



1-1-2013

# Adaptation of the Retina to Stimulus Correlations

Kristina D. Simmons

University of Pennsylvania, [kristina.d.simmons@gmail.com](mailto:kristina.d.simmons@gmail.com)

Follow this and additional works at: <http://repository.upenn.edu/edissertations>

 Part of the [Neuroscience and Neurobiology Commons](#)

---

## Recommended Citation

Simmons, Kristina D., "Adaptation of the Retina to Stimulus Correlations" (2013). *Publicly Accessible Penn Dissertations*. 800.  
<http://repository.upenn.edu/edissertations/800>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/edissertations/800>  
For more information, please contact [libraryrepository@pobox.upenn.edu](mailto:libraryrepository@pobox.upenn.edu).

---

# Adaptation of the Retina to Stimulus Correlations

## **Abstract**

Visual scenes in the natural world are highly correlated. To efficiently encode such an environment with a limited dynamic range, the retina ought to reduce correlations to maximize information. On the other hand, some redundancy is needed to combat the effects of noise. Here we ask how the degree of redundancy in retinal output depends on the stimulus ensemble. We find that retinal output preserves correlations in a spatially correlated stimulus but adaptively reduces changes in spatio-temporal input correlations. The latter effect can be explained by stimulus-dependent changes in receptive fields. We also find evidence that horizontal cells in the outer retina enhance changes in output correlations. GABAergic amacrine cells in the inner retina also enhance differences in correlation, albeit to a lesser degree, while glycinergic amacrine cells have little effect on output correlation. These results suggest that the early visual system is capable of adapting to stimulus correlations to balance the challenges of redundancy and noise.

## **Degree Type**

Dissertation

## **Degree Name**

Doctor of Philosophy (PhD)

## **Graduate Group**

Neuroscience

## **First Advisor**

Vijay Balasubramanian

## **Keywords**

adaptation, decorrelation, efficient coding, lateral inhibition, retina

## **Subject Categories**

Neuroscience and Neurobiology

ADAPTATION OF THE RETINA TO STIMULUS CORRELATIONS

Kristina Diane Simmons

A DISSERTATION

in

Neuroscience

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2013

Supervisor of Dissertation

---

Vijay Balasubramanian, Professor of Physics and Astronomy

Graduate Group Chairperson

---

Joshua I. Gold, Associate Professor of Neuroscience

Dissertation Committee

David H. Brainard, Professor of Psychology

Joshua I. Gold, Associate Professor of Neuroscience

Larry A. Palmer, Professor of Neuroscience

Michael J. Berry II, Associate Professor of Molecular Biology, Princeton University

# Acknowledgements

First and foremost, I want to thank my advisor, Vijay Balasubramanian, for his support and advice throughout my time in graduate school. Vijay has not only helped shape my scientific thinking but has also motivated me during the hardest parts of graduate school and encouraged me to follow my own passion. Thanks to Josh Gold for his help as chair of my thesis committee and as chair of the Neuroscience Graduate Group. I have also benefited greatly from the suggestions and guidance of the rest of my thesis committee — Michael Berry, David Brainard, and Larry Palmer. I want to thank the Neuroscience Graduate Group, in particular Jane Hoshi for her constant assistance and Mikey Nusbaum for his advice both as graduate group chair and afterwards.

I have been lucky to be part of a collaborative group of retina researchers at Penn. Thanks to Peter Sterling for his advice early in my graduate career and for his long-lasting influence on the lab. Michael Freed, Rob Smith, and Noga Vardi have been an amazing source of knowledge about all aspects of the retina and have challenged me to think about my research from different perspectives. For this I am extremely grateful. I also want thank Phil Nelson for many fruitful discussions.

Many past and current members of the retina lab have contributed to my research and education. In particular, Jason Prentice has been indispensable at every stage and contributed equally to the work in Chapter 2. Thanks to Bart Borghuis for first

introducing me to retinal electrophysiology and sharing the excitement of my first recordings, Gašper Tkačik for many helpful discussions and consistently insightful feedback, and Zhiyin Liang for advice on all aspects of my experiments. Marie Fina has made my research experience smoother in countless ways, and Sergei Nikonov made many substantial improvements to the experimental equipment. Thanks to Jan Homann and John Briguglio for many engaging conversations. I also thank Ann Hermundstad and David Schwab for interesting discussions about the implications of my research.

I would not have been able to succeed in graduate school without the support of my classmates and friends. Thanks to my college friends — especially Julia Cohen, Saurav Dhital, and Michelle Tomasik — for their constant support and friendship when I have needed it most. Thanks to Emily Chen, Meng-meng Fu, Olivia Perng, Melissa Wang, and Lindsay Yurcaba for many great times and for being there throughout the ups and downs of the past six years. Thanks to Stephanie Majkut, Maria Murray, and Ann Vanleer for sharing in my successes as well as my frustrations. Special thanks to Garth Baughman for the innumerable pep talks and, especially, for his patience with me during the last stages of my thesis work.

I also want to thank Paul Katz, Jim Knierim, and all of my teachers and classmates at the Marine Biological Laboratory for giving me great memories along with a wider understanding of neuroscience and for helping me gain clarity about my future.

Finally, I am forever indebted to my parents for their unwavering love and encouragement. I would not be here without their support.

*My training was supported by the Systems and Integrative Biology (NIH T32 GM 7517-31), Computational Neuroscience (NIH T90 DA 22763-4), and Vision (NIH T32 EY 7035-33) Training Grants and by the Hearst Foundation.*

# ABSTRACT

## ADAPTATION OF THE RETINA TO STIMULUS CORRELATIONS

Kristina Diane Simmons

Vijay Balasubramanian

Visual scenes in the natural world are highly correlated. To efficiently encode such an environment with a limited dynamic range, the retina ought to reduce correlations to maximize information. On the other hand, some redundancy is needed to combat the effects of noise. Here we ask how the degree of redundancy in retinal output depends on the stimulus ensemble. We find that retinal output preserves correlations in a spatially correlated stimulus but adaptively reduces changes in spatio-temporal input correlations. The latter effect can be explained by stimulus-dependent changes in receptive fields. We also find evidence that horizontal cells in the outer retina enhance changes in output correlations. GABAergic amacrine cells in the inner retina also enhance differences in correlation, albeit to a lesser degree, while glycinergic amacrine cells have little effect on output correlation. These results suggest that the early visual system is capable of adapting to stimulus correlations to balance the challenges of redundancy and noise.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Overview of the retina . . . . .	1
1.1.1	Organization . . . . .	2
1.1.2	Ganglion cell responses . . . . .	3
1.2	Adaptation in the retina . . . . .	3
1.2.1	Adaptation to contrast . . . . .	4
1.2.2	Adaptation to complex stimulus features . . . . .	5
1.3	Natural scenes and correlations . . . . .	5
1.3.1	Statistics of natural scenes . . . . .	6
1.3.2	Visual responses to natural stimuli . . . . .	6
1.3.3	Correlations in retinal output . . . . .	8
1.4	Efficient coding . . . . .	9
1.4.1	Theories of efficient coding . . . . .	9
1.4.2	Efficient coding in sensory systems . . . . .	11
<b>2</b>	<b>Transformation of stimulus correlations by the retina</b>	<b>14</b>
2.1	Abstract . . . . .	14
2.2	Introduction . . . . .	15
2.3	Results . . . . .	18

2.3.1	Simultaneous measurements of ganglion cell responses . . . . .	18
2.3.2	Output correlations in response to different stimuli . . . . .	21
2.3.3	Adaptation of temporal filters . . . . .	28
2.3.4	Adaptation of spatial receptive fields . . . . .	33
2.3.5	Adaptation of nonlinearity gain . . . . .	35
2.3.6	Output correlations in an adapting model . . . . .	36
2.4	Discussion . . . . .	37
2.5	Methods . . . . .	41
<b>3</b>	<b>Lateral inhibition maintains differences in stimulus correlations</b>	<b>58</b>
3.1	Abstract . . . . .	58
3.2	Introduction . . . . .	59
3.2.1	Outer retina mechanisms . . . . .	59
3.2.2	Inner retina mechanisms . . . . .	63
3.2.3	Flexibility in retinal circuitry . . . . .	65
3.3	Results . . . . .	69
3.3.1	Feedback in outer retina enhances stimulus correlations . . . . .	69
3.3.2	Inner retina may enhance extreme stimulus correlations . . . . .	77
3.4	Discussion . . . . .	78
3.5	Methods . . . . .	81
<b>4</b>	<b>General discussion</b>	<b>84</b>
4.1	Summary . . . . .	84
4.2	Studying natural vision . . . . .	85
4.3	Adaptation to complex stimulus features . . . . .	87
4.4	Efficient coding of correlated stimuli . . . . .	89
4.5	Functional classification of ganglion cells . . . . .	90



4.6	Population models of retinal correlations . . . . .	92
4.7	Concluding remarks . . . . .	93
<b>A</b>	<b>Identifying spikes in multielectrode array data</b>	<b>95</b>
A.1	Abstract . . . . .	95
A.2	Introduction . . . . .	96
A.3	Results . . . . .	99
A.3.1	Preliminary visualization of our data . . . . .	101
A.3.2	Summary of our method . . . . .	103
A.3.3	Tests of our method . . . . .	104
A.4	Discussion . . . . .	111
A.5	Methods . . . . .	112
A.5.1	STEP 1: Data preparation and segmentation . . . . .	113
A.5.2	STEP 2: Cluster identification and template building . . . . .	115
A.5.3	STEP 3: Spike fitting . . . . .	118
A.5.4	STEP 4: Evaluation of template reliability . . . . .	127

# List of Tables

2.1	Number of cells recorded in each condition . . . . .	25
2.2	Model validation of receptive field analysis . . . . .	49
2.3	Variants on receptive field analysis method . . . . .	54

# List of Figures

2.1	Natural and artificial stimuli vary in correlation structure . . . . .	17
2.2	Retinal ganglion cell receptive fields measured using a multi-electrode array . . . . .	20
2.3	Retinal output correlations are largely constant between stimulus con- ditions . . . . .	22
2.4	Analysis of pairwise correlations . . . . .	26
2.5	Adaptation of the linear temporal filter. . . . .	30
2.6	Adaptation of the spatio-temporal receptive field and gain . . . . .	34
2.7	Correlation measurement controls . . . . .	46
2.8	Receptive field results are validated with model neurons and are robust to analysis method changes . . . . .	52
3.1	HEPES removes stimulus dependence of output correlations . . . . .	70
3.2	HEPES does not abolish light responses . . . . .	72
3.3	Effects of HEPES on information . . . . .	75
3.4	Blocking GABA <sub>A</sub> receptors may reduce changes in output correlation	78
A.1	Recording chamber and typical data . . . . .	99
A.2	Visualization of data and flowchart of our method . . . . .	102
A.3	Tests of the method . . . . .	107

A.4	Template fitting to complex events . . . . .	110
A.5	Template building . . . . .	119
A.6	After fitting spikes, only noise remains . . . . .	124

# Chapter 1

## Introduction

In this thesis, I seek to address two questions. First, does the retina adapt to stimulus correlations? Second, seeing that such adaptation occurs, where in the retina does it arise? In this chapter, I review the basic structure and function of the retina. I then discuss what we know about the retina's ability to adapt to stimuli and put forward a motivation for this thesis using the framework of efficient coding theory. In [Chapter 2](#), I show that correlations in retinal output differ less than correlations in input and explain this finding as a result of receptive field adaptation. In [Chapter 3](#), I examine the circuitry underlying these effects and suggest a role for the outer retina in combating noise. Finally, in [Chapter 4](#), I discuss my findings in the context of visual processing as we currently understand it and suggest a number of directions for further study.

### 1.1 Overview of the retina

While there are many open questions about retinal circuitry and functional, its basic structure and output are well understood. Here we provide a general introduction.

### 1.1.1 Organization

The retina has a highly organized laminar structure: incoming light is detected by photoreceptors, which synapse onto bipolar cells in the outer plexiform layer. Bipolar cells then synapse onto retinal ganglion cells (RGCs) in the inner plexiform layer. In each of these stages, the feedforward signals are shaped by inhibition from horizontal cells in the outer retina and amacrine cells in the inner retina. Details of each connection will be discussed in [Chapter 3](#); here we focus on the general structure and the nature of signals as they travel through the retina.

The retina is able to achieve a great deal of image processing due in part to the large number of cell subtypes. In the outer retina, photoreceptors come in two varieties. Rods are responsible for vision in low light conditions, while cones allow vision at higher background light levels. In most mammals, there are two types of cones that differ in their spectral sensitivity, thus forming the basis of color vision. Humans and other primates have a third type of cone that gives us red-green color vision. Rods, on the other hand, only come in one subtype but are sensitive enough to detect a single photon. All photoreceptors are depolarized in the dark and hyperpolarized in response to light. At the first synapse, from photoreceptor to bipolar cell, the signal is split into two pathways. ON bipolar cells depolarize in response to light onset, whereas OFF bipolar cells hyperpolarize. The opposite happens at light offset: ON cells hyperpolarize and OFF cells depolarize. Bipolar cells are further subdivided into approximately 10 subtypes with different functions and synaptic connections ([Masland, 2001](#)). Horizontal cells come in only two types, one that feeds back onto cones only and one that sends feedback to cones and rods ([Masland, 2001](#)). As with bipolar cells, there are between 10 and 15 types of ganglion cells ([Wassle, 2004](#); [Roska et al., 2006](#)). Many of these maintain the division into ON and OFF pathways, but

some ON-OFF ganglion cells respond to both light onset and light offset. There are also about 30 types of amacrine cell that enhance the specific functions of each RGC type. This diversity of bipolar, amacrine, and ganglion cells gives the retina a great deal of flexibility in responding to a large number of stimulus features.

### 1.1.2 Ganglion cell responses

Different types of retinal ganglion cells respond very differently to stimuli, depending on their inputs and intrinsic properties (Rockhill et al., 2002; Roska et al., 2006; Segev et al., 2006). Within ON and OFF types, many RGCs can be grouped by how quickly they respond to a stimulus and how long their responses last. For example, brisk-transient cells respond quickly with a small number of spikes, whereas brisk-sustained cells also have a short latency but respond with a longer train of action potentials. Sluggish cells in turn have a slower response time. Other cells have a more specialized purpose: direction-selective cells respond only to a stimulus moving in a specific direction, while local edge detectors respond to borders but not to spatially uniform stimuli. On the other hand, uniformity detectors respond well to a constant stimulus and are silent following a change in stimulus.

## 1.2 Adaptation in the retina

Using the general structure and functions discussed above, the retina is able to adapt its processing to a variety of stimulus features. In this section, we review such adaptation.

### 1.2.1 Adaptation to contrast

One of the most commonly studied forms of adaptation in the visual system is adaptation to contrast, or the range of intensities in a stimulus ensemble. When stimulus contrast increases, ganglion cell firing rates increase at first and then gradually drop (Smirnakis et al., 1997). Similarly, some ganglion cells show depolarization shortly after a contrast increment, followed by slow hyperpolarization (Baccus and Meister, 2002). This adaptation can be examined in more detail by fitting responses to a linear-nonlinear model (Chichilnisky, 2001). In this model, the visual stimulus is filtered with a linear kernel and then the output is passed through a nonlinear transfer function to generate a predicted firing rate (or membrane current if applicable). The linear filter represents the spatio-temporal receptive field of the cell, and the nonlinearity encompasses any thresholding and saturation. The cell’s sensitivity to stimuli can be incorporated into either stage as the amplitude of the kernel or the slope of the nonlinearity. In this thesis, to avoid ambiguity, we will always take the linear filters to be normalized so that all control of sensitivity is in the gain, or slope, of the nonlinear transfer function. Typically, one will fit responses from different adaptation states separately and compare the estimated filters and nonlinearities between adaptation states.

After an increase in contrast, ganglion cell filters become faster and more biphasic (Baccus and Meister, 2002; Demb, 2002; Smirnakis et al., 1997; Beaudoin et al., 2007; Kim and Rieke, 2001; Chander and Chichilnisky, 2001; Zaghloul et al., 2005). In other words, cells are more high-pass when signal strength is higher. When contrast (and hence signal strength) is lower, filters are slower and ganglion cells integrate over a longer time to counteract the effects of noise in the signal. In addition, the gain decreases when contrast is higher, allowing the cell to use its full dynamic range



without saturating (Baccus and Meister, 2002; Demb, 2002; Beaudoin et al., 2007; Kim and Rieke, 2001; Zaghloul et al., 2005).

### 1.2.2 Adaptation to complex stimulus features

Adaptation to stimulus features more complicated than contrast has been observed but is less fully understood. Hosoya et al. (2005) found adaptation in retinal ganglion cell filters when the stimulus alternated between horizontal and vertical gratings, but they did not explore the nature of the changes. RGCs also adapt to spatial scale. After a switch from a spatially uniform stimulus to a spatially modulated checkerboard, some RGC firing rates immediately increase and then gradually decrease (Smirnakis et al., 1997), and receptive fields display adaptation (Hosoya et al., 2005). The retina can also adapt to temporal patterns. When Schwartz et al. (2007) displayed a series of identical flashes, RGCs displayed high firing rates after the sequence ended. This omitted stimulus response occurred at the time when another flash would be expected and was stronger than responses to each of the flashes in the series. Thus, the retina is able to adapt to a pattern of stimuli, as well as orientation and spatial scale. However, the mechanisms and circuitry underlying such phenomena have not been identified.

## 1.3 Natural scenes and correlations

Natural scenes, unlike commonly used artificial stimuli, are highly correlated. This striking feature leads us to ask whether the retina can adapt to stimulus correlations. Here we review general properties of natural scenes and the current state of knowledge about natural vision.

### 1.3.1 Statistics of natural scenes

Natural visual stimuli vary greatly in content but typically have a scale-invariant structure (for a review, see [Geisler, 2008](#); [Simoncelli and Olshausen, 2001](#)). That is, in a given image, the power spectrum does not change depending on the magnification or viewing distance ([Ruderman, 1997](#); [Lee et al., 2001](#); [Balboa et al., 2001](#)). Specifically, across natural scenes, the amplitude spectrum is approximately of the form  $1/f$ , where  $f$  is the spatial or temporal frequency ([Burton and Moorhead, 1987](#); [Field, 1987](#); [Ruderman and Bialek, 1994](#)). The relatively high power at low spatial frequencies may be due to nearby areas of a single object having a similar appearance. The fact that illumination is usually somewhat constant (in space) within a scene is also likely to induce large-scale correlations.

In addition to being highly correlated, natural images have a luminance distribution with a strong positive skew. Many pixels are darker than the mean luminance and a smaller number are very bright ([Brady and Field, 2000](#); [Laughlin, 1981](#); [Ruderman et al., 1998](#)). Local contrast is generally of low magnitude ([Tadmor and Tolhurst, 2000](#)), presumably due to correlations between nearby points in the same object. In addition, contrast and luminance are largely independent ([Frazor and Geisler, 2006](#)).

### 1.3.2 Visual responses to natural stimuli

Many studies have suggested that visual responses change as a result of changes in stimulus correlation structure. In the fly eye, [de Ruyter van Steveninck et al. \(1997\)](#) found that responses of the motion-sensitive neuron H1 to naturalistic motion could not be predicted well by models based on responses to random stimuli, hinting at a change in the basic processing in a fly eye depending on the stimulus.

In vertebrates, [Nirenberg et al. \(2010\)](#) found that some retinal ganglion cells

whiten naturalistic stimuli. In addition, RGCs show only a weak surround when probed with white noise (Chichilnisky and Kalmar, 2002), in contrast to the strong surround measured by Derrington and Lennie (1982) using the more highly correlated sine-wave gratings. This finding suggests that spatial correlations may already shape receptive fields in retina.

Output of the lateral geniculate nucleus (LGN) of the thalamus is also whitened in response to natural movies but has increased high-frequency power when presented with white noise (Dan et al., 1996). Specifically, Lesica et al. (2007) recorded from LGN in response to uncorrelated white noise stimuli and correlated natural scene movies. They found that responses are faster and have lower gains and stronger surrounds when probed with natural scenes (Lesica et al., 2007). Reliability and sparseness also increase with stimulus correlations.

Similarly, Sharpee et al. (2006) measured responses of simple cell in primary visual cortex (V1) to white noise and to natural scenes. The receptive fields measured with each stimulus were better at predicting the response to that stimulus than to the other ensemble. Moreover, receptive fields measured under natural stimuli had increased sensitivity at high frequencies relative to white noise receptive fields, thus compensating for the low stimulus power at these frequencies. This indicates that V1 can adapt to counteract changes in stimulus correlation structure. These differences are stronger in the inhibitory portions of the receptive fields, suggesting that the difference in stimulus structure has an effect on how the receptive fields act to reduce redundancy (David et al., 2004). Thus, individual neurons in the early visual system are able to adapt to natural scenes in a way that whitens the power spectra. These results lead us to ask whether the retina also adapts to such stimuli on a population level.

### 1.3.3 Correlations in retinal output

Retinal ganglion cells often show correlations in their spike trains. Seminal work by [Mastronarde \(1983a,b,c, 1989\)](#) characterized these correlations in cat X- and Y-type cells. In the absence of any modulated light stimulus, pairs of cells have narrow peaks and wells in their cross-correlograms: a cell is more likely to fire if a nearby cell of the same polarity fires and less likely to fire when a nearby cell of the opposite polarity fires. These correlations appear to be mainly a result of common input, possibly from spiking amacrine cells, but may also include a contribution from reciprocal connections between ganglion cells. At low light levels, correlations with less temporal precision are also seen (similar results have been measured in primate by [Greschner et al., 2011](#)). These broad correlations are stronger between pairs of ON cells than between pairs of OFF cells and may depend in part on tonic signals from bipolar cells. Thus, it appears that correlations are stimulus-dependent and may depend on multiple sites within the retinal circuitry.

Similarly, rabbit RGCs have cross-correlograms with a single peak at zero, adding further support to Mastronarde’s conclusions that correlations are largely the result of shared input. Reciprocal connections among OFF cells may also be important, as OFF brisk-transient ganglion cells (analogous to Y-type cells) show two peaks in correlation separated by a few milliseconds ([DeVries, 1999](#)). When time-varying stimuli are presented, temporally broad correlations between neurons in primate retina increase ([Greschner et al., 2011](#)). The spatial extent of these correlations increases with the extent of stimulus correlation, consistent with the idea that common input contributes to correlations. At a fixed distance, however, the strength of correlations does not precisely reflect the degree of input correlation. Rather, output correlations are higher than stimulus correlations for white noise but lower than stimulus corre-

lations for naturalistic stimuli, resulting in similar output correlation strengths for the two stimulus categories (Pitkow and Meister, 2012). This stimulus dependence of correlations could be in part due to changes in the timing of single-cell responses relative to the stimulus such as those discussed above (Lesica et al., 2007).

Patterns in firing among more than two neurons have not been studied as extensively, in part due to the exponential increase in the number of possible patterns. However, computational models that incorporate pairwise correlations are able to account for patterns of groups of cells (Shlens et al., 2008). Thus, characterizing pairwise correlations may allow us to better understand structured responses of large populations of RGCs.

## 1.4 Efficient coding

Why would we expect to see adaptation to stimulus correlations? As described below, such adaptation would benefit an organism by reducing the redundancy of the visual signal. This change may be needed to accommodate environments with different correlation structures using the limited bandwidth of the optic nerve.

### 1.4.1 Theories of efficient coding

As a result of the strong correlations in natural scenes, visual signals are highly redundant, with most information contained at points in space and time where the stimulus changes (Attneave, 1954). However, the optic nerve has limited spatial and energetic resources and so representations of the stimulus must be efficient (Niven and Laughlin, 2008; Perge et al., 2009). Removing redundancy present in the stimulus is one way to increase the amount of information conveyed with a limited capacity (Atick, 1992).

It has been speculated that an optimal sensory system would have evolved in a way that accomplishes this task (Geisler, 2008; Barlow, 1961; Atick, 1992). Specifically, neurons ought to maximize the mutual information between their responses and the stimulus (Simoncelli and Olshausen, 2001). For example, a receptive field with an excitatory center and an inhibitory surround could contribute to redundancy reduction in sensory systems by decorrelating the signals of neighboring cells (Barlow, 1961). However, the optimal coding for a given environment depends on features of the stimulus ensemble (Simoncelli and Olshausen, 2001; Deweese, 1996). Thus, it would be advantageous for the retina to adapt its processing to maximize information in each environment.

As an example, consider contrast adaptation. The range of stimuli to which sensory systems are exposed is generally several orders of magnitude larger than the dynamic range of any encoding neuron. To process the entire stimulus space with a fixed coding scheme the neuron would have to either saturate for strong stimuli or lose sensitivity to small changes. A better solution is to adapt to the recent history of stimuli. If the variation of a stimulus increases, then a neuron encoding that stimulus should decrease its gain, thereby using its limited output range to encode the full range of input it is likely to receive. Similarly, environments with different spatial structures could dictate different coding strategies.

While redundancy reduction appears to be an important goal, there are other considerations in designing an optimal visual system. In particular, noise can obscure the signal from a single unit. One way to compensate is to transmit the same signal through multiple channels. The signal strength improves, while independent noise is averaged out. Thus, the optimal degree of correlation between neurons depends on the noise in the system (Tkačik et al., 2010). Similarly, in the temporal domain, an optimal sensory system should attenuate the highest frequencies to minimize noise

(Atick, 1992).

In addition, correlations in the noise present in nearby neurons can increase precision of spike timing. In midget ganglion cells responding to a repeated stimulus, trial-to-trial noise in excitatory and inhibitory currents is correlated (Cafaro and Rieke, 2010). Moreover, injecting these currents using dynamic clamp leads to spike trains that have higher reliability when the injected excitatory and inhibitory conductances are taken from a simultaneous recording rather than different recordings using the same stimulus. Thus, correlations in inputs may play an important role in shaping output. It is possible that the pattern of correlations in the retina similarly provides a signal to downstream neurons.

#### 1.4.2 Efficient coding in sensory systems

The predictions of efficient coding theory have been confirmed in a number of experimental systems. Much of this evidence demonstrates a match between stimulus contrast and neural dynamic range (for a review, see Wark et al., 2007). For example, in the fly eye, the contrast response function is very similar to the cumulative distribution of contrast values in natural scenes (Laughlin, 1981). That is, the gain is high for the range of contrasts that occurs frequently but low for rare contrasts. As a result, the distribution of responses to a natural stimulus is uniform, using the entire dynamic range effectively. Moreover, when a fly is exposed to a stimulus ensemble in which the range of velocities suddenly increases, the motion-sensitive neuron H1 adapts by decreasing its gain. More precisely, the H1 input-output transfer function in each condition is scaled by the standard deviation of the input velocity ensemble. After compensating for this scaling, the transfer functions for the two stimulus velocities align (Brenner et al., 2000; Fairhall et al., 2001). The gain used by this system

is the one that maximizes the entropy of the output firing rates. As the noise entropy does not change between conditions, this amounts to maximizing the information. The first synapse in the fly eye also seems to adapt to intensity changes in a way that optimally reduces redundancy and noise (Laughlin, 1994).

In mammals, contrast responses in LGN and V1 appear to match the range of contrasts seen in the natural world (Tadmor and Tolhurst, 2000). Contrast adaptation is independent of luminance adaptation, matching the independence of the two measurements in natural scenes (Mante et al., 2005). In rat barrel cortex, a change in the distribution of vibrissa displacements leads to adaptation such that input-output curves are scaled by standard deviation and, as a result, information transmission is equal for the two variances (Maravall et al., 2007). Contrast adaptation could be relevant to the processing of stimulus correlations in that variance normalization can lead to statistical independence (Schwartz et al., 2001). Consider two neurons  $A$  and  $B$ , where the variance of neuron  $A$ 's response is greater when neuron  $B$  has a high average response. These two neurons could be uncorrelated on average but clearly still convey redundant information. Their independence would be increased if  $A$ 's responses were normalized by their variance (Simoncelli and Schwartz, 1999). Such a transformation is thought to underlie contrast gain control, thus implicating decorrelation as a possible purpose of contrast adaptation.

Direct measures of neural processing of stimulus correlations also appear consistent with the redundancy reduction hypothesis. In particular, classical measurements of retinal receptive fields show a center-surround structure (Kuffler, 1953). This structure can be modeled as a difference of two Gaussians: a narrow, excitatory center with a broader, inhibitory surround (Enroth-Cugell and Robson, 1966; Rodieck, 1965). The surround is thought to reduce redundancy by subtracting the mean luminance from the intensity in the center so that the visual system encodes deviations from the mean



(Barlow, 1961). Srinivasan et al. (1982) and Atick (1992) computed theoretical receptive fields that would remove the spatio-temporal correlations in natural images. These predictions approximate the structure of visual receptive fields, including surround inhibition. Moreover, receptive fields at different light levels can be predicted by considering the effect of the signal-to-noise ratio on the optimal balance between redundancy reduction and noise reduction (Atick et al., 1990; van Hateren, 1992). As background intensity increases, the visual signal becomes stronger and noise reduction is less of a concern. Hence, the dominant effect is to remove redundancy by increasing surround strength. One might also predict that correlations decrease with light level, and this turns out to be the case (Greschner et al., 2011). Downstream from the retina, receptive fields adapt to naturalistic stimuli, as discussed above. Such adaptation increases mutual information between the stimulus and the response (Sharpee et al., 2006).

Other features of the retina are consistent with information theoretic predictions as well. For instance, the greater number of OFF cells gives the retina more resources for processing negative contrast, which is more common than positive contrast (Balasubramanian and Sterling, 2009), and the level of overlap observed in dendritic fields is optimal for maximizing information (Borghuis et al., 2008). Cell types vary in noise level and information rate, but these two features compensate so that all types have similar coding efficiency (Koch et al., 2006). Such examples suggest that efficient coding is a driving force in the evolution of the visual system.

Thus, various lines of evidence lead us to believe that the retina may be able to adapt its processing to optimally encode stimuli with different correlation structures.

# Chapter 2

## Transformation of stimulus correlations by the retina

The material in this chapter has been submitted for publication as [Simmons et al. \(2013\)](#).

### 2.1 Abstract

Redundancies and correlations in the responses of sensory neurons seem to waste neural resources but can carry cues about structured stimuli and may help the brain to correct for response errors. To assess how the retina negotiates this tradeoff, we measured simultaneous responses from populations of ganglion cells presented with natural and artificial stimuli that varied greatly in correlation structure. We found that pairwise correlations in the retinal output remained similar across stimuli with widely different spatio-temporal correlations including white noise and natural movies. Responding to more correlated stimuli, ganglion cells had faster temporal kernels and tended to have stronger surrounds. These properties of individual cells,

along with gain changes that opposed changes in effective contrast at the ganglion cell input, largely explained the similarity of pairwise correlations across stimuli. Meanwhile, purely spatial stimulus correlations tended to evoke concomitant increases in response correlations.

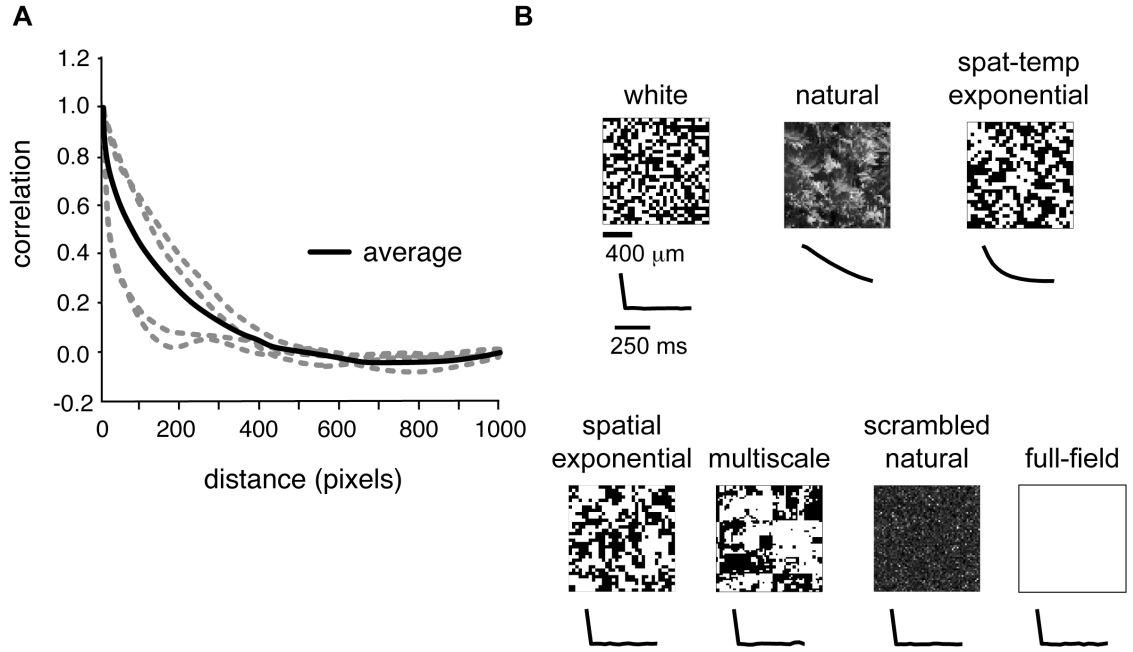
## 2.2 Introduction

An influential theory of early sensory processing argues that sensory circuits should conserve scarce resources in their outputs by removing correlations present in their inputs (Barlow, 1961; Srinivasan et al., 1982; Atick, 1992). However, recent work has clarified that some redundancy in retinal output is useful for hedging against noise (Borghuis et al., 2008; Tkačik et al., 2010). Moreover, sensory outputs with varying amounts of correlation can engage cortical circuits differently and thus result in a different sensory “code” (Estebanez et al., 2012). Thus, some degree of redundancy appears to be useful to the brain when dealing with response variability and making decisions based on probabilistic input (Barlow, 2001). Indeed, correlations between neurons in visual cortex are largely unchanged between unstructured and naturalistic visual stimuli (Fiser et al., 2004). We therefore hypothesized that *retina may adjust to the spatio-temporal structure of stimuli not to decorrelate but to maintain a relatively invariant degree of output correlation*. Previous studies have examined pairwise correlations amongst retinal ganglion cell spike trains in specific stimulus conditions (Ganmor et al., 2011; Puchalla et al., 2005; Schneidman et al., 2003; Trong and Rieke, 2008; Greschner et al., 2011; Pitkow and Meister, 2012) but did not study the changes in correlation for the same pairs across stimuli.

Are there mechanisms that might allow the retina to adjust its functional properties when stimulus correlations change? Traditionally, retinal ganglion cells (RGCs)

have been described by a fixed linear receptive field followed by a static nonlinearity (Rodieck, 1965), where surround inhibition acts linearly to suppress pairwise correlations in natural visual input (Srinivasan et al., 1982; Atick, 1992). In this view, the receptive field and nonlinearities might vary dynamically with stimulus correlations, possibly by changing the strength of lateral inhibition to maintain a fixed amount of output correlation. Indeed, correlation-induced changes in receptive fields have been observed in the LGN and visual cortex (Lesica et al., 2007; Sharpee et al., 2006).

To test these ideas, we performed a series of experiments in which we presented the retina with several stimuli with varying degrees of spatial and temporal correlations. The retina never fully decorrelated its input; even for white noise stimuli some correlations were present between pairs of retinal ganglion cell spike trains. Responding to natural movies, however, output correlations were only slightly greater than they were while responding to white noise, despite the dramatic difference in input correlations. We found a similar result for a spatio-temporal exponentially correlated stimulus, with the increase in output correlations being smaller still. For stimuli with high spatial (but not temporal) correlations, output correlation increased with input correlation to a larger degree than in natural movies. Thus, pairwise output correlations are similar over a broad range of spatio-temporal correlations but increase with spatial correlation in the absence of temporal correlation. Additionally, we observed a faster response timecourse and a skew towards stronger inhibitory surrounds in response to correlated stimuli. These changes were sufficient to largely explain the observed suppression of pairwise correlations in the retinal output.



**Figure 2.1: Natural and artificial stimuli vary in correlation structure.** (A) Spatial correlation functions from four natural images, in *gray*. *Black line* shows average correlation function over a large database of natural images. Although all images' correlation functions have the same general shape, there are clear differences between images. (B) Examples of the stimuli used in this work. Traces above frames show the temporal correlation function of each stimulus. Stimuli were displayed at 30 Hz in alternating 10-minute blocks.

## 2.3 Results

### 2.3.1 Simultaneous measurements of ganglion cell responses

We used a multi-electrode array to measure simultaneous responses from groups of  $\sim 40$  retinal ganglion cells in guinea pig. Each recording interleaved 10-minute blocks of white noise checkerboard stimuli with 10-minute blocks of correlated stimuli. Example frames from each stimulus are shown in Figure 2.1B, together with their respective temporal correlation functions. We probed retinal responses to natural movies, which allowed us to determine properties of ganglion cell population activity during natural vision. However, natural movies contain strong correlations in time (trace under “natural” stimulus in Fig. 2.1B) and space (Fig. 2.1A, B). There are challenges with reliably estimating receptive fields from natural stimuli due to these strong correlations as well as the highly skewed natural intensity distribution (see [Methods](#)). We therefore also assessed the effect of spatio-temporal correlations in a more controlled stimulus with short-range exponential correlations in time and space and a binary intensity distribution (Fig. 2.1B, “spat-temp exponential”).

Additional stimuli allowed us to vary the spatial correlation over a broad range, without temporal structure, in order to test the hypothesis that surround strength adapts to remove correlations in nearby parts of an image. Thus, we examined spatial correlations of increasing extent: spatially exponential, a “multiscale” naturalistic stimulus featuring structure over many spatial scales, and full-field flicker (Fig. 2.1B, bottom row). The multiscale stimulus was designed to mimic the scale invariance of natural scenes in a controlled binary stimulus; featuring both small and large patches of correlated checks (such as the white area near the center). Its construction is detailed in [Methods](#). In one experiment, we also compared responses to low-contrast white and multiscale stimuli to their high-contrast versions. Finally, to control for

the effect of the skewed natural intensity distribution, we also conducted experiments presenting scrambled natural movies lacking spatial or temporal correlations but preserving the intensity distribution. The mean luminance and single-pixel standard deviation were matched across all stimuli other than low-contrast stimuli, which had a mean matched to the others with half the standard deviation, and natural intact and scrambled movies, which were matched to each other. Over 30 minutes of recording in each stimulus condition, the typical cell fired  $\sim 7000$  spikes. This was sufficient to assess spike train correlations and to measure receptive fields for the white and exponentially correlated stimuli.

For preliminary analyses, we measured the spike-triggered average (STA) from each ganglion cell’s response to white noise. The resulting receptive fields typically gave good coverage of the sampled visual field (Fig. 2.2A) and clustered into types on the basis of their response polarity and temporal properties (Fig. 2.2B; details in [Methods](#)). The four basic types that we consistently identified across experiments were fast-ON and fast-OFF, distinguished by the transient and biphasic nature of their temporal filter, and slow-ON and slow-OFF, which had longer integration times and often less prominent biphasic filter lobes. Separating cells by type did not qualitatively change many of the results reported below; in these cases, we combined all cells to improve statistical power.

To probe the effect of stimulus correlation on ganglion cell response properties in detail, we applied a standard functional model, the linear-nonlinear (LN) model. In this model, the visual stimulus is filtered with a linear kernel that represents the spatio-temporal receptive field (STRF) of the cell. The filter output is then passed through a nonlinear transfer function to generate a predicted firing rate. The nonlinearity encompasses thresholding and saturation, as well as any gain on the linear response. For white noise stimuli, the STA is a good estimator of the STRF

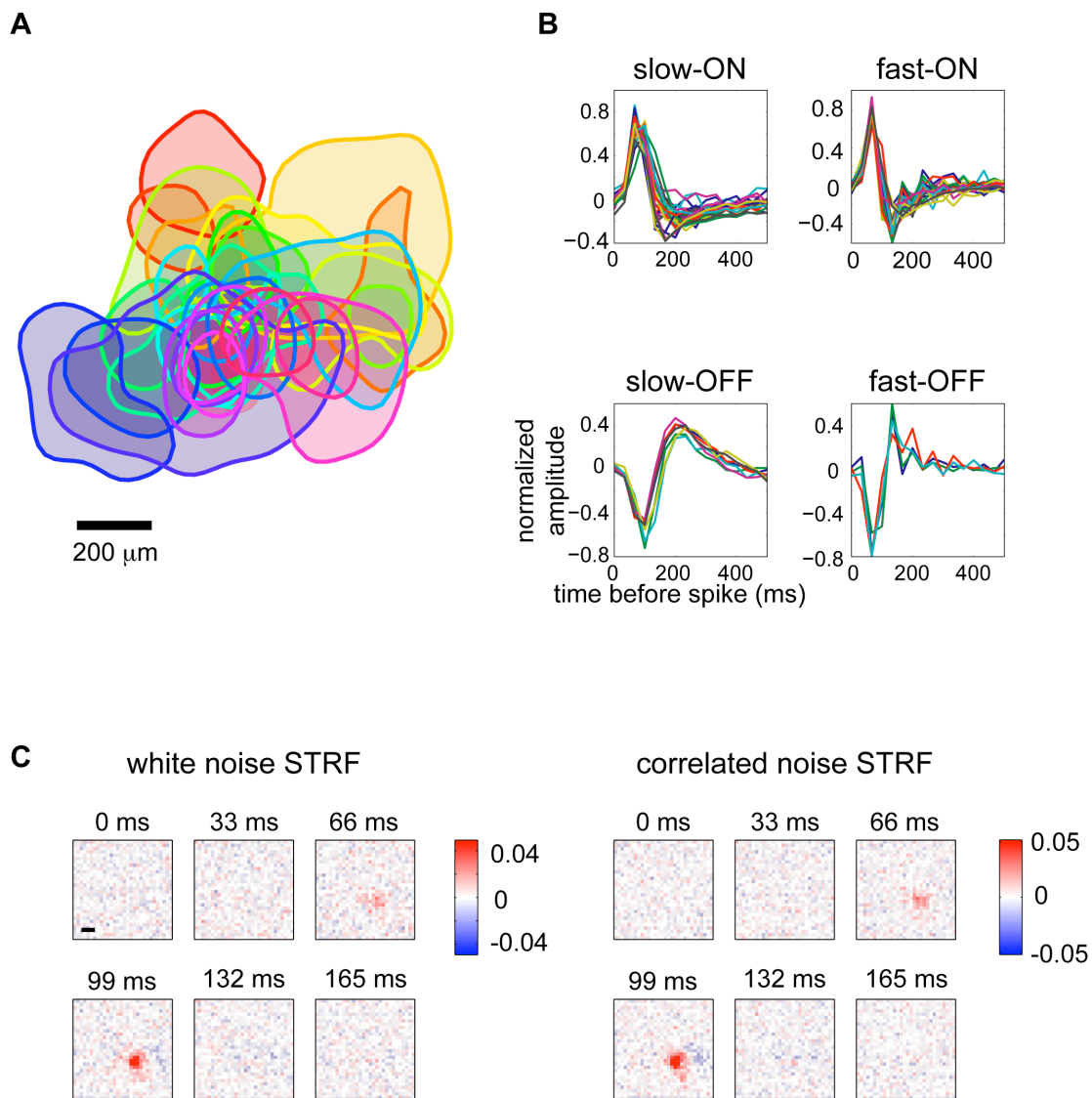


Figure 2.2



**Figure 2.2: Retinal ganglion cell receptive fields measured using a multi-electrode array.** (A) Receptive field locations of 31 cells recorded simultaneously from guinea pig retina. Each curve shows the 70% contour line of one receptive field. (B) Best-fitting temporal kernels for 75 cells, clustered into four types. Types were obtained by manually clustering temporal filters on the basis of the projection onto their first three principal components. (C) Maximum likelihood estimates of spatio-temporal receptive fields (STRFs) for an example cell. STRFs were computed separately using responses to white noise (*left*) or exponential spatio-temporally correlated stimuli (*right*). Scale bar is 200  $\mu\text{m}$ .

(Chichilnisky, 2001). However, this simple property does not hold for correlated stimuli, and so we fit the STRFs and other LN model parameters by maximum likelihood estimation (see [Methods](#)). For the weakly correlated spatio-temporal exponential stimulus, this technique reliably extracted receptive fields (Fig. 2.2C).

### 2.3.2 Output correlations in response to different stimuli

We computed the correlation coefficient between spike trains (binned at 33ms; see [Methods](#) and Fig. 2.7B) for all pairs of simultaneously recorded neurons. In response to natural movies, correlations between most pairs of cells increased in magnitude when compared with the correlations between the same pairs when viewing white noise (Fig. 2.3A). We quantified the size of this increase by finding the least-squares best fit line (Fig. 2.3B, gray lines) and defining the “excess correlation” of a population as the slope of this line minus one (see [Methods](#)). If all cell pairs had, on average, the same correlation in both stimulus conditions, the excess correlation would be zero. In the case of natural movies, the excess correlation was  $0.32 \pm 0.20$  (95% confidence interval computed using bootstrap resampling, as explained in [Methods](#)), modestly different from zero.

Since the retinal ganglion cell output is a highly transformed representation of its input, it is not trivial to formulate a naïve expectation for the magnitude of output

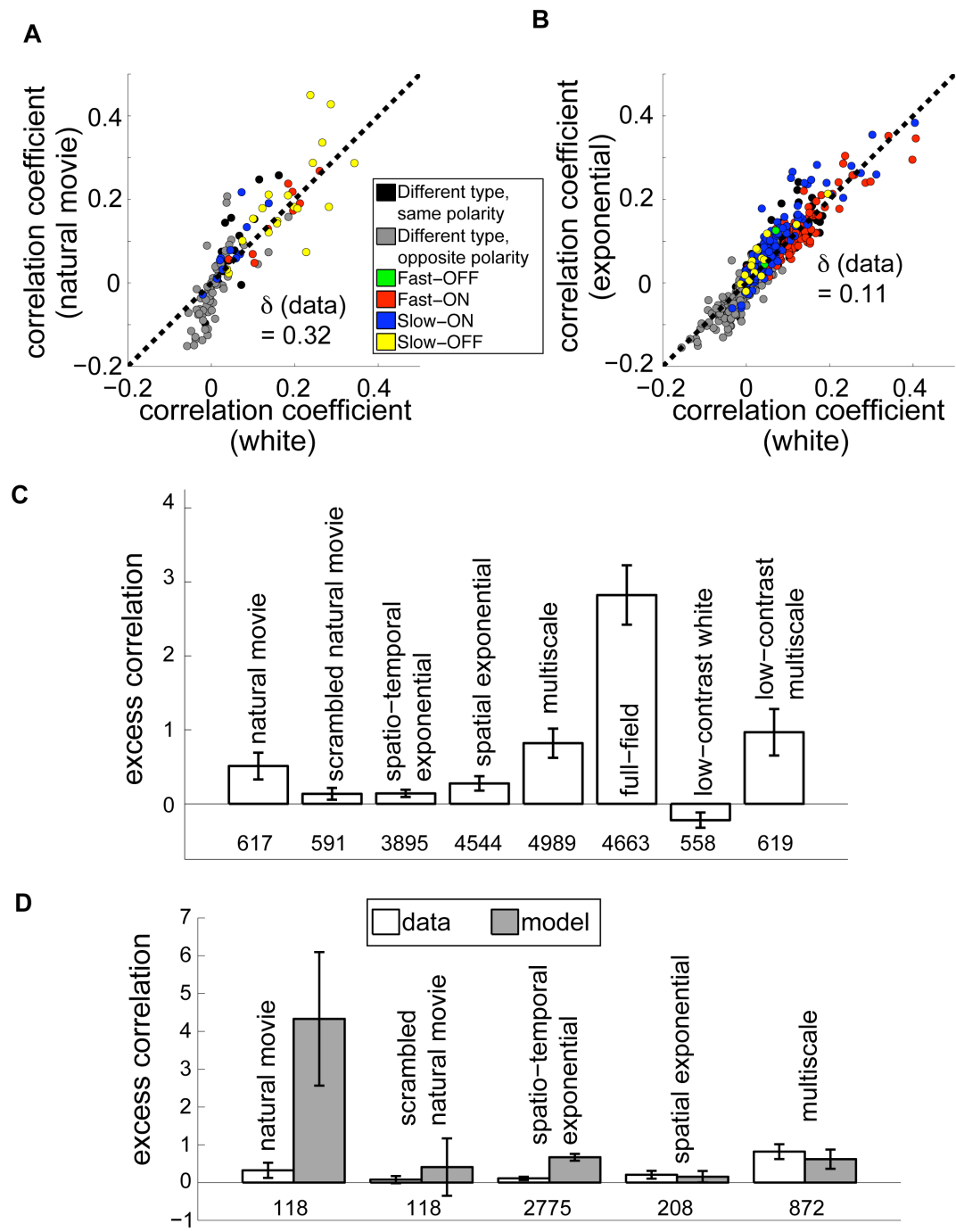


Figure 2.3

**Figure 2.3: Retinal output correlations are largely constant between stimulus conditions.** (A) Instantaneous spike train correlation coefficients between pairs of ganglion cells, comparing responses to natural movies and to white noise. *Dashed black line* is the diagonal. Cell pairs of the same type are indicated by colors in the legend. Different-type pairs are separated into ON-OFF (*gray*) and ON-ON or OFF-OFF pairs (*black*). The excess correlation,  $\delta$ , is the deviation of the slope of the best fit line (*gray*) from the diagonal. (B) Same as A but for spatio-temporal exponential stimulus. (C) Excess correlation measured from ganglion cells responding to the indicated stimulus, compared to white noise. Numbers below bars indicate the number of cell pairs in each condition; all recorded cells are included. Error bars are 95% bootstrap confidence intervals computed over 50,000 random samples with replacement from the set of cell pairs. (D) Comparison of measured excess correlation (*white*) to non-adapting model predictions (*gray*) for the indicated stimuli. Model values were derived from LN neurons with parameters fit to white noise data. Only cells whose receptive fields met a quality threshold are used here.

correlation given an input correlation. We therefore chose to quantify this expectation in a simple null model: the LN model fit to the white noise responses. This model captures correlation due to receptive field overlap and simple nonlinear processing while neglecting correlations due to shared circuitry and more complex nonlinear behavior, such as adaptation. For cells which had sufficiently well-estimated white noise LN model parameters (as described in [Methods](#)), we were able to compare the excess correlation predicted by the model to that observed in the data. We adjusted the threshold of each model neuron separately under each stimulus to match predicted average firing rates to their empirical values, which differed between stimuli. All other parameters, namely the spatio-temporal receptive field and the gain, were unchanged between stimuli. This “non-adapting” model predicted a significantly larger excess correlation in response to natural movies (gray bars in Fig. 2.3D and Fig. 2.4A), suggesting that the low observed excess correlation value under natural stimulation is a consequence of nontrivial processing in the retina.

In addition to strong correlations, however, natural stimuli are also characterized

by a skewed distribution with many dark pixels and a few extremely bright pixels, whereas our white noise stimulus included equal numbers of bright and dark pixels. To disentangle effects of correlations from effects due to intensity distribution, we presented the same retinae with a scrambled natural movie. In this stimulus, we started with natural movies and randomly shuffled the pixels in space and time to maintain the intensity distribution but remove correlations. The excess correlation in response to this stimulus was consistent with zero in both the measured and simulated responses (Fig. 2.4A, left bars). Moreover, comparing the natural movie and scrambled natural movie directly, we again found a small excess correlation consistent with that in the natural movie vs. white noise case. The non-adapting model again predicts that this low excess correlation is nontrivial (Fig. 2.4A, right bars). Thus, the retina greatly suppresses increases in correlation when natural visual stimuli are presented.

We found a similar set of results for the more weakly correlated spatio-temporal exponential stimulus (Fig. 2.3B). In particular, the excess correlation was low ( $0.12 \pm 0.05$ ) whereas simulated responses from the non-adapting model showed an increase (excess correlation of 0.67; Fig. 2.3C). We also examined the results of experiments in which we presented stimuli with varying degrees of spatial correlation (see Table 2.1). As shown in Figure 2.3C, many stimuli produced only a modest increase in output correlations. Some stimuli with strong spatial correlations, particularly the multiscale and full-field flicker stimuli, resulted in a clear increase in output correlations when compared to white noise. For stimuli where we varied the contrast (namely white and multiscale noise), output correlations decreased when the contrast was lowered while all other stimulus properties were kept fixed. Thus, the degree of correlation in the retinal output is not a reflection of stimulus correlations alone.

For further analysis, we focused on the subset of stimuli shown in Figure 2.3D, all

**Table 2.1: Number of cells recorded in each condition.**

stimulus	retinae	all cells	high-quality RFs <sup>b</sup>	cell types <sup>c</sup>			
				fast-ON	fast-OFF	slow-ON	slow-OFF
natural movie	3	84	34	12	0	9	13
scrambled natural movie	3	82	34	12	0	9	13
spatio-temporal exponential	5	212	75	29	4	31	8
spatial exponential <sup>a</sup>	17	510	46	-	-	-	-
multiscale <sup>a</sup>	16	513	62	-	-	-	-
full-field <sup>a</sup>	14	483	-	-	-	-	-
low-contrast white	1	49	-	-	-	-	-
low-contrast multiscale	1	49	-	-	-	-	-

<sup>a</sup> For our measurements of output correlation (Fig. 2.3C), we include additional data from experiments performed as part of other studies in which receptive field structure was not probed. For model correlations and other analyses, we only used the subset of retinae and cells for which we obtained robust receptive field estimates.

<sup>b</sup> We used a stringent requirement that receptive fields (RFs) be of high quality for any analyses in which we used receptive field estimates.

<sup>c</sup> Cells were only divided into subtypes if they had high-quality receptive fields and were recorded in response to stimuli chosen for detailed analysis.

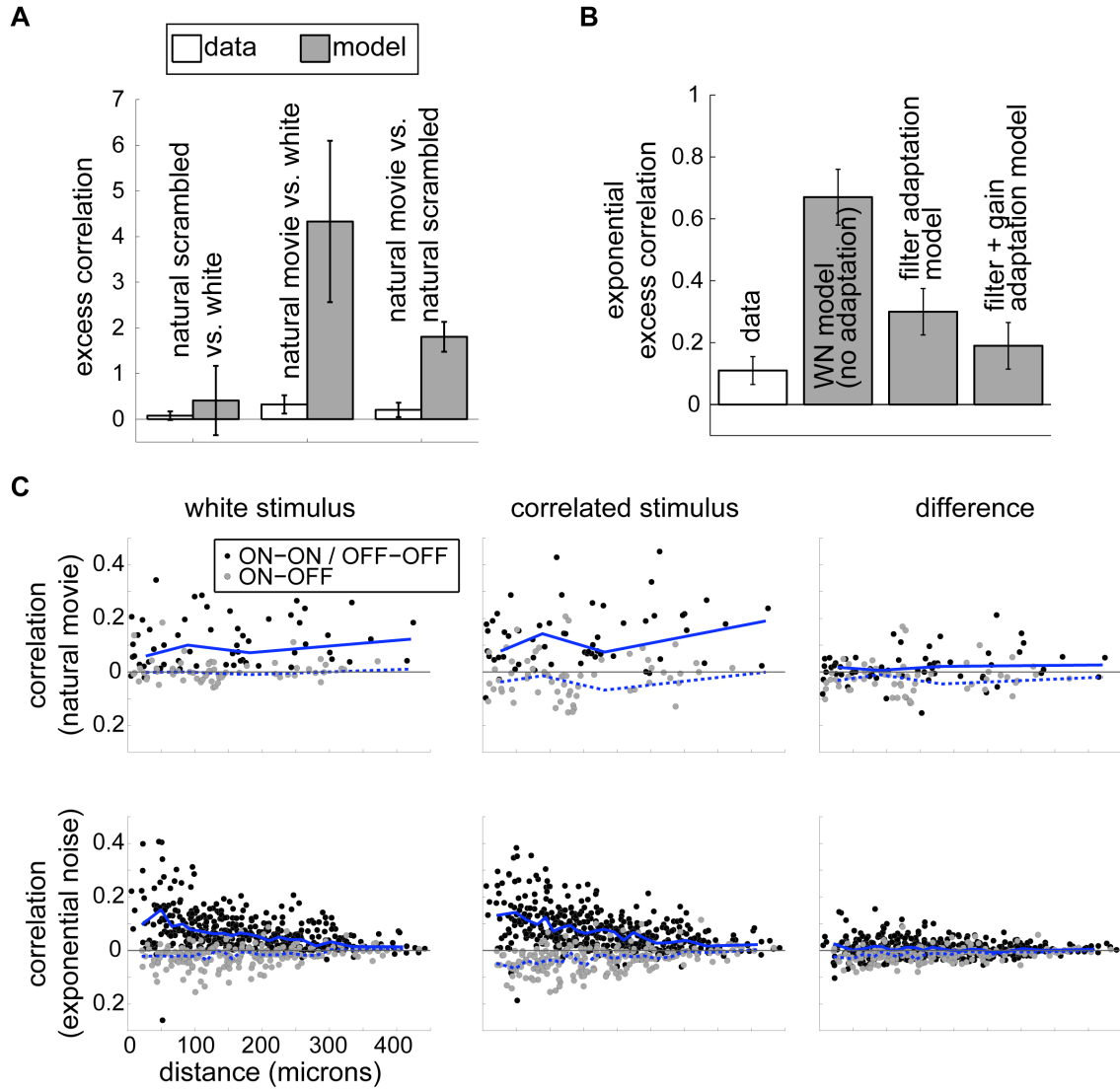


Figure 2.4

**Figure 2.4: Analysis of pairwise correlations.** (A) Excess correlations for natural stimuli. *Left* and *middle bars* show excess correlation when scrambled natural movies and intact natural movies, respectively, are compared to white noise in the data and in a population of non-adapting model neurons. *Right bars* show excess correlation when responses to natural movies are compared to scrambled natural movies directly. A non-adapting model predicts larger changes in output correlations than seen in the data in response to the correlated natural input. (B) Output correlations under the spatio-temporal exponential stimulus compared with white noise as predicted by LN models with parameters fit to the data. The two *leftmost bars* (“data” and “WN model (no adaptation)”) reproduce the spatio-temporal “data” and “model” bars in Figure 2.3D. (Note the difference in scale.) For the other bars, we simulated a population of neurons using linear filters measured from each stimulus but gains measured only from white noise (“filter adaptation model”) or using experimentally derived estimates of both linear filters and gains for each stimulus (“filter + gain adaptation model”). In the fully adapted model, excess correlations are consistent with the data. (C) Pairwise output correlation as a function of the distance between receptive field centers in the natural movie (*top*) and spatio-temporal exponential (*bottom*) datasets. Each dataset contained responses of the same cell population to white noise (*left*) and to a correlated stimulus (*middle*). The difference in output correlation in the correlated stimulus over the white noise stimulus is also shown for each cell pair (*right*). Each point corresponds to one simultaneously recorded cell pair; the *blue lines* are the median correlation within bins chosen to contain 30 cells each. *Solid lines* are median correlations for same-polarity cell pairs; *dashed lines* are for opposite-polarity pairs.

of which were presented in experiments where we also obtained robust estimates of white noise receptive fields. Here we again simulated responses of an LN model using fixed receptive fields measured under white noise. For all stimuli, the model neurons showed changes in correlation at least as large as those observed in recordings. However, unlike the spatio-temporally correlated exponential and natural stimuli discussed above, the stimuli which had correlations in space only (spatial exponential and multiscale) or no correlations (scrambled natural movie) produced similar excess correlation values in recorded cells and in our non-adapting model. This suggests that a fixed linear filter, as in the non-adapting model, is largely sufficient to account for

the effect of *spatial* stimulus correlations on output correlations, whereas higher-order processing is necessary to suppress the impact of *temporal* stimulus correlations.

As discussed above, we were able to identify the cell types for many of our recorded neurons. In response to spatio-temporal exponentially correlated noise and natural movies, cell type had a modest effect on output correlations (Fig. 2.3A and 2.3B). Cells with opposite ON- or OFF- polarities (gray points) tended to have negative correlations, whereas cells of the same polarity (black and colored points) generally had positive correlations. Moreover, pairs with opposite polarity showed a greater-than-average excess correlation, particularly in response to natural movies. Under natural movies, opposite-polarity pairs had an excess correlation of 1.5; under the spatio-temporal exponential stimulus their excess correlation was 0.38. Within same-type pairs, slow-ON and slow-OFF pairs (blue and yellow) tended to show a greater excess correlation than fast-ON and fast-OFF pairs (red and green). Pairs of slow cells had an excess correlation of 0.29 in the natural stimulus and 0.28 in the spatio-temporal exponential, while fast pairs were measured as 0.01 and -0.02 for the two stimuli, respectively. These type-dependent excess correlations are all small compared to the overall non-adapting model predictions (excess correlations of 4.33 and 0.67 for natural and spatio-temporal exponential stimuli). We also assessed the relationship between receptive field separation and output correlation (Fig. 2.4C). Pairwise correlations tended to decay with distance, but the change in output correlation between the correlated and white stimulus was small for all receptive field separations.

### 2.3.3 Adaptation of temporal filters

We next sought to determine whether receptive fields adapt to stimulus correlations and whether this adaptation can explain the observed pattern of output correlations.



As noted above, we were able to obtain STAs from responses to white noise. STAs computed in response to correlated stimuli, however, will be artificially blurred by the stimulus correlations. To obtain a better estimate of the spatio-temporal receptive field (STRF), we used maximum likelihood estimation to fit an LN model separately for the white and exponentially correlated stimuli (Theunissen et al., 2001). Examples of STRFs obtained in this way for one cell are shown in Figure 2.2C. The strongly correlated structure of the multiscale stimulus and the natural movies precluded robust, unbiased STRF estimation with limited data (see Methods). For this reason, we restricted any STRF computations to white noise and exponentially correlated noise. The latter stimulus is only weakly correlated and thus we would expect at most weak changes in the receptive fields between the conditions; indeed, receptive fields are hard to distinguish by eye for many cells. Measuring such weak changes requires high-quality receptive fields whose locations can be unambiguously determined, as was the case for 75 neurons ( $\sim 60\%$  of the neurons recorded under exponentially correlated conditions). Cells that did not meet this standard were likely to include types that do not respond as well to checkerboard stimuli, e.g., direction selective ganglion cells and uniformity detectors. We included such cells in the analysis of Figure 2.3C in order to maximize the generality of our results and to allow for the possibility that these neurons had qualitatively different output correlations. For the neurons that did pass the quality threshold, we found that the parameters of the LN model (for each neuron, a linear filter and a nonlinearity gain and threshold) changed with the stimulus.

Spike trains with sparse, transient firing events tend to be more decorrelated (Pitkow and Meister, 2012). Motivated by this finding, together with our observation that spatio-temporally correlated stimuli yielded high excess correlation in the non-adapting model, we analyzed adaptation in the temporal filtering properties of retinal

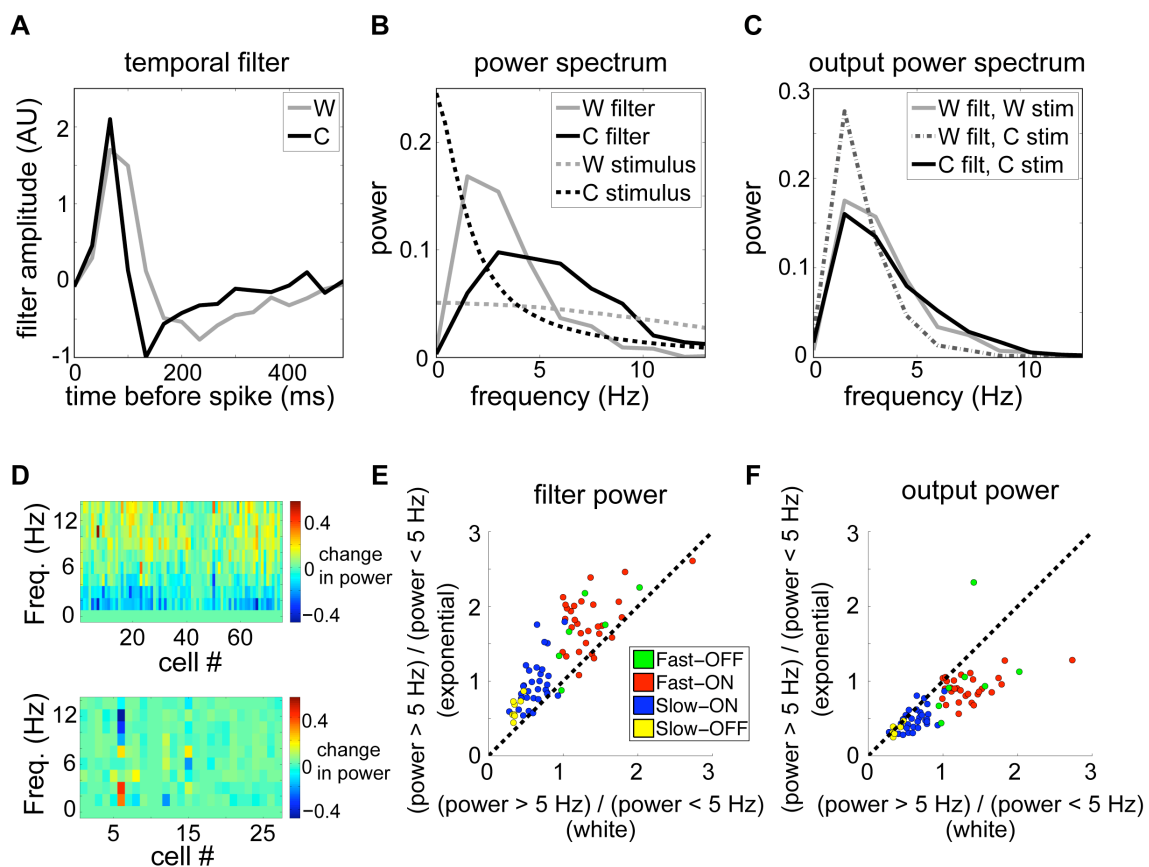


Figure 2.5

**Figure 2.5: Adaptation of the linear temporal filter.** **(A)** Temporal filters are faster under exponentially correlated noise (C) than white noise (W). **(B)** Power spectrum of correlated noise input (C, *black dashed line*) has more low frequency power than white noise (W, *gray dashed line*). The power spectrum of the temporal filter for correlated noise (C, *black solid line*) has more high frequency power than the filter for white noise (W, *gray solid line*). **(C)** Power spectra of filter outputs: White-noise filter acting on white stimulus (*solid gray*); White-noise filter acting on correlated stimulus (*dashed*); Adapting correlated-noise filter acting on correlated stimulus (*solid black*). In adapted cases, output power spectra are similar between stimuli — i.e., temporal kernels compensate to maintain invariant output autocorrelation. **(D)** The difference in normalized filter power spectra between the correlated and white stimuli, for spatio-temporal (*top*) and spatial (*bottom*) exponential experiments. The power spectra of all filters in each stimulus were normalized by removing the DC component and dividing by the sum of squared amplitudes. The population change in temporal filters shows a consistent increase in high-frequency power relative to low-frequency power for the spatio-temporal, but not the spatial, stimulus. **(E)** Total power above 5 Hz divided by total power below 5 Hz for filters computed in response to correlated vs. white noise stimuli shows a shift towards high-pass signaling across the population. **(F)** Same analysis as in **E** applied to the filter output as in **C**. Points near the diagonal indicate near-complete compensation for stimulus changes; points below the diagonal indicate incomplete compensation.

ganglion cells. To isolate changes in temporal processing, we examined each neuron’s STRFs (estimated separately under the white and exponentially correlated stimulus conditions) and extracted the temporal components (see [Methods](#)). These temporal profiles were faster for the correlated stimulus than for white noise (Fig. 2.5A). To quantify this difference, we computed the power spectrum of each neuron’s temporal filter under each stimulus (Fig. 2.5B and 2.5D, top) and found a systematic increase in high frequencies under the correlated stimulus, indicating a shift toward high-pass filtering (Fig. 2.5E). As the correlated stimulus had relatively more power at low frequencies compared to the white stimulus (Fig. 2.5B), this form of adaptation compensates for differences in the power spectrum and, hence, tends to equalize output auto-correlations. In contrast, a non-adapting model with a filter estimated

from white noise acting on the correlated stimulus predicts large changes in the output power spectrum (Fig. 2.5C). Indeed, this compensation was nearly exact for many cells (Fig. 2.5C), though generally incomplete over the full population (Fig. 2.5F).

Next, we found separate temporal profiles for the center and surround and computed the latency, measured as time to peak, of each. Surround latencies did not differ between white noise and spatio-temporally exponentially correlated noise (t-test,  $p = .7$ ,  $N = 75$ ). However, center latencies  $l$  were shorter for correlated noise. We quantified the shift in terms of an adaptation index  $(l_{corr} - l_{white}) / (l_{corr} + l_{white})$ . The histogram of the adaptation index (Fig. 2.6A; mean =  $-0.03$ , std =  $0.03$ ; t-test  $p < 10^{-12}$ ,  $N = 75$ ; Wilcoxon signed rank test  $p < 10^{-10}$ ) showed a robust tail toward shorter center latency for correlated stimuli (skewness =  $-0.53$ ). Moreover, almost every cell from which we obtained receptive fields had a longer latency for white noise than for correlated noise (Fig. 2.6B; mean change =  $6.1$  ms). This was true across cell types.

To determine whether these changes in temporal filtering were due to the presence of temporal correlations in this particular stimulus (unlike many of the other stimuli we examined), we also measured receptive fields from a separate population of ganglion cells responding to white noise and to a stimulus that was exponentially correlated in space but not in time. In this case, filters did not show a systematic change in power spectra (Fig. 2.5D, bottom), but the center latencies were still shorter for the correlated stimulus (Fig. 2.6B; mean change =  $7.2$  ms). Again, computing adaptation indices indicated that this effect was significant (mean =  $-0.04$ , std =  $0.03$ ; t-test  $p < 10^{-10}$ ,  $N = 37$ ; Wilcoxon signed rank test  $p < 10^{-7}$ ). Thus, spatial correlations in the stimulus affect the temporal filtering properties of neurons, albeit to a lesser degree than spatio-temporal correlations. These results, combined with those of Pitkow and Meister (2012), may indicate that the observed consistency of

correlations is produced in part by an increase in response transience when stimulus correlations increase.

### 2.3.4 Adaptation of spatial receptive fields

The conventional view of retinal circuitry suggests that adaptive decorrelation arises from stronger or wider surround inhibition induced by viewing correlated stimuli. We thus computed the amplitudes of the surround and center components of each neuron’s STRFs in both white noise and spatio-temporal exponentially correlated noise. Defining the relative surround strength,  $k$ , as the ratio of surround amplitude to center amplitude (details in [Methods](#)), we computed an adaptation index for each cell as  $(k_{corr} - k_{white}) / (k_{corr} + k_{white})$ . This adaptation index has a modestly positive mean (Fig. 2.6D; mean = 0.075, std = 0.24; two-tailed t-test,  $p = 0.008$ ,  $N = 75$ ; Wilcoxon signed rank test  $p = .008$ ), as do the changes in surround strength themselves (Fig. 2.6E). In addition, the cells with the greatest degree of surround adaptation had a robust tendency to increase in surround strength (skewness = 0.15). There was no discernible dependence on cell type. Interestingly, the surround strength showed only a marginally significant change when spatial correlations (but not temporal correlations) were added to white noise (Fig. 2.6F; mean adaptation index =  $-0.087$ , std = 0.26; two-tailed t-test,  $p = 0.05$ ,  $N = 37$ ; Wilcoxon signed rank test  $p = .02$ ). Thus, while we do find some evidence for an increase in surround strength with stimulus correlation, the effect is subtle. This outcome is surprising given the prevailing view since the work of Barlow ([Barlow, 1961](#); [Srinivasan et al., 1982](#)) that surround inhibition is primarily responsible for decorrelation of visual stimuli. However, it is possible that the exponential stimulus that permitted us to estimate receptive fields is too weakly correlated to evoke strong lateral inhibition.

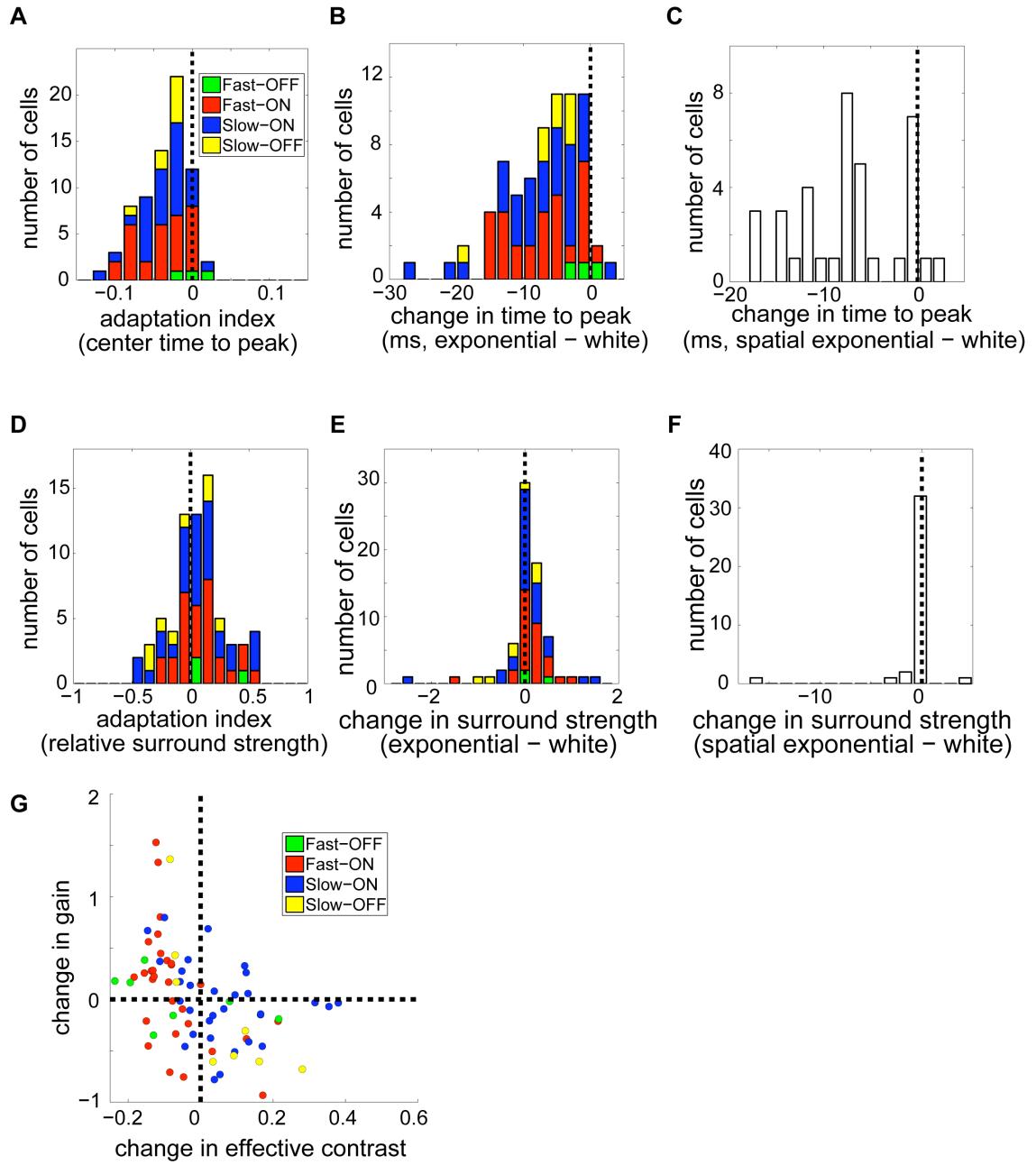


Figure 2.6

**Figure 2.6: Adaptation of the spatio-temporal receptive field and gain.**

(A) Center latency  $l$  (time to peak of the temporal kernel) is shorter for spatio-temporal exponentially correlated noise. Histogram shows adaptation indices  $(l_{corr} - l_{white}) / (l_{corr} + l_{white})$  for center latency. (B, C) Changes in center latency  $l_{corr} - l_{white}$  for spatio-temporally correlated (B) and temporally correlated (C) stimuli, in milliseconds. Almost all cells have a decreased time to peak when responding to a correlated stimulus. (D) Adaptation indices  $(k_{corr} - k_{white}) / (k_{corr} + k_{white})$  for relative surround strength  $k$  (surround/center ratio) show a slight skew toward a stronger surround for spatio-temporally correlated noise. (E, F) Difference in surround strength  $k_{corr} - k_{white}$  for the spatio-temporal (E) and spatial (F) exponential stimuli. (G) Gain adaptation. Gains were defined as the slope of the LN model nonlinearity and were obtained separately for the response to white noise and to the spatio-temporally correlated exponential stimulus. Effective contrast  $\sigma$ , the standard deviation of the linear filter output, was similarly measured in both stimuli. The difference in gain  $g_{corr} - g_{white}$  is plotted against the difference in effective contrast  $\sigma_{corr} - \sigma_{white}$ . Increases in effective contrast tend to invoke compensating decreases in gain.  $N = 75$  cells.

### 2.3.5 Adaptation of nonlinearity gain

We also examined the gain  $g$  of each neuron, defined as the maximum slope of the logistic nonlinearity fit to each neuron’s response (see [Methods](#)). Since the gain enters the nonlinearity after the stimulus passes through the linear filter, we normalized the filter to unit euclidean norm in order to obtain an unambiguous definition of  $g$ . We found that the gains of individual neurons changed when the stimulus was more correlated, but there was not a systematic change between stimuli. Recall that the gain of many sensory neurons, including retinal ganglion cells, is known to change with the contrast of the stimulus ([Smirnakis et al., 1997](#); [Baccus and Meister, 2002](#)). To test for a possibly related mechanism at work in our data, we first defined “effective contrast,”  $\sigma_{white}$  and  $\sigma_{corr}$ , as the standard deviation of the normalized linear filter output in each stimulus, respectively. This notion of effective contrast roughly captures the variability of the ganglion cells’ input, taking presynaptic processing into account. Any nonlinear gain control mechanism in the ganglion cell layer should

therefore be sensitive to this quantity. For some cells  $\sigma_{corr}$  exceeded  $\sigma_{white}$ , while for others the reverse was true. Measuring the gains in both stimulus conditions ( $g_{white}$  and  $g_{corr}$ ), however, we found systematic adaptation opposing the change in effective contrast: gain tended to increase when effective contrast decreased and vice-versa. Specifically, the quantities  $\Delta g = g_{corr} - g_{white}$  and  $\Delta \sigma = \sigma_{corr} - \sigma_{white}$  were significantly anticorrelated (Fig. 2.6G; Spearman’s  $\rho = -0.54$ ,  $p = 10^{-6}$ ,  $N = 75$ ).

### 2.3.6 Output correlations in an adapting model

Finally, we assessed whether the receptive field changes reported above could account for the observed similarity of output correlations between white noise and the spatio-temporal exponential stimulus. For experiments using these two stimuli, we measured the adaptation in LN model parameters fit to each stimulus as discussed above. We then separately examined the effect of adaptation in different parameters on the excess correlations predicted by the LN models. Including adaptation of the linear filters, but not the gain, produced a significantly improved match between the model and the data (Fig. 2.4B, “filter adaptation model”). Additionally allowing the gain to adapt produced output correlations consistent with the data (Fig. 2.4B, “filter + gain adaptation model”). The contribution of gain adaptation to decorrelation is interesting in light of our observation that output correlations are lower for stimuli with lower contrast (Fig. 2.3C). Low contrast stimuli generally evoke lower firing rates, which could result in decreased pairwise correlations absent any change in linear filtering properties. (See [Methods](#) for a derivation of this result.) At the same time, changes in contrast lead to gain control, wherein gain is higher for lower stimulus contrast. This gain adaptation could also affect output correlations, as in Figure 2.4B. It would be interesting to know how gain control interacts with changes



in other properties, such as the nonlinearity threshold and the shape of the linear filter, to set the correlations in the retinal response.

Note that the LN model is fit to each neuron independently, without taking correlations between neurons into account. Its successful prediction of the change in pairwise correlations, without explicit introduction of inter-neural interactions, is therefore noteworthy. We conclude that observed adaptation in receptive fields and gains is adequate to explain the output correlations in responses to a spatio-temporally correlated stimulus.

## 2.4 Discussion

Our principal finding is that the retina maintains a moderate, and relatively constant, level of output correlation across a wide range of spatio-temporally correlated stimuli ranging from white noise to natural movies. Our data also suggest a differential effect of spatial vs. temporal correlations on the functional properties of the retinal output. We focused here on spatial variations in our control stimuli, but it would be interesting to design future studies to explore the space-time differences more systematically. In response to exponentially correlated noise, where the receptive fields could be estimated, we showed that the relative invariance of output correlations is largely accounted for by the observed changes in the linear receptive field (faster temporal kernels and slightly stronger surround inhibition for more correlated stimuli) and by changes in the nonlinear gain (anti-correlated to changes in effective contrast). While the latter findings give an interpretation of the results in terms of a conventional functional model (here a linear-nonlinear cascade), the measurement of output correlations is model-independent.

Classifying cells into types revealed a slight dependence of excess correlation on

cell type: most robustly, opposite polarity ON-OFF pairs showed the greatest increase in correlation magnitude when stimulus correlation increased. Indeed, if the retinal output is split across parallel functional channels, redundancy is likely to be highest within a channel due to shared circuit inputs. It may thus be advantageous, from an information encoding perspective, for decorrelation to act *within* a channel, with residual correlations across types signaling relevant relationships between the information in different channels for use by downstream areas.

Pitkow and Meister (2012) showed that the retina partly decorrelates naturalistic inputs but that the response to white noise is more correlated than the input, in part due to receptive field overlap between ganglion cells. Consistent with their results, we found that changes in output correlations were often smaller than changes in input correlations. We also extended their findings by showing that this partial decorrelation occurs in individual pairs of neurons. Pitkow and Meister (2012) suggested that the linear receptive field measured from white noise was insufficient to explain the amount of decorrelation seen for naturalistic stimuli and the bulk of the decorrelation was attributed to changes in the threshold of a functional model of ganglion cells. However, the authors did not directly measure the (possibly different) receptive fields of ganglion cells responding to correlated stimuli, nor did they follow particular cell pairs across different stimuli. Our measurements suggest that the nonlinear processing proposed in Pitkow and Meister (2012) can be described in terms of adaptation of the linear receptive field and nonlinear gain with the net effect that output correlations are roughly constant for each cell pair across a range of correlated stimuli, as was observed in visual cortex by Fiser et al. (2004). Our results also recall those of Lesica et al. (2007), David et al. (2004), Sharpee et al. (2006), and Touryan et al. (2005), who showed that receptive fields in LGN and primary visual cortex differ in structure when probed with natural movies versus random stimuli.

We also found that the gain of retinal ganglion cells responding to correlated stimuli changes with “effective contrast”  $\sigma_{white}$  and  $\sigma_{corr}$ , i.e. with the standard deviation of the input to the nonlinearity in a linear-nonlinear model of ganglion cells. In classical contrast gain control, firing rates and response kinetics adapt to temporal contrast and to the spatial scale of stimuli (Smirnakis et al., 1997; Baccus and Meister, 2002). As increased stimulus correlation may produce a qualitatively similar input to the inner plexiform layer as increased contrast, some of the cellular mechanisms underlying contrast adaptation might also contribute to the phenomena we have uncovered. This provides an avenue for future study of the functional mechanisms underlying adaptation to correlations.

We have focused in the present work on the failure of a non-adapting linear-nonlinear model to capture the small scale of observed excess correlations and have seen that adaptation in the linear filter might remedy this discrepancy. Alternatively, shared circuitry in the population of neurons may be engaged by correlated inputs and require explicit inclusion in any functional model of retinal responses to different classes of correlated stimuli (Ganmor et al., 2011; Pillow et al., 2008). Such shared circuitry leads to noise in one neuron being passed to multiple nearby neurons and is thus measured by “noise correlations.” While addition of fixed, stimulus-independent noise correlation would not greatly change our results, a *change* in noise correlation with stimulus would provide a different candidate mechanism to account for our data (Granot-Atedgi et al., 2013). This is another avenue for future work.

We have focused here on the effects of spatial correlations in an experimental design where we could compare receptive fields computed from responses to two different stimuli. Thus, we used relatively weak exponential correlations to ensure that we were not measuring artifacts of the stimulus correlations themselves. Recovering receptive fields from strongly correlated stimuli can require long recording times; be-

cause our experimental design involved comparisons between several different stimuli, we were only able to recover receptive fields for moderately correlated stimuli. Future work could simply present each stimulus for a longer duration to assess receptive field changes at a population level rather than analyzing multiple stimuli in one experiment. Further work could also include parallel studies with stimuli including temporal correlations only to complement our findings on responses to spatial correlations.

Finally, it would be interesting to determine the timecourse of the adaptations observed here. Knowing whether a change in stimulus correlations induces changes in receptive fields and output correlations within seconds, tens of seconds, or longer would help to clarify the relationship between processing of correlations and adaptation to other stimulus features such as contrast. Again, the design of our experiments precluded making these measurements – we focused on long segments to measure steady-state processing of correlations, whereas assessing the timecourse of changes requires finer and more systematic sampling of transitions between stimuli.

Why would the retina need to adapt, in the behaving animal, to variations in spatial correlations? While natural scenes are scale-invariant on average, the specific correlations do vary depending on the scene and the viewing distance (see Fig. 2.1A). Barlow originally suggested that sensory systems should decorrelate their inputs to make efficient use of limited neural bandwidth (Barlow, 1961). Consistent with this idea, we found that retina removes redundancies, in spatio-temporally correlated stimuli, but also that the retinal output is not completely decorrelated. Rather, the output correlations are reduced to an approximately fixed level, roughly matching correlations in responses to white noise checkerboards. What drives this tradeoff? Recall that redundancy can be useful to protect against noise, to facilitate downstream computations, or to enable separate modulation of information being routed to distinct cortical targets. Thus, it may be that a certain degree of output correlation

between retinal ganglion cells represents a good balance of the benefits of decorrelation with the benefits of redundancy (Tkačik et al., 2010). Sensory outputs with varying amounts of correlation may also be decoded differently by cortex (Estebanez et al., 2012), in which case maintaining a fixed visual code might require that retinal output correlations are within the range expected by downstream visual areas. In these interpretations, it makes sense that the retina adapts to maintain correlation within a narrow range across stimulus conditions, as we have found.

## 2.5 Methods

**Ethics statement.** All procedures were in keeping with the guidelines of the University of Pennsylvania, the NIH, and the AVMA.

**Neural recording.** We recorded retinal ganglion cells from Hartley guinea pig using a 30-electrode array (30  $\mu\text{m}$  spacing, Multi Channel Systems MCS GmbH, Reutlingen, Germany). After anesthesia with ketamine/xylazine (100/20 mg/kg) and pentobarbital (100 mg/kg), the eye was enucleated and the animal was euthanized by pentobarbital overdose. The eye was hemisected and dark adapted. The retina was separated from the pigment epithelium, mounted on filter paper, and placed onto the electrode array, ganglion cells closest to the electrodes. Extracellular signals were recorded at 10 kHz. Spike times were extracted with the spike-sorting algorithm described in Appendix A; briefly, a subset of data was manually clustered to generate spike templates that were then fit to the remaining data using a Bayesian goodness-of-fit criterion.

**Stimulus generation.** We displayed checkerboard stimuli (see Fig. 2.1B) at 30 Hz on a Lucivid monitor (MicroBrightField, Inc., Colchester, VT) and projected the im-

age onto the retina. The mean luminance on the retina was  $9000 \text{ photons/s} \cdot \mu\text{m}^2$  (low photopic) for most stimuli; each check occupied an area between  $50 \mu\text{m} \times 67 \mu\text{m}$  and  $100 \mu\text{m} \times 133 \mu\text{m}$ . To make white noise and exponentially correlated stimuli, we first produced random checkerboards with intensities drawn from a Gaussian distribution. Spatio-temporally correlated stimuli were produced by filtering the Gaussian stimulus with an exponential filter with a time constant of three stimulus frames (99 ms) and a space constant of one check to match the scale of typical receptive fields. Stimuli with only spatial exponential correlations were constructed similarly, but with a time constant set to zero. To create the multiscale stimulus, we first generated gaussian white noise checkerboards at several power-of-two scales. The largest scale consisted of a single check filling the entire stimulus field, the next largest was a  $2 \times 2$  checkerboard (with check size equal to half the stimulus field), the third largest was a  $4 \times 4$  checkerboard (check size one quarter of the stimulus field), and so on. The checkerboards at all scales were then summed and thresholded to obtain a binary stimulus qualitatively mimicking the scale-invariant structure of spatial correlations in natural scenes (Fig. 2.1B). This stimulus did not contain temporal correlations.

Natural movies of leaves and grasses blowing in the wind were collected with a Prosilica GE 1050 high-speed digital camera with a  $1/2''$  sensor (Allied Vision Technologies GmbH, Stadtroda, Germany) connected to a laptop running StreamPix software (NorPix Inc., Montreal, Canada) to grab frames at 60 fps. The camera resolution was set to  $512 \times 512$  pixels, and movies were filmed from a fixed tripod about 5 feet from the trees and grass. Natural light was used to illuminate our outdoor scenes, and exposure time was set ( $300 - 3000 \mu\text{s}$ ) to capture variation in shadows and avoid saturation of our 8-bit luminance depth. Videos were collected for up to 30 minutes; 10 second to 5 minute segments with continuous motion were selected. Videos were downsampled to match the resolution and frame rate of our stimulus

monitor and normalized to have a mean luminance of  $2800 \text{ photons/s} \cdot \mu\text{m}^2$  with a standard deviation of  $1700 \text{ photons/s} \cdot \mu\text{m}^2$  (overall contrast of 48%). To produce a scrambled control for natural movies, pixels were randomly shuffled in space and time to remove all correlations. All stimuli other than natural movies (intact and scrambled) were thresholded at the median to fix the mean luminance and single-pixel variance and to maximize contrast. This binarization did not affect the power spectra significantly. For low-contrast stimuli, all deviations from the mean luminance were halved to give an overall contrast of 50%. Typically, we alternated 10-minute blocks of white noise with 10-minute blocks of a correlated stimulus.

**Cell typing.** We used reverse correlation to compute the spike-triggered average (STA) for each cell responding to white noise. We performed principal component analysis (PCA) on the best-fitting temporal kernels and split cells into two clusters based on the sign of the first component; the clusters were identified as ON and OFF types based on the sign of their temporal kernels. (Our under-sampling of OFF cells, relative to the findings of [Borghuis et al., 2008](#), and [Ratliff et al., 2010](#), may be due to electrode bias, as individual OFF cells are smaller and therefore less likely to be detected by our electrode array.) PCA was repeated for the ON and OFF groups separately. We manually identified clusters based on the projections onto the first three principal components; in this way we identified four functional types, including slow-OFF, fast-OFF, fast-ON, and slow-ON (see Fig. [2.2B](#)).

**Maximum likelihood estimation of linear-nonlinear models.** In order to obtain spatio-temporal receptive fields (STRFs) for both white and exponentially correlated stimuli, we used publicly available code ([strflab.berkeley.edu](http://strflab.berkeley.edu); [Theunissen et al., 2001](#)) to carry out maximum likelihood estimation. We parameterized the model

by a linear filter acting on the stimulus and a logistic nonlinearity, so that firing probability is  $p(s) = 1/(1 + \exp(-g * (s - b)))$ , where  $s$  represents the linear filter output, and  $g$  and  $b$  are gain and offset parameters. Gradient ascent with early stopping was used to compute a maximum likelihood estimate of the linear filter that best fit the data. We initialized the algorithm for each neuron using the spike-triggered average recorded in response to white noise so that any incomplete fitting would bias our results towards a null effect. Many cells do not yield clear receptive fields when probed with white noise, either because this stimulus does not evoke a sufficiently strong response or because the response is not well modeled as a single linear filter. To avoid potential artifacts that could arise from including such cells in our receptive field and model analyses, we selected cells whose receptive fields had clearly visible centers. This classification of receptive fields as high-quality was done before any other data analysis in order to avoid biasing the selection. In datasets where we obtained receptive fields for both white noise and a correlated stimulus the designations of high-quality agreed between the two stimuli for 98% of cells. The subset of cells identified in this way also had center locations that were clearly delineated by our automated receptive field analysis, giving confirmation of our visual threshold.

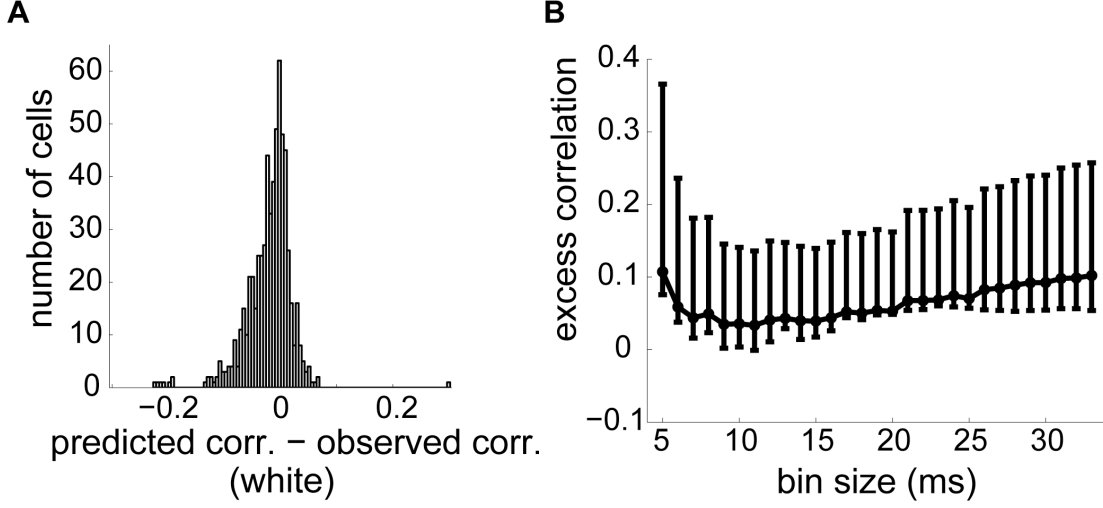
The STRF baseline was poorly constrained by the maximum likelihood procedure, since an additive change in the STRF has a similar effect to a proportional shift in the offset parameter of the nonlinearity. We therefore normalized the STRFs by subtracting an estimate of the baseline: we allowed the fit to include components extending 100 ms after the spike — where the true filter must be zero by causality — and subtracted the mean of these frames. Inclusion of these post-spike frames also allowed us to verify that the temporal autocorrelations in our stimuli did not produce any acausal artifacts in the recovered STRFs. We normalized the estimated linear filters to have unit Euclidean norm (square root of the sum of squares of



filter values) and then used gradient ascent to separately fit the gain and offset of a logistic nonlinearity. Since the likelihood function in this case is convex, there was no possibility of local maxima. While we were able to compute unbiased estimates of STRFs from responses to stimuli with exponential correlations, our multiscale and natural movie stimuli were too correlated to estimate unbiased receptive fields with the number of spikes we were able to obtain in a single recording.

**Correlation analysis.** Correlations were measured as the correlation coefficient between pairs of simultaneously recorded neurons. Spike trains were divided into 33 ms bins; we assigned a bin a 1 if it had one or more spikes and a zero otherwise. The results reported above did not change if we used spike counts in each bin rather than binarizing. Indeed, 98% of timebins had one or fewer spikes and less than 0.05% had more than three spikes. In addition, using more finely grained binning did not change our results (Fig. 2.7B). We summarized the results by finding the best fit line of the form  $\rho_{corr} = (1 + \delta)\rho_{white}$ , where  $\rho_{white}$  and  $\rho_{corr}$  are the pairwise correlations under the white and correlated stimuli, respectively. We refer to  $\delta$  as the excess correlation; we estimated  $\delta$  by the total least squares regression method and computed 95% bootstrap confidence intervals from 50,000 bootstrap resamples of the set of ganglion cell pairs.

Such instantaneous correlations are thought to combine slow stimulus-driven effects with fast intrinsic effects due to shared noise (Greschner et al., 2011). To verify that this did not affect our results, we isolated the stimulus-driven component, by noting that our cross-correlation functions can feature a short-timescale peak riding on a slow component and extracting the latter. Specifically, we binned the spike trains into 1 ms bins and computed cross-covariance functions between pairs. To isolate the stimulus-induced component, we smoothed the cross-covariance functions by fitting a cubic B-spline curve with knots spaced at 20 ms to suppress the fast noise



**Figure 2.7: Correlation measurement controls.** (A) As a control on the quality of the non-adapting LN model, we examined the difference between its predicted pairwise correlations under the white noise stimulus (the stimulus to which the model was fit) and the observed correlations. Since the model is a single-neuron model that does not attempt to capture pairwise correlations, we do not expect it to reproduce these correlations perfectly. Nevertheless, many cell pairs are well-predicted, indicating that their correlation is largely due to receptive field overlap. There is, however, a slight tendency for the model to underestimate correlations: this is likely due to its neglect of noise correlations. We note that such a bias will not effect the model’s predicted *excess correlation*, unless it is very different in the two stimulus conditions. But such an effect would represent a form of non-trivial processing in its own right. (B) Our correlation measurements were based on binned spike trains. We measured excess correlation, in the spatio-temporal exponential dataset, for a variety of bin sizes. Its value is largely independent of bin size. Error bars represent 95% bootstrap confidence intervals. All correlations reported in the main text were estimated from spike trains binned at 33 ms.

component. We then found the shift, within a 200 ms window, which maximized the absolute value of the smoothed cross-covariance and estimated the correlation coefficient as the cross-covariance at this shift normalized by the product of the standard deviations. This gave excess correlation values consistent with those reported above (not shown).

To put the observed correlations in context, we also measured the correlations in the output of model neurons. For each each recorded neuron form which we obtained a robust STRF estimate, we produced unadapted model responses by passing each stimulus through the STRF derived from white noise and a nonlinearity function where the gain was estimated from white noise and the offset from the corresponding stimulus. That is, for a correlated stimulus, we chose the offset to match each cell’s firing rate in response to the correlated stimulus. We did this so that our simulations would represent the correlation we would expect given overall stimulus-dependent activity levels. For some stimuli, we also produced adapted model neurons in which the STRF and/or gain used to simulate correlated stimulus responses were estimated from recorded responses to the correlated stimulus. As shown in Figure 2.7A, the correlations in our model population were within the physiological range of true retinal output.

We also computed the power spectra of the stimuli, the best-fitting temporal kernels, and the filter outputs (i.e. stimulus power spectra multiplied by filter power spectra). We summarized each power spectrum by computing the total power above 5 Hz divided by the total power below 5 Hz.

**Measures of receptive field characteristics.** Given a STRF estimated for one cell under one of the stimulus conditions, we first performed principal component analysis on the timecourses of the individual pixels. From the resulting set of “princi-

pal timecourses” we selected the one most similar to the timecourse of the pixel that achieves the peak value in the full STRF. The complete linear filter was collapsed into a single frame by finding the projection of each pixel onto this principal timecourse. This procedure is equivalent to finding the best (least squares) spatio-temporally separable approximation to the STRF:  $K(x, t) = k(x)w(t)$ , where  $k(x)$  and  $w(t)$  are the spatial and temporal components of the approximation. From the single frame  $k(x)$ , we extracted the center and surround regions. To find the center, we began with the peak pixel and then recursively expanded the region in a contiguous patch to include any pixels whose timecourses had at least a 50% correlation with already included pixels. We ended the recursive process after the first pass in which no pixels were added to the center. At this point, all pixels not included in the center were considered part of the surround for the purpose of assessing the surround strength.

Taking the center defined in this way as a mask for the full STRFs, we summed all pixel values within the center at each time point to generate a temporal profile for the central receptive field. To obtain temporal kernels with greater precision than the 30 Hz time scale of our STRFs, we used cubic spline interpolation with knots spaced every 33 ms. From the interpolated timecourses, we measured the time to peak under each stimulus for the center. The peak value of the temporal profile was taken to be the center weight of the receptive field. Similar computations yielded the surround time to peak and surround weight. We then quantified the relative surround strength as the ratio of surround weight to center weight.

In addition, we measured the gain  $g$  of each neuron under each stimulus condition. To test for contrast gain control, we defined “effective contrast,”  $\sigma$ , as the standard deviation of the linear filter output. To avoid ambiguity between linear filter amplitude and gain, we normalized each STRF to have unit Euclidean norm before computing the gain and the effective contrast.

Table 2.2: Model validation of receptive field analysis.

	AI (relative surround strength)				AI (center time to peak)			
	mean	std	$p$	skew	mean	std	$p$	skew
Standard model	-0.02	0.06	.004	-1.10	-0.0003	0.002	.04	-4.70
Small surround weight	-0.06	0.13	< .0001	-0.12	-0.0003	0.002	.04	-4.69
Large surround weight	-0.05	0.26	.08	-2.44	-0.0014	0.004	.0003	-2.66
Small surround radius	-0.01	0.08	.19	-4.56	-0.0009	0.003	.0007	-2.49
Large surround radius	-0.02	0.05	< .0001	-0.24	-0.0005	0.002	.01	-3.71

Adaptation index (AI) in surround strength and center latency for different non-adapting control models. Columns labeled “mean,” “std,” and “skew” show the mean, standard deviation, and skewness of the adaptation indices for the indicated analysis; columns labeled “ $p$ ” show the  $p$ -values from  $t$ -tests of each distribution against the null hypothesis of zero mean. **Standard model:** Surround radii (relative to center radii) had mean 2 and standard deviation 0.3; surround weights (relative to center weights) had mean 1 and standard deviation 0.1. **Small surround weight:** Surround weights had mean 0.5; all other parameters were the same as in the standard model. **Large surround weight:** Surround weights had mean 2. **Small surround radius:** Surround radii had mean 1. **Large surround radius:** Surround radii had mean 3.

**Model validation of receptive field analysis.** To validate our STRF analysis methods, we generated synthetic data using a linear-nonlinear (LN) model. We then applied STRF extraction and analysis methods identical to those applied to real data to check that the known LN parameters were extracted in an unbiased fashion. The linear filter was chosen to be spatio-temporally separable, with the temporal component taken from measured ganglion cell responses. The spatial filter was modeled as a difference-of-Gaussians, where the size and strength of the surround Gaussian relative to the center Gaussian were chosen to mimic receptive fields of real neurons. We ran five separate simulations, each of which included 100 model neurons whose parameters were chosen independently. The results are summarized in Table 2.2.

In our first simulation, the surround radius (relative to center radius) was chosen from a Gaussian distribution with mean 2 and standard deviation 0.3, the relative surround strength from a Gaussian distribution with mean 1 and standard deviation 0.1, and the offset coordinates from Gaussian distributions with mean 0 and standard deviation 2 (“Standard model” in Table 2.2). For each model neuron, the same filter was applied to the exponentially correlated and uncorrelated stimuli in order to simulate cases without adaptation. Across the population, our model neurons showed only a slight bias in center latency between the two stimuli (Fig. 2.8A). While this effect reaches significance (for  $\alpha = .05$ ), the effect size is orders of magnitude smaller than that seen in the data and thus could not explain our experimental results. We also observed a tendency toward a slightly stronger relative surround strength under white noise than under correlated noise (Fig. 2.8B). Note that this is opposite the effect observed in our experimental results (Fig. 2.6D–F). Thus, if anything our results may be stronger than reported in the main text.

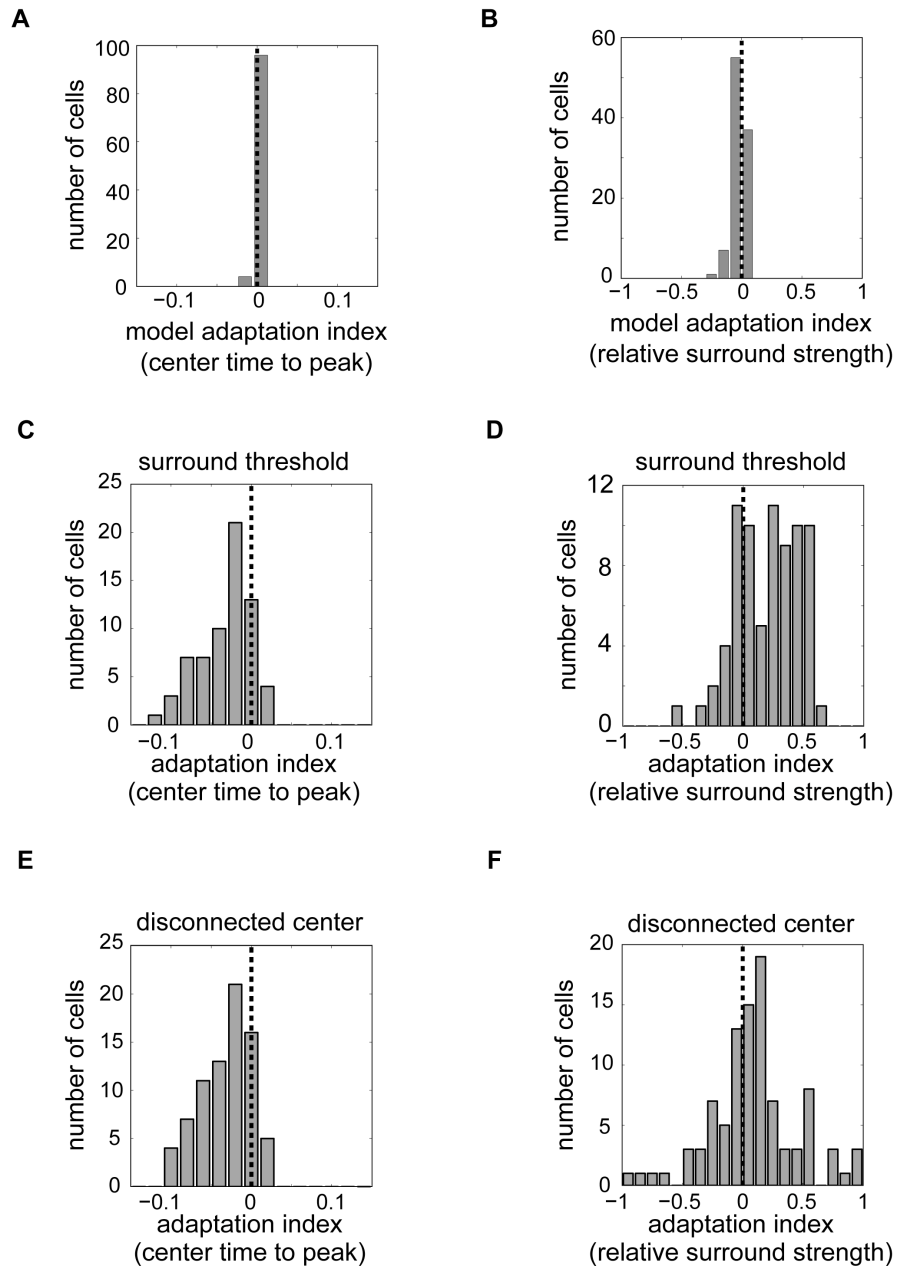
To further validate our analysis we ran simulations with an even wider range of model parameters. We first constructed model neurons with surround radii measured

from Gaussian distributions with means of 1 (“Small surround radius” in Table 2.2) or 3 (“Large surround radius”), both with standard deviation 0.3, and all other parameters the same as in our original simulation. In separate simulations, we kept the mean surround radius at 2 but chose the relative surround strength from a Gaussian distribution with mean 0.5 (“Small surround weight”) or 2 (“Large surround weight”), both with standard deviation 0.1. As with our original set of parameters, the models recovered from STRF analysis had at most slight biases toward weaker surrounds and shorter center times to peak under correlated noise.

**Tests of robustness.** We used the analysis method described above because it gave the most robustly unbiased results in our simulations, but we also wanted to verify that our results did not change dramatically with slightly different methods. To this end, we made a series of modifications to our receptive field extraction method and repeated the analyses described above for each modification.

In our experiments, we alternated white noise and correlated stimuli. The retina is known to adapt to a variety of stimulus features on timescales ranging from hundreds of milliseconds to a few seconds. Thus, it is possible that estimating receptive fields using the entire trials confounds different states of adaptation. To control for this possibility, we repeated our analyses leaving out the first 10 seconds of each trial and found that the results did not change.

In addition, we varied the details of the analysis method, as summarized in Table 2.3. We first considered the possibility that we were including too many pixels in the surround. To address this, we repeated our analysis but placed a threshold criterion on the surround so that only pixels positively correlated with the peak surround pixel were included. Including only these pixels, center latency was still shorter for correlated noise than for white noise (Fig. 2.8C; “Surround threshold”



**Figure 2.8**



**Figure 2.8: Receptive field results are validated with model neurons and are robust to analysis method changes.** (A) Center time to peak for a population of non-adapting model neurons, plotted as in Figure 2.6A. Receptive fields were constructed as a difference of Gaussians. Surround radii (relative to center radii) had a mean of 2 and a standard deviation of 0.3. Surround weights (relative to center weights) had a mean of 1 and a standard deviation of 0.1. The model neurons do not show a large difference between stimuli in center time to peak. (B) Model neurons described in A show a slight bias toward smaller recovered relative surround strength under correlated noise compared to white noise. (C) Center time to peak is longer for white noise than for correlated noise when the surround only includes non-center pixels whose time courses are positively correlated with the time course of the peak negative pixel. (D) Relative surround strengths is marginally higher for correlated noise than for white noise under the same analysis as in C. (E) Center time to peak is longer for white noise than for correlated noise when the center is not required to form a single contiguous component. The figure omits a few outliers that lie outside the range of the horizontal axis and have longer time courses under correlated noise. (F) Relative surround strengths is marginally higher for correlated noise than for white noise under the same analysis as in E.

in Table 2.3), while the increase in surround strength became more robust than in the original analysis (Fig. 2.8D). We then tested whether requiring the center to be contiguous was too stringent. Removing this criterion did not change our overall results, although the center time to peak statistics are skewed by a few outliers (Fig. 2.8E, F; “Disconnected center”).

When we computed STRFs, we included three frames after every spike so that we could measure any baseline offset in the estimated STRFs. We generally subtracted the mean of these three frames from each STRF, but skipping this step did not affect our results (“No mean subtraction”). As an additional check, we collapsed the full STRF into a single frame by projecting onto the first principal timecourse rather than the principal timecourse most similar to the peak pixel. Making this change does not affect our results (“First principal component”). (We obtained similar results when we chose the principal timecourse corresponding to the peak surround pixel rather

Table 2.3: Variants on receptive field analysis method.

	AI (relative surround strength)				AI (center time to peak)			
	mean	std	p	skewness	mean	std	p	skewness
Standard analysis	0.01	0.24	.008	0.15	-0.03	0.03	< .0001	-0.53
Disconnected center	0.10	0.36	.01	-0.01	-0.03	0.22	.2	-2.81
Surround threshold	0.21	0.25	< .0001	-0.40	-0.03	0.03	< .0001	-0.59
No mean subtraction	0.12	0.23	< .0001	0.08	-0.03	0.03	< .0001	-0.72
First principal component	0.07	0.24	.01	0.11	-0.03	0.03	< .0001	-0.48
Masks from WN	0.02	0.20	.4	0.04	-0.04	0.04	< .0001	-2.19
Masks from CN	-0.04	0.19	.08	-0.07	-0.03	0.03	< .0001	-0.45

Adaptation index (AI) in surround strength and center latency measured by variations in our analysis method.

Columns labeled “mean,” “std,” and “skew” show the mean, standard deviation, and skewness of the adaptation indices for the indicated analysis; columns labeled “*p*” show the *p*-values from *t*-tests of each distribution against the null hypothesis of zero mean. **Standard analysis:** Receptive fields were analyzed as presented in the main text.

**Disconnected center:** The center was not required to form a contiguous region. **Surround threshold:** A threshold criterion was used to find the surround so that only pixels positively correlated with the peak surround pixel were included. **No mean subtraction:** The mean of the frames after each spike was not subtracted from the STRFs. **First principal component:** The full STRF was collapsed onto a single frame by projecting onto the first principal timecourse rather than the principal timecourse most similar to the peak pixel. **Masks from WN:** Center and surround regions measured from white noise were used to analyze STRFs from both stimuli. **Masks from CN:** Center and surround regions measured from exponentially correlated noise were used to analyze STRFs from both stimuli.

than the peak center pixel.)

To investigate whether the changes we measured in receptive fields came from a change in the size or location of the receptive field center and surround or from a change in the receptive field strength at individual points in space, we repeated our standard analysis with the same masks for both stimuli. That is, for each cell we first found the center region based on the STRF measured from white noise and then computed the time courses of this region under each stimulus condition from the full STRFs. The surround time courses were computed similarly (“Masks from white noise”). We then did the reverse, finding the center and surround regions from correlated noise STRFs and applying them to both stimuli (“Masks from correlated noise”). In either case, the center latencies were still larger for white noise, indicating that the time courses of individual pixels differ when stimulus correlations change. On the other hand, the relative surround strength adaptation indices were centered around zero when masks were kept fixed. Thus, any changes in surround strength observed in our main analysis were likely due to a subset of pixels switching from center to surround or vice versa.

**Dependence of output correlation on gain and firing rate.** In the main text we found that gain adaptation contributes to decorrelation in a population of model LN neurons. To gain further insight into why this should be the case, consider a pair of simplified LN neurons with logistic nonlinearity,

$$P_i(\text{spike}|s) = 1 / (1 + \exp(-g_i(s_i - b_i))) \quad (2.1)$$

and small firing probabilities in each time bin. For neuron  $i$ , the gain of the model is  $g_i$ , the linear filter output is  $s_i$ , and  $b_i$  is an offset that will be adjusted to fix the

average firing rate. Assuming a small firing probability amounts to assuming that  $g_i(s_i - b_i) \ll 0$  with high probability. Thus, the exponential term in  $P_i(spike|s)$  dominates, and the stimulus-dependent firing probability simplifies to  $P_i(spike|s) = \exp(g_i(s_i - b_i))$ . The average firing probability of one neuron is now

$$P_i = \langle \exp(g_i(s_i - b_i)) \rangle, \quad (2.2)$$

where the average is over the distribution of the filter output  $s_i$ , which can be approximated by the central limit theorem as a zero-mean Gaussian. (A nonzero mean could simply be absorbed into a redefinition of  $b_i$ .) Using standard properties of Gaussian integrals, the averaging gives

$$P_i = \exp \left[ 1/2(g_i\sigma_i)^2 - g_ib_i \right], \quad (2.3)$$

where  $\sigma_i$  is the standard deviation of  $s_i$ . Note that, by the low firing probability assumption,  $P_1 \ll 1$  and  $P_2 \ll 1$ .

The average probability of simultaneous firing of two neurons is then given by

$$P_{12} = \langle \exp(g_1(s_1 - b_1)) \exp(g_2(s_2 - b_2)) \rangle \quad (2.4)$$

$$= \langle \exp(g_1s_1 + g_2s_2 - g_1b_1 - g_2b_2) \rangle. \quad (2.5)$$

Assuming that the filter outputs are jointly Gaussian with correlation  $\rho_s$ , the variance of  $g_1s_1 + g_2s_2$  is  $(g_1\sigma_1)^2 + (g_2\sigma_2)^2 + 2(g_1\sigma_1)\rho_s(g_2\sigma_2)$ . The expectation can therefore

be computed as

$$P_{12} = \exp \left[ 1/2((g_1\sigma_1)^2 + (g_2\sigma_2)^2 + 2(g_1\sigma_1)\rho_s(g_2\sigma_2)) - g_1b_1 - g_2b_2 \right] \quad (2.6)$$

$$= P_1P_2 \exp \left[ (g_1\sigma_1)\rho_s(g_2\sigma_2) \right]. \quad (2.7)$$

The variance in firing of each neuron is  $P_i(1 - P_i)$ , and the covariance between the two is  $P_{12} - P_1P_2$ . The correlation coefficient of the two spike trains is then given by

$$\rho = \frac{P_{12} - P_1P_2}{\sqrt{P_1(1 - P_1)P_2(1 - P_2)}}, \quad (2.8)$$

which, using the above result for  $P_{12}$  and taking the limit of small  $P_i$ , simplifies to

$$\rho = \sqrt{P_1P_2} \left( e^{(g_1\sigma_1)\rho_s(g_2\sigma_2)} - 1 \right). \quad (2.9)$$

Thus there are three ways to reduce output correlations in this simple model: lower the overall firing rates, decrease  $\rho_s$  by filter adaptation, or lower the rescaled gains  $g_i\sigma_i$ .

# Chapter 3

## Lateral inhibition maintains differences in stimulus correlations

### 3.1 Abstract

We showed in [Chapter 2](#) that changes in stimulus correlations do not produce equal changes in output correlations. Here we ask how the retinal circuitry achieves this relative invariance in correlation. We first examine outer retina processing and determine that feedback from horizontal cells *enhances* stimulus correlations. Moving to inner retina, we find that feedforward GABAergic inhibition onto ganglion cells may also enhance strong correlations in the input, while feedback inhibition onto bipolar cells does not affect output correlations. These results show that, contrary to naïve expectations, lateral inhibition in the retina does not necessarily act to decorrelate visual signals.

## 3.2 Introduction

Processing of visual stimuli by the retina, from photoreceptors to bipolar cells to ganglion cells, is shaped by several mechanisms of lateral information flow. Here we review the connections at each stage of processing.

### 3.2.1 Outer retina mechanisms

Visual processing begins in the outer plexiform layer, where photoreceptors convert light into a chemical signal passed to bipolar cells. This signal transmission is shaped by glutamate receptors on bipolar cells and by lateral connections, mainly via horizontal cells.

**Feedforward signaling to bipolar cells.** In the dark, photoreceptors chronically release glutamate from ribbon synapses. In response to a light stimulus, glutamate release is decreased. Glutamate is detected differently by OFF and ON bipolar cells. OFF bipolar cells have ionotropic glutamate receptors that produce membrane depolarization in the presence of glutamate. ON bipolar cells, on the other hand, have the metabotropic glutamate receptor mGluR6. Activation of mGluR6 by glutamate results in hyperpolarization of the bipolar cell. Thus, at light onset, photoreceptors hyperpolarize and release less glutamate. As a result, OFF bipolar cells also hyperpolarize, while ON bipolar cells depolarize. The opposite pattern occurs at light offset: photoreceptors depolarize, resulting in depolarization of OFF bipolars and hyperpolarization of ON bipolars.

**Photoreceptor coupling.** The first lateral connections in the retina occur between photoreceptors. Mammalian cones are electrically coupled through gap junction channels, which allow current to flow directly between connected cells ([DeVries](#)

et al., 2002). These gap junctions, formed by connexin-36 in cone pedicles (Lee et al., 2003; Feigenspan et al., 2004), spread the signal in each cone to its neighbors. This coupling can reduce acuity through blurring but also greatly increases signal strength by diluting photoreceptor noise that is not shared between neighboring cones. Rods are also coupled to each other and to cones to decrease noise and improve detection at scotopic light levels (Bloomfield and Völgyi, 2009). Gap junctions between photoreceptors are modulated throughout the day, with greater electrical coupling at night (Li et al., 2009). This cycle enables summing of weak signals at night while avoiding saturation during the day.

**Feedback inhibition onto photoreceptors.** An important source of lateral inhibition in the retina is feedback from horizontal cells to photoreceptors. Horizontal cell feedback has been shown to contribute to the surround measured in ganglion cells (Davenport et al., 2008). When cones hyperpolarize in response to light, they release less glutamate, leading to hyperpolarization of horizontal cells. Feedback from horizontal cells then results in a shift in the activation curve of  $\text{Ca}^{2+}$  channels on cone pedicles, so that the channels open (thus producing an inward, depolarizing current) at more hyperpolarized membrane potentials. The exact mechanism of this feedback has been the subject of much investigation (Kamermans and Fahrenfort, 2004). One explanation, the *ephaptic hypothesis*, posits a dependence on the intricate nature of synapses in the outer retina. In these invaginating synapses, small processes from bipolar and horizontal cells are surrounded by the cone pedicle. Thus, the synaptic cleft is relatively small and isolated from other extracellular spaces. Under the ephaptic hypothesis, hyperpolarization of horizontal cells leads to an inward current through hemichannels on the horizontal cells. This current will in turn hyperpolarize the intersynaptic space between the cone and horizontal cell. From the perspective of



calcium channels on the cone membrane near the ribbon synapse, hyperpolarization of the intersynaptic space appears the same as depolarization of the cone. As a result, calcium channels are activated at more negative membrane voltages than they would be otherwise. The resulting influx of calcium depolarizes the cone, thus opposing the direct light-induced hyperpolarization. (Kamermans and Fahrenfort, 2004)

An opposing explanation, the *pH hypothesis*, again relies on hyperpolarization of horizontal cells. In this view, hyperpolarization increases the driving force for an influx of protons into horizontal cells, leading to alkalization of the synaptic cleft. This change in pH then affects activation curves of  $\text{Ca}^{2+}$  channels in the cone membrane so that calcium flows into cones at more hyperpolarized membrane potentials. This mechanism, like the ephaptic mechanism, would lead to depolarization of cones (Kamermans and Fahrenfort, 2004). Davenport et al. (2008) found that application of pH buffers such as HEPES decreased feedback from horizontal cells to cones, giving some support to the pH hypothesis. However, because changes in pH can also affect the conductance of hemichannels, these results are still ambiguous (Fahrenfort et al., 2009).

**Horizontal cell coupling.** Horizontal cells are coupled to each other by gap junctions involving connexin-57 (Shelley et al., 2006). This coupling produces an extensive syncytium that greatly increases the receptive field size of horizontal cells (Bloomfield and Völgyi, 2009). As in photoreceptors, gap junctions between horizontal cells are modulated by light level. In this case, prolonged exposure to dark *or* light conditions reduces coupling; horizontal cells are connected most effectively at intermediate light levels (Xin and Bloomfield, 1999).

**Feedforward inhibition onto bipolar cells.** There is also some evidence for feedforward inhibition from horizontal cells to bipolar cells, although this connection has not been studied as thoroughly as feedback inhibition to photoreceptors. Depolarization of horizontal cells (as occurs after a light decrement) is known to depolarize nearby ON bipolar cells and hyperpolarize nearby OFF bipolar cells (Toyoda and Kujiraoka, 1982). While this effect could potentially be an indirect result of feedback from horizontal cells to photoreceptors, roughly half of ON bipolar cells in salamander retina appear to receive surround inhibition from horizontal cells even when glutamate input is blocked (Fahey and Burkhardt, 2003). This phenomenon has not been tested extensively in mammalian retina, however, and is dependent on background light level (Thoreson and Mangel, 2012).

Feedforward input from horizontal cells to bipolar cells could be mediated by GABA. Although horizontal cells do not contain the typical machinery for GABA synthesis and release, they may be able to produce GABA from glutamate (Deniz et al., 2011). ON and OFF bipolar cell dendrites express the same type of GABA<sub>A</sub> receptors and at least one GABA<sub>C</sub> receptor subunit (Vardi and Sterling, 1994). The same receptor could produce opposite effects in ON and OFF cells if the two cell types have different intracellular chloride concentrations, resulting in different chloride equilibrium potentials (Thoreson and Mangel, 2012). In at least some bipolar cell types, the chloride gradient is consistent with this explanation (Duebel et al., 2006). Moreover, ON and OFF bipolar cells express different chloride cotransporters that produce chloride efflux in ON cells and chloride influx in OFF cells (Vardi et al., 2000; Vu et al., 2000). While this evidence points to direct GABA-mediated inhibition from horizontal cells to bipolar cells, experiments using GABA antagonists have so far given inconclusive results (Thoreson and Mangel, 2012). Thus, caution must be used in interpreting studies of feedforward inhibition in outer retina.

### 3.2.2 Inner retina mechanisms

In the inner plexiform layer, bipolar cells provide excitatory input to ganglion cells. Amacrine cells inhibit both bipolar and ganglion cells. These two pathways may serve overlapping functions. For example, stimulation of the receptive field surround decreases ganglion cell response both by directly inhibiting RGCs and by decreasing excitatory input. Inhibition in the inner retina can come in many forms, due in part to the large number of amacrine cell sub-types.

**Feedforward excitation onto ganglion cells.** Bipolars release glutamate onto ganglion cells to produce excitation, which is shaped by presynaptic inhibition, discussed below. Although most glutamate receptors on RGCs are ionotropic ([Massey and Miller, 1988](#)), the mixture of subtypes (NMDA, AMPA, and kainate) on a given cell gives rise to its particular response properties ([Manookin et al., 2010](#); [Peng et al., 1995](#)). In addition, the rate of glutamate clearance varies between cell types and may contribute to the range of RGC response properties ([Lukasiewicz, 2005](#)).

**Feedback inhibition onto bipolar cells.** The indirect contribution to lateral inhibition of ganglion cells is mediated by GABA<sub>A</sub>, GABA<sub>C</sub>, and glycine receptors on bipolar cell terminals ([Lukasiewicz, 2005](#); [Wassle et al., 1998](#)). GABA<sub>C</sub> receptors contribute more than GABA<sub>A</sub> receptors. In addition, feedback from sustained OFF amacrine cells affects many ganglion cells indirectly through GABA<sub>B</sub> receptors on bipolar cell terminals ([de Vries et al., 2011](#)).

In addition, amacrine and bipolar cells can be connected by gap junctions. The “direct” or “primary” pathway of rod signaling depends on these connections ([Bloomfield and Völgyi, 2009](#)). Signals from rods are inverted and passed to rod bipolar cells. These bipolar cells do not synapse directly onto ganglion cells but instead pass their

signals to AII amacrine cells. AII cells then communicate with cone bipolars through gap junctions involving connexin-36 and possibly connexin-45. As a result, the rod signal travels from rod bipolars to AII amacrine cells to cone bipolars and then to the ganglion cell layer.

**Feedforward inhibition onto ganglion cells.** Direct chemical inhibition in the inner retina occurs through GABA<sub>A</sub> and glycine receptors on RGC dendrites. Depending on the type of amacrine cell, this inhibition can generally be divided into three groups: local (narrow-field) glycinergic, local (medium-field) GABAergic, and broad (wide-field) GABAergic (Chen et al., 2010; Werblin, 2011). Ganglion cells also have GABA<sub>B</sub> but not GABA<sub>C</sub> receptors (Wassle et al., 1998), and amacrine and ganglion cells are frequently coupled by gap junctions (Bloomfield and Völgyi, 2009).

Glycine released by narrow-field amacrine cells is largely limited to local crossover inhibition between the ON and OFF pathways (Werblin, 2011) and does not appear to contribute to surround inhibition (Flores-Herr et al., 2001). Medium-field amacrine cells, on the other hand, release GABA over a range of  $\sim 250\mu\text{m}$ , similar to the size of a ganglion cell receptive field. GABAergic amacrine cells are thought to provide direct lateral inhibition of ganglion cells (Flores-Herr et al., 2001) and have relatively long response latencies (Chen et al., 2010).

The specific properties of each form of inhibition can shape ganglion cell responses differently. For example, OFF ganglion cells receive direct crossover inhibition from the ON pathway in response to light, allowing responses to decrements in light level but not to increments. ON ganglion cells, on the other hand, are inhibited indirectly by a decrease in excitation from bipolar cells and thus can respond to both increments and decrements. As a result, OFF brisk-transient cells show greater rectification in membrane and spiking responses than their ON counterparts (Zaghloul et al., 2003).

In other cases, increasing the strength of amacrine cell responses driven by a visual stimulus increases ganglion cell firing rates ([Manu and Baccus, 2011](#)). This may be the result of a release from inhibition of the RGC when the amacrine cell goes from depolarization to hyperpolarization frequently, as may happen with a time-varying input.

**Amacrine cell coupling.** In addition to their coupling with bipolar cells mentioned above, AII amacrine cells are coupled to each other with connexin-36 gap junctions ([Bloomfield and Völgyi, 2009](#)). These connections serve to reduce uncorrelated noise, thus strengthening the signal-to-noise ratio of the rod pathway and preventing changes at low light levels from being obscured by noise. There are also glycinergic amacrine-to-amacrine synapses, but GABAergic inhibition of amacrine cells is less common. For example, narrow-field amacrine cells receive transient GABAergic inhibition and sustained glycinergic inhibition, whereas wide-field amacrine cells receive sustained glycinergic inhibition only ([Chen et al., 2011](#)).

**Ganglion cell coupling.** Finally, gap junctions connect ganglion cells directly. This electrical coupling appears to be restricted to connections between ganglion cells of the same type. Gap junctions between ganglion cells typically have low conductance, thus giving the direct coupling a relatively limited scope ([Bloomfield and Völgyi, 2009](#)). Ganglion cells can also be coupled indirectly through amacrine cells ([Trong and Rieke, 2008](#)).

### 3.2.3 Flexibility in retinal circuitry

In [Chapter 2](#), we showed that retinal processing is sensitive to correlations in the input. Ganglion cell responses also adapt to the luminance and contrast of a stimulus

ensemble. Here we review what is known about the circuit mechanisms underlying these forms of flexibility.

**Dependence on light level.** The different mechanisms of lateral inhibition are thought to contribute under different light levels, with GABAergic amacrine cells playing a greater role under dim illumination and horizontal cells becoming more important under bright illumination ([Lukasiewicz, 2005](#)). Part of this dependence on light level may be due to dopamine and nitric oxide levels, as both of these molecules are modulated by light levels and reduce gap-junction coupling of horizontal cells ([Bloomfield and Völgyi, 2009](#)). Horizontal cell coupling can be decreased by prolonged darkness or prolonged brightness. Maximal horizontal cell coupling occurs at intermediate background lighting, partially explaining the greater role of horizontal cells in surround inhibition at these light levels. A similar dopamine-mediated dependence on illuminance is seen in AII-AII coupling, with the greatest conductances again observed at intermediate light levels ([Bloomfield and Völgyi, 2009](#)).

Unlike connections among horizontal cells and amacrine cells, gap junctions between alpha ganglion cells increase conductance in response to dopamine. As a result, ganglion cells are more strongly coupled in the light-adapted state when retinal dopamine levels are higher ([Bloomfield and Völgyi, 2009](#)).

**Adaptation to contrast.** A number of studies have sought to determine where in the retina temporal contrast adaptation originates. The earliest stage of processing at which contrast adaptation has been observed is in bipolar cells ([Rieke, 2001](#); [Baccus and Meister, 2002](#)). This adaptation does not depend on inhibition from horizontal or amacrine cells or on voltage-dependent processes ([Rieke, 2001](#)), suggesting a mechanism within dendrites. At the same time, horizontal cell feedback may provide a

mechanism for gain control in photoreceptors (VanLeeuwen et al., 2009). Adaptation observed in RGCs may then be due in part to adaptation of bipolar cells, as contrast gain control occurs for changes in the location of a small stimulus within the RGC receptive field center (Brown and Masland, 2001) and is stronger for a large contrast change in a small spot than for a small contrast change over the entire center (Beaudoin et al., 2007). Furthermore, contrast adaptation is not abolished by voltage clamping of ganglion cells (Beaudoin et al., 2007) or pharmacological blockers of amacrine cells (Beaudoin et al., 2007; Brown and Masland, 2001; Manookin and Demb, 2006). These findings indicate that inner retina processing, while likely contributing to contrast adaptation, is not required.

Most amacrine cells do not adapt to contrast, but some do change potential slowly after a contrast switch (Baccus and Meister, 2002). Beaudoin et al. (2008) measured a gain reduction in both excitatory and inhibitory conductances in ganglion cells following an increase in contrast, suggesting that both bipolar and amacrine cells adapt to stimulus variance. On the other hand, Brown and Masland (2001) did not see any weakening of contrast gain control when GABA and glycine were blocked. Thus, while direct inhibition to ganglion cells may have some role in contrast adaptation, it does not appear to be crucial.

Finally, adaptation also occurs in ganglion cells in the conversion of current input to spike output. This is demonstrated most clearly by the observation that gain control is stronger in spiking responses than in input currents. In addition, RGC spiking adapts to the variance of current injections (Kim and Rieke, 2001). This additional contrast gain control may be due in part to inactivation of sodium channels. Fluctuations in input current to RGCs lead to inactivation of sodium channels with a slow recovery, and simulations show that this inactivation could explain adaptation of spiking to input variance (Kim and Rieke, 2003). Thus, intrinsic mechanisms

within ganglion cells appear to enhance the adaptation to contrast that already occurs presynaptically.

**Sources of correlated firing.** Our current understanding of the mechanisms shaping correlations between ganglion cells is fairly limited, although the presence of such correlations is well-known ([Mastrorade, 1989](#); [DeVries, 1999](#)).

Electrical coupling throughout the retina likely shapes output correlations. The gap junctions between ganglion cells have a low conductance and can thus produce coherent firing in a small patch of cells without diluting the signal across wide swaths of retina. This coupling can lead to reciprocal excitation between two neighboring ganglion cells, resulting in short (1-3 ms) spike latencies between the cells ([Bloomfield and Völgyi, 2009](#)). Shared electrical input from a common amacrine cell can also lead to synchronous activity in two nearby ganglion cells ([DeVries, 1999](#)). Coupling elsewhere in the retina could similarly enhance correlations between nearby cells. As discussed above, many gap junctions are modulated by light level. If any such modulation occur on a fast time-scale or is sensitive to complex features of visual scenes, then the resulting correlations could be stimulus-dependent.

In addition, chemical synapses appear to contribute to retinal correlations. Synaptic inputs to pairs of ganglion cells show correlations that are affected by the presence or absence of a visual stimulus ([Trong and Rieke, 2008](#)). These correlations are observed when looking at pairs of excitatory inputs and pairs of inhibitory inputs, suggesting that at least some correlations in RGC output are determined pre-synaptically. Furthermore, ganglion cell types that display correlated spiking are not always electrically coupled ([DeVries, 1999](#)). Although both of these lines of evidence could be explained if correlations present in bipolar cell dendrites are simply carried forward to the inner retina, gap junctions alone are unlikely to account for all correlated firing



in the retina.

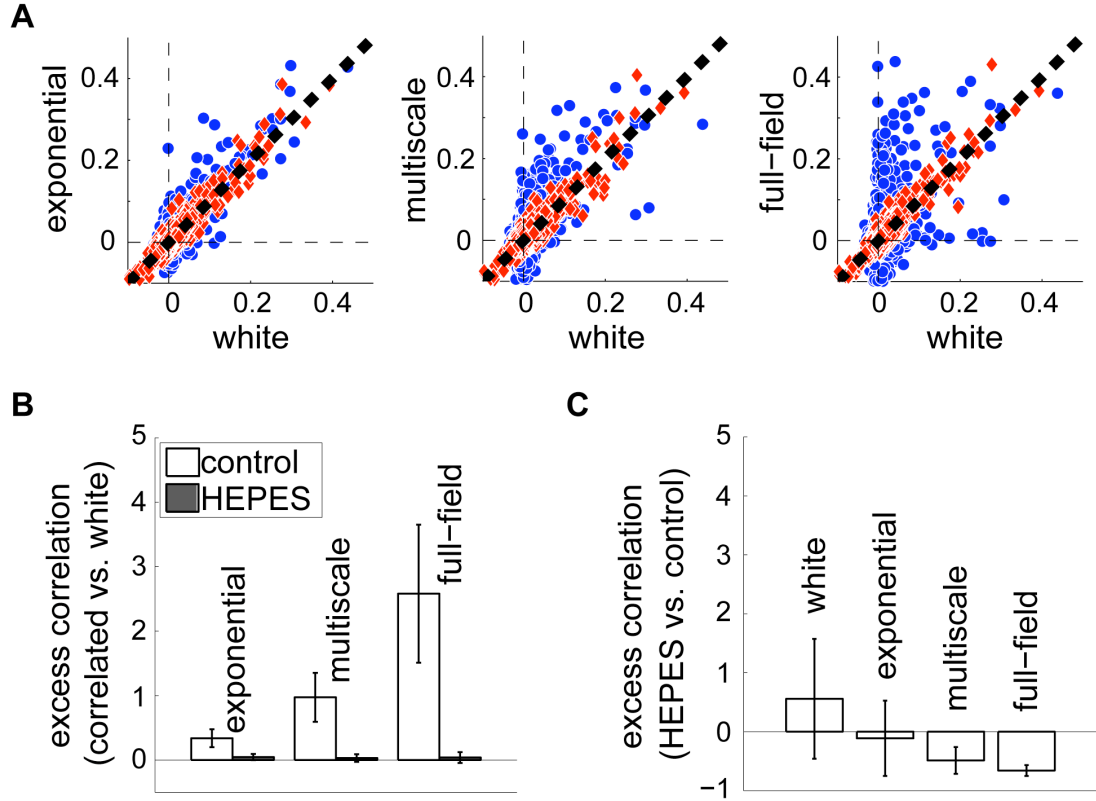
In the experiments reported here, we sought to understand how retinal circuitry shapes correlations, particularly the dependence on stimulus correlations. We first blocked outer retina and uncovered a possible role for horizontal cells in maintaining redundancy. We then blocked lateral connections in inner retina. Our findings suggest that direct inhibition of ganglion cells by GABA<sub>A</sub> receptors might also enhance changes in output correlations but GABA<sub>C</sub> and glycine receptors are less likely to be involved.

## 3.3 Results

### 3.3.1 Feedback in outer retina enhances stimulus correlations

We began by testing the contribution of the outer retina to stimulus-dependent correlations between ganglion cells. To do this, we added the pH buffer HEPES to the perfusate. Although the exact mechanism of horizontal cell feedback to photoreceptors is debated, HEPES has been shown to block this feedback and to decrease the surround measured in ganglion cells (Davenport et al., 2008). We initially hypothesized that blocking this source of lateral inhibition would increase correlations in retinal output and make them more heavily dependent on stimulus correlations.

**Invariance of spike train correlations.** To measure the effects of horizontal cell feedback on retinal output correlations, we binned spike trains with 33-ms bins and found correlation coefficients between spike trains of pairs of simultaneously recorded ganglion cells, as in Chapter 2. In control conditions, as discussed above, output correlations tended to be stronger for more correlated stimuli (Fig. 3.1A, B). When HEPES was applied, the output correlations became almost completely invariant with



**Figure 3.1: HEPES removes stimulus dependence of output correlations.**

(A) Pairwise correlations for three types of correlated stimuli compared to an uncorrelated flickering white noise checkerboard. *Blue circles* represent measurements made in control conditions; *red diamonds* are in HEPES. HEPES makes correlations more invariant between stimuli.  $N = 1953$  pairs. (B) Excess correlation computed from the data in A. All excess correlations are computed with respect to the white noise checkerboard responses. Error bars are calculated by bootstrapping. (C) Excess correlation, computed as in B but comparing HEPES against control conditions separately for each stimulus.

respect to the stimulus. On a population level, HEPES tended to weaken correlations in response to multiscale and full-field stimuli without significantly changing the strength of correlations for spatially white and exponentially correlated noise (Fig. 3.1C). The more striking effect, however, was on individual pairs of cells. Some pairs maintained approximately equal correlations across stimuli by increasing white noise checkerboard correlations when HEPES applied; for others, HEPES decreased output correlations in response to highly correlated stimuli.

Thus, contrary to our naïve hypothesis, it appears that horizontal cell feedback does not explain the relative decorrelation observed in Chapter 2. In fact, horizontal cells may actually *increase* the stimulus dependence of output correlations. One might ask whether the addition of HEPES made the retina insensitive to light. If this were the case, then our measurements of output correlations would reflect spontaneous activity and hence be independent of the stimulus class. For several reasons, we do not believe this is the case. First, we measured the average firing rates in each condition. For most stimuli, activity levels increased when HEPES was applied (Fig. 3.2A), but the firing rates were still within a physiological range. This is one indication that the retina was still healthy in the presence of HEPES.

We also looked directly at visual responses to alternating flashes of light and dark. We recorded responses to such a stimulus immediately before and after each of our main experimental trials (that is, before and after measuring correlations in the control condition and before and after making the same measurements in HEPES). Figure 3.2B shows the average firing rate of a few cells in response to such a stimulus. We can see from these responses that cells maintained their general responsiveness throughout the experiment, even after addition of HEPES. To summarize this observation, we averaged across time bins to obtain, for each RGC, the mean firing rate during the ON phase and the mean firing rate during the OFF phase. Comparing

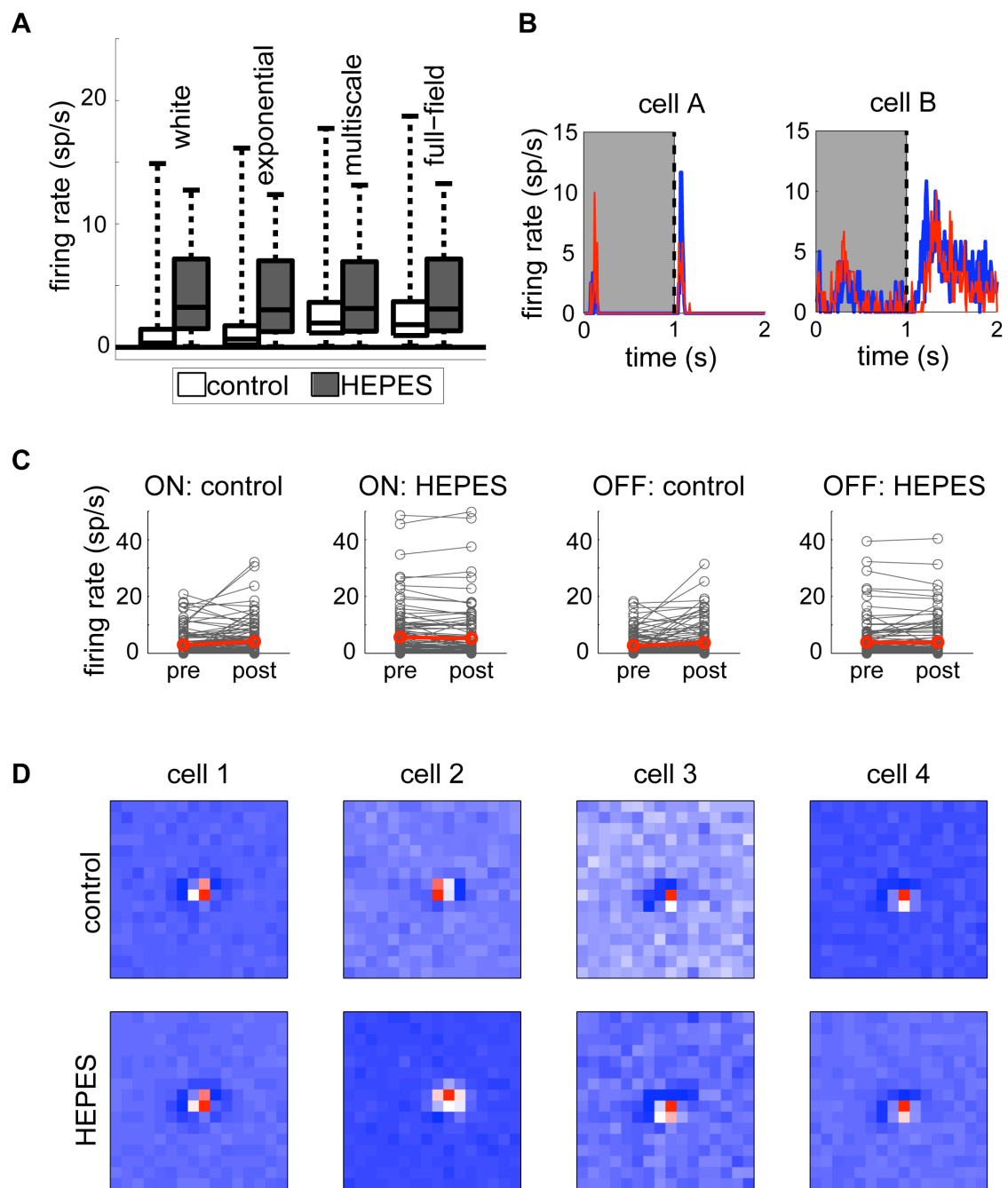


Figure 3.2

**Figure 3.2: HEPES does not abolish light responses.** **(A)** Average firing rates for each stimulus in control condition (*white*) and in the presence of HEPES (*gray*). Each box spans the 25th to 75th percentiles for the indicated stimulus with the central line denoting the median; whiskers indicate the full range. **(B)** Representative traces of a ganglion cell responding to a 1-second alternating flash of dark (*shaded area*) and light (*white area*) in the control condition (*blue*) and in HEPES (*red*). **(C)** Average firing rate for all recorded cells in response to a 1-second flash of light. Recordings were taken immediately before and after the main experimental trials (i.e.  $\sim 50$  minutes apart). Although HEPES affects firing rate, activity is stable throughout the HEPES recording. **(D)** Representative examples of receptive fields measured from white noise responses in control condition (*top row*) and in HEPES (*bottom row*). Each image shows the spatial component of one receptive field measured by reverse correlation. Note that the spatial component is defined here to have a positive peak; thus all frames appear to show ON responses even for OFF cells.

these values, we see that responses again tended to be stronger under HEPES (Fig. 3.2C) but were still in a reasonable physiological range. Moreover, the stability of the responses measured before and after our experimental trials suggests that the retina was still healthy and responsive at the end of the experiment.

Even with these robust light responses, the retina still could have become insensitive to the small-scale structure of our stimuli. To test for this, we estimated the spatio-temporal receptive field of each cell in response to white noise by reverse correlation. As shown for a few cells in Figure 3.2D, we could extract receptive fields in the presence of HEPES, indicating responsiveness to spatially and temporally varying checkerboards. Thus, we are confident that our findings in the presence of HEPES are not an artifact of insensitivity to visual stimuli but instead indicate that the retina is able to equilibrate output correlations when horizontal cell feedback is blocked. While surprising, this result could be evidence of a need for redundancy to be maintained in early visual processing. This possibility is returned to in the [Discussion](#).

**Information rates.** To further understand the effects of HEPES in our experiments, we turned to analyses of single cell responses. Since our results thus far suggest that horizontal cell feedback may increase redundancy by allowing stimulus correlations to trigger more correlated output, we sought to determine whether horizontal cell feedback was beneficial in terms of any other information theoretic measures.

We first looked at the information carried by a single spike in each condition (see [Methods](#) for details). In the control condition, single-spike information values had a narrower distribution for a more correlated stimulus. That is, neurons with low information rates (below  $\sim 3$  bits/spike) when probed with a checkerboard had increased information rates under full-field flicker, whereas cells with high information rates for the checkerboard had decreased information for the full-field flicker. The same general pattern held when horizontal cell feedback was blocked with HEPES ([Fig. 3.3A](#)). The addition of HEPES increased single-spike information about the spatially uniform stimulus (t-test  $p < .01$ ,  $N = 103$  cells) without significantly affecting information during the white noise checkerboard presentation ( $p = .10$ ). The difference in information rate between full-field and spatially modulated stimuli increased when HEPES was applied, although this effect was small for most cells. Thus, it is possible that horizontal cell feedback acts to maintain similar information rates across stimuli.

**Signal-to-noise ratio.** In addition, we computed the signal-to-noise ratio (SNR; see [Methods](#)) for each cell under each of the four conditions. As shown in [Figure 3.3B](#), SNR was greater for the highly correlated full-field stimulus than for the checkerboard stimulus in almost all cells, both in the control solution and in HEPES. Within a given stimulus type, HEPES tended to increase SNR (t-tests  $p < .05$  for each stimulus,  $N = 96$  cells). HEPES also tended to decrease the fractional change in

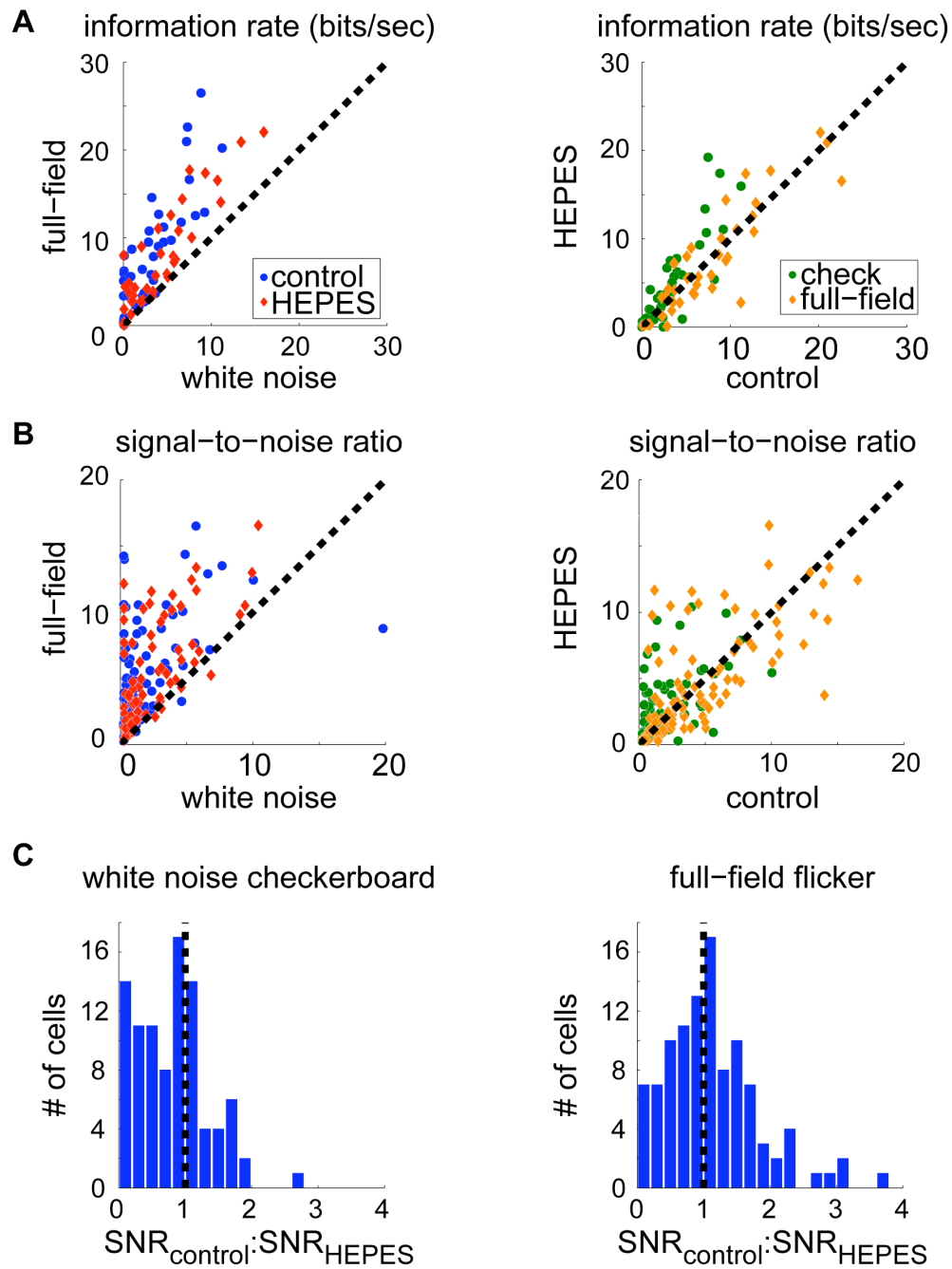


Figure 3.3

**Figure 3.3: Effects of HEPES on information.** (A) *Left*, single-spike information in control conditions (*blue*) and in HEPES (*red*) is consistently higher for full-field flicker than for a white noise checkerboard. The stimulus consisted of 30 seconds of a flickering white noise checkerboard and 30 seconds of full-field flicker. The same stimulus sequence was repeated 60 times. *Right*, HEPES increases information rates for full-field flicker (*orange*) but not for the spatio-temporally modulated checkerboard (*green*).  $N = 108$  cells. (B) Signal-to-noise ratio (SNR), plotted as in A. SNR tends to increase under HEPES and is consistently greater for the full-field stimulus than for spatially uncorrelated white noise. A few cells with large SNR values are left off at this scale but are included in all analyses. (C) Ratios of SNR in control condition to SNR with HEPES applied. *Black dotted line* indicates a 1:1 ratio.

SNR produced by the increase in stimulus correlation (Fig. 3.3C; t-test  $p < .005$ ). That is, SNR values for white noise and full-field stimuli are more similar when horizontal cell feedback is blocked. This suggests that horizontal cell feedback serves to make SNR more dependent on stimulus type, as we saw with pairwise correlations.

**Receptive fields.** As noted above, we can estimate receptive fields from ganglion cells recorded in the presence of HEPES. Ideally, we would want to measure responses of a population of cells to stimuli with varying correlations in the presence and absence of HEPES. Due to the amount of data needed for STRF estimation and limits on recording stability, we were not able to record from all conditions in the same cells. Instead, we presented white and exponentially correlated checkerboard stimuli in HEPES only and found that the receptive field adaptations in HEPES are consistent with those reported for the control medium in Chapter 2. Namely, center latencies were faster in response to correlated noise (t-test on adaptation index,  $p < .001$ ,  $N = 22$ ) and relative surround strengths tended to be stronger for correlated noise (t-test  $p = .02$ ). However, the small number of recorded cells makes it difficult to compare the magnitude of these changes. Future experiments test the effects of



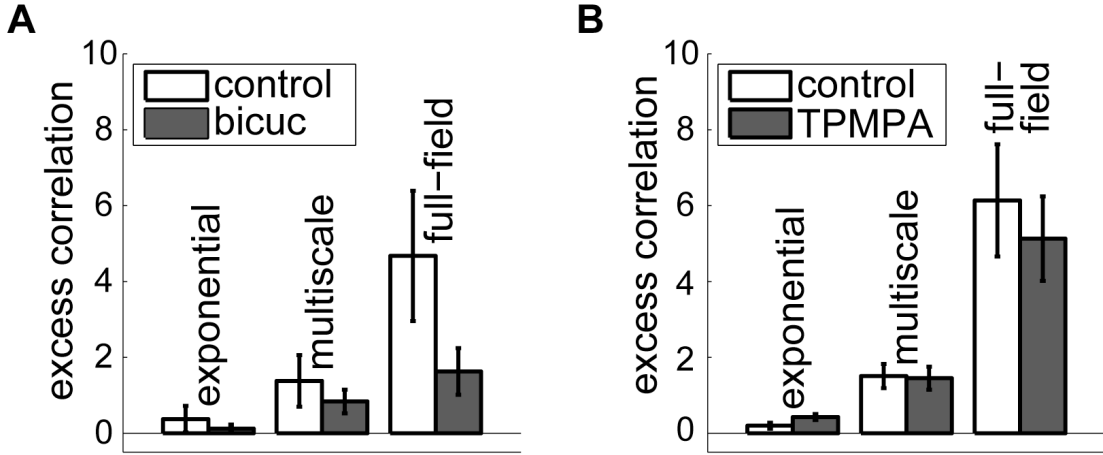
HEPES on receptive fields measured under each stimulus condition independently.

### 3.3.2 Inner retina may enhance extreme stimulus correlations

**GABA.** Lateral inhibition in the inner retina is largely a result of GABA released from amacrine cells onto bipolar and ganglion cells (Werblin, 2011). Thus, we suspected that blocking GABA receptors might affect the partial decorrelation seen in Chapter 2. To test this hypothesis, we followed the same procedure as described above but with the GABA<sub>A</sub> receptor antagonist bicuculline or the GABA<sub>C</sub> receptor antagonist TPMPA in the solution instead of HEPES.

In the presence of bicuculline, we observed a tendency for pairs of ganglion cells to have greater excess correlation than in the control condition, especially when comparing full-field flicker to spatially uncorrelated white noise (Fig. 3.4A;  $N = 87$  cells). TPMPA, on the other hand, did not appear to have an effect on the stimulus dependence of output correlations (Fig. 3.4B;  $N = 66$ ). Neither drug had a strong effect on the signal-to-noise ratios or information rates. These results suggest that amacrine cell inhibition may preserve correlations in highly correlated stimuli. This effect is most likely due to direct inhibition of ganglion cells, as GABA<sub>A</sub> receptors play a more minor role in indirect inhibition through bipolar cells (Lukasiewicz, 2005). The lack of an effect from blocking GABA<sub>C</sub> receptors, which are present in bipolar cells but not RGCs (Wassle et al., 1998), also lends support to this interpretation.

**Glycine.** In the retina, glycine is used largely to create crossover inhibition from the ON pathway to the OFF pathway and vice versa. The processes of glycinergic amacrine cells cover a narrow field, and thus glycine only spreads a short distance



**Figure 3.4: Blocking GABA<sub>A</sub> receptors may reduce changes in correlation.** (A) Excess correlation in control condition (*white*) and in the presence of the GABA<sub>A</sub> receptor blocker bicuculline (*gray*) are similar in response to stimuli with exponential or multiscale correlations (*left* and *middle bars*), but bicuculline decreases excess correlation for full-field flicker (*right bars*). Excess correlation is computed as in Figure 3.1B, with each correlated stimulus compared to a flickering white noise checkerboard.  $N = 87$  pairs. (B) Excess correlation is similar in control condition and in the presence of the GABA<sub>C</sub> receptor blocker TPMPA.  $N = 66$  pairs.

from the point of release. While crossover inhibition could affect correlations between neighboring ON and OFF cells, we did not expect glycine to have a large role in shaping correlations between ganglion cells in general. Consistent with this, we did not observe an effect of strychnine on output correlations, although the small number of cells recorded (21) prevents us from making a strong conclusion.

### 3.4 Discussion

We have shown that blocking horizontal cell feedback with HEPES makes pairwise correlations in retinal output invariant with respect to stimulus correlations. That

is, lateral inhibition in the outer retina appears to *preserve* stimulus correlations. Similarly, the tendency to decrease excess correlation in the the presence of GABA blockers is consistent with the possibility that lateral inhibition in the inner retina may partially remove stimulus correlations.

Why might lateral connections in the retina induce stimulus-dependent correlations? Redundancy may be important for processing noisy signals. Indeed, early theories of redundancy reduction stipulate that the optimal code is one that minimizes the loss of information (Barlow, 1961; Atick, 1992). More precisely, Tkačik et al. (2010) examined the coding strategies for a maximum-entropy network with pairwise correlations. When noise levels are high and the stimulus is correlated, the optimal strategy is to have positive correlations between neurons. Such redundancy strengthens the signal, whereas independent noise is averaged out. Thus, correlations in the retina could be needed to overcome noise and transfer signals to the rest of the visual system. Horizontal and amacrine cells could conceivably be responsible for adjusting these correlations as the noise level of the signal changes. Any overall decorrelation might occur later in the visual system. In addition, it is possible that the degree of retinal correlation itself provides information to later visual areas about the structure of the current stimulus ensemble.

*How* horizontal and amacrine cells enhance differences in stimulus correlations is less clear. Stimulus correlations could lead to correlations in nearby horizontal cell responses and thus correlated inhibition of nearby photoreceptors. This explanation would imply correlated output of nearby photoreceptors onto horizontal cells, however, and thus the necessity of horizontal cells is not obvious. Horizontal cells might also shape the timing of responses in photoreceptors, and thus in ganglion cells as well. While this may be an important factor in shaping retinal output, the results reported here hold when correlations are measured in bins of 200 ms where jitter in

spike timing should not have an effect.

Another possibility is that connections between horizontal cells change when stimulus correlations change. Coupling strength between horizontal cells changes depending on the background light level due to changes in dopamine and nitric oxide levels. When dopamine or nitric oxide levels are increased in horizontal cell culture, uncoupling peaks after two minutes and recovers after five minutes or more (Lasater and Dowling, 1985; Lu and McMahon, 1997); such changes occur in about half as much time in the intact retina (Pottet et al., 1997). Thus, our five-minute stimulus segments were likely adequate for dopamine or nitric oxide to exert some effect, but the slow time course of these agents makes it unlikely that they are the sole mediators of the observed effects. HEPES may also have non-specific effects in the retina due to its global effect on pH. Future experiments should use other blockers of horizontal cell feedback (such as cobalt and carbenoxolone) to test for specificity of the results reported here and should test the combined effects of HEPES and other pharmacological agents.

Similarly, it is unclear how lateral inhibition in the inner retina may affect output correlations. Our results thus far suggest a mechanism whereby direct GABAergic inhibition from amacrine cells to ganglion cells maintains changes in stimulus correlations. This effect appears to be modest in amplitude, although GABAergic signaling might be expected to play a greater role at low light levels. Alternatively, amacrine cells may play a role in shaping output correlations in some cell types but not others, leading to a small average effect.

Future work should seek to clarify the roles of GABAergic and ephaptic inhibition in retinal response to stimulus correlations. In particular, it will be important to test the effects of GABA<sub>A</sub> receptor antagonists and horizontal cell blockers at different background light levels. It would also be useful to examine the effects of GABA<sub>A</sub>

blockers when horizontal cell feedback is blocked and vice versa to determine whether inner and outer retina shape output correlations synergistically or independently. Recording from horizontal and bipolar cells, while technically challenging, would also be helpful in elucidating the effects of outer retinal processing on redundancy and correlations in retinal output. Finally, gap junctions are widespread in the retina and likely to play a role in coordinated output. We have not examined their role in this thesis, but such measurements will be crucial to fully understanding the mechanisms shaping correlations in retinal output.

### 3.5 Methods

**Neural recording.** The general procedure was the same as that described in [Chapter 2](#). Briefly, we recorded from isolated guinea pig retina from using a multi-electrode array with two panels of 30 electrodes (Multi Channel Systems MCS GmbH, Reutlingen, Germany). Stimuli were displayed on a Lucivid monitor (MicroBrightField, Inc., Colchester, VT) and projected onto the retina.

In some trials, the perfusate included 20 mM HEPES, 100  $\mu$ M bicuculline, 50  $\mu$ M (1,2,5,6-tetrahydropyridin-4-yl)methylphosphinic acid (TPMPA), or 2  $\mu$ M strychnine. TPMPA was obtained from Tocris (Minneapolis, MN, USA); all other drugs were from Sigma-Aldrich (St. Louis, MO, USA).

**Stimulus generation.** In a typical experiment we displayed four types of stimuli: spatio-temporal white noise, exponentially correlated noise, multiscale correlated noise, and full-field flicker. For all stimuli except for multiscale, we began by choosing intensity values from independent Gaussian distributions for each check. In the case of spatio-temporal white, exponentially correlated, or multiscale noise, our stimulus

had 64 x 64 checks, each of which had an area of 42  $\mu\text{m}$  x 55  $\mu\text{m}$ . Full-field flicker, of course, had only one check that covered an area of 2.67 mm x 3.5 mm, well beyond the extent of the electrode array. To generate exponentially correlated noise, we filtered the Gaussian noise with an exponential filter with a space constant of one check. Note that here, unlike in [Chapter 2](#), the exponentially correlated stimulus does not have temporal correlations. The multiscale stimulus was produced by adding Gaussian noise at several nested scales, as described in [Chapter 2](#). All four stimuli were again thresholded at their mean intensity to yield binary stimuli. We generally presented 10 minutes of each stimulus in each condition (control and drug).

In some experiments, we presented repeated stimuli instead of unique sequences. In these cases, we generated 30–60 seconds each of white noise checkerboard and full-field flicker and repeated each segment 60 times. To control for any drifts in response, we alternated blocks of 10 repeats of each stimulus.

Before and after each trial, we presented a square wave stimulus for 2 minutes to test for tissue health and light responses. This stimulus was presented at 0.5 Hz (1 second ON, 1 second OFF) and used the maximum and minimum intensities available.

**Analysis.** [Chapter 2](#) describes the method for measuring pairwise spike train correlations and estimating receptive fields. Here we also computed information rates and signal-to-noise ratios. For each stimulus, we computed the information contained in a single spike following [Fairhall et al. \(2006\)](#). By averaging across repeats, we found the firing rate  $r(t)$  of each cell as a function of time  $t$  measured in 5-ms bins. The single-spike information is then

$$I_{\text{spike}} = \left\langle \frac{r(t)}{\bar{r}} \log \left( \frac{r(t)}{\bar{r}} \right) \right\rangle_t, \quad (3.1)$$

where  $\bar{r}$  is the overall mean firing rate of the cell and the average is taken across time bins.

The signal-to-noise ratio (SNR) quantifies the strength of the stimulus-dependent response relative to the trial-to-trial variability in response. We define  $s(t)$  to be the response at time  $t$  averaged across repeats and  $n_i(t)$  to be the residual response on trial  $i$ , both computed in 33-ms bins. Their Fourier transforms at frequency  $k$  are denoted  $S(k)$  and  $N_i(k)$ , respectively. The signal-to-noise ratio is then

$$SNR = \int \frac{S(k)S^*(k)}{\langle N_i(k)N_i^*(k) \rangle_i} dk, \quad (3.2)$$

where  $(\cdot)^*$  denotes the complex conjugate. The integral is taken over frequencies up to the Nyquist frequency of the firing rate, which for a bin size of 33 ms was 15 Hz.

# Chapter 4

## General discussion

### 4.1 Summary

In the preceding chapters I presented a series of results relating to the processing of correlated scenes. In [Chapter 2](#), I showed first that, when presented with stimuli that differ in spatio-temporal correlation structure, the retina shows reduced changes in output correlation relative to changes in input correlation. I then demonstrated that the receptive fields of ganglion cells are more high-pass and show a skew toward stronger surrounds when the stimulus is more correlated. Moreover, these receptive field changes, along with adaptation in the gain, are sufficient to explain the observed pattern of output correlations. In [Chapter 3](#), I showed that horizontal cell feedback and, to a lesser extent, feedforward inhibition from amacrine cells *increase* the stimulus dependence of output correlations and of the signal-to-noise ratios of individual ganglion cells. Here I discuss these results in a broader context and suggest a number of future research directions.



## 4.2 Studying natural vision

Historically, much of our knowledge of retinal processing has come from experiments in which a stimulus such as a spot of light (Kuffler, 1953) or drifting grating (Enroth-Cugell and Robson, 1966) is presented and action potentials in the optic nerve are recorded. More recently, white noise analysis (Chichilnisky, 2001) has become popular for mapping receptive fields. Such studies allow us to systematically examine the sensitivity of RGCs to a wide range of stimuli. However, single spots of light and drifting gratings do not occur frequently outside the experimental setting. Rather, natural scenes contain a large number of objects and textures with a distinct  $1/f^2$  power spectrum (Field, 1987). The ability of the retina to adapt to its visual environment leads to the speculation that its response properties during natural vision may be different from those observed in many experiments. In this thesis, we have investigated one aspect of this question: Does the retina adapt to stimulus correlations? In particular, does the retina become increasingly decorrelating as stimulus correlations increase? As we have seen, the answer is that ganglion cells do adapt in a way that partially mitigates the redundancy introduced by highly correlated stimuli. Further exploration of questions such as this will be crucial to fully understanding how we perceive the world.

Here we have focused on the first few steps in visual processing. There are a number of reasons why the retina is ideal for studying population coding. At a conceptual level, the inputs and outputs of the system are well-defined. Input to the retina consists of photons of visible light, and there is very little feedback from the rest of the brain (but see Reperant et al., 2006, for a review of rare exceptions). Thus, we know to a close approximation what information the retina receives. In turn, all communication from the retina to the rest of the nervous system goes through

the optic nerve, which is made up of ganglion cell axons. Thus, if we recorded the responses of all RGCs in a retina we could be confident that we had measured all output information. This is in contrast to areas such as cortex, where bidirectional connections between regions make it difficult to determine the inputs and outputs of each area.

On a practical level, the retina is a thin sheet, making it unnecessary to slice the tissue prior to recording. While some cuts must be made to overcome the curvature of the eye, we can make these cuts far from the area of recording so that only the most distant of connections are disrupted. In addition, the main output cells are on the surface and thus relatively close to the (planar) electrodes in a multi-electrode array. Moreover, the fact that retina is inherently photosensitive means that we can provide a good approximation of its natural input. (Of course, to have a truly naturalistic input we would need to filter the stimuli to mimic the blur that occurs when light passes through the eye before reaching the retina, but such a manipulation could easily be included in experimental design.) These properties make population studies in the retina particularly exciting because any changes in output must be due either to the stimulus, which we control, or to changes within the system itself. Thus, while many details of retinal circuitry remain a mystery, questions of network behavior are tractable.

Studies of retinal population coding not only contribute to our understanding of the visual system but also enrich our knowledge of neural circuits in general. While the retina is a specialized tissue with unique cell types, many of the synaptic mechanisms are found elsewhere in the brain. Thus, understanding circuitry and population coding in the retina can give us candidates for behavior in other circuits, leading to improved computational models and more targeted experiments.

The experiments reported in this thesis used more naturalistic stimuli than many

previous studies, but there is still a degree of artificiality in that we record from isolated retina. The ideal case would be to record from the retinæ of awake, behaving animals as they viewed their natural environments. Such a project would be technically difficult but is becoming more of a reality with the development of advanced recording techniques. For example, adaptive optics can be used to measure calcium-based light responses of ganglion cells in mice injected with fluorescent proteins (Yin et al., 2013). This and other techniques will hopefully be used in the future to study vision in increasingly natural environments.

### 4.3 Adaptation to complex stimulus features

The ability of retinal ganglion cells to adapt to simple stimulus feature such as luminance and contrast is well documented (e.g., Shapley and Enroth-Cugell, 1984; Smirnakis et al., 1997). We also know that changes in stimulus correlations lead to altered receptive fields in lateral geniculate nucleus (LGN; Lesica et al., 2007) and primary visual cortex (David et al., 2004; Felsen et al., 2005). On the other hand, adaptation of the retinal population as a whole to stimulus correlations has not been studied as extensively. Pitkow and Meister (2012) found that the retina decorrelates pink ( $1/f$ ) noise and adds correlations to white noise. The results reported here replicate and strengthen this finding by showing that correlations in an individual pair of neurons change between stimuli. On the surface, our finding that linear properties of ganglion cell responses could explain such adaptation are in conflict with Pitkow and Meister (2012)’s conclusion that nonlinear processing is needed. However, they only tested this claim with receptive fields measured during white noise presentation, whereas we measured receptive fields under both white and correlated stimuli to account for any stimulus-dependent changes in linear processing. Thus, while the

output of a fixed linear filter for stimuli of varying correlation levels is not adequate to account for correlation invariance, the output of an adapting filter is sufficient.

The similarity of adaptation to correlation to contrast adaptation suggests that the two forms of adaptation may be related. Adaptation to contrast can produce receptive fields with shorter latencies and lower sensitivity ([Baccus and Meister, 2002](#)), as we report here for adaptation to correlations. It is plausible, for instance, that naturalistically correlated stimuli have higher effective contrast than uncorrelated stimuli, even if the two stimuli have similar contrast when they enter the eye. In LGN, changes in effective contrast due to receptive field adaptation can explain the higher reliability of responses to correlated stimuli ([Lesica et al., 2007](#)); a similar phenomenon could occur in retina. Whether such similarities are the result of a shared mechanism is currently unknown. In this thesis we have presented evidence of a large role for horizontal cells in adaptation to stimulus correlations. While horizontal cell input to bipolar cells does not appear to contribute to contrast adaptation ([Rieke, 2001](#)), horizontal cells can produce gain control in photoreceptors and could thus play a role in adaptation to contrast and correlation ([VanLeeuwen et al., 2009](#)). Amacrine cells are generally thought to play a minor role in contrast adaptation ([Brown and Masland, 2001](#)), although they may contribute in some cases ([Beaudoin et al., 2008](#)). Similarly, our results are consistent with a role for direct inhibition of ganglion cells, but any such effect is weak at best. Thus, while there are some suggestions that correlation and contrast may produce adaptation via shared mechanisms, further work is clearly needed. Efforts to measure changes in pairwise correlations when contrast changes and to compare any such effects with those that follow changes in stimulus correlations are currently ongoing.

## 4.4 Efficient coding of correlated stimuli

According to theories of efficient coding, the visual system ought to remove redundancy present in the input whenever doing so would not reduce the amount of information conveyed (Barlow, 1961). Here we have proposed and tested a corollary of this theory: retina adapts to reduce the effect of input correlations on output correlations. In this view, it may be adaptive at time to increase correlations relative to the stimulus. In particular, population responses to a flickering checkerboard show correlations at distances greater than the size of a single check (see Figure 2.4C). Pitkow and Meister (2012) similarly found that retinal output can show greater correlations than those present in a white noise stimulus. If the goal is to minimize changes in correlations, rather than the correlations themselves, then it makes sense that responses to an uncorrelated stimulus might be more correlated than the stimulus.

The presence of some correlations in retinal output is expected because redundancy reduction is not the sole purpose of the visual system. If ganglion cells were fully independent, it would be impossible for the thalamus and cortex to distinguish between variation in a neuron’s response due to a change in stimulus and variation due to noise. Clearly, then, some redundancy is needed to combat the effects of noise. What is less obvious is how the population of RGCs should respond to a given stimulus in order to obtain the optimal balance of noise reduction and redundancy reduction. In a simulation of two coupled neurons, Tkačik et al. (2010) found that the optimal strategy is to average the signal (i.e. enhance correlations) when neural noise is high and reduce correlations when noise is low. Moreover, the optimal magnitude of averaging or decorrelation increases for a more correlated stimulus, with the stimulus-dependence increasing at high noise levels. It would be interesting to see whether manipulating the amount of noise (for example, by changing the background

light level) affects the output correlations in retinal responses.

## 4.5 Functional classification of ganglion cells

It would be highly informative to characterize stimulus-dependent correlations within and between cell types. Indeed, [Mastronarde \(1983a,b,c\)](#) found that correlations between cat Y cells are stronger and more likely to be due to reciprocal connections than correlations between X cells. In addition, the strength of correlation between and within types of ganglion cells in monkey depends not only on the polarity but also on the cell types of the pair in questions ([Greschner et al., 2011](#)).

Ganglion cells can be classified into types based on morphology, physiology, or a combination of the two. To determine morphological cell types, we would need to image our cells at high resolution. Typical setups for multi-electrode array recordings use microscopes with no more than 10x magnification, which is adequate to focus the stimulus on the same plane as the retina but not to see individual cells. For the latter, we would most likely want to be able to switch between a low-magnification lens for stimulus presentation and a high-magnification immersion lens for imaging. Such a procedure is possible in theory (and is already done in many single-electrode experiments on retina) but would require significant modification of our experimental setup.

A more feasible pathway would be to identify cell types based on their response properties. Such physiological catalogs have been developed in mouse ([Farrow and Masland, 2011](#)), rabbit ([DeVries and Baylor, 1997](#)), and salamander ([Segev et al., 2006](#)), but a careful functional classification of guinea pig RGCs has not been done. Responses to stimuli such as checkerboards and drifting gratings are typically used to cluster cells into types. We performed a rudimentary classification of cells by type

in [Chapter 2](#), but our precision was hampered by the relatively small size of our array. To differentiate cell types with similar functional properties, we would want to verify that each type tiles the retina ([DeVries and Baylor, 1997](#)). With an area of only  $120\text{ }\mu\text{m} \times 150\text{ }\mu\text{m}$  per array, we cannot achieve adequate sampling to make this determination. With a larger array (for example, one with 252 electrodes), we could sample from a larger area and thus examine coverage factors to verify our cell typing. To obtain accurate results, the ideal process would be to develop such a classification system in an independent series of experiments and then to use a shortened set of classification stimuli as part of each of our experiments.

Classifying cells by type may enable us to better understand how the retina adapts to stimulus correlations. For example, we might think that the downstream effects of redundancy depend on whether the redundancy is within a single channel (i.e. cell type) or between channels. Indeed, we found that correlations in pairs with one OFF cell and one ON cell depend on stimulus correlations to a greater degree than pairs where both cells are of the same polarity. It would be interesting to test such within- and between-channel effects in a more thorough manner. One might also suspect that the horizontal cell contribution to correlations is common across RGC types. While there are about a dozen types of bipolar cell, most mammals have only two types of cone and two types of horizontal cell. Thus, the specificity of any processing in the outer retina may be limited. On the other hand, we might expect the amacrine cell contribution to be more specialized due to the large diversity of amacrine cells. This hypothesis would be easy to test if we had a detailed functional classification system.

## 4.6 Population models of retinal correlations

Another way to understand our data is to build a population model of the retina, fit it to data, and look at which aspects of the fit depend on stimulus correlations and which are static across stimuli. While it has become common to fit models of individual RGCs (for example, the linear-nonlinear model used in this thesis), population models of the retina are also being developed. Two approaches are Ising models and Restricted Boltzmann Machines.

The Ising model is taken from statistical mechanics. The goal is to find a model that fits the measured data (in our case, firing rates and pairwise correlations in firing rate) as well as possible without making any additional assumptions. The minimal model that can be produced from partial knowledge of the true network state is the *maximum entropy distribution* (Jaynes, 1957). In this approach, the uncertainty about the system is maximized to allow the model population to attain any of a wide range of states consistent with the data. Specifically, each neuron can be spiking or silent given a stimulus. For a model with known pairwise correlations, the closed-form solution to this problem is known. Each pair of neurons has a coupling coefficient that represents the connection strength, and each individual neuron has a bias that accounts for its spontaneous firing rate (for a more detailed description, see Tkačik et al., 2010; Stephens et al., 2011). These parameters are found using recordings such as those described in this thesis. Ising models have been useful in predicting the behavior of a population of neurons. For example, Tkačik et al. (2010) examined the effects of noise level on the optimal coding strategy for a pair of neurons and showed that correlations ought to be enhanced at high noise levels and reduced at low noise levels. Ising models have also been fit to recordings from populations of up to ten retinal ganglion cells (Ann Hermundstad, personal communication), and efforts to



determine how the parameters in these models depend on the stimulus presented are currently ongoing.

A Restricted Boltzmann Machine is a form of neural network that can learn representations. It is generally constructed with a layer of visible units (RGCs in our case) and one or more layers of hidden units (processing within the retina) with connections between units in different layers but not between units within the same layer (Hinton, 2007). The motivation for using such a model is that it contains no direct interactions between ganglion cells. Thus, correlations are not explicitly fit; any patterns between neurons must arise from fits to single neurons' firing rates. Despite this, when such a network is trained on firing rates of individual neurons from our recordings, it can capture the pairwise correlations in the data (David Schwab, personal communication).

Efforts are currently underway to fit both these models to recordings of retinal ganglion cells responding to stimuli with different correlation structures. Such projects will hopefully give us an understanding of how properties of individual cells and their connections can lead to the observed changes in network behavior. In addition, any interesting behavior in the models should lead to predictions that we can then test experimentally.

## 4.7 Concluding remarks

Natural visual stimuli are highly correlated. It has long been speculated that an optimal sensory system would remove these correlations to reduce redundancy. On the other hand, some degree of redundancy is needed to combat noise. Here I presented evidence that the retina balances these two needs by reducing differences in stimulus correlation to maintain a relatively stable degree of redundancy. I then suggested a

circuit mechanism wherein correlations are enhanced, especially in the outer retina, to preserve signal strength. Future studies should further explore the contributions of horizontal and amacrine cells, particularly with regard to correlations within and between cell types. It will also be interesting to determine how adaptation to stimulus correlations is balanced with other goals of visual processing and how downstream areas of the visual pathway interpret correlations in retinal output.

# Appendix A

## Identifying spikes in multielectrode array data

The material in this Appendix was published as [Prentice et al. \(2011\)](#).

### A.1 Abstract

We present an algorithm to identify individual neural spikes observed on high-density multi-electrode arrays (MEAs). Our method can distinguish large numbers of distinct neural units, even when spikes overlap, and accounts for intrinsic variability of spikes from each unit. As MEAs grow larger, it is important to find spike-identification methods that are *scalable*, that is, the computational cost of spike fitting should scale well with the number of units observed. Our algorithm accomplishes this goal, and is fast, because it exploits the spatial locality of each unit and the basic biophysics of extracellular signal propagation. Human interaction plays a key role in our method; but effort is minimized and streamlined via a graphical interface. We illustrate our method on data from guinea pig retinal ganglion cells and document its performance

on simulated data consisting of spikes added to experimentally measured background noise. We present several tests demonstrating that the algorithm is highly accurate: it exhibits low error rates on fits to synthetic data, low refractory violation rates, good receptive field coverage, and consistency across users.

## A.2 Introduction

The vertebrate retina is an important model system in neuroscience because it is amenable to detailed study despite having a complex structural and functional architecture (Sterling, 2004). Population coding and collective behavior in the retinal output is studied by use of multi-electrode arrays (MEAs) to record extracellularly from many retinal ganglion cells (RGCs) simultaneously (Meister et al., 1994; DeVries and Baylor, 1997). Similar recordings can now also be made in other brain areas (Buzsáki, 2004). MEAs offer unprecedented possibilities to obtain both single neuron and single action potential resolution from large tissue samples. However, recordings obtained in this way are useful only if most spikes can be assigned, with sufficient accuracy, to the neurons that generated them. Even if each neuron spikes with a unique waveform signature, we must still determine all those “template” waveforms present in a dataset, separating them from each other and from noise. Moreover, in practice there can be wide variation in the spike waveforms from a given unit (for instance in amplitude), complicating the task of determining from data which units fired and when.

This “spike sorting problem” is therefore a bottleneck in the use of high density arrays with hundreds or thousands of electrodes. Methods that were manageable with tetrodes (Gray et al., 1995) do not generally scale up to the massive datasets that large arrays generate. For example, some standard methods cluster data by manually

examining two-dimensional projections in a feature space of a few tens of dimensions. This approach is infeasible when the feature space contains thousands of dimensions.

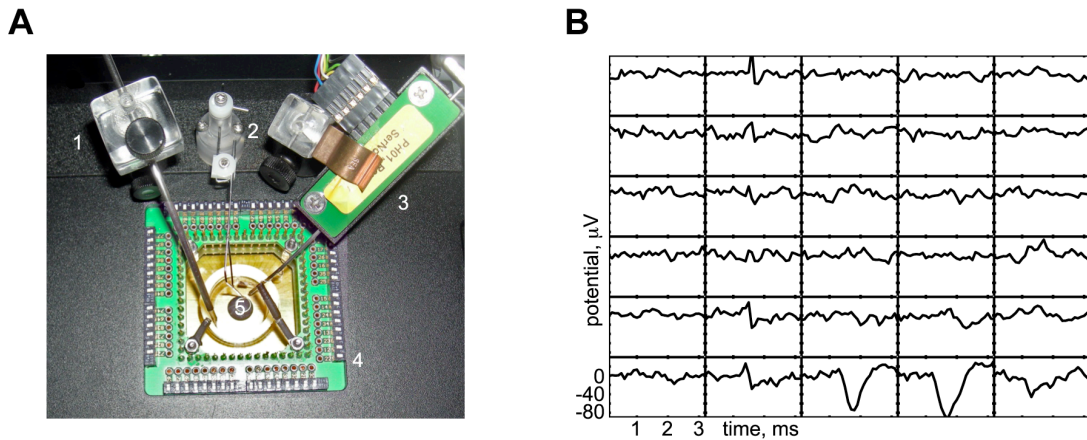
Another challenge with large arrays is that the chance of seeing a single isolated spike becomes negligible, simply because there is so much activity. Thus we must find template waveforms corresponding to the activity of single neural units (as extracellular recording cannot unambiguously identify distinct single neurons, we will refer to our recovered putative neurons as “neural units”) without ever seeing a pristine example of one, and we must be prepared to decompose temporally overlapping spikes in essentially every recorded event. Overlaps in both space and time are less frequent, but they nevertheless must be resolved if we wish to unravel the patterns of collective neural activity. Resolution methods that rely on exhaustively checking all possible combinations suffer a combinatorial explosion for large arrays. Further, any spike decomposition method must stop before every spike has been found, because there will be some units whose intrinsic amplitude is not larger than recording noise. We need a principled approach to terminating each fit and to deciding later which units’ activities have been reliably captured.

Thus, to be most useful for large arrays, a spike identification algorithm should ideally both *scale well* and also *be able to decompose overlapping events*. This appendix outlines a method that accomplishes these goals (MATLAB code available upon request.) It first clusters a small subset of a larger dataset, using a partially automated, human-guided clustering technique. This manual step is efficient, and scalable, because (i) the clustering is based on an ordering that arranges event data by similarity along a single dimension, (ii) the ordered data display band-like features that are visually very salient for human operators, making cluster cutting unambiguous, and (iii) the algorithm is robust to variations and outliers in the cluster-cutting procedure. The algorithm then fits the full dataset to the spike templates thus obtained, using a

modified Bayesian approach. In our data (from guinea pig retina) most of the intrinsic variability of spikes from a given unit consists of amplitude variation only, whereas other variability can be summarized as a universal (spike-independent) noise process. By carefully modeling these circumstances we greatly reduce our computational burden.

After characterizing the spatiotemporal character of the noise, our algorithm identifies spikes iteratively in a matching-pursuit (or “greedy”) scheme (Mallat and Zhang, 1993; Segev et al., 2004). Fitting terminates when addition of another spike does not improve the posterior probability of a fit; a natural overfitting penalty is provided by prior probabilities of firing and of spike amplitude. No assumptions are made about spike time cross-correlations; in particular, we do not require *a priori* any refractory “hole” in the spike time autocorrelation functions. Nevertheless, all of the inferred spike trains corresponding to otherwise acceptable spike types do exhibit such a hole, which serves as a check on our results. Fitting is followed by several qualitative post-hoc validations of the templates. Where our method requires human judgment, we structure our techniques and develop tools to facilitate interaction and proofreading.

Our approach combines successful elements from previous techniques: the empirical characterization of the noise (Pouzat et al., 2002); separation of clustering and fitting steps and the iterative subtraction scheme for handling overlaps (Segev et al., 2004); and division of the clustering task by leader electrode address (Litke et al., 2004). Novel features of our approach include systematic exploitation of the spatial organization of the signals, the use of an ordering algorithm to greatly simplify clustering, the observation that the noise temporal correlation is well represented by a simple function, an explicit model of spike amplitude variation, and the introduction of a principled Bayesian likelihood criterion for terminating spike fitting. Each of these innovations adds a critical element to the success of our spike sorting method.



**Figure A.1: Recording chamber and typical data.** (A) Typical MEA apparatus. A tissue sample was mounted in an inverted microscope, with images projected onto it via a small video monitor at the camera port (not visible). *Clockwise from left*, 1: suction; 2: tissue hold-down ring; 3: perfusion inflow, with temperature control; 4: preamplifier; 5: location of the multi-electrode array. (B) Example of a single-spike event. Each subpanel shows the time course of electrical potential ( $\mu\text{V}$ ) on a particular electrode in the  $5 \times 6$  array. The electrodes are separated by  $30 \mu\text{m}$  (similar to RGC spacing). A spike from one unit is visible in the lower right corner and an axonal spike can be seen running vertically in the second column of electrodes. Data were acquired at 10 kHz. After baseline subtraction and high-pass filtering, a spatial whitening filter was applied (see [Methods](#), Step 1).

Although we focused on data taken from vertebrate retina, the methods should be equally applicable to other kinds of MEA data, such as those collected from different brain areas ([Buzsáki, 2004](#)).

### A.3 Results

To illustrate our method, we tested our spike sorting algorithm on 120 minutes of recordings from guinea pig retinal ganglion cells (RGC), acquired with a 30-electrode, dense MEA covering about  $0.018 \text{ mm}^2$  of tissue (Fig. [A.1A](#)). The analysis described

in this appendix identified 1,260,475 spikes in the dataset. A typical firing event took the form illustrated in Figure A.1B, where each panel shows 3 ms of the electrical potential recorded by each electrode (or “channel”). We identified spiking events as voltages crossing a threshold of  $-40\text{ }\mu\text{V}$ , taking into account the fact that simultaneous threshold crossings on neighboring channels represent the same spike event (see Methods for details). The duration of each spike event was taken to be 3.2 ms centered on the event’s peak.

In addition to identifiable spikes, each electrode had background activity with a standard deviation of  $\sim 10\text{ }\mu\text{V}$  that we will collectively refer to as “noise.” Potential sources for this activity include true (Johnson) noise in the electrode and electronics, electrical pickup from the environment, and a hash of background activity from weak or distant neurons (Fee et al., 1996). A challenge for spike identification is that in general there is no way to separate these three classes of “noise” cleanly from each other, nor from the spikes of interest to us. Nevertheless, we will propose a technique for identifying spikes that is very accurate for firing events with intrinsic amplitude at least 4 times the standard deviation of the noise.

Figure A.1B illustrates that each single firing unit will be “heard” on multiple electrodes, and that those electrodes form a spatially localized group. Our method is scalable because it systematically exploits this simple observation: even on a large electrode array, most firing units will involve only a handful of electrodes. (Some of our signals were *not* local, and stretched over the entire electrode array in a line (e.g., Fig. A.1B). We ignored such axonal signals, which were also distinguished by their low amplitude and triphasic shape.)



### A.3.1 Preliminary visualization of our data

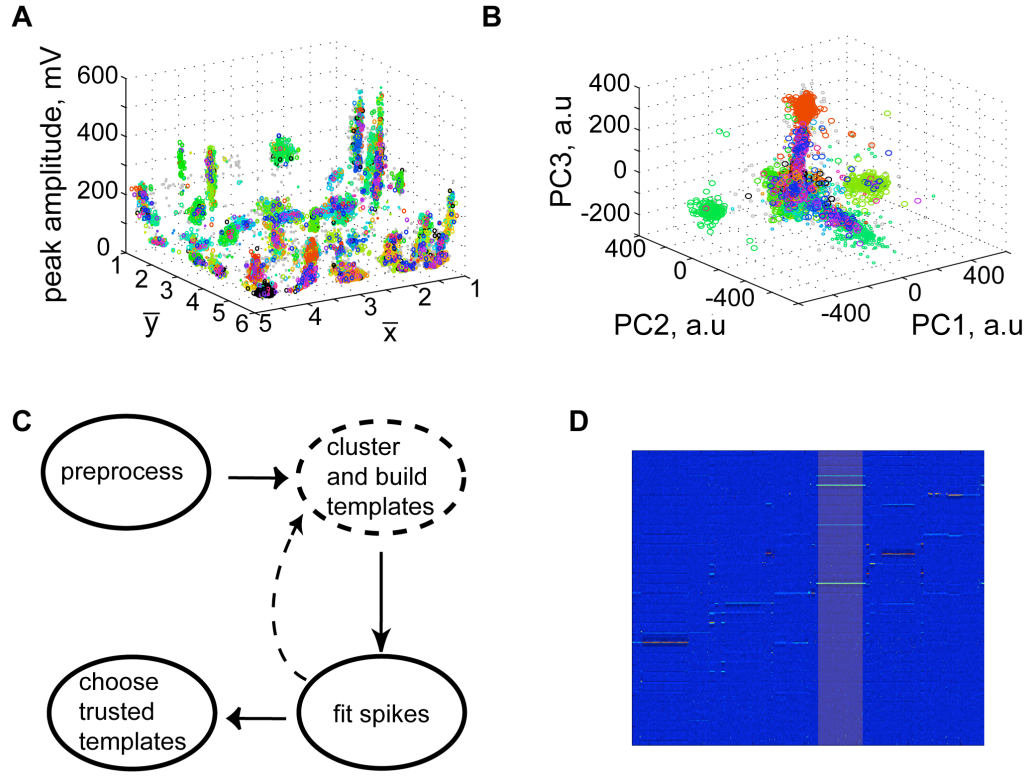
We first attempted a “geographical clustering”: from each event we found the minimum of the potential on each channel and the channel containing the deepest minimum (“leader channel”). We then used the absolute values of the minima as weights in a weighted average of the locations of the 9 electrodes neighboring (and including) the leader channel. This weighted average gave a particularly salient two-dimensional feature, the event’s *barycenter*:

$$\bar{\mathbf{x}} = \frac{\sum_{i=1}^9 \mathbf{x}_i |\min V_i(t)|}{\sum_{i=1}^9 |\min V_i(t)|}.$$

The sums run over the 9 electrodes neighboring the leader, and  $\mathbf{x}_i$  is the location of the  $i^{th}$  neighbor electrode. We then augmented  $\bar{\mathbf{x}}$  with a third feature,  $z$ , equal to minus the peak potential (because peaks were generally negative,  $z$  was almost always the absolute value of the peak). These three features gave a scatterplot that clearly showed many well-separated clusters (Fig. A.2A), without any need to deduce the “best” features by principal component analysis (Fig. A.2B).

This extension of the “triangulation” method developed for tetrode recordings (Gray et al., 1995) already shows key aspects of the data: (a) many clusters are highly dispersed in amplitude, and (b) some cluster pairs appear at nearly the same spatial locations but are nevertheless well separated by amplitude. The first circumstance means that we must allow for variable amplitude when fitting spikes to templates representing the clusters. The second warns us that a simple least-squares fit to amplitude could confound two distinct units. For this reason our spike-fitting method employs a Bayesian prior for each cluster’s amplitude variation, allowing us to make such discriminations.

Although the simple clustering based on spatial location in Figure A.2A looks



**Figure A.2: Visualization of data and flowchart of our method.** (A) 22,234 firing events cluster well in terms of their *barycenter* (voltage weighted average spatial location) and absolute peak voltage (see text), despite wide amplitude dispersion in some groups; each combination of color and marker size corresponds to one spiking unit identified by the clustering method developed in the text. *Gray points* were not assigned to any cluster. A total of 107 clusters are marked. (B) Events cluster poorly when projected onto the three principal features uncovered by principal component analysis (PCA). Coloring as in A. (C) Schematic of our spike sorting method. *Dashed lines* involve a small subset of a full dataset. The *backwards arrow* describes the introduction of new spike templates found after the first pass of fitting; a total of two passes are performed. (D) The OPTICS algorithm orders all firing events into a linear sequence based on a distance measure (see text). Events are lined up in this order (x-axis), and represented in terms of the 960 voltage samples recorded by all the electrodes during a 3.2 ms firing event (y-axis; from top to bottom, 32 consecutive time samples from one channel, then 32 time samples from a second channel, and so on). The human operator highlights bands of events (typically very clear to an observer) that appear to constitute a single cluster; one such band is shown. Later automated diagnostics refine and check these assessments.

promising, it can be misleading. Indeed, the restriction of the weighted average to the 9 electrodes around the leader can artificially separate clusters by biasing the barycenter to be located near a particular electrode. This problem could be alleviated by using a larger neighborhood, but on large arrays there will inevitably be temporal collisions of spikes from distinct units. The barycentric features in Figure A.2A will register such collisions as a haze of seemingly random spots. Thus, at a minimum the MEA voltage traces must be segmented by exploiting the spatial locality of recorded responses. Despite these shortcomings, Figure A.2A points out why the more sophisticated method developed in this appendix can succeed: the “geographic” information encoded by the MEA is a powerful intrinsic clue to each unit’s identity.

### A.3.2 Summary of our method

Our sorting method is outlined in Figure A.2C (details in [Methods](#)). From a subset of the raw data, we made a preliminary classification of spike events in terms of the electrode on which they achieved their peak voltage. All events sharing a given leader channel were cropped to the 9 electrodes neighboring the leader (i.e. the leader and the 8 bordering electrodes), then ordered with the OPTICS algorithm ([Ankerst et al., 1999](#)) into a linear sequence. The OPTICS algorithm places similar events nearby in the sequence; distance was measured by a normalized Euclidean distance between event voltage traces (see [Methods](#)). The linear sequence of events was displayed to the user, along with all the recorded voltage samples for each event (Fig. A.2D), and manually clustered. Although the ordering was based on events cropped to 9 channels, the full waveforms were displayed to the user (Fig. A.2D). Because the data are ordered in one dimension, and because precision is not required in view of later refinement, this manual step remains rapid. An automated method for cluster cutting

could be implemented, but in view of the inevitable desire for human proofreading we preferred to simply carry out this step by hand. From each preliminary cluster, we estimated a template waveform representing the corresponding neural unit and then fit the templates to the remaining data. Fitting was accomplished by a Bayesian algorithm based on a probabilistic model capturing the dominant sources of variability we observed in our data: background noise, spike amplitude variation, and overlapping spikes from distinct units. After finding, for each event, the most probable template which accounts for the event, we subtracted it and then iterated. Finally, the fit results were used in a post-hoc validation of the initial clustering, and we repeated the procedure in a second pass if necessary. Details of each step are presented in [Methods](#).

### A.3.3 Tests of our method

OPTICS-based clustering of a subset of our dataset led to 107 potential templates for events from distinct neural units. Many of these templates had low amplitudes; such low-amplitude templates were sometimes mistakenly fit to noise by our algorithm. We therefore rejected units that were likely to contain substantial noise fits because they were of amplitude less than or comparable to the noise (details in [Methods](#)). This left fifty potentially reliable units that were accepted in our dataset.

**Comparison with geographical clusters.** Our OPTICS-based procedure identified 107 potential clusters of events in a subset of the data. To check that the procedure gave reasonable results we plotted each event in the barycentric coordinates of Figure [A.2A](#) and colored the events according to the assigned cluster label. The clusters were spatially localized and separated in peak amplitude, as they should be if they were produced by distinct single neurons. Gray dots in Figure [A.2A](#) were not

assigned to any cluster. Some of these events contained overlaps of spikes that were not resolved by the initial spatial segmentation of data during the preprocessing step. The subsequent spike fitting step in our algorithm resolved most such ambiguities.

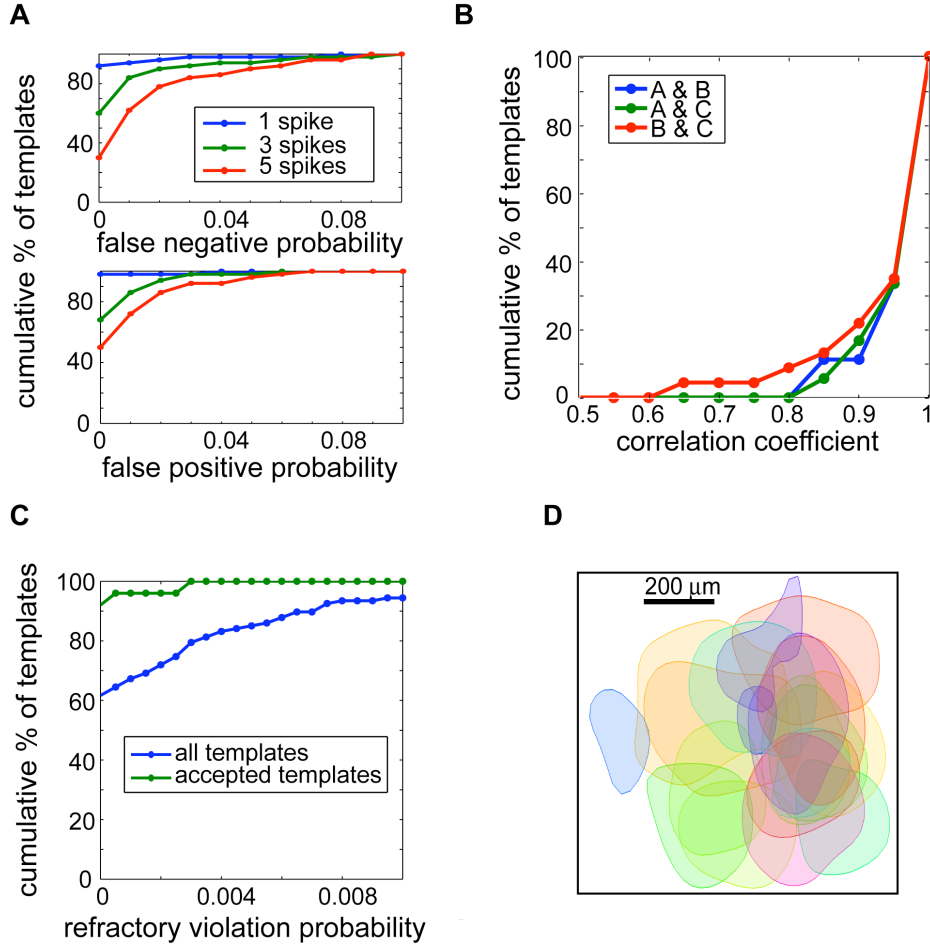
**Error rates on synthetic data.** To validate our algorithm we tested its performance on synthetic data created by adding spikes to experimentally measured background noise clips, then fitting templates to each clip. We took noise clips to be 3.2 ms segments of time during which no spikes were recorded on any channel; we identified 14,000 such clips. For each clip, we randomly chose a fixed number (1, 3, or 5) of templates from the initial set of 107, with uniform probability and without replacement. We then added these templates to the noise clip at random times, leaving a margin of 0.6 ms on either side of the clip to prevent waveforms from being cut off. (Our typical template waveforms extended approximately 0.5 ms to either side of the peak.) We gave each spike an amplitude drawn from a Gaussian distribution with mean equal to its template amplitude and standard deviation 10% of the template amplitude (similar to the observed distribution).

The template fitting algorithm was then run over this synthetic dataset and analyzed for false positive and false negative rates (Fig. A.3A). We counted a false negative for a template every time that template was present in an event but not fit correctly to within 1 ms; we counted a false positive every time a template was fit to the data without actually being present. The error rates increased with the number of template overlaps; thus, for the fifty templates with amplitudes that exceed the noise, we separately plotted error rate histograms for each degree of overlap. Error rates were robustly low — even within extremely complex events with 5 overlapping spikes (rare in the data), the majority of spike templates had an error rate of a few percent or less. To gain perspective on these values, we measured the number of templates

fit to each event in our recorded data: 60% of events contained 1 spike, 94% had 3 or fewer spikes, and 98% had 5 spikes or fewer. Most of the errors were made on lower amplitude templates for which amplitude variations can lead to confusion with noise.

**Refractory violations.** When sorting spikes recorded extracellularly, ground truth can be assessed if simultaneous intracellular recordings are available (e.g. [Harris et al., 2001](#)). Since we do not have such recordings, in order to validate our algorithm on real data we examined the rate of refractory violations — i.e., the fraction of interspike intervals of duration less than 1.5 ms. Refractory violations can appear in our sorted data if spikes from distinct neural units are mis-assigned to the same unit or if noise fluctuations are mistaken for spikes. Of the 107 templates constructed from the initial clustering, 84% had less than 0.5% refractory violations and all had less than 2.5%, providing evidence that the templates produced by the initial clustering rarely merge distinct neural units (Fig. [A.3C](#)). More significantly, all fifty templates describing units that rose reliably over the noise level had less than 0.5% refractory violations. Furthermore, 96% of these had less than 0.1% refractory violations (Fig. [A.3C](#)). Note that the percentage of refractory violations only provides a lower bound on the number of misidentified spikes — nevertheless, the low refractory violation rates we observed provide strong evidence that our algorithm makes few fitting mistakes on the units otherwise identified as reliable.

**Coverage.** While the absence of refractory violations gives evidence that our algorithm does not merge different neural units together, it might still split spikes from the same unit into two distinct clusters if, e.g. there was substantial amplitude variation. To test for this, for each unit that was above the noise level we measured the linear receptive field by taking the spike triggered average (STA) of the flickering



**Figure A.3: Tests of the method.** (A) (Top) The cumulative percentage of templates having false negative probabilities less than the indicated values. Error rates were measured in fits to synthetic data as the fraction of times a fit was *not* reported for a template when it was actually present. (Bottom) As above, but showing false positive probabilities (fraction of times a fit was reported for a template when it was *not* actually present). Results reported separately for fits to events with different numbers of overlapping spikes (*inset colors*). (B) Correlation in spike trains across fits by three different users (A, B, and C). Each curve corresponds to one pair of users and gives the cumulative fraction of templates having lower correlation than indicated. See main text for further details. (C) Cumulative fraction of templates having fewer refractory violations than indicated. Refractory violations are rare (see text). (D) The centers of 19 OFF cell receptive fields recorded from a single piece of tissue. To map a neuron's receptive field center, we first find the peak (in space and time) of the spike-triggered average stimulus. Restricting to the peak time, we apply cubic spline interpolation in space and then draw contour lines at 75% of the peak value.

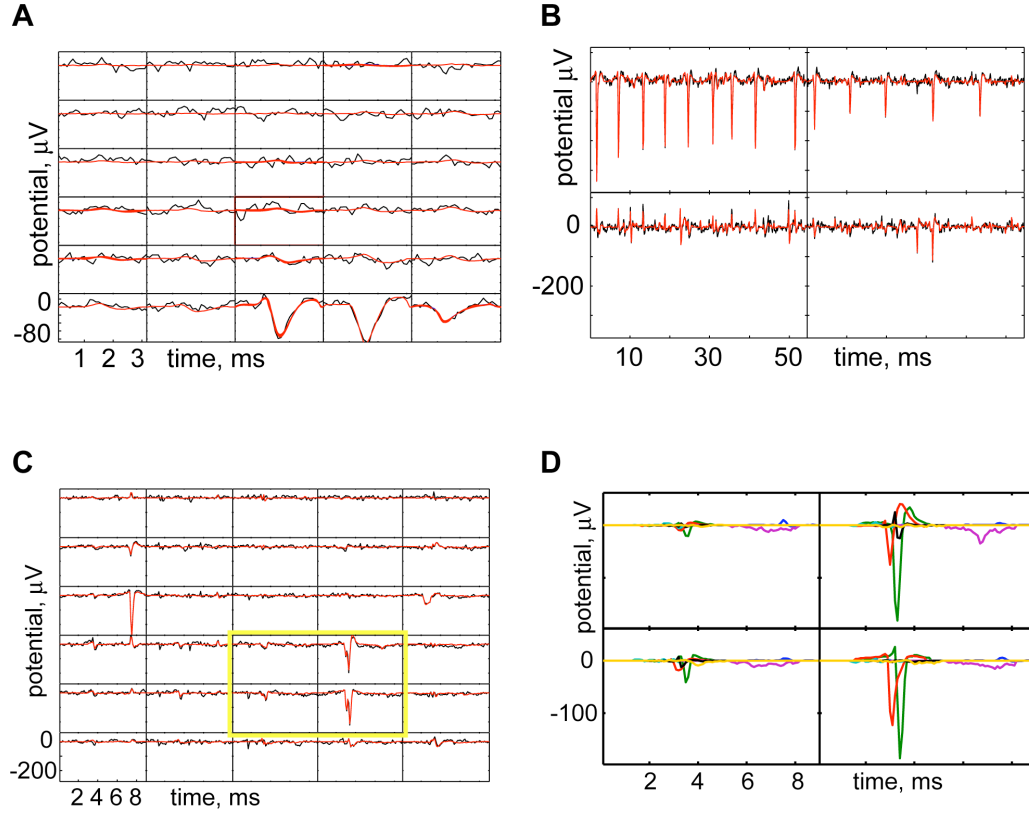
checkerboard stimulus. We expect that such receptive fields will be connected regions of the visual field, roughly elliptical in shape, and that no two units will have identical receptive fields. 31 of the 50 reliably identified units had enough spikes to give reliable estimates of the spatial receptive field. Of these, examination of the temporal kernel showed that 19 were OFF cells (responding to light decrement) and 12 were ON cells (responding to light increment), consistent with the expected excess of OFF ganglion cells (Ratliff et al., 2010; Chichilnisky and Kalmar, 2002; Peichl et al., 1987; Peichl, 1989; Dacey and Petersen, 1992; Tauchi et al., 1992). None of these receptive fields were identical, giving evidence that our algorithm did not split single units into multiple clusters. Further, all of the receptive fields were connected, suggesting that none of our clusters are mixtures of different RGC. In addition, essentially all of the recorded area was covered by at least one receptive field (coverage of OFF cells shown in Fig. A.3D). The density of RGCs in guinea pig varies from  $250 \text{ mm}^{-2}$  to  $1500 \text{ mm}^{-2}$  (Do-Nascimento et al., 1991). We record signals from a region slightly larger than the electrode array, roughly  $0.065 \text{ mm}^2$ . Thus the expected number of RGCs is 16–97, consistent with our total of 31 receptive fields, keeping in mind that many of the sluggish cell types would not have enough spikes to yield a good spike triggered average.

**Different users.** Because our method involves human intervention, there exists a risk that different users could obtain divergent results. One possible discrepancy is that different users could identify distinct sets of templates. This would not pose a serious problem in and of itself, as long as further analysis did not demand a complete population. Moreover, the completeness of the template set is ultimately limited only by the time invested by the operator in clustering. However, it would be problematic if the spike trains fit to *the same* template by different users differed significantly.



To assess the robustness of our method, we therefore had three different users (here referred to as A, B, and C) carry out our spike sorting procedure on the same data set and examined the correlation of the results. User A found 20 templates with large enough spike counts to assess cross-correlation, B found 25, and C found 28. While these numbers were smaller than the 50 we had identified previously, the difference consisted of units with extremely low firing rates. We compared the three sets of templates to identify those which were found by multiple users, by minimizing the euclidean distance between template pairs. 18 templates agreed between A and B, 18 between A and C, and 23 between B and C. A total of 17 templates were found by all three users. For each template that was found by at least two users, we computed the cross-correlation coefficient of the corresponding spike trains. Across all pairs of users, two-thirds of such templates had a spike train correlation higher than 0.95 (Fig. [A.3B](#)).

**Complex events.** A major challenge for a spike sorting algorithm is dealing with variability in spikes produced by individual neural units. An even greater challenge arises from spatio-temporal overlaps between spikes from different neural units. Our low error rate in analysis of synthetic data containing both of these complexities (Fig. [A.3](#)) provides evidence that our algorithm is effective at resolving overlaps and identifying variable spikes from given units. To test this further, we manually examined many events in the real data which a human observer could identify as representing overlaps or neural variability. Our algorithm typically did an excellent job of dealing with variable-amplitude bursts (Fig. [A.4B](#)), as well as events that overlap in space and time (Fig. [A.4C](#)).



**Figure A.4: Template fitting to complex events.** (A) Example of a single-spike event. Each subpanel shows the time course of electrical potential (in  $\mu\text{V}$ , *black curves*), on a particular electrode in the  $5 \times 6$  array. After baseline subtraction and high-pass filtering, a spatial whitening filter was applied (see [Methods](#)). *Red curves* show the result of our fitting algorithm, in this case a single template waveform representing an individual neural unit. (B) Detail of a more complex event and its fit, in which a single unit fires a burst of 9 spikes of varying amplitudes (upper left channel), while a different unit fires 5 additional spikes (upper right channel). Simultaneous data from four neighboring electrodes are shown. (C) Example of an overlap event and its fit, which now is a linear superposition of 7 templates. (D) Detail of C, showing signals on four of the electrodes. This time individual fit spikes are displayed. The *red* and *green traces* show fit templates that, although similar, differ significantly in their overall strength, and in the relative strengths of their features. The *black trace* shows a fit to a low-amplitude template that was later classified as unusable, and hence discarded, by the procedure in [Methods](#), Step 4.

**Speed.** Currently the main fitting code, written in MATLAB and run on a commercial 2.5 GHz computer, requires about 5 ms of real computer time per fit spike for each of the two passes. This is fast enough for our purposes; considerable further improvement is possible with existing software (Mex) and hardware (GPU) techniques.

## A.4 Discussion

A review of early work on spike identification, prior to the widespread use of MEAs, can be found in [Lewicki \(1998\)](#). Like some earlier work, our method separates spike identification into distinct steps of clustering and fitting. The clustering step uses all the waveform features, and makes no assumption about the cluster structure (e.g., that it is a mixture of Gaussian distributions). The fitting step acknowledges that each neural unit’s signals are subject to intrinsic, multiplicative variation as well as additive noise and uses a Bayesian approach to infer the identity of the most likely firing unit.

Our approach is intentionally not fully automatic, since human proof-reading of the results of automated clustering is generally essential. However, we have been careful to use human judgement only where it is indispensable. Further, both the human and machine steps are organized so as to scale well with array area (or number of units monitored). For example, cluster cutting was greatly simplified by representing spikes in an ordered one-dimensional array. This feature, along with systematic exploitation of the spatio-temporal locality of spikes (Fig. [A.2A](#)), and the use of a simple but powerful generative model, make our method scalable to large arrays. Furthermore, we observed that our more ambiguous templates tended to be located on the boundary of the array due to recording of units located some distance from the electrodes. These “boundary effects” should become less important for larger ar-

rays; we thus anticipate that the methods described in this appendix will yield more accurate spike sorting for large, dense arrays.

Our method can be extended in many ways. For example, it would be straightforward to update the priors continually as fitting proceeds, allowing non-stationarity and stimulus dependence to be handled more gracefully. In some applications it may be preferable to report spike identification probabilistically, rather than just listing the most-likely spike events; our formulas already provide this information. The method can also be extended to non-planar arrangements of electrodes and neural tissues, for, e.g., cortical applications. Finally, the generative model in the present appendix does not take into account correlations within and between spike trains, or the receptive field structure and stimulus dependence of responses. Performance could be further improved on complex overlapping spike events via a bootstrapping procedure. We could first use the simple independent, Poisson generative model of this appendix to produce an accurate preliminary assignment of spikes to units. From this assignment we could construct a more detailed model of correlated activity with pairwise interactions (e.g., [Schneidman et al., 2006](#); [Pillow et al., 2008](#); [Tkačik et al., 2010](#)). This more complex generative model could then be used to further refine spike assignments for applications requiring a very high degree of accuracy.

## A.5 Methods

The recordings used to develop our spike-sorting method were those reported in [Chapter 2](#) of this thesis. Here we present the method of identifying spikes in data of this type. Our procedure had four steps, each detailed below: (1) Preprocessing, where spatial locality was exploited to segment the data, (2) Clustering and template building, where a subset of the data was clustered to separate the responses of likely neural

units, template waveforms for each neural unit were built, and their variability characterized, (3) Spike fitting, where every firing event was separated into a superposition of responses from different neural units, and (4) Validation of templates, where each template and the spikes identified with it were tested for reliability.

### A.5.1 STEP 1: Data preparation and segmentation

The first step in our procedure (Fig. A.2C) was to prepare the data for clustering of events from different neural units by separating firing events from noise and segmenting spatio-temporally distinct regions of spiking activity on the electrode array.

Data from the array were sampled at 10KHz, high-pass filtered below 200 Hz with a finite impulse response filter to remove low frequency baseline fluctuations, and then packaged into 3.2 ms clips: (a) “noise clips” in which the potential never fell below  $-30 \mu\text{V}$  and (b) “spike events” surrounding moments at which the potential crossed  $-40 \mu\text{V}$ . The duration of the clip was chosen to include the full extent of most spike events with a small margin on either side. The threshold was set such that there were no discernible spikes among the subthreshold voltage deflections, although this low-amplitude noise likely contained contributions from spikes of distant neurons. Clips with potentials between  $-30 \mu\text{V}$  and  $-40 \mu\text{V}$  were neither used to characterize noise (since they might contain small spikes) nor used to identify spikes (since they were very noisy). The threshold for spikes was set to  $\sim 4$  times the standard deviation of the potential in the noise clips. Each spike event thus consisted of  $N = 3.2 \text{ ms} \times 10 \text{ kHz} \times 5 \times 6 = 960$  numbers, the potentials on a  $32 \times 5 \times 6$  grid of space-time pixels (“stixels”). Spike events sometimes overlapped each other, for example if a burst of spikes lasted longer than 3.2 ms. Cluster identification and spike template building (Step 2) used three or four 30-second segments sampled from

different times, but subsequent spike fitting and sorting (Step 3) used all the data.

Electrodes can share signals because of instrumental cross-talk and because the activity of neurons spreads passively to nearby electrodes. We modeled both effects by a linear filter that spatially blurs signals. To find this filter empirically, we noted that it also applies to the noise. Accordingly, we measured the spatial covariance of noise clips, finding that it was spatially isotropic and had a roughly exponential falloff, with a correlation length of  $\sim 30 \mu\text{m}$ . We applied the square root of the inverse of this covariance matrix to all data, and used the resulting “spatially whitened” data for all analysis. In some datasets this transformation sharpened the individual spikes spatially, improving our ability to distinguish them in the clustering stage. In other datasets the transformation had little effect. Our data also exhibited temporal correlations, but these have a different physical origin from the essentially instantaneous passive spatial spread. We found that *temporal* whitening prior to clustering (Pouzat et al., 2002) worsened our signal-to-noise ratio and impeded cluster determination. Thus we addressed temporal correlations later, during the spike fitting.

**Segmentation.** Each spike event is a superposition of spikes from an unknown number of distinct neural units with stereotyped waveforms that we sought to identify. We first spatially segmented the data to isolate waveforms from individual units and their immediate neighbors. To this end, we identified all stixels at which the potential was more negative than the threshold of  $-40 \mu\text{V}$  and divided this set into connected components (two stixels were considered connected if they were nearest neighbors in either time or space). Within each connected patch we identified the absolute peak electrode and time, then extracted a 3.2 ms region centered temporally on the peak time and cropped spatially to a neighborhood of nine channels surrounding (and including) the leader electrode. Thus each spike event was segmented into one or more

cropped events; each of which was then classified according to its leader electrode. A similar segmentation method has recently been applied to the spike identification problem by J. Schulman (unpublished); see <http://caton.googlecode.com>. In subsequent clustering, only those events having the same leader electrode were directly compared to each other (Litke et al., 2004).

Some cropped events might be composites of two spike types corresponding to neighboring, but distinct, neural units. However, this step at least decomposes composite events whose components are well separated in space or in time, and hence reduces the combinatorial burden inherent in large arrays; later steps handle composites missed at this stage. The method also ensures that, if spikes from two well-separated units frequently co-occur, the two units will nevertheless be correctly handled as separate.

### A.5.2 STEP 2: Cluster identification and template building

The second step in our procedure (Fig. A.2C) was to cluster spiking events in a subset of the data (four 30-second segments) into groups that had similar waveforms and thus probably came from the same neural unit. For each cluster, we produced a template waveform describing the typical spike and determined the distribution of amplitude rescalings that best matched spikes to this template.

**Cluster identification.** In order to group events into clusters based on the similarity of their waveforms, some previous approaches have sought a low-dimensional set of discriminable waveform “features,” and have assumed that variability between events in the same cluster arises only from additive noise. In practice, systematic variation in the shape of spikes from single units is often observed that is not well described by additive noise. Furthermore, identifying the correct set of salient waveform

features that discriminate between units is challenging (Fig. A.2B; Quiroga, 2007). Thus, seeking a technique that did not require feature extraction, we adapted the OPTICS algorithm (Ankerst et al., 1999). Briefly, OPTICS computes distances between all pairs of waveforms, then orders the waveforms such that similar ones are placed close together in a single linear sequence. OPTICS makes no assumption that clusters have a Gaussian distribution in feature space, nor does it set any threshold density in that space to trigger cluster identification. The linear ordering allows for easy visualization and cutting of clusters.

We applied this algorithm to cropped and segmented spike events (obtained in Step 1) which were upsampled by a factor of 5 (using MATLAB’s cubic spline interpolation) and then temporally aligned to place the absolute peak of the waveform at a common position before downsampling again. The interpolation was necessary to compensate for apparent variations in spike waveforms due to discrete sampling (Lewicki, 1994). To reduce the fuzziness of the clusters, we masked spike events by setting voltage samples to zero if they were less negative than  $-15\mu V$ . As a distance metric between the masked potentials  $\mathbf{V}$  and  $\mathbf{V}'$  of two spike events, we chose  $d(\mathbf{V}, \mathbf{V}') = \left( \sum_{i=1}^N [(V_i - V'_i)^2 / (k \sqrt{|V_i| + |V'_i|})] \right)^{1/2}$  where  $i$  indexes the potentials at each channel and timepoint, and  $k$  is the total number of nonzero potentials after masking of either  $\mathbf{V}$  or  $\mathbf{V}'$ . Division by  $k$  normalized for the effective dimensionality (given by the number of dimensions containing nonzero entries). We observed that higher voltage traces tended to have a higher variance; the factor  $\sqrt{|V_i| + |V'_i|}$  partially compensated for this, leading to more homogeneous clusters.

We constructed a graphical user interface (GUI) that allowed a human operator to visualize each spike event in the OPTICS sequence as a vertical column of pixels color-coded by voltage (see Fig. A.2D). Transitions between distinct spike types were usually obvious to the operator, who could quickly find and select bands corresponding



to each spike type. (For the data in this appendix, the operator found over 100 such clusters in about 30 minutes.) The software then wrote the corresponding cropped events to a set of data files. The ease of separation likely occurred because clusters could already be fairly well delineated with just the “geographical” features in Fig. [A.2A](#).

Up to this point, the events being clustered were still segregated into batches according to their leader channel  $x_0, y_0$ . Thus it was possible for a single unit to be multiply identified: If it stimulated two neighboring electrodes nearly equally, the unit could generate events in both of the corresponding batches. We tested for duplicates by manually examining pairs of clusters whose medians had a large cross-correlation and merged the clusters if necessary. There was also a possibility that the initial clustering would assign multiple units to one cluster. In these cases, visual examination of the superposed waveforms of the cluster often showed it to be a composite of multiple units. This was resolved by doing a principal components analysis on the waveforms in that cluster: if the cluster was composite, at least one of the first few principal component weights had a multimodal histogram. The cluster was split by thresholding at the valleys of the histogram; we then tested whether any of the split components ought to be merged with an existing cluster. We developed a graphical user interface to assist the operator in performing these merging and splitting steps.

Generally it was clear to the human operator when a band in the GUI output was clean enough and wide enough (contained enough events) to generate a good cluster; thus there was no need to specify *a priori* the desired number of clusters, an advantage over many automated clustering procedures. Marginally significant clusters were either eliminated during template building (see below), or else generated fits that were themselves discarded during spike fitting (Step 3) and evaluation of template

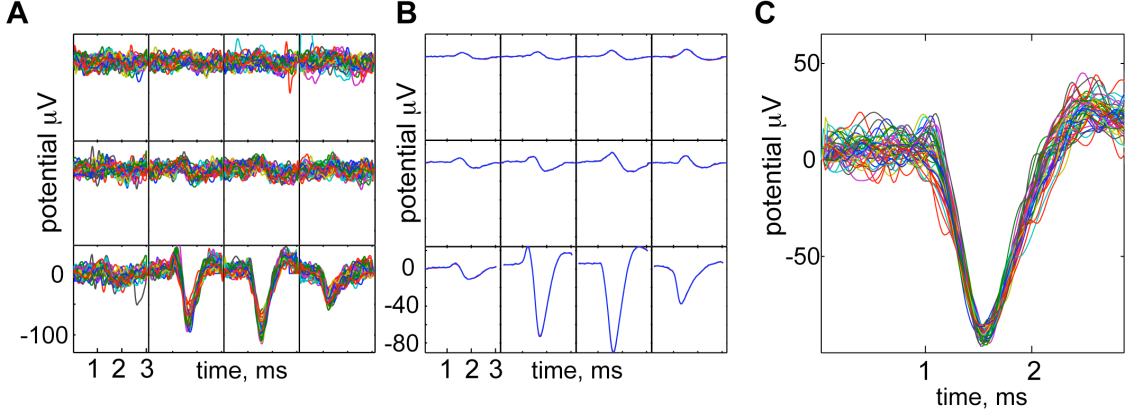
reliability (Step 4). Any significant clusters *missed* at this stage, for example because of the small fraction of the data used in this step, were found and reincorporated later during spike fitting (Step 3).

**Template building.** Next we created a consensus waveform (“template”) summarizing each cluster of cropped, upsampled events and characterized meaningful deviations from that consensus. We created a draft template by finding the pointwise median over all events in a cluster and then aligned each event to the draft template by maximizing their cross-correlation over time shifts, which we found to be more accurate than aligning to each event’s peak time. Finally, we found the pointwise median (to suppress the effects of outliers) of the aligned events; this waveform was our template (Fig. A.5B).

A key step was to realize that, in our data, the most significant sources of variation of individual spikes from the template were (a) additive noise, and (b) overall multiplicative rescaling of the spike’s amplitude (Fig. A.5C). To quantify (b), we found the rescaling factor that optimized the overlap of each spike with its template, then stored the mean and variance of those factors in a lookup table for later use as a prior probability for amplitude variation. We also logged the number of events associated to each template, converted to an approximate firing rate, and saved those rates, again for later use as a prior.

### A.5.3 STEP 3: Spike fitting

The third step in our procedure was to fit the spike templates constructed in Step 2 to each firing event in the data in order to determine which neural units were responsible for the activity. To this end, we constructed a simple generative model of firing events, and included distributions of firing rate and of amplitude variation



**Figure A.5: Template building.** (A) Detail of 40 of the aligned events used to compute a template, upsampled and shifted into alignment as described in the text. Some outlier traces reflect events in which this unit fired together with some other unit; the unwanted peaks occur at random times relative to the one of interest and thus do not affect the template. (B) *Blue*, detail of a template waveform, showing the potential on 12 neighboring electrodes. Time in ms runs horizontally; the vertical axis is potential in  $\mu\text{V}$ . *Red*, for comparison, the pointwise mean of the 430 waveforms used to find this template (nearly indistinguishable from the blue curve). (C) Detail of A, showing only the leader channel. Each trace has been rescaled by a constant to emphasize their similarity apart from variation in overall amplitude.

for each template. The firing probabilities and amplitude distributions were inferred from the previously clustered data, and therefore served as priors in the template fitting procedure described below. The fitting procedure iteratively identified and subtracted the most likely templates in each firing event.

The cluster templates were produced using an upsampled 50 kHz sample rate, but for fitting to data we downsampled back to the actual 10 kHz, in each of 5 “reading frames”; that is, we created five versions of each template corresponding to subsample shifts. Let  $F_{\mu;x,y}(t)$  be the potential of template  $\mu$  on the electrode with address  $x, y$ , at time  $t$ , with time measured in units of the sampling time  $\delta t = 0.1$  ms, and the template peak centered within the 3.2 ms template frame at the central point  $t = 16$ .

We use the vector notation  $\mathbf{F}_{\mu_i t_i}$  for the template  $\mu_i$  shifted to time  $t_i$ , i.e. its  $x, y, t$  component is  $F_{\mu_i, x, y}(t - t_i)$ .

**Generative model** The goal of spike fitting is to identify, for each spike event, all the units  $\{\mu_i\}$  which contribute to the event and their firing times  $\{t_i\}$  irrespective of their amplitudes  $\{A_i\}$ . Thus we assumed a probabilistic generative model of the data (Atiya, 1992; Lewicki, 1994; Sahani, 1999; Pouzat et al., 2002) and computed the posterior probability of  $\{(\mu_i, t_i)\}$  given the observed data. We assumed that a spike event  $\mathbf{V}$  could be explained by a linear combination of templates  $\mathbf{F}_{\mu_i t_i}$  with variable amplitudes  $A_i$  and correlated, zero-mean Gaussian noise  $\delta\mathbf{V}$ :

$$\mathbf{V} = \sum_{i=1}^K A_i \mathbf{F}_{\mu_i t_i} + \delta\mathbf{V}. \quad (\text{A.1})$$

Here  $K$  is the (a priori unknown) number of units contributing to the event. Given this model, to obtain the posterior probability that a firing event consists of a particular set of templates, we need to specify the prior probabilities of  $\mu$ ,  $t$ , and  $A$ . We chose a Gaussian prior for the amplitude  $A$ , a Poisson prior for  $\mu$ , and a uniform prior for  $t$ . Although  $A$  is a strictly positive quantity, we modeled its distribution with a Gaussian for analytical tractability. In practice, the distribution of  $A$  was tightly concentrated around its mean of approximately 1 and the Gaussian approximation had negligible weight at negative values (Fig. A.6D).

Our generative model assumes that spike waveforms from a given neural unit are stereotyped, apart from their amplitude. We did observe considerable variation in spike amplitude (Fig. A.2A), in part due to bursting (Fee et al., 1996; Harris et al., 2001), and thus included it in the model as a distribution of amplitude rescaling factors. Allowing for the possibility of slight variations in spike width also slightly

improved our results. But there was little additional variability to be modelled (Fig. A.5C). Our model also assumes that signals from different units combine linearly, as does the noise. This is reasonable, because the biophysics of extracellular recording is governed by the equations of electrodynamics, Ohm’s law, and other linear relations. A third assumption is that noise and the variability of spike amplitude are each well described by Gaussian distributions. Assuming Gaussianity (well-confirmed in some settings (Pouzat et al., 2002), but not others (Shoham et al., 2003)) allows for a fast, partially analytic approach to fitting. We validate this assumption quantitatively below.

Our generative model has a Poisson prior probability for firing by each neural unit, i.e. a prior that is as unstructured as possible while being consistent with the mean firing rate. The prior probability could be made somewhat more accurate by including refractory periods, the likelihood of bursting, and correlations between neural units. But this would significantly increase the complexity of the model, and inferring the distribution would require much more data (Schneidman et al., 2006).

Finally, we assumed that all statistical distributions that enter into the model are stationary and independent of the stimulus. While our retinal preparation does not suffer electrode drift (as might implanted electrodes), there are occasionally shifts in spike amplitudes and firing rates over the course of a lengthy experiment. Although in principle our fixed priors could lead to biased estimates, these biases are small when spike identification is robust, i.e. when the likelihood function dominates the prior in the posterior probability of a neural unit (Ventura, 2009).

**Noise characterization.** In the context of our generative model, in order to assess the probability that the residual after subtracting a putative spike is indeed noise, we first need to measure the distribution of noise. After applying the spatial whitening

filter (described in Step 1), our noise clips are decorrelated in space, but not in time (Fig. A.6A). Assuming that the noise has a correlated Gaussian distribution, we need the inverse of the noise covariance matrix,  $\mathbf{C}^{-1}$ . One approach to finding  $\mathbf{C}^{-1}$  is to invert the empirical covariance matrix  $\mathbf{C}$  of a large set of noise clips. Besides being intractable for larger arrays, this approach has the disadvantage that a numerically stable evaluation requires a very large noise sample.

For these reasons we instead took a parametric approach. We evaluated the covariance  $\mathbf{C}(x, y, t; x', y', t')$  and noted that it was approximately diagonal and translation-invariant in space (i.e. proportional to  $\delta_{x,x'}\delta_{y,y'}$  and independent of  $x$  and  $y$ ). It was also approximately stationary, i.e. invariant under time shifts  $t \rightarrow t + \Delta t$ ,  $t' \rightarrow t' + \Delta t$ , and thus only depended on  $t - t'$ . Finally, we observed that the time dependence of  $\mathbf{C}$  was roughly an exponential,  $\approx \eta e^{-|t-t'|/\tau}$  (Fig. A.6A). This gave:

$$\mathbf{C}^{-1}(x, y, t; x', y', t') = \eta^{-1} \times \delta_{x,x'}\delta_{y,y'} \begin{cases} (1 + \xi^2)/(1 - \xi^2) , & \text{if } t = t' \\ -\xi/(1 - \xi^2) , & \text{if } t = t' \pm \delta t \\ 0 , & \text{otherwise.} \end{cases} \quad (\text{A.2})$$

Here  $\delta t = 0.1$  ms is the sample time, and  $\delta_{x,x'}$  is the Kronecker symbol.  $\eta$  and  $\xi = e^{-\delta t/\tau}$  are obtained from the noise covariance. The dataset used in Results yields noise strength  $\eta \approx 57 \mu\text{V}^2$  and  $\xi \approx 0.58$ .

By construction, our noise model reproduces the 2-point correlations in the noise clips. However, real noise may not be Gaussian distributed. One check on this is to construct the transformed quantities  $\mathbf{U} = \mathbf{C}^{-1/2}\mathbf{V}$  empirically, find their full distribution as  $\mathbf{V}$  ranges over noise clips (the “one-point marginal” distribution), and compare to a normal distribution. Figure A.6B shows this comparison, lumping together every element of  $\mathbf{U}$ . The empirical noise deviates from a Gaussian only in

the far tails that contain very little weight.

**Fitting algorithm for single spikes.** Given the above characterization of the noise distribution, and our Gaussian prior for spike amplitude variation, the generative model Eqn. A.1 defined the posterior probability  $P(\{\mu_i, t_i, A_i\}|\mathbf{V})$  for templates  $\{\mu_i\}$  to be present at times  $\{t_i\}$  with amplitude scale factors  $\{A_i\}$ , given the recorded potentials  $\mathbf{V}$ . We ideally would have marginalized  $P(\{\mu_i, t_i, A_i\}|\mathbf{V})$  over the nuisance parameters  $\{A_i\}$  and then maximized with respect to  $\{(\mu_i, t_i)\}$  to identify the most probable set of units and spike times. In practice, this maximization is computationally expensive to perform on many templates simultaneously. Instead, we used a greedy approximation which fit one template at a time.

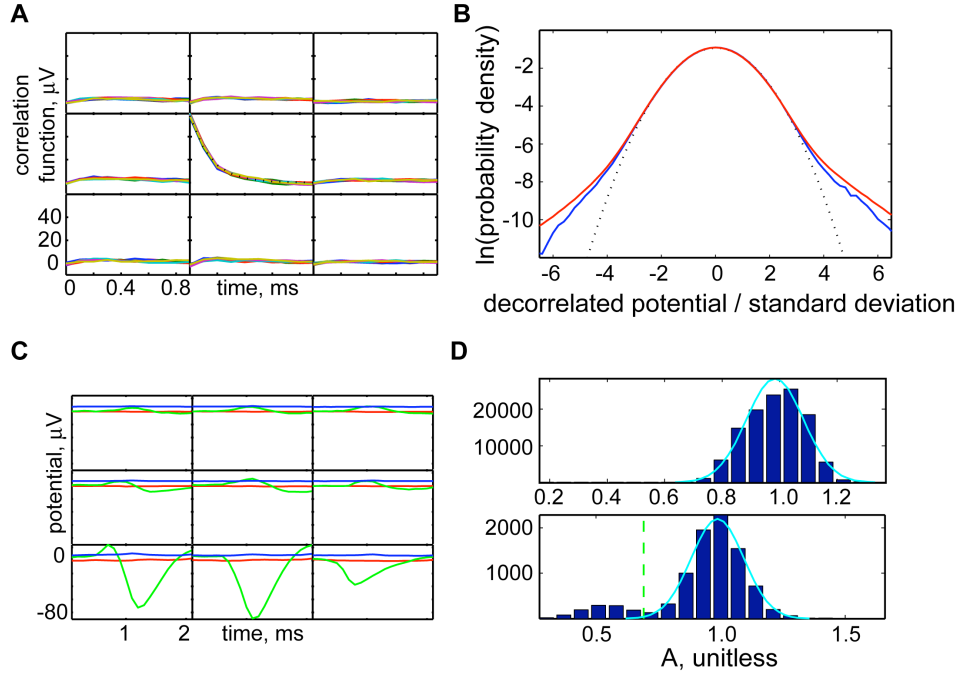
We first assumed that the event contained exactly one spike and identified the spike's type  $\mu$  and time of occurrence  $t$ . Bayes' formula gives for the posterior probability:

$$P(\mu, t, A|\mathbf{V})dtdA \propto P(\mathbf{V}|\mu, t, A)P(\mu, t, A)dtdA, \quad (\text{A.3})$$

up to a constant independent of  $\mu$ ,  $t$ , and  $A$ . Here  $P(\mu, t, A)$  is the prior probability of the template  $\mu$  appearing at time  $t$  with an amplitude  $A$ :

$$P(\mu, t, A)dtdA = r_\mu dt (2\pi\sigma_\mu^2)^{-1/2} \exp\left(-(A - \gamma_\mu)^2/2\sigma_\mu^2\right) dA, \quad (\text{A.4})$$

where  $\gamma_\mu$  is the mean and  $\sigma_\mu^2$  the variance of the scale factor for cluster  $\mu$ ;  $r_\mu$  is the estimated overall rate of firing for this cluster. The generative model gave the probability of the observed potential  $\mathbf{V}$  given  $\mu, t, A$  (the likelihood) as  $P(\mathbf{V}|\mu, t, A) = P_{\text{noise}}(\mathbf{V} - A\mathbf{F}_{\mu,t})$ , where  $P_{\text{noise}}$  is a Gaussian distribution with zero mean and covariance  $\mathbf{C}$  (Eqn. A.2). Combining the likelihood and prior, then integrating out  $A$ , gave



**Figure A.6: After fitting spikes, only noise remains.** (A) Noise covariance after spatial whitening. *Subpanels*: spacetime covariance  $C(x, y, t; x_*, y_*, t + \Delta t)$  between the central channel and its neighbors as a function of  $\Delta t$ , for various fixed  $t$  (colored curves). *Central panel* (dotted line): the function  $(57 \mu V^2) \exp(-\Delta t / (0.18 \text{ ms}))$ . (The various  $t$  lines and the dotted line are too similar to discriminate visually.) Horizontal axes:  $\Delta t$  in ms; Vertical axes:  $C$  in  $\mu V^2$ . (B) *Blue curve*, Semilog plot of the one point marginal probability density function of decorrelated noise samples. *Red curve*, same quantity, evaluated on residuals after spikes have been removed from spike events. *Dotted curve*, Gaussian chosen to represent this distribution. (C) *Green*, detail of the same template waveform shown in Figure A.5. *Red*, pointwise mean of the residuals after the fit spike is subtracted from 4,906 one-spike events of this type. The fact that it is nearly flat validates our assumption that spikes vary only in overall amplitude, and that noise is independent of spiking. *Blue*, pointwise standard deviation of the residuals, again evidence that only noise remains after fitting and subtracting spikes. (D) *Top*, histogram of fit values of the scale factor  $A$  for a template with peak amplitude  $-168 \mu V$  (well above noise) obtained without a prior on  $A$ , superposed with a Gaussian of the same mean and variance. *Bottom*, similar histogram for a low amplitude template. A secondary bump appears, due to noise fits, but is well separated from the main peak; a cutoff is shown as a dashed green line. The superposed Gaussian has mean and variance computed from the part of the empirical distribution lying above the cutoff.



the formula we ultimately used in our fitting algorithm:

$$\frac{P(\mu, t | \mathbf{V})}{P(\text{no spike} | \mathbf{V})} = \frac{r_\mu dt}{\sqrt{1 + \sigma_\mu^2 \mathbf{F}_{\mu,t}^t \mathbf{C}^{-1} \mathbf{F}_{\mu,t}}} \exp \left( \frac{1}{2\sigma_\mu^2} \frac{(\gamma_\mu + \sigma_\mu^2 \mathbf{V}^t \mathbf{C}^{-1} \mathbf{F}_{\mu,t})^2}{1 + \sigma_\mu^2 \mathbf{F}_{\mu,t}^t \mathbf{C}^{-1} \mathbf{F}_{\mu,t}} - \frac{\gamma_\mu^2}{2\sigma_\mu^2} \right). \quad (\text{A.5})$$

We have expressed Eqn. A.5 as a ratio with the probability that no templates contributed to the event. Finally, we varied  $\mu$  and  $t$  to find the template and time which maximized Eqn. A.5. This gave the first fit template,  $\mu_*$ , its firing time,  $t_*$ , and the probability ratio of the fit.

We improved scalability by a slight approximation. Starting from a spike event, we first identified the time and electrode address of its absolute peak and restricted the matrix products in expression Eqn. A.5 to only sum over a spatiotemporal neighborhood surrounding this peak. The size of the neighborhood was chosen to match the typical spatial extent and temporal duration of the templates.

**Multiple spikes.** In principle, we could have extended the single template procedure described above to compare the probabilities of all possible combinations of two or more spikes. Such an exhaustive approach, however, would quickly have become impractical. We instead noted that, even if an event contains multiple spikes, the single-spike fit described above still identified that template whose removal would lead to the largest increase in the probability that the remaining waveform is noise. Thus we adopted an iterative (matching-pursuit or “greedy”) approach: starting with a spike event, we found the absolute peak, fit it by the method described above, subtracted the fit, and then repeated the process (Segev et al., 2004).

The single-spike procedure found the most probable spike type  $\mu_*$ ; we then computed the scale factor  $A$  and spike time  $t'$  that would allow the fit spike to be subtracted as fully as possible. We thus held  $\mu$  fixed to  $\mu_*$  and minimized the ordinary

norm  $\|\mathbf{V} - A\mathbf{F}_{\mu_*, t'}\|^2$  over  $A$  and  $t'$ . The scaled and shifted template obtained in this way was subtracted before repeating the single-spike fitting procedure. The parameter  $t'$  was only used for template subtraction. The spike time which was actually reported by the algorithm was not  $t'$ , but  $t_*$  described above. In practice,  $t'$  was always very close to  $t_*$ .

To determine when to stop fitting spikes, we adopted a likelihood ratio test. At each step of the fitting loop, we summed Eqn. A.5 over  $t$ , obtaining the probability ratio that an additional spike of type  $\mu$  is present relative to the probability that *no* additional spike was present. We can then say that fitting an additional spike is justified if the ratio exceeds unity for some  $\mu_*$ . The fitting loop terminates when the significance test fails. Note that the factor  $r_\mu dt$  (the prior probability of a spike from unit  $\mu$ ) in Eqn. A.5 is typically much smaller than one. In order for the fitting algorithm to accept a candidate spike, the remaining terms in Eqn. A.5, which relate to goodness-of-fit, must be large enough to overcome this small factor. Furthermore, the amplitude prior is typically concentrated tightly around its mean. The marginalization over amplitudes thus suppresses templates that would have to stretch by a large factor in order to fit the data. Both of these effects counteract overfitting. Figure A.4C, D shows an example of the successful decomposition of a multiple-spike event using our method.

**Second pass.** The spike fitting algorithm might exit prematurely if a spike is present that does not appear in the list of templates initially extracted from the small subset of data. In this case, the fit will terminate, even though other identifiable spikes of lower amplitude may remain. To check this, if the residual exceeds  $V_{\text{trust}}^* = -44 \mu\text{V}$  after termination, the code declares an “incomplete fit” and writes the residual to a file; the small set of resulting waveforms are then reintroduced into

our clustering code and used for a second round of fitting. In this way we can be assured of finding even rare spike types, without having to perform clustering on the complete dataset. Using this method, only 0.02% of fits in the second pass were classified as incomplete. It can also happen that the small data sample used for clustering gives a poor estimate of some firing rates and amplitude distributions that enter our priors for spike fitting. Thus, before the second pass of fitting the priors are updated based on the outcome of the first pass.

#### A.5.4 STEP 4: Evaluation of template reliability

After spike identification, we performed a final evaluation to test whether templates and their sorted spike trains were trustworthy. The primary criteria were: (1) residuals after spike removal should resemble noise, (2) the histogram of amplitude scale factors should be unimodal, and (3) the inter-spike interval (ISI) distributions should display “refractory holes,” whereas the cross-correlation functions should not. Reliable templates were taken to be those that passed all these tests. Most unreliable templates failed multiple tests.

**Residuals.** For single-spike events, the residual signal after subtracting the fit should resemble pure noise. In particular, it should be stationary in time and translation-invariant in space. Figure A.6C shows that these expectations were met, validating our assumptions. For example, if the unit in Figure A.6C had significant variations other than amplitude rescaling, or if there were an amplitude-dependent noise process, then we would expect significant non-stationarity in the residual curves (Lewicki, 1994). To test that, after termination, the residual of a spike event consists only of noise, we computed the one-point marginal distribution of waveforms after all known spikes had been removed. Figure A.6B shows that this distribution closely

resembled that of noise clips, indicating that our code indeed found the significant spikes. Of particular note, the standard deviation of the residuals matched that of the noise: For each template, we found the standard deviation of the residuals of events to which only that template was fit. This value ranged from  $6.92 \mu\text{V}$  to  $8.63 \mu\text{V}$ , while the noise standard deviation was  $7.57 \mu\text{V}$ .

**Amplitude.** For large amplitude templates, the distribution of amplitude scale factors  $A$  obtained during spike fitting was typically close to Gaussian (Fig. A.6D). On the other hand, for low amplitude templates, accidental noise fits can sometimes lead to a histogram of  $A$  values with a secondary, low-amplitude peak well separated from the expected peak near  $A \approx 1$  (Fig. A.6D). Examining the amplitude histograms allowed us to quickly set an individual threshold for each reliable template. Fit spikes with  $A$  value below this threshold were discarded. If two peaks were discernible but overlapped significantly, the entire cluster was deemed unreliable and its spikes were not used in further analysis. In addition, our trigger rejected any spike event that did not cross  $-40 \mu\text{V}$ . Thus, any cluster whose amplitude histogram extended closer to zero than this was probably missing some true spikes and was not used.

**ISI distribution and cross-correlation.** Interspike interval distributions for single units are expected to have a refractory hole; our analysis of these distributions was described in Results. Two distinct neural units need not respect any mutual refractory period. Their spike-time cross-correlation function is therefore *not* expected to display any hole. We looked for such unexpected behavior and, when found, reexamined the corresponding templates. If the templates appeared to be duplicates, we merged the corresponding spike trains (Litke et al., 2004; Segev et al., 2004). Another diagnostic for duplicate templates is a coincident receptive field.

# Bibliography

- ANKERST M, BREUNIG MM, KRIEGEL HP, SANDER J (1999). OPTICS: Ordering points to identify the clustering structure. *SIGMOD Rec.* 28, 2:49–60.
- ATICK JJ (1992). Could information theory provide an ecological theory of sensory processing? *Network* 3:213–251.
- ATICK JJ, LI Z, REDLICH AN (1990). Color coding and its interaction with spatiotemporal processing in the retina. Technical report, Institute for Advanced Studies.
- ATIYA A (1992). Recognition of multiunit neural signals. *IEEE Transactions on Biomedical Engineering* 39, 7:723–729.
- ATTNEAVE F (1954). Some informational aspects of visual perception. *Psychol Rev* 61, 3:183–93.
- BACCUS SA, MEISTER M (2002). Fast and slow contrast adaptation in retinal circuitry. *Neuron* 36, 5:909–919.
- BALASUBRAMANIAN V, STERLING P (2009). Receptive fields and functional architecture in the retina. *J Physiol* 587, Pt 12:2753–67.
- BALBOA RM, TYLER CW, GRZYWACZ NM (2001). Occlusions contribute to scaling in natural images. *Vision Res* 41, 7:955–964.
- BARLOW HB (1961). Possible principles underlying the transformations of sensory messages. In: Sensory Communication (ROSENBLITH W, ed), pp 217–234. Cambridge, MA: MIT Press.
- BARLOW HB (2001). Redundancy reduction revisited. *Network* 12, 3:241–253.
- BEAUDOIN DL, BORGHUIS BG, DEMB JB (2007). Cellular basis for contrast gain control over the receptive field center of mammalian retinal ganglion cells. *J Neurosci* 27, 10:2636–2645.

- BEAUDOIN DL, MANOOKIN MB, DEMB JB (2008). Distinct expressions of contrast gain control in parallel synaptic pathways converging on a retinal ganglion cell. *J Physiol* 586, Pt 22:5487–502.
- BLOOMFIELD SA, VÖLGYI B (2009). The diverse functional roles and regulation of neuronal gap junctions in the retina. *Nat Rev Neurosci* 10, 7:495–506.
- BORGHUIS BG, RATLIFF CP, SMITH RG, STERLING P, BALASUBRAMANIAN V (2008). Design of a neuronal array. *J Neurosci* 28, 12:3178–89.
- BRADY N, FIELD DJ (2000). Local contrast in natural images: normalisation and coding efficiency. *Perception* 29, 9:1041–1055.
- BRENNER N, BIALEK W, DE RUYTER VAN STEVENINCK R (2000). Adaptive rescaling maximizes information transmission. *Neuron* 26, 3:695–702.
- BROWN SP, MASLAND RH (2001). Spatial scale and cellular substrate of contrast adaptation by retinal ganglion cells. *Nat Neurosci* 4, 1:44–51.
- BURTON GJ, MOORHEAD IR (1987). Color and spatial structure in natural scenes. *Applied Optics* 26, 1:157–170.
- BUZSÁKI G (2004). Large-scale recording of neuronal ensembles. *Nat Neurosci* 7, 5:446–51.
- CAFARO J, RIEKE F (2010). Noise correlations improve response fidelity and stimulus encoding. *Nature* 468, 7326:964–967.
- CHANDER D, CHICHILNISKY EJ (2001). Adaptation to temporal contrast in primate and salamander retina. *J Neurosci* 21, 24:9904–9916.
- CHEN X, HSUEH H, GREENBERG K, WERBLIN F (2010). Three forms of spatial temporal feedforward inhibition are common to different ganglion cell types in rabbit retina. *J Neurophysiol* 103, 5:2618–32.
- CHEN X, HSUEH HA, WERBLIN FS (2011). Amacrine-to-amacrine cell inhibition: Spatiotemporal properties of GABA and glycine pathways. *Vis Neurosci* 28, 3:193–204.
- CHICHILNISKY EJ (2001). A simple white noise analysis of neuronal light responses. *Network* 12, 2:199–213.
- CHICHILNISKY EJ, KALMAR RS (2002). Functional asymmetries in ON and OFF ganglion cells of primate retina. *J Neurosci* 22, 7:2737–47.
- DACEY D, PETERSEN M (1992). Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proc Natl Acad Sci U S A* 89, 20:9666.

- DAN Y, ATICK JJ, REID RC (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *J Neurosci* 16, 10:3351–3362.
- DAVENPORT CM, DETWILER PB, DACEY DM (2008). Effects of pH buffering on horizontal and ganglion cell light responses in primate retina: Evidence for the proton hypothesis of surround formation. *J Neurosci* 28, 2:456–64.
- DAVID SV, VINJE WE, GALLANT JL (2004). Natural stimulus statistics alter the receptive field structure of V1 neurons. *J Neurosci* 24, 31:6991–7006.
- DE RUYTER VAN STEVENINCK RR, LEWEN GD, STRONG SP, KOBERLE R, BIALEK W (1997). Reproducibility and variability in neural spike trains. *Science* 275, 5307:1805–8.
- DE VRIES SEJ, BACCUS SA, MEISTER M (2011). The projective field of a retinal amacrine cell. *J Neurosci* 31, 23:8595–8604.
- DEMB JB (2002). Multiple mechanisms for contrast adaptation in the retina. *Neuron* 36, 5:781–783.
- DENIZ S, WERSINGER E, SCHWAB Y, MURA C, ERDELYI F, SZABO G, RENDON A, SAHEL JA, PICAUD S, ROUX MJ (2011). Mammalian retinal horizontal cells are unconventional GABAergic neurons. *J Neurochem* 116, 3:350–362.
- DERRINGTON AM, LENNIE P (1982). The influence of temporal frequency and adaptation level on receptive field organization of retinal ganglion cells in cat. *J Physiol* 333, 0022-3751 (Print):343–66.
- DEVRIES SH (1999). Correlated firing in rabbit retinal ganglion cells. *J Neurophysiol* 81, 2:908–920.
- DEVRIES SH, BAYLOR DA (1997). Mosaic arrangement of ganglion cell receptive fields in rabbit retina. *J Neurophysiol* 78, 4:2048–60.
- DEVRIES SH, QI X, SMITH R, MAKOUS W, STERLING P (2002). Electrical coupling between mammalian cones. *Curr Biol* 12, 22:1900–7.
- DEWEESE M (1996). Optimization principles for the neural code. *Network* 7, 2:325–331.
- DO-NASCIMENTO JLM, DO-NASCIMENTO RSV, DAMASCENO BA, SILVEIRA LCL (1991). The neurons of the retinal ganglion cell layer of the guinea pig: Quantitative analysis of their distribution and size. *Brazilian J. Med. Biol. Res.* 24:199–214.

- DUEBEL J, HAVERKAMP S, SCHLEICH W, FENG G, AUGUSTINE GJ, KUNER T, EULER T (2006). Two-photon imaging reveals somatodendritic chloride gradient in retinal ON-type bipolar cells expressing the biosensor Clomeleon. *Neuron* 49, 1:81–94.
- ENROTH-CUGELL C, ROBSON JG (1966). The contrast sensitivity of retinal ganglion cells of the cat. *J Physiol* 187, 3:517–52.
- ESTEBANEZ L, EL BOUSTANI S, DESTEXHE A, SHULZ DE (2012). Correlated input reveals coexisting coding schemes in a sensory cortex. *Nat Neurosci* 15, 12:1691–1699.
- FAHEY PK, BURKHARDT DA (2003). Center-surround organization in bipolar cells: symmetry for opposing contrasts. *Vis Neurosci* 20, 1:1–10.
- FAHRENFORT I, STEIJAERT M, SJOERDSMA T, VICKERS E, RIPPS H, VAN ASSELT J, ENDEMAN D, KLOOSTER J, NUMAN R, TEN EIKELDER H, VON GERSDORFF H, KAMERMANS M (2009). Hemichannel-mediated and pH-based feedback from horizontal cells to cones in the vertebrate retina. *PLoS ONE* 4, 6:e6090.
- FAIRHALL AL, BURLINGAME CA, NARASIMHAN R, HARRIS RA, PUCHALLA JL, BERRY II MJ (2006). Selectivity for multiple stimulus features in retinal ganglion cells. *J Neurophysiol* 96, 5:2724–2738.
- FAIRHALL AL, LEWEN GD, BIALEK W, DE RUYTER VAN STEVENINCK RR (2001). Efficiency and ambiguity in an adaptive neural code. *Nature* 412:787–792.
- FARROW K, MASLAND RH (2011). Physiological clustering of visual channels in the mouse retina. *J Neurophysiol* 105, 4:1516–1530.
- FEE MS, MITRA P, KLEINFELD D (1996). Automatic sorting of multiple unit neuronal signals in the presence of anisotropic and non-Gaussian variability. *J Neurosci Methods* 69:175–188.
- FEIGENSPAN A, JANSSEN-BIENHOLD U, HORMUZDI S, MONYER H, DEGEN J, SOHL G, WILLECKE K, AMMERMULLER J, WEILER R (2004). Expression of connexin36 in cone pedicles and OFF-cone bipolar cells of the mouse retina. *J Neurosci* 24, 13:3325–3334.
- FELSEN G, TOURYAN J, HAN F, DAN Y (2005). Cortical sensitivity to visual features in natural scenes. *PLoS Biol* 3, 10:e342.
- FIELD DJ (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J Opt Soc Am A* 4, 12:2379–94.
- FISER J, CHIU C, WELIKY M (2004). Small modulation of ongoing cortical dynamics by sensory input during natural vision. *Nature* 431, 7008:573–578.



- FLORES-HERR N, PROTTI DA, WASSLE H (2001). Synaptic currents generating the inhibitory surround of ganglion cells in the mammalian retina. *J Neurosci* 21, 13:4852–63.
- FRAZOR RA, GEISLER WS (2006). Local luminance and contrast in natural images. *Vision Res* 46, 10:1585–1598.
- GANMOR E, SEGEV R, SCHNEIDMAN E (2011). The architecture of functional interaction networks in the retina. *J Neurosci* 31, 8:3044–3054.
- GEISLER WS (2008). Visual perception and the statistical properties of natural scenes. *Annu Rev Psychol* 59:167–192.
- GRANOT-ATEDGI E, TKAČIK G, SEGEV R, SCHNEIDMAN E (2013). Stimulus-dependent maximum entropy models of neural population codes. *PLoS Comput Biol* 9, 3.
- GRAY CM, MALDONADO PE, WILSON M, MCNAUGHTON B (1995). Tetrodes markedly improve the reliability and yield of multiple single-unit isolation from multi-unit recordings in cat striate cortex. *J Neurosci Methods* 63, 1-2:43–54.
- GRESCHNER M, SHLENS J, BAKOLITSA C, FIELD GD, GAUTHIER JL, JEPSON LH, SHER A, LITKE AM, CHICHILNISKY EJ (2011). Correlated firing among major ganglion cell types in primate retina. *J Physiol* 589, Pt 1:75–86.
- HARRIS KD, HIRASE H, LEINEKUGEL X, HENZE DA, BUZSÁKI G (2001). Temporal interaction between single spikes and complex spike bursts in hippocampal pyramidal cells. *Neuron* 32, 1:141–9.
- HINTON GE (2007). Learning multiple layers of representation. *Trends Cogn Sci* 11, 10:428–434.
- HOSOYA T, BACCUS SA, MEISTER M (2005). Dynamic predictive coding by the retina. *Nature* 436, 7047:71–77.
- JAYNES ET (1957). Information theory and statistical mechanics. *Physical Review* 106, 4:620.
- KAMERMANS M, FAHRENFORT I (2004). Ephaptic interactions within a chemical synapse: hemichannel-mediated ephaptic inhibition in the retina. *Curr Opin Neurobiol* 14, 5:531–541.
- KIM KJ, RIEKE F (2001). Temporal contrast adaptation in the input and output signals of salamander retinal ganglion cells. *J Neurosci* 21, 1:287–299.
- KIM KJ, RIEKE F (2003). Slow Na<sup>+</sup> inactivation and variance adaptation in salamander retinal ganglion cells. *J Neurosci* 23, 4:1506–1516.

- KOCH K, MCLEAN J, SEGEV R, FREED MA, BERRY II MJ, BALASUBRAMANIAN V, STERLING P (2006). How much the eye tells the brain. *Curr Biol* 16, 14:1428–34.
- KUFFLER SW (1953). Discharge patterns and functional organization of mammalian retina. *J Neurophysiol* 16, 1:37–68.
- LASATER EM, DOWLING JE (1985). Dopamine decreases conductance of the electrical junctions between cultured retinal horizontal cells. *Proc Natl Acad Sci U S A* 82, 9:3025–3029.
- LAUGHLIN SB (1981). A simple coding procedure enhances a neuron’s information capacity. *Z Naturforsch* 36, 9-10:910–912.
- LAUGHLIN SB (1994). Matching coding, circuits, cells, and molecules to signals: Principles of retinal design in the fly’s eye. *Prog Retin Eye Res* 13, 1:165–196.
- LEE AB, MUMFORD D, HUANG J (2001). Occlusion models for natural images: A statistical study of a scale-invariant dead leaves model. *International Journal of Computer Vision* 41, 1:3559.
- LEE EJ, HAN JW, KIM HJ, KIM IB, LEE MY, OH SJ, CHUNG JW, CHUN MH (2003). The immunocytochemical localization of connexin 36 at rod and cone gap junctions in the guinea pig retina. *Eur J Neurosci* 18, 11:2925–2934.
- LESICA NA, JIN J, WENG C, YEH CI, BUTTS DA, STANLEY GB, ALONSO JM (2007). Adaptation to stimulus contrast and correlations during natural visual stimulation. *Neuron* 55, 3:479–91.
- LEWICKI MS (1994). Bayesian modeling and classification of neural signals. *Neural Comput* 6, 5:1005–1030.
- LEWICKI MS (1998). A review of methods for spike sorting: The detection and classification of neural action potentials. *Network* 9, 4:R53–78.
- LI H, CHUANG AZ, O’BRIEN J (2009). Photoreceptor coupling is controlled by connexin 35 phosphorylation in zebrafish retina. *J Neurosci* 29, 48:15178–86.
- LITKE AM, BEZAYIEFF N, CHICHILNISKY EJ, CUNNINGHAM W, DABROWSKI W, GRILLO AA, GRIVICH M, GRYBOS P, HOTTOWY P, KACHIGUINE S, KALMAR RS, MATHIESON K, PETRUSCA D, RAHMAN M, SHER A (2004). What does the eye tell the brain?: Development of a system for the large scale recording of retinal output activity. *IEEE Trans. Nucl. Sci.* 51, 4:1434–1439.
- LU C, MCMAHON DG (1997). Modulation of hybrid bass retinal gap junctional channel gating by nitric oxide. *J Physiol* 499 ( Pt 3):689–699.

- LUKASIEWICZ PD (2005). Synaptic mechanisms that shape visual signaling at the inner retina. *Prog Brain Res* 147, 0079-6123 (Print):205–18.
- MALLAT SG, ZHANG ZF (1993). Matching Pursuits with Time-Frequency Dictionaries. *IEEE Transactions on Signal Processing* 41, 12:3397–3415.
- MANOOKIN MB, DEMB JB (2006). Presynaptic mechanism for slow contrast adaptation in mammalian retinal ganglion cells. *Neuron* 50, 3:453–464.
- MANOOKIN MB, WEICK M, STAFFORD BK, DEMB JB (2010). NMDA receptor contributions to visual contrast coding. *Neuron* 67, 2:280–293.
- MANTE V, FRAZOR RA, BONIN V, GEISLER WS, CARANDINI M (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nat Neurosci* 8, 12:1690–1697.
- MANU M, BACCUS SA (2011). Disinhibitory gating of retinal output by transmission from an amacrine cell. *Proc Natl Acad Sci U S A* 108, 45:18447–18452.
- MARAVALL M, PETERSEN RS, FAIRHALL AL, ARABZADEH E, DIAMOND ME (2007). Shifts in coding properties and maintenance of information transmission during adaptation in barrel cortex. *PLoS Biol* 5, 2:e19.
- MASLAND RH (2001). The fundamental plan of the retina. *Nat Neurosci* 4, 9:877–886.
- MASSEY SC, MILLER RF (1988). Glutamate receptors of ganglion cells in the rabbit retina: evidence for glutamate as a bipolar cell transmitter. *J Physiol* 405:635–655.
- MASTRONARDE DN (1983a). Correlated firing of cat retinal ganglion cells. I. Spontaneously active inputs to X- and Y-cells. *J Neurophysiol* 49, 2:303–324.
- MASTRONARDE DN (1983b). Correlated firing of cat retinal ganglion cells. II. Responses of X- and Y-cells to single quantal events. *J Neurophysiol* 49, 2:325–349.
- MASTRONARDE DN (1983c). Interactions between ganglion cells in cat retina. *J Neurophysiol* 49, 2:350–365.
- MASTRONARDE DN (1989). Correlated firing of retinal ganglion cells. *Trends Neurosci* 12, 2:75–80.
- MEISTER M, PINE J, BAYLOR DA (1994). Multi-neuronal signals from the retina: acquisition and analysis. *J Neurosci Methods* 51, 1:95–106.
- NIRENBERG S, BOMASH I, PILLOW JW, VICTOR JD (2010). Heterogeneous response dynamics in retinal ganglion cells: the interplay of predictive coding and adaptation. *J Neurophysiol* 103, 6:3184–3194.

- NIVEN JE, LAUGHLIN SB (2008). Energy limitation as a selective pressure on the evolution of sensory systems. *J Exp Biol* 211, Pt 11:1792–1804.
- PEICHL L (1989). Alpha and delta ganglion cells in the rat retina. *J Comp Neurol* 286, 1:120–139.
- PEICHL L, BUHL E, BOYCOTT B (1987). Alpha ganglion cells in the rabbit retina. *J Comp Neurol* 263, 1:25–41.
- PENG YW, BLACKSTONE CD, HUGANIR RL, YAU KW (1995). Distribution of glutamate receptor subtypes in the vertebrate retina. *Neuroscience* 66, 2:483–497.
- PERGE JA, KOCH K, MILLER R, STERLING P, BALASUBRAMANIAN V (2009). How the optic nerve allocates space, energy capacity, and information. *J Neurosci* 29, 24:7917–7928.
- PILLOW JW, SHLENS J, PANINSKI L, SHER A, LITKE AM, CHICHILNISKY EJ, SIMONCELLI EP (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454, 7207:995–9.
- PITKOW X, MEISTER M (2012). Decorrelation and efficient coding by retinal ganglion cells. *Nat Neurosci* 15, 4:628–635.
- POTTEK M, SCHULTZ K, WEILER R (1997). Effects of nitric oxide on the horizontal cell network and dopamine release in the carp retina. *Vision Res* 37, 9:1091–1102.
- POUZAT C, MAZOR O, LAURENT G (2002). Using noise signature to optimize spike-sorting and to assess neuronal classification quality. *J Neurosci Methods* 122, 1:43–57.
- PRENTICE JS, HOMANN J, SIMMONS KD, TKAČIK G, BALASUBRAMANIAN V, NELSON PC (2011). Fast, scalable, bayesian spike identification for multi-electrode arrays. *PLoS ONE* 6, 7:e19884.
- PUCHALLA JL, SCHNEIDMAN E, HARRIS RA, BERRY II MJ (2005). Redundancy in the population code of the retina. *Neuron* 46, 3:493–504.
- QUIROGA RQ (2007). Spike sorting. *Scholarpedia* 2, 12:3583 Rev. #73204.
- RATLIFF CP, BORGHUIS BG, KAO YH, STERLING P, BALASUBRAMANIAN V (2010). Retina is structured to process an excess of darkness in natural scenes. *Proc Natl Acad Sci U S A* 107, 40:17368–17373.
- REPERANT J, WARD R, MICELI D, RIO JP, MEDINA M, KENIGFEST NB, VESSELKIN NP (2006). The centrifugal visual system of vertebrates: a comparative analysis of its functional anatomical organization. *Brain Res Rev* 52, 1:1–57.

- RIEKE F (2001). Temporal contrast adaptation in salamander bipolar cells. *J Neurosci* 21, 23:9445–9454.
- ROCKHILL RL, DALY FJ, MACNEIL MA, BROWN SP, MASLAND RH (2002). The diversity of ganglion cells in a mammalian retina. *J Neurosci* 22, 9:3831–3843.
- RODIECK RW (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Res* 5, 11:583–601.
- ROSKA B, MOLNAR A, WERBLIN FS (2006). Parallel processing in retinal ganglion cells: how integration of space-time patterns of excitation and inhibition form the spiking output. *J Neurophysiol* 95, 6:3810–3822.
- RUDERMAN DL (1997). Origins of scaling in natural images. *Vision Res* 37, 23:3385–3398.
- RUDERMAN DL, CRONIN TW, CHIAO CC (1998). Statistics of cone responses to natural images: implications for visual coding. *J Opt Soc Am A* 15, 8:2036–2045.
- RUDERMAN DL, BIALEK W (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters* 73, 6:814–817.
- SAHANI M (1999). Latent variable models for neural data analysis. Ph.D. diss., California Institute of Technology, Pasadena CA.
- SCHNEIDMAN E, BERRY II MJ, SEGEV R, BIALEK W (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440, 7087:1007–1012.
- SCHNEIDMAN E, BIALEK W, BERRY II MJ (2003). Synergy, redundancy, and independence in population codes. *J Neurosci* 23, 37:11539–11553.
- SCHWARTZ G, HARRIS R, SHROM D, BERRY II MJ (2007). Detection and prediction of periodic patterns by the retina. *Nat Neurosci* 10, 5:552–554.
- SCHWARTZ O, CHICHILNISKY EJ, SIMONCELLI EP (2001). Characterizing neural gain control using spike-triggered covariance. In: Advances in neural information processing systems (DIETTERICH TG, BECKER S, GHAHRAMANI Z, eds), Vol. 14, pp 269–276. MIT Press.
- SEGEV R, GOODHOUSE J, PUCHALLA J, BERRY II MJ (2004). Recording spikes from a large fraction of the ganglion cells in a retinal patch. *Nat Neurosci* 7, 10:1154–61.
- SEGEV R, PUCHALLA J, BERRY II MJ (2006). Functional organization of ganglion cells in the salamander retina. *J Neurophysiol* 95, 4:2277–2292.

- SHAPLEY RM, ENROTH-CUGELL C (1984). Visual adaptation and retinal gain controls. *Prog Retin Res* 3:263–346.
- SHARPEE TO, SUGIHARA H, KURGANSKY AV, REBRIK SP, STRYKER MP, MILLER KD (2006). Adaptive filtering enhances information transmission in visual cortex. *Nature* 439, 7079:936–942.
- SHELLEY J, DEDEK K, SCHUBERT T, FEIGENSPAN A, SCHULTZ K, HOMBACH S, WILLECKE K, WEILER R (2006). Horizontal cell receptive fields are reduced in connexin57-deficient mice. *Eur J Neurosci* 23, 12:3176–3186.
- SHLENS J, RIEKE F, CHICHILNISKY E (2008). Synchronized firing in the retina. *Curr Opin Neurobiol* 18, 4:396–402.
- SHOHAM S, FELLOWS MR, NORMANN RA (2003). Robust, automatic spike sorting using mixtures of multivariate t-distributions. *J Neurosci Methods* 127, 2:111–22.
- SIMMONS KD, PRENTICE JS, TKAČIK G, HOMANN J, YEE HK, PALMER SE, NELSON PC, BALASUBRAMANIAN V (2013). Transformation of stimulus correlations by the retina. *PLoS Comput Biol* [under review].
- SIMONCELLI EP, OLSHAUSEN BA (2001). Natural image statistics and neural representation. *Annu Rev Neurosci* 24:1193–1216.
- SIMONCELLI EP, SCHWARTZ O (1999). Image statistics and cortical normalization models. In: *Advances in neural information processing systems* (KEARNS MS, SOLLA SA, COHN DA, eds), Vol. 11, pp 153–159. MIT Press.
- SMIRNAKIS SM, BERRY II MJ, WARLAND DK, BIALEK W, MEISTER M (1997). Adaptation of retinal processing to image contrast and spatial scale. *Nature* 386, 6620:69–73.
- SRINIVASAN MV, LAUGHLIN SB, DUBS A (1982). Predictive coding: a fresh view of inhibition in the retina. *Proc R Soc Lond B Biol Sci* 216, 1205:427–59.
- STEPHENS GJ, OSBORNE LC, BIALEK W (2011). Searching for simplicity in the analysis of neurons and behavior. *Proc Natl Acad Sci U S A* 108 Suppl 3:15565–15571.
- STERLING P (2004). How retinal circuits optimize the transfer of visual information. In: *The Visual Neurosciences* (CHALUPA LM, WERNER JS, eds), pp 234–259. Cambridge MA: MIT Press.
- TADMOR Y, TOLHURST DJ (2000). Calculating the contrasts that retinal ganglion cells and LGN neurons encounter in natural scenes. *Vision Res* 40, 22:3145–57.

- TAUCHI M, MORIGIWA K, FUKUDA Y (1992). Morphological comparisons between outer and inner ramifying alpha cells of the albino rat retina. *Experimental Brain Research* 88, 1:67–77.
- THEUNISSEN FE, DAVID SV, SINGH NC, HSU A, VINJE WE, GALLANT JL (2001). Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network: Comput Neural Syst* 12:289–316.
- THORESON WB, MANGEL SC (2012). Lateral interactions in the outer retina. *Prog Retin Eye Res* 31, 5:407–441.
- TKAČIK G, PRENTICE JS, BALASUBRAMANIAN V, SCHNEIDMAN E (2010). Optimal population coding by noisy spiking neurons. *Proc Natl Acad Sci U S A* 107, 32:14419–14424.
- TOURYAN J, FELSEN G, DAN Y (2005). Spatial structure of complex cell receptive fields measured with natural images. *Neuron* 45:781–791.
- TOYODA J, KUJIRAOKA T (1982). Analyses of bipolar cell responses elicited by polarization of horizontal cells. *J Gen Physiol* 79, 1:131–145.
- TRONG PK, RIEKE F (2008). Origin of correlated activity between parasol retinal ganglion cells. *Nat Neurosci* 11, 11:1343–1351.
- VAN HATEREN JH (1992). A theory of maximizing sensory information. *Biol Cybern* 68, 1:23–29.
- VANLEEUWEN M, FAHRENFORT I, SJOERDSMA T, NUMAN R, KAMERMANS M (2009). Lateral gain control in the outer retina leads to potentiation of center responses of retinal neurons. *J Neurosci* 29, 19:6358–6366.
- VARDI N, STERLING P (1994). Subcellular localization of GABAA receptor on bipolar cells in macaque and human retina. *Vision Res* 34, 10:1235–1246.
- VARDI N, ZHANG LL, PAYNE JA, STERLING P (2000). Evidence that different cation chloride cotransporters in retinal neurons allow opposite responses to GABA. *J Neurosci* 20, 20:7657–7663.
- VENTURA V (2009). Traditional waveform based spike sorting yields biased rate code estimates. *Proc Natl Acad Sci U S A* 106, 17:6921–6.
- VU TQ, PAYNE JA, COPENHAGEN DR (2000). Localization and developmental expression patterns of the neuronal K-Cl cotransporter (KCC2) in the rat retina. *J Neurosci* 20, 4:1414–1423.
- WARK B, LUNDSTROM BN, FAIRHALL A (2007). Sensory adaptation. *Curr Opin Neurobiol* 17, 4:423–429.

- WASSLE H, KOULEN P, BRANDSTATTER JH, FLETCHER EL, BECKER CM (1998). Glycine and GABA receptors in the mammalian retina. *Vision Res* 38, 10:1411–1430.
- WASSLE H (2004). Parallel processing in the mammalian retina. *Nat Rev Neurosci* 5, 10:747–757.
- WERBLIN FS (2011). The retinal hypercircuit: a repeating synaptic interactive motif underlying visual function. *J Physiol* 589, Pt 15:3691–3702.
- XIN D, BLOOMFIELD SA (1999). Dark- and light-induced changes in coupling between horizontal cells in mammalian retina. *J Comp Neurol* 405, 1:75–87.
- YIN L, GENG Y, OSAKADA F, SHARMA R, CETIN AH, CALLAWAY EM, WILLIAMS DR, MERIGAN WH (2013). Imaging light responses of retinal ganglion cells in the living mouse eye. *J Neurophysiol* 109, 9:2415–2421.
- ZAGHLOUL KA, BOAHEN K, DEMB JB (2003). Different circuits for ON and OFF retinal ganglion cells cause different contrast sensitivities. *J Neurosci* 23, 7:2645–2654.
- ZAGHLOUL KA, BOAHEN K, DEMB JB (2005). Contrast adaptation in subthreshold and spiking responses of mammalian Y-type retinal ganglion cells. *J Neurosci* 25, 4:860–8.