The Mechanisms of Reliable Coding in Mouse Visual Cortex

By

Rajeev V. Rikhye

M. Eng. Biomedical Engineering
Imperial College London, 2010

SUBMITTED TO THE DEPARTMENT OF BRAIN AND COGNITIVE SCIENCES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN NEUROSCIENCE
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

August 2016 [September 2016]

Signature of Author: Signature redacted
Department of Brain and Cognitive Sciences
August 05, 2016

Certified By: Signature redacted
Mriganka Sur
Paul E. and Lilah Newton Professor of Neuroscience
Thesis Supervisor

Approved By: Signature redacted
Matthew A. Wilson
Sherman Fairchild Professor of Neuroscience and Picower Scholar
Director of Graduate Education for Brain and Cognitive Sciences

The Mechanisms of Reliable Coding in Mouse Visual Cortex

By

Rajeev V. Rikhye

Submitted to the Department of Brain and Cognitive Sciences on August 05, 2016
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Neuroscience.

Abstract

As we interact with the environment, our senses are constantly bombarded with information. Neurons in the visual cortex have to transform these complex inputs into robust and parsimonious neural codes that effectively guide behavior. The ability of neurons to efficiently convey information is, however, limited by intrinsic and shared variability. Despite this limitation, neurons in primary visual cortex (V1) are able to respond with high fidelity to relevant stimuli. My thesis proposes that high fidelity encoding can be achieved by dynamically increasing trial-to-trial response reliability. In particular, in this thesis, I use the mouse primary visual cortex (V1) as a model to understand how reliable coding arises, and why it is important for visual perception. Using a combination of novel experimental and computational techniques, my thesis identifies three main factors that can modulate intrinsic variability.

My first goal was to understand the extrinsic, stimulus-dependent, factors responsible for reliably coding (**Chapter 3**). Natural scenes contain unique statistical properties that could be leveraged by the visual cortex for efficient coding. Thus, the first aim is to elucidate how image statistics modulate reliable coding in V1. To this end, I developed a novel noise masking procedure that allowed us to specifically perturb the spectral content of natural movies without altering the edges. Using high-speed two-photon calcium imaging in mice, I discovered that movies with stronger spatial correlations are more reliably processed by V1 neurons than movies lacking these correlations. In particular, perturbing spatial correlations in the movie dynamically altered the structure of interneuronal correlations. Movies with more naturalistic correlations typically recruited large neuronal ensembles that were weakly noise correlated. Using computational modeling, I discovered that these ensembles were able reduce shared noise through divisive normalization. Together, these findings demonstrate that natural scene statistics dynamically recruit neuronal ensembles to ensure reliable coding.

Microcircuits of inhibitory interneurons lie at the heart of all cortical computations. It has been proposed that these interneurons are responsible for reliable spiking by controlling the temporal window over which synaptic inputs are integrated. However, no study has yet conclusively investigated the role of different interneuron subtypes. Thus, my second goal was to establish **how** natural scenes are reliably encoded by dissecting

the inhibitory mechanisms underlying reliable coding (**Chapter 4**). Specifically, I investigated the role of somatostatin-expressing dendrite targeting interneurons (SST) and parvalbumin-expressing soma targeting interneurons (PV), which are known to provide distinct forms of inhibition onto pyramidal neurons. Using a novel combination of dual-color calcium imaging and optogenetic manipulation, I have discovered that the SST→PV inhibitory circuit plays a crucial role in modulating pyramidal cell reliability. In particular, by transiently suppressing PV neurons, SST neurons are able to route inhibition rapidly from the soma to the dendrites. Strong dendritic inhibition allows noisy inputs to be filtered out by the dendrites, while weaker somatic inhibition allows these inputs to be integrated to produce reliable spikes. In agreement with these results, I found that selectively deleting MeCP2 from these interneurons resulted in unreliable visual processing and other circuit-specific deficits, which are commonly observed in Rett Syndrome (**Chapter 5**). These results underscore the importance of intact inhibitory microcircuits in reliable processing.

Finally, my goal was to determine **why** reliable coding is necessary for visual processing (**Chapter 6**). To this end, I trained head-fixed mice to perform a natural movie discrimination task. Mice were able to learn how to discriminate between two movies after a short training period. By perturbing the amplitude spectrum of these movies, I discovered that mice used structural information in the phase spectrum to discriminate between the different movies. This suggests that mice also use similar strategies as higher mammals for scene recognition. Inspired by this result, we trained mice on a harder *target categorization task*, where mice had to identify the movies from an ensemble that were more similar to the target movie to gain a water reward. We developed this movie ensemble by blending together the phase spectrum of a target and non-target movie at different fractions. Optically activating SST neurons in V1 improved the ability of mice to correctly identify "target-like" movies. This increase in behavioral performance correlated well with an increase in V1 coding reliability. Thus, reliable codes are a prerequisite for accurate visual perception.

Taken together, this work bridges the gap between cells, circuits and behavior, and provides mechanistic insight into how complex visual stimuli are encoded with high fidelity in the visual cortex.

Thesis Supervisor: Prof. Mriganka Sur
Title: Paul E. and Lilah Newton Professor of Neuroscience

Acknowledgements

First and foremost, I would like to thank my thesis advisor, Prof. Mriganka Sur, for all the guidance and support that he has given me these past 5 years. Most importantly, Mriganka gave me the independence and freedom to realize my potential as a researcher and was always willing to give me sound advice whenever I needed it. He was constantly challenging me to take my research to the next level by posing tough and intriguing questions. He has single-handedly sculpted me into the researcher that I am today.

In addition to Mriganka, I have benefitted immensely from the mentorship and guidance of Dr. Sami El-Boustani. Sami taught me everything in visual neurophysiology, from computational to experimental techniques. He was always available to answer any questions and quell any doubts that I had regarding my projects. I am grateful for all the time that we spent planning experiments together. He was constantly pushing me to find new and interesting questions to pursue. Without Sami's encouragement and advice, I doubt I would be where I am today. It is a great honor for me to have met and learned from a scientist of his caliber.

One of the greatest advantages of being at MIT is interacting with some of the most brilliant minds in the field of neuroscience. I am especially humbled to have Prof. Jim DiCarlo and Prof. Matt Wilson on my thesis committee. Their tough questions forced me to rethink my project and this allowed me to maximize my potential as a scientist. I would like to thank Prof. Mark Andermann from Harvard for serving as my external examiner.

I would also like to acknowledge the Howard Hughes Medical Institute. None of this work would have been possible without the International Student Predoctoral Fellowship that supported my work from 2014-2016.

I am greatly thankful tor all the members of the Sur Lab who have made the 5 years that I spent in the lab so much more enjoyable. I am especially indebted to Dr. Ming Hu, without whom the behavior would not have been possible. Gerald Pho for vetting my speeches and presentations, and for his constant advice and discussions. Drs. Vincent Breton-Provencher, Murat Yildrim, Jaque Ip and Rafiq Huda, thank you both for making the lab such a lively place to work in.

My family has been the rock that has supported and guided me through the tumultuous and dark times of graduate school. My wife Annabel, thank you for being there with me at every step of the way. Finally, none of this would have been possible with out the support and love of my parents.

This page intentionally left blank

Table of contents

Chapter 1

Thesis Overview

One of the most iconic moments in rugby history, Joel Stransky's (number 10) drop goal wins the 1995 World Cup for South Africa. Picture credit: telegraph.co.uk

1.1 Introduction

It is extra time in the 1995 Rugby World Cup between the All Blacks and the Springboks, scores are tied 12-12 and the players are setup in a scrum. Joost van der Wusthizen wins the ball and passes it to fly-half Joel Stransky. Stransky now has to make a split second decision to either run with the ball or to attempt a 30m-drop goal. Within this time frame, the sensory areas in his brain have to rapidly integrate several sensory cues – the speed and direction of the wind, and the charging All Black defenders. He has to scan the field to identify open Springbok players to whom he can pass the ball. To attempt a drop goal, Stransky has to locate and segment the goal post from a background of fans frantically waving their flags and blowing their vuvuzelas. With seconds left on the clock, Stransky takes a chance and kicks the ball. The stadium erupts as the ball sails gracefully over the goal line. South Africa wins the Rugby World Cup 15-12.

This is a perfect example of how the brain is able to rapidly and precisely convert sensory information into goal-directed actions. Most intelligent organisms do this effortless on a daily basis. Our sensory systems enable us to distinguish a car horn from a cacophony of other sounds and identify a familiar face from a sea other faces. Understanding and emulating how these computations take place in the brain is a central goal in systems neuroscience. In this thesis, I will focus on elucidating the computations underlying visual perception.

We know that all information from the visual scene is packaged in the retina, the first stage of visual processing. It is the job of the visual cortex to reformat this information into a form that better suits the needs of the organism. For example, in Stansky's case, raw pixel information (luminance, contrast, color, etc.) would not help him much. Instead, he made his decision based on salient features, such as the shape of the ball or the location of the goal post relative to the background, which were reformatted and extracted from this raw pixel information and combined with his past experiences. These same transformations allow us to effortlessly identify the inverted triangle in **Figure 1-1A**. Since the Nobel Prize winning efforts of Hubel and Wiesel, we

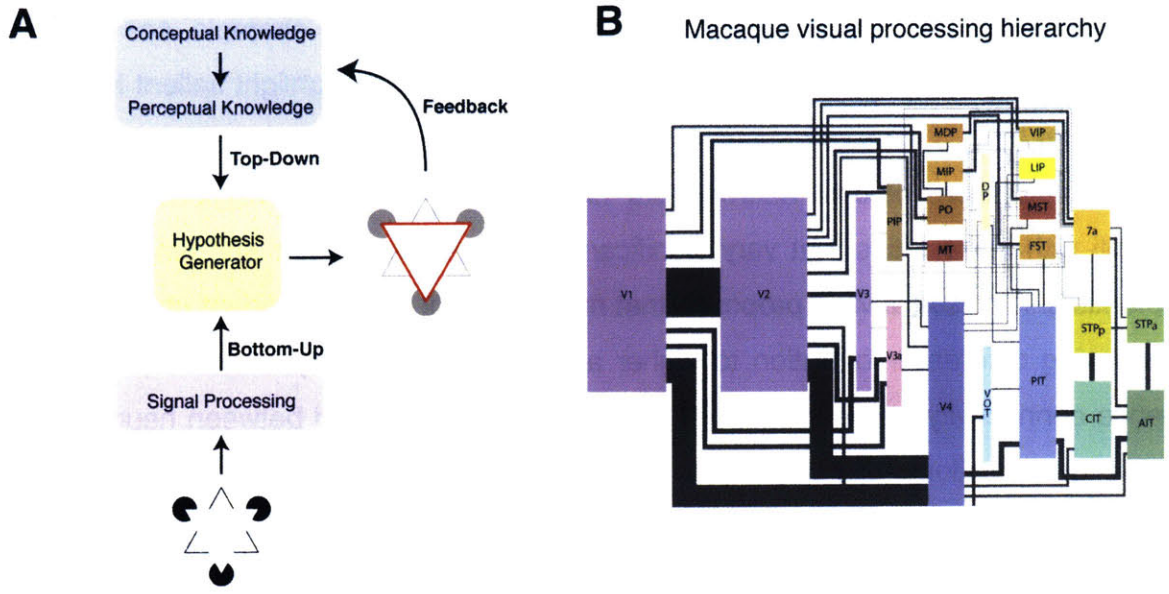

Figure 1-1. Visual perception. (A) A schematic illustrating the main concepts behind visual perception. Modified from Fregnac and Bathellier (Neuron, 2015) and inspired from the work of Richard Gregory. **(B)** Wiring diagram showing the flow of information between visual areas in macaque. Adapted from Wallisch and Movshon (Neuron, 2008) and based on work by J. Maunsell.

know that visual processing takes place over a hierarchy of different visual areas (**Figure 1-1B**), each performing distinct computations on the stream of raw pixel information. The details of these computations and the neuronal architecture that implements them remains an open question that is currently being actively pursued (reviewed in **Chapter 2**).

Thesis Goals

In this thesis, I am interested in answering a more fundamental question – how are signals arising from the natural environment processed in the primary visual cortex (V1)? Put in another way, I am interested in understanding the mechanisms of bottom-up perception. It is this bottom-up information that the visual hierarchy uses to generate models and inferences about the natural environment. Thus, it is in the best interest of the organism to represent information about salient features with the highest possible fidelity. However, the natural environment contains a vast amount of information, and as

a consequence many salient features tend to be occluded (for example, see image on Page 1). As a consequence, reformatting this information to highlight salient features is a daunting feat.

In addition, cortical neurons are highly variable as they respond to identical stimuli with spike trains that vary significantly in both the number and timing of action potentials. A classical view proposes that neurons work as independent channels, each conveying different information to higher areas in the hierarchy. However, due to the dense connectivity between neurons, variability is often shared between neuron, which leads to interactions, or correlations, between these channels. This variability is known to limit the information processing ability of neurons. Thus, given that signal processing is noisy, how then are we still able to generate reliable percepts?

The goal of my thesis is to understand how neurons are able to dynamically modulate intrinsic variability to serve the behavioral needs of the animal. In this thesis, I will discuss how neuronal variability can be controlled by both extrinsic and intrinsic factors. My work demonstrates that neurons in mouse V1 are able to overcome intrinsic variability to generate temporally precise spikes for naturalistic stimuli. Using a novel behavior, I also demonstrate that reliable coding at the level of the visual cortex is required for accurate visual perception. Overall, my thesis investigates the mechanisms by which neural circuits in V1 are able dynamically modulate intrinsic variability to improve the perception of complex visual stimuli.

More specifically, inspired by David Marr[1], the three fundamental questions that my thesis aims to address are the following.

1. **Computational Theory**. *What is the goal of reliable coding, why is it relevant to the behavior of mice? How does reliable coding ensure efficient coding? Why are natural scenes more reliably processed, what statistics allow this?*

[1] David Marr (1982). Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. New York: Freeman.

2. **Representation and Algorithm.** *How can reliable coding be implemented in the brain? What conditions do neurons and neural populations need to satisfy to ensure reliable responses?*

3. **Hardware implementation**. *How do neurons integrate noise synaptic inputs to produce reliable spikes? What is the role of different cell types, namely inhibitory interneurons?*

1.2 Thesis organization

To address these questions, I have organized this thesis into the following six chapters.

- In **Chapter 2**, I will briefly summarize the literature that has influenced this thesis. I will start this chapter by discussing the main sources of noise and variability in the cortex. I will also discuss natural scene statistics and how they can be used for image discrimination. Finally, I will elaborate on the computational role of various cell types in the cortex and how they can contribute towards efficient coding.

- In **Chapter 3**, I will demonstrate how stimulus statistics, namely the correlation between pixels, influences and modulates population coding within the mouse visual cortex. In this chapter, I will show that stimulus correlations play a critical role in determining the reliability of the neural code by recruiting specific ensembles of neurons. This work has been published in the Journal of Neuroscience.

- In **Chapter 4**, I will dissect the neural circuits underlying reliable coding of natural scenes focusing on the role of various inhibitory interneurons. Using in vivo calcium imaging, combined with high temporal resolution optical manipulations, I will identify a novel inhibitory neural circuit that can ensure reliable coding by modulating the signal-to-noise ratio within pyramidal neurons. This work is currently in preparation for submission to Cell.

- What happens when these neural mechanisms of reliable coding go awry? In **Chapter 5**, I will demonstrate that many of the cortical and behavioral deficits associated with Rett Syndrome, a monogenic disorder on the Autism spectrum, are due to a dysfunction in cortical information processing. This work is currently in review at PNAS.

- In **Chapter 6**, using a novel head-fixed behavior, I will demonstrate that mice require reliable neural codes to accurately perceive and discriminate complex natural movies. In particular, I will show that activation of somatostatin-expressing inhibitory interneurons improves the ability of mice to categorize different scenes. In this chapter, I also present a concluding overview and discuss future directions that can be pursued to better elucidate how reliable coding emerges in the visual cortex.

Together, the work that I present in this thesis lays the groundwork for understanding how high fidelity sensory processing takes place within the mouse visual cortex. I believe that this is a necessary first step towards decoding the mystery that is visual processing.

This page intentionally left blank

Chapter 2

What do we already know?

Summary

A fundamental goal in systems neuroscience is to understand how natural signals are transformed into neural representations. Studies have shown that the visual system is more efficient at processing information from natural scenes compared to simplistic stimuli. In particular, natural movies are processed using reliable codes with low trial-to-trial variability. However, the neuronal mechanisms that are responsible for this transformation remain poorly understood. The overarching aim of this thesis is to understand the mechanisms responsible for this efficient coding. In this chapter, I discuss recent advances in understanding how neuronal variability is modulated in the cortex. I also justify why mice are a good model system for studying natural scene processing. Finally, I speculate on how inhibitory interneurons could work together to sculpt the reliability of neural responses. This literature review raises several questions (highlighted below) that I would like to answer in this thesis.

Questions that remain unanswered:

1) How do natural scene statistics modulate response reliability?
2) What inhibitory mechanisms can suppress variability?
3) What statistics do mice use for visual perception?
4) How are cortical computations altered when inhibition goes awry?

2. 1 General Introduction

As we interact with the natural environment, our senses are inundated with information. A fundamental role of the visual cortex is to transform signals arising from the natural environment into more efficient and useful forms. A central goal of systems neuroscience, and this thesis, is to understand how these transformations take place. One longstanding guiding principle is Barlow's efficient coding hypothesis, which asserts that sensory systems have evolved to maximize information transmitted to the brain from the natural environment (Barlow, 1994). In further detail, Barlow proposed that efficient coding is achieved by allocating more resources to encode more frequently occurring signals (Simoncelli and Olshausen, 2001). This would help to remove statistical redundancy from the inputs. In particular, the efficiency of a neural code depends critically on three factors. (1) The statistics of signals that are to be encoded. (2) The variability (or the noise) induced during signal transformation. (3) The metabolic cost of generating the code.

If the efficient coding hypothesis is correct, what behaviors should we expect to see in the response properties of neurons? The biggest hurdle to overcome in any biological system is intrinsic variability. Thus, in order to efficiently encode any stimulus, responses to that stimulus must be relatively noise free. Moreover, at the population level the efficient coding hypothesis postulates that neurons should be independent channels of information transmission. This is, however, limited by the fact that neurons share strong correlations between each other in both space (across the cortex) and in time (in activation synchrony). As a consequence, efficient codes should have responses that are decorrelated. Indeed, numerous studies (reviewed here) have shown that natural movies are processed responses that are reliable from trial-to-trial. The goal of this thesis is to understand what mechanisms the cortex uses to ensure reliable coding. In particular, I will focus on two aspects: (1) the role of stimulus statistics and (2) the role of inhibitory interneurons in modulating coding reliability. In this chapter, I will discuss recent literature that relevant towards achieving this aim.

2.2 Neuronal Variability: Feature or Bug?

Cortical neurons transmit information using sequences of action potentials. To an observer, these responses often appear variable as repeated measurement often give very different spike patterns, even under identical conditions. From the experimenter's perspective this variability is interpreted as noise, or uninformative and unpredictable signals that carry no information. However, from the neurons perspective, these unexplained spikes could provide information about the state of the network. Thus, when characterizing the response properties of neurons it is important to understand neuronal variability and the sources of this variability.

Remarkably, neuronal variability is dynamic, as it appears more dominant in some conditions than others. For example, complex natural movies are process with relatively noise-free codes compared to simple stimuli (Baudot et al., 2013; Haider et al., 2010). Why does this happen and what are the mechanisms responsible? The goal of my thesis is to answer these questions. Before I begin to answer these questions, I provide a definition of variability, and review the how variability arises in the cortex and how it can be controlled.

Measuring variability

The most commonly used method to measure *stimulus-evoked* and *stimulus-independent* components of a sensory neuron's activity is to measure the response of this neuron's response (i.e. spikes or calcium response) to repeated presentations of the identical stimulus. Let us define the spike count over some time bin (denoted by the shaded boxes in **Figure 2-1A**) to the x^{th} trial as $Y_x(t)$. Then, the across-trial average firing rate of this neuron will thus be $r(t) = E_x[Y_x(t)]$, where $E_x[\cdot]$ is the expectation or mean taken over trials. This component is most commonly referred to as signal. Similarly, the variability of this neuron across trials will be $v(t) = Var_x[Y_x(t)]$. Using the law of total variance, this across trial variability can be further broken down into stimulus-evoked variance plus a stimulus-independent variance as follows:

$$v(t) = Var_t(r) + E_t[Var_t(Y|r)]$$

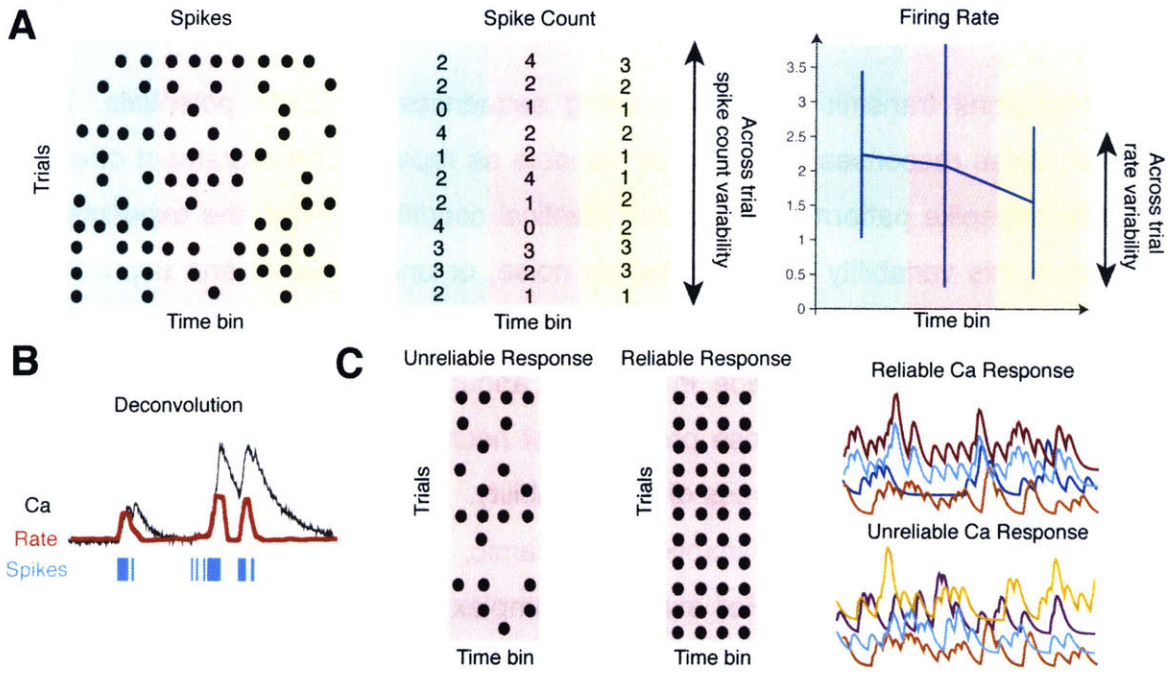

Figure 2-1. How neuronal variability is calculated. (A) Cartoon illustrating spike count and firing rate variability. **(B)** In calcium imaging experiments, we do not have access to spike times, thus we must deal with computing PSTH variability across trials. **(C)** Examples contrasting a reliable response from an unreliable response.

Thus, the variance of the firing rate ($Var_t(r)$) distribution provides a measure of the fraction of the variance of the response that can be explained by the stimulus and the average variance of the response ($E_t[Var_t(Y|r)]$), provides a measure of the variance that cannot be explained by the stimulus. This component is most commonly referred to as noise (Goris et al., 2014). The variance of the firing rate is easy to measure and interpret as it simply a measure of the peri-stimulus time histogram (PSTH) of a neuron. However, this equation also shows that the variance of the firing rate (PSTH) underestimates a neuron's total variance. In contrast, the fraction of the variance that cannot be explained by the stimulus is much harder to measure and interpret. As we will discuss below, this unexplained variability arises from several intrinsic and extrinsic sources.

One powerful measure of a neuron's response variability is the signal-to-noise ratio (SNR). Essentially, the SNR given by the equation below provides a measure of what fraction of the variability of a neuron can explained by the stimulus.

$$SNR = \frac{Var_t(r)}{E_t[Var_t(Y|r)]}$$

Thus, if SNR is high, then more of the variance in the response can be explained by properties of the stimulus. Otherwise, if the SNR is low, then most of the variance cannot be explained by the stimulus. An alternate SNR measure is the reciprocal of the coefficient of variation (CV), which provides a measure of how much the mean of the firing rate differs from its standard deviation. In other words, the CV measures of the firing rate variability with respect to the mean.

$$SNR = CV^{-1} = \frac{E_t[r(t)],}{\sqrt{Var_t(r)}}$$

In calcium imaging experiments, however, we do not have access to single spike times. This is primarily limited by both the dynamics of the indicator and the speed of imaging (Chen et al., 2013). As a consequence, the observed calcium transients could result from a numerous spikes, and thus calculating the exact spike count within a particular time bin is close to impossible (**Figure 2-1B**). Thus, following temporal deconvolution (Theis et al., 2016; Vogelstein et al., 2010), what we observe on a single trial is a predicted PSTH of that neuron on that trial. Nevertheless, we can still use the law of total variance to estimate the fraction of the variance that cannot be explained by the stimulus as follows.

$$Var_t(r) = Var_t(E_t[r]) + E_t[Var_t(r)]$$

In this equation, the first term is the measured PSTH variance across trials (i.e. stimulus-driven variance) and the second term ($E_t[Var_t(r)]$) denotes the fraction of the variance that cannot be explained by the stimulus. In this thesis, I will refer to this second term as noise. Thus, using this equation, a more appropriate SNR measure for calcium imaging is

$$SNR_{Ca} = \frac{Var_t(E_t[r])}{Var_t(r)} \in [0,1]$$

This SNR measure ranges between 0 and 1, with a value of 1 indicating that there is no unexplained variance in the response. In this **Chapter 3** of this thesis, I use

CV^{-1} to measure the effect of image statistics on SNR, and in **Chapter 4**, I use SNR_{Ca} to measure how increasing or decreasing inhibition reduces unexplained responses.

Another powerful measure of neuronal variability is response reliability. Mathematically, the response reliability is computed by averaging the Pearson's correlation coefficient between all pair-wise combinations of trials (Bathellier et al., 2012; Haider et al., 2010).

$$Rel = \frac{2}{T^2 - T}\sum_{i=1}^{T}\sum_{j\neq i}^{T} Corr\left(r_i(t), r_j(t)\right)$$

$$Corr\left(r_i(t), r_j(t)\right) = \frac{E\left[r_i(t)r_j(t)\right] - E\left[r_j(t)\right]E\left[r_i(t)\right]}{\sqrt{Var\left(r_j(t)\right)}\sqrt{Var\left(r_i(t)\right)}}$$

Thus, reliability provides a measure the degree of similarity in a neuron's response across trials. A neuron is said to have high reliability if it responds consistently over several stimulus repetitions (**Figure 2-1C**). Intuitively, a neuron that has high reliability will produce precise spikes to every repetition to the same stimulus feature, and consequentially will have low unexplained variance (within the time bin of interest). Also, it is much easier to decode stimulus identity from a neuron that responds consistently to the same stimulus feature (Tiesinga et al., 2008).

In theory, an ideal model neuron should respond reliably and strongly to a stimulus that matches it's receptive field, because the stimulus remains unchanged between trials. However, adding noise to this model, either in the stimulus or in the neuron itself, will induce unreliability in the response. Thus, by keeping the stimulus fixed, and by measuring both reliability and SNR, we can get a sense of how much noise the intrinsic factors are adding to the response. In the next section I will review some of these sources of intrinsic variability.

Sources of Stimulus-independent variability

In now what has become a seminal set of experiments, Mainen and Sejnowski discovered that repeatedly injecting a white noise current into a cortical neuron led to temporally precise spiking across repetition. In stark contrast, the same neuron

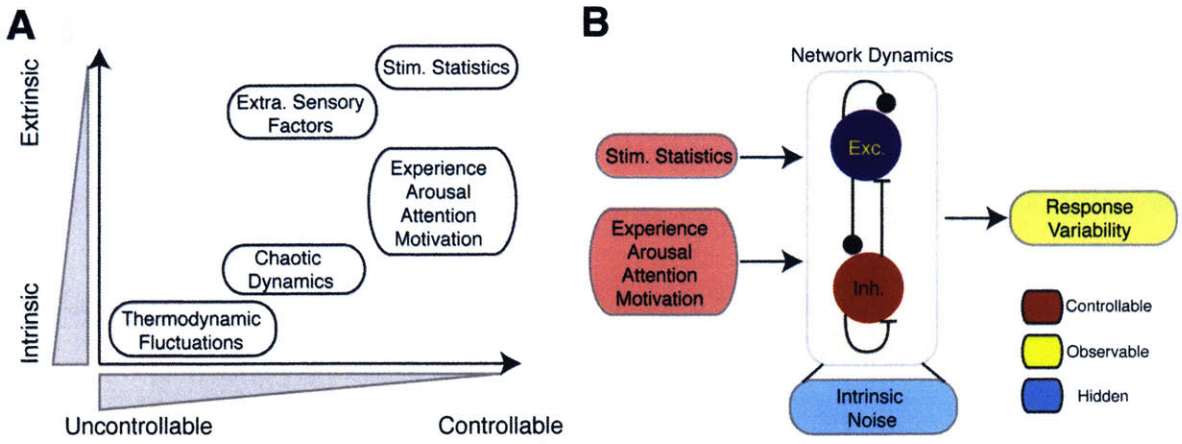

Figure 2-2. Sources of noise in the sensory system. (**A**) Categorization of sources of noise according to controllability. The y-axis categorizes these sources into whether they are intrinsic or extrinsic to the sensory system. The x-axis classifies these sources according to their controllability, that is, to what extent they are accessible to experimental control or monitoring. (**B**) Cartoon illustrating how controllable sources of noise act through network dynamics to alter response variability.

responded to a constant current with highly irregular spiking (Mainen and Sejnowski, 1995). Interestingly, both the reliability (measured using the equation above) and the precision of spiking increased with increasing variance of the white noise current. This work demonstrated that neurons are able to integrate stochastic inputs to produce stochastic inputs. In other words, the same neuron can change its reliability depending on the statistics of the inputs that it receives.

In contrast, when measured *in vivo*, cortical neurons have been found to respond to repeated stimulus presentations with spike trains that vary greatly in both the number and the timing of spikes (Azouz and Gray, 1999; Heggelund and Albus, 1978; Tolhurst et al., 1983). Surprisingly, cortical variability is not inherited from subcortical structures, which appear to fire consistently to identical stimuli (Butts et al., 2007; Kara and Reid, 2003). These observations have led to the conclusion that it is spontaneous variations in the cortical state that drives cortical variability, in particular the arrival of synchronous synaptic barrages (Softky and Koch, 1993; Stevens and Zador, 1998). These spontaneous fluctuations are believed to reflect the internal state of the animal (Gur et al., 1997) and arise due to fluctuations in attention (McAdams and Maunsell, 1999) or arousal (Reimer et al., 2014). In particular, it has been shown that neuronal variability covaries with behavioral variability (Britten et al., 1996). Interestingly, neuroimaging

studies have revealed excessive noise in patients with Autism (Dinstein et al., 2015). Together this suggests that neuronal variability might indeed be maladaptive (Parker and Newsome, 1998; Shadlen and Newsome, 1998). Several theoretical studies have now show that reliable neurons can be driven to fire chaotically when embedded in networks where inhibition and excitation is balanced (i.e. balanced network) (Denève and Machens, 2016; Litwin-Kumar and Doiron, 2012). These studies reconcile these observations of variable firing *in vivo* with the findings of Mainen and Sejnowski, and show that it is the network state that influences neuronal reliability.

What are the additional sources of variability and can they be controlled? In a recent review, Renart and Machens (2014) succinctly classified cortical noise into four categories, which are summarized in **Figure 2-2A**. Most sources of cortical noise can be either intrinsic or extrinsic to the system of interest. The best example of intrinsic noise is the stochastic fluctuation of ion channels in neurons, which is governed by thermodynamic principles. Further examples include the random walk diffusion of neurotransmitters at synapses or second-messenger signaling molecules within neurons. These thermodynamic-driven fluctuations can neither be observed nor can they be controlled. As a consequence, these intrinsic sources of noise set a lower bound on the fidelity with which neural computations can occur (Faisal et al., 2008; Renart and Machens, 2014).

A classical manifestation of intrinsic noise is the spontaneous postsynaptic potential, which can be seen even in the absence of an input to a neuron. These random synaptic events can sum to cause excursions of membrane potential above the threshold potential, resulting in spontaneously occurring spikes (Graupner and Reyes, 2013). Worst still, due to various nonlinearities imposed by network connectivity, these noise-driven spikes can accumulate to cause chaotic dynamics within the network (van Vreeswijk and Sompolinsky, 1996). Thus, stimulus driven responses ride on top of these stochastic dynamics, which could explain why spike timing *in vivo* is imprecise. Given the high degree of trial-to-trial variability, it is commonly believed that the cortex uses a *rate-based* code rather than a *spike-time-based* code to transmit information. In rate-based codes, information about the stimulus is encoded in the weighted average firing

rate of the neuron. Supporting this idea, several studies have shown that cortical activity patterns are very sensitive to small perturbations (London et al., 2010).

Also, because neurons are so densely interconnected, these sources of noise are shared by most neurons. As a consequence, neurons appear to be *noise correlated*, in the sense that neuronal activities co-fluctuate in response to a fixed stimulus ((Averbeck et al., 2006). Noise correlations capture the fraction of the response that is variable from trial to trial, and thus stem from relationships between neurons that are not locked to the external stimulus. These factors can include the strength and dynamics of synaptic connectivity between neurons. These noise correlations to set a fundamental lower bound to the precision and speed at which neurons can encode and convey information (Moreno-Bote et al., 2014; Zohary et al., 1994).

Alternatively, noise could also be caused by factors that cannot be explained. By this token, a noise-free response is one in which every aspect of the response can be explained by a detailed model of both the stimulus presented to the animal and the animal's internal state. Hence, controllable sources of noise are those that can be accounted for by factors that can be measured (or explained) by the experimenter. As a consequence, more variables the experimenter can observe, the lower will be unexplainable fraction of the response. In recent study, Hires and colleagues found that aligning spikes from L4 neurons in the primary somatosensory cortex (S1) to touch on set resulted in an almost *noise-free* representation, with very low SNR (measured as CV^{-1}). In contrast, when spikes were aligned to the stimulus onset they were much more variable between trials (see Fig. 1 in (Hires et al., 2015)). This study neatly showed that many of the spurious spikes in the recording could be explained by observing the dynamics of whisker deflections. Interestingly, they did not find a similar phenomenon in L5, presumably because L5 neurons receive inputs from many other cortical areas, which in turn provide more unexplainable variance.

Similarly, work from the Tolias lab used pupillometery to show that response variability in mouse V1 was lower when the pupil was actively dilating (Reimer et al., 2014). It has been well established that pupil diameter correlated with the arousal state of the animal (Aston-Jones and Cohen, 2005), thus pupillometery provides a powerful

tool to measure, and account for, the internal arousal state of the animal in real time. Increases in pupil diameter correlated with an increase in cortical activity and a decrease in low-frequency oscillations. Thus, the increase in reliability seen during periods of dilation could be due to a decrease in the frequency of stochastic synaptic barrages. Additionally Reimer et al. showed that changes in pupil diameter led to the differential activation of VIP and SST interneurons, implying that arousal-driven changes in reliability could be enforced by local interneurons. We explore the notion of how interneurons can sculpt response reliability in **Chapter 4.**

In summary, these studies show that fluctuations in the sensor (pupil in the case of vision and whisker in the case of somatosenstation) and the arousal state of the animal both strongly impact the level of variability in a neurons response. In other words, spuriously occurring spikes are due both stochastic membrane fluctuations and inputs into a neuron that are not due to the stimulus. In order to effectively encode a stimulus, the cortex should have mechanisms to dynamic suppress this noise and maximize stimulus drive. Uncovering how this is achieved is an aim of this thesis.

Controllability of noise

Noise in neuronal systems is dynamic and can either change slowly over learning or rapidly through shifts in attention. For example in the motor cortex, neural activity patterns start out as variable and chaotic as the animal learns a new motor action. But as the motor actions are refined through learning, neural activity becomes more reliable and exhibits reproducible spatiotemporal patterns of activation (Georgopoulos et al., 1982; Peters et al., 2014). In the visual cortex, repeated presentation of the identical natural movie sequence causes neurons to become more reliable (Yao et al., 2007). This study argued that NMDAR-dependent plasticity and adaptation work to refine the neuronal circuit and improve response reliability. Similarly, activation of basal forebrain cholinergic projections also can improve reliability by reducing correlated noise between neurons (Goard and Dan, 2009). In fact, high variability during learning has been noted in numerous systems, and such variability is believed to be the basis for exploration of all possible network states and facilitate learning (Rokni et al., 2007). In computer

science, literature stochastic gradient descent is often used for optimization when the parameter space is large. Thus, early variability allows *agents* or neurons to explore all possible motor outcomes and select the motor behavior that is associated with the best outcome (Fiete and Seung, 2006). Hence, whereas large variability is harmful to signal processing, it might be beneficial during early learning.

In addition to stabilizing plasticity, changing redundancy between neurons might also be a critical means for controlling neuronal variability. In particular, when large populations of neurons encode stimuli, averaging across the population can minimize the effect of individual variability. However, the efficiency of this averaging process depends on noise correlations between neurons (Averbeck et al., 2006; Moreno-Bote et al., 2014). These noise correlations, or correlations between neurons for a fixed stimulus, are known to play a significant role in limiting information processing (Abbott and Dayan, 1999; Zohary et al., 1994). Interestingly, it has been shown that both perceptual learning and attention both can reduce noise correlations between neurons in sensory areas (Cohen and Maunsell, 2009; Gu et al., 2011). Also, learning can change the sign of the relationship between signal and noise correlations (Jeanne et al., 2013). Surprisingly, however, a study found that attention did not change the reliability of a neuronal response (McAdams and Maunsell, 1999) even though it changed response gain (Maunsell, 2015). Together, this work suggests that attention functions primarily to alter the way large ensembles of neurons sum together to minimize noise. In doing so, attention is able to select ensembles that encode behaviorally-relevant information and enhance the information contained within them. Thus, these studies show that active processes, such as attention, perceptual learning and plasticity, can dynamically modulate neuronal variability.

In summary, I have shown in this section that neuronal variability can be portioned into stimulus-driven variability and stimulus-independent variability. The stimulus-independent variability can arise from multiple sources and mainly reflect the attentional state of the animal. In the next section, I will explore stimulus-driven variability and how this variability can change depending on the statistics of the visual stimulus.

2.3 Stimulus statistics and their influence on visual processing

In the previous section, we reviewed the sources of cortical variability and how altering network dynamics can control them. In this section, we will focus on the influence that stimulus statistics have on cortical processing. Specifically, we investigate how natural scenes are processed in the visual system and explain why natural scenes are more reliably processed than simpler visual stimuli.

Overview of natural scene statistics

Before we begin the discussion of how natural scenes are processed in the visual cortex, it is important to understand the statistical properties of natural scenes. Natural scenes are broadband stimuli and contain an entire spectrum of spatial and temporal frequencies. Although natural scenes appear complex, they do contain several statistical regularities (Dong and Atick, 1995; Ruderman, 1997). Due to the presence of textures, contours and edges, neighboring pixels in natural images are strongly correlated with each other. Fourier analysis is a useful method of quantifying and characterizing the spatial statistics of natural images. Using the two-dimensional Fourier transform, any natural image can be decomposed into amplitude (or power spectrum) and a phase spectrum, each with unique properties.

Mathematically, following the Wiener–Khinchin theorem, the power spectrum is the Fourier transform of the autocorrelation matrix of an image (**Figure 2-2**). Thus, the power spectrum succinctly describes second-order spatial correlations between pixels. Several studies have shown that the rotationally-averaged power spectrum (P) of a broad class of natural scenes decays with increasing spatial frequency (SF) according to the power law: $P \sim SF^{-\alpha}$, with α, known as the spectral slope taking values in the range 1.5 – 2 (Ruderman and Bialek, 1994). This power law relationship implies that spatial correlations between pixels are stronger at lower SFs than at higher SFs. For example, if an image contained no spatial correlations (white noise) then its power

spectrum would be flat and $\alpha = 0$ as a consequence. More importantly, the spectral

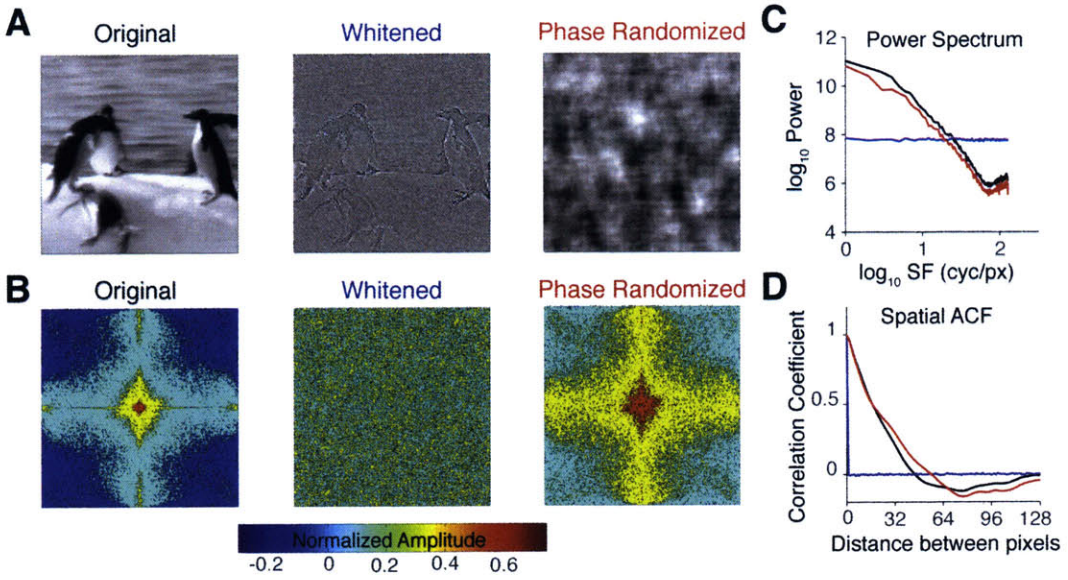

Figure 2-2. Perturbing the Fourier components alters the appearance of natural images. **(A)** Example of a natural image subjected to amplitude spectrum whitening (middle) or phase randomization (right). **(B)** Fourier amplitude spectra of these images, illustrating that phase randomization keeps the amplitude spectrum intact. **(C, D)** Quantification of Power spectrum (C) and spatial ACF (D) of these images. This illustrates that whitening the amplitude spectrum preserves edges but removes second-order spatial correlations between pixels.

slope can be used as parameter to describe spatial correlations in natural scenes.

Another important statistical feature is the phase spectrum. Since edges and line-like features arise from the phase congruency of different SF components, the phase spectrum contains information of salient features that define the structure of an image (Thomson and Foster, 1997). It has been shown that randomizing the phase of an image, preserves the power spectrum but destroys all salient visual information, making the image unrecognizable (Burr, 1980; McCotter et al., 2005). In contrast, removing pair-wise spatial correlations in an image by whitening (i.e. making $\alpha = 0$) preserves visual features (see example in **Figure 2-2**). This implies that information used for the identification of images resides in the phase spectrum (Wichmann and Gegenfurtner, 2010). Thus, it is generally agreed that the amplitude spectrum determines the overall appearance of an image, while the phase spectrum determines the structure of objects in the image. However, unlike the amplitude spectrum, the phase spectrum cannot be

parameterized. As a consequence, it has hard to relate changes in the phase spectrum to changes in neural response or behavior.

How is the amplitude and phase information in images used for scene recognition? Many theories of scene classification rely solely on information provided by the phase spectrum (Serre et al., 2007; Wichmann et al., 2006), suggesting that the amplitude spectrum has little to no significance. Indeed, psychophysical experiments have shown that human observers are unable to discriminate between phase randomized images (Párraga et al., 2005). Similarly, using natural images with different degrees of phase randomization, Wichmann found that disrupting image structure leads to a significant deficit in classification performance. These results are not surprising given that the phase spectrum contains all the salient structural information.

However, the fact that human observers are able to make rapid image classifications (Li et al., 2002; Thorpe et al., 1996) suggests that the visual system might be using other information to make broad generalizations before using detailed structural information contained in the phase spectrum. One possibility is that the visual system uses information contained in the amplitude spectrum to extract the gist of an image. In particular, the specific shape of the two-dimensional Fourier amplitude spectrum is strongly determined by the distributions of orientations and SFs in the image. For example, an image with more vertical and horizontal edges would have a more "diamond-like" amplitude spectrum since more power is concentrated in the cardinal orientations. Whereas an image with few edges would yield a "circle-like" edges as power is isotropic over the different orientations. In an influential study, Torralba and Oliva (2003) showed that different image categories have different spectral signatures due to their unique edge orientation and SF distributions. For example, man-made images contain a greater prevalence of horizontal and vertical contours while natural environments have a broader variation in spectral shapes (see Fig. 3 in (Torralba and Oliva, 2003)). This leads to the spectral-based categorization hypothesis, which states that the amplitude spectrum of an image conveys contextual information that can facilitate the recognition of salient visual features (Oliva and Torralba, 2007). In support of this hypothesis, numerous studies have shown that perturbing the amplitude

spectrum, either by normalization across different categories or by swapping between categories, significantly reduce classification accuracy (Gaspar and Rousselet, 2009; Joubert et al., 2009).

Taken together, these studies demonstrate that both amplitude and phase play important roles in image discrimination and recognition. However, there has still not been a systematic study looking at how perturbations to amplitude or phase influence cortical processing (see **Chapter 3**).

Cortical coding of natural scenes

The first demonstration that the statistics of natural movies can influence the efficiency of the neural code came from experiments by Vinje and Gallant. They found that stimulating the classical receptive field (CRF) of neurons in alert monkeys with a vignetted naturalistic movie, led to dense and correlated neuronal firing (Vinje and Gallant, 2000). In contrast, increasing the size aperture around the movie, such that it now stimulated the CRF and surrounding region (nCRF), increased response sparseness. This sparse coding strategy increases the information conveyed by each spike (Vinje and Gallant, 2002) and also functions to conserve metabolic energy (Laughlin, 2001). This allowed them to conclude that wide-field natural movies are more efficiently represented in V1 using sparse codes. Since then numerous studies have noted that natural movies sparse and temporally decorrelated responses, which are often not much stronger than spontaneous activity (Fiser et al., 2004; Onat et al., 2011).

Although the activity is weak, responses to natural scenes are reliable in the sense that repeatedly presenting the same movie results in very similar responses. For example, Haider and colleagues showed that during wide-field natural movie stimulation, neurons in cat V1 were not only sparser and but were more reliable between trials (**Figure 2-3B**). Interestingly, even sub-threshold membrane fluctuations, which reflect synaptic inputs into the neuron, were more reliable when stimulated with wide-field movies (Haider et al., 2010). Indeed, compared to RF-sized stimulation, wide-field presentation of natural movies activates a larger proportion of the network. Thus, the observed increase in sparseness and reliability of the response could be due to

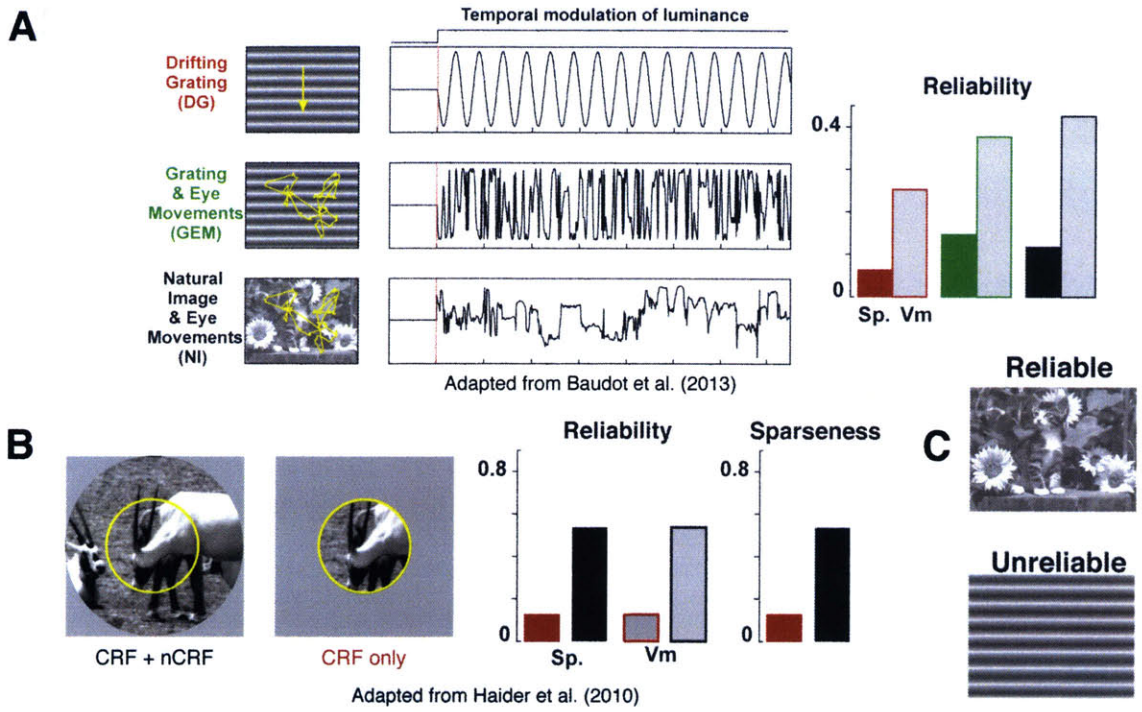

Figure 2-3. Spatiotemporally complex stimuli are more reliably processed in V1. (A) Data from Baudot et al., showing that NI are more reliably processed than DG, and that reliability scales with temporal complexity. **(B)** Data from Haider et al. showing that wide-field movies are more reliably processed than movies that only activate the receptive field of a neuron. **(C)** Summary.

"network effects" acting through local inhibitory interneurons (more detail in **Section 2.5**). However, the specific contribution of different interneurons remains unexplored. We explore some of these inhibitory mechanisms in **Section 2.5** and delve further in **Chapter 4**.

Another key contributor to response variability is the statistics of the visual stimulus. For example, a stimulus that is not optimally matched to a neuron's filter would elicit weaker signals than the ambient variability, and hence would lead to a much lower SNR. Natural scenes statistics are matched to the filters found in visual cortex (Olshausen and Field, 1996), implying that the visual cortex is optimized to efficiently process natural scenes. Also, these filters are not fixed and change linearity depending on the statistics of the stimulus (Fournier et al., 2011). Surprisingly, a study in ferret visual cortex found that flashed natural images failed to evoke sparse or reliable responses (Tolhurst et al., 2009). This is likely due to the fact that these stimuli lack the temporal correlations between pixels that are found naturally in movies, or that are

imposed by saccadic eye movements. To more formally relate response variability relate to the statistics of the visual stimulus, Baudot and colleagues performed detailed signal and noise analysis of intracellular recordings from anesthetized cats that were presented with stimuli of different spatiotemporal statistics (**Figure 2-3A**). They found that natural movies were more reliably and more precisely processed than drifting gratings (DG) or dense white noise (DN). Statistically, DN contains a uniform distribution of spatial (SF) and temporal frequencies (TF), whereas DGs contain only one spatial and one temporal frequency. This is in stark contrast to natural scenes, which contain a power law distribution of both spatial and temporal frequencies (described above). Interestingly, Baudot found that imposing a naturalistic temporal structure to gratings (GEM), through a model of eye movements, improved increased SNR (Baudot et al., 2013). Specifically, GEM had a lower noise and a comparable level of signal to DG. The sub-threshold noise level was however still higher for GEM than natural movies. Taken together, these studies show that response variability in V1 changes depending on both the spatial and temporal statistics of the visual stimulus.

Importantly, Baudot's study is an *in vivo* analog of Mainen and Sejnowski's experiments. Both studies show that stimuli with temporally complex properties (GEM, natural movies or white noise current, see examples in **Figure 2-3A**) are more reliably processed than stimuli with more regular temporal properties (DG or DC current). Baudot and colleagues argued that broadband stimuli elicit much more complex repertoire of neural interactions, which in turn constrains the network in a more reliable coding regime (Borst and Theunissen, 1999). Thus, the conditions required for reliable coding are: **(1)** stimuli with broadband spatial and temporal statistics; and **(2)** network-dependent mechanisms which reduce shared intrinsic noise.

In this thesis, I will investigate both these conditions in mouse visual cortex. In **Chapter 3**, I show that changing spatial correlations in a natural movie by perturbing the amplitude spectrum, changes neuronal reliability and restructures correlations between neurons. In **Chapter 6**, I show using a novel head-fixed behavior that mice rely on phase information to discriminate between different natural scenes. Also in **Chapter 4**, I will investigate the inhibitory mechanisms that are required for reliable coding.

2.4 The structure and organization of rodent visual cortex

Most studies of natural scene processing described in the pervious section have used cats or monkeys. However in order to fully understand natural scene processing, it is crucial to elucidate the various cellular and network mechanisms responsible for reliable coding. Mice are rapidly becoming the model system of choice in systems neuroscience research due to their genetic malleability. In particular, modern genetic tools allow researchers to study the function of different cell types in the visual system and also perturb their activity in a reversible manner. These studies have the promise of allowing us to link defined categories of neurons with different aspects of visual perception. These unparalleled advantages outweighs the obvious disadvantage of using mice in vision research – the fact that mice have several orders of magnitude lower spatial resolution than the macaque monkeys (Prusky and Douglas, 2004). In this section, I review what is currently known about the structure and function of the mouse visual system, focusing primarily on the visual cortex.

Orientation processing in mouse V1

In mice, as in all other mammals, visual information is first processed in the retina. Interestingly, the mouse retina contains primarily rods and thereby is specialized for vision under low light conditions (Jeon et al., 1998). Additionally, mice also lack a fovea, which means that mice primarily have very low visual acuity when compared to macaques (Huberman and Niell, 2011). Without a fovea, mice are only able to detect large objects at a distance and smaller objects at closer range. Nevertheless, studies have shown that the photoreceptor mosaic in mouse retina is organized to allow mice to efficiently sample the visual scene. Interestingly, a recent study showed that genetically modifying the mouse retina with human photoreceptors was sufficient to introduce color vision in mice (Jacobs et al., 2007). Remarkably, these mice were able to discriminate between red and green colored stimuli. This suggests that the mouse visual hardware is actually well equipped to process complex visual inputs, with the main limitation being the lack of photoreceptors in its retina (Huberman and Niell, 2011).

From the retina, visual information flows to numerous subcortical targets. These include the accessory optic system, which controls saccadic eye movements, the superior colliculus, which works as a looming detector to recognize predatory-like stimuli (Yilmaz and Meister, 2013), and the dorsal lateral geniculate nucleus (LGN), which then feeds information into the cortex. Overall, the mouse LGN has been found to be analogous to cat or monkeys, with most neurons showing the classical ON-OFF center-surround receptive fields (Grubb and Thompson, 2003). Additionally, the mouse LGN also contains a substantial number of orientation and direction selective neurons (Marshel et al., 2012). A recent study, used two-photon calcium imaging, to show that LGN axons in L4 of visual cortex carry both orientation and directionally-tuned information (Sun et al., 2016). This result suggests that neurons in V1 might inherit their orientation tuning properties directly from thalamic afferents, which goes counter to Hubel and Wiesel's original proposal. Indeed, another study found that L4 of visual cortex functions primarily to alter the gain of tuned responses from the LGN (Lien and Scanziani, 2013). Together, this work implies that extraction of more complex visual features, such as orientated line segments, begins as early as the LGN in the mouse.

Similar to other mammals, the mouse primary visual cortex (and neocortex, in general) is organized into six distinct layers. Expectedly, the mouse V1 is retinotopically organized, with neurons sampling similar parts of the visual scene located close to each other (Wang and Burkhalter, 2007). Classically, V1 is viewed as a bank of nonlinear filters that function to extract oriented edge information from a visual scene. In cats and monkeys, receptive field mapping has indeed shown that V1 receptive fields (RFs) are oriented Gabor functions (Ringach et al., 2002). Surprisingly, very similar RF properties have also been observed in mouse V1 (Dräger, 1975; Gao et al., 2010; Niell and Stryker, 2008). In particular, neurons in mouse V1 show both linear and nonlinear responses to drifting gratings, consistent with the classical definition of simple and complex cells respectively (Skottun et al., 1991). Specifically, researchers have observed a marked difference in RF linearity depending on cortical depth. Neurons in L2/3 and L4 are predominantly simple cells, whereas neurons in L5 are primarily complex (Niell, 2015). Moreover, inhibitory interneurons in mouse V1 also have

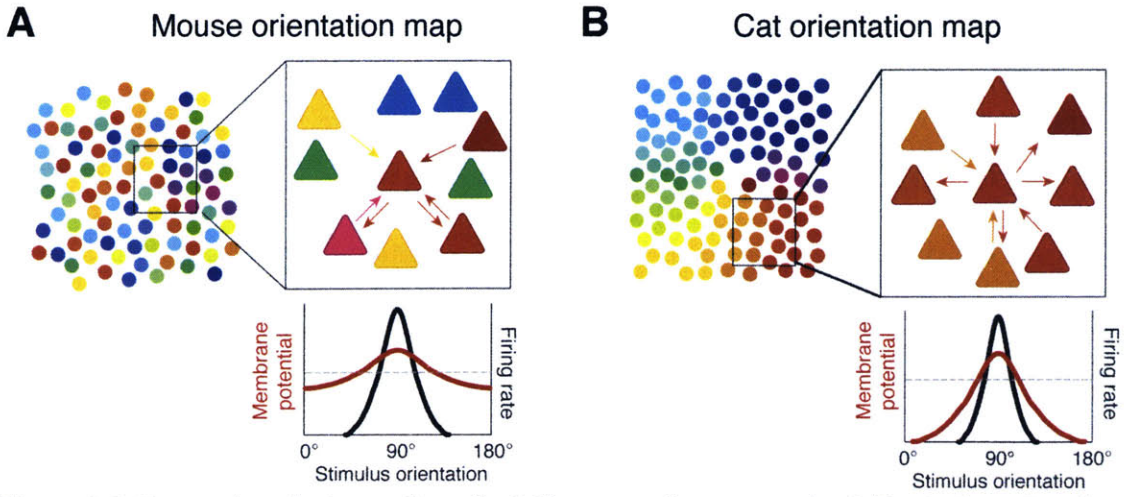

Figure 2-4. Comparison between the orientation map of mouse and cat. Due to the salt-and-pepper organization in mice (A), the summed membrane potential is broader (A, bottom). Despite this, mice are able to generate sharply tuned cells. By contrast, the neurons in the cat (B) get inputs from similarly tuned cells. Adapted from Harris and Mrsic-Flogel (2013, Nature).

complex-like RFs (Liu et al., 2009; Ma et al., 2010). Thus, RF structure is not dramatically different between "more visual mammals" and mice. However, the most striking difference is in the spatial acuity. Mice have a median preferred spatial frequency of 0.04 cycles/degree (cpd), whereas cats prefer 0.9cpd and monkeys prefer >1.5cpd (Van Hooser, 2007). Interestingly, despite this lower spatial acuity, the range of orientation selectivity in the mouse is not very different from cats or monkeys (Niell and Stryker, 2008; Tan et al., 2011). Although, there is a slight bias towards the cardinal orientations in mouse V1 (Roth et al., 2012). Together, these results suggest that mice are performing similar computations as cats/monkeys on the visual scene, albeit at a much lower spatial resolution. Ethologically, this lack of spatial acuity makes sense as mice use their vision primarily to avoid predators, which appear as large imposing shapes (Choi and Kim, 2010). In contrast, cats use their vision to hunt prey, which are often well camouflaged.

Although the mouse V1 has a clear topographic map of visual space (retinotopy), it lacks organized orientation or spatial frequency maps. In many higher species, including members of the rodent family (tree shrew), preferred orientation is neatly organized into stereotypical patterns of pinwheels (Nauhaus and Nielsen, 2014) (**Figure 2-4**). Instead, mouse/rat V1 has a "salt-and-pepper" organization in which neighboring

neurons prefer very different orientations (Andermann et al., 2013; Ohki and Reid, 2007). One possible explanation for this disorganized structure is the lack of segregation of ON-center and OFF-center neurons in the LGN (Piscopo et al., 2013). One advantage of having an organized map is that it helps to minimize the wiring length between neurons (Koulakov and Chklovskii, 2001). Thus, since the mouse visual cortex is so small, there might not be a need for wiring length minimization.

Importantly, the fact that neurons in mouse V1 have sharp orientation tuning and structured receptive fields, suggest that the presence of well-structured orientation maps is not necessary to generate selectivity. In fact, several studies have now demonstrated that selectivity in mouse V1 arises due to a very precise organization at the synaptic level. For example, orientation selective neurons tend to receive more similarly tuned synaptic inputs (Chen et al., 2013; Jia et al., 2010). Using functional *in vivo* imaging followed by *in vitro* patch recoding, Mrsic-Flogel's laboratory demonstrated a remarkable degree of functional organization at the synaptic level. Neurons that shared similar orientation selectivity and similar RF properties had a higher probability of being synaptically connected to each other (Cossell et al., 2015; Ko et al., 2011). These studies clearly show that, even though there is no mesoscopic organization, neurons that share similar orientation are synaptically coupled to each other. Specifically, they show that in mouse visual cortex, very specific neural ensembles can be recruited to process visual scenes (Harris and Mrsic-Flogel, 2013).

Organization of higher visual areas

After V1, further visual processing occurs in numerous extra-striate areas, each working to extract a variety of higher order visual features, with the overall aim of object recognition and visual perception (Orban, 2008). These extra-striate areas (or higher visual areas, HVAs) are hierarchically organized and are often broadly categorized into tow main pathways – the dorsal stream ("where" pathway) and the ventral stream ("what" pathway) (Felleman and Van Essen, 1992) (**Figure 2-5A**). Neurons along the dorsal stream process motion and can represent the location of an object irrespective of its identity. On the other hand, neurons along the ventral stream represent objects and

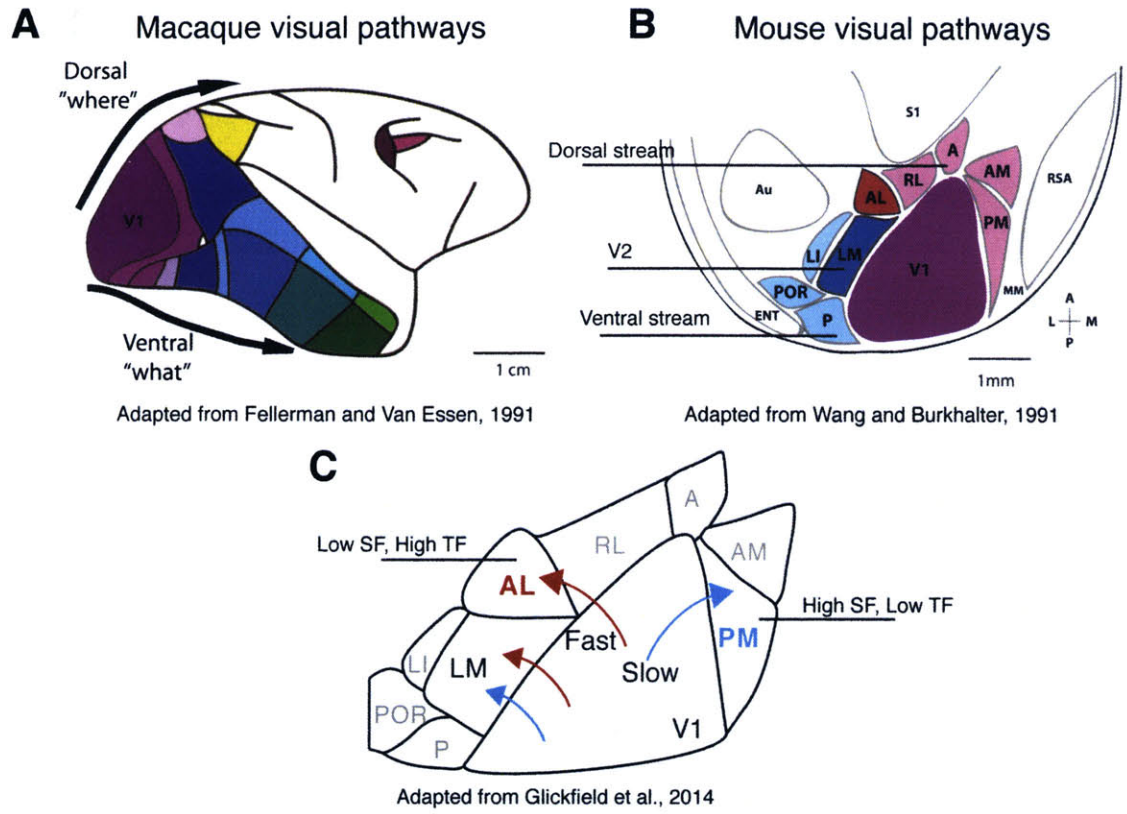

Figure 2-5. Structure and organization of extrastriate cortex in mouse. (A) Organization of dorsal and ventral streams in macaque. **(B)** Map showing location of mouse extrastriate areas. This map is color coded to show areas believed to correspond to the dorsal and ventral stream. **(C)** Segregation of visual information from V1 to AL and PM. Modified from (Glickfeld et al., 2014).

are invariant to viewing angle. Thus, this form of hierarchical processing is critical for visual perception as it ensures the parallelization of different computations, which in turn allows the visual scene to be parsed rapidly. Given that mouse V1 is analogous to monkey V1, to some extent, in the computations that it performs, it currently remains an open question if mice also have specialized HVAs for recognizing visual objects.

Using novel anterograde tracing methods, Wang and Burkhalter (2007) identified nine distinct cortical areas that received projections from V1. These nine areas, shown in **Figure 2-5B**, were also retinoropically organized, suggesting that these HVAs also contained a complete representation of visual space (Wang et al., 2011). More recent studies have identified up to 11 different HVAs (Garrett et al., 2014). Further, using a combination intrinsic imaging technology to identify retinotopically organized areas and cellular-resolution two-photon calcium imaging; the cellular properties of these HVAs

were recently characterized (Andermann et al., 2011; Marshel et al., 2011). In particular, they characterized orientation/direction selectivity, spatial frequency and temporal frequency tuning and found a striking dichotomy within these different areas. Area AL preferred stimuli with lower spatial frequency and higher temporal frequency (i.e. faster speeds), whereas Area PM preferred stimuli with higher spatial frequency and lower temporal spatial frequency (i.e. slower speeds, see **Figure 2-5C**). This was the first demonstration of a visual processing dichotomy in mice. To elucidate how these visual areas inherited these tuning properties, Glickfeld and colleagues used calcium imaging to study the axons of V1 neurons in both AL an PM. Interestingly, this study found that V1→AL projections and V1→PM projections also had similar tuning properties to their target sites (Glickfeld et al., 2013a). Together, these studies demonstrate the existence of specialized information processing pathways in the mouse visual system.

Given the precise routing and segregation of information from V1 into the HVA, it is likely that mice also have dorsal and ventral streams. It is possible that Area AL forms part of the dorsal stream, due to their preference to higher speeds. By responding to faster stimuli, Area AL can compute optic flow, or the movement of objects in the visual field relative to the motion of the mouse. This optic flow information can then by used by the mouse to compute running speed for navigation, or identify visual stimuli as it is running. In contrast, because Area PM prefers high spatial frequency stimuli, is likely to be involved in object recognition. This is because fine structural details are contained in the high SF end of the power spectrum (see **Section 2.3**). Interestingly, analysis of cortico-cortical projections revealed that both AL and PM project to the motor, parietal and limbic systems, whereas LM and P project to the hippocampal areas (Wang et al., 2012). This it is possible that PM is also part of the ventral stream. In fact, there is still no conclusive evidence that links these HVAs to the dorsal or the ventral streams.

Visual psychophysics in mice

So mice can process edges like monkeys and also have whole hierarchy for processing complex visual features. But can mice actually see? In other words, can mice make decisions using visual information alone? It is known that rats are able perform object

recognition tasks (Zoccolan et al., 2009), and in fact guide their decisions based on specific visual features (Zoccolan, 2015). A group was even able to train rats to categorize natural images into ones containing an image of a rat and into other objects (Vinken et al., 2014). This daunting task took close to 6 months for rats to master. Mice are also able to use their visual systems to navigate complex virtual environments (Harvey et al., 2012). Interestingly, a recent study showed that mice performed better at navigation in a brightly lit room compared to a dark room (Sofroniew et al., 2014), suggesting that mice supplement somatosensory information with visual information.

In more recent years, numerous head-fixed and non-head fixed paradigms for visual psychophysics have been developed. For example Busse and colleagues were able to train freely moving mice to perform contrast detection tasks and were able to obtain reliable psychometric functions (Busse et al., 2011). Similarly, mice were able to learn a head-fixed version of this task within 3-4 weeks (Glickfeld et al., 2013b). Apart from contrast, mice can also be trained to discriminate between orientations (Andermann et al., 2010; Lee et al., 2012). Surprisingly, despite their low spatial acuity, mice are able to discriminate between gratings as fine as 15^{o} apart. In **Chapter 6**, I will show that mice are also able to discriminate between complex natural scenes. Taken together, these studies suggest that mice are able to use their vision for goal directed behaviors.

Natural scene processing in mouse visual cortex

It is now becoming clear that computations performed by neurons in the mouse visual cortex are very similar to those in more visual mammals. As such, mice are the ideal model system to study the neural mechanisms of natural scene processing. Unfortunately, there have only been a handful of studies that have looked at how natural scenes are processed in mouse visual cortex. Most commonly, the studies that I have described above relied on simple stimuli, such as gratings, to characterize the tuning properties of neurons. The main advantage of using these stimuli is that they are easily parameterized, which means that neuronal activity can be easily related to one stimulus feature. Natural scenes, on the other hand, are broad-band stimuli, containing an entire

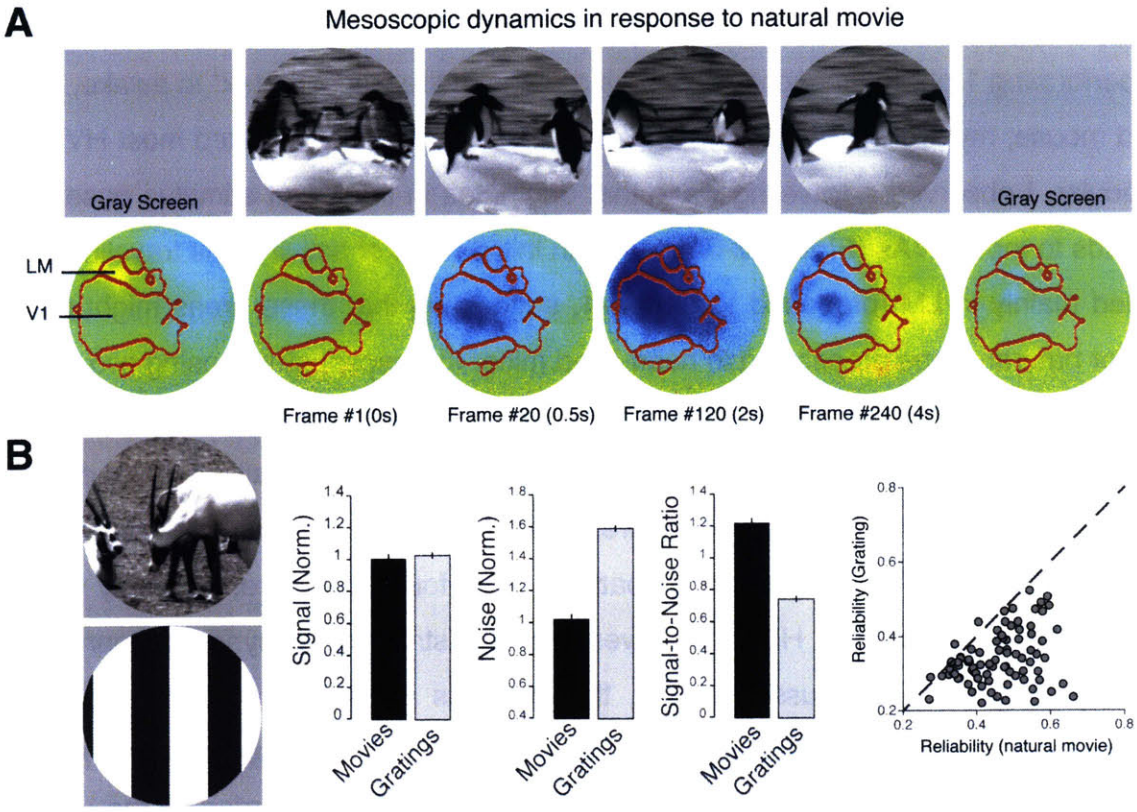

Figure 2-6. Response properties of mouse visual cortex to natural movies. (A) Mesoscale dynamics of entire visual cortex (V1 and LM highlighted) to a natural movie. A spreading wave of activity that originates and ends in V1 can be seen following stimulus onset. Borders of visual areas are shown in red. Data from 1 mouse **(B)** Natural movies evoke much more reliable responses than gratings. Unpublished data from 3 mice.

spectrum of spatial frequencies and orientations. As a consequence are likely to recruit neurons in a more complex way. Studies in monkey and cat have revealed that natural scenes are processed with sparse and temporally reliable codes (see Section 2.3). In particular, RF models based on responses to simple stimuli fail to predict response to natural scenes (David et al., 2004). This suggests that the RF properties of neurons are optimized to process natural scenes (Olshausen and Field, 2004). Thus, given the similarities in RF properties between mice and monkeys, natural scenes should also be similarly processed in mice.

Expectedly neurons in mouse visual cortex respond robustly to natural movies. By performing 1-photon mesoscopic imaging of the entire visual cortex[1] in awake, head-fixed mouse, my preliminary data revealed that natural movies activated most HVAs. In particular, I observed a traveling wave that originates in the V1 at stimulus onset, and spreads to most HVAs (**Figure 2-6A**). After an initial wave of activity, this traveling wave settled mainly in LM (proposed V2) and V1, suggesting that these areas might be the loci for natural scene processing in head-fixed mice. In a recent study, using two-photon calcium imaging in anesthetized mice, Kampa and colleagues found that natural scenes recruited very specific ensembles of neurons in mouse V1. Specifically, a linear decoder could successfully predict the movie presented based on neuronal activity patterns (Kampa et al., 2011). Similar activation patterns were found in awake, passively viewing mice (Miller et al., 2014). However, given the fine structure of synaptic connectivity between neurons as discussed above, this result is not unexpected. Nevertheless, these studies show that there exists specialized circuitry for processing complex visual stimuli in mice. It would be interesting to see what information is carried by the V1→LM projections and how natural scenes are processed in the various HVAs.

In addition, neurons in mouse V1 are very sensitive to the statistics of the visual stimuli. My preliminary data revealed that, although natural scenes evoked responses with similar magnitude to gratings (at preferred orientation), the responses to natural scenes are much less variable and are more reliable between trials (**Figure 2-6B**). In **Chapter 3**, I will explore the network mechanisms behind why natural scenes are more reliably processed. Other studies have also investigated how neurons in mouse V1 encode natural scene statistics. For example, randomizing the phase spectrum of natural movies increased the variability, decreased sparseness and increased correlations between neurons in V1 (Froudarakis et al., 2014). This study also demonstrated that phase randomized movies were much less discriminable from the original movie based on population activity patterns. Similarly, another study also found

[1] A 5mm craniotomy was opened over the entire visual cortex of mice expressing GCaMP6f. The mesoscope was designed by Dr. Ming Hu and Dr. Michael Goard and allows epifluoresnece imaging of large cortical regions. This work is unpublished and will be presented at the 2016 SfN meeting.

that randomizing the phase of the surround, but keeping the center intact, decreased response sparseness in V1 pyramidal neurons (Pecka et al., 2014). Importantly, both studies show that neurons in mouse V1 are capable of sparse coding (Vinje and Gallant, 2000). Physically, the phase spectrum contains higher order structural information, such as edges and contours, which define salient features in the image. Randomizing the phase destroys these salient features but keeps the power spectrum intact. Together, both this work indicates that it is the spatial structure of natural images that drives sparse coding in the visual cortex. In **Chapter 3**, I will show that the power spectrum of an image also plays an important role in natural scene processing, primarily by altering the spatial structure of inter-neuron correlations.

Thus, the work that I have summarized in this chapter provides a compelling argument that mice are an ideal model system for studying visual processing. While a lot of effort has gone into charactering the response properties of neurons, very little information still exists on how natural scenes or more complex stimuli are processed. My thesis bridges this gap. By using a combination of genetic tools to perturb and record neural activity, I will relate neural circuit function with specific aspects of natural scene processing. In doing so, this thesis advances our knowledge of the computations performed by the visual system.

2.5 The role of cortical interneurons in modulating signal and noise

In the previous section, we established that natural scenes are processed with sparse and temporally reliable codes. Sparser and more reliable responses means that neurons are able to efficiently integrate noisy inputs and produce an output that has fewer and more precise action potentials. It is very likely that these transformations are sculpted by GABAergic neurons in the visual cortex. In this section, I describe the computations performed by the different cortical interneurons.

GABAergic neurons constitute 20-30% of neurons in the mammalian neocortex. These inhibitory neurons (INs) show a huge diversity in their morphology, electrophysiological properties and chemical composition (DeFelipe et al., 2013; Kepecs and Fishell, 2014; Markram et al., 2004) . Based on chemical markers, INs can be broadly categorized into three non-overlapping subclasses of cells: (1) Parvalbumin (PV) expressing neurons that make up approximately 36% of all INs, (2) Somatostaitn (SST) expressing neurons that make up 30% of all INs and (3) vasoactive intestinal peptide (VIP) or serotonin-receptor ($5HT_{3A}$-R) expressing neurons, which make up 17% of all INs (Rudy et al., 2011). Recent molecular profiling studies have shown that each subclass can be further divided into several additional cell types (Tasic et al., 2016), each with a unique genetic and morphological profile. For instance, calbidin-expressing Martinotti cells are a subclass of SST neurons that reside mainly in the superficial layers and provide robust inhibition onto the distal dendrites of pyramidal neurons (Silberberg and Markram, 2007). PV neurons also can be subdivided into basket cells, which target the perisomatic compartment (including proximal dendrites, (Klausberger and Somogyi, 2008)), and chandelier cells, which target mainly the axon initial segment (Taniguchi et al., 2013). Based on these connectivity patterns, it is likely that the computations performed by pyramidal neurons are sculpted by distinct inhibitory mechanisms – one providing dendritic inhibition to control input, and the other providing somatic inhibition to control the output (Bloss et al., 2016) (**Figure 2-7A**). Thus in order to fully understand the computations performed by pyramidal neurons, it is critical to elucidate the function of these different interneuron subtypes.

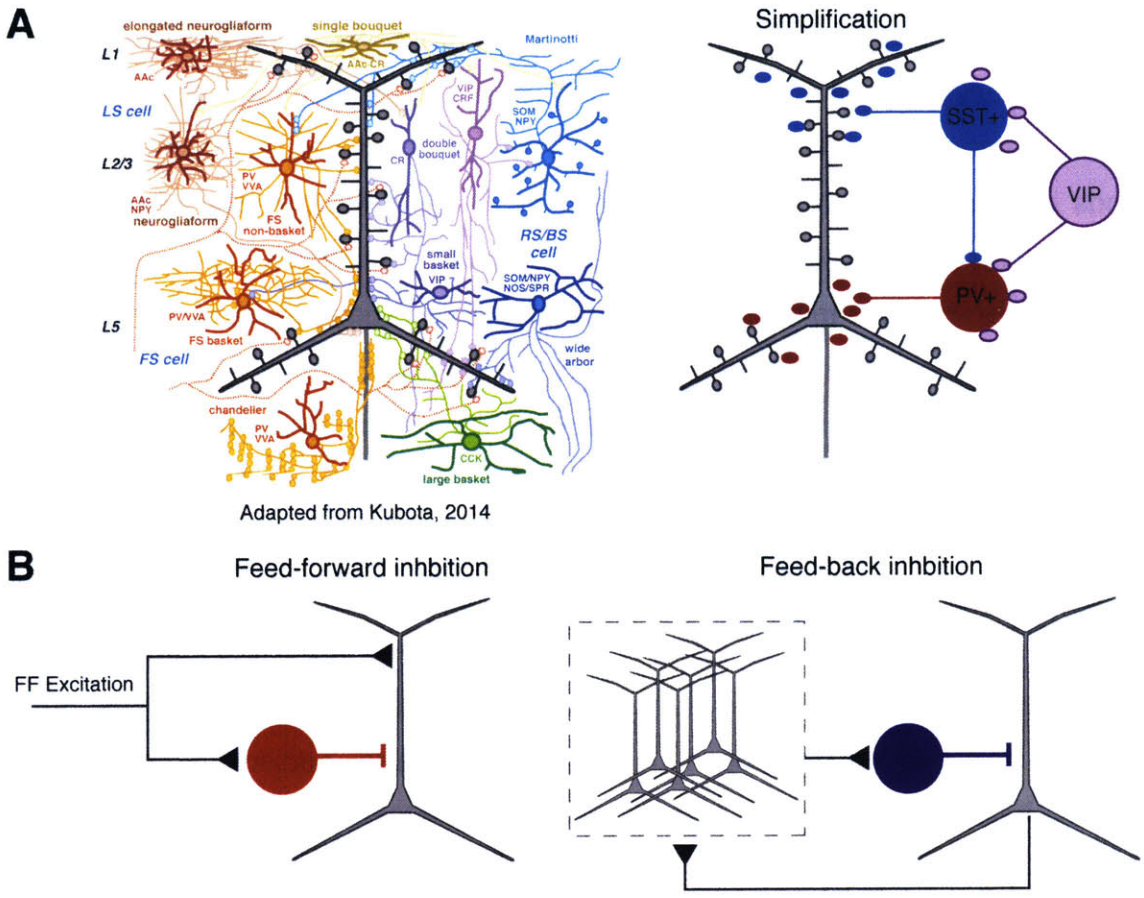

Figure 2-7. Overview of inhibitory microcircuits in the cortex. (A) Left, cartoon showing the complexity of inhibitory neuron subtypes. Right, a simplified version of this circuit showing the three main interneuron subtypes. Adapted from (Kubota, 2014). **(B)** The two main inhibitory circuit motifs – feed-forward and feed-back inhibition

Inhibitory circuit motifs

Synaptic inhibition is critical for information processing within cortical circuits (Isaacson and Scanziani, 2011). Despite differences in the anatomy, cortical area or the functions of neuronal circuits, two basics motifs of inhibition repeatedly emerge: feed-forward (FF) and feedback inhibition (FB) (**Figure 2.7B**). In FF inhibition, excitation originating from a different brain area or layer (e.g. thalamus or L4) impinges on INs before targeting pyramidal neurons. In other words, the source of inhibition is outside the pool of neurons targeted by the IN. On the other hand, in FB inhibition, interneurons receive recurrent excitation from the same population of neurons that they are inhibiting. Both motifs are able to powerfully alter spiking in their postsynaptic targets. In particular, because FF

inhibition is controlled by upstream neurons it permits a much faster modulation of postsynaptic activity. Thus, the FF inhibition works to balance the strength of FF excitation to prevent runaway excitation. FB inhibition, in contrast, is much slower as it depends on the strength of recurrent activity. As a result, FB inhibition often appears much later than FF inhibition (i.e. late onset). Because FB inhibition controls recurrent dynamics, it can be used to synchronize network activity. Together, these two inhibitory motifs represent the fundamental building blocks of cortical computations.

Inhibition is inseparable from excitation in the cortex. Numerous studies have now noted that excitation and inhibition is tightly coupled, giving rise to the notion of E-I balance. Specifically, in all sensory modalities, stimuli that evoke the largest mean excitation also evoked the strongest inhibition (Anderson et al., 2000; Denève and Machens, 2016; Monier et al., 2003; Wilent and Contreras, 2005). By measuring synaptic conductance, Okun and Lampl (2008) observed that inhibition was tightly temporally correlated with excitation on a trial-to-trial basis. Additionally, researchers have observed that inhibition is delayed relative to excitation – a sharp increase in excitation is normally followed a few milliseconds later by a sharp increase in inhibition (Tan et al., 2011; Wehr and Zador, 2003). Mechanistically, this short delay between inhibition and excitation is believed to open a brief "window of opportunity" for the neuron to integrate inputs and generate spikes (Haider and McCormick, 2009). Based on these observations, it has been proposed that cortical FF inhibition functions to prevent runaway excitation, and thus homeostatically regulates network activity (Okun and Lampl, 2008). In agreement with this, disrupted cortical inhibition, in particular PV inhibition, has been shown to cause seizures and cognitive impairment seen in neuropsychiatric diseases, such as Schizophrenia (Lewis et al., 2005) and Autism (see discussion in **Chapter 5**). Interestingly, a recent study showed that inhibitory currents dominate excitation in the awake cortex, and are much weaker in the anesthetized state (Haider et al., 2013). Thus, taken together these studies show that strong and temporally balanced inhibition is required to quench excitatory drive in the cortex. In doing so, inhibition helps to sharpen stimulus selectivity (Isaacson and Scanziani, 2011).

Computations performed by PV neurons

Which form of inhibition do PV neurons provide? Studies in the hippocampus and the cortex, have demonstrated that fast-spiking PV neurons are critical in mediating FF inhibition (Pouille and Scanziani, 2001). In particular, PV neurons are recruited with short latency following thalamocortical FF excitation (Gabernet et al., 2005; Tan et al., 2008) or visual stimulation (Ma et al., 2010).

By providing FF inhibition, PV neurons are prime candidates for maintaining E-I balance in the cortex. PV neurons are the most numerous IN subtype and consequentially provide the bulk of cortical inhibition. Also, because they target the somatic compartment of pyramidal cells, PV neurons are poised to powerfully alter the synaptic integration window. This unique feature allows PV neurons to control the temporal precision of spiking by filtering out synaptic inputs falling outside this integration window ((Gabernet et al., 2005), see also discussion in **Chapter 4**). For this reason PV neurons are believed to control pyramidal neuron reliability ((Zhu et al., 2015). However, in **Chapter 4**, I argue that the circuit for reliable coding is not as simple as the one proposed by Zhu and colleagues.

A recent also found that the strength of PV inhibition onto L4 pyramidal neurons in visual cortex is precisely tuned to match the strength of thalamic excitation (Xue et al., 2014). This finding implies that PV neurons homeostatically adjust their activity to balance pyramidal cell activity, and in doing so are able to equalize E-I ratios across the network. Plasticity of PV synapses onto pyramidal neurons could also play a crucial role in preserving E-I balance and ensuring that sensory information is processed with high fidelity (Froemke, 2015).

By scaling inhibition to match the strength of input excitation, PV neurons can alter the dynamic range of pyramidal neurons through a process known as divisive normalization. Specifically, because PV synapses are close to the site of action potential initiation, their activation changes somatic conductance resulting in a divisive change in the neuron's response (Chance and Abbott, 2000). This divisive inhibition changes the slope of the neuronal input-output function (**Figure 2-8**), and consequentially provides a gain control mechanism that regulates the normalization of activity in neuronal circuits

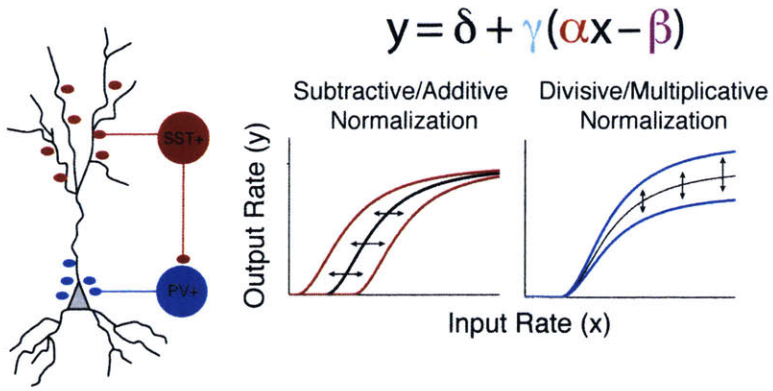

Figure 2-8. Inhibition can alter neuronal input-output functions. By changing the gain (gamma), PV neurons are can prevent neuronal firing to weaker inputs. By changing the threshold (alpha, or beta) SST neurons can change the threshold required to elicit a response.

(Pouille et al., 2009; Silver, 2010). In particular, divisive normalization is an essential component of neural computation that scales a response by a weighted sum of a pool of neurons in the network (Carandini and Heeger, 2012). Divisive normalization allows neurons to respond without saturation to a wide range of stimuli, and functions to remove redundancies between neurons. Recent studies have demonstrated that optically activating PV neurons sharpens the orientation tuning of pyramidal neurons, by altering the gain of the response (Atallah et al., 2012; Wilson et al., 2012). Additionally, increasing PV inhibition increases the functional coupling between neurons in auditory cortex (Hamilton et al., 2013). However, other studies have argued that PV activation can result in a more mixed effect on neuronal tuning (El-Boustani and Sur, 2014; Seybold et al., 2015), suggesting the effect of PV neurons on somatic conductance also depends on other network factors. Nevertheless, it cannot be argued that PV activation improves feature selectivity across numerous sensory modalities.

Taken together, these studies show that PV neurons are prime candidates for providing FF inhibition. In addition, PV neurons are densely connected to pyramidal neurons and are broadly tuned to sensory cues (Fino and Yuste, 2011; Kerlin et al., 2010; Packer et al., 2013; Sohya et al., 2007). Hence, PV neurons provide strong, blanket-like inhibition in the cortex (Karnani et al., 2014). In doing so, PV neurons can stabilize the intrinsically unstable dynamics of the cortex. This idea, however, has been challenged by the discovery of orientation selective PV neurons in mouse visual cortex (Runyan et al., 2010), suggesting that PV neurons can provide structured inhibition onto their targets. One possibility is that co-tuned PV and pyramidal neurons form functional

cellular assemblies (Harris and Mrsic-Flogel, 2013; Hofer et al., 2011). Within these assemblies, tuned inhibition from PV neurons functions to equalize E-I balance in an input dependent manner. However, more experiments are required to elucidate the connectivity between PV and pyramidal neurons, and how this contributes to stimulus selectivity.

Computations performed by SST neurons

In stark contrast to PV neurons, the dendrite targeting SST neurons provide slower FB inhibition (Pouille and Scanziani, 2004). In particular, SST neurons are activated after a substantial delay (Ma et al., 2010; Tan et al., 2008), suggesting that they are driven to fire only after recurrent network activity reaches a sufficient level. Indeed, rabies-mediated circuit tracing has revealed that SST neurons receive mostly short-ranged intra-cortical inputs (Wall et al., 2016), mainly from recurrently connected pyramidal neurons (Adesnik et al., 2012). By targeting the proximal dendrites of pyramidal neurons, activation of SST neurons produce a change in charge within the dendrites resulting in shunting inhibition (Holt and Koch, 1997; Vu and Krasne, 1992). This shunting inhibition results in an offset of the neuronal input-output relationship, and for this reason is commonly referred to as subtractive inhibition (**Figure 2-8**). Subtractively scaling the input-output function of a neuron raises the threshold for sensory input to trigger action potentials (Silver, 2010). One consequence of this is a sparsening of neuronal activity and a corresponding increase in stimulus discriminability (Sturgill and Isaacson, 2015; Wilson et al., 2012).

Numerous recent studies have linked SST neurons to surround suppression in the visual cortex (Adesnik et al., 2012; Nienborg et al., 2013). Surround suppression is the phenomenon whereby stimuli located outside the classical receptive field (CRF) of a neuron suppress responses to stimuli located within the CRF. By performing intracellular recordings in anesthetized cats, Haider and colleagues observed that that increasing the size of a natural movie such that it stimulated both the CRF and surround increased the amplitudes of IPSPs without increasing EPSPs (Haider et al., 2010). These results suggest that surround suppression occurs via a sustained increase in

intracortical inhibition. Using a similar technique, but with drifting gratings instead, Ozeki and colleagues observed a transient increase in inhibition followed by a coupled decrease in both excitation and inhibition (Ozeki et al., 2009). Thus, in the Ozeki model, surround suppression is the consequence of a net decrease in both excitation and inhibition. Despite their differences, both studies conclusively show that surround suppression is the by-product of an increase (either sustained or transient) in intra-cortical inhibition. However, neither study identified the source of inhibition. By leveraging the genetic tractability of mice, Adesnik et al. (2012) found that SST neurons increased their firing rate to increasing stimulus size and received strong inputs mainly from recurrently connected L2/3 pyramidal neurons. In contrast, PV neurons, similar to pyramidal neurons, were also suppressed by larger stimuli. Optically inhibiting SST neurons also prevented surround suppression in pyramidal neurons. Together, these results suggest that SST neurons are prime candidates for enforcing surround suppression. In particular, larger visual stimuli recruit a larger fraction of the network, which in turn drives an increase in SST activity. Increased SST inhibition then leads to a suppression of pyramidal neurons, which have receptive fields in the center. Thus, by providing stimulus-dependent inhibition, SST neurons can function to sparsen neuronal responses and reduce redundancy between neurons. In doing so, surround suppression maximizes the amount of information conveyed per spike (Vinje and Gallant, 2002; Pecka et al., 2014), and can improve the efficiency natural scene coding.

Rather than simply acting as a brake on spiking activity, SST neurons are poised to sculpt neuronal computations at a more subcellular level by inhibiting the apical and distal dendrites of pyramidal neurons. Nonlinear dendritic mechanisms, such as dendritic conductance, allow neurons to optimally integrate synaptic inputs from diverse sources (Spruston, 2008). In particular, it is believe that related synapses cluster together within the dendritic arbor forming functionally distinct subunits. Thus, it is likely that one function of dendritic inhibition could be to block inputs from these subunits from being integrated at the soma (Gidon and Segev, 2012). In agreement with this idea, a recent study found that SST neurons indeed provide compartmentalized dendritic inhibition (Chiu et al., 2013). In doing so, SST neurons are able to locally transform

integration and also control synaptic plasticity. In addition, SST neurons can also control synaptic integration by changing active dendritic nonlinearities (Lovett-Barron et al., 2012). A recent study in the hippocampus demonstrated that inhibition from SST neurons changes the gain of the dendritic input-output function by altering NMDAR-dependent active processes. As a result, SST neurons can control how pyramidal neurons integrate both sensory information and top-down inputs (Murayama et al., 2009; Palmer et al., 2012). With respect to neuronal computations, increasing dendritic inhibition has been shown to increase the reliability of spiking in the barrel cortex (Egger et al., 2015). Thus, using precisely placed synapses in the distal dendritic arbor, SST neurons are in perfect position to control which inputs get integrated by the soma. Consequentially, SST neurons can filter out noisy synaptic inputs to ensure robust sensory-driven activity.

Disinhbitory circuits: inhibition of inhibitory neurons

So far we have shown that cortical inhibition can sculpt neural computations through a domain specific division of labor – with PV neurons providing strong, but broadly tuned FF inhibition onto the soma and SST neurons providing FB inhibition onto specific compartments of the dendritic arbor. Do these separate channels of inhibition interact with each other? Studies have shown that inhibition can be dynamically routed along the somato-dendritic axis. For example in CA1 pyramidal neurons, it has been shown that inhibition shifts from soma to the dendrites depending on the firing rate of the pyramidal neuron (Pouille and Scanziani, 2004). This study suggests that early during stimulation, inhibition is primarily focused on the soma, and functions primarily to prevent saturation of the neuron's response. Later on in the stimulus, as recurrent activity increases, inhibition is dynamically routed into the dendrites. This presumably allows the neuron to control integration of these recurrent inputs. However, how this routing is control remains unknown. A possibility is that disinhibitory mechanisms determine how and when this routing takes place (Lovett-Barron et al., 2012).

The abovementioned inhibitory neurons are interconnected with a remarkable degree of specificity. These connections give rise to disinhibitory circuits, in which the

activity of one interneuron is modulated by another interneuron embedded within the same network. Recent studies have identified are three main dis-inhibitory circuits in the visual cortex, illustrated in **Figure 2-7**. Namely, SST neurons strongly inhibit PV neurons (SST→PV circuit), while VIP neurons strongly inhibit both SST neurons (VIP→SST circuit) and PV neurons (VIP→PV circuit). Surprisingly, PV neurons do not target any other interneuron subtype except for themselves (Pfeffer et al., 2013). This PV→PV connection allows PV neurons to self regulate their activity, increasing or decreasing FF inhibition depending on the level of FF excitation. Unfortunately, the computational role of these other disinhibitory circuits remains enigmatic primarily due to a lack of tools to interrogate their function.

A crucial first step towards understanding these disinhibitory circuits is to elucidate the cortical and subcortical inputs that drive these interneurons. It is likely that top-down cortical inputs or neuro-modulatory inputs are responsible for altering local microcircuits by recruiting these disinhibitory mechanisms. Using rabies-mediated viral tracing, Wall et al. (2016) were able to map the various brain areas that project to these interneurons. Interestingly, this study found that all interneuron subtypes received comparable levels of thalamic inputs. This suggests that all interneurons play some role in FF inhibition. The largest fraction of inputs to VIP neurons came from subcortical regions and these inputs were mainly neuromodulatory in nature (cholinergic, (Fu et al., 2014)). Both SST and VIP neurons also received a large projection from higher cortical areas, and from other sensory modalities (e.g. auditory and motor cortices). Taken together, these results suggest that many other brain areas can modulate local circuit computations by via these disinhibitory pathways. For example, a recent study showed that top-down projections from cingulate cortex, act via VIP and SST neurons to modulate the gain of pyramidal neurons in V1 (Jackson et al., 2016; Zhang et al., 2014).

VIP neurons are prime candidates for dynamically modulating the magnitude of cortical inhibition. A recent study found that VIP neurons in the auditory cortex were robustly recruited by reinforcement signals, such as air puff or foot shock, which in turn resulted in a robust lifting of pyramidal inhibition (i.e. disinhibition) (Pi et al., 2013). Similarly, another study also showed that cholinergic inputs, caused by a foot-shock,

robustly increased inhibition in L1 (mainly VIP neurons) of auditory cortex (Letzkus et al., 2011). This disinhibition enabled auditory fear learning. Taken together, these studies suggest that in the auditory cortex, VIP neurons are recruited by long-range neuromodulatory signals to release cortical inhibition by suppressing both PV and SST neurons. In doing so, VIP neurons allow behaviorally relevant events, like reinforcement signals, to drive learning in the auditory cortex. A similar circuit has been found in the visual cortex, where movement related cholinergic signals activate the VIP→SST circuit to drive disihibition (Fu et al., 2014). Surprisingly, there is very little evidence of a functional VIP→PV connection in V1 (Pfeffer et al., 2013), even though anatomical studies have shown that VIP neurons make contacts onto the somatic compartments of PV neurons (Hioki et al., 2013). Thus it is likely that VIP-mediated disinhibition in V1 functions mainly to lift dendritic inhibition.

In addition to these cholinergic reinforcement signals, long-range inputs from the auditory cortex and motor cortex impinging on VIP neurons can also cause cross-modality effects on sensory processing. For example, during active whisking, M1 inputs strongly drive an increase in VIP neuron activity, which is followed by a suppression of SST neurons in S1 (Lee et al., 2013). This suppression of SST neurons leads to an increase in nonlinear dendritic events in S1 pyramidal cells, which in turn enables better sensory processing (Xu et al., 2012). In the visual cortex, auditory stimuli strongly increased VIP inhibition, which in turn sharpened the orientation tuning of L2/3 pyramidal neurons (Ibrahim et al., 2016). Thus, cross-modality top-down modulation, which is mediated by the VIP→SST circuit, can sharpen sensory processing. In turn, this disinhibitory mechanism is likely to be advantageous to the survival of an animal as it can use information from multiple sensory modalities to compensate for noisy sensory cues. Taken together, the studies reviewed here provide strong evidence that VIP neurons function primarily as disinhibitory specialists.

The second major disinhibitory circuit is the SST→PV connection that has received far less attention than the VIP disinhibitory circuits. Anatomical studies have shown that SST neurons also inhibit the dendrites of PV neurons (Hioki et al., 2013). As mentioned above, PV neurons play a major role in gain control by changing the spiking

threshold of pyramidal neurons. Thus, the SST→PV pathway can be exploited to route inhibition from the soma to the dendrites. For example, strong activity in the SST network, caused by full-field stimuli, would inhibit PV neurons, and in turn route inhibition away from the soma and into the dendrites. Hypothetically, this reduction of somatic inhibition and increase in dendritic inhibition, would allow the neuron to spike more reliably, because only the most reliable inputs are filtered through the dendrites (see discussion in **Chapter 4**). However, the role of the SST→PV circuit in modulating pyramidal activity remains unexplored. A recent study used *in vivo* electrophysiology to show that SST neurons provide much stronger inhibition to PV than to pyramidal neurons (Cottam et al., 2013). This study proposed that, by suppressing PV neurons, SST neurons are able to sharpen orientation selectivity by selectively suppressing untuned inputs. However, how this can be achieved however, was not explored by this study. More recently, the role of the SST→PV circuit during development and disease has been highlighted. Specifically, it has been shown that early-born SST neurons play an important role in shaping the development of PV circuits (Tuncdemir et al., 2016). Also, hyper-excitability of SST neurons in ALS and dementia leads to a pathological excitotoxicity in the cortex (Zhang et al., 2016). Taken together, it is clear that the SST→PV circuit plays a very important role in modulating FF inhibition. However, many more experiments are still required to properly dissect the role of this circuit in information processing.

In conclusion, although cortical inhibition is diverse, we are now beginning to gain a better picture of the role of each circuit component in cortical information processing. New tools, such as double transgenic mice (see **Chapter 4**), will allow us to better study the interactions between these circuit elements. These experiments will bring new clarity to our understanding of how the various cortical interneurons regulate information processing. Furthermore, once we elucidate their function, then we will be able to understand how these neurons go awry in neurodevelopmental and psychiatric disorders (see **Chapter 5**).

2.6 Open questions

The studies presented in this chapter provide a strong argument that we still know very little about how and why natural scenes are processed with more efficient codes in the visual cortex. Specifically, I have highlighted four open questions that my thesis hopes to address.

1. How do spatial statistics of natural scenes modulate response reliability?

While the work summarized in Section 2.3 suggests that stimuli with broadband spatiotemporal statistics are more reliably processed than simpler stimuli, it still remains unknown what features of natural scenes are necessary for reliable coding and how. One hypothesis is that amplitude spectrum, which contains information about the spatial correlations contained within an image. Due to the retinotopic organization of the visual cortex, spatial correlations within the image can coactivate distinct ensembles of neurons, thereby changing the network dynamics and altering response reliability. This hypothesis is tested in **Chapter 3**. Specifically, I develop a new noise-masking algorithm to perturb the amplitude spectrum of natural movies. Using two-photon calcium imaging I demonstrate that the amplitude spectrum does play an important role in structuring correlations between neurons. This restructuring gives rise to ensembles of neurons, which share low noise correlations, and pool to produce reliable responses. This work shows for the first time that perturbing the amplitude spectrum changes stimulus coding, primarily by influencing reliability. More importantly, this work shows that population coding through correlated ensembles helps to minimize intrinsic variability.

2. What inhibitory circuit mechanisms does the cortex use to modulate reliability?

Given the plethora of interneurons in the cortex and the impact that changing network dynamics on neuronal variability, it is likely that inhibition is the key mechanism by which the cortex is able to dynamically regulate noise. However, no study has yet shown how this can be achieved. In Section 2.5, I described that, by inhibiting mainly

the somatic compartment, PV neurons play a critical role in preserving the balance between excitation and inhibition through gain control. In contrast SST neurons work primarily to alter dendritic integration by inhibiting specific compartments of the dendritic arbor. Thus, it is possible that both PV and SST neurons work in concert to maximize SNR, by reducing noise in the input and increasing signal gain of the output. In **Chapter 4**, I will demonstrate, using a combination of calcium imaging and optogenetic manipulation, that the SST→PV circuit does indeed play a major role in controlling neuronal reliability. Specifically, by simultaneously inhibiting PV neurons and pyramidal neuron dendrites, SST neurons are able to create a window of opportunity which reliable spiking can occur. The work that I will present in **Chapter 4** shows for the first time how the disinhibitory circuits modulate coding efficiency in mouse visual cortex.

3. What statistics do mice use for visual perception?

Before we can use mice as a model system for studying visual perception, we need to understand how mice use their vision to guide goal-directed actions. While mice can use their vision to discriminate contrast and orientation, it remains unknown what strategies they use for discriminating between more complex natural scenes. Such experiments will reveal crucial insights into how mice use vision to perceive their environment. In **Chapter 6**, I use a head-fixed Go/NoGo paradigm to show that mice an learn within a short period of time to discriminate between different natural movies. By perturbing the statistics of these movies, I show that mice use phase information to discriminate between movies. Finally, I demonstrate that activating SST neurons improves the reliability of the neuronal code and also increases discriminability. Conclusively, this result shows reliable processing leads to better perception

4. How are cortical computations altered when inhibition goes awry?

Numerous neurodevelopmental and psychiatric disorders are related to changes in inhibition. However, relatively few studies have related chronic perturbations in specific subsets of neurons to circuit level deficits associated with the disease. In **Chapter 5**, we investigate how network dynamics changes when inhibition is chronically

perturbed. We also focus on understanding the circuit level deficits in Rett Syndrome, a monogenic disorder on the Autism spectrum disorder. While many studies have looked at processing deficits in Rett Syndrome, no study has yet shown how the processing of complex stimuli is affected. Using cell-specific knock-out mice and calcium imaging, we show in **Chapter 5** that deleting MeCP2 from PV neurons results in decorrelated and unreliable responses to natural scenes. In contrast, deleting MeCP2 from SST neurons reduces the ability of mice to discriminate between different scenes. This work reconciles recent studies that showed unreliable visual processing in human patients with Rett Syndrome. Together, our finding in this chapter demonstrates conclusively that loss of MeCP2 from specific interneurons lead to specific deficits observed in Rett Syndrome.

In sum, the work that I present in this thesis investigates the mechanisms responsible for the reliable coding of natural scenes in the mouse visual cortex. By focusing on both healthy and diseased mice, I describe the conditions required for reliable coding in the visual cortex.

2.7 References

Abbott, L.F., Dayan, P., 1999. The effect of correlated variability on the accuracy of a population code. Neural Comput. 11, 91–101.

Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z.J., Scanziani, M., 2012. A neural circuit for spatial summation in visual cortex. Nature 490, 226–31. doi:10.1038/nature11526

Andermann, M.L., Gilfoy, N.B., Goldey, G.J., Sachdev, R.N.S., Wölfel, M., McCormick, D.A., Reid, R.C., Levene, M.J., 2013. Chronic cellular imaging of entire cortical columns in awake mice using microprisms. Neuron 80, 900–13. doi:10.1016/j.neuron.2013.07.052

Andermann, M.L., Kerlin, A.M., Reid, R.C., 2010. Chronic cellular imaging of mouse visual cortex during operant behavior and passive viewing. Front. Cell. Neurosci. 4, 3. doi:10.3389/fncel.2010.00003

Andermann, M.L., Kerlin, A.M., Roumis, D.K., Glickfeld, L.L., Reid, R.C., 2011. Functional specialization of mouse higher visual cortical areas. Neuron 72, 1025–39. doi:10.1016/j.neuron.2011.11.013

Anderson, J.S., Carandini, M., Ferster, D., 2000. Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. J. Neurophysiol. 84, 909–26.

Aston-Jones, G., Cohen, J.D., 2005. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu. Rev. Neurosci. 28, 403–50. doi:10.1146/annurev.neuro.28.061604.135709

Atallah, B. V, Bruns, W., Carandini, M., Scanziani, M., 2012. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. Neuron 73, 159–70. doi:10.1016/j.neuron.2011.12.013

Averbeck, B.B., Latham, P.E., Pouget, A., 2006. Neural correlations, population coding and computation. Nat. Rev. Neurosci. 7, 358–66. doi:10.1038/nrn1888

Azouz, R., Gray, C.M., 1999. Cellular mechanisms contributing to response variability of cortical neurons in vivo. J. Neurosci. 19, 2209–2223.

Barlow, H., 1994. What Is the Computational Goal of the Neocortex? MIT Press, Cambridge, MA.

Bathellier, B., Ushakova, L., Rumpel, S., 2012. Discrete neocortical dynamics predict behavioral categorization of sounds. Neuron 76, 435–49. doi:10.1016/j.neuron.2012.07.008

Baudot, P., Levy, M., Marre, O., Monier, C., Pananceau, M., Frégnac, Y., 2013. Animation of

natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. Front. Neural Circuits 7, 206. doi:10.3389/fncir.2013.00206

Bloss, E.B., Cembrowski, M.S., Karsh, B., Colonell, J., Fetter, R.D., Spruston, N., 2016. Structured Dendritic Inhibition Supports Branch-Selective Integration in CA1 Pyramidal Cells. Neuron 89, 1016–30. doi:10.1016/j.neuron.2016.01.029

Borst, A., Theunissen, F.E., 1999. Information theory and neural coding. Nat. Neurosci. 2, 947–57. doi:10.1038/14731

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

Burr, D.C., 1980. Sensitivity to spatial phase. Vision Res. 20, 391–6.

Busse, L., Ayaz, A., Dhruv, N.T., Katzner, S., Saleem, A.B., Schölvinck, M.L., Zaharia, A.D., Carandini, M., 2011. The detection of visual contrast in the behaving mouse. J. Neurosci. 31, 11351–61. doi:10.1523/JNEUROSCI.6689-10.2011

Butts, D. a, Weng, C., Jin, J., Yeh, C.-I., Lesica, N. a, Alonso, J.-M., Stanley, G.B., 2007. Temporal precision in the neural code and the timescales of natural vision. Nature 449, 92–95. doi:10.1038/nature06105

Carandini, M., Heeger, D.J., 2012. Normalization as a canonical neural computation. Nat Rev Neurosci 13, 51–62. doi:10.1038/nrn3136

Chance, F.S., Abbott, L.F., 2000. Divisive inhibition in recurrent networks. Network 11, 119–29.

Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., Looger, L.L., Svoboda, K., Kim, D.S., 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295–300. doi:10.1038/nature12354

Chiu, C.Q., Lur, G., Morse, T.M., Carnevale, N.T., Ellis-Davies, G.C.R., Higley, M.J., 2013. Compartmentalization of GABAergic inhibition by dendritic spines. Science 340, 759–62. doi:10.1126/science.1234274

Choi, J.-S., Kim, J.J., 2010. Amygdala regulates risk of predation in rats foraging in a dynamic fear environment. Proc. Natl. Acad. Sci. U. S. A. 107, 21773–7. doi:10.1073/pnas.1010079108

Cohen, M.R., Maunsell, J.H.R., 2009. Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. 12, 1594–600. doi:10.1038/nn.2439

Cossell, L., Iacaruso, M.F., Muir, D.R., Houlton, R., Sader, E.N., Ko, H., Hofer, S.B., Mrsic-

Flogel, T.D., 2015. Functional organization of excitatory synaptic strength in primary visual cortex. Nature. doi:10.1038/nature14182

Cottam, J.C.H., Smith, S.L., Häusser, M., 2013. Target-specific effects of somatostatin-expressing interneurons on neocortical visual processing. J. Neurosci. 33, 19567–78. doi:10.1523/JNEUROSCI.2624-13.2013

David, S. V, Vinje, W.E., Gallant, J.L., 2004. Natural stimulus statistics alter the receptive field structure of v1 neurons. J. Neurosci. 24, 6991–7006. doi:10.1523/JNEUROSCI.1422-04.2004

DeFelipe, J., López-Cruz, P.L., Benavides-Piccione, R., Bielza, C., Larrañaga, P., Anderson, S., Burkhalter, A., Cauli, B., Fairén, A., Feldmeyer, D., Fishell, G., Fitzpatrick, D., Freund, T.F., González-Burgos, G., Hestrin, S., Hill, S., Hof, P.R., Huang, J., Jones, E.G., Kawaguchi, Y., Kisvárday, Z., Kubota, Y., Lewis, D.A., Marín, O., Markram, H., McBain, C.J., Meyer, H.S., Monyer, H., Nelson, S.B., Rockland, K., Rossier, J., Rubenstein, J.L.R., Rudy, B., Scanziani, M., Shepherd, G.M., Sherwood, C.C., Staiger, J.F., Tamás, G., Thomson, A., Wang, Y., Yuste, R., Ascoli, G.A., 2013. New insights into the classification and nomenclature of cortical GABAergic interneurons. Nat. Rev. Neurosci. 14, 202–16. doi:10.1038/nrn3444

Denève, S., Machens, C.K., 2016. Efficient codes and balanced networks. Nat. Neurosci. 19, 375–82. doi:10.1038/nn.4243

Dinstein, I., Heeger, D.J., Behrmann, M., 2015. Neural variability: friend or foe? Trends Cogn. Sci. 19, 322–8. doi:10.1016/j.tics.2015.04.005

Dong, D., Atick, J., 1995. Statistics of natural time-varying images. Netw. Comput. Neural Syst. 6, 345–358. doi:10.1088/0954-898X/6/3/003

Dräger, U.C., 1975. Receptive fields of single cells and topography in mouse visual cortex. J. Comp. Neurol. 160, 269–90. doi:10.1002/cne.901600302

Egger, R., Schmitt, A.C., Wallace, D.J., Sakmann, B., Oberlaender, M., Kerr, J.N.D., 2015. Robustness of sensory-evoked excitation is increased by inhibitory inputs to distal apical tuft dendrites. Proc. Natl. Acad. Sci. U. S. A. 112, 14072–7. doi:10.1073/pnas.1518773112

El-Boustani, S., Sur, M., 2014. Response-dependent dynamics of cell-specific inhibition in cortical networks in vivo. Nat. Commun. 5, 5689. doi:10.1038/ncomms6689

Faisal, A.A., Selen, L.P.J., Wolpert, D.M., 2008. Noise in the nervous system. Nat. Rev. Neurosci. 9, 292–303. doi:10.1038/nrn2258

Felleman, D.J., Van Essen, D.C., n.d. Distributed hierarchical processing in the primate cerebral

cortex. Cereb. Cortex 1, 1–47.

Fiete, I.R., Seung, H.S., 2006. Gradient learning in spiking neural networks by dynamic perturbation of conductances. Phys. Rev. Lett. 97, 048104. doi:10.1103/PhysRevLett.97.048104

Fino, E., Yuste, R., 2011. Dense inhibitory connectivity in neocortex. Neuron 69, 1188–203. doi:10.1016/j.neuron.2011.02.025

Fiser, J., Chiu, C., Weliky, M., 2004. Small modulation of ongoing cortical dynamics by sensory input during natural vision. Nature 431, 573–8. doi:10.1038/nature02907

Fournier, J., Monier, C., Pananceau, M., Fregnac, Y., 2011. Adaptation of the simple or complex nature of V1 receptive fields to visual statistics. Nat Neurosci 14, 1053–1060. doi:10.1038/nn.2861

Froemke, R.C., 2015. Plasticity of cortical excitatory-inhibitory balance. Annu. Rev. Neurosci. 38, 195–219. doi:10.1146/annurev-neuro-071714-034002

Froudarakis, E., Berens, P., Ecker, A.S., Cotton, R.J., Sinz, F.H., Yatsenko, D., Saggau, P., Bethge, M., Tolias, A.S., 2014. Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. Nat Neurosci 17, 851–857. doi:10.1038/nn.3707

Fu, Y., Tucciarone, J.M., Espinosa, J.S., Sheng, N., Darcy, D.P., Nicoll, R.A., Huang, Z.J., Stryker, M.P., 2014. A cortical circuit for gain control by behavioral state. Cell 156, 1139–52. doi:10.1016/j.cell.2014.01.050

Gabernet, L., Jadhav, S.P., Feldman, D.E., Carandini, M., Scanziani, M., 2005. Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. Neuron 48, 315–27. doi:10.1016/j.neuron.2005.09.022

Gao, E., DeAngelis, G.C., Burkhalter, A., 2010. Parallel input channels to mouse primary visual cortex. J Neurosci 30, 5912–5926. doi:10.1523/JNEUROSCI.6456-09.2010

Garrett, M.E., Nauhaus, I., Marshel, J.H., Callaway, E.M., 2014. Topography and areal organization of mouse visual cortex. J. Neurosci. 34, 12587–600. doi:10.1523/JNEUROSCI.1124-14.2014

Gaspar, C.M., Rousselet, G.A., 2009. How do amplitude spectra influence rapid animal detection? Vision Res. 49, 3001–12. doi:10.1016/j.visres.2009.09.021

Georgopoulos, A.P., Kalaska, J.F., Caminiti, R., Massey, J.T., 1982. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. J. Neurosci. 2, 1527–37.

Gidon, A., Segev, I., 2012. Principles Governing the Operation of Synaptic Inhibition in

Dendrites. Neuron 75, 330–341. doi:10.1016/j.neuron.2012.05.015

Glickfeld, L.L., Andermann, M.L., Bonin, V., Reid, R.C., 2013a. Cortico-cortical projections in mouse visual cortex are functionally target specific. Nat. Neurosci. 16, 219–26. doi:10.1038/nn.3300

Glickfeld, L.L., Histed, M.H., Maunsell, J.H.R., 2013b. Mouse primary visual cortex is used to detect both orientation and contrast changes. J. Neurosci. 33, 19416–22. doi:10.1523/JNEUROSCI.3560-13.2013

Glickfeld, L.L., Reid, R.C., Andermann, M.L., 2014. A mouse model of higher visual cortical function. Curr. Opin. Neurobiol. 24, 28–33. doi:10.1016/j.conb.2013.08.009

Goard, M., Dan, Y., 2009. Basal forebrain activation enhances cortical coding of natural scenes. Nat. Neurosci. 12, 1444–9. doi:10.1038/nn.2402

Goris, R.L.T., Movshon, J.A., Simoncelli, E.P., 2014. Partitioning neuronal variability. Nat. Neurosci. 17, 858–65. doi:10.1038/nn.3711

Graupner, M., Reyes, A.D., 2013. Synaptic input correlations leading to membrane potential decorrelation of spontaneous activity in cortex. J. Neurosci. 33, 15075–85. doi:10.1523/JNEUROSCI.0347-13.2013

Grubb, M.S., Thompson, I.D., 2003. Quantitative characterization of visual response properties in the mouse dorsal lateral geniculate nucleus. J. Neurophysiol. 90, 3594–607. doi:10.1152/jn.00699.2003

Gu, Y., Liu, S., Fetsch, C.R., Yang, Y., Fok, S., Sunkara, A., DeAngelis, G.C., Angelaki, D.E., 2011. Perceptual learning reduces interneuronal correlations in macaque visual cortex. Neuron 71, 750–61. doi:10.1016/j.neuron.2011.06.015

Gur, M., Beylin, A., Snodderly, D.M., 1997. Response variability of neurons in primary visual cortex (V1) of alert monkeys. J. Neurosci. 17, 2914–20.

Haider, B., Häusser, M., Carandini, M., 2013. Inhibition dominates sensory responses in the awake cortex. Nature 493, 97–100. doi:10.1038/nature11665

Haider, B., Krause, M.R., Duque, A., Yu, Y., Touryan, J., Mazer, J. a, McCormick, D. a, 2010. Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. Neuron 65, 107–21. doi:10.1016/j.neuron.2009.12.005

Haider, B., McCormick, D. a, 2009. Rapid neocortical dynamics: cellular and network mechanisms. Neuron 62, 171–89. doi:10.1016/j.neuron.2009.04.008

Hamilton, L.S., Sohl-Dickstein, J., Huth, A.G., Carels, V.M., Deisseroth, K., Bao, S., 2013.

Optogenetic activation of an inhibitory network enhances feedforward functional connectivity in auditory cortex. Neuron 80, 1066–76. doi:10.1016/j.neuron.2013.08.017

Harris, K.D., Mrsic-Flogel, T.D., 2013. Cortical connectivity and sensory coding. Nature 503, 51–8. doi:10.1038/nature12654

Harvey, C.D., Coen, P., Tank, D.W., 2012. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature 484, 62–8. doi:10.1038/nature10918

Heggelund, P., Albus, K., 1978. Response variability and orientation discrimination of single cells in striate cortex of cat. Exp. brain Res. 32, 197–211.

Hioki, H., Okamoto, S., Konno, M., Kameda, H., Sohn, J., Kuramoto, E., Fujiyama, F., Kaneko, T., 2013. Cell type-specific inhibitory inputs to dendritic and somatic compartments of parvalbumin-expressing neocortical interneuron. J. Neurosci. 33, 544–55. doi:10.1523/JNEUROSCI.2255-12.2013

Hires, S.A., Gutnisky, D.A., Yu, J., Connor, D.H.O., Svoboda, K., 2015. Low-noise encoding of active touch by layer 4 in the somatosensory cortex 1–18. doi:10.7554/eLife.06619

Hofer, S.B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., Lein, E., Lesica, N. a, Mrsic-Flogel, T.D., 2011. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. Nat. Neurosci. 14, 1045–52. doi:10.1038/nn.2876

Holt, G.R., Koch, C., 1997. Shunting inhibition does not have a divisive effect on firing rates. Neural Comput. 9, 1001–13.

Huberman, A.D., Niell, C.M., 2011. What can mice tell us about how vision works? Trends Neurosci. 34, 464–73. doi:10.1016/j.tins.2011.07.002

Ibrahim, L.A., Mesik, L., Ji, X.-Y., Fang, Q., Li, H.-F., Li, Y.-T., Zingg, B., Zhang, L.I., Tao, H.W., 2016. Cross-Modality Sharpening of Visual Cortical Processing through Layer-1-Mediated Inhibition and Disinhibition. Neuron 89, 1031–45. doi:10.1016/j.neuron.2016.01.027

Isaacson, J.S., Scanziani, M., 2011. How inhibition shapes cortical activity. Neuron 72, 231–43. doi:10.1016/j.neuron.2011.09.027

Jackson, J., Ayzenshtat, I., Karnani, M.M., Yuste, R., 2016. VIP+ interneurons control neocortical activity across brain states. J. Neurophysiol. jn.01124.2015. doi:10.1152/jn.01124.2015

Jacobs, G.H., Williams, G.A., Cahill, H., Nathans, J., 2007. Emergence of novel color vision in mice engineered to express a human cone photopigment. Science 315, 1723–5. doi:10.1126/science.1138838

Jeanne, J.M., Sharpee, T.O., Gentner, T.Q., 2013. Associative learning enhances population

coding by inverting interneuronal correlation patterns. Neuron 78, 352–63. doi:10.1016/j.neuron.2013.02.023

Jeon, C.J., Strettoi, E., Masland, R.H., 1998. The major cell populations of the mouse retina. J. Neurosci. 18, 8936–46.

Jia, H., Rochefort, N.L., Chen, X., Konnerth, A., 2010. Dendritic organization of sensory input to cortical neurons in vivo. Nature 464, 1307–12. doi:10.1038/nature08947

Joubert, O.R., Rousselet, G.A., Fabre-thorpe, M., Fize, D., 2009. Rapid visual categorization of natural scene contexts with equalized amplitude spectrum and increasing phase noise 9, 1–16. doi:10.1167/9.1.2.Introduction

Kampa, B.M., Roth, M.M., Göbel, W., Helmchen, F., 2011. Representation of visual scenes by local neuronal populations in layer 2/3 of mouse visual cortex. Front. Neural Circuits 5, 18. doi:10.3389/fncir.2011.00018

Kara, P., Reid, R.C., 2003. Efficacy of retinal spikes in driving cortical responses. J. Neurosci. 23, 8547–57.

Karnani, M.M., Agetsuma, M., Yuste, R., 2014. A blanket of inhibition: functional inferences from dense inhibitory connectivity. Curr. Opin. Neurobiol. 26, 96–102. doi:10.1016/j.conb.2013.12.015

Kepecs, A., Fishell, G., 2014. Interneuron cell types are fit to function. Nature 505, 318–26. doi:10.1038/nature12983

Kerlin, A.M., Andermann, M.L., Berezovskii, V.K., Reid, R.C., 2010. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. Neuron 67, 858–71. doi:10.1016/j.neuron.2010.08.002

Klausberger, T., Somogyi, P., 2008. Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. Science 321, 53–7. doi:10.1126/science.1149381

Ko, H., Hofer, S.B., Pichler, B., Buchanan, K.A., Sjostrom, P.J., Mrsic-Flogel, T.D., 2011. Functional specificity of local synaptic connections in neocortical networks. Nature 473, 87–91. doi:10.1038/nature09880

Koulakov, A.A., Chklovskii, D.B., 2001. Orientation preference patterns in mammalian visual cortex: a wire length minimization approach. Neuron 29, 519–27.

Kubota, Y., 2014. Untangling GABAergic wiring in the cortical microcircuit. Curr. Opin. Neurobiol. 26, 7–14. doi:10.1016/j.conb.2013.10.003

Laughlin, S.B., 2001. Energy as a constraint on the coding and processing of sensory information. Curr. Opin. Neurobiol. 11, 475–80.

Lee, S., Kruglikov, I., Huang, Z.J., Fishell, G., Rudy, B., 2013. A disinhibitory circuit mediates motor integration in the somatosensory cortex. Nat. Neurosci. 16, 1662–70. doi:10.1038/nn.3544

Lee, S.-H., Kwan, A.C., Zhang, S., Phoumthipphavong, V., Flannery, J.G., Masmanidis, S.C., Taniguchi, H., Huang, Z.J., Zhang, F., Boyden, E.S., Deisseroth, K., Dan, Y., 2012. Activation of specific interneurons improves V1 feature selectivity and visual perception. Nature 488, 379–83. doi:10.1038/nature11312

Letzkus, J.J., Wolff, S.B.E., Meyer, E.M.M., Tovote, P., Courtin, J., Herry, C., Lüthi, A., 2011. A disinhibitory microcircuit for associative fear learning in the auditory cortex. Nature 480, 331–5. doi:10.1038/nature10674

Lewis, D.A., Hashimoto, T., Volk, D.W., 2005. Cortical inhibitory neurons and schizophrenia. Nat. Rev. Neurosci. 6, 312–24. doi:10.1038/nrn1648

Li, F.F., VanRullen, R., Koch, C., Perona, P., 2002. Rapid natural scene categorization in the near absence of attention. Proc. Natl. Acad. Sci. U. S. A. 99, 9596–601. doi:10.1073/pnas.092277599

Lien, A.D., Scanziani, M., 2013. Tuned thalamic excitation is amplified by visual cortical circuits. Nat. Neurosci. 16, 1315–23. doi:10.1038/nn.3488

Litwin-Kumar, A., Doiron, B., 2012. Slow dynamics and high variability in balanced cortical networks with clustered connections. Nat. Neurosci. 15, 1498–505. doi:10.1038/nn.3220

Liu, B., Li, P., Li, Y., Sun, Y.J., Yanagawa, Y., Obata, K., Zhang, L.I., Tao, H.W., 2009. Visual receptive field structure of cortical inhibitory neurons revealed by two-photon imaging guided recording. J. Neurosci. 29, 10520–32. doi:10.1523/JNEUROSCI.1915-09.2009

London, M., Roth, A., Beeren, L., Häusser, M., Latham, P.E., 2010. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. Nature 466, 123–7. doi:10.1038/nature09086

Lovett-Barron, M., Turi, G.F., Kaifosh, P., Lee, P.H., Bolze, F., Sun, X.-H., Nicoud, J.-F., Zemelman, B. V, Sternson, S.M., Losonczy, A., 2012. Regulation of neuronal input transformations by tunable dendritic inhibition. Nat. Neurosci. 15, 423–430. doi:10.1038/nn.3024

Ma, W., Liu, B., Li, Y., Huang, Z.J., Zhang, L.I., Tao, H.W., 2010. Visual representations by cortical somatostatin inhibitory neurons--selective but with weak and delayed responses. J. Neurosci. 30, 14371–9. doi:10.1523/JNEUROSCI.3248-10.2010

Mainen, Z.F., Sejnowski, T.J., 1995. Reliability of spike timing in neocortical neurons. Science

268, 1503–1506. doi:10.1126/science.7770778

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., Wu, C., 2004. Interneurons of the neocortical inhibitory system. Nat. Rev. Neurosci. 5, 793–807. doi:10.1038/nrn1519

Marshel, J.H., Garrett, M.E., Nauhaus, I., Callaway, E.M., 2011. Functional specialization of seven mouse visual cortical areas. Neuron 72, 1040–54. doi:10.1016/j.neuron.2011.12.004

Marshel, J.H., Kaye, A.P., Nauhaus, I., Callaway, E.M., 2012. Anterior-posterior direction opponency in the superficial mouse lateral geniculate nucleus. Neuron 76, 713–20. doi:10.1016/j.neuron.2012.09.021

Maunsell, J.H.R., 2015. Neuronal Mechanisms of Visual Attention. Annu. Rev. Vis. Sci. 1, 373–391. doi:10.1146/annurev-vision-082114-035431

McAdams, C.J., Maunsell, J.H., 1999. Effects of attention on the reliability of individual neurons in monkey visual cortex. Neuron 23, 765–73.

McCotter, M., Gosselin, F., Sowden, P., Schyns, P., 2005. The use of visual information in natural scenes. Vis. cogn. 12, 938–953. doi:10.1080/13506280444000599

Miller, J.-E.K., Ayzenshtat, I., Carrillo-Reid, L., Yuste, R., 2014. Visual stimuli recruit intrinsically generated cortical ensembles. Proc. Natl. Acad. Sci. U. S. A. doi:10.1073/pnas.1406077111

Monier, C., Chavane, F., Baudot, P., Graham, L.J., Frégnac, Y., 2003. Orientation and direction selectivity of synaptic inputs in visual cortical neurons: a diversity of combinations produces spike tuning. Neuron 37, 663–80.

Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., Pouget, A., 2014. Information-limiting correlations. Nat. Neurosci. 17, 1410–7. doi:10.1038/nn.3807

Murayama, M., Pérez-Garci, E., Nevian, T., Bock, T., Senn, W., Larkum, M.E., 2009. Dendritic encoding of sensory stimuli controlled by deep cortical interneurons. Nature 457, 1137–41. doi:10.1038/nature07663

Nauhaus, I., Nielsen, K.J., 2014. Building maps from maps in primary visual cortex. Curr. Opin. Neurobiol. 24, 1–6. doi:10.1016/j.conb.2013.08.007

Niell, C.M., 2015. Cell types, circuits, and receptive fields in the mouse visual cortex. Annu. Rev. Neurosci. 38, 413–31. doi:10.1146/annurev-neuro-071714-033807

Niell, C.M., Stryker, M.P., 2008. Highly selective receptive fields in mouse visual cortex. J. Neurosci. 28, 7520–36. doi:10.1523/JNEUROSCI.0623-08.2008

Nienborg, H., Hasenstaub, a., Nauhaus, I., Taniguchi, H., Huang, Z.J., Callaway, E.M., 2013.

Contrast Dependence and Differential Contributions from Somatostatin- and Parvalbumin-Expressing Neurons to Spatial Integration in Mouse V1. J. Neurosci. 33, 11145–11154. doi:10.1523/JNEUROSCI.5320-12.2013

Ohki, K., Reid, R.C., 2007. Specificity and randomness in the visual cortex. Curr. Opin. Neurobiol. 17, 401–7. doi:10.1016/j.conb.2007.07.007

Okun, M., Lampl, I., 2008. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nat. Neurosci. 11, 535–7. doi:10.1038/nn.2105

Oliva, A., Torralba, A., 2007. The role of context in object recognition. Trends Cogn. Sci. 11, 520–7. doi:10.1016/j.tics.2007.09.009

Olshausen, B.A., Field, D.J., 2004. Sparse coding of sensory inputs. Curr. Opin. Neurobiol. 14, 481–7. doi:10.1016/j.conb.2004.07.007

Olshausen, B.A., Field, D.J., 1996. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nature 381, 607–9. doi:10.1038/381607a0

Onat, S., König, P., Jancke, D., 2011. Natural scene evoked population dynamics across cat primary visual cortex captured with voltage-sensitive dye imaging. Cereb. Cortex 21, 2542–54. doi:10.1093/cercor/bhr038

Orban, G.A., 2008. Higher order visual processing in macaque extrastriate cortex. Physiol. Rev. 88, 59–89. doi:10.1152/physrev.00008.2007

Ozeki, H., Finn, I.M., Schaffer, E.S., Miller, K.D., Ferster, D., 2009. Inhibitory stabilization of the cortical network underlies visual surround suppression. Neuron 62, 578–92. doi:10.1016/j.neuron.2009.03.028

Packer, A.M., McConnell, D.J., Fino, E., Yuste, R., 2013. Axo-dendritic overlap and laminar projection can explain interneuron connectivity to pyramidal cells. Cereb. Cortex 23, 2790–802. doi:10.1093/cercor/bhs210

Palmer, L., Murayama, M., Larkum, M., 2012. Inhibitory Regulation of Dendritic Activity in vivo. Front. Neural Circuits 6, 26. doi:10.3389/fncir.2012.00026

Parker, A.J., Newsome, W.T., 1998. Sense and the single neuron: probing the physiology of perception. Annu. Rev. Neurosci. 21, 227–77. doi:10.1146/annurev.neuro.21.1.227

Párraga, C. a, Troscianko, T., Tolhurst, D.J., 2005. The effects of amplitude-spectrum statistics on foveal and peripheral discrimination of changes in natural images, and a multi-resolution model. Vision Res. 45, 3145–68. doi:10.1016/j.visres.2005.08.006

Pecka, M., Han, Y., Sader, E., Mrsic-Flogel, T.D., 2014. Experience-Dependent Specialization of Receptive Field Surround for Selective Coding of Natural Scenes. Neuron.

doi:10.1016/j.neuron.2014.09.010

Peters, A.J., Chen, S.X., Komiyama, T., 2014. Emergence of reproducible spatiotemporal activity during motor learning. Nature 510, 263–7. doi:10.1038/nature13235

Pfeffer, C.K., Xue, M., He, M., Huang, Z.J., Scanziani, M., 2013. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat. Neurosci. 1–12. doi:10.1038/nn.3446

Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J.I., Huang, Z.J., Kepecs, A., 2013. Cortical interneurons that specialize in disinhibitory control. Nature 503, 521–4. doi:10.1038/nature12676

Piscopo, D.M., El-Danaf, R.N., Huberman, A.D., Niell, C.M., 2013. Diverse visual features encoded in mouse lateral geniculate nucleus. J. Neurosci. 33, 4642–56. doi:10.1523/JNEUROSCI.5187-12.2013

Pouille, F., Marin-Burgin, A., Adesnik, H., Atallah, B. V, Scanziani, M., 2009. Input normalization by global feedforward inhibition expands cortical dynamic range. Nat. Neurosci. 12, 1577–85. doi:10.1038/nn.2441

Pouille, F., Scanziani, M., 2004. Routing of spike series by dynamic circuits in the hippocampus. Nature 429, 717–723. doi:10.1038/nature02615

Pouille, F., Scanziani, M., 2001. Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. Science 293, 1159–1163. doi:10.1126/science.1060342

Prusky, G.T., Douglas, R.M., 2004. Characterization of mouse cortical spatial vision. Vision Res. 44, 3411–8. doi:10.1016/j.visres.2004.09.001

Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D., Denfield, G.H., Tolias, A.S., 2014. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. Neuron 84, 355–62. doi:10.1016/j.neuron.2014.09.033

Renart, A., Machens, C.K., 2014. Variability in neural activity and behavior. Curr. Opin. Neurobiol. 25, 211–220. doi:10.1016/j.conb.2014.02.013

Ringach, D.L., Hawken, M.J., Shapley, R., 2002. Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. J. Vis. 2, 12–24. doi:10:1167/2.1.2

Rokni, U., Richardson, A.G., Bizzi, E., Seung, H.S., 2007. Motor learning with unstable neural representations. Neuron 54, 653–66. doi:10.1016/j.neuron.2007.04.030

Roth, M.M., Helmchen, F., Kampa, B.M., 2012. Distinct functional properties of primary and posteromedial visual area of mouse neocortex. J. Neurosci. 32, 9716–26.

doi:10.1523/JNEUROSCI.0110-12.2012

Ruderman, D., Bialek, W., 1994. Statistics of natural images: Scaling in the woods. Phys. Rev. Lett. 73, 814–817.

Ruderman, D.L., 1997. Origins of scaling in natural images. Vision Res. 37, 3385–98.

Rudy, B., Fishell, G., Lee, S., Hjerling-Leffler, J., 2011. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev. Neurobiol. 71, 45–61. doi:10.1002/dneu.20853

Runyan, C.A., Schummers, J., Van Wart, A., Kuhlman, S.J., Wilson, N.R., Huang, Z.J., Sur, M., 2010. Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. Neuron 67, 847–57. doi:10.1016/j.neuron.2010.08.006

Serre, T., Kreiman, G., Kouh, M., Cadieu, C., Knoblich, U., Poggio, T., 2007. A quantitative theory of immediate visual recognition. Prog. Brain Res. 165, 33–56. doi:10.1016/S0079-6123(06)65004-8

Seybold, B.A., Phillips, E.A.K., Schreiner, C.E., Hasenstaub, A.R., 2015. Inhibitory Actions Unified by Network Integration. Neuron 87, 1181–92. doi:10.1016/j.neuron.2015.09.013

Shadlen, M.N., Newsome, W.T., 1998. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. 18, 3870–96.

Silberberg, G., Markram, H., 2007. Disynaptic inhibition between neocortical pyramidal cells mediated by Martinotti cells. Neuron 53, 735–46. doi:10.1016/j.neuron.2007.02.012

Silver, R.A., 2010. Neuronal arithmetic. Nat. Rev. Neurosci. 11, 474–89. doi:10.1038/nrn2864

Simoncelli, E.P., Olshausen, B.A., 2001. Natural image statistics and neural representation. Annu. Rev. Neurosci. 24, 1193–216. doi:10.1146/annurev.neuro.24.1.1193

Skottun, B.C., De Valois, R.L., Grosof, D.H., Movshon, J.A., Albrecht, D.G., Bonds, A.B., 1991. Classifying simple and complex cells on the basis of response modulation. Vision Res. 31, 1079–86.

Sofroniew, N.J., Cohen, J.D., Lee, A.K., Svoboda, K., 2014. Natural whisker-guided behavior by head-fixed mice in tactile virtual reality. J. Neurosci. 34, 9537–50. doi:10.1523/JNEUROSCI.0712-14.2014

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

Sohya, K., Kameyama, K., Yanagawa, Y., Obata, K., Tsumoto, T., 2007. GABAergic neurons are less selective to stimulus orientation than excitatory neurons in layer II/III of visual

cortex, as revealed by in vivo functional Ca2+ imaging in transgenic mice. J. Neurosci. 27, 2145–9. doi:10.1523/JNEUROSCI.4641-06.2007

Spruston, N., 2008. Pyramidal neurons: dendritic structure and synaptic integration. Nat. Rev. Neurosci. 9, 206–21. doi:10.1038/nrn2286

Stevens, C.F., Zador, A.M., 1998. Input synchrony and the irregular firing of cortical neurons. Nat. Neurosci. 1, 210–217. doi:10.1038/659

Sturgill, J.F., Isaacson, J.S., 2015. Somatostatin cells regulate sensory response fidelity via subtractive inhibition in olfactory cortex. Nat. Neurosci. 18, 531–5. doi:10.1038/nn.3971

Sun, W., Tan, Z., Mensh, B.D., Ji, N., 2016. Thalamus provides layer 4 of primary visual cortex with orientation- and direction-tuned inputs. Nat. Neurosci. 19, 308–15. doi:10.1038/nn.4196

Tan, A.Y.Y., Brown, B.D., Scholl, B., Mohanty, D., Priebe, N.J., 2011. Orientation selectivity of synaptic input to neurons in mouse and cat primary visual cortex. J. Neurosci. 31, 12339–50. doi:10.1523/JNEUROSCI.2039-11.2011

Tan, Z., Hu, H., Huang, Z.J., Agmon, A., 2008. Robust but delayed thalamocortical activation of dendritic-targeting inhibitory interneurons. Proc. Natl. Acad. Sci. U. S. A. 105, 2187–92. doi:10.1073/pnas.0710628105

Taniguchi, H., Lu, J., Huang, Z.J., 2013. The spatial and temporal origin of chandelier cells in mouse neocortex. Science 339, 70–4. doi:10.1126/science.1227622

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., Bertagnolli, D., Goldy, J., Shapovalova, N., Parry, S., Lee, C., Smith, K., Bernard, A., Madisen, L., Sunkin, S.M., Hawrylycz, M., Koch, C., Zeng, H., 2016. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat. Neurosci. 19, 335–46. doi:10.1038/nn.4216

Theis, L., Berens, P., Froudarakis, E., Reimer, J., Román Rosón, M., Baden, T., Euler, T., Tolias, A.S., Bethge, M., 2016. Benchmarking Spike Rate Inference in Population Calcium Imaging. Neuron 90, 471–82. doi:10.1016/j.neuron.2016.04.014

Thomson, M.G. a., Foster, D.H., 1997. Role of second- and third-order statistics in the discriminability of natural images. J. Opt. Soc. Am. A 14, 2081. doi:10.1364/JOSAA.14.002081

Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. Nature 381, 520–2. doi:10.1038/381520a0

Tiesinga, P., Fellous, J.-M., Sejnowski, T.J., 2008. Regulation of spike timing in visual cortical

circuits. Nat. Rev. Neurosci. 9, 97–107. doi:10.1038/nrn2315

Tolhurst, D.J., Movshon, J.A., Dean, A.F., 1983. The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res. 23, 775–85.

Tolhurst, D.J., Smyth, D., Thompson, I.D., 2009. The sparseness of neuronal responses in ferret primary visual cortex. J. Neurosci. 29, 2355–70. doi:10.1523/JNEUROSCI.3869-08.2009

Torralba, A., Oliva, A., 2003. Statistics of natural image categories. Network 14, 391–412.

Tuncdemir, S.N., Wamsley, B., Stam, F.J., Osakada, F., Goulding, M., Callaway, E.M., Rudy, B., Fishell, G., 2016. Early Somatostatin Interneuron Connectivity Mediates the Maturation of Deep Layer Cortical Circuits. Neuron 89, 521–35. doi:10.1016/j.neuron.2015.11.020

Van Hooser, S.D., 2007. Similarity and diversity in visual cortex: is there a unifying theory of cortical computation? Neuroscientist 13, 639–56. doi:10.1177/1073858407306597

van Vreeswijk, C., Sompolinsky, H., 1996. Chaos in neuronal networks with balanced excitatory and inhibitory activity. Science 274, 1724–6.

Vinje, W.E., Gallant, J.L., 2002. Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. J Neurosci 22, 2904–2915. doi:20026216

Vinje, W.E., Gallant, J.L., 2000. Sparse coding and decorrelation in primary visual cortex during natural vision. Science (80-.). 287, 1273–1276.

Vinken, K., Vermaercke, B., Op de Beeck, H.P., 2014. Visual categorization of natural movies by rats. J. Neurosci. 34, 10645–58. doi:10.1523/JNEUROSCI.3663-13.2014

Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R., Paninski, L., 2010. Fast nonnegative deconvolution for spike train inference from population calcium imaging. J. Neurophysiol. 104, 3691–704. doi:10.1152/jn.01073.2009

Vu, E.T., Krasne, F.B., 1992. Evidence for a computational distinction between proximal and distal neuronal inhibition. Science 255, 1710–2.

Wall, N.R., De La Parra, M., Sorokin, J.M., Taniguchi, H., Huang, Z.J., Callaway, E.M., 2016. Brain-Wide Maps of Synaptic Input to Cortical Interneurons. J. Neurosci. 36, 4000–9. doi:10.1523/JNEUROSCI.3967-15.2016

Wang, Q., Burkhalter, A., 2007. Area map of mouse visual cortex. J. Comp. Neurol. 502, 339–57. doi:10.1002/cne.21286

Wang, Q., Gao, E., Burkhalter, A., 2011. Gateways of ventral and dorsal streams in mouse visual cortex. J. Neurosci. 31, 1905–18. doi:10.1523/JNEUROSCI.3488-10.2011

Wang, Q., Sporns, O., Burkhalter, A., 2012. Network analysis of corticocortical connections reveals ventral and dorsal processing streams in mouse visual cortex. J. Neurosci. 32,

4386–99. doi:10.1523/JNEUROSCI.6063-11.2012

Wehr, M., Zador, A.M., 2003. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. Nature 426, 442–446. doi:10.1038/nature02116

Wichmann, F. a, Braun, D.I., Gegenfurtner, K.R., 2006. Phase noise and the classification of natural images. Vision Res. 46, 1520–9. doi:10.1016/j.visres.2005.11.008

Wichmann, F.A., Gegenfurtner, K.R., 2010. Animal detection in natural scenes : Critical features revisited 10, 1–27. doi:10.1167/10.4.6.Introduction

Wilent, W.B., Contreras, D., 2005. Stimulus-dependent changes in spike threshold enhance feature selectivity in rat barrel cortex neurons. J. Neurosci. 25, 2983–91. doi:10.1523/JNEUROSCI.4906-04.2005

Wilson, N.R., Runyan, C. a, Wang, F.L., Sur, M., 2012. Division and subtraction by distinct cortical inhibitory networks in vivo. Nature 488, 343–8. doi:10.1038/nature11347

Xu, N., Harnett, M.T., Williams, S.R., Huber, D., O'Connor, D.H., Svoboda, K., Magee, J.C., 2012. Nonlinear dendritic integration of sensory and motor input during an active sensing task. Nature 492, 247–51. doi:10.1038/nature11601

Xue, M., Atallah, B. V, Scanziani, M., 2014. Equalizing excitation-inhibition ratios across visual cortical neurons. Nature 511, 596–600. doi:10.1038/nature13321

Yao, H., Shi, L., Han, F., Gao, H., Dan, Y., 2007. Rapid learning in cortical coding of visual scenes. Nat. Neurosci. 10, 772–8. doi:10.1038/nn1895

Yilmaz, M., Meister, M., 2013. Rapid innate defensive responses of mice to looming visual stimuli. Curr. Biol. 23, 2011–5. doi:10.1016/j.cub.2013.08.015

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.C., Jenvay, S., Miyamichi, K., Luo, L., Dan, Y., 2014. Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science (80-.). 345, 660–665. doi:10.1126/science.1254126

Zhang, W., Zhang, L., Liang, B., Schroeder, D., Zhang, Z.-W., Cox, G.A., Li, Y., Lin, D.-T., 2016. Hyperactive somatostatin interneurons contribute to excitotoxicity in neurodegenerative disorders. Nat. Neurosci. 19, 557–9. doi:10.1038/nn.4257

Zhu, Y., Qiao, W., Liu, K., Zhong, H., Yao, H., 2015. Control of response reliability by parvalbumin-expressing interneurons in visual cortex. Nat. Commun. 6, 1–11. doi:10.1038/ncomms7802

Zoccolan, D., 2015. Invariant visual object recognition and shape processing in rats. Behav. Brain Res. 285, 10–33. doi:10.1016/j.bbr.2014.12.053

Zoccolan, D., Oertelt, N., DiCarlo, J.J., Cox, D.D., 2009. A rodent model for the study of invariant visual object recognition. Proc. Natl. Acad. Sci. U. S. A. 106, 8748–53. doi:10.1073/pnas.0811583106

Zohary, E., Shadlen, M.N., Newsome, W.T., 1994. Correlated neuronal discharge rate and its implications for psychophysical performance. Nature 370, 140–3. doi:10.1038/370140a0

This page intentionally left blank

Chapter 3

How do stimulus correlations modulate reliable coding in V1[1]?

Summary

Intrinsic neuronal variability significantly limits information encoding in the primary visual cortex (V1). Certain stimuli can suppress this inter-trial variability to increase the reliability of neuronal responses. In particular, responses to natural scenes, which have broadband spatiotemporal statistics, are more reliable than responses to stimuli such as gratings. However, very little is known about which stimulus statistics modulate reliable coding and how this occurs at the neural ensemble level. Here, we sought to elucidate the role that spatial correlations in natural scenes play in reliable coding. We developed a novel noise masking method to systematically alter spatial correlations in natural movies, without altering their edge structure. Using high-speed two-photon calcium imaging *in vivo*, we found that responses in mouse V1 were much less reliable at both the single neuron and population level when spatial correlations were removed from the image. This change in reliability was due to a reorganization of between-neuron correlations. Strongly correlated neurons formed ensembles that reliably and accurately encoded visual stimuli, whereas reducing spatial correlations reduced the activation of these ensembles, leading to an unreliable code. Together with an ensemble-specific normalization model, these results suggest that the coordinated activation of specific subsets of neurons underlies the reliable coding of natural scenes.

Highlights

- Spatial correlations in natural scenes can be perturbed without changing salient image features.
- Response strength and reliability decrease as spatial correlations are removed.
- Spatial correlations restructure neuronal correlations to create ensembles with reliable responses.
- Neuronal ensembles carry out response normalization, a key mechanism for reliable coding.

[1] The findings of this chapter are published in The Journal of Neuroscience as "*Rikhye and Sur (2015) Spatial Correlations in Natural Scenes Modulate Response Reliability in Mouse Primary Visual Cortex*".

3.1 Introduction

A challenge faced by the visual system is to rapidly and accurately extract salient features from rich natural scenes while discarding redundant information. The fidelity with which visual information is processed is limited by both the intrinsic variability of neurons and the correlation structure of the network (Averbeck et al., 2006; Azouz and Gray, 1999; Moreno-Bote et al., 2014; Shadlen and Newsome, 1998). Despite these sources of noise, evidence from several studies indicates that natural scenes are processed efficiently in V1 (Olshausen and Field, 2004; Simoncelli and Olshausen, 2001). This suggests that response variability is reduced for natural scenes, permitting information to be represented efficiently and with high fidelity (Borst and Theunissen, 1999). How this is achieved in V1 remains unclear.

The degree of trial-to-trial variability in a response is commonly measured in terms of reliability. A neuron is said to be reliable if it fires the same number of precisely-timed spikes on every repetition of a stimulus (Tiesinga et al., 2008). Several physiological studies in cats have shown that both spike trains and sub-threshold potentials are reliable and sparse when stimulated with full-field natural movies (Baudot et al., 2013; Haider et al., 2010). This coding strategy increases the amount of information conveyed per spike (Pecka et al., 2014; Vinje and Gallant, 2002). In contrast, vignetting these movies such that they only stimulate the classical receptive fields of neurons degrades both reliability and sparseness. These results hint at a paradoxical population-coding regime where response variability decreases when larger numbers of neurons are activated (Renart and Machens, 2014; Shadlen and Newsome, 1994). Supporting this idea, both attention (Cohen and Maunsell, 2009) and neuromodulatory mechanisms (Goard and Dan, 2009), which change correlations between neurons, also improve response reliability. However, the relationship between inter-neuronal correlations and reliable coding of natural scenes remains unexplored.

Apart from population coding, stimulus statistics are also known to influence reliability. A recent study showed that reliability was highest for natural scenes and

weakest for simple stimuli, such as gratings and dense noise (Baudot et al., 2013). Statistically, sinusoidal gratings only contain one spatial frequency (SF), whereas dense noise contains a spectrum of different SFs with constant power spectral density. This is in contrast to natural scenes, which contain a power law distribution of SFs. Particularly; most natural scenes have higher power in the low SF bands (broad image features) and weaker power in high SF bands (fine structural details) (Bar, 2004; Ruderman and Bialek, 1994). It has been hypothesized that the visual system uses these unique statistical properties to efficiently encode information (Barlow, 2001; Simoncelli and Olshausen, 2001). This theory is bolstered by several psychophysical studies, which demonstrate that image statistics are critical for rapid and accurate discrimination (McCotter et al., 2005; Torralba and Oliva, 2003). However, how stimulus statistics modulate the reliability, and in turn, the efficiency of the neural code remains unknown. One hypothesis is that specific features of natural scenes improve response reliability by reducing correlated variability (Kohn and Smith, 2005; Snyder et al., 2014). To test this hypothesis, it is important to parametrically relate stimulus statistics to changes in population coding and neural reliability.

In this study, we sought to answer the following two questions. **(1)** What properties of natural scenes influence response reliability? **(2)** What population coding mechanisms improve reliability? To address these questions, we developed a novel stimulus set where we systematically perturbed spatial correlations in natural movies, while preserving their edge structure. By applying *in vivo* two-photon calcium imaging to layer 2/3 neurons in mouse V1, we found that spatial correlations in natural movies strongly influenced response reliability. Cluster analysis of neuronal responses showed that reliable coding was achieved via a stimulus-driven restructuring of inter-neuronal correlations. These findings, supported by a normalization model, demonstrate that a hallmark of reliable coding is the coordinated activation of specific neuronal ensembles.

3.2 Predictions and Hypothesis

We first sought to understand the relationship between natural scene statistics and neuronal processing in mouse V1. To this end, we presented multiple repetitions of gray-scale natural movies (five in total) to awake, passively viewing mice, while simultaneously performing high-speed calcium imaging (see Appendix 3A for methods). Interestingly, we discovered a significant difference in the number of reliably responding neurons recruited by each movie (**Figure 3-1A**). For instance, Movie-5 reliably drove only 8% of all recorded neurons (88/1,011), whereas Movie-4 elicited reliable responses from 28% of the recorded neurons (310/1,011). In addition, we observed a significant difference in pyramidal cell reliability between these five movies. Specifically, Movie-5 always evoked responses with the lowest reliability (**Figure 3-1B**, $F_{(4,5466)}= 24.86$, $p < 10^{-8}$ Bonferroni-corrected one-way ANOVA).

What could the source of this between-movie variability in response reliability be? Despite having the same mean luminance and contrast, each movie had unique structural properties (**Figure 3-1C, D**). These structural differences arise due to differences in spatial content, which arise due to local differences in edge and texture distributions within each image. To determine if these structural differences were sufficient to explain the observed difference in response reliability, we performed linear regression between differences in the structural similarity index and reliability for all movie pairs. Although movies that were most different tended to have bigger differences in reliability, we found no significant correlation between SSIM and reliability (**Figure 3-1E**). For example, we found that, although Movies-3 and -5 were most similar, Movie-3 was more reliably processed than Movie-5. Together, these suggest that local structural differences in luminance and contrast are unable to explain the movie-wise difference in reliability.

Next, we first sought to investigate the relationship between the SF spectrum of a movie and pyramidal cell reliability. To do so, we extracted two measures from the SF power spectral density: (1) the slope of the power spectrum and (2) the spatial power,

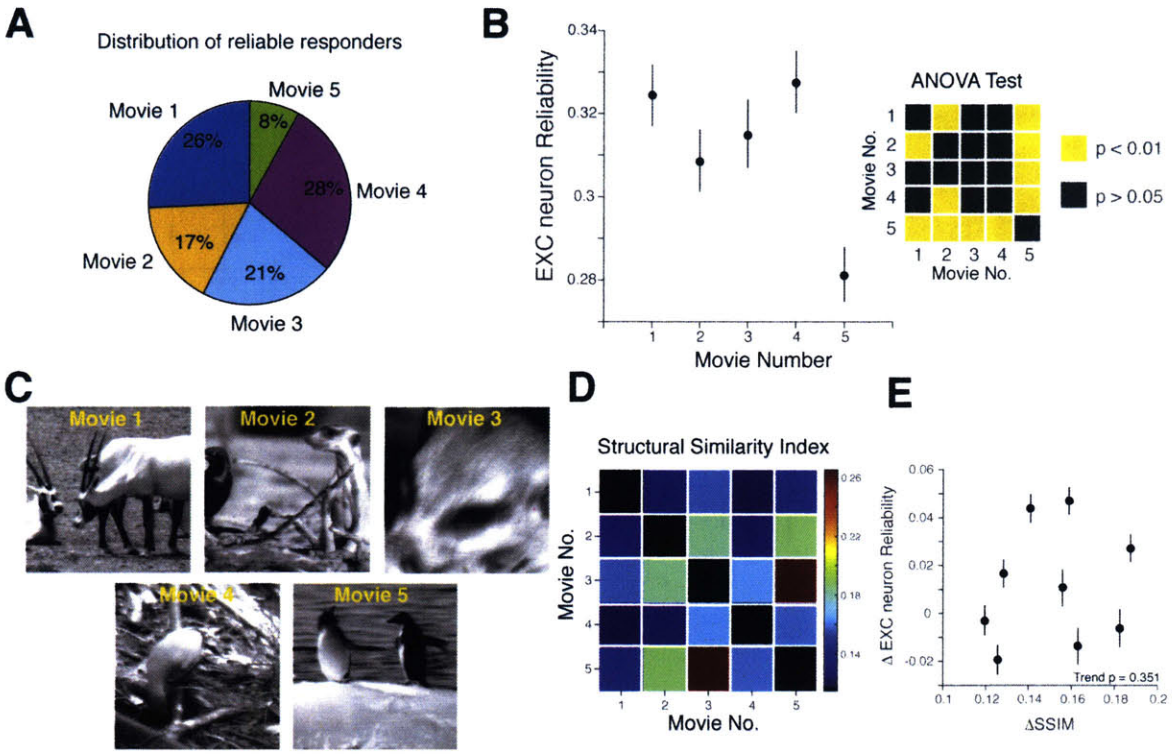

Figure 3-1. Diverse response reliability values between natural movies. (A) Pie chart showing distribution of reliably responding neurons (Rel > 0.25) for each of the five movies. **(B)** Quantification of mean reliability for each of the 5 movies. Movie-5 has significantly lower reliability than all other movies (Bonferroni-corrected one-way ANOVA). **(C)** Example movie frames. **(D)** Matrix showing pair-wise structural similarity between the movies. The higher the similarity index the more similar the movies are. **(E)** Regression analysis showing that differences in reliability cannot be explained from structural differences alone. Data collected from 10 mice (1,101 neurons). Error-bars, 95% CI of mean.

which is obtained by integrating the power spectral density (**Figure 3-2A**). The SF spectral slope provides a measure of the ratio of the low SF content relative to the high SF content. In general, most natural movies have more power in the low SF bands than in the high SF bands, resulting in slope values between 1.5-2 (**see Chapter 2**). On the other hand, the spatial power provides a measure of the total SF content within that image. We limited the integral to the bands relevant to the visual acuity of mice. Using multivariate linear regression analysis between pyramidal neuron reliability and these measures, we observed that the spectral slope was a better predictor of the differences in reliability ($p = 0.067$, F-test) than the spatial power (**Figure 3-2B**). In essence, this analysis implies that movies with stronger low SF content than high SF content (i.e. a steeper slope) should be more reliably processed.

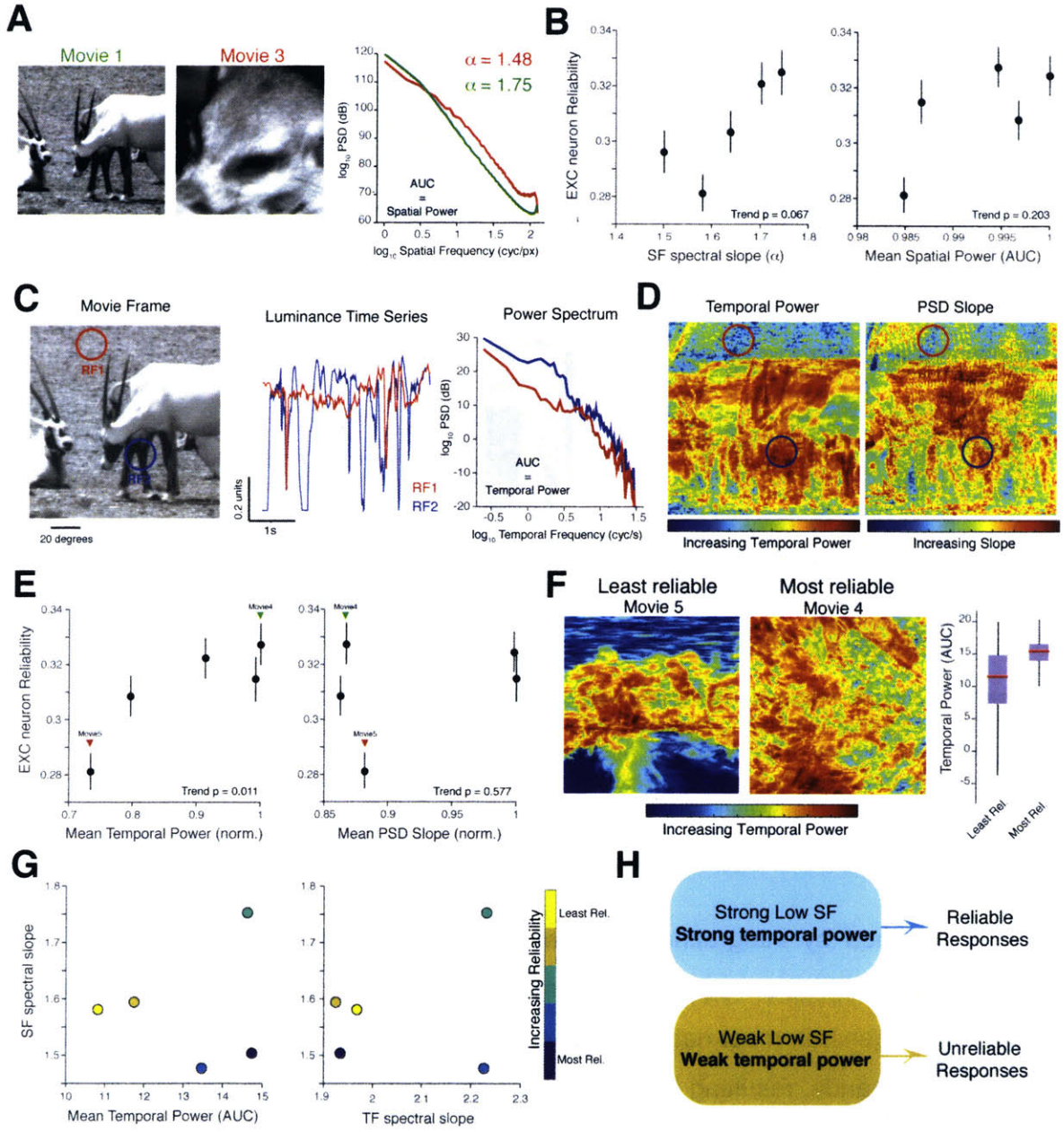

Figure 3-2. The influence of spatial and temporal statistics on reliability. (A) Representative example showing the spatial frequency (SF) power spectral density of two natural movies. The PSD was integrated over the gray rectangle to compute the total spatial content. **(B)** Regression analysis showing that SF spectral slope, but not total spatial content, is a good predictor of reliability. **(C)** Analysis of temporal properties of natural movies. Red receptive field indicates a region of the movie that remains constant over time, whereas blue receptive field indicates pixels that change rapidly over time. The right most plot shows differences in the temporal frequency (TF) PSD between these receptive fields. **(D)** Map of temporal power and PSD slope, showing regions of the movie in **(C)** that change the most over time. **(E)** Regression analysis showing that total temporal content is a good predictor of reliability. **(F)** Comparison between the temporal power maps of least reliable and most reliable movies. **(G)** Multivariate regression results. **(H)** Summary. Data same as **Figure 3-1**.

In addition to being correlated over space, pixels in natural movies are strongly correlated over time (Dan et al., 1996; Dong and Atick, 1995). These temporal differences are known to influence the reliability of neurons in the LGN (Dan et al., 1996) and in the visual cortex (Baudot et al., 2013). Thus, we next sought to determine if these observed differences in reliability could be explained by differences in temporal content between the movies. To analyze temporal content, we first parceled the each movie into 20 degree receptive fields with a 10 degree overlap, as shown in **Figure 3-2C**. Within each movie, we noticed regions where luminance did not change much over time (red line, **Figure 3-3C**) and also regions where luminance changed rapidly over time (blue line, **Figure 3-3C**), with distinct high and low temporal frequency (TF) components. We quantified the TF content of this luminance change by first computing a TF power spectral density for each RF location (**Figure 3-3C**) and then extracting TF spectral slope and total temporal power information from each power spectrum. As shown in **Figures 3-3C** and **3-3D**, regions of the image with large luminance fluctuations in general have a higher total temporal power, but do not necessarily have a steeper TF spectral slope. Thus, this suggests that the total temporal power is a good way to discriminate movie regions that change rapidly over time from those that remain relatively static. Again using multivariate linear regression, we found that total temporal power ($p = 0.011$, F-test) was a much better predictor of pyramidal neuron reliability than the TF spectral slope (**Figure 3-3E**). This implies that neurons with receptive fields over dynamically changing parts of a movie will respond more reliably than neurons sampling from relatively static parts of the movie. In agreement with this idea, we found that the least reliably processed movie (Movie-5) had a smaller fraction of dynamically changing pixels than the most reliably processed movie (Movie-4, **Figure 3-3F**). Importantly, this analysis neatly demonstrates the relationship between the temporal statistics of the stimulus and the reliability of neural responses in mouse V1 as predicted by Mainen and Sejnowski (1995).

Finally, we augmented our regression model to include interaction terms between both spatial and temporal properties of these movies. This allowed us to investigate the predictive power of spatiotemporal stimulus statistics on reliable coding. Interestingly,

we found that interactions between SF spectral slope and total temporal power ($p < 0.012$, F-test) were sufficient to explain the trend in reliability (**Figure 3-2G**). In contrast, TF spectral slope and total spatial content were poor predictors of neuronal reliability. Taken together, this regression analysis demonstrates that both spatial and temporal properties of natural movies are responsible for creating reliable responses. In particular, movies with strong low SF content (steeper spectral slope) and dynamically changing luminance over time will be more reliably processed. Thus, reliable firing is the consequence of the unique spatial and temporal properties of visual stimuli.

Hypothesis

Our regression models predict that movies broadband spatiotemporal statistics, in particular large fluctuations in temporal dynamics will be more reliably processed than stationary stimuli. It has already been very well established that the efficiency temporal integration by sensory neurons depends strongly on the statistics of the incoming sensory stimuli (Atick and Redlich, 1990). For example, in now what has become a classic experiment, Mainen and Sejnowski (1995) demonstrated neurons integrated temporally irregular current waveforms to produce reliable and precise spike responses, whereas responses to constant current steps were much more variable (Mainen and Sejnowski, 1995). Importantly, this experiment demonstrates that temporally fluctuating stimuli elicit reliable responses because neurons are unable to adapt to these stimuli. However, this result does not immediately transpose to the *in vivo* condition where stochastic synaptic barrages result in much higher trial-to-trial variability (Stevens and Zador, 1998). In an analogous experiment, Baudot and colleagues demonstrated that natural movies, which are spatiotemporally broadband stimuli, are more reliably processed than gratings via a reduction in noise (measured here as the unexplained variance, see **Chapter 2**). These authors argued that broadband stimuli recruit more neurons and allows the network to explore more dynamic regimes, which is in turn important for efficient coding (Borst and Theunissen, 1999). Interestingly, this work also implies that saccadic eye movements can impose temporal nonlinearities in the input signal to V1 neurons, and thus can function to improve the reliability of neural responses

(Segal et al., 2015). Also, the lack of broadband temporal dynamics could explain why flashed natural images do not evoke reliable responses in ferret V1 (Tolhurst et al., 2009).

Both these results underscore the importance of rapidly fluctuating temporal dynamics in the stimulus for neural coding. These luminance fluctuations are also known to drive adaptation (Hosoya et al., 2005) and influence the reliability of neurons in the early visual system (Desbordes et al., 2008). Importantly, fluctuations that occur within the timescale of neuronal integration allow neurons to produce spike trains with millisecond precision (Butts et al., 2007). We will explore this further in **Chapter 4**. Interestingly, in addition to temporal statistics, our regression modeling also demonstrates that the spatial statistics of the stimulus are also important in modulating the reliability of neural responses. We know that the visual cortex is retinotopically organized, with neighboring neurons sampling similar parts of the visual scene (see references in **Chapter 2**). Thus, due to the strong spatial correlations in natural scenes, natural movies are likely to co-activate numerous neurons simultaneously. This spatially coordinated form of population coding could also work to reduce shared noise between neurons.

Based on these findings, we hypothesize that, while temporal statistics coordinate activity in time, spatial statistics coordinate activity in space. This allows the formation of neuronal ensembles. This strategy permits the cortex to optimize its neural code by multiplexing over multiple filter sets to cooperatively reduce redundancy in the stimulus. However, the role of spatial statistics in coordinating modulating population coding in V1 has been poorly studied. In this chapter, we investigate these mechanisms by selectively perturbing spatiotemporal correlations in natural movies.

3.3 Results

Noise-masking: a novel method to alter spatial correlations in natural scenes

With the aim of understanding the properties of natural scenes that influence reliability, we developed a novel noise-masking technique that allowed us to selectively increase or decrease second-order spatial correlations in time-varying natural scenes without altering their underlying edge structure (**Figure 3-3**). To do so, we took advantage of the fact that power (P) in most natural images decreases with increasing SF (k) according to the power law: $P \sim k^{-\alpha}$, where the parameter α is the spectral slope (van Hateren and van der Schaaf, 1998). White-noise images, for example, lack correlations between pixels, possess constant power in all SF bands and have a zero spectral slope. Thus, the spectral slope can succinctly parameterize both spatial correlations and the distribution of SFs in natural images. Furthermore, because the phase spectrum contains information about edges, spatial correlations can be altered without significantly changing the saliency of an image (Wichmann et al., 2010). Thus, in our noise-masking technique, we preserved the phase spectrum and added noise masks with precisely defined spectral slopes to each frame of a set of natural movies (**Figure 3-3A**, 5 movies, see Appendix 3A for details).

We used this method to generate movies with four different spectral slopes: 0, 1, 1.5 and 2, henceforth referred to as *K0 – K2* respectively. Example frames from these noise-masked movies together with their power spectra are shown in **Figure 3-3B**. Since changes in pixel statistics, especially mean luminance and contrast, are known to affect coding in the visual system (Bonin et al., 2006), we adjusted each noise-masked frame to have the same mean luminance, contrast, kurtosis and skewness (**Figure 3-3C**). As a result, our noise masked movies differed only in their spectral slope and not in their relative distribution of light and dark pixels. For example, *K0* movies possessed no spatial correlations between pixels and, consequentially, had a flat power spectrum (**Figure 3-3B**) and a flat spatial autocorrelation function (ACF, **Figure 3-3D**). Visually,

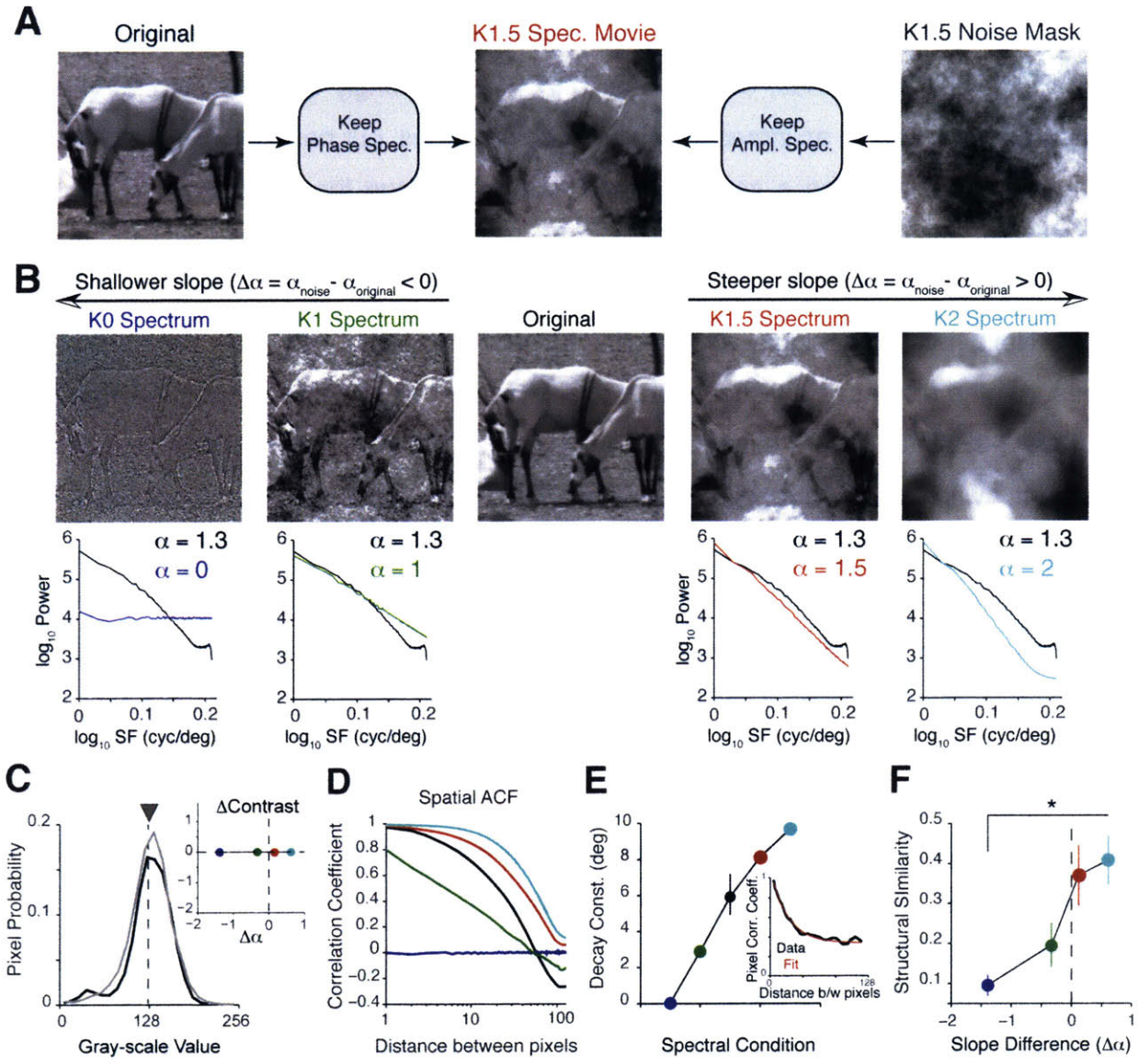

Figure 3-3. Perturbing spatial correlations in natural movies. (A) Illustration of noise-masking procedure. Briefly, we combined the Fourier phase spectrum of a natural movie frame with the amplitude spectrum of a noise movie to create a noise-masked movie frame. **(B)** *Top,* Example frames from noise-masked movies with four different levels of spatial correlations. *Bottom,* Comparison between the power spectra of these movies (colored lines) and the original movie (black line). **(C)** Pixel intensity distribution from one original movie (black) and its noise-masked variants (gray). All noise-masked movies were adjusted to have the same luminance (gray triangle), contrast (see inset), kurtosis and skewness as the original movie. As a result, all noise-masked movies had the same pixel intensity distribution. *Inset,* Plot of change in contrast against difference in spectral slope, computed relative to the original movie. **(D)** Representative spatial autocorrelation functions (ACF) of the different noise-masked movies. **(E)** Percentage change in the ACF integral of the noise movies relative to the original movie. **(F)** Structural Similarity index (SSIM) between noise-masked movies and their original versions. SSIM was averaged over frames for each movie (1200 frames). P-value for **(F)** computed using Cochran-Armitage test for trend. All error bars indicate SEM. Colors indicate noise-masked movies labeled in (**B**). Data from 5 different movies.

edges in *K0* movies appeared sharper than the original movie because of an increase in high SF power. *K2* movies, on the other hand, were strongly spatially correlated and appeared blurry due to an increase in low SF power. Thus, our noise-masking technique is equivalent to applying a zero-phase SF filter to change the SF content of natural movies. Specifically, in *K0* and *K1* movies, low SF content is attenuated, while high SF content is attenuated in *K1.5* and *K2* movies.

Next, by fitting single exponential functions to each ACF, we parametrically measured the distance, in visual space, over which pixels remained strongly correlated (**Figure 3-3E**, inset shows example fit). As expected, *K2* movies had the largest decay constant, with pixels strongly correlated with each other up to a distance of approximately 10 degrees. In contrast, pixels in *K1* movies became decorrelated beyond 3 degrees. Thus, pixels in both *K0* and *K1* movies decorrelate within the span of a typical mouse V1 receptive field [5-7 degrees, (Niell and Stryker, 2008)]. Notably the remarkably low variability in noise-masked movie data points reinforces the fact that these movies are spatially homogeneous: all noise-masked movies have the same power spectra regardless of differences in their phase spectrum.

In addition, we also assessed changes in image quality by computing a frame-wise structural similarity index (SSIM) between each noise movie and its original version. The SSIM uses image structural information, such as luminance and contrast, to assess the degree of similarity between two images. Images that are similar in appearance have a SSIM close to one, while statistically dissimilar images have a SSIM close to zero. We found that SSIM decreased as the difference in spectral slope became more negative ($p = 0.036$, Cochran-Armitage test for trend). This implies that *K0* and *K1* movies were the most dissimilar to their original version despite having edges that were visually sharper. Thus, our noise-masking technique revealed that perturbing the spectral slope alters the statistical structure of natural scenes by changing correlations between pixels.

Neurons respond unreliably to movies with decorrelated spectral properties.

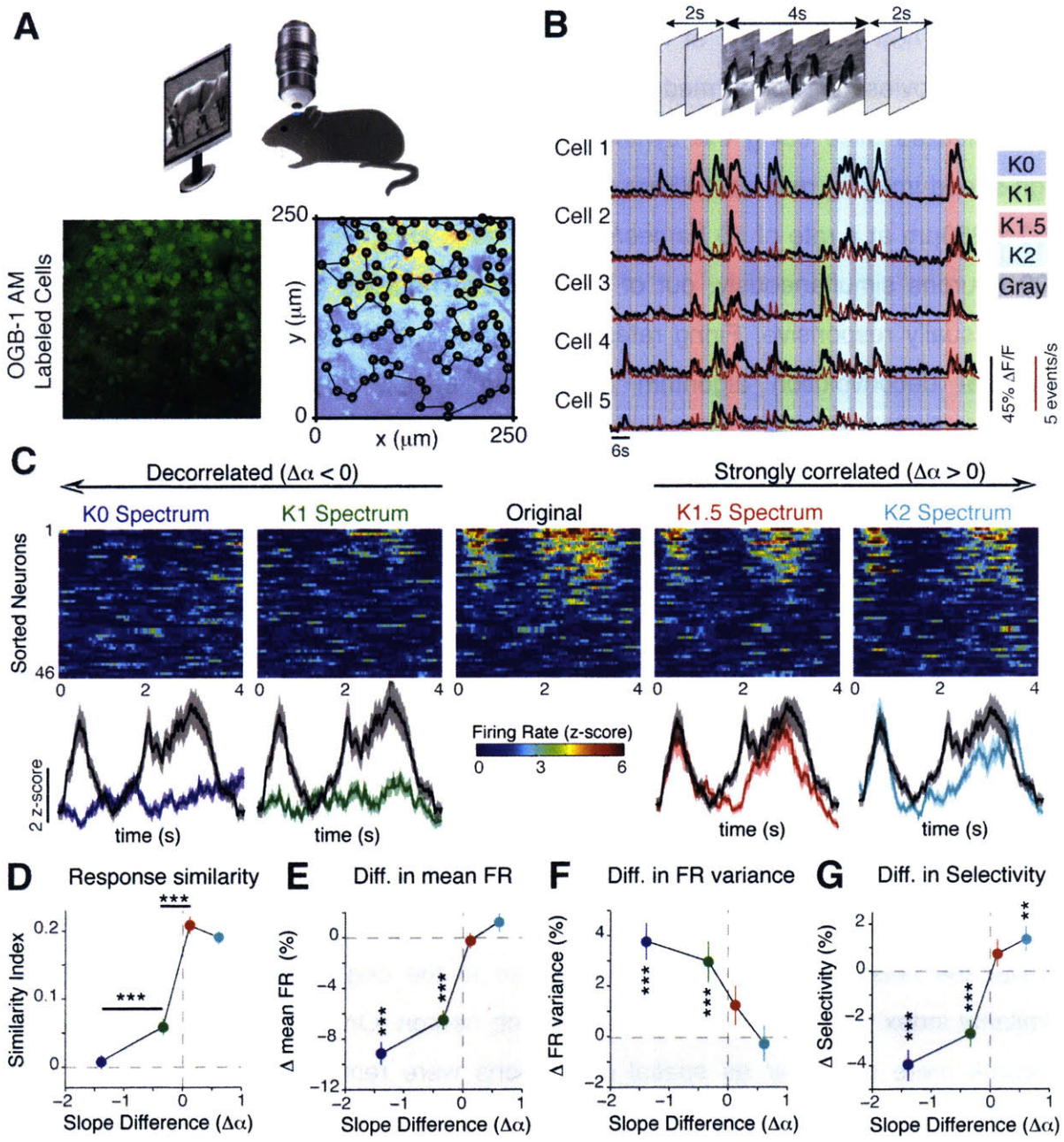

Figure 3-4. V1 neurons are sensitive to changes in spatial correlations. (A) *Top,* Schematic of experimental set up. *Bottom left,* Representative neural population labeled with OGB-1 AM. *Bottom right,* Neurons were detected with an automatic image segmentation algorithm and scanned with a custom line scan path (see Materials and Methods for details). **(B)** *Top,* Illustration of stimulus timing. Each movie was presented for 4s and was flanked by 2s gray screens. *Bottom,* Example calcium transients (black) and deconvolved firing rates (red) from 5 cells in response to a randomized sequence of stimuli. Each colored bar denotes the 4s period over which a movie was presented. **(C)** *Top,* Heat map showing trial-averaged firing rates (z-scored relative to spontaneous activity) from a population of 46 neurons. Neurons were sorted according to the mean response to the original movie and the order was preserved in each heat map. *Bottom,* Population-averaged responses. Shaded area denotes SEM.

To assess how populations of neurons in layer 2/3 of V1 represented these noise-masked movies, we preformed high-speed two-photon calcium imaging in lightly anaesthetized mice using the synthetic calcium indicator OGB1 (**Figure 3-4A**). Our high-speed imaging method (Wilson et al., 2013) allowed us to scan a cortical area of 250 x 250μm at a rate of 50 frames/s. Within each area, we were able to record up to 100 neurons simultaneously, out of which 46±20 neurons (21 populations, 16 mice) were visually responsive. Firing rates were inferred from the calcium signals using a temporal deconvolution algorithm (Vogelstein et al., 2010) (**Figure 3-4B**). The parameters used for accurate deconvolution were determined in previous studies (El-Boustani and Sur, 2014; Wilson et al., 2012). Expectedly, we found a significant correlation between inferred firing rate and average fluorescence change for all spectral conditions, implying a strong correspondence between the measured calcium signal and inferred firing rate.

The majority of neurons responded strongly and synchronously to the original movie at distinct epochs (**Figure 3-4C**, *center panel*). Surprisingly, the same neurons responded weakly, and at random times, to the decorrelated stimuli (*K0* and *K1* movies), which resulted in flat population-averaged responses (**Figure 3-4C**, *bottom row*). In contrast, these neurons responded to *K1.5* and *K2* movies with similar amplitudes and during the same epochs as the original movie. We assessed how similar the noise movie firing dynamics were to the original movie by computing a similarity index (SI, see Appendix 3A) for each neuron. On average, neural responses became more dissimilar as spatial correlations were removed from the movie, with responses to *K0* movies being the most different ($p < 10^{-6}$, Friedman test, 1006 neurons, 16 mice, **Figure 3-4D**). The SI was also significantly lower for responses to *K0* movies than responses to *K1* movies ($p = 8.12 \times 10^{-5}$, Bonferroni-corrected rank-sum test). In

← **Figure 3-2 (Continued). (D)** Firing rate similarity index between neural responses to noise-masked movies and the original movie as a function of the change in spectral slope.
(E-G) Percent change in mean firing rate (FR, **E**), trial-to-trial FR variance (**F**) and selectivity (**G**) relative to the original movie. Data are presented as median ±95% confidence interval (CI) from 1006 neurons (16 mice). All P-values computed via Friedman's test followed by Bonferroni-corrected rank-sum tests. Colors labeled in (**C**).

contrast, both *K1.5* and *K2* movies evoked very similar responses to the original movie, even though they had stronger spatial correlations ($p = 0.27$, Bonferroni-corrected rank-sum test). When compared with image SSIM (**Figure 3-3D**), these observations suggest that neurons in mouse V1 are highly sensitive to subtle changes in natural movie spatial correlations.

Flattening the power spectrum strongly decreased the mean firing rate of *K0* and *K1* movies relative to the original movie ($p < 10^{-6}$, Friedman test with Bonferroni correction, 16 mice, **Figure 3-4E**). Surprisingly, even though *K1.5* and *K2* movies had stronger spatial correlations than the original movie, they failed to evoke stronger firing rates (*K1.5*: $p = 0.95$, *K2*: $p = 0.07$, Bonferroni-corrected rank-sum tests). In addition, we found that the trial-to-trial variance for *K0* and *K1* movies also increased ($p < 10^{-6}$, Friedman test with Bonferroni correction, 16 mice, **Figure 3-4F**). Thus, V1 neurons have reduced and highly variable responses when spatial correlations are removed from natural movies.

Next, we computed lifetime sparseness from the inferred firing rates as a measure of selectivity to the different stimuli. By definition, a neuron with high lifetime sparseness responds selectively to only a few frames in the stimulus (Willmore et al., 2011). Not surprisingly, *K0* and *K1* movies evoked highly unselective responses, while increasing stimulus correlations increased lifetime sparseness (*K1.5*, $p = 0.23$; *K2*, $p = 2.1 \times 10^{-3}$, rank-sum test relative to original movie, 1006 neurons, **Figure 3-4G**). Thus, the selectivity of V1 neurons is also strongly modulated by stimulus spatial correlations. Taken together, these observations support the conclusion that attenuating the low SF power of natural scenes has a detrimental effect on neural responses.

Decorrelating movies reduces response reliability

We next asked to what degree removing spatial correlations affected response reliability. Reliability is traditionally measured in terms of the number of spikes produced by a neuron on every stimulus repetition (Tiesinga et al., 2008). However, because we did not have single spike resolution in our recordings, we reasoned that a highly reliable

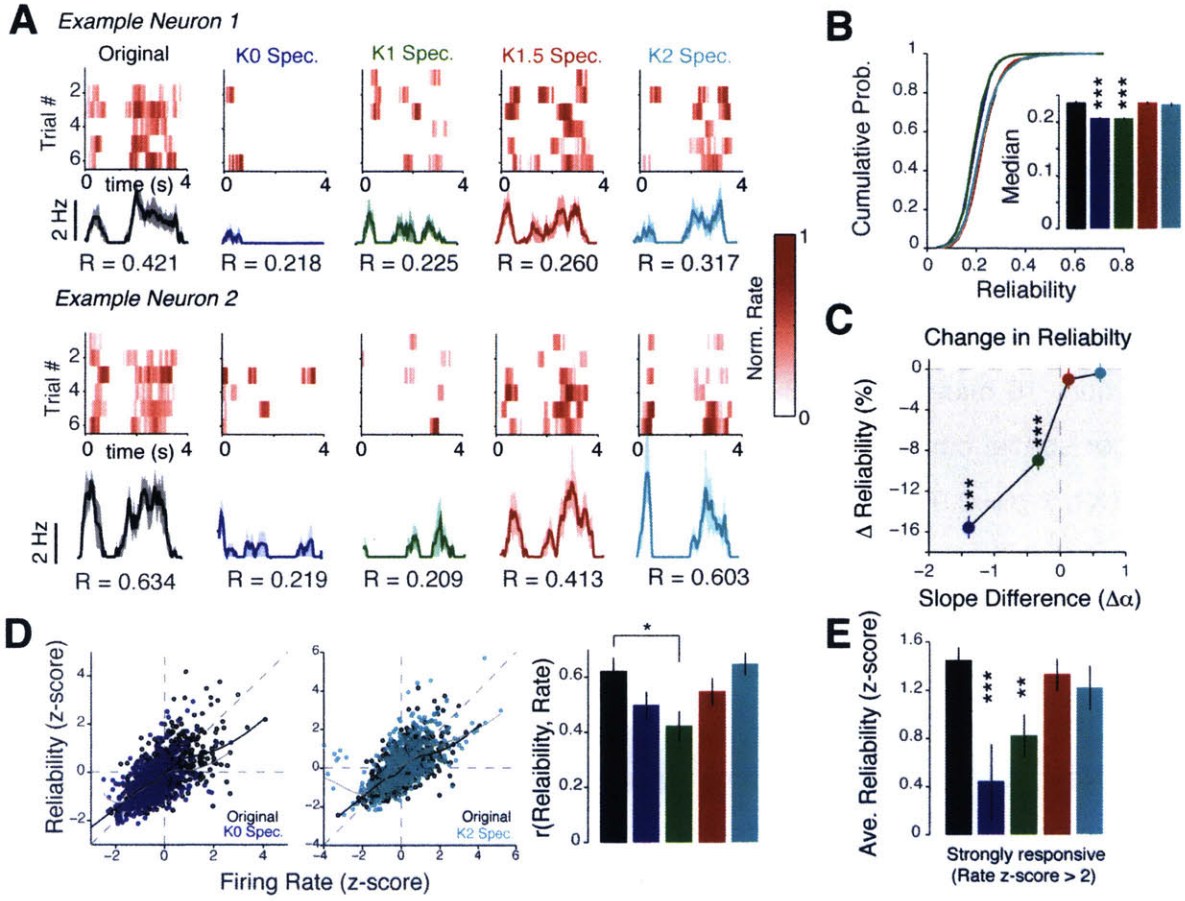

Figure 3-5. Spatially decorrelating natural movies decreases trial-to-trial reliability. (A) Example raster plots from two neurons, where each row is the normalized response to a single presentation of a movie. The trial-averaged response is shown below together with the reliability (R) values. Shaded area indicates SEM. **(B)** Cumulative distribution of reliability values. Inset shows bootstrapped estimates of the median reliability value. Data from 650 neurons from 16 mice. **(C)** Percent change in reliability relative to the original movie. Data presented as median±95% CI. **(D)** Scatter plot showing the relationship between reliability and mean firing rate, illustrated for *K0* (*left*) and *K2* movies (*middle*). The thick black and gray lines are the LOESS regression line for the original movie and the noise movie respectively. *Right*, Quantification of Spearman rank correlation coefficient between response reliability and mean firing rate for all noise movie conditions. Only *K1* movies had a significantly smaller correlation ($p = 0.035$, two-tailed rank-sum test). **(E)** Quantification of response reliability within strongly responding neurons. Strongly responding neurons were defined as neurons, which had a firing, rate more than 2 standard deviations above the original movie (z-score > 2). Reliability was similarly calculated to allow us to pool data from different experiments. Data in **C** - **E** from 650 neurons. P-values computed using Bonferroni-corrected two-tailed rank-sum tests relative to original movie. Error bars in **D** and **E** denote SEM.

neuron would respond with stereotypical responses on every trial and, as a result, would have a high, positive correlation coefficient between trials. Thus, we defined response reliability as the average Pearson's correlation coefficient between every pair-wise combination of trials (see Appendix 3A). Since this metric quantifies the degree of

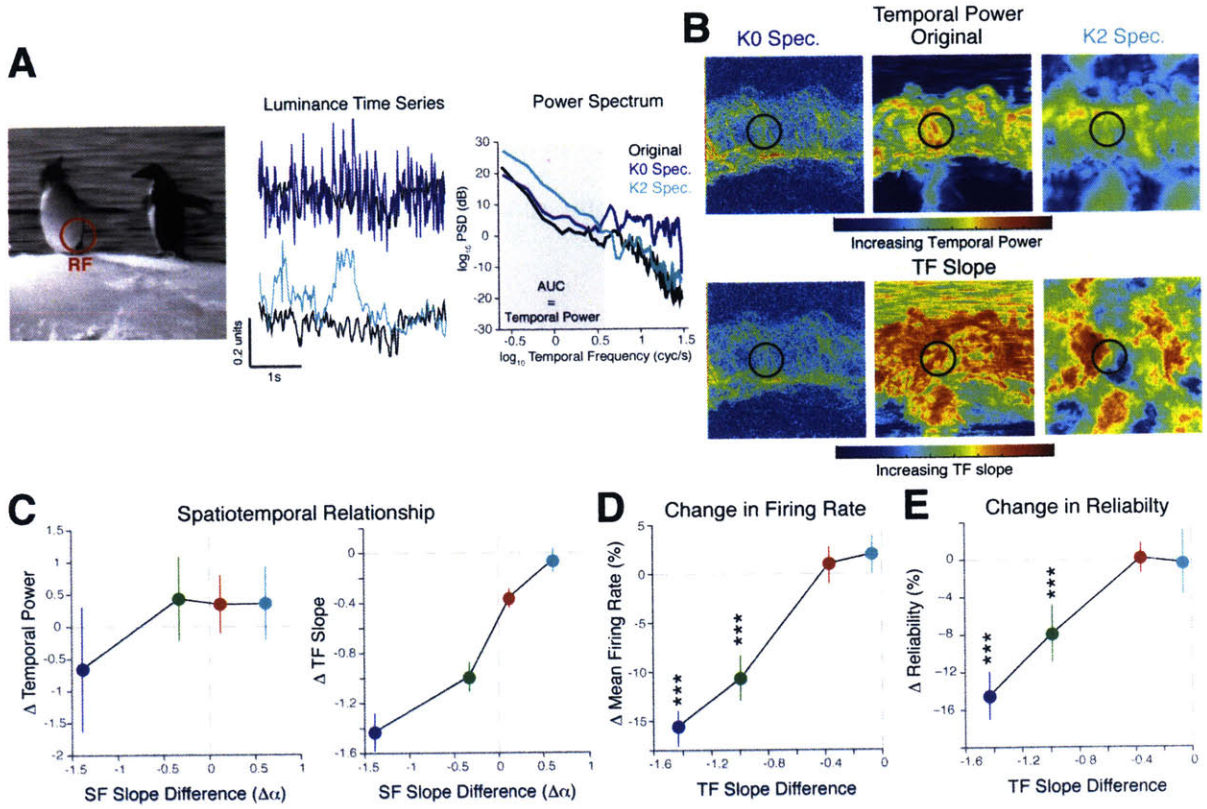

Figure 3-6. A change in spatiotemporal correlations explains the reduction in reliability. (A) Analysis of temporal frequency (TF) content of luminance changes. **(B)** Representative example showing the difference in the total temporal power (top) and TF slope (below) for K0 and K2 spectrum movies relative to the original movie. **(C)** *Left,* No significant difference in total temporal power caused by noise masking. *Right,* reducing correlations spatially between pixels also causes a temporal decorrelation. **(D-E)** Movies with reduced spatiotemporal correlations elicit weaker (D) and more unreliable responses (E) than movies with strong spatiotemporal correlations.

response similarity between trials, it provides a measure of both reliability and temporal precision (Bathellier et al., 2012; Baudot et al., 2013; Haider et al., 2010).

Neurons responded reliably to the unperturbed movies, as well as their *K1.5* and *K2* spectrum versions (**Figure 3-5A**). Relative to the original movie, we found no difference between *K1.5* and *K2* movie reliability (*K1.5*, $p = 0.565$; *K2*, $p = 0.624$, Bonferroni-corrected rank-sum test, **Figure 3-5B**). In contrast, *K0* and *K1* movies evoked responses with significantly lower reliability ($p < 10^{-6}$, Friedman Test, 650 neurons, 16 mice, **Figure 3-5B**) due to the high variability between trials. Thus, removing spatial correlations and attenuating low SF power caused a strong monotonic decrease in response reliability (**Figure 3-5C**). In contrast, attenuating high SF power

did not change reliability. Interestingly, we found an increase in the fraction of reliable responders to *K2* movies (*Orig.* vs. *K2*: 15.7% vs. 30.9%, $p = 3.4 \times 10^{-3}$, Kruskal-Wallis one-way ANOVA with Bonferroni correction, 16 mice). These results suggest that, for the same absolute level of reliability, slightly different sets of neurons were reliably activated by either a *K1.5* or *K2* movie.

Was the reduced reliability to decorrelated movies a consequence of reduced firing rates? We found a linear relationship between average firing rate and reliability, for all stimulus conditions (Spearman's rho = 0.65±0.22, $p = 8.9 \times 10^{-3}$, 650 neurons, **Figure 3-5D**). Partitioning data according to firing rates, we found a small fraction of neurons that responded strongly to both *K0* and *K1* movies (132/650, z-score > 2σ above original movie response). Interestingly, despite having higher firing rates, these neurons still responded less reliably to both decorrelated movies (*K0*: $p < 10^{-4}$, *K1*: $p = 0.035$, 132 neurons, rank-sum test, **Figure 3-5E**).

In addition to spatial properties, the temporal dynamics of visual stimuli are also known to strongly influence coding along the visual pathway. For example, rapid changes in contrast evoke changes in the gain and temporal dynamics of neurons in retina (Hosoya et al., 2005) and LGN (D. A. Butts et al., 2007; Lesica et al., 2007). Thus, we sought to investigate how the relationship between temporal statistics in these noise masked movies and neuronal reliability. To analyze the temporal statistics of these movies, we computed temporal frequency (TF) power spectral densities from pixel-wise luminance intensity changes as shown in **Figure 3-6A**. For the example shown, we notice that pixels *K0* movies have slightly higher high TF than the original movie, whereas *K2* movies have slightly higher low TF. Surprisingly, even though a different random noise mask was used for each frame, *K2* movies had more regions of temporally correlated pixels than *K0* movies (**Figure 3-6B**). On average, we found no significant difference in the total temporal power (integrated power spectral density) between the noise movies and the original movie (**Figure 3-6C**). Instead, we noticed that spatially decorrelated movies had a smaller TF spectrum slope than the original movie, which is indicative of an increase in high TF power. Together, these results

indicate that noise masking not only changes the spatial statistics but also changed the temporal statistics of natural movies.

As expected, we found a monotonic decrease in both mean firing rate (**Figure 3-6D)** and reliability (**Figure 3-6E**) with increasing TF slope difference relative to the original movie. Taken together, these results suggest that *K0* and *K1* movies evoke unreliable responses because they do not possess the appropriate spatial and temporal spectral content. Thus, reliable coding of natural scenes requires stimuli that have intact low SF and low TF information.

Stimulus-induced changes in neuronal responses in awake mice

Our results so far show that the response properties of neurons in mouse V1 are sensitive to the spectral properties of the visual stimulus; specifically, removing spatial correlations decreases both response amplitudes and between-trial reliability. Since anesthesia is known to influence cortical computations (Haider et al., 2013), we next sought to extend the generality of these findings to awake mice by repeating the same experiments in passively viewing, head-fixed mice using the genetically encoded calcium indicator GCaMP6f instead of OGB-1 (**Figure 3-7A**, see Appendix 3A).

Neurons in awake mice also responded to both *K0* and *K1* movies with significantly weaker responses than the original movie ($p < 10^{-3}$, Friedman test with Bonferroni correction, 230 neurons, see examples in **Figure 3-7B**). Surprisingly, firing rates to *K1* movies were slightly higher than in anesthetized mice ($p = 6.7 \times 10^{-3}$, one-tailed rank-sum test relative to anesthetized data). However, despite this increase in firing rate, *K1* movie responses were significantly less reliable (**Figure 3-7C**) and unselective (**Figure 3-7D**) than responses to original movies. Similar to data from anesthetized mice, we also found that removing spatial correlations led to a monotonic decrease in response similarity in awake mice (**Figure 3-7E**).

Thus, our awake imaging experiments show that weak and unreliable responses to decorrelated stimuli are unlikely to be caused by a change in brain state imposed by

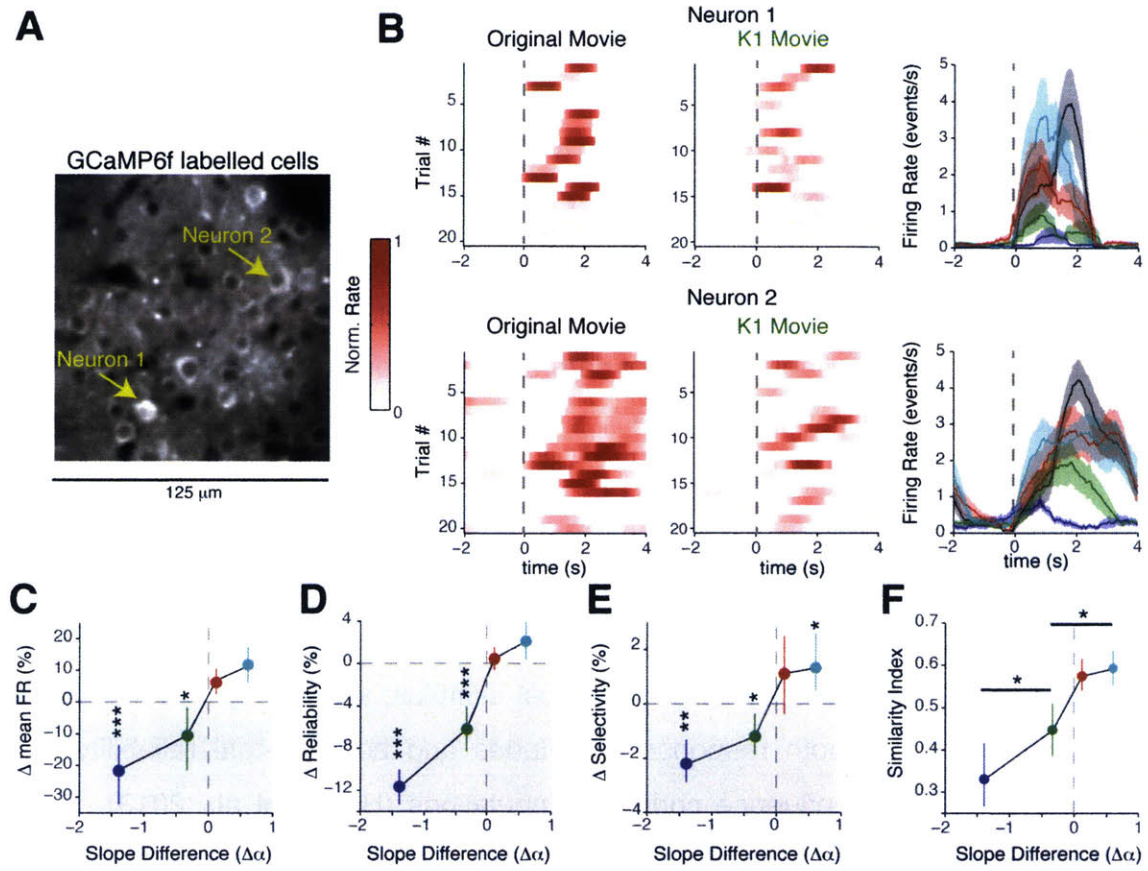

Figure 3-7. Decorrelated movies evoke weak and unreliable responses in awake mice. (A) Two-photon image of a population of 29 neurons expressing GCaMP6f. See Materials and Methods for details. **(B)** *Left and Middle,* Example raster plots of calcium responses from the cells marked in (**A**) to the original movie and its *K1* version. *Right,* Trial-averaged calcium responses from these cells. Both cells responded strongly to the spatially correlated stimuli (original, *K1.5* and *K2* movies), but less so to *K1* and very weakly to *K0* movies. Shaded region indicates SEM computed over 20 trials. **(C)** Percentage change of mean firing rate relative to the original movie. Neurons responded with significantly lower firing rates to spatially decorrelated stimuli. The decrease in firing rate is commensurate with OGB1 data in anesthetized mice (**Fig. 3-2E**). **(D)** Change in reliability relative to the original movie shows a strong decrease in reliability as spatial correlations are removed, similar to anesthetized animals (**Fig. 3-3C**). **(E)** Significant reduction in selectivity (measured from lifetime sparseness) for *K0* movies ($p < 10^{-5}$), but less so for *K1* movies ($p = 0.012$, Bonferroni-corrected rank-sum test). **(F)** Quantification of similarity index relative to the original movie. Similar to **Fig. 3-2D**, we observed a monotonic decrease in response similarity as spatial correlations were removed in the awake brain state. P-values in **(F)** computed using post-hoc rank-sum test relative to *K2* movies. All other P-values computed via Bonferroni-corrected rank-sum test. Data for **C-F** is from 230 neurons (4 mice) and is represented as median ± 95% confidence interval, obtained by bootstrapping.

anesthesia. Rather, decorrelated stimuli, which lack crucial low spatiotemporal information, are not effective at eliciting reliable responses from V1 neurons.

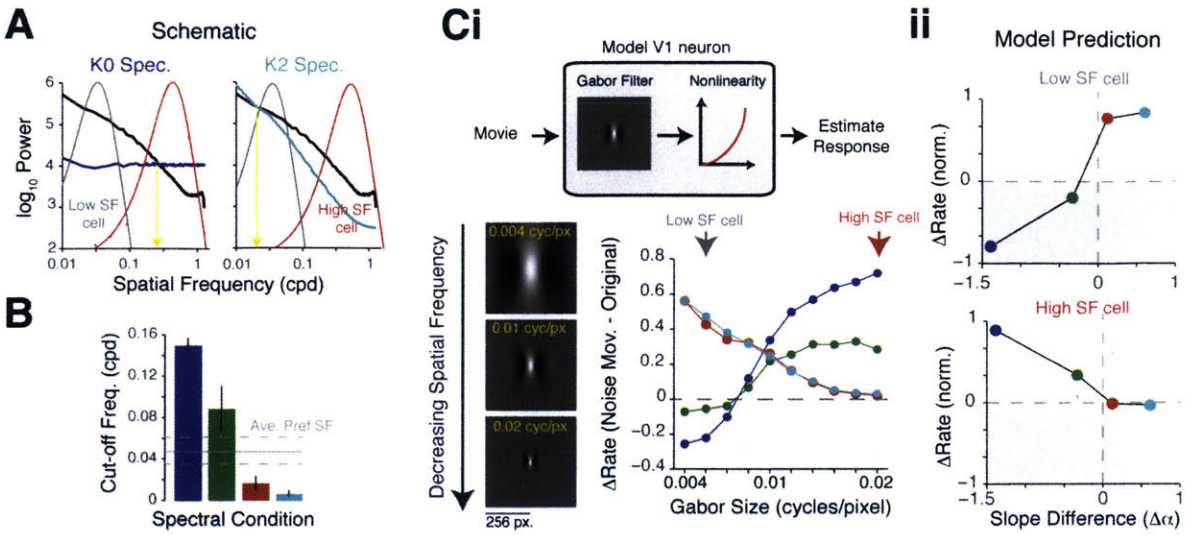

Figure 3-8. LN model to predict responses to noise-masked movies. (A) Schematic illustrating the portion of the power spectrum sampled by a typical low SF- (gray) or high SF-preferring (red) cell. The yellow arrow indicates the cut off frequency, which is defined as the SF at which the power in the noise-masked movie is the same as the original movie. **(B)** Quantification of the cut-off frequency (cpd, cycles per degree) for the different spectral conditions. As expected, cut-off frequency is the highest for *K0* and lowest for *K2* movies. Data pooled from 5 movies. The gray lines show the mean±SEM preferred SF of 308 neurons (4 mice). **(Ci)** *Top*, schematic of a linear-nonlinear (LN) model of a V1 neuron used to predict responses to the noise movies. Briefly, a Gabor filter was convolved with a noise-masked movie and the resultant was rectified with a point-wise nonlinearity to obtain a predicted firing rate. *Bottom*, plot showing the predicted firing rate difference (ΔRate = noise movie rate - original movie rate) for Gabor filters of different sizes. Decreasing the size of the Gabor filter is equivalent to increasing the preferred spatial frequency. (**Cii**) Plots showing predicted firing rate difference of a typical low SF-preferring cell (top, SF denoted by gray arrow in **Ci**) and high SF-preferring cell (bottom, red arrow).

Responses to noise-masked movies cannot be predicted from spatial frequency tuning of neurons

V1 neurons are known to be strongly selective for SF, in addition to orientation (Gao et al., 2010; Niell and Stryker, 2008). Since our noise masking technique alters the SF content of natural movies by redistributing power in different bands, we next asked if our observed results could be attributed to SF tuning properties alone (**Figure 3-8A**). Specifically, because *K2* movies have higher power in the low SF band (SF < 0.012 cycles/degree (cpd), **Figure 3-8B**), we expected low SF-preferring neurons to respond more strongly to *K2* movies than the original movie. Conversely, because *K0* movies had more power than the original movie at SFs greater than 0.14 cpd, we expected high SF preferring neurons to prefer *K0* movies.

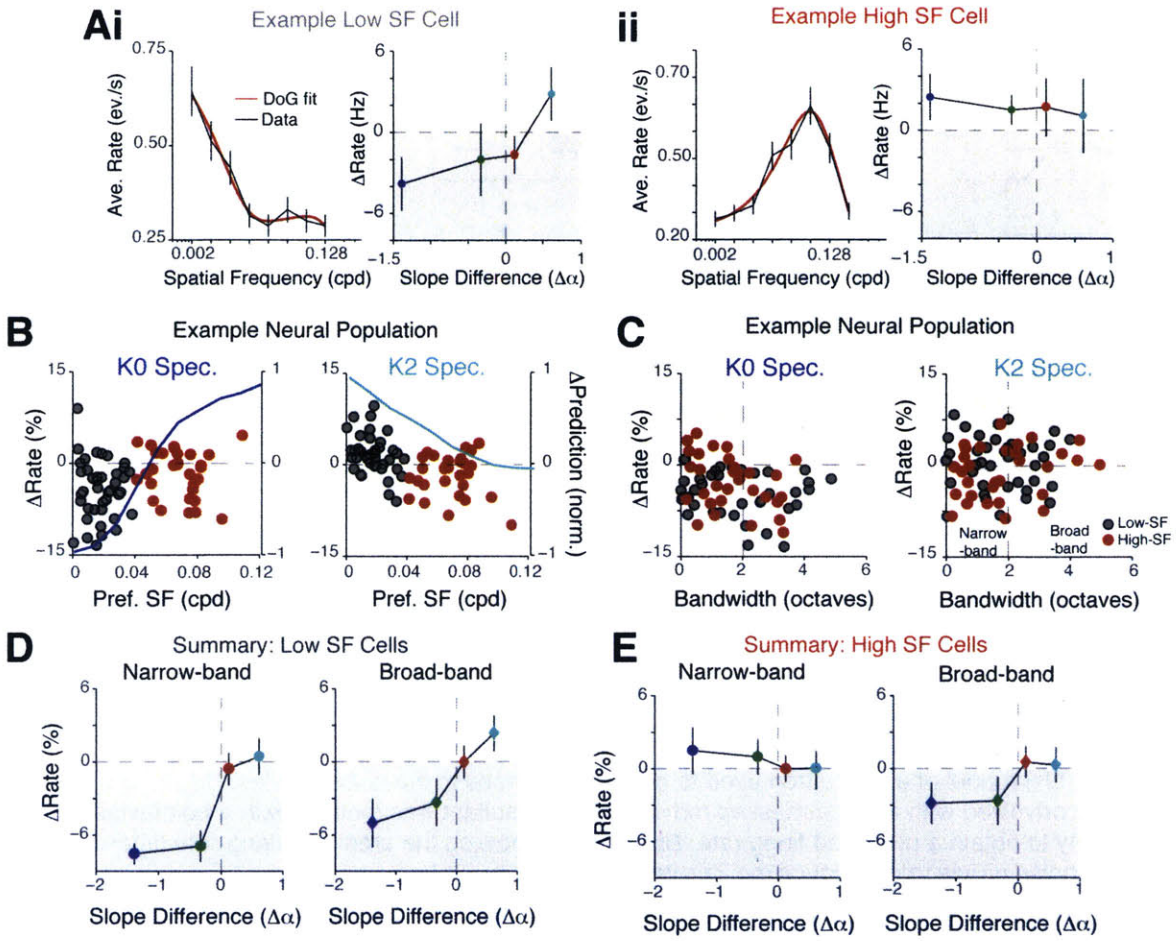

Figure 3-9. SF tuning of neurons cannot explain responses to noise-masked movies. (Ai) *Left,* Representative tuning curve of a low SF preferring cell, which responds more to the spatially correlated movies. The SF tuning curve of this cell is well fit with a difference of Gaussians model (red line). *Right,* Plot of firing rate difference of this neuron demonstrating that this cell responds more to *K2* movies. **(Aii)** Same plot as **(Ai)** but for a high SF preferring cell that responds more to spatially decorrelated movies instead. **(B)** Plots of difference in firing rate against preferred SF. Data from 78 simultaneously recorded neurons. Dark grey dots indicate low SF preferring neurons (SF < 0.04 cpd) and red dots indicate high SF preferring neurons. Also plotted on the same axes is the difference in firing rate predicted from the model in **(Fig. 3-5C)** to illustrate the expected trend. There was no statistically significant trend in the data for either spectral condition (*K0*: p = 0.65, *K2*: p = 0.071, Spearman's rank correlation coefficient). **(C)** Plots of difference in firing rate against SF tuning bandwidth. Dark grey dots indicate low SF preferring neurons and red dots indicate high SF preferring neurons. The dashed vertical line separates narrowband neurons (bandwidth < 2 octaves) from broadband neurons. **(D, E)** Difference in firing rates plotted against difference in slopes for low SF preferring cells **(D)** and high SF preferring cells **(E)**. Cells were further partitioned into narrowband and broadband based on tuning bandwidth.

The spatial summation properties of V1 neurons have been traditionally modeled as a linear-nonlinear (LN) cascade (Carandini et al., 1997; Movshon et al., 1978). In this model, the linear component resembles the spatiotemporal receptive field (RF), while

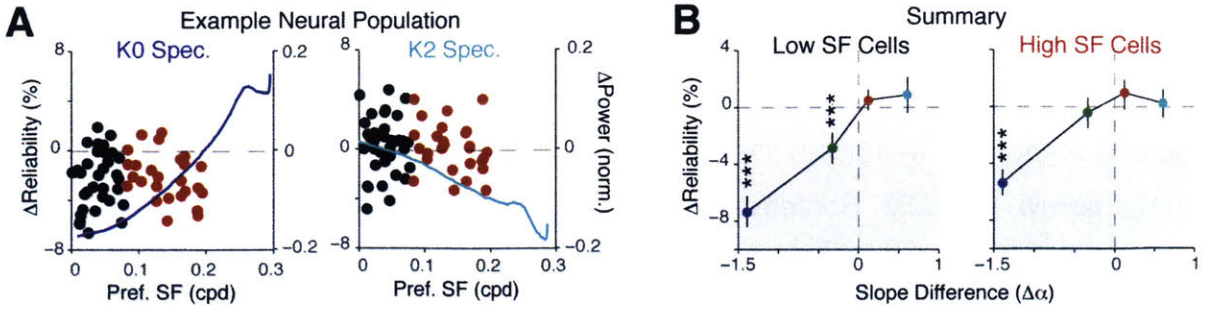

Figure 3-10. SF tuning properties cannot explain decrease in reliability for decorrelated movies. (A) Plot of difference in reliability (ΔReliability = noise movie reliability - original movie reliability) against preferred SF. The solid line denotes predicted difference in reliability from the model shown in **Figure 3-8C**. Reliability was calculated from predicted firing rate using the gradient of the regression between firing rate and reliability shown in **Figure 3-5G**. There was no statistically significant relationship between reliability and preferred SF for either spectral condition (*K0*: p = 0.85, *K2*: p = 0.89, Spearman's rank correlation coefficient). **(B)** Reliability is low for *K0* movies in both high and low SF preferring cells. Data was not separated according to bandwidth. Compared to original movie, no significant difference was found for *K1* movie reliability in high SF preferring cells. All data and error bars represent median±95% CI from 308 neurons (4 mice).

the nonlinear component models the spike generating mechanism. We used this simple model to predict the responses of neurons preferring different SFs (**Figure 3-8Ci**, top). As expected, the model predicted that high SF preferring neurons would respond with higher firing rates to decorrelated movies, whereas low SF preferring neurons would prefer the more strongly correlated movies (**Figure 3-8Ci**, bottom).

To test the veracity of this prediction, we determined SF tuning properties of neurons empirically from tuning curves that were well fit by a difference of Gaussians function (average R^2 = 82.4±8.22%, 308 neurons, 4 mice). Surprisingly, only very few neurons obeyed the predicted trend. Representative responses of two neurons that obeyed the model's prediction are shown in **Figure 3-9A**. Rather, on average, both high SF (neurons with preferred SF > 0.04 cycles/degree) and low SF preferring cells responded weakly to *K0* movies (**Figure 3-9B**, *left*). Similarly, there was only a weak correlation between the actual and predicted firing rate difference for *K2* movies (r = 0.153, Spearman's rank correlation coefficient, **Figure 3-9B**, *right*).

V1 neurons also have different tuning bandwidths, that is, they respond to a range of SFs, in addition to their preferred SF. Thus, we next investigated the relationship between bandwidth and firing rate difference (**Figure 3-9C**). This allowed us

to further categorize neurons into narrow- (bandwidth < 2 octaves, 106/308 neurons) and broad-band (202/308 neurons). On average, both narrow- and broad-band low SF neurons responded marginally more to *K2* movies than the original movie (broad, $p = 0.042$; narrow, $p = 0.53$, Bonferroni corrected rank sum test, **Figure 3-9C**). In contrast, only narrow-band neurons responded with marginally higher rates to *K0* movies (**Figure 3-9D**). Surprisingly, high SF preferring neurons responded equally to *K2* movies ($p = 0.871$), even though high SF power was significantly attenuated in these movies. Thus, these results are notably different from the prediction of the simple model (**Figure 3-8Cii**), suggesting that responses to the noise-masked movies cannot be predicted from SF tuning alone. Rather, neural activity is modulated by the entire power spectrum of natural scenes.

Additionally, we also found no correlation between reliability and preferred SF for either spectral condition (**Figure 3-10A**). Both high and low SF preferring neurons responded unreliably to *K0* movies (both $p < 10^{-5}$, Bonferroni corrected rank sum test, 308 neurons) and reliably to *K2* movies (**Figure 3-10B**). Interestingly, *K1* movies evoked reliable responses from high, but not low SF-preferring cells (high-SF: $p = 0.21$, 130 neurons; low-SF: $p < 10^{-5}$, 178 neurons, Bonferroni corrected rank-sum test, **Figure 3-10B**). These results support the hypothesis that reliability is a network-dependent phenomenon and may not be explained solely from the SF tuning properties of neurons.

Perturbing spatial correlations bi-directionally modulates signal and noise correlations between neurons

Computations in L2/3 of visual cortex depend on an intricate balance between feed-forward stimulus drive and recurrent network dynamics. As a result, the ability of V1 to encode information is influenced by both the structure and magnitude of correlations between neurons (Ecker et al., 2011; Moreno-Bote et al., 2014; Zohary et al., 1994). Thus, we sought to characterize how perturbing stimulus correlations altered inter-neuronal correlations. Specifically, we analyzed both signal correlation, which reflects similarities in evoked responses, and noise correlations, which captures the

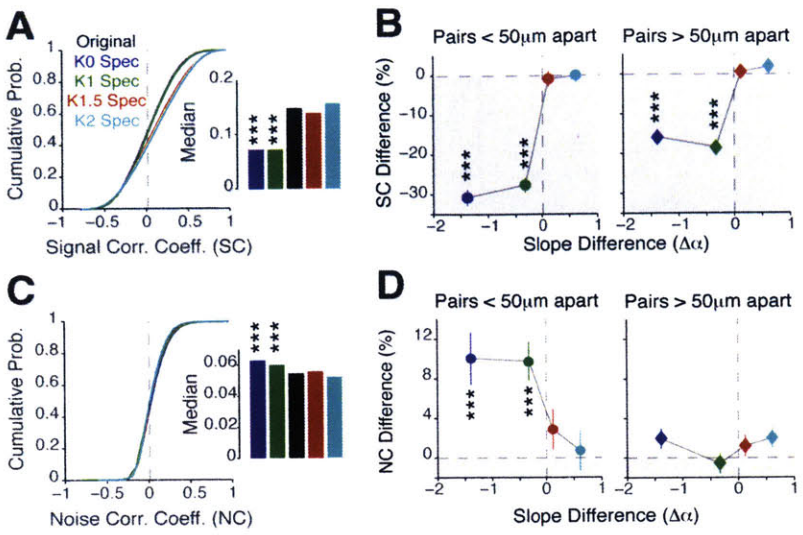

Figure 3-11. Bi-directional modulation of inter-neuronal correlations. (A) Cumulative distribution of signal correlation (SC) for all spectral conditions. Inset shows median SC. **(B)** Pairwise difference in signal correlation as a function of spectral slope difference for neuron pairs less than 50μm apart (left) and more than 50μm apart (right). **(C-D)** Same as **(A-B)** but for noise correlation (NC) instead. All data presented as median±95% CI from 16 mice (650 neurons, 210925 pairs).

dependencies between neurons that are not locked to the external stimulus (Cohen and Kohn, 2011).

In response to unperturbed natural movies, neurons were on average positively signal-correlated (mean ± SEM, 0.116±0.003, 650 neurons, **Figure 3-11A**), but only weakly noise-correlated (0.031±0.005, **Figure 3-11C**), comparable to previous reports using similar preparations (Cossell et al., 2015; Hofer et al., 2011). Presenting decorrelated movies strongly decreased signal correlation and increased noise correlation (SC: $p = 2.34 \times 10^{-6}$, NC: $p = 3.45 \times 10^{-5}$, Bonferroni-corrected rank-sum test relative to original, **Figure 3-11A, C**). In contrast, neurons maintained the same levels of signal and noise correlations in response to strongly correlated movies. Thus, removing spatial correlations from natural scenes not only increased trial-to-trial variability, but also increased correlated variability between neurons. These data also imply that perturbing spatial correlations in natural scenes is sufficient to drive the network into a weakly responsive and noisy coding regime (Churchland et al., 2010).

To better quantify how correlations changed between neurons, we computed, for each cell pair, a bootstrapped estimate of the percentage change in correlation relative to the original movie (**Figure 3-11B**). We partitioned our data into pairs that were either located <50μm ("neighboring pairs") or >50μm apart ("distant pairs") to reflect the fact that neurons separated by less than 50μm tend to have higher connectivity (Ko et al.,

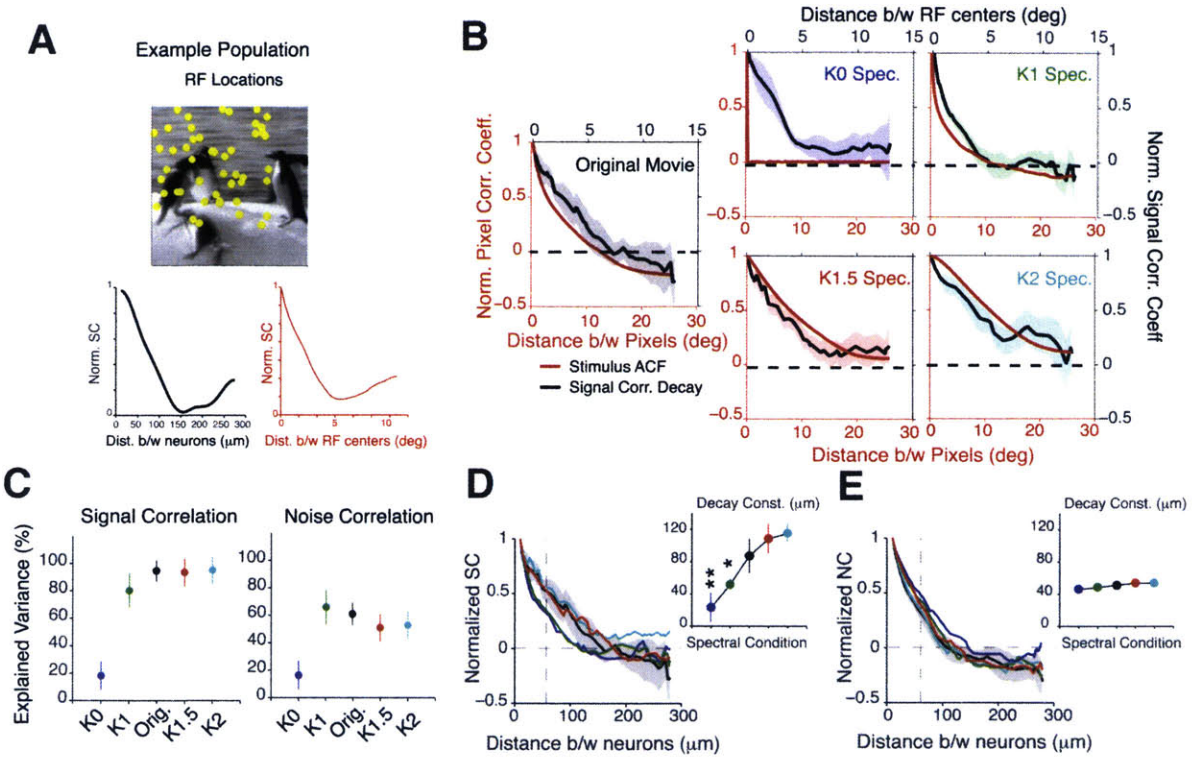

Figure 3-12. Spatial organization of inter-neuronal correlations closely track stimulus correlations. **(A)** *Top,* example receptive field (RF) locations of a population of 50 neurons. Note, the RF centers are not drawn to scale. *Bottom,* signal correlations (SC) between neurons in this example population decay with increasing distance between neurons (*left*) and increasing separation between RFs (Euclidean distance between centers, *right*). **(B)** Comparison between signal correlation decay (black lines) and stimulus ACF (red lines) for all spectral conditions. The shaded area denotes 95% CI of the median computed from 16 mice. In most conditions, the decay of signal correlations (except *K0* movies) between neurons closely tracks the stimulus ACF. **(C)** Explained variance of the fit between the stimulus ACF and the signal correlation decay function. The ACF is able to explain up to 80% of the decay in signal correlations and 60% of the decay in noise correlations. Data presented as median±95% CI. **(D, E)** Signal correlation **(D)** and noise correlation **(E)** as a function of distance between neurons for all spectral conditions. The inset shows the average decay constants estimated from single exponential fit. Data in the inset shown as mean±SEM. Shaded area in **(D)** in indicates 95% CI of the median of original movie. All data in (**B-E**) from 16 mice.

2011). Removing spatial correlations led to a significant reduction in signal correlations in both neighboring and distant pairs of neurons ($p < 10^{-8}$, Friedman test with Dunn's test, 16 mice, **Figure 3-11D**). In particular, *K0* movies most strongly decorrelated neighboring neurons (30% vs. 12% reduction). In contrast, noise correlation only increased for *K0* and *K1* movies in neighboring pairs but remained stimulus-invariant in distant pairs (**Figure 3-11C**).

Next, we investigated how the spatial organization of inter-neuronal correlations changed for the different stimuli. In particular, we were interested in understanding the relationship between inter-neuronal correlations and stimulus correlations. To this end, we first used a sparse noise stimulus to map receptive field (RF) centers of neurons in order to compute a cortical magnification factor (CMF, see Materials and Methods for details). In the example population shown in **Figure 3-12A**, $1deg^2$ of visual space corresponded to $0.992x10^{-3}$ mm^2 of cortical space (CMF in rostro-caudal axis = $32.3x10^{-3}$ mm/deg, CMF in medio-lateral axis = $30.7x10^{-3}$ mm/deg), which is consistent with previous studies of mouse visual cortex (Garrett et al., 2014). Importantly, within this example population, cells with neighboring RFs were more strongly signal correlated when stimulated with a natural movie than cells with distant RFs (**Figure 3-12A**, *bottom*). This is primarily because neurons with neighboring RFs sample similar parts of the visual scene (Bonin et al., 2011; Cossell et al., 2015). As a consequence, signal correlations in all populations (**Figure 3-12B**, 16 mice) decayed as the separation between neurons increased.

We then used the CMF to topographically map visual space into cortical distance. This permitted us to directly compare both signal correlations and the spatial ACF of the different movies on the same scale (**Figure 3-12B**). For all imaged populations, the spatial decay of signal correlations closely matched the spatial ACF. In particular, the change in pixel correlations was able to explain more than 80% of the variance in the spatial organization of signal correlations for all spectral conditions, with the exception of *K0* movies (**Figure 3-12C**, 16 mice). Therefore, the spatial ACF provides a good model for signal correlations between V1 neurons, implying that correlations between neurons fall off as expected from the visual stimulus.

To further explore this notion, we quantified the spatial decay constant by fitting single exponentials to the signal correlation decay functions. Expectedly, this analysis revealed that signal correlations decayed the fastest between neurons in for both *K0* and *K1* movies (*K0*, $p = 1.86x10^{-2}$, *K1*, $p = 3.81x10^{-2}$, rank-sum test relative to original, 16 mice, **Figure 3-12D**), which was commensurate with the fact that pixel correlations also decayed quickly with distance in these stimuli (also see **Figure 3-1E**). In stark

contrast, these correlations persisted over a longer range when K2 movies ($p = 6.27x10^{-2}$) were displayed. In particular, coupling between neurons remained strong even among neurons located >200μm apart. Interestingly, comparing **Figures 3-12D** (inset) **and 3-1E**, we noticed that neurons were less correlated than expected when stimulated with *K2* movies as the signal correlation decay constant appeared to asymptote at approximately 120μm. This could be due to either decorrelation occurring earlier in the visual pathway or the limited spatial scale of our imaging method. Despite this, our results demonstrate that strongly correlated movies couple neurons located far apart in the cortex, without changing the coupling magnitude.

In contrast to signal correlations, noise correlations decayed exponentially between neurons (average R^2 of exponential fit = 87±1.6%, 16 mice), with a decay constant of close to 50μm, regardless of the stimulus condition ($p = 0.67$, Cochran-Armitage test for trend, **Figure 3-12E**). As a consequence, the spatial ACF poorly predicted the decay in noise correlations (**Figure 3-12C**). Thus, unlike the magnitude (**Figure 3-11C**), the spatial structure of noise correlations is not significantly altered when stimulus correlations are removed. Taken together, these results indicate that perturbing spatial correlations bi-directionally alters both the spatial structure and magnitude of neuronal co-activation. Specifically, neighboring neurons that are located within 50μm of each other, become less signal correlated and more noise correlated when presented with weakly correlated stimuli.

Clustering analysis reveals ensembles of neurons that share similar functional properties

These findings raise an important question: how do changes in inter-neuronal correlations influence reliability? Recent studies have shown that neurons with similar orientation tuning or receptive fields are recurrently interconnected (Cossell et al., 2015; Ko et al., 2011). As a consequence, distinct neural ensembles are recruited by different visual stimuli (Harris and Mrsic-Flogel, 2013; Miller et al., 2014). However, it remains unclear how the activation of these ensembles improves coding reliability. Thus, we sought to elucidate the relationship between response reliability and inter-neuronal

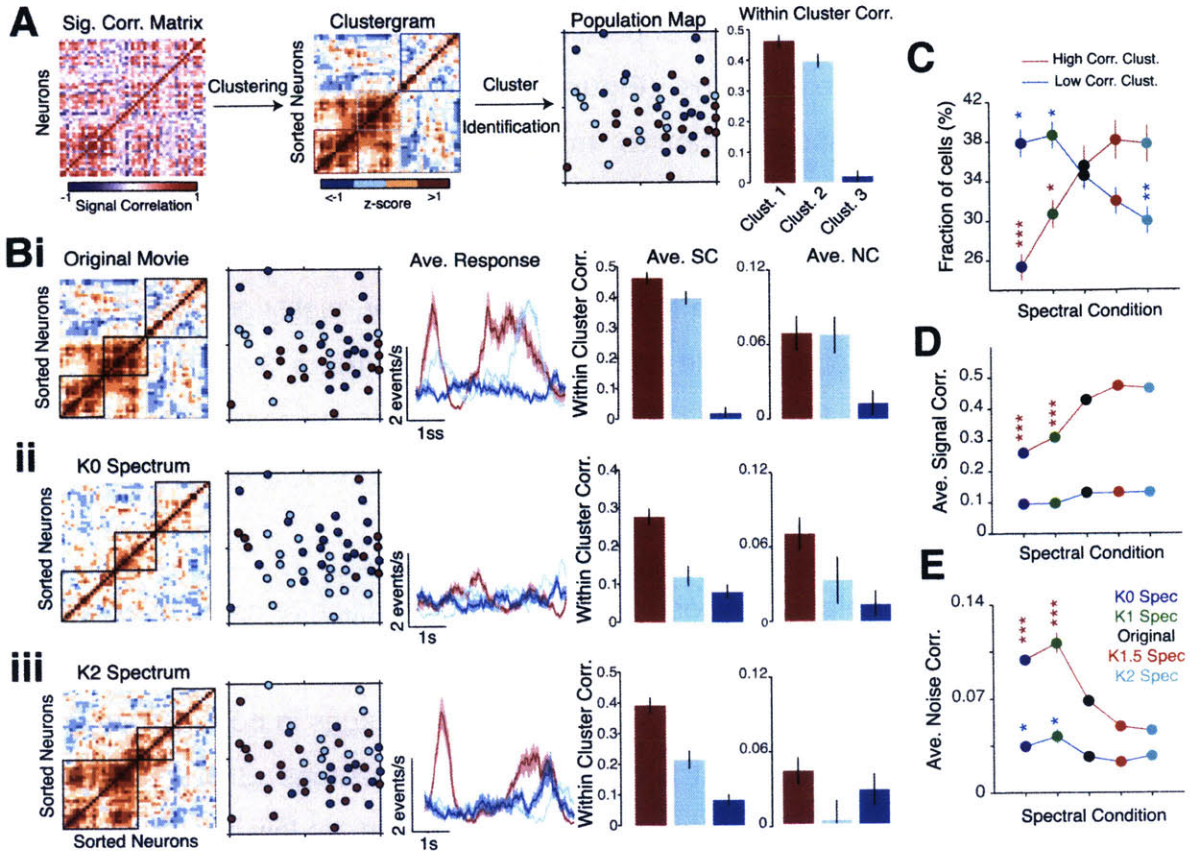

Figure 3-13. Clustering analysis reveals distinct ensembles of highly correlated neurons. (A) Illustration of the clustering method used to group neurons based on their signal correlation (SC) coefficients. Histogram shows average SC within each cluster, error bars denote SEM. **(B)** Clustering analysis applied to the same neural population in response to an original movie **(Bi)**, and its *K0* spectrum **(Bii)** and *K2* spectrum versions **(Biii)**. The color-coded population maps show a reorganization of the clusters to the different stimuli. The size of the population map is 250μm x 250μm. Subsequent panels show the average response of neurons in each clusters and a quantification of within-cluster signal and noise correlation coefficients (SC and NC respectively). Error bars and shaded areas denote SEM across neurons in each cluster. **(C)** Fraction of neurons in the high (dark red) and low (dark blue) correlation clusters for all stimulus conditions. **(D, E)** Average signal correlation **(D)** and noise correlation **(E)** between neurons in each cluster. All data in (**D, E**) presented as mean±SEM from 16 mice. P-values computed relative to original movie using post-hoc Bonferroni-corrected rank-sum test. Colors of symbols indicate the correlation cluster.

correlations. Specifically, we hypothesized that an increase in signal correlations should help to reduce response variability and increase reliability.

To test this hypothesis, we used an agglomerative hierarchical clustering technique to group neurons according to similarities in their correlation coefficients (**Figure 3-13A**, see Appendix 3A for further details). This method allowed us to identify neurons that were either strongly or weakly correlated with their neighbors. We found

that clusters within the same neural population reorganized depending on the level of spatial correlations in the stimulus (**Figure 3-13B**). Neurons within each cluster responded with similar temporal dynamics, implying that all clustered neurons were simultaneously co-activated by the stimulus. Specifically, strongly correlated neurons in the original movie (**Figure 3-13Bi**) became decorrelated when presented with a *K0* movie (**Figure 3-13Bii**). In response to *K2* movies, however, highly correlated neurons were more dispersed throughout the imaged population (**Figure 3-13Biii**), consistent with the result shown in **Figure 3-12D**. This example also demonstrates that, even though *K2* movies evoke the same absolute level of intra-neuronal correlations as original movies, they recruit distinct ensembles of neurons.

The strongly correlated movies recruited a larger fraction of neurons in the high correlation cluster (38% *K2* vs. 27% *K0* movies, 16 mice, **Figure 3-13C**). Surprisingly, the original movie recruited almost the same number of neurons in both high- and low-correlation clusters, suggesting that a sparse subset of neurons (36%) is active for natural movies. Additionally, *K2* movies had fewer neurons in the low correlation cluster compared to original movie ($p = 1.03 \times 10^{-4}$, Kruskal-Wallis one-way ANOVA, 16 mice). This result confirmed that increasing spatial correlations in natural movies increased the proportion of active neurons. This also highlights the different coding strategy used by the visual cortex to process movies with different spectral properties.

Weakly correlated ensembles did not show a strong stimulus-dependence in signal correlations (low correlation cluster: $p = 0.43$, Kruskal-Wallis one-way ANOVA, 16 mice, **Figure 3-13D**), as they were poorly driven by the different movies (**Figure 3-13Bi-iii**). We were surprised to find that decorrelated movies elicited stronger noise correlations than the original movie, even within the cluster of strongly responding cells ($p = 1.24 \times 10^{-5}$, Kruskal-Wallis one-way ANOVA with Bonferroni correction, 16 mice, **Figure 3-13E**). Thus, movies that are unreliably processed are dominated by higher noise correlations.

Correlated ensembles of neurons reliably represent visual stimuli

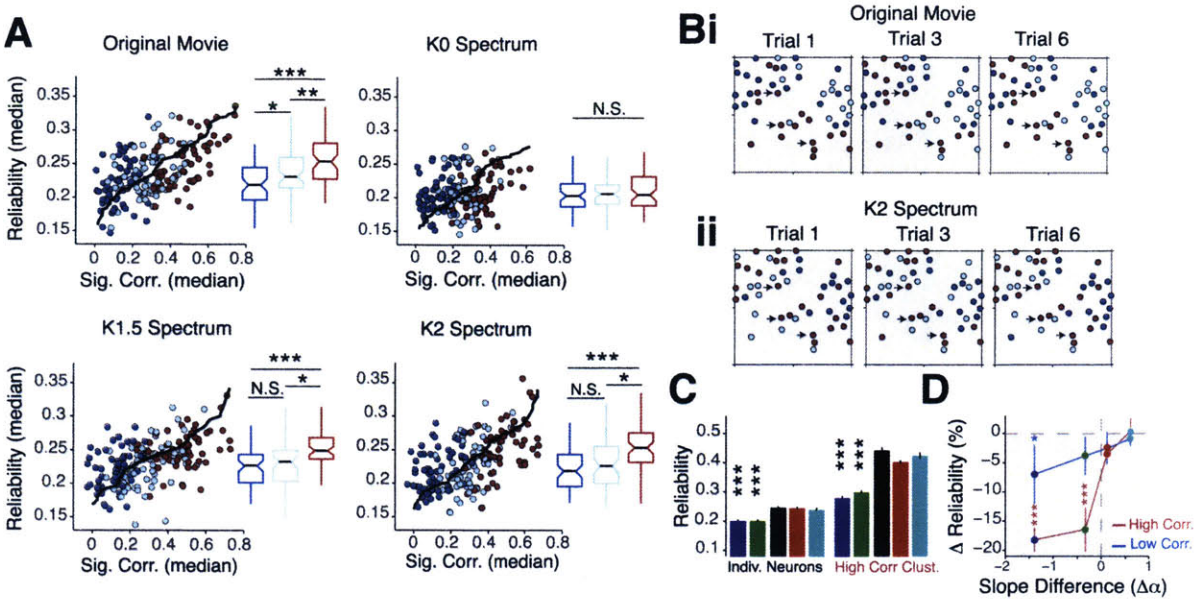

Figure 3-14. Highly correlated neural ensembles are reliably activated over multiple stimulus repetitions. (A) Scatter plot of median reliability against median signal correlation for each population and movie. The solid black line denotes trend computed via LOESS regression. The box-whisker plots quantify reliability in each cluster. **(B)** Cluster maps derived from inter-neuronal correlations on each trial. Representative examples are shown for three trials (1, 3 and 6) for both original movie **(Bi)** and *K2* movie (**Bii**). Neurons colored in red are strongly correlated with each other on each repetition of the stimulus. Almost the same clusters are re-activated on each trial for both stimuli (examples indicated by black arrows). **(C)** Quantification of median reliability within the high correlations cluster. For all stimulus conditions, pooling responses from neurons in a cluster improves reliability ($p < 10^{-10}$ Kruskal-Wallis ANOVA between individual neuron and high correlation cluster reliability). **(D)** Percentage change in high (dark red) or low (dark blue) correlation cluster reliability plotted against spectral slope difference relative to the original movie. Data presented as median±95% CI from 16 mice.

Next, we analyzed how response reliability differed between the clusters of either strongly or weakly correlated neurons (**Figure 3-14A**). In almost all recorded populations, the high correlation cluster contained a larger fraction of reliable responders than the low correlation cluster (mean±SD: 29.8±1.1% vs. 14.6±1.5%, $p < 10^{-5}$, Wilcoxon two-tailed rank-sum test, 650 neurons). Consistent with this, strongly correlated neurons had more reliable responses than weakly correlated neurons for original, *K1.5* and *K2* movies, but not for the decorrelated movies ($p < 10^{-5}$, Kruskal-Wallis one-way ANOVA with Bonferroni correction, 16 mice, **Figure 3-14A**). Specifically, all clusters from the *K0* movie condition were weakly reliable.

Are these clusters reliably activated from one trial to the next? To answer this question, we repeated this clustering analysis on responses obtained from individual

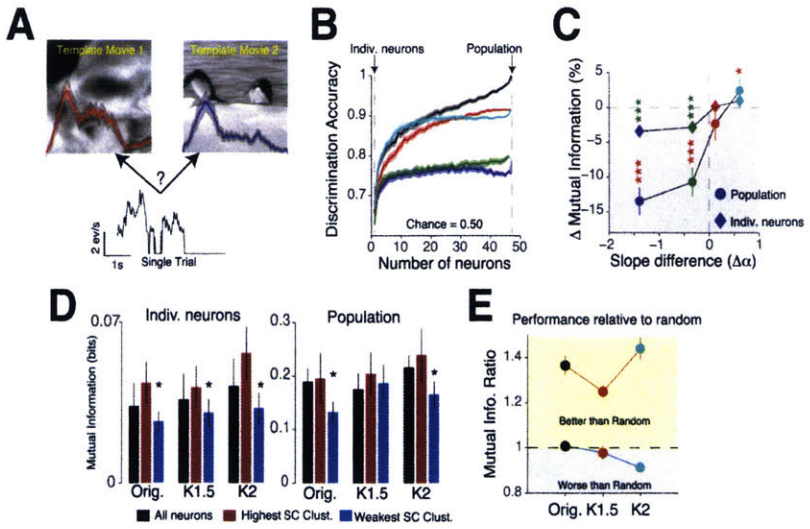

Figure 3-15. Ensembles of correlated neurons efficiently encode natural scenes. (A) Illustration of the decoding method. A single trial response was classified into one of two movies (repeated for all possible pairs of movies) on the basis of its Euclidean distance to a population template (see Appendix 3A). **(B)** Plot illustrating that decoding accuracy (% of correctly classified trials) increases as the number of neurons contributing to the single-trial response increases. Data in this plot is from a single population of neurons (46 neurons). **(C)** Percentage changes in mutual information for both individual neurons (diamonds) and whole populations (circles). Data presented as median±95% CI. P-values computed using post-hoc Bonferroni-corrected rank-sum test relative to original movie. Green stars, P-values for individual neurons. Red stars, P-values for population. **(D)** Average discriminability for the individual neurons (left) and whole population (right) decomposed according to correlation cluster. P-values computed using one-tailed rank-sum test relative to original movie. Error bars denote SEM. **(E)** Comparison of mutual information for high and low correlation clusters relative to randomly sampled subpopulations of the same size. Data in (**C-E**) from 16 mice, error bars denote SEM.

trials. "Snapshots" of the same neural population taken during the first, third and sixth stimulus repetitions (**Figure 3-14B**) showed that, for both the unperturbed and the *K2* movie, neurons in the high correlation cluster were reliably activated during each stimulus repetition. In particular, neurons that were strongly correlated on the first trial remained strongly correlated during even the sixth stimulus repetition. Examples of these neurons are indicated by black arrows in **Figure 3-14B**. Consistent with the idea that population coding decreases variability (Shadlen and Newsome, 1994), the high correlation cluster was always more reliably activated than the individual neurons themselves ($p < 10^{-10}$, Bonferroni-corrected rank-sum test between individual neurons and high correlation cluster, **Figure 3-14C**). However, this population coding strategy did not improve the reliability for either *K0* or *K1* movies. In stark contrast, the low correlation cluster was unreliably activated by all spectral conditions (**Figure 3-14D**).

Taken together, our clustering analysis revealed that: (1) neuronal ensembles dynamically reorganize depending on the spatial properties of the stimulus. (2) Coding within these ensembles is highly reliable. (3) These ensembles are also reliably and stably activated over multiple stimulus repetitions.

Reliability enables accurate discrimination of strongly correlated stimuli

In principle, the highly reliable responses of correlated neurons should efficiently represent visual information by ensuring greater discriminability between the different movies. To further explore this idea, we performed linear decoding analysis on neural responses to assess the ability of neurons to discriminate different movies in each spectral condition. Specifically, we used a nearest means classifier (**Figure 3-15A**, described in Appendix 3A) to predict from single-trial responses which movie had been presented (Goard and Dan, 2009; Kampa et al., 2011).

We measured discrimination accuracy by computing the fraction of correct classifications made by the classifier. **Figure 3-15B** exemplifies the decoding performance of a single population of neurons (46 neurons). In this population, discrimination accuracy improved as the number of neurons in the template increased. Interestingly, the classifier's performance saturated at approximately 12 neurons, suggesting that all neuron groups larger than 12 provided the same information about the stimulus. We found a similar trend in all imaged populations (16 mice): increasing the number of neurons in the template, up to a limit, improved discrimination accuracy for correlated movies (*Orig.*: 59.8±0.05% vs.71.6±0.11%; *K2*: 60.7±0.06% vs. