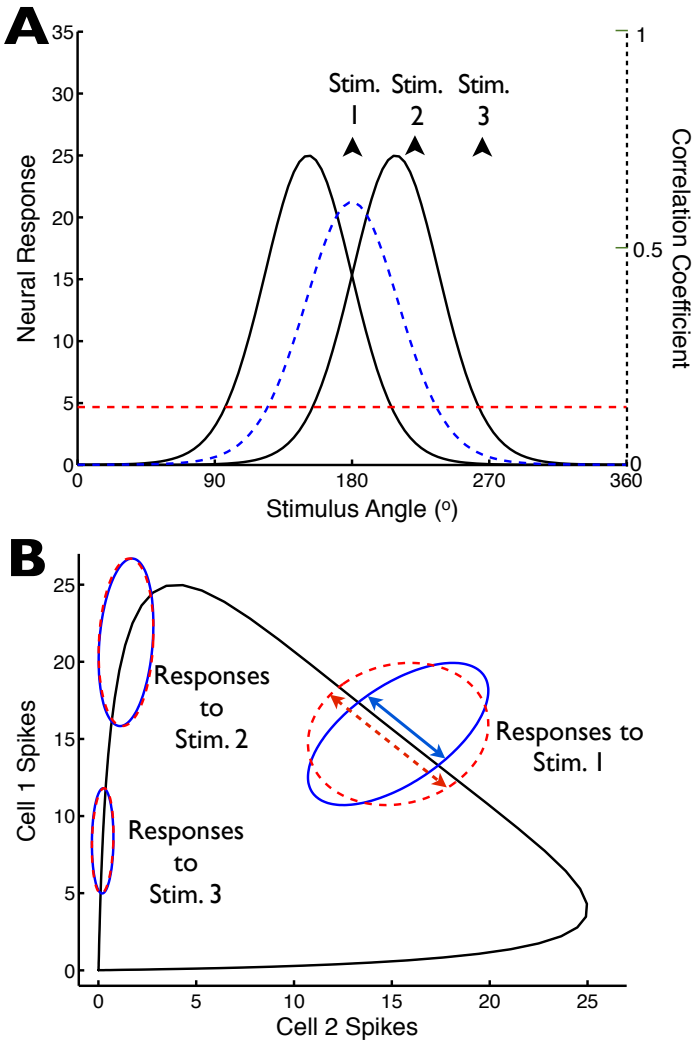

Figure 5.2: **Stimulus-dependence of noise correlations may change our perspectives on neural population coding.** (A) Tuning curves of two model neurons (solid lines), along with the noise correlation coefficient (dashed blue line), which is assumed to be proportional to the geometric mean of the neural tuning curves (as in Fig. 5.1F). For comparison, we show the average of that correlation coefficient over stimuli (horizontal dashed red line). (B) For the tuning curves in (A), we computed the pairs of mean responses elicited by all possible stimulus angles: this "signal" is indicated by the black curve. To investigate the relationship between the trial-to-trial variability and this signal curve, we assumed Poisson-like variability (variance equal to mean), and computed the stimulus-conditioned response distributions assuming either either stimulus-dependent correlations (as in (A); solid blue ellipses), or stimulus-independent correlations that matched the average of the stimulus-dependent correlations (dashed red ellipses). Ellipses show 1 standard-deviation probability contours for 3 different stimulus values (stimulus values indicated on panel (A)). By changing the projections of the stimulus-conditioned distributions onto the signal curve (indicated, for example, by the double-headed arrows for the responses to stimulus 1), the stimulus-dependence of the correlations impacts the population code.

currents (Fig. 5.11).

This approach allows measurement of all relevant means, variances, and covariances for inputs to two ooDS cells. We start by describing the means and (co)variances of inputs to single cells (Fig. 5.3). We then measure the covariances of the inputs to different cells (Fig. 5.4).

Variability in synaptic inputs: single cells

The mean excitatory (E) and inhibitory (I) conductances had opposite preferred bar directions but I inputs were more strongly and consistently tuned (Fig. 5.3C). Consequently, E and I inputs do not remain balanced as the stimulus varies.

We estimated the variability in the synaptic inputs elicited by a moving bar by subtracting an estimate of the mean response from the responses recorded on individual trials (Fig. 5.3C; see Experimental Procedures for details). The conductance residuals were used to compute the variances and covariances as a function of time delay (Fig. 5.3D). All three single cell (co)variances were modulated with bar direction - much as noise correlations in the spike outputs depended on direction (Figure 5.1D,E).

Based on the stimulus dependence of the means and (co)variances, as well as prior work on the sharing of upstream noise in the retina [3, 224], we hypothesized a circuit architecture wherein a multiplicative stimulus-dependent gain acts on separate E and I pathways that share common, noisy input (Fig. 5.3E). In the model, upstream noise diverges into parallel E and I pathways and hence is shared between them. Each pathway includes a stimulus-dependent multiplicative gain term (g_E & g_I) equal to the mean conductance in that channel for the given stimulus. Finally, additional noise (N_E & N_I), independent in the two pathways and not subject to the multiplicative gain, is added.

The model of Fig. 5.3E predicts a linear dependence of the variance of E or I inputs on the square of the respective gain, and similarly a linear dependence of the covariance of E and I inputs on the product of the E and I gains. We estimated the means of the E or I conductances during short windows about the peak response to the stimulus (see Experimental Procedures), set the gains in our model equal to these mean conductances, and

measured the peak (co)variability in the synaptic inputs during the same time window. Our data supports the predicted linear dependence of (co)variance on squared gains (Fig. 5.3F,G; Fig. 5.12). These results suggest that the ooDS cell synaptic inputs are well-described by the model of Fig. 5.3E, in which common input noise, modulated by stimulus-dependent gain factors, underlies the co-fluctuations of E and I inputs to single neurons.

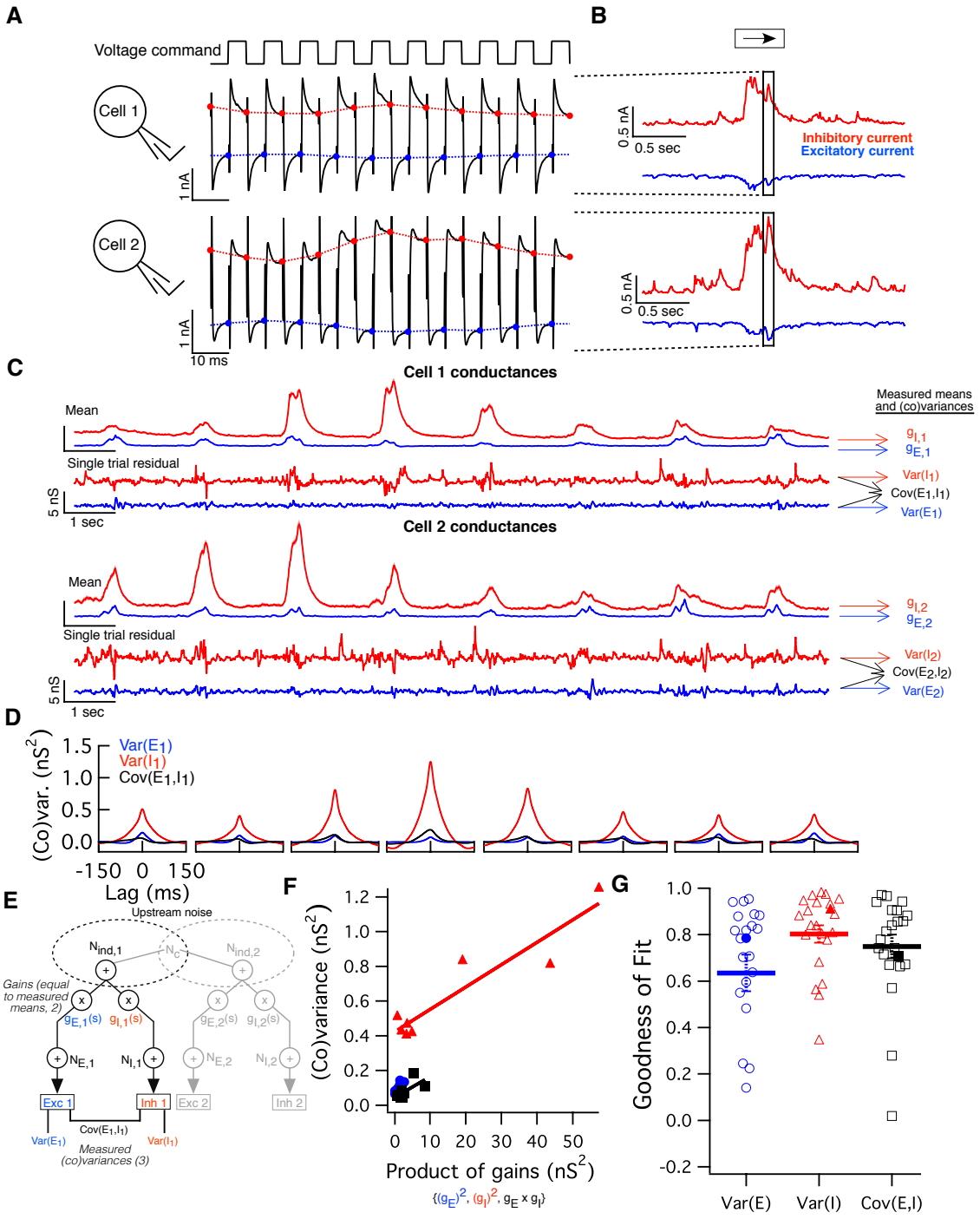
A
Voltage command
Cell 1
1 nA
Cell 2
1 nA
10 ms
B
0.5 nA
0.5 sec
Inhibitory current
Excitatory current
0.5 nA
0.5 sec
C
Cell 1 conductances
Mean
Single trial residual
5 nS
1 sec
Measured means and (co)variances
gI,1
gE,1
Var(I1)
Cov(E1,I1)
Var(E1)
Cell 2 conductances
Mean
Single trial residual
5 nS
1 sec
gI,2
gE,2
Var(I2)
Cov(E2,I2)
Var(E2)
D
(Co)var. (nS2)
1.5
1.0
0.5
0.0
Var(E1)
Var(I1)
Cov(E1,I1)
-150 0 150
Lag (ms)
E
Upstream noise
Nind,1
Nc
Nind,2
Gains (equal to measured means, 2)
gE,1(s)
gI,1(s)
gE,2(s)
gI,2(s)
NE,1
NI,1
NE,2
NI,2
Exc 1
Inh 1
Exc 2
Inh 2
Cov(E1,I1)
Var(E1)
Var(I1)
Measured (co)variances (3)
F
(Co)variance (nS2)
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0 10 20 30 40 50
Product of gains (nS2)
{(gE)2, (gI)2, gE x gI}
G
Goodness of Fit
1.0
0.8
0.6
0.4
0.2
0.0
-0.2
Var(E)
Var(I)
Cov(E,I)

Figure 5.3: (Previous page) **Alternating voltage clamp measurements reveal sources of variability in single-cell synaptic inputs.** (A) Method of paired alternating voltage clamp recordings: The voltage command (top) for both cells recorded simultaneously was alternated between the excitatory and inhibitory reversal potentials while the current (black traces) was recorded. Interpolating between the last value measured on each cycle yields an estimate of each cell's excitatory (blue) and inhibitory (red) current nearly simultaneously. (B) We simultaneously measured these four synaptic currents (excitatory and inhibitory inputs to each cell) while presenting moving bar stimuli. Highlighted region corresponds to the traces in (A). (C) Excitatory (blue) and inhibitory (red) conductances in a pair of ooDS cells measured simultaneously in response to moving bars. Mean responses (shading indicates S.E.M.) reveal tuning of input conductances. Subtracting the mean conductance traces (over nearby trials) from single-trial traces, we obtained residuals, which are shown for a single trial. (D) Using conductance residuals, we computed three single-cell (co)variances of the inputs (as a function of time delay) for each stimulus and each cell in the pair: inhibitory variance (red), excitatory variance (blue) and single-cell excitatory-inhibitory covariance (black). Shown are the (co)variance functions for cell 1 of the example pair. (E) Schematic of stimulus-dependent gain model. A portion of upstream noise is shared between cells and among excitatory and inhibitory channels. Upstream noise is multiplied by a stimulus-dependent gain factor: $g_{E,1}(s)$ or $g_{I,1}(s)$, defined as the measured mean conductance in each channel for each stimulus. Thereafter, independent, post-gain, noise is added to each channel ($N_{E,1}$ and $N_{I,1}$, respectively), yielding the noisy conductance inputs to the cell. Measured simultaneously, these inputs reveal the three possible (co)variances into a single cell. (F) Single cell peak (co)variance as a function of gain product for the responses of cell 1 above to 8 different stimuli. Solid line shows the linear best fit: data confirm the model's prediction that (co)variance should be linearly related to the relevant gain product. (G) Linear correlation coefficients between (co)variance and gain product for the three single-cell input (co)variances measured in 21 ooDS cells. Thick lines and error bars indicate mean $\pm$ S.E.M. and filled points indicate the example cell.

Related models that explicitly couple gain (equal to the mean response to a given stimulus in our case) and (co)variance fluctuations across time or stimuli have been recently explored [88, 131]. Our model is equivalent to the "multiplicative" one considered by [131], where the gain is proportional to the variance, and is similar to the model of [88], where the gain and variance are related but not strictly proportional. Our model differs from that of [64], which has no explicit relationship between gain and variance.

Variability in synaptic inputs: cell pairs

In addition to the converging (co)variances of the inputs to single cells, the paired alternating-voltage recordings characterized the four pairwise input covariances (Fig. 5.4A):

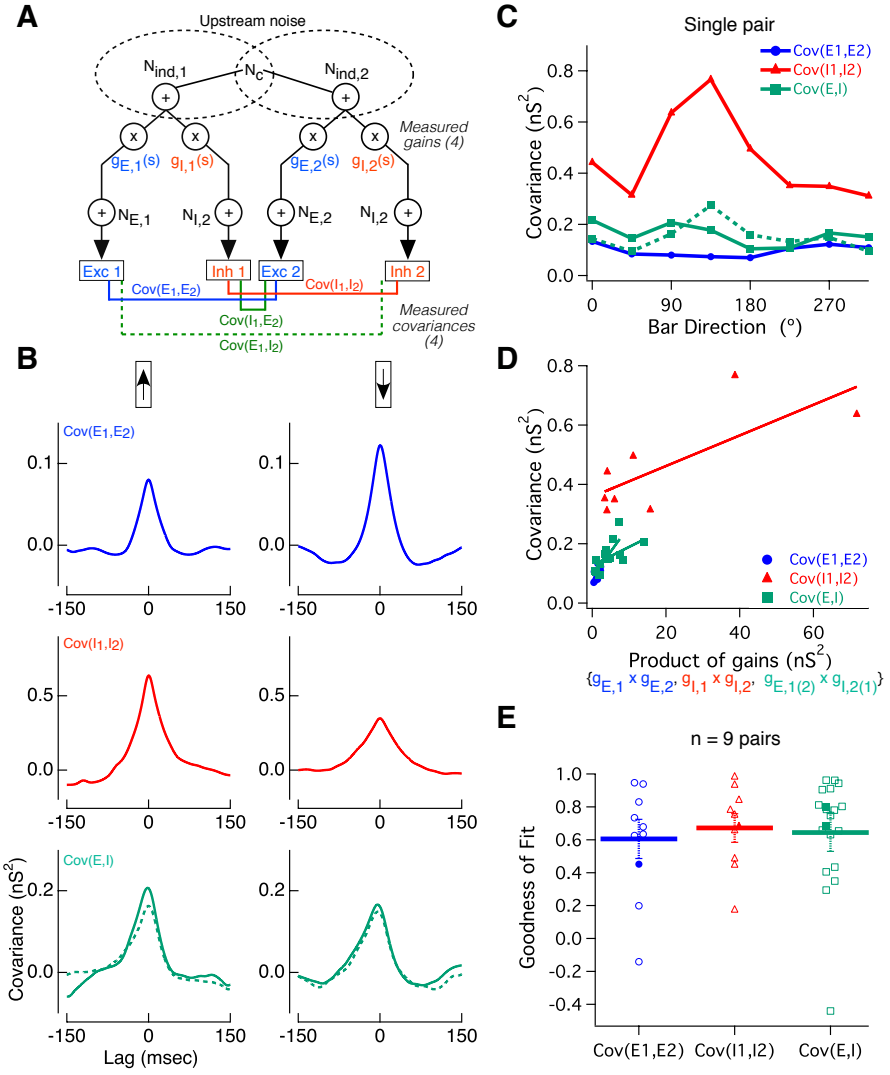

Figure 5.4: **Alternating voltage clamp experiments reveal the sources of co-variability between synaptic inputs to ooDS cell pairs.** (A) Paired recordings as in Fig. 5.3 allow measurement of all four types of pairwise input covariances: EE (blue), II (red), EI (dotted green), and IE: (solid green). The network model in Fig. 5.3 was generalized to account for pairwise input correlations. A fraction of upstream noise is shared between cells in a pair ("Common noise", N_c) (B) In response to two illustrative bar directions, the pairwise input covariance functions are shown for the example cell pair in Fig. 5.3. As with single-cell input (co)variances, pairwise covariances are modulated by bar direction. (C) Peak covariance of all four types of input correlations as a function of bar angle for the example cell pair. (D) The stimulus-dependent gain model in (A) predicts a linear relationship between pairwise input covariance and the relevant gain product - this prediction is borne out in this example cell pair. Solid lines show linear best fit. EI and IE covariances have been combined for reasons of symmetry. (E) Population data from 9 ooDS cell pairs showing linear correlation coefficients between the three pairwise covariances and their respective gain products. Thick lines and error bars indicate mean ± S.E.M. and filled points indicate the example pair.

covariance between excitatory input into cell 1 and excitatory input into cell 2 ("EE" covariance), covariance between inhibitory inputs ("II"), excitatory-inhibitory ("EI") covariance, and inhibitory-excitatory ("IE") covariance. We computed these covariances using the conductance residuals as in Fig. 5.3A. The pairwise input covariances, like the single cell covariances, showed a clear dependence on bar direction (Fig. 5.4B,C).

We generalized the network model used to capture the stimulus dependence of the inputs to single cells to cell pairs (Fig. 5.4A). This generalized model similarly predicts that the covariance of each of the pairwise inputs will depend linearly on the product of the relevant gains. This prediction is confirmed by our paired recordings (Fig. 5.4D,E): the circuit architecture in Fig. 5.4A offers a parsimonious circuit-level explanation for shared stimulus-dependent fluctuations in the inputs to neural populations; shared upstream noise, modulated by the stimulus-dependent gain, leads to co-fluctuations in the synaptic inputs experienced by ooDS cell pairs.

Relating input and output correlations

How do the different sources of input correlation collectively generate output correlations? The transfer of input correlations to output correlations depends on the relative strength of excitatory and inhibitory inputs and on nonlinearities in synaptic integration and spike generation [57] (Fig. 5.13). Of particular relevance here, the effects of EE and II input correlations (which positively correlate the cells' spiking responses) can be at least partially canceled by EI input correlations [91, 95, 133, 176, 198, 199]. Such cancellation is particularly effective for neurons operating in a balanced regime where E and I inputs are similar in magnitude, and this is often cited as a reason for the weak average correlations exhibited in cortical circuits. However, the tuning of ooDS cells relies on modulation of I relative to E (Fig. 5.4A), that in turn modulates the EI correlations (Fig. 5.4B,C). This suggests that EI correlations may not always cancel EE and II input correlations.

Linear predictions of output correlations based on the measured synaptic inputs should reveal a role of EI correlations in limiting output correlation strength. However, such linear predictions failed to capture output correlations (Fig. 5.13). Hence we used dynamic

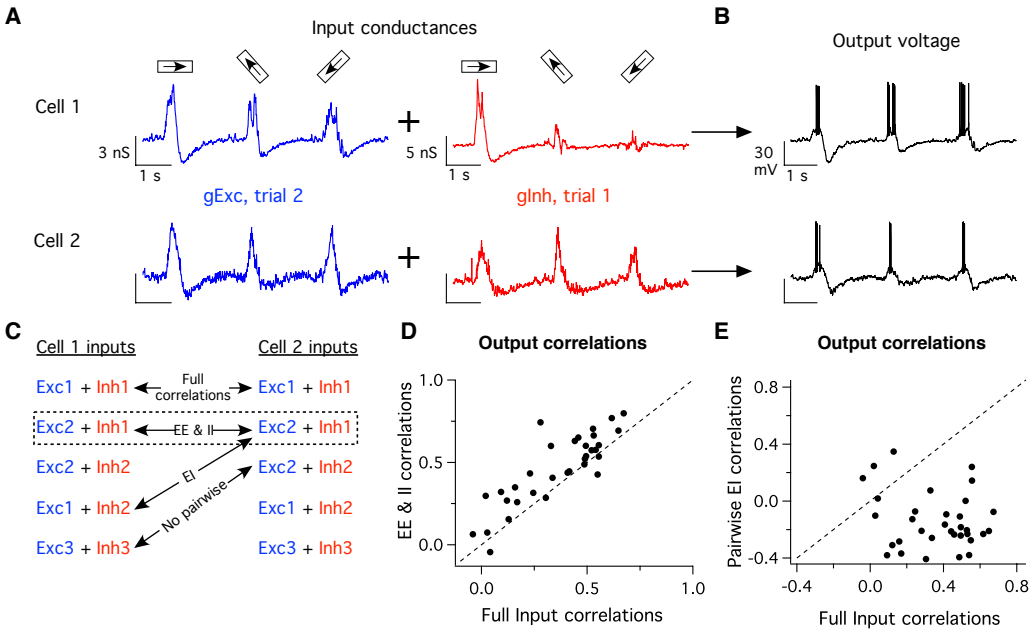

Figure 5.5: **Pairwise input correlations differentially shape output correlations.** (A) Example input conductance traces measured in alternating-voltage experiments using three different bar directions, and injected into recipient cells in dynamic-clamp experiments. By shuffling conductance combinations we can exclude or include specific sets of input correlations. For these inputs to cell 1 (top) and cell 2 (bottom) on this trial, we have taken the excitatory conductances measured on trial 2 and combined them with the inhibitory conductances measured on trial 1, leaving only pairwise EE & II input correlations intact, while breaking all other sources of input correlations. (B) Combined conductances are injected into a recipient ooDS cell, whose output voltage is measured, and spike correlations are measured as in Fig. 5.1. (C) Schematic showing the logic of shuffling input conductances to include specific sets of pairwise input correlations. (D) Correlation coefficients for dynamic clamp trials where all input correlations were left intact versus those where only pairwise EE & II correlations remained. Each point corresponds to one stimulus for one pair, shown are data from 11 cells, three stimuli each (52-88 trials, mean 72 trials). Dotted line is unity. Removing pairwise EI correlations increased output correlations ($p = 2 \times 10^{-5}$, paired t-test) (E) Same as (D) but showing trials where all input correlations remained versus those wherein only pairwise EI correlations were intact. Removal of pairwise EE & II correlations decreased output correlations (t-test, $p = 2 \times 10^{-10}$, paired t-test).

clamp techniques to determine how manipulating input correlations affected output correlations [34, 207]. The dynamic clamp approach measures a cell's voltage while injecting current calculated from the measured time-varying conductances. We manipulated the correlations in the inputs by injecting either simultaneously-recorded conductances (correlated) or non-simultaneously-recorded conductances (uncorrelated), measured in response to the same stimulus (e.g. Fig. 5.5A,B). By comparing spike responses recorded under different conditions (Fig. 5.5C), we isolated the impacts of different types of input correlations on the correlations in the cells' spiking responses. Spike count correlations were measured as in Fig. 5.1.

Spikes produced when EI correlations were removed and EE and II input correlations were retained showed higher output noise correlations than did control data in which all input correlations were left intact (Fig. 5.5D). Spike responses in which pairwise EI input correlations were retained and EE and II input correlations were removed showed near-zero or negative correlations (Fig. 5.5E). These observations support the general notion that EE and II correlations increase, whereas EI correlations reduce, correlations in spiking responses. However, the suppressive impact of EI correlations was relatively small compared to the role of EE and II correlations - the inclusion of EI correlations (Fig. 5.5D) reduced output correlations by just 34 $\pm$ 8% (mean $\pm$ S.E.M.). Thus, the impact of EI correlations may be relatively modest during stimuli that modulate the EI balance.

5.3.4 Impact of stimulus-dependent correlations on direction coding

The mechanistic understanding from the above experiments allowed us to construct a model of the ooDS cell population in which we could manipulate the noise correlations and investigate their impact on direction coding. We then generalized our study of the impact of stimulus-dependent correlations on population coding to arbitrary heterogeneous neural populations.

A mechanistic model of a ooDS cell population

Ganglion cell tiling and dendritic overlap [4] suggest that a given region of visual space

is sampled by at least eight ooDS cells (two of each of the four sub-types). We constructed a mechanistic model to understand direction coding by these 8-cell populations. The model allowed us to explore stimulus space more completely than we could experimentally and to manipulate otherwise inaccessible features like the stimulus dependence of spike correlations.

The model follows the architecture revealed by our experiments (Fig. 5.4A; Fig. 5.14). The model has 13 free parameters describing basic network and intrinsic cellular properties (the amplitudes of inhibitory and excitatory inputs, the stimulus dependence of inhibitory input, the variances of common vs. independent noise sources, etc.). We fit these parameters to match 15 different experimentally-measured quantities describing the single-cell and pairwise response statistics (Fig. 5.6A,B).

The model has realistic tuning curves and realistic levels of trial-to-trial response variability (Figs. 5.6A,C). Moreover, although the model was not directly fit to the rate-correlation relationship, (Fig. 5.1F), it recreates it with high fidelity (Fig. 5.6D). We next used the model to study how the noise correlation structure affects direction encoding.

Correlations, stimulus dependence, and direction coding by ooDS cell populations

We used the model to generate responses to 500 different directions of motion; these responses captured the experimentally observed first- and second-order spiking statistics. Using these statistics (stimulus-dependent means and covariances), we computed the linear Fisher information (see Experimental Procedures), which quantifies the population's coding ability. The Fisher information places an upper bound on the precision with which the stimulus can be recovered from the neural responses by an unbiased linear estimator [102, 171], and is a standard way to assess neural population coding [1, 104, 148, 200]. We first compared coding performance for model responses with the stimulus-dependent correlation structure with that for trial-shuffled uncorrelated data (Fig 5.6E); correlated responses provided >100% more information. Analyses of coding based on simultaneous recordings from larger populations [80] show similar improvements in coding performance.

What are the key features of the correlations that give rise to their beneficial impact on coding performance? Observed noise correlations were strongly modulated with the stimulus

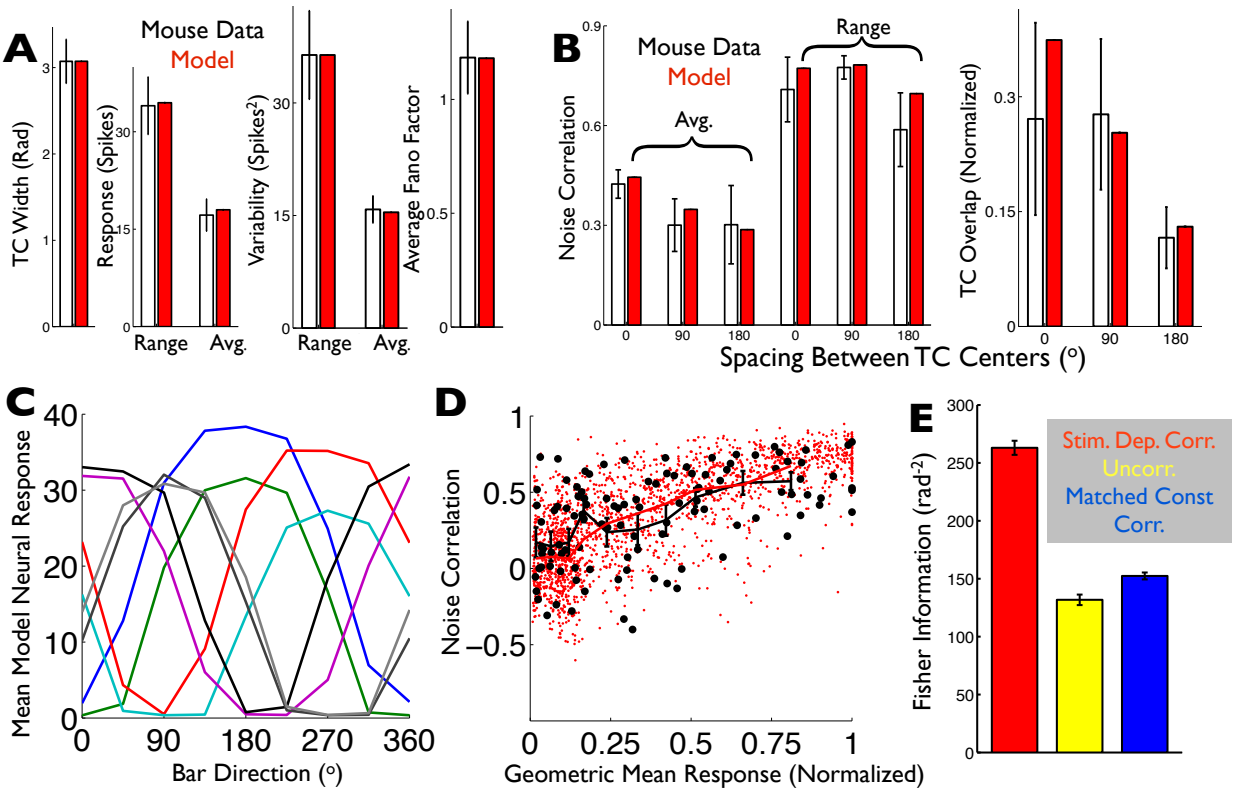

Figure 5.6: **A computational model captures the ooDS cell population response statistics and reveals that stimulus-dependent correlations significantly improve the population's direction code.** (AB) Spiking statistics (black: ooDS cell population mean ± S.E.M.) to which the model (red: average over 25 8-cell model populations) was fit. (A) Single-cell statistics. (B) Pairwise statistics. (C) Tuning curves of an example 8-cell population generated by the fitted model. (D) The rate-correlation relationship (Fig. 5.1F) was not used in training the model; it serves as an independent test. Correlation coefficient and geometric mean response for 250 model cell pairs (red dots) and 14 experimentally observed ooDS cell pairs (black circles), each in response to 8 different stimuli. The two distributions are not significantly different (2-dimensional KS test, KS statistic 0.15, $p = 0.2$). Overlain are the mean correlation coefficients in the experimental data (black symbols: mean ± S.E.M.) in different bins of geometric mean response and in the computational model (solid red curve). (E) Fisher information provided by model 8-cell population responses about the stimulus direction. Colors indicate the assumed correlation structure: stimulus-dependent correlations, as in the experimental data (red); no correlations (yellow); or "matched" stimulus-independent correlations that, for each cell pair, match the stimulus-average of their stimulus-dependent correlation coefficients (blue). Error bars in (E) are the S.E.M. over ensembles of 10 randomly-generated model populations.

(Fig. 5.1F). To investigate the role of this stimulus dependence, we generated responses in which the noise correlations for a given cell pair, for all stimuli, were maintained at the average value of the correlation for that pair (Fig. 5.6E "Matched Stim.-Indep. Corr."). Like the uncorrelated data, these responses had relatively low information. This suggests that the stimulus dependence of the correlations is an important feature of the population code, and that the average (over stimuli) level of correlation does not necessarily capture how correlations affect coding.

One limitation of the analysis above is that linear Fisher information cannot extract information encoded in, for example, the stimulus-dependent variances of the neural activities [201]. To verify the robustness of our findings (Fig. 5.6), we repeated the above investigations using different measures of coding performance. We quantified the error of the Maximum Likelihood Estimator (MLE) and the Optimal Linear Estimator [188] (OLE) of the stimulus given the neural activities. The results (Fig. 5.15) qualitatively match that which we obtained using Fisher information.

Generalization to larger and more heterogeneous neural populations

The 8-ooDS-cell model indicates that stimulus-dependent noise correlations can significantly boost the neural direction code. How do these observations apply to larger and possibly more heterogeneous populations? To answer this question, we first performed calculations of the amount of stimulus information conveyed by neural populations with identically-shaped and evenly-spaced tuning curves (Fig. 5.7A-C). For each population and stimulus, the mean responses were given by the tuning curves, and the variability was assumed to be Poisson-like (variances equal to means). Finally, we assumed that the correlations were stimulus dependent, and proportional to the geometric mean of the neural responses (as in the experimental data, Fig. 5.1F). The magnitude of the correlations was parameterized by ρmax, which determines the largest possible correlation coefficient in the population. Our general setup and computation of information in the case of stimulus-dependent correlations follows [112] (their IFmean), where explicit analytical formulae are derived. The results for these homogeneous populations (Fig. 5.7B) indicate that stimulus-dependent correlations

lead to much more information than is contained in independent populations with the same tuning curves and noise levels.

To directly test the importance of stimulus dependence, we repeated our calculations while holding the correlation coefficient for each cell pair constant at that pair's stimulus-averaged value (as in red curve in Fig. 5.2A). With these "matched" stimulus-independent correlations, coding performance is typically worse than is obtained by independent cells (Fig. 5.7C). Similar results are seen for populations with randomly-shaped and randomly-located tuning curves (Fig. 5.7D-F). For this case of stimulus-independent correlations, the correlation coefficients have a "limited-range" structure [1, 10, 43, 66, 200, 213, 247] in which the correlations between pairs of cells decrease as the cells' tuning curves become more widely spaced. These limited-range correlations are typically thought to be harmful to neural population coding, although that result can depend on the heterogeneity of the neural tuning curves, and the speed with which correlations fall off as tuning curve overlap decreases [66, 200, 202]. Our results indicate that, if the limited-range structure arises due to stimulus dependence of the correlation coefficients (as in our DS cell recordings; Fig. 5.1E,F), then those correlations can significantly improve the population code; this effect might be missed if the stimulus dependence of the correlations were ignored.

Do the coding benefits of stimulus-dependent correlations occur in populations with small levels of correlation, as often exhibited (on average) by cortical neurons [17, 43, 65, 64, 85, 173, 200]? While the largest coding benefits are obtained for large ρmax values, the average correlation coefficients are much smaller than ρmax. To highlight this, we show in the legend to Fig. 5.7B,E the average correlation coefficient (averaged over stimuli and cell pairs) obtained for each ρmax value. Even for ρmax=0.8, the average correlation coefficient is only 0.11 (for the homogeneous population), which is similar to the values typically reported in cortical recordings.

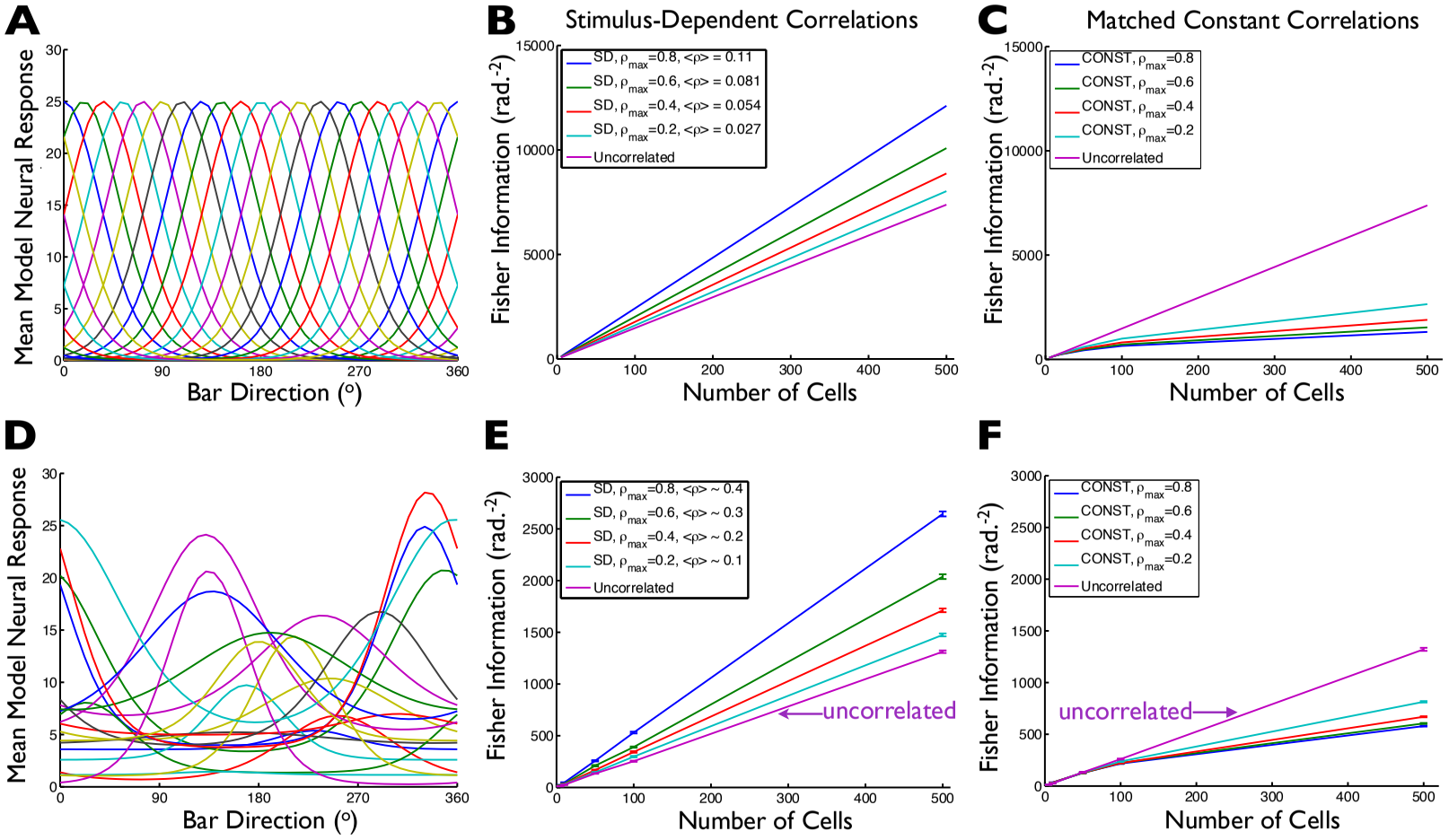
A
Mean Model Neural Response
Bar Direction (°)
B
Stimulus-Dependent Correlations
Fisher Information (rad.-2)
SD, ρmax=0.8, <ρ> = 0.11
SD, ρmax=0.6, <ρ> = 0.081
SD, ρmax=0.4, <ρ> = 0.054
SD, ρmax=0.2, <ρ> = 0.027
Uncorrelated
Number of Cells
C
Matched Constant Correlations
CONST, ρmax=0.8
CONST, ρmax=0.6
CONST, ρmax=0.4
CONST, ρmax=0.2
Uncorrelated
D
E
SD, ρmax=0.8, <ρ> ~ 0.4
SD, ρmax=0.6, <ρ> ~ 0.3
SD, ρmax=0.4, <ρ> ~ 0.2
SD, ρmax=0.2, <ρ> ~ 0.1
uncorrelated
F

Figure 5.7: (Previous page) **Stimulus-dependent correlations typically improve population coding.** We generated model neural populations of different sizes, with either identically-shaped and evenly displaced tuning curves (example population shown in panel (A)), or with randomly shaped and located tuning curves (example population shown in panel (D)). For each population, we assumed Poisson-like variability, and let the noise correlations coefficients be stimulus-dependent (being proportional to the geometric means of the cells' tuning curves), with the overall magnitude of correlations in the population being determined by the parameter ρmax (maximum possible correlation coefficient, obtained for cell pairs where both cells are firing at their maximum possible rates). For populations of different sizes, and with different ρmax values, we then computed the Fisher information those populations provide about the stimulus. On balance, stronger stimulus-dependent correlations yield better coding performance, over a wide rage of population sizes (B,E). (C,F) To understand how the stimulus-dependence of the noise correlations affects the population code, we repeated our calculations from panels (B,E), but instead of keeping the "full" stimulus-dependent correlation structure, we replaced the correlation coefficients for each cell pair with the average over stimuli of that pair's stimulus-dependent correlations: average correlations are "matched" between (B) and (C), and between (E) and (F). With stimulus-independent correlations (in stark contrast to stimulus-dependent ones), larger correlations correspond to weaker population codes (C,F). Error bars in (E,F) are the S.E.M. over ensembles of 20 randomly-chosen model populations. Average correlations values in the legends of (B,E) are averages over all stimuli and all cell pairs. The values are larger for the heterogeneous population because the heterogeneous population has, on average, wider tuning curves.

While [112] analyzed the role of stimulus-dependent correlations in neural population coding in a similar setting, that work [112] compared the stimulus-dependent correlations to the case where the correlation coefficients were stimulus independent and uniform across the population. That investigation did not resolve whether coding improvement in the stimulus-dependent case should be attributed to the stimulus dependence of the correlations, or to the diversity of correlations across the population. For contrast, we compared populations with stimulus-dependent correlations to populations with matched stimulus-independent correlations, where those correlations varied between cells so as to match the average (over stimuli) of the correlations in the stimulus-dependent case. Our observations (Fig. 5.7) thus resolve the ambiguity left by [112], and point to significantly larger effects.

In summary, our calculations show that stimulus-dependent noise correlations may be a generally beneficial feature of neural population codes - and a feature that could be missed

by averaging correlations over stimuli when reporting spike correlations.

Geometrical intuition: orthogonality between signal and noise

Why does the stimulus dependence of the noise correlations have such a striking impact on the neural population code? The impact of noise depends on its projection onto the signal space (Fig. 5.2). First, consider the signal space itself. For a population of cells with tuning curves uniformly tiling the space of stimulus direction (e.g. Fig. 5.6C, Fig. 5.7A), each stimulus direction leads to strong activation of some cells and weak activation of others. As the stimulus direction changes, the identities of the active and inactive cells change, but the overall level of activity in the population is relatively constant. Geometrically, this means that the signal space will lie on or near the surface of a (hyper-) spherical shell in the space of neural responses (Fig. 5.8A). Vectors pointing radially are orthogonal to this spherical signal space. Thus, any trial-by-trial fluctuations that are radial in response space (i.e. leading to changes in the magnitude but not direction of the response vector) will minimally impact the neural code for direction.

Increasing the magnitude of the stimulus-dependent correlations increases the fraction of the trial-to-trial variability that is in the radial direction (Fig. 5.8B,C). This effect can be observed in the pairwise model example in Fig. 5.2 and can be seen in our recorded ooDS pair population (Fig. 5.16). Thus, stimulus-dependent correlations enhance the population code by shaping the noise in the population responses such that it is orthogonal to the signal space.

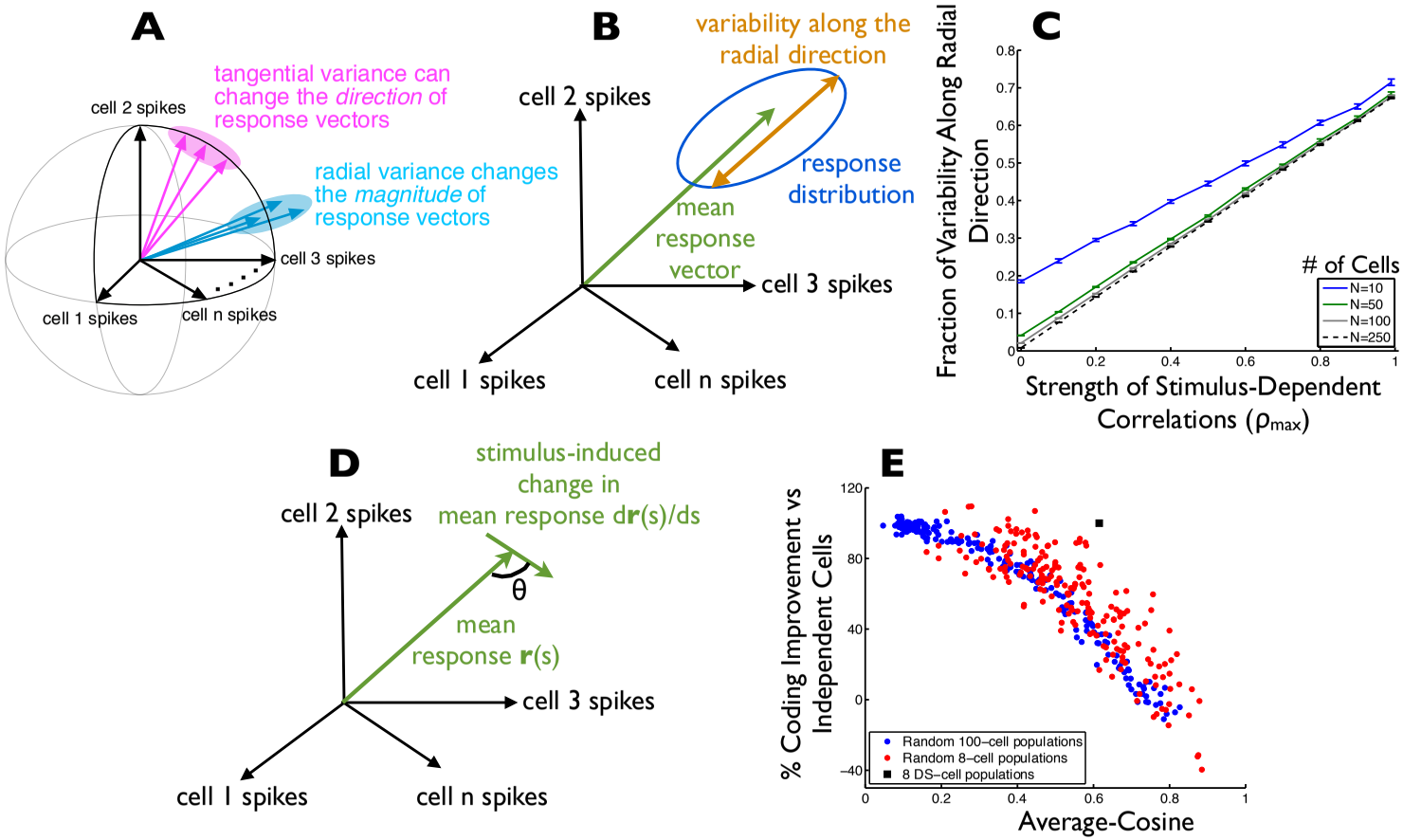
A
tangential variance can change the direction of response vectors
radial variance changes the magnitude of response vectors
cell 2 spikes
cell 3 spikes
cell 1 spikes
cell n spikes
B
variability along the radial direction
response distribution
mean response vector
cell 2 spikes
cell 3 spikes
cell 1 spikes
cell n spikes
C
Fraction of Variability Along Radial Direction
Strength of Stimulus-Dependent Correlations (ρmax)
# of Cells
N=10
N=50
N=100
N=250
D
stimulus-induced change in mean response dr(s)/ds
θ
mean response r(s)
cell 2 spikes
cell 3 spikes
cell 1 spikes
cell n spikes
E
% Coding Improvement vs Independent Cells
Average-Cosine
Random 100–cell populations
Random 8–cell populations
8 DS–cell populations

Figure 5.8: (Previous page) **Why (and when) do stimulus-dependent correlations improve neural population codes?** (A) The space of possible neural activities is shown. Each axis in this space is the firing rate of a single cell. Within that space, each stimulus elicits a distribution of responses (ellipses show 1 standard-deviation probability contours). For populations (like those in Fig. 5.7A,D and 5.6C) whose tuning curves tile the response space, the stimulus-averaged population responses (and thus the "signal" curve) tends to lie near a spherical shell in the space of possible neural activities. In this case, trial-to-trial variability in the radial direction is relatively benign to the population code, whereas variability in the tangential direction (along the signal curve) is especially pernicious. (B) The mean response to a given stimulus is indicated by the green arrow. The amount of variability along the direction indicated by the mean response vector ("radial direction") is shown in orange. (C) For heterogeneously tuned populations (as in Fig. 5.7D), we computed the fraction of overall variability in the population responses that lies along the radial direction as a function of the strength of the stimulus-dependent correlations (ρmax). The calculation was repeated for several different population sizes; error bars are $\pm$ 1 S.E.M. over 20 randomly-drawn populations. (D) To predict which populations will show coding benefits from stimulus-dependent correlations, we measured the cosine of the angle (θ) between the stimulus-induced changes in neural population responses, and the mean (trial-averaged) population response vector. (E) We considered 200 randomly-generated populations of 100 cells each (blue data points). For each population, the stimulus-averaged cosine value (defined in D) is shown (horizontal axis). For the same neural populations, we also computed the Fisher information, either in the presence of stimulus-dependent correlations ($\rho max = 0.8$) or with uncorrelated noise, and computed the percentage by which the population code formed in the presence of stimulus-dependent correlated noise exceeds the performance of the population code formed in the presence of independent noise. For comparison, we repeated this calculation with 8-cell populations (red data points), and we show the corresponding values for our 8 ooDS-cell model populations (black square; model is described in Fig. 5.6; data shown is averaged over 10 model ooDS cell populations, of 8 cells each). The ooDS cell population appears as an outlier to the point clouds in (E) because the ooDS cell populations have a larger ρmax value (of roughly 1; see Fig. 5.6D) than do the other populations shown here.

Generality for populations with heterogeneous tuning curves

The geometrical picture discussed above identifies the conditions under which stimulus-dependent correlations will improve the population code: the benefits of stimulus-dependent correlations arise whenever the "signal" curve is orthogonal to the radial direction (i.e. when signals lie along a spherical shell). To test this intuition, we developed a metric to measure the extent to which the signals are orthogonal to the radial direction and tested this metric across several neural populations with differing degrees of tuning curve regularity.

The metric we used measures the cosine of the angle between the mean response vector

for a particular stimulus and the stimulus-induced changes in the mean response vector (Fig. 5.8D). This cosine is averaged over all stimuli to yield our metric. When the signal direction is typically orthogonal to the radial direction, the average-cosine metric is near 0 (cosine of 90°), and stimulus-dependent noise correlations are expected to enhance the population code. When the signals lie predominantly along the radial direction in the response space, the average cosine metric will be near 1 and the stimulus-dependent correlations may be damaging to the population code.

We assembled many neural populations with randomly-drawn tuning curves (as in Fig. 5.7C). For each population, we computed the Fisher Information (as a measure of the ability of the population to encode stimulus direction) either in the presence of stimulus-dependent noise correlations (with $\rho max = 0.8$), or in the presence of uncorrelated noise. For each population, we computed the percentage improvement in population coding performance that is obtained in the presence of stimulus-dependent correlations vs uncorrelated noise. By comparing the coding benefit of stimulus-dependent correlations (vertical axis of Fig. 5.8E) to the average-cosine metric (horizontal axis of Fig. 5.8E) for these populations, we see that the average-cosine metric is a good indicator of when stimulus-dependent correlations improve neural population codes: low average-cosine values correspond to large coding benefits from stimulus-dependent correlations. We repeated our calculations on populations of both 8 and 100 cells, and observed the same trends. This serves to confirm the geometric intuition (Fig. 5.8A) underlying the coding benefits of stimulus-dependent correlations.

We emphasize that, in generating Fig. 5.8E, we took extra care to draw model populations that had a wide range of average-cosine values. To do this, we forced the tuning curves to be clustered together (not fully spanning the stimulus space) to varying degrees. If we had let the tuning curves be randomly positioned within the stimulus space (as in Fig. 5.7D), we would have observed predominantly low values of the average-cosine metric (not shown): when the tuning curves randomly tile the stimulus space, the overall population activity level varies relatively little as the stimulus changes. This effect is magnified in larger populations. Thus, in large neural populations, the geometrical picture presented in Fig.

5.8A is more likely to apply, and stimulus-dependent noise correlations are more likely to improve population coding.

5.4 Discussion

Studies of population coding in retinal ooDS ganglion cells provide a unique opportunity to investigate the extent, origins and impacts of correlated variability because the relevant circuit mechanisms can be studied directly in the context of physiological stimuli and because the space of relevant stimulus parameters encoded by these cells is clear. Our central finding is that the circuitry shaping the synaptic inputs to ooDS cells generates stimulus-dependent correlations in the cells' trial-to-trial variability; these correlations enhance the precision of direction coding roughly 2-fold compared to the case of independent noise (see also [80]).

Spike outputs of ooDS cells exhibit positive and strongly stimulus-dependent noise correlations (Fig. 5.1; see also [5, 80]). The seemingly complex correlation structure falls naturally out of a direction-tuned circuit with shared upstream noise (Figs. 5.3 & 5.4). Thus divergence of noisy signals produces correlated fluctuations in downstream neurons; stimulus dependence of those correlations arises due to circuit nonlinearities that control the gain of shared noise. Such circuit mechanisms are common, and other neural circuits are likely to share the beneficial shaping of noise described here. Indeed, recent work shows that noise correlations in visual cortex resemble those of the ooDS cells and that the resulting shaping of noise can benefit coding [131].

In agreement with previous work [91, 176], correlations between excitatory and inhibitory synaptic inputs decorrelated spike outputs (Fig. 5.5D,E). However, this decorrelating mechanism did not prevent large output correlations in ooDS cells because excitatory and inhibitory synaptic inputs are not always balanced (Fig. 5.3). Indeed, retinal direction selectivity is known to rely on a strongly modulated ratio of excitatory to inhibitory input [83, 218]. Temporal EI imbalance has similarly been shown to shape stimulus selectivity in many areas of the brain [73, 241], and an EI imbalance has been suggested to play a role in gating noise correlations in cortex [95].

The observed stimulus-dependent correlations enhance the ability of the neural population to convey direction information, and these findings generalize to large heterogeneous populations (Figs. 5.7B,E; also see [112]). Moreover, the strong population coding effects we see in the retinal ooDS cell population emerged only when the stimulus dependence of the correlations was accounted for. Thus, correlations that look (on average) to be small and inconsequential for coding may in fact be important features of the population code.

One clear way for noise correlations to improve population coding is if, for all cell pairs, the noise correlations have opposite signs relative to the signal correlations [9, 104, 109, 156, 158] ("sign rule"). This phenomenon is not responsible for the strong boost in coding performance observed here (Fig. 5.6 & 5.7). For example, in the 8-cell population, with two cells of each sub-type, there will be four cell pairs in which both cells have the same direction tuning (like the pair shown in Fig. 5.1B-E), and correspondingly positive signal correlations. For these cell pairs, the "sign rule" indicates that negative noise correlations would improve the population code. However, we typically observed positive noise correlations for these cell pairs. Thus, the sign rule is strongly violated for some cell pairs, yet noise correlations still improved the population code on balance (Fig. 5.6E).

What matters most in understanding how correlations impact population coding in cases like these is the orientation of signal and noise within the larger population, not within cell pairs [1, 104, 200, 202, 50]. In other words, the coding effect occurs at the population level, and cannot always be understood by looking at cell pairs in isolation. Indeed, while some pairs showed reductions in coding precision when noise correlations were removed (Fig. 5.17, cf. schematic in Fig. 5.2), the effect for pairs of cells was modest. Related theoretical work shows that the sign rule identifies only one of the myriad patterns of correlation that can substantially improve population coding [80, 104, 200, 50] relative to the case of independent cells. The current study provides a clear example of how populations in which some pairwise responses violate the sign rule may nevertheless have noise correlations that are beneficial to the population code.

The structure of the signal space depends on what is being encoded. Here, we considered

the encoding of direction of motion and found that the stimulus-dependent correlations were quite beneficial. This is not generally true for all stimulus parameters. For encoding contrast information (where the tuning of all ooDS cells is similar), correlations may well be harmful. This emphasizes that correlations do not remove noise from the population responses: rather, they orient noise along axes that may be beneficial for some computations, but potentially harmful to others.

The circuit mechanisms that underlie the observed stimulus-dependent noise correlations – diverging, stimulus-tuned excitatory and inhibitory inputs to multiple cells – are not unique to the retina. Moreover, many sensory populations are composed of neurons with heterogeneous tuning properties that together span the relevant parameter space. These facts suggest that the relation between the circuit mechanisms shaping noise correlations and their impact on coding that we explored here may have correlates in other circuits in the nervous system.

5.5 Experimental Procedures

5.5.1 Tissue preparation

We used 6-10 week old mice (c57/BL6 or TRHR-GFP [179] or BDxTSY [118]). These mice express GFP in ooDS cells, facilitating cell identification. BD mice were injected with tamoxifen (100 μg, Sigma) intraperitoneally at postnatal day 0-1. The mice were dark adapted at least 2 hours. Under IR illumination, animals were euthanized by cervical dislocation and the eyes were removed and hemisected. The vitreous humor was removed mechanically and eye cups were stored in a light-tight container with 32°C bicarbonate-buffered Ames' solution (Sigma), continuously bubbled with 95% O2/5% CO2. For recording, wedges of retina were removed from the sclera and retinal pigment epithelium and mounted photoreceptor-side down onto a poly-D-lysine coated glass coverslip (BD Biosciences). Retinal mounts were perfused with 30-36°C oxygenated Ames' at 6-9 mL/min.

5.5.2 Visual stimulation

Visual stimuli were designed using MATLAB (MathWorks) and the Psychophysics Toolbox and delivered through an OLED monitor (eMagin, Bellevue, WA) focused on the photoreceptor outer segments. All light-driven responses were recorded at a background generating approximately 50 rhodopsin isomerizations/rod/second. To probe direction selectivity of ooDS cell pairs (Figs 5.1, 5.3, 5.4, 5.5), stimuli were centered midway between the two somas and 100% contrast bars were swept through their receptive fields. Bars were 120-180 μm wide (perpendicular to the axis of motion) and moved $\sim$850 μm/s on the retina (except two cell pairs: 575 μm/s).

5.5.3 Cell identification and selection

In recordings made in wildtype retina, ooDS RGCs were identified by their soma size and shape, low maintained firing rate, transient ON and OFF responses to a light step, and direction selective spike responses to moving bars stimuli (see Fig 5.9).

In recordings made in TRHR-GFP and BDxTSY retinas, where two-photon targeting also aided ooDS cell identification, laser exposure was kept to a minimum, typically $<$ 1 minute exposure, $<$ 2 mW post-objective laser power at 960 nm. Recording from ON alpha RGCs at the light levels for these experiments ($\sim$50 R*) before and after typical laser exposure suggested that contrast sensitivity is not substantially affected by this targeting. For loose-patch (cell attached) spike recordings, only retinal mounts that met certain health and sensitivity requirements were used. Tissue health was judged by visual inspection, and contrast sensitivity was measured using loose patch recordings from ON alpha RGCs, which, in sufficiently sensitive tissue, produced robust spike responses to 10% contrast spots centered over their receptive fields.

5.5.4 Alternating voltage-clamp recording

DS RGC pairs and single cell conductances were measured using a previously described technique [34] that allows simultaneous measurement of both excitatory and inhibitory synaptic conductances (see Fig. 5.11). Cells were whole-cell patched and voltage-clamped using a MultiClamp 700B amplifier (Molecular Devices) compensated at 50%. Series resistances were typically between 10 and 20 MΩ. Pipettes for voltage-clamp recordings were filled with a Cs-based internal solution (105 mM CsCH3SO3, 10 mM TEA-Cl, 20 mM HEPES, 10 mM EGTA, 5mM Mg-ATP, 0.5mM Tris-GTP and 2mM QX-314, pH 7.3, 280mOsm). Excitatory and inhibitory reversal potentials were assessed empirically during a full-field light step. Holding potentials used were chosen to be close to the assessed reversal potentials while remaining conservatively between the two, in order avoid noise correlation measurement error, as previously described. Sets of moving bars were then presented in three blocks while synaptic currents were recorded. The cell was first voltage-clamped near the inhibitory reversal potential for one trial, and then near the excitatory reversal potential, and then the holding potential was alternated every 5 ms between the excitatory and inhibitory reversal potentials. These blocks of three were repeated for 10-45 trials.

Excitatory and inhibitory synaptic currents were extracted from the current recorded during the alternating voltage trials by interpolating between the last sample point of current recorded before the next voltage change. This process effectively measures excitatory and inhibitory synaptic currents simultaneously at a sample rate of 100 Hz and is able to record the majority of both signal and noise.

Conductances were calculated from both non-simultaneous currents, recorded during the single hold voltage trials, and simultaneous currents, extracted from the alternating voltage trials. Conductances were calculated by dividing the measured excitatory and inhibitory currents by an assumed driving force, -62 mV and 62 mV respectively. The average excitatory and inhibitory synaptic conductances measured during the single hold trials are very similar to those measured during the alternating voltage trials, supporting previous results.

5.5.5 Dynamic clamp experiments

Synaptic input was mimicked and manipulated during dynamic clamp experiments [207]. In brief, cells were whole-cell patched and current-clamped. Pipettes for dynamic-clamp experiments were filled with a K-based internal solution (110mM K aspartate, 1mM MgCl, 10mM HEPES, 5mM NMDG, 0.5mM CaCl2, 10mM phosphocreatine, 4mM Mg-ATP and 0.5mM Tris-GTP, pH 7.2, 280mosM). Current injected into a cell (I) during dynamic clamp experiments was calculated using equation (1) below.

$$I(t) = G_{exc} \times (V(t - \Delta t) - E_{Exc}) + G_{inh}(t) \times (V(t - \Delta t) - E_{inh}) \quad (1)$$

Where G_{exc} and G_{inh} are a set of conductances recorded during light stimulation, V is the cell's membrane potential, and Eexc and Einh are reversal potentials set at 0 mV and -60mV respectively. The exact inhibitory reversal potential did not substantially impact the highlighted results. The absolute amplitude of the conductances were adjusted for each cell to achieve spike numbers near those recorded from the DSGCs from which the conductances were recorded and a constant current was often injected to help the cell maintain the appropriate resting potential.

5.5.6 Data analysis

Custom analysis scripts were written in MATLAB. In cell-attached recordings, we measured spike counts while the bar was swept over the receptive field. The response window was 2-3 seconds long. Because ooDS cells have a very low maintained (spontaneous) spike rate, our spike counts typically reflect the entirety of the response to each stimulus presentation.

Direction selectivity was quantified as follows. For each stimulus angle θ, we defined the two-dimensional direction vector $\mathrm{d}(\theta) = (\cos(\theta),\sin(\theta))$. We then multiplied these direction vectors by the mean firing rate induced in the by that stimulus, to get the vector $\mathrm{v}(\theta) = \mathrm{r}(\theta)\mathrm{d}(\theta)$. The length of this vector is the cell's mean firing rate (for this particular stimulus), and the direction indicates the direction of stimulus motion. We averaged these vectors over stimuli and computed the norm of the result. If the cell fired equally to stimuli in all

directions, the horizontal and vertical components of the v(θ) vectors would cancel, leading to the average vector having small magnitude. Finally, we scaled the mean vector by the cell's overall average firing rate (averaged over all stimuli). This yields a number between 0 (cell responds equally to stimuli in all directions) and 1 (cell responds only to motion in 1 direction), that we called the Direction Selectivity Index (DSI).

$$DSI = \langle v(\theta)\rangle_\theta / \langle r(\theta)\rangle_\theta \quad (2)$$

Over the course of paired recordings, both spike and conductance recordings are subject to slow, long-term drift that, if left uncorrected, can introduce artifactual noise correlation measurements. To correct for long term drift in cell attached recordings (Fig. 5.1), we subtracted the window (radius 2 trials) averaged spike count from each trial. By inspection of spike counts across trials and the trial autocorrelation function [17], this window-correction method effectively removed long-term drift from reported data. Qualitative results agree using uncorrected spike counts, window-corrected data (reported in main paper), and Bair et al.'s trial cross-covariance correction method [17]. Similarly, for synaptic conductance (co)variance measurements, we used window-averaged conductance (radius 1 trial) for calculation of the residuals.

The simultaneously-measured synaptic currents from our alternating voltage experiments were used to estimate the synaptic gains and (co)variances in our network model (Fig. 5.3 & 5.4). For each input and each stimulus, we measured the mean conductance during the window of time that the mean conductance exceeded a value of 25% of its maximum for that stimulus. This ensured that we isolated the light-driven component of the responses. The mean conductance was used as the estimate of the gain for that input (in response to that stimulus), in our network model. We computed each (co)variance as a function of time delay by taking the cross-correlation, using MATLAB's built-in xcorr function, between the appropriate residuals at each trial, and averaging across trials. The estimated (co)variance measure for our network model was defined as the peak of this (co)variance function.

The strength of the linear relationship between the appropriate gain product and (co)variance measurement (see Fig. 5.3 & 5.4), depended on the modulation of the (co)variance by the

stimulus (Fig. 5.12). Fit quality was also affected by limited numbers of trials in these technically challenging paired alternating-voltage experiments: when only subsets of the total number of trials were used to fit the linear model, the fit qualities for all pairwise covariances declined (data not shown).

5.5.7 Information Calculations

To compute the information content of the model ooDS cell populations (Fig. 5.6), and the larger model populations (Fig. 5.7 & 5.8), we used the linear Fisher information. This is a fairly standard way to quantify neural population codes [10, 25, 104], and it determines the precision with which the neural activities can be decoded to recover the stimulus. Mathematically, the Fisher information uses the vector of tuning curves (where, for a given stimulus angle θ, each element of the vector is one neuron's mean firing rate), and the covariance matrix $\Sigma(\theta)$ of the trial-to-trial variability in responses to stimulus θ. The Fisher information $I(\theta)$ is then calculated as

$$I(\theta) = \frac{d\overline{f(\theta)}}{d\theta}^T \Sigma(\theta)^{-1} \frac{d\overline{f(\theta)}}{d\theta} \quad (3)$$

For our information calculations, $I(\theta)$ was computed for many different stimuli (500 for the data in Fig. 5.6; 100 for the data in Fig. 5.7) uniformly spanning the range of [0, 2π], and the reported information quantities are averages over all such stimuli.

5.5.8 Model Containing Stimulus-Dependent Correlations

For the data shown in Fig. 5.7, we generated the neural tuning curves via Von Mises functions [66, 104], and assumed Poisson variability (vector of spike count variance equal to vector of mean spike counts given by tuning curves f(θ)). The correlation coefficient between cells i and j was assumed to follow the functional form

$$\rho_{ij} = a_{ij}\sqrt{f_i(\theta)f_j(\theta)},$$
$$a_{ij} = \rho_{max}[\sqrt{max_\theta(f_i(\theta))max_\theta(f_j(\theta))}]^{-1} \quad (4)$$

where the $max_\theta(.)$ operation selects the peak amplitude of the tuning curve. This function ensures that the correlation coefficient varies between 0 and 1, and increases with increasing

geometric mean firing rate (as in our experimental data).

5.6 Acknowledgments

We thank Bruno Averbeck, Braden Brinkman, Mike DeWeese, Brent Doiron, Charlie Hass, Kresimir Josic, Steven Lisberger, Gabe Murphy, Michael Shadlen, and Ben Strowbridge for helpful discussions, insights, and suggestions. Excellent technical assistance was provided by Paul Newman, Shellee Cunnington, and Mark Cafaro. Support provided by NSF CRCNS grant 1208027 (ESB and FR), a Simons Fellowship in Mathematics (ESB), NEI grants EY07031 (MHT) and EY11850 (FR), and HHMI (FR).

5.7 Supporting Information

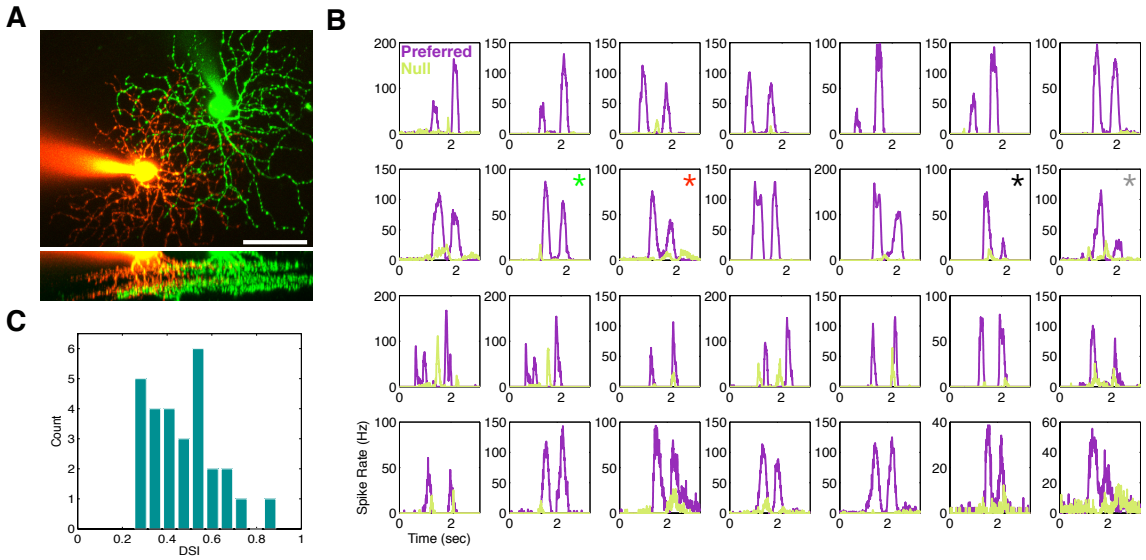

Figure 5.9: **Identification of ON-OFF dsRGCs.** (A) Confocal image of a pair of ON-OFF Direction Selective RGCs that were filled with dye using a patch pipette and imaged after a paired loose-patch spike recording. Top: maximum intensity projection image showing dendritic overlap. Bottom: cross-section maximum intensity projection showing characteristic bistratified dendritic arbors. (B) For each cell in the cell-attached dataset, average PSTH traces show ON-OFF spike responses. Responses are shown to bars moving in the preferred (purple) and null (yellow) directions. Panels marked with an (*) indicate example cells: red and green correspond to the dye-filled pair in (A), which is the 0°pair in Fig. 5.1D; black and grey correspond to the example cells from the 180°pair in Fig. 5.1B-E. (C) Histogram of vector-sum DSI values (see Experimental Procedures) for all 28 cells in the cell-attached dataset.

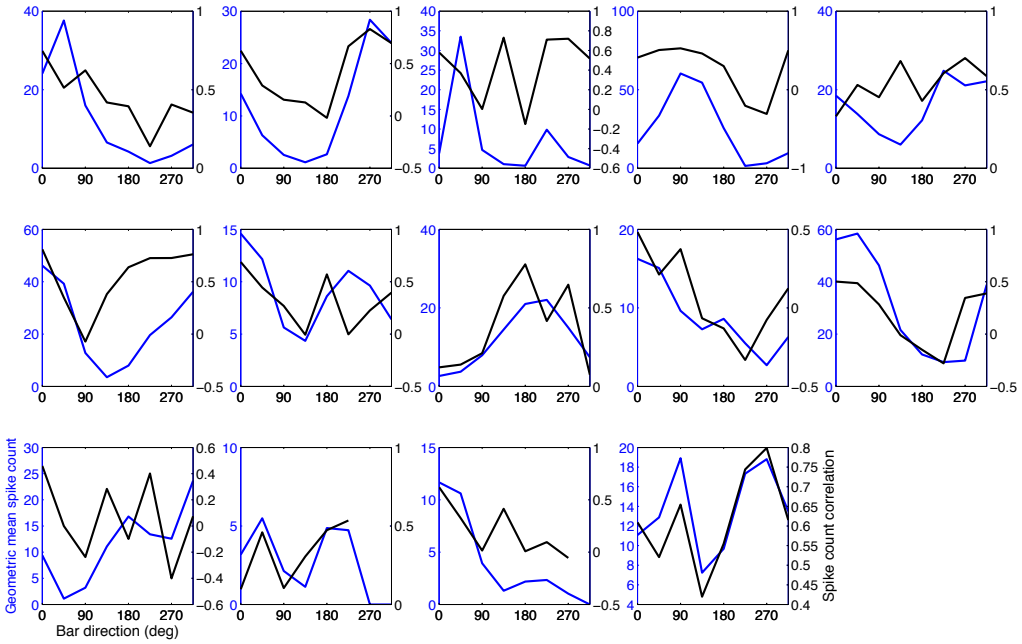

Figure 5.10: **Direction dependence of responses and correlations for all recorded cell pairs.** Each panel corresponds to one cell-attached recorded ooDS cell pair. Blue traces are the geometric mean of the mean spike counts and black traces are the spike count noise correlation coefficients, each as a function of bar angle (as in Fig. 5.1E). Panels are organized by relative preferred direction. Top row shows pairs (n = 5 pairs total) with roughly the same direction tuning ("0°pairs"); middle row shows 90°pairs (n = 5), and bottom row shows 180°pairs (n = 4).

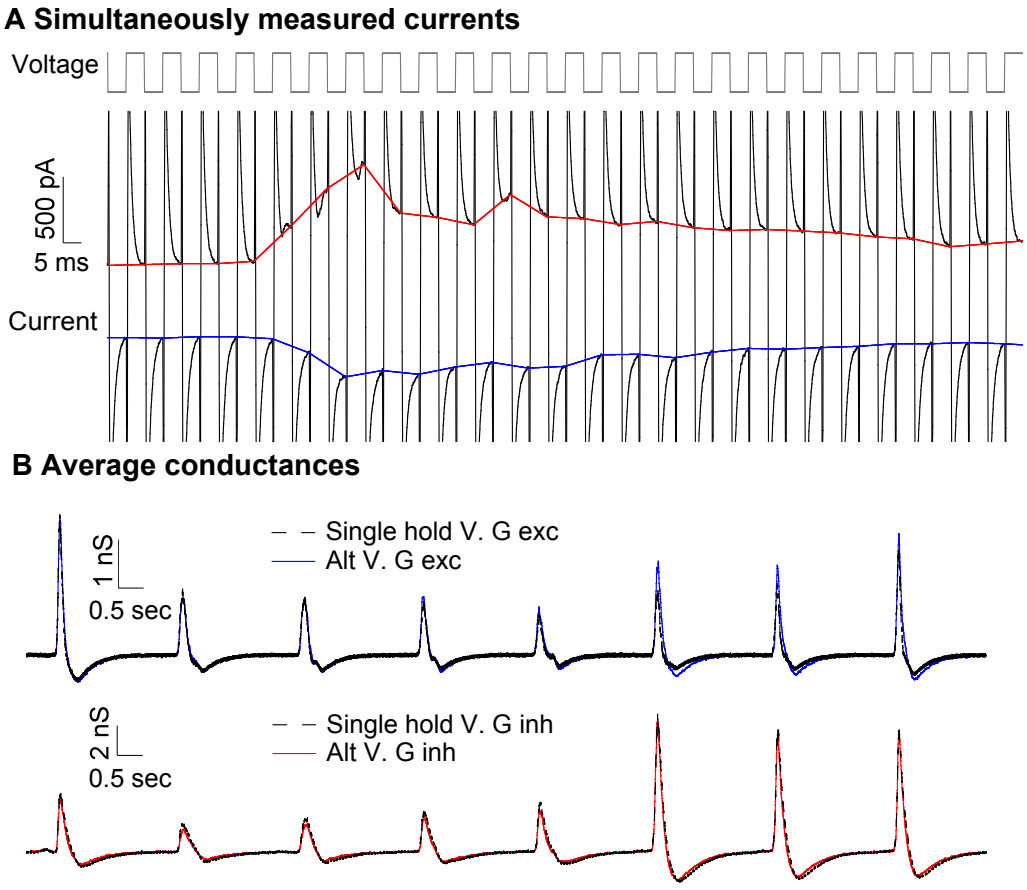

Figure 5.11: **Alternating voltage-clamp conductance measure.** (A) The voltage command (top) was alternated between the excitatory and inhibitory reversal potentials while the current (gray line) was measured. Interpolating between the last current values measured on each cycle (colored lines) provides an estimate of the neurons excitatory and inhibitory currents. (B) The average conductances calculated using this procedure were very similar to the average conductances measured using a single holding potential at both the excitatory and inhibitory reversal potentials.

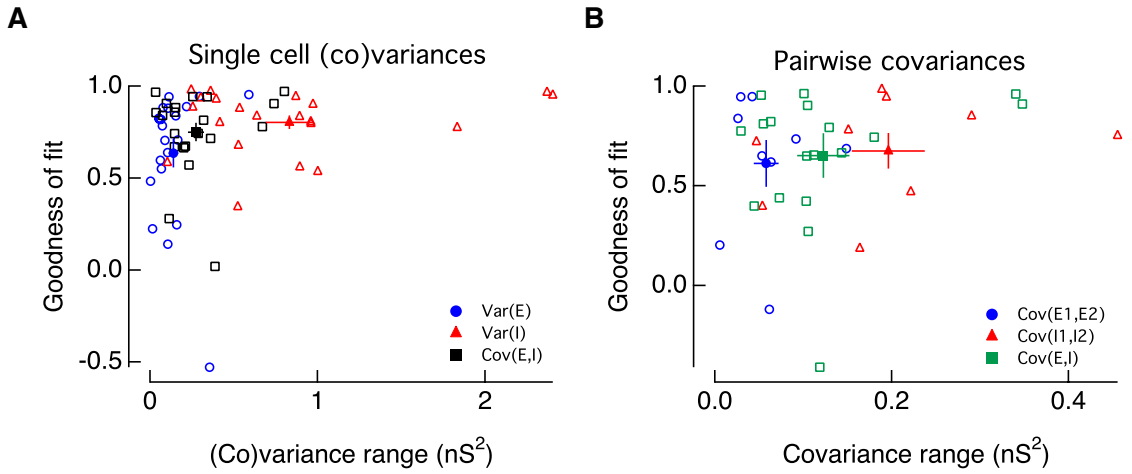

Figure 5.12: **Gain model goodness of fit as a function of (co)variance range.** The linear correlation between measured conductance gain and (co)variance is shown as a function of the (co)variance range for each single cell (A) and pair (B) in the alternating voltage clamp experiments. The linear correlation is stronger for conductances that are more strongly modulated by the stimulus. Error bars represent the S.E.M.

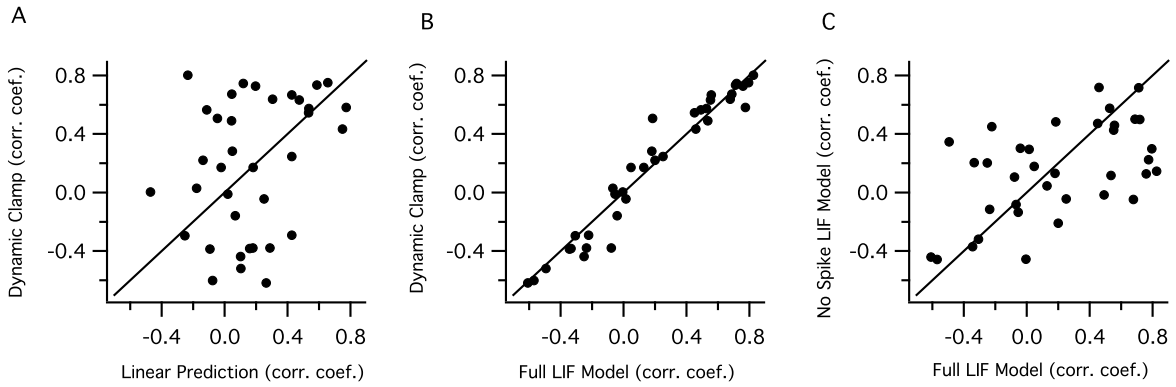

Figure 5.13: **Impact of spike generation on spike count correlations.** (A) The linear prediction (see below) failed to predict the dynamic clamp output spike count correlations. (B) A simple LIF model was highly predictive of the dynamic clamp output spike count correlations. (C) Preventing spike generation in the LIF model caused both increases and decreases in the spike count correlation compared with the spiking model.

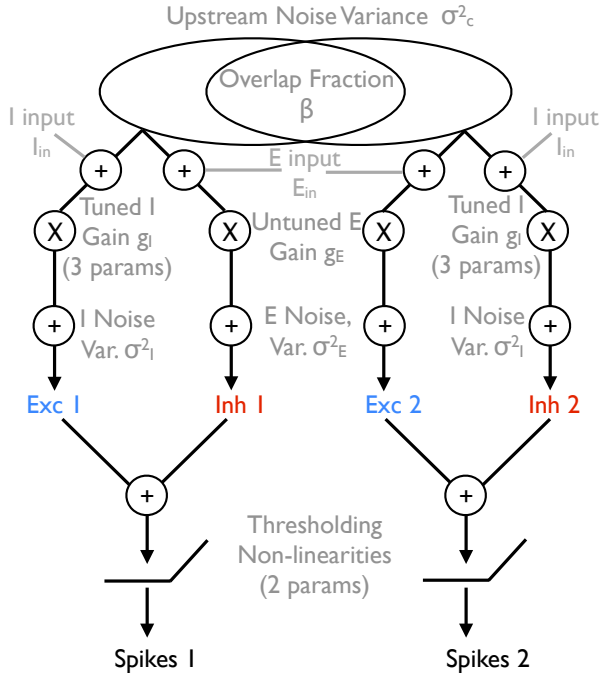

Figure 5.14: **Diagrammatic description of the computational model.** The model follows the architecture (Fig. 5.4A) revealed by our alternating voltage clamp experiments, and is shown here schematically for 2 cells, although the "full" model can contain any number of cells. Each cell pair received correlated input noise, with the correlation dictated by parameter β (which varies from pair-to-pair). This noise, and (constant for all stimuli) excitatory and inhibitory signals are then added, and multiplied by the stimulus-dependent gain factors. As such, the stim.-dependent gain factors are proportional to the mean inputs of each type for each stimulus. The E gain is untuned, and the I gain is parameterized by 3 parameters that describe the amplitude of the tuning, the (constant) offset in the gain, and the sharpness of the tuning. After the multiplicative gain stage, independent noise is added to all channels. Finally, excitation and inhibition are combined, to yield the net inputs, which are 3E - I. The factor of 3 comes in because the E or I inputs describe conductances, and excitation typically has a larger driving force, as its reversal potential is further from the resting potential than is the inhibitory reversal potential. The inputs are then passed through thresholding nonlinearities, parameterized by the threshold value and the slope of the linear portion of the curve. The parameters can vary from cell-to-cell in the model.

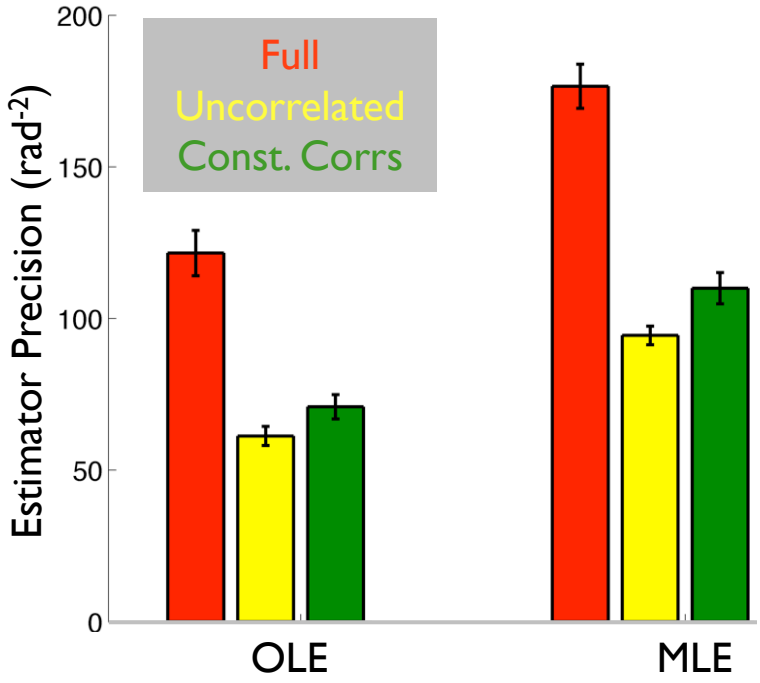

Figure 5.15: **Coding performance of the 8-cell model using measures other than Fisher information.** Similar to Fig. 5.6E, we simulated responses of our 8-ooDS-cell model to 500 different directions of stimulus motion. We then attempted to decode those responses to estimate the stimulus, and quantified the precision (inverse of mean-squared error) of the estimator. We used both the optimal linear estimator (OLE; [188]), and the maximum likelihood estimator (MLE). We repeated this calculation for the responses with stimulus-dependent correlations generated by our model ("Full"; red), for uncorrelated (trial-shuffled) responses (yellow), and for responses with stimulus-independent correlations that matched, pair-by-pair, the average of the stimulus-dependent correlations ("const corrs"; green). In all cases, the estimators were optimized for the given data type (for example, for the MLE, we used the joint distribution of population responses from the 8-cell model), and the performance was computed on held-out neural responses, in a cross-validation procedure. Error bars are the S.E.M. over 10 randomly-drawn model populations.

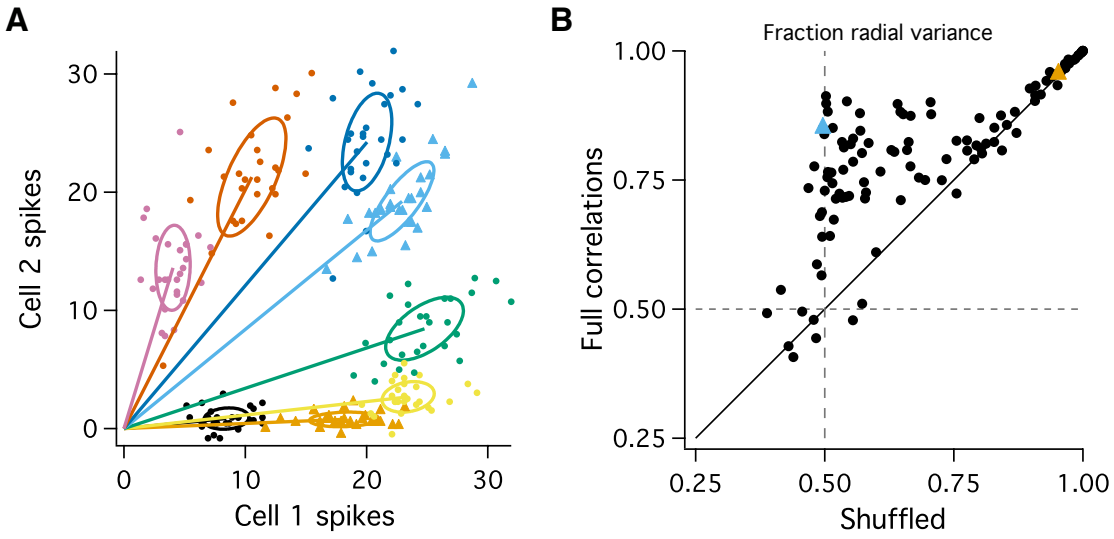

Figure 5.16: **Variability in ooDS spike responses lies along the radial direction in response space.** (A) Spike data from an example ooDS cell pair showing the conditional response distributions to 8 different bar directions, each indicated by a different color. The ellipses are the 1 standard deviation contours of these distributions, and the solid lines denote the mean response vector corresponding to each stimulus. The pairwise correlations tend to orient response variability in the direction of the mean response vector. (B) We computed the variability for each conditional distribution along the radial direction (aligned with the mean response vector) and divided it by the total variance in that distribution for the original data and for trial-shuffled data. The radial variance, V_r, was computed as $V_r = u^T \Sigma u$, where u is the mean response vector, normalized to be unit length (u^T denotes its transpose) and Σ is the covariance matrix of the conditional response distribution. Geometrically, we are computing the variance of the distribution projected along the mean response vector. The full pairwise correlated data has a greater fraction of its variance oriented radially than shuffled data ($p = 3 \times 10^{-16}$; for 14 recorded pairs, 8 stimuli each). Colored triangles denote the corresponding conditional response distributions in (A).

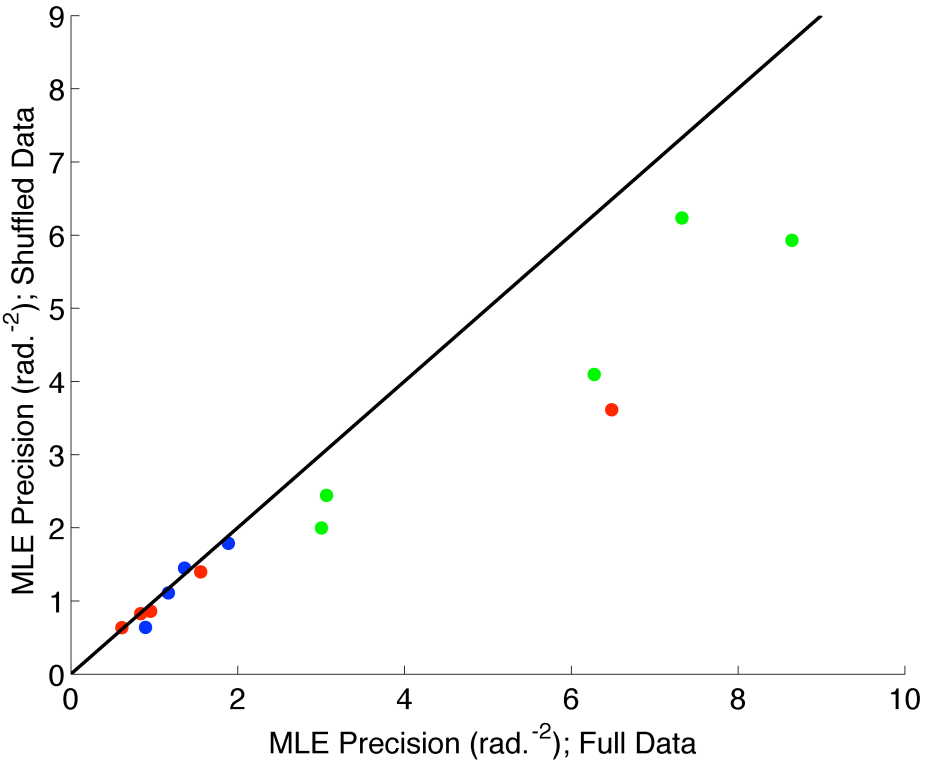

Figure 5.17: **Impact of correlations on direction coding by ooDS cell pairs.** Precision (inverse variance) of the MLE estimates of stimulus direction from the measures ooDS cell pair responses. Different colors indicate pairs with different spacings between their tuning curve peaks: either 0°(red), 90°(green), or 180°(blue). Correlations slightly improve direction coding (points tend to fall below unity line; p = 0.01, paired t-test).

Chapter 6

CONCLUDING REMARKS

6.1 Retinal ganglion cell receptive field structure and natural scene encoding

6.1.1 Findings and relation to other work

In Chapters 3 & 4, I presented findings that inform our conception of the retinal ganglion cell (RGC) receptive field, especially as it relates to the encoding of naturalistic visual stimuli. In both chapters, a key finding was that the degree of rectification applied to the output of each receptive field subunit is a key regulator of the apparent organization of the receptive field center. When subunit outputs are sharply rectified, a RGC integrates nonlinearly over visual space, and the nonlinear subunit structure of the receptive field is more likely to be important to include when constructing predictive models.

This work, as well as past work in the retinas of other species, has shown that a variety of factors can influence the rectification of subunit output and thus regulate nonlinear spatial integration by RGCs. For example, Chapter 3 showed that differences in rectification exist between On and Off parasol RGCs in macaque monkey retina (see Figs. 3.1 & 3.3). Rectification can also change with lighting conditions (see Fig. 3.4, as well as [192] for an example in mouse RGCs). Chapter 4 showed that in addition to these steady-state changes in subunit rectification, subunit rectification can be dynamically modulated by inputs to the receptive field surround.

This latter finding shows that the function of the receptive field surround may be more complicated than once thought. By combining with the center before this crucial nonlinear element of retinal circuity, the surround can modulate the spatial contrast sensitivity of the receptive field center depending on its activation. More broadly, these results support the notion that apparent receptive field structure can depend on the stimulus presented (Chapter

2).

The findings presented in Chapters 3 & 4 should help guide the development of predictive models that aim to account for RGC responses to complex visual inputs, including natural stimuli. In particular, Chapter 3 suggests that the inclusion of nonlinear subunits into receptive field models will be an important part of models for some types of RGCs. Chapter 4 extends these findings to suggest that the receptive field subunits themselves should have surrounds, rather than the surround combining with the center after formation of the center (as is currently the norm in predictive models of RGCs). As alluded to in Chapter 2, current predictive models in the retina fail to account for responses to natural scenes [99]. I would predict that these changes to model architecture will improve the ability of predictive models to generalize to natural scenes, especially for RGCs with strongly rectified subunits.

6.1.2 Remaining questions

The most puzzling result in Chapter 3 was the observation that On parasol RGCs spatially integrated natural images approximately linearly (Figs. 3.1 & 3.3), despite their strongly nonlinear responses to gratings stimuli (see Fig. 3.5). Part of this was due to the relatively shallow rectification of these cells' excitatory subunit outputs, but there seem to be additional mechanisms at play. In particular, our experiments with DOVES movies suggested that the excitatory inputs to On parasol RGCs are more nonlinear than the spike output of these cells (Figs. 3.1 & 3.2).

Preliminary results from experiments designed to explore this issue are shown in Fig. 6.1. We presented flashed natural images and linear equivalent discs, as in Fig. 3.3, while measuring spike outputs and then synaptic currents in the same cell. As we saw in Chapter 3, On parasol RGCs show more spatially linear spike responses to natural images than do Off parasol RGCs (compare Fig. 6.1A and B, left). Excitatory charge transfer measurements, from voltage-clamp experiments, reveal that excitatory input to both of these cells is nonlinear (Fig. 6.1, middle). The degree of nonlinearity is greater in Off parasol RGCs, as expected from the sharper rectification of the excitatory inputs to these cells. Nevertheless

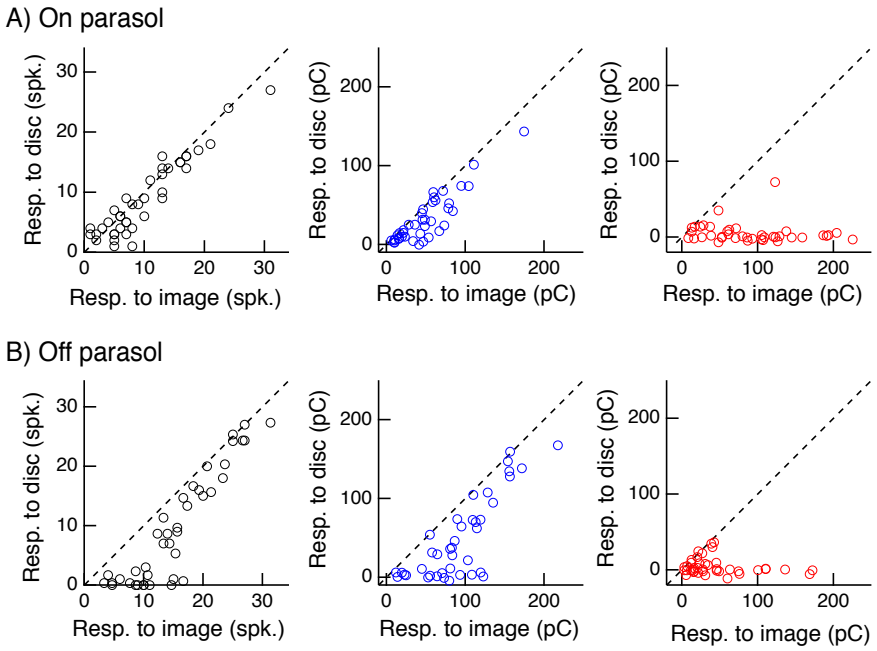

Figure 6.1: **On parasol RGC spike output is more linear than either of its synaptic inputs.** We measured spike output (left, black points), excitatory synaptic input (middle, blue), and inhibitory synaptic input (right, red) in response to flashed natural image patches and corresponding linear equivalent disc stimuli to the receptive field center, as in Fig. 3.3. (A) Responses in an example On parasol RGC. Each point is a different natural image patch. (B) Same as (A) for an example Off parasol RGC.

the spike output of On parasol RGCs is more linear than its excitatory input (compare 6.1A left and middle). The inhibitory responses to these same image patches are very nonlinear, especially in the case of the On parasol RGC (6.1, right).

I hypothesized that crossover inhibitory input into On parasol RGCs was acting to "linearize" spike responses to natural images. The same image patches that drive nonlinear responses in the cell's excitatory inputs would also be expected to drive nonlinear responses in the inhibitory inputs as well. The result, then, would be that the excitatory input, relative to the inhibitory input, would be approximately equal between a natural image and its linear equivalent stimulus. To test this, I used a leaky integrate-and-fire (LIF) modeling approach.

As inputs to the LIF model, I used measured conductances from On and Off parasol

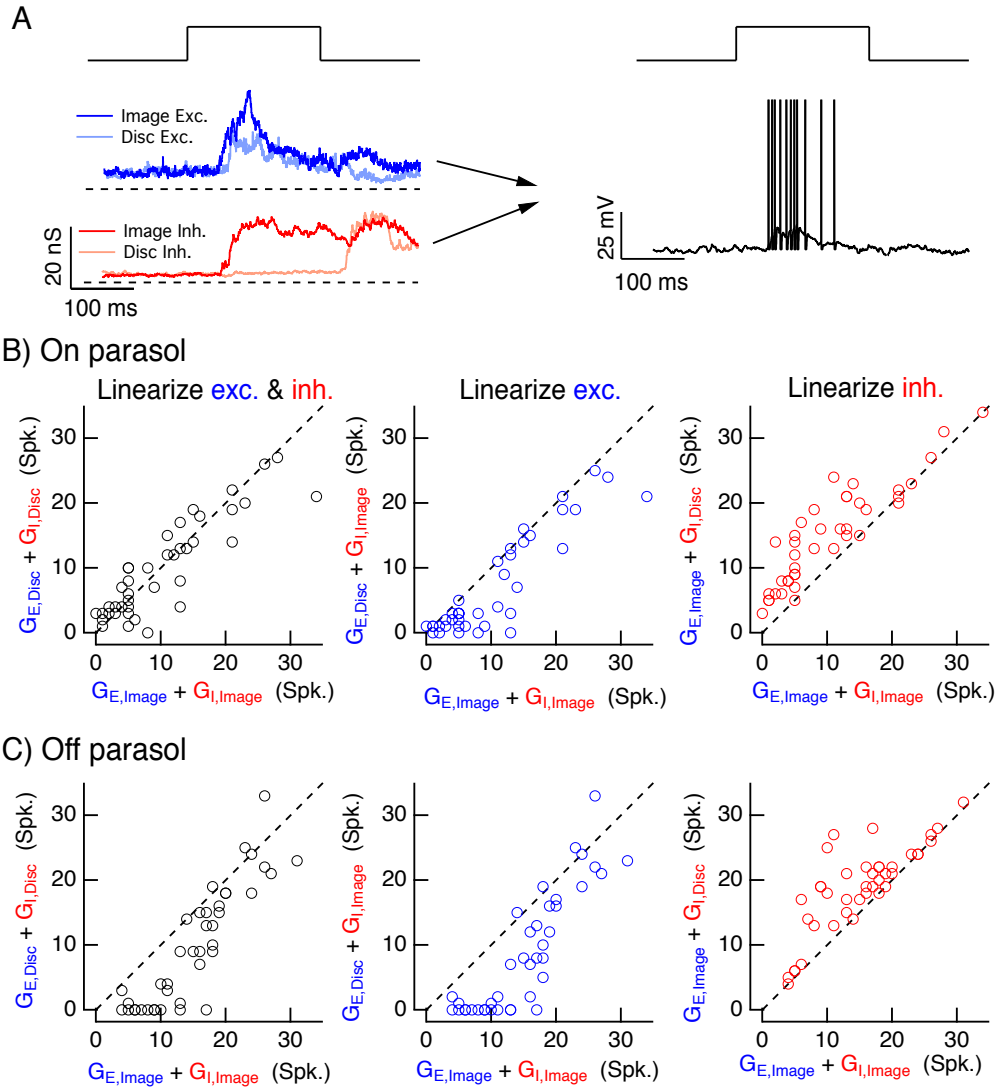

Figure 6.2: **Crossover inhibition may shape nonlinear responses to natural images.** We used measured synaptic conductances and a leaky integrate-and-fire (LIF) model to explore the role of crossover inhibition in linearizing responses to natural images. (A) Example conductance traces measured from an On parasol RGC and used as inputs to the model. For each image patch, we measured an excitatory and inhibitory response to (1) the natural image, and (2) the associated linear equivalent disc. (B) LIF model responses to various On parasol RGC conductance inputs. Each point is a different image patch. Left (black points) shows the model responses to "matched" image versus disc conductances, meant to mimic the measured spike responses in Fig. 6.1. We then linearized the excitatory channel by using the excitatory response to the linear equivalent disc paired with the inhibitory response to the natural image (middle). Conversely, we linearized the inhibitory channel by using the inhibitory response to the disc and the excitatory response to the image (right). (C) Same as (B) using conductances measured in an Off parasol RGC.

RGCs in response to flashed natural images as well as each image's linear equivalent disc stimulus. Synaptic currents were measured in voltage clamp recordings, and conductances were estimated by dividing the currents by an estimate of the driving force. The sub-threshold membrane potential, V_m, of the LIF neuron is given by equation (1) below.

$$V_m < V_\theta : \frac{dV_m}{dt} = -\frac{g_e(V_m - E_e) + g_i(V_m - E_i) + g_{leak}(V_m - E_{leak})}{C} \quad (1)$$

Where V_θ is the threshold for firing an action potential; g_e and g_i are the measured excitatory and inhibitory conductance traces, respectively; g_{leak} is a constant leak conductance; E_e, E_i, and E_{leak} are reversal potentials for the excitatory, inhibitory, and leak currents. When the membrane potential exceeds threshold, a "spike" occurs, the membrane is reset, and there is the imposition of a 2 msec absolute refractory period as well as a temporary elevation of threshold (i.e. a relative refractory period), which is given by equation (2).

$$V_\theta(t) = V_\theta + A_\theta \times e^{-t/\tau_\theta} \quad (2)$$

Where τ_θ is the time constant of the relative refractory period and A_θ is the amplitude of threshold elevation.

This allows us to ask how each synaptic input might contribute to shaping parasol RGC spike responses. We "linearize" a synaptic input by using its measured responses to the linear equivalent disc stimulus rather than responses to the natural image stimulus. For instance, we can linearize both excitatory and inhibitory inputs to mimic measured spike responses to the linear equivalent disc stimulus (Fig. 6.2BC, left, vertical axes), or we can linearize neither and mimic spike responses to natural image stimuli (Fig. 6.2BC, left, horizontal axes). The power of this modeling approach comes from the ability to simulate the spike output that would result from combinations of excitatory and inhibitory synaptic inputs that do not naturally occur, such as a "linearized" excitatory input alongside the image-driven inhibitory input. For On and Off parasol RGC conductances, we linearized excitatory inputs (Fig. 6.2BC, middle) and inhibitory inputs (Fig. 6.2BC, right) in isolation, comparing each to the image-driven spike response.

These preliminary results show that "linearizing" the excitatory inputs alone was sufficient to reproduce the measured disc responses for Off parasol RGCs. This is consistent with

the earlier finding in Chapter 3 that the excitatory inputs to Off parasol RGCs are sufficient to explain their nonlinear spiking behavior. This was not true for On parasol RGCs. Indeed, the spike responses of the On parasol model when the excitatory inputs alone were linearized were much more nonlinear than the spike responses when both excitatory and inhibitory inputs were linearized. For On parasol RGCs, it appears that nonlinear responses in both the excitatory and inhibitory inputs are important in shaping the spike outputs of the cells.

Future studies should perform similar experiments using a dynamic clamp approach, which would allow us to use a cell's real spike generating mechanism to probe synaptic integration. One limitation of the LIF modeling approach is that there are many hand-tuned parameters, and we have no systematic way to constrain many of them.

If these findings persist under more detailed study, we will still be left with the question of why the inhibitory inputs to On parasol cells so much more strongly shape spike output compared to the Off parasol cells. Using a dynamic clamp approach, we could manipulate certain features of the simulated conductances to tease apart which features of synaptic inputs are important for this linearization of spike output. A likely culprit is the amplitude of the inhibitory relative to excitatory inputs. Relative timing between excitation and inhibition may also play a role.

The possibility that crossover inhibition may serve to linearize RGC responses has been suggested as a general role of crossover inhibition in the retina [240], but no real exploration of this issue has been presented beyond correlative observations. These future studies would provide a firm test of this idea in the context of natural visual stimuli.

Another outstanding question from this work relates to the function of the observed context-dependent modulation of spatial contrast encoding. I showed that (de)activation of the receptive field surround can have a large effect on the spatial integration properties of the center. However, it is unclear how how often this modulation might be important in the context of natural scene encoding, and more generally what might its function be in encoding natural stimuli?

To address this question, we could pursue a modeling approach similar to that presented

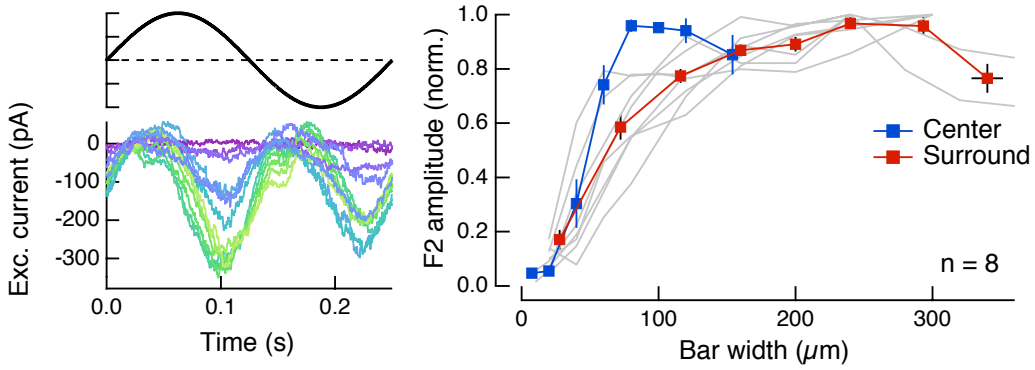

Figure 6.3: **Nonlinear spatial integration within the receptive field surround.** We measured excitatory synaptic inputs to On parasol RGCs in voltage clamp recordings while presenting contrast-reversing grating stimuli restricted to the receptive field surround. Left: each trace is the excitatory current response of an On parasol RGC to a grating of different spatial frequency. Right: the (normalized) frequency-doubled (F2) response amplitude as a function of the width of the bars that compose the grating, across a population of On parasol RGCs. Blue trace shows results for gratings restricted to the center, red for gratings in the surround, and gray traces show individual cell results for the surround.

in Chapter 3 (see Fig. 3.8). That could give some insight into the question of where in natural scenes this kind of modulation might be expected to be important. Because the peak sensitivity to spatial contrast in the center seemed to be when the surround was very weakly activated, I would predict that regions of a scene where there is strong contrast at the center-surround spatial scale would drive relatively linear responses in the center. It may be that sensitivity to fine (sub-RF center) spatial structure is a somewhat special case of parasol RGC encoding during naturalistic stimulation. It will be important to understand under what natural conditions this mode of encoding is prevalent.

Finally, we explored nonlinear interactions between the center and surround in Chapter 4, but much less is known about spatial nonlinearities within the receptive field surround itself. We have evidence that spatial nonlinearities are present in the surround of On and Off parasol RGCs, when probed using contrast-reversing grating stimuli (see Fig. 6.3 for On parasol RGC data). The excitatory synaptic inputs to On parasol RGCs show nonlinear spatial integration in the surround at a larger spatial scale than the center. This suggests

that the subunits that mediate nonlinear integration in the surround may have a different circuit origin than subunits in the center. Further work will be needed to characterize (1) the circuit basis of spatial nonlinearity in the surround and (2) the implications of a spatially nonlinear surround for the encoding of natural stimuli.

6.2 Efficient direction encoding in the retina

The work presented in Chapter 5 showed how a behaviorally-relevant feature of the visual world is efficiently encoded in the retina despite limitations imposed by noisy neural responses. The strategy for mitigating the effect of noise on encoding fidelity relies on "shaping" the noise in such a way that it does not interfere with direction encoding. It is interesting to note that this noise shaping strategy may make noise more harmful for the encoding of other stimulus features like visual contrast. This supports the notion that direction-selective circuits in the retina are very much designed to extract direction information from the visual world, and are likely not used for encoding other stimulus features.

Other observations are consistent with this interpretation of retinal direction selectivity as well. For instance, a considerable body of work shows that direction computation persists under diverse stimulus conditions including noisy backgrounds [37], very low contrasts [166, 197], and a range of luminance levels [103]. It is also robust to seemingly drastic insult, like the genetically-targeted disruption of GABA release from starburst amacrine cells, which are thought to be central to direction computation [161]. This, along with the work shown in Chapter 5, all suggests that direction-selective RGCs are highly specialized for encoding one feature of the visual world, sometimes to the detriment of the encoding of other features. Future work could test this by exploring the efficiency with which direction-selective RGCs can encode non-direction features of a scene compared to other classes of RGCs.

Recent work by one of the co-authors of the study presented in Chapter 5 showed that although the radial noise shaping seen in direction-selective RGCs was well-suited for information transmission out of the retina, other constraints on information propagation make predictions about what noise should look like in subsequent stages of visual processing [248].

In particular, the response variance introduced by downstream neurons should be "matched" to the covariance structure present in the retinal output. This means that the additional variance at these subsequent stages should also be oriented orthogonally to the stimulus space. On-Off direction-selective RGCs project to specific regions in the lateral geniculate nucleus (LGN) [49]. Future studies are needed to examine noise correlations in direction-selective LGN neurons to test whether the downstream visual circuits can make use of the beneficial noise shaping done in the retina.

BIBLIOGRAPHY

[1] Larry F. Abbott and Peter Dayan. The effect of correlated variability on the accuracy of a population code. *Neural Computation*, 11(1):91–101, 1999.

[2] E.H. Adelson and J.R. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A, Optics and image science*, 2(2):284–299, 1985.

[3] Petri Ala-Laurila, Martin Greschner, E. J. Chichilnisky, and Fred Rieke. Cone photoreceptor contributions to noise and correlations in the retinal output. *Nature Neuroscience*, 14(10):1309–1316, 2011.

[4] F R Amthor and C W Oyster. Spatial organization of retinal information about the direction of image motion. *Proceedings of the National Academy of Sciences*, 92(9):4002–4005, 1995.

[5] Franklin R. Amthor, John S. Tootle, and Norberto M. Grzywacz. Stimulus-dependent correlated firing in directionally selective retinal ganglion cells. *Visual Neuroscience*, 22(6):769–787, 2005.

[6] Juan M. Angueyra and Fred Rieke. Origin and effect of phototransduction noise in primate cone photoreceptors. *Nature Neuroscience*, 16(11):1692–1700, 2013.

[7] Joseph J. Atick. Could information theory provide an ecological theory of sensory processing? *Network: Computation in Neural Systems*, 3(2):213–251, 1992.

[8] Joseph J. Atick and A. Norman Redlich. What Does the Retina Know About Natural Scenes? *Neural Computation*, 210:196–210, 1992.

[9] Bruno B. Averbeck, Peter E. Latham, and Alexandre Pouget. Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7(5):358–366, 2006.

[10] Bruno B. Averbeck and Daeyeol Lee. Effects of noise correlations on information encoding and decoding. *Journal of Neurophysiology*, 95(6):3633–44, 2006.

[11] Stephen A. Baccus. Timing and Computation in Inner Retinal Circuitry. *Annual Review of Physiology*, 69(1):271–290, 2007.

[12] Stephen A. Baccus, Bence P. Olveczky, Mihai Manu, and Markus Meister. A retinal circuit that computes object motion. *The Journal of Neuroscience*, 28(27):6807–6817, 2008.

[13] Roland Baddeley, L. F. Abbott, Michael C. A. Booth, Frank Sengpiel, Tobe Freeman, Edward A. Wakeman, and Edmund T. Rolls. Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proceedings of the Royal Society of London B*, 264:1775–1783, 1997.

[14] Roland J. Baddeley and Benjamin W. Tatler. High frequency edges (but not contrast) predict where we fixate: A Bayesian system identification analysis. *Vision Research*,

46(18):2824–2833, 2006.

[15] Tom Baden, Philipp Berens, Katrin Franke, Miroslav Román Rosón, Matthias Bethge, and Thomas Euler. The functional diversity of retinal ganglion cells in the mouse. *Nature*, 529(7586):345–50, 2016.

[16] Wyeth Bair. Visual receptive field organization. *Current Opinion in Neurobiology*, 15(4):459–464, 2005.

[17] Wyeth Bair, Ehud Zohary, and William T. Newsome. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *The Journal of Neuroscience*, 21(5):1676–97, 2001.

[18] H. B. Barlow. Possible principles underlying the transformations of sensory messages. *Sensory communication*, 6(2):217–234, 1961.

[19] H. B. Barlow. Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, 1:371–394, 1972.

[20] H. B. Barlow, R. Fitzhugh, and S. W. Kuffler. Change of organization in the receptive fields of the cat's retina during dark adaptation. *The Journal of Physiology*, 37:338–354, 1957.

[21] H. B. Barlow, R. M. Hill, and W. R. Levick. Retinal ganglion cells responding selectively to direction and speed of image motion in the rabbit. *The Journal of Physiology*, 173:377–407, 1964.

[22] H. B. Barlow and W. R. Levick. The mechanism of directionally selective units in rabbit's retina. *The Journal of Physiology*, 178(3):477–504, 1965.

[23] Horace Barlow. Redundancy reduction revisited. *Network: Computation in Neural Systems*, 12(3):241–253, aug 2001.

[24] D. A. Baylor, B. J. Nunn, and J. L. Schnapf. Spectral sensitivity of cones of the monkey Macaca fascicularis. *The Journal of Physiology*, 390:145–160, 1987.

[25] Jeffrey Beck and Alexandre Pouget. Insights from a Simple Expression for Linear Fisher Information in a Recurrently Connected Population of Spiking Neurons. *Neural Computation*, 23(6):1484–1502, 2011.

[26] Anthony J. Bell and Terrence J. Sejnowski. The "independent components"of natural scenes are edge filters. *Vision Research*, 37(23):3327–3338, 1997.

[27] Michael J. Berry, David K. Warland, and Markus Meister. The structure and precision of retinal spike trains. *Proceedings of the National Academy of Sciences*, 94:5411–5416, 1997.

[28] Marc D. Binder and Randall K. Powers. Relationship between simulated common synaptic input and discharge synchrony in cat spinal motoneurons. *Journal of Neurophysiology*, 86:2266–2275, 2001.

[29] N. Boeddeker, J. P. Lindemann, M. Egelhaaf, and J. Zeil. Responses of blowfly motion-sensitive neurons to reconstructed optic flow along outdoor flight paths. *Journal of Comparative Physiology*, 191:1143–1155, 2005.

[30] Marco Boi, Martina Poletti, Jonathan D. Victor, and Michele Rucci. Consequences of the Oculomotor Cycle for the Dynamics of Perception. *Current Biology*, 27(9):1268–

1277, 2017.

[31] Daniel Bölinger and Tim Gollisch. Closed-Loop Measurements of Iso-Response Stimuli Reveal Dynamic Nonlinear Stimulus Integration in the Retina. *Neuron*, 73(2):333–346, 2012.

[32] Bart G. Borghuis, Jonathan S. Marvin, Loren L. Looger, and Jonathan B. Demb. Two-photon imaging of nonlinear glutamate release dynamics at bipolar cell synapses in the mouse retina. *The Journal of Neuroscience*, 33(27):10972–85, 2013.

[33] Alexander Borst, Juergen Haag, and Dierk F. Reiff. Fly motion vision. *Annual Review of Neuroscience*, 33:49–70, 2010.

[34] Jon Cafaro and Fred Rieke. Noise correlations improve response fidelity and stimulus encoding. *Nature*, 468(7326):964–7, 2010.

[35] Jon Cafaro and Fred Rieke. Regulation of spatial selectivity by crossover inhibition. *The Journal of Neuroscience*, 33(15):6310–20, 2013.

[36] Matteo Carandini, Jonathan B. Demb, Valerio Mante, David J. Tolhurst, Yang Dan, Bruno A. Olshausen, Jack L. Gallant, and Nicole C. Rust. Do We Know What the Early Visual System Does? *The Journal of Neuroscience*, 25(46):10577–10597, 2005.

[37] Qiang Chen, Zhe Pei, David Koren, and Wei Wei. Stimulus-dependent recruitment of lateral inhibition underlies retinal direction selectivity. *eLife*, 5:e21053, 2016.

[38] E. J. Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, 12(2):199–213, 2001.

[39] E. J. Chichilnisky and R. S. Kalmar. Functional asymmetries in ON and OFF ganglion cells of primate retina. *The Journal of Neuroscience*, 22(7):2737–2747, 2002.

[40] Damon A. Clark, James E. Fitzgerald, Justin M. Ales, Daryl M. Gohl, Marion A. Silies, Anthony M. Norcia, and Thomas R. Clandinin. Flies and humans share a motion estimation strategy that exploits natural scene statistics. *Nature Neuroscience*, 17(2):296–303, 2014.

[41] Ruben Coen-Cagli, Peter Dayan, and Odelia Schwartz. Cortical surround interactions and perceptual salience via natural scene statistics. *PLoS Computational Biology*, 8(3), 2012.

[42] Ruben Coen-Cagli, Adam Kohn, and Odelia Schwartz. Flexible gating of contextual influences in natural vision. *Nature Neuroscience*, 18(11):1648–1655, 2015.

[43] Marlene R. Cohen and Adam Kohn. Measuring and interpreting neuronal correlations. *Nature Neuroscience*, 14(7):811–819, 2011.

[44] Paul B. Cook and John S. McReynolds. Lateral inhibition in the inner retina is important for spatial tuning of ganglion cells. *Nature Neuroscience*, 1(8):714–719, 1998.

[45] R .James Cotton, Emmanouil Froudarakis, Patrick Storer, Peter Saggau, and Andreas S. Tolias. Three-dimensional mapping of microcircuit correlation structure. *Frontiers in Neural Circuits*, 7:151, 2013.

[46] Joanna D. Crook, Michael B. Manookin, Orin S. Packer, and Dennis M. Dacey. Horizontal Cell Feedback without Cone Type-Selective Inhibition Mediates "Red-Green" Color Opponency in Midget Ganglion Cells of the Primate Retina. *The Journal of*

Neuroscience, 31(5):1762–1772, 2011.

[47] Joanna D. Crook, Orin S. Packer, and Dennis M. Dacey. A synaptic signature for ON- and OFF-center parasol ganglion cells of the primate retina. *Visual Neuroscience*, 31:57–84, 2014.

[48] Joanna D. Crook, Beth B. Peterson, Orin S. Packer, Farrel R. Robinson, John B. Troy, and Dennis M. Dacey. Y-Cell Receptive Field and Collicular Projection of Parasol Ganglion Cells in Macaque Monkey Retina. *The Journal of Neuroscience*, 28(44):11277–11291, 2008.

[49] Alberto Cruz-Martín, Rana N. El-Danaf, Fumitaka Osakada, Balaji Sriram, Onkar S. Dhande, Phong L. Nguyen, Edward M. Callaway, Anirvan Ghosh, and Andrew D. Huberman. A dedicated circuit links direction-selective retinal ganglion cells to the primary visual cortex. *Nature*, 507(7492):358–61, 2014.

[50] Rava Azeredo da Silveira and Michael J. Berry. High-Fidelity Coding with Correlated Neurons. *PLoS Computational Biology*, 10(11), 2014.

[51] Dennis Dacey, Orin S. Packer, Lisa Diller, David Brainard, Beth Peterson, and Barry Lee. Center surround receptive field structure of cone bipolar cells in primate retina. *Vision Research*, 40(14):1801–1811, 2000.

[52] Dennis M. Dacey, Beth B. Peterson, Farrel R. Robinson, and Paul D. Gamlin. Fireworks in the primate retina: In vitro photodynamics reveals diverse LGN-projecting ganglion cell types. *Neuron*, 37(1):15–27, 2003.

[53] Yang Dan, Joseph J. Atick, and R. Clay Reid. Efficient Coding of Natural Scenes in the Lateral Geniculate Nucleus: Experimental Test of a Computational Theory. *The Journal of Neuroscience*, 16(10):3351–3362, 1996.

[54] Christopher M. Davenport, Peter B. Detwiler, and Dennis M. Dacey. Effects of pH buffering on horizontal and ganglion cell light responses in primate retina: evidence for the proton hypothesis of surround formation. *The Journal of Neuroscience*, 28(2):456–464, 2008.

[55] Stephen V. David, William E. Vinje, and Jack L. Gallant. Natural Stimulus Statistics Alter the Receptive Field Structure of V1 Neurons. *The Journal of Neuroscience*, 24(31):6991–7006, 2004.

[56] Peter Dayan and Larry F. Abbott. *Theoretical Neuroscience*. Cambridge, MA: MIT Press, 2002.

[57] Jaime de la Rocha, Brent Doiron, Eric Shea-Brown, Kresimir Josić, and Alex Reyes. Correlation between neural spike trains increases with firing rate. *Nature*, 448(7155):802–6, 2007.

[58] Russell L. De Valois, Herman Morgan, and D. Ma Snodderly. Psychophysical studies of monkey Vision-III. Spatial luminance contrast sensitivity tests of macaque and human observers. *Vision Research*, 14(1):75–81, 1974.

[59] Jonathan B. Demb. Cellular Mechanisms for Direction Selectivity in the Retina. *Neuron*, 55(2):179–186, 2007.

[60] Jonathan B. Demb, Loren Haarsma, Michael A. Freed, and Peter Sterling. Functional

circuitry of the retinal ganglion cell's nonlinear receptive field. *The Journal of Neuroscience*, 19(22):9756–9767, 1999.

[61] Jonathan B. Demb, Kareem Zaghloul, and Peter Sterling. Cellular basis for the response to second-order motion cues in Y retinal ganglion cells. *Neuron*, 32(4):711–721, 2001.

[62] Dawei W. Dong and Joseph J. Atick. Statistics of Natural Time-Varying Images. *Network: Computation in Neural Systems*, 6(3):345–358, 1995.

[63] Felice A. Dunn, Martin J. Lankheet, and Fred Rieke. Light adaptation in cone vision involves switching between receptor and post-receptor sites. *Nature*, 449(7162):603–606, 2007.

[64] Alexander S. Ecker, Philipp Berens, R. James Cotton, Manivannan Subramaniyan, George H. Denfield, Cathryn R. Cadwell, Stelios M. Smirnakis, Matthias Bethge, and Andreas S. Tolias. State dependence of noise correlations in macaque primary visual cortex. *Neuron*, 82(1):235–248, 2014.

[65] Alexander S. Ecker, Philipp Berens, Georgios A. Keliris, Matthias Bethge, Nikos K. Logothetis, and Andreas S. Tolias. Decorrelated neuronal firing in cortical microcircuits. *Science*, 327(5965):584–587, jan 2010.

[66] Alexander S. Ecker, Philipp Berens, Andreas S. Tolias, and Matthias Bethge. The Effect of Noise Correlations in Populations of Diversely Tuned Neurons. *The Journal of Neuroscience*, 31(40):14272–14283, 2011.

[67] Martin Egelhaaf, Roland Kern, Holger G. Krapp, Jutta Kretzberg, Rafael Kurtz, and Anne Kathrin Warzecha. Neural encoding of behaviourally relevant visual-motion information in the fly. *Trends in Neurosciences*, 25(2):96–102, feb 2002.

[68] Justin Elstrott, Anastasia Anishchenko, Martin Greschner, Alexander Sher, Alan M. Litke, E. J. Chichilnisky, and Marla B. Feller. Direction Selectivity in the Retina Is Established Independent of Visual Experience and Cholinergic Retinal Waves. *Neuron*, 58(4):499–506, 2008.

[69] Christina Enroth-Cugell and J. G. Robson. the Contrast Sensitivity of Retinal Ganglion Cells of the Cat. *The Journal of Physiology*, 187(17):517–552, 1966.

[70] Thomas Euler, Silke Haverkamp, Timm Schubert, and Tom Baden. Retinal bipolar cells: elementary building blocks of vision. *Nature Reviews Neuroscience*, 15(8):507–519, jul 2014.

[71] Karl Farrow and Richard H. Masland. Physiological clustering of visual channels in the mouse retina. *Journal of Neurophysiology*, 105:1516–1530, 2013.

[72] Gidon Felsen, Jon Touryan, Feng Han, and Yang Dan. Cortical sensitivity to visual features in natural scenes. *PLoS Biology*, 3(10), 2005.

[73] David Ferster. Spatially opponent excitation and inhibition in simple cells of the cat visual cortex. *The Journal of Neuroscience*, 8(4):1172–1180, 1988.

[74] David J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America. A*, 4(12):2379–94, 1987.

[75] David J. Field. What Is the Goal of Sensory Coding? *Neural Computation*, 6(4):559–601, 1994.

[76] G. D. Field and E. J. Chichilnisky. Information processing in the primate retina: circuitry and coding. *Annual Review of Neuroscience*, 30(1):1–30, 2007.

[77] Greg D. Field, Jeffrey L. Gauthier, Alexander Sher, Martin Greschner, Timothy A. Machado, Lauren H. Jepson, Jonathon Shlens, Deborah E. Gunning, Keith Mathieson, Wladyslaw Dabrowski, Liam Paninski, Alan M. Litke, and E. J. Chichilnisky. Functional connectivity in the retina at the resolution of photoreceptors. *Nature*, 467(7316):673–677, 2010.

[78] James E. Fitzgerald and Damon A. Clark. Nonlinear circuits for naturalistic visual motion estimation. *eLife*, 4:e09123, 2015.

[79] Nicholas Flores-Herr, Dario A. Protti, and Heinz Wässle. Synaptic currents generating the inhibitory surround of ganglion cells in the mammalian retina. *The Journal of Neuroscience*, 21(13):4852–4863, 2001.

[80] Felix Franke, Michele Fiscella, Maksim Sevelev, Botond Roska, Andreas Hierlemann, and Rava Azeredo da Silveira. Structures of Neural Correlation and How They Favor Coding. *Neuron*, 89(2):409–422, 2016.

[81] Katrin Franke, Philipp Berens, Timm Schubert, Matthias Bethge, Thomas Euler, and Tom Baden. Balanced excitation and inhibition decorrelates visual feature representation in the mammalian inner retina. *Nature*, 542:439–444, 2017.

[82] Jeremy Freeman, Greg D. Field, Peter H. Li, Martin Greschner, Deborah E. Gunning, Keith Mathieson, Alexander Sher, Alan M. Litke, Liam Paninski, Eero P. Simoncelli, and E. J. Chichilnisky. Mapping nonlinear receptive field structure in primate retina at single cone resolution. *eLife*, e05241:1–21, 2015.

[83] Shelley I. Fried, Thomas A. Münch, and Frank S. Werblin. Mechanisms and circuitry underlying directional selectivity in the retina. *Nature*, 420(6914):411–414, 2002.

[84] Emmanouil Froudarakis, Philipp Berens, Alexander S. Ecker, R. James Cotton, Fabian H. Sinz, Dimitri Yatsenko, Peter Saggau, Matthias Bethge, and Andreas S. Tolias. Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. *Nature Neuroscience*, 17(6):851–7, 2014.

[85] Timothy J. Gawne and Barry J. Richmond. How independent are the messages carried by adjacent inferior temporal cortical neurons? *The Journal of Neuroscience*, 13(7):2758–71, 1993.

[86] Tim Gollisch and Markus Meister. Rapid neural coding in the retina with relative spike latencies. *Science*, 319(5866):1108–1111, 2008.

[87] Tim Gollisch and Markus Meister. Eye Smarter than Scientists Believed: Neural Computations in Circuits of the Retina. *Neuron*, 65(2):150–164, 2010.

[88] Robbe L. T. Goris, J. Anthony Movshon, and Eero P. Simoncelli. Partitioning neuronal variability. *Nature Neuroscience*, 17(6):858–65, 2014.

[89] Daniel J. Graham, Damon M. Chandler, and David J. Field. Can the theory of "whitening" explain the center-surround properties of retinal ganglion cell receptive fields?

Vision Research, 46(18):2901–2913, 2006.

[90] Einat Granot-Atedgi, Gašper Tkačik, Ronen Segev, and Elad Schneidman. Stimulus-dependent Maximum Entropy Models of Neural Population Codes. *PLoS Computational Biology*, 9(3), 2013.

[91] Michael Graupner and Alex D. Reyes. Synaptic Input Correlations Leading to Membrane Potential Decorrelation of Spontaneous Activity in Cortex. *The Journal of Neuroscience*, 33(38):15075–15085, 2013.

[92] William N. Grimes, Logan R. Graves, Mathew T. Summers, and Fred Rieke. A simple retinal mechanism contributes to perceptual interactions between rod-and cone-mediated responses in primates. *eLife*, 4:e08033, 2015.

[93] William N. Grimes, Gregory W. Schwartz, and Fred Rieke. The synaptic and circuit mechanisms underlying a change in spatial encoding in the retina. *Neuron*, 82(2):460–473, apr 2014.

[94] Bilal Haider, Matthew R. Krause, Alvaro Duque, Yuguo Yu, Jonathan Touryan, James A. Mazer, and David A. McCormick. Synaptic and Network Mechanisms of Sparse and Reliable Visual Cortical Activity during Nonclassical Receptive Field Stimulation. *Neuron*, 65(1):107–121, 2010.

[95] Bryan J. Hansen, Mircea I. Chelaru, and Valentin Dragoi. Correlated Variability in Laminar Cortical Circuits. *Neuron*, 76(3):590–602, 2012.

[96] H. K. Hartline. The Response of Single Optic Nerve Fibers of the vertebrate eye to illumination of the retina. *American Journal of Physiology*, pages 400–415, 1937.

[97] H. K. Hartline. The receptive fields of optic nerve fibers. *American Journal of Physiology*, pages 690–699, 1940.

[98] Klaus Hausen and Martin Egelhaaf. Neural mechanisms of visual course Control in Insects. In Doekele G. Stavenga and Roger C Hardie, editors, *Facets of Vision*, chapter 18, pages 391–424. Springer London, 1989.

[99] Alexander Heitman, Nora Brackbill, Martin Greschner, Alexander Sher, Alan M. Litke, and E. J. Chichilnisky. Testing pseudo-linear models of responses to natural scenes in primate retina. *bioRxiv*, 2016.

[100] S. Hochstein and R. M. Shapley. Linear and nonlinear spatial subunits in Y cat retinal ganglion cells. *The Journal of Physiology*, 262(2):265–84, 1976.

[101] S. Hochstein and R. M. Shapley. Quantitative analysis of retinal ganglion cell classifications. *The Journal of Physiology*, 262(2):237–264, 1976.

[102] Paul G. Hoel, J. Wolfowitz, and Harald Cramer. Mathematical Methods of Statistics. *Journal of the American Statistical Association*, 42(237):174, 1947.

[103] Alex Hoggarth, Amanda J. McLaughlin, Kara Ronellenfitch, Stuart Trenholm, Rishi Vasandani, Santhosh Sethuramanujam, David Schwab, Kevin L. Briggman, and Gautam B. Awatramani. Specific wiring of distinct amacrine cells in the directionally selective retinal circuit permits independent coding of direction and size. *Neuron*, 86(1):276–291, 2015.

[104] Yu Hu, Joel Zylberberg, and Eric Shea-Brown. The Sign Rule and Beyond: Bound-

ary Effects, Flexibility, and Noise Correlations in Neural Population Codes. *PLoS Computational Biology*, 10(2), 2014.

[105] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148:574–591, 1959.

[106] D. H. Hubel and T. N. Wiesel. Integrative Action in the Cat's Lateral Geniculate Body. *The Journal of Physiology*, 155:385–398, 1961.

[107] Tomomi Ichinose and Peter D. Lukasiewicz. Inner and outer retinal pathways both contribute to surround inhibition of salamander ganglion cells. *The Journal of Physiology*, 565:517–535, 2005.

[108] Tim Jarsky, Mark Cembrowski, Stephen M. Logan, William L. Kath, Hermann Riecke, Jonathan B. Demb, and Joshua H. Singer. A synaptic mechanism for retinal adaptation to luminance and contrast. *The Journal of Neuroscience*, 31(30):11003–11015, 2011.

[109] James M. Jeanne, Tatyana O. Sharpee, and Timothy Q. Gentner. Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron*, 78(2):352–363, 2013.

[110] Lauren H. Jepson, Pawel Hottowy, Geoffrey A. Weiner, Władysław Dabrowski, Alan M. Litke, and E. J. Chichilnisky. High-fidelity reproduction of spatiotemporal visual signals for retinal prosthesis. *Neuron*, 83(1):87–92, 2014.

[111] Na Ji, Jeremy Freeman, and Spencer L Smith. Technologies for imaging neural activity in large volumes. *Nature Neuroscience*, 19(9):1154–64, 2016.

[112] Kresimir Josić, Eric Shea-Brown, Brent Doiron, and Jaime de la Rocha. Stimulus-dependent correlations and population codes. *Neural Computation*, 21(10):2774–2804, oct 2009.

[113] K. Karmeier, J. H. van Hateren, R. Kern, and M. Egelhaaf. Encoding of Naturalistic Optic Flow by a Population of Blowfly Motion-Sensitive Neurons. *Journal of Neurophysiology*, 96(3):1602–1614, 2006.

[114] David B. Kastner and Stephen A. Baccus. Spatial segregation of adaptation and predictive sensitization in retinal ganglion cells. *Neuron*, 79(3):541–554, 2013.

[115] David B. Kastner and Stephen A. Baccus. Insights from the retina into the diverse and general computations of adaptation, detection, and prediction. *Current Opinion in Neurobiology*, 25:63–69, 2014.

[116] Justin Keat, Pamela Reinagel, R. Clay Reid, and Markus Meister. Predicting Every Spike: A Model for the Responses of Visual Neurons. *Neuron*, 30(3):803–817, 2001.

[117] D. H. Kelly. Adaptation effects on spatio-temporal sine-wave thresholds. *Vision Research*, 12(1):89–101, 1972.

[118] In-Jung Kim, Yifeng Zhang, Markus Meister, and Joshua R. Sanes. Laminar Restriction of Retinal Ganglion Cell Dendrites and Axons: Subtype-Specific Developmental Patterns Revealed with Transgenic Markers. *The Journal of Neuroscience*, 30(4):1452–1462, 2010.

[119] In-Jung Kim, Yifeng Zhang, Masahito Yamagata, Markus Meister, and Joshua R. Sanes. Molecular identification of a retinal cell type that responds to upward motion.

Nature, 452(7186):478–482, 2008.

[120] Kristin Koch, Judith McLean, Ronen Segev, Michael A. Freed, Michael J. Berry, Vijay Balasubramanian, and Peter Sterling. How Much the Eye Tells the Brain. *Current Biology*, 16(14):1428–1434, 2006.

[121] Adam Kohn and Matthew A. Smith. Stimulus Dependence of Neuronal Correlation in Primary Visual Cortex of the Macaque. *The Journal of Neuroscience*, 25(14):3661–3673, 2005.

[122] Holger G. Krapp, Roland Hengstenberg, and Martin Egelhaaf. Binocular contributions to optic flow processing in the fly visual system. *Journal of neurophysiology*, 85(2):724–734, 2001.

[123] Vidhyasankar Krishnamoorthy, Michael Weick, and Tim Gollisch. Sensitivity to image recurrence across eye-movement-like image transitions through local serial inhibition in the retina. *eLife*, page e22431, 2017.

[124] Xutao Kuang, Martina Poletti, Jonathan D. Victor, and Michele Rucci. Temporal encoding of spatial information during active visual fixation. *Current Biology*, 22(6):510–514, 2012.

[125] Steven W. Kuffler. Discharge Patterns and Functional Organization of Mammalian Retina. *Journal of Neurophysiology*, 16(1):37–68, 1953.

[126] Sidney P. Kuo, Gregory W. Schwartz, and Fred Rieke. Nonlinear Spatiotemporal Integration by Electrical and Chemical Synapses in the Retina. *Neuron*, 90(2):320–332, 2016.

[127] Lampl I., Reichova I., and D. Ferster. Synchronous membrane potential fluctuations in neurons of the cat visual cortex. *Neuron*, 22(2):361–374, 1999.

[128] Simon Laughlin. A simple coding procedure enhances a neuron's information capacity. *Zeitschrift fur Naturforschung - Section C Journal of Biosciences*, 36:910–912, 1981.

[129] Nicholas A. Lesica, Toshiyuki Ishii, Garrett B. Stanley, and Toshihiko Hosoya. Estimating receptive fields from responses to natural stimuli with asymmetric intensity distributions. *PLoS ONE*, 3(8):23–25, 2008.

[130] J.Y. Lettvin, H.R. Maturana, W.S. McCulloch, and W.H. Pitts. What the Frog's Eye Tells the Frog's brain. *Proceedings of the National Academy of Sciences*, pages 1940–1951, 1959.

[131] I. Chun Lin, Michael Okun, Matteo Carandini, and Kenneth D. Harris. The Nature of Shared Cortical Variability. *Neuron*, 87(3):644–656, 2015.

[132] Yvonne H. L. Luo and Lyndon Da Cruz. A review and update on the current status of retinal prostheses (bionic eye). *British Medical Bulletin*, 109(1):31–44, 2014.

[133] Cheng Ly, Jason W. Middleton, and Brent Doiron. Cellular and Circuit Mechanisms Maintain Low Spike Co-Variability and Enhance Population Coding in Somatosensory Cortex. *Frontiers in Computational Neuroscience*, 6:1–26, 2012.

[134] Niru Maheswaranathan, Stephen A. Baccus, and Surya Ganguli. Inferring hidden structure in multilayered neural circuits. *bioRxiv*, pages 1–28, 2017.

[135] Stuart C. Mangel. Analysis of the horizontal cell contribution to the receptive field

surround of ganglion cells in the rabbit retina. *The Journal of Physiology*, 442:211–234, 1991.

[136] Adam Mani and Gregory W. Schwartz. Circuit Mechanisms of a Retinal Ganglion Cell with Stimulus-Dependent Response Latency and Activation Beyond Its Dendrites. *Current Biology*, 27(4):471–482, 2017.

[137] D Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society of London B*, 207:187–217, 1980.

[138] Susana Martinez-Conde, Stephen L. Macknik, and David H. Hubel. The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience*, 5(3):229–240, 2004.

[139] Susana Martinez-Conde, Stephen L. Macknik, Xoana G. Troncoso, and David H. Hubel. Microsaccades: a neurophysiological analysis. *Trends in Neurosciences*, 32(9):463–475, 2009.

[140] Susana Martinez-Conde, Jorge Otero-Millan, and Stephen L Macknik. The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nature Reviews Neuroscience*, 14(2):83–96, 2013.

[141] Richard H. Masland. The many roles of starburst amacrine cells. *Trends in Neurosciences*, 28(8):395–396, 2005.

[142] Richard H. Masland. The Neuronal Organization of the Retina. *Neuron*, 76(2):266–280, 2012.

[143] Richard H. Masland and Paul R. Martin. The unsolved mystery of vision. *Current Biology*, 17(15):577–582, 2007.

[144] Lane T. Mcintosh, Niru Maheswaranathan, Aran Nayebi, S. Ganguli, and Stephen A. Baccus. Deep Learning Models of the Retinal Response to Natural Scenes. *Advances in Neural Information Processing Systems (NIPS)*, 30:1–9, 2016.

[145] Matthew J. McMahon, Orin S. Packer, and Dennis M. Dacey. The Classical Receptive Field Surround of Primate Parasol Ganglion Cells Is Mediated Primarily by a Non-GABAergic Pathway. *The Journal of Neuroscience*, 24(15):3736–3745, 2004.

[146] Markus Meister, Leon Lagnado, and Denis A. Baylor. Concerted signaling by retinal ganglion cells. *Science*, 270(5239):1207–1210, 1995.

[147] Fernando Montani, Adam Kohn, Matthew A. Smith, and Simon R. Schultz. How do stimulus-dependent correlations between V1 neurons affect neural coding? *Neurocomputing*, 70(10-12):1782–1787, 2007.

[148] Rubén Moreno-Bote, Jeffrey Beck, Ingmar Kanitscheider, Xaq Pitkow, Peter Latham, and Alexandre Pouget. Information-limiting correlations. *Nature Neuroscience*, 17(10):1410–1417, 2014.

[149] Lena S. Mortensen, Silvia J. H. Park, Jiang Bin Ke, Benjamin H. Cooper, Lei Zhang, Cordelia Imig, Siegrid Lowel, Kerstin Reim, Nils Brose, Jonathan B. Demb, Jeong Seop Rhee, and Joshua H. Singer. Complexin 3 Increases the Fidelity of Signaling in a Retinal Circuit by Regulating Exocytosis at Ribbon Synapses. *Cell Reports*, 15(10):2239–2250, 2016.

[150] Amurta Nath and Gregory W Schwartz. Cardinal Orientation Selectivity Is Represented by Two Distinct Ganglion Cell Types in Mouse Retina. *The Journal of Neuroscience*, 36(11):3208–21, 2016.

[151] D. C. O'Carroll, N. J. Bidwell, S. B. Laughlin, and E. J. Warrant. Insect motion detectors matched to visual ecology., 1996.

[152] Bruno A. Olshausen and David J. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583):607–609, 1996.

[153] Bruno A. Olshausen and David J. Field. Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research*, 37(23):3311–3325, 1997.

[154] Bruno A. Olshausen and David J. Field. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487, 2004.

[155] Bence P. Olveczky, Stephen A. Baccus, and Markus Meister. Segregation of object and background motion in the retina. *Nature*, 423(6938):401–408, 2003.

[156] Mike W. Oram, Peter Földiák, David I. Perrett, and Frank Sengpiel. The 'ideal homunculus': Decoding neural population signals. *Trends in Neurosciences*, 21(6):259–265, 1998.

[157] Clyde W. Oyster and Horace B. Barlow. Direction-selective units in rabbit retina: distribution of preferred directions. *Science*, 155(764):841–842, 1967.

[158] Stefano Panzeri, Simon R. Schultz, Alessandro Treves, and Edmund T. Rolls. Correlations and the encoding of information in the nervous system. *Proceedings of the Royal Society of London B*, 266(1423):1001–12, 1999.

[159] Mijung Park and Jonathan W. Pillow. Receptive field inference with localized priors. *PLoS Computational Biology*, 7(10):e1002219, 2011.

[160] Michael Pecka, Yunyun Han, Elie Sader, and Thomas D. Mrsic-Flogel. Experience-Dependent Specialization of Receptive Field Surround for Selective Coding of Natural Scenes. *Neuron*, 84(2):457–469, 2014.

[161] Zhe Pei, Qiang Chen, David Koren, Benno Giammarinaro, Hector Acaron Ledesma, and Wei Wei. Conditional Knock-Out of Vesicular GABA Transporter Gene from Starburst Amacrine Cells Reveals the Contributions of Multiple Synaptic Mechanisms Underlying Direction Selectivity in the Retina. *The Journal of Neuroscience*, 35(38):13219–32, 2015.

[162] Dumitru Petrusca, Matthew I. Grivich, Alexander Sher, Greg D. Field, Jeffrey L. Gauthier, Martin Greschner, Jonathon Shlens, E. J. Chichilnisky, and Alan M. Litke. Identification and characterization of a Y-like primate retinal ganglion cell type. *The Journal of Neuroscience*, 27(41):11019–11027, 2007.

[163] Jonathan W. Pillow, Liam Paninski, Valerie J. Uzzell, Eero P. Simoncelli, and E. J. Chichilnisky. Prediction and Decoding of Retinal Ganglion Cell Responses with a Probabilistic Spiking Model. *The Journal of Neuroscience*, 25(47):11003–11013, 2005.

[164] Jonathan W. Pillow, Jonathon Shlens, Liam Paninski, Alexander Sher, Alan M. Litke, E. J. Chichilnisky, and Eero P. Simoncelli. Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, 454(7207):995–9, 2008.

[165] Xaq Pitkow and Markus Meister. Decorrelation and efficient coding by retinal ganglion cells. *Nature Neuroscience*, 15(4):628–635, 2012.

[166] Alon Poleg-polsky and Jeffrey S. Diamond. Retinal Circuitry Balances Contrast Tuning of Excitation and Inhibition to Enable Reliable Computation of Direction Selectivity. *The Journal of Neuroscience*, 36(21):5861–5876, 2016.

[167] Adrián Ponce-Alvarez, Alexander Thiele, Thomas D Albright, Gene R Stoner, and Gustavo Deco. Stimulus-dependent variability and noise correlations in cortical MT neurons. *Proceedings of the National Academy of Sciences*, 110(32):13162–13167, 2013.

[168] Jason L. Puchalla, Elad Schneidman, Robert A. Harris, and Michael J. Berry. Redundancy in the population code of the retina. *Neuron*, 46(3):493–504, 2005.

[169] Christian Puller, Michael B. Manookin, Jay Neitz, Fred Rieke, and Maureen Neitz. Broad Thorny Ganglion Cells: A Candidate for Visual Pursuit Error Signaling in the Primate Retina. *The Journal of Neuroscience*, 35(13):5397–5408, 2015.

[170] Theresa Puthussery, Sowmya Venkataramani, Jacqueline Gayet-primo, Robert G. Smith, and W. Rowland Taylor. NaV1.1 Channels in Axon Initial Segments of Bipolar Cells Augment Input to Magnocellular Visual Pathways in the Primate Retina. *The Journal of Neuroscience*, 33(41):16045–16059, 2013.

[171] C. Radhakrishna Rao. Information and the Accuracy Attainable in the Estimation of Statistical Parameters. *Bull. Calcutta Math. Soc*, 37(3):81–91, 1945.

[172] Esteban Real, Hiroki Asari, Tim Gollisch, and Markus Meister. Neural Circuit Inference from Function to Structure. *Current Biology*, 27(2):189–198, 2017.

[173] Daniel S. Reich, Ferenc Mechler, and Jonathan D. Victor. Independent and Redundant Information in Nearby Cortical Neurons. *Science*, 294(5551):2566–2568, 2001.

[174] Pamela Reinagel. How do visual neurons respond in the real world? *Current Opinion in Neurobiology*, 11(4):437–442, 2001.

[175] Pamela Reinagel and Anthony M. Zador. Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, 10(4):341–350, 1999.

[176] Alfonso Renart, Jaime de la Rocha, Peter Bartho, Liad Hollender, Nestor Parga, Alex Reyes, and Kenneth D. Harris. The asynchronous state in cortical circuits. *Science*, 327(5965):587–90, 2010.

[177] Fred Rieke, Davd Warland, Rob de Ruyter van Steveninck, and William Bialek. *Spikes: exploring the neural code*. MIT Press, 1999.

[178] Dario L Ringach, Michael J Hawken, and Robert Shapley. Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *Journal of Vision*, 2(1):12–24, 2002.

[179] Michal Rivlin-Etzion, Kaili Zhou, Wei Wei, Justin Elstrott, Phong L. Nguyen, Ben A. Barres, Andrew D. Huberman, and Marla B. Feller. Transgenic Mice Reveal Unexpected Diversity of On-Off Direction-Selective Retinal Ganglion Cell Subtypes and Brain Structures Involved in Motion Processing. *The Journal of Neuroscience*, 31(24):8760–8769, 2011.

[180] Edmund T. Rolls and Martin J. Tovee. Sparseness of the neuronal representation of

stimuli in the primate temporal visual cortex. *Journal of Neurophysiology*, 73(2):713–26, 1995.

[181] Ranulfo Romo, Adrián Hernández, Antonio Zainos, and Emilio Salinas. Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron*, 38(4):649–657, 2003.

[182] Michele Rucci. Fixational eye movements, natural image statistics, and fine spatial vision. *Network: Computation in Neural Systems*, 19:253–285, 2008.

[183] Michele Rucci and Jonathan D. Victor. The unsteady eye: An information-processing stage, not a bug. *Trends in Neurosciences*, 38(4):195–206, 2015.

[184] Daniel L Ruderman. The statistics of natural images. *Network: Computation in Neural Systems*, 5(4):517–548, 1994.

[185] Daniel L Ruderman and William Bialek. Statistics of natural images: Scaling in the woods. *Advances in neural information processing systems*, 73(6):551–558, 1994.

[186] Nicole C. Rust and J. Anthony Movshon. In praise of artifice. *Nature Neuroscience*, 8(12):1647–1650, 2005.

[187] Nicole C. Rust, Odelia Schwartz, J. Anthony Movshon, and Eero P. Simoncelli. Spatiotemporal elements of macaque V1 receptive fields. *Neuron*, 46(6):945–956, 2005.

[188] Emilio Salinas and L. F. Abbott. Vector Reconstruction from Firing Rates. *Journal of computational neuroscience*, 107(1-2):89–107, 1994.

[189] Jason M. Samonds and A. B. Bonds. Gamma Oscillation Maintains Stimulus Structure-Dependent Synchronization in Cat Visual Cortex. *Journal of Neurophysiology*, 93(1):223–236, 2005.

[190] J. L. Schnapf, B. J. Nunn, M. Meister, and D. A. Baylor. Visual transduction in cones of the monkey Macaca fascicularis. *The Journal of Physiology*, 427:681–713, 1990.

[191] Elad Schneidman, Susanne Still, Michael J Berry, and William Bialek. Network information and connected correlations. *Physical Review Letters*, 91(23):238701, 2003.

[192] Gregory W. Schwartz, Haruhisa Okawa, Felice A. Dunn, Josh L. Morgan, Daniel Kerschensteiner, Rachel O. Wong, and Fred Rieke. The spatial structure of a nonlinear receptive field. *Nature Neuroscience*, 15(11):1572–1580, 2012.

[193] Gregory W. Schwartz and Fred Rieke. Perspectives on: Information and coding in mammalian sensory physiology: Nonlinear spatial encoding by retinal ganglion cells: when 1 + 1 != 2. *The Journal of General Physiology*, 138(3):283–290, 2011.

[194] Odelia Schwartz, Jonathan W. Pillow, Nicole C. Rust, and Eero P. Simoncelli. Spike-triggered neural characterization. *Journal of Vision*, 6(4):484–507, 2006.

[195] Odelia Schwartz and Eero P Simoncelli. Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4(8):819–25, 2001.

[196] Irina Yonit Segal, Chen Giladi, Michael Gedalin, Michele Rucci, Mor Ben-Tov, Yam Kushinsky, Alik Mokeichev, and Ronen Segev. Decorrelation of retinal response to natural scenes by fixational eye movements. *Proceedings of the National Academy of Sciences*, 112(10):3110–5, 2015.

[197] Santhosh Sethuramanujam, Amanda J. McLaughlin, Geoffery DeRosenroll, Alex Hog-

garth, David J. Schwab, and Gautam B. Awatramani. A Central Role for Mixed Acetylcholine/GABA Transmission in Direction Coding in the Retina. *Neuron*, 90(6):1243–1256, 2016.

[198] Michael N. Shadlen and William T. Newsome. Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4(4):569–579, 1994.

[199] Michael N. Shadlen and William T. Newsome. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *The Journal of Neuroscience*, 18(10):3870–96, 1998.

[200] Maoz Shamir. Emerging principles of population coding: In search for the neural code. *Current Opinion in Neurobiology*, 25:140–148, apr 2014.

[201] Maoz Shamir and Haim Sompolinsky. Nonlinear population codes. *Neural Computation*, 16(6):1105–1136, 2004.

[202] Maoz Shamir and Haim Sompolinsky. Implications of neuronal diversity on population coding. *Neural Computation*, 18(8):1951–86, 2006.

[203] C. E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[204] Robert Shapley. Linear and Nonlinear Systems Analysis of the Visual System: Why does it seem so linear? *Vision Research*, 49(9):907–921, 2010.

[205] Robert Shapley and V. Hugh Perry. Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences*, 9:229–235, 1986.

[206] Robert M Shapley and Jonathan D Victor. The effect of contrast on the transfer properties of cat retinal ganglion cells. *The Journal of Physiology*, 285:275–98, 1978.

[207] Andrew A. Sharp, Michael B. O'Neil, Larry F. Abbott, and Eve Marder. Dynamic clamp: computer-generated conductances in real neurons. *Journal of neurophysiology*, 69(3):992–995, 1993.

[208] Tatyana O Sharpee. Computational Identification of Receptive Fields. *Annual Review of Neuroscience*, 36(1):103–120, 2013.

[209] Tatyana O Sharpee, Kenneth D Miller, and Michael P Stryker. On the importance of static nonlinearity in estimating spatiotemporal neural filters with natural stimuli. *Journal of neurophysiology*, 99(5):2496–509, 2008.

[210] Marion Silies, Daryl M. Gohl, and Thomas R. Clandinin. Motion-Detecting Circuits in Flies: Coming into View. *Annual Review of Neuroscience*, 37(1):307–327, 2014.

[211] Eero P. Simoncelli and Bruno A. Olshausen. Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24:1193–1216, 2001.

[212] Darragh Smyth, Ben Willmore, Gary E Baker, Ian D Thompson, and David J Tolhurst. The Receptive-Field Organization of Simple Cells in Primary Visual Cortex of Ferrets under Natural Scene Stimulation. *The Journal of Neuroscience*, 23(11):4746–4759, 2003.

[213] Haim Sompolinsky, Hyoungsoo Yoon, Kukjin Kang, and Maoz Shamir. Population coding in neuronal systems with correlated noise. *Physical Review E*, 64, 2001.

[214] M. V. Srinivasan, S. B. Laughlin, and A. Dubs. Predictive coding: a fresh view of

inhibition in the retina. *Proceedings of the Royal Society of London B*, 216(1205):427–59, 1982.

[215] Charlene Stone and Lawrence H. Pinto. Response properties of ganglion cells in the isolated mouse retina. *Visual Neuroscience*, 10(1):31–39, 1993.

[216] Wenzhi Sun, Qiudong Deng, William R. Levick, and Shigang He. ON direction-selective ganglion cells in the mouse retina. *The Journal of Physiology*, 576(1):197–202, 2006.

[217] W. Rowland Taylor. TTX attenuates surround inhibition in rabbit retinal ganglion cells. *Visual Neuroscience*, 16(2):285–90, 1999.

[218] W. Rowland Taylor, Shigang He, William R. Levick, and David I. Vaney. Computation of Direction Selectivity By Retinal Ganglion Cells. *Science*, 289(2000):2347–2350, 2000.

[219] Frederic E. Theunissen, Steven V. David, Nandini C. Singh, Anne Hsu, William E. Vinje, and Jack L. Gallant. Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network: Computation in Neural Systems*, 12(3):289–316, 2001.

[220] M. G. A. Thomsom. Visual coding and the phase structure of natural scenes. *Network: Computation in Neural Systems*, 10:123–132, 1999.

[221] Jon Touryan, Gidon Felsen, and Yang Dan. Spatial structure of complex cell receptive fields measured with natural images. *Neuron*, 45(5):781–791, 2005.

[222] Stuart Trenholm, Kyle Johnson, Xiao Li, Robert G. Smith, and Gautam B. Awatramani. Parallel mechanisms encode direction in the retina. *Neuron*, 71(4):683–694, 2011.

[223] Stuart Trenholm, Amanda J. McLaughlin, David J. Schwab, and Gautam B. Awatramani. Dynamic tuning of electrical and chemical synaptic transmission in a network of motion coding retinal neurons. *The Journal of Neuroscience*, 33(37):14927–14938, 2013.

[224] Philipp Khuc Trong and Fred Rieke. Origin of correlated activity between parasol retinal ganglion cells. *Nature Neuroscience*, 11(11):1343–1351, 2008.

[225] Christopher W. Turner, Lina A J Reiss, and Bruce J. Gantz. Combined acoustic and electric hearing: Preserving residual acoustic hearing. *Hearing Research*, 242(1-2):164–171, 2008.

[226] Maxwell H. Turner and Fred Rieke. Synaptic Rectification Controls Nonlinear Spatial Integration of Natural Visual Inputs. *Neuron*, 90(6):1257–1271, 2016.

[227] Ian Van Der Linde, Umesh Rajashekar, Alan C. Bovik, and Lawrence K. Cormack. DOVES: a database of visual eye movements. *Spatial vision*, 22(2):161–77, 2009.

[228] J. H. van Hateren, R. Kern, G. Schwerdtfeger, and M. Egelhaaf. Function and Coding in the Blowfly H1 Neuron during Naturalistic Optic Flow. *The Journal of Neuroscience*, 25(17):4343–4352, 2005.

[229] J. Hans Van Hateren and A. van der Schaaf. Independent Component Filters of Natural Images Compared with Simple Cells in Primary Visual Cortex. *Proceedings of the Royal Society of London B*, 265(1394):359–366, 1998.

[230] David I Vaney, Benjamin Sivyer, and W Rowland Taylor. Direction selectivity in the

retina: symmetry and asymmetry in structure and function. *Nature Reviews Neuroscience*, 13(3):194–208, 2012.

[231] Jan Verweij, Eric P. Hornstein, and Julie L. Schnapf. Surround antagonism in macaque cone photoreceptors. *The Journal of Neuroscience*, 23(32):10249–10257, 2003.

[232] Jonathan D Victor and Robert Shapley. Receptive field mechanisms of cat X and Y retinal ganglion cells. *The Journal of General Physiology*, 74(2):275–298, 1979.

[233] Jonathan D. Victor, Robert M. Shapley, and Bruce W. Knight. Nonlinear analysis of cat retinal ganglion cells in the frequency domain. *Proceedings of the National Academy of Sciences*, 74(7):3068–3072, 1977.

[234] William E. Vinje and Jack L. Gallant. Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science*, 287:1273–1276, 2000.

[235] William E. Vinje and Jack L. Gallant. Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. *The Journal of Neuroscience*, 22(7):2904–2915, 2002.

[236] Brett Vintch, J. Anthony Movshon, and Eero P. Simoncelli. A Convolutional Subunit Model for Neuronal Responses in Macaque V1. *The Journal of Neuroscience*, 35(44):14829–41, 2015.

[237] Brett Vintch, Andrew D Zaharia, J Anthony Movshon, and Eero P Simoncelli. Efficient and direct estimation of a neural subunit model for sensory coding. *Neural Information Processing Systems*, 25:1–9, 2012.

[238] Heinz Wässle. Parallel processing in the mammalian retina. *Nature Reviews Neuroscience*, 5(10):747–57, 2004.

[239] Shijun Weng, Wenzhi Sun, and Shigang He. Identification of ONOFF direction-selective ganglion cells in the mouse retina. *The Journal of Physiology*, 5623(3):915–923, 2005.

[240] Frank S. Werblin. Six different roles for crossover inhibition in the retina: correcting the nonlinearities of synaptic transmission. *Visual Neuroscience*, 27:1–8, 2010.

[241] W. Bryan Wilent and Diego Contreras. Stimulus-Dependent Changes in Spike Threshold Enhance Feature Selectivity in Rat Barrel Cortex Neurons. *The Journal of Neuroscience*, 25(11):2983–2991, 2005.

[242] Stefan D Wilke and Christian W Eurich. Representational accuracy of stochastic neural populations. *Neural Computation*, 14(1):155–189, 2002.

[243] Ben D B Willmore, James a Mazer, and Jack L Gallant. Sparse coding in striate and extrastriate visual cortex. *Journal of Neurophysiology*, 105:2907–2919, 2011.

[244] Stefan Winkler and Ramanathan Subramanian. Overview of Eye Tracking Datasets. *Fifth International Workshop on Quality of Multimedia Experience*, pages 212–217, 2013.

[245] Si Wu, Shun Ichi Amari, and Hiroyuki Nakahara. Information processing in a neuron ensemble with the multiplicative correlation structure. *Neural Networks*, 17(2):205–214, 2004.

[246] Chun-I Yeh, Dajun Xing, Patrick E. Williams, and Robert M. Shapley. Stimulus

ensemble and cortical layer determine V1 spatial receptive fields. *Proceedings of the National Academy of Sciences*, 106(34):14652–14657, 2009.

[247] Ehud Zohary, Michael N. Shadlen, and William T. Newsome. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature*, 370(6485):140–3, 1994.

[248] Joel Zylberberg, Alexandre Pouget, Peter E. Latham, and Eric Shea-Brown. Robust information propagation through noisy neural circuits. *PLoS Computational Biology*, 13(4):e1005497, 2017.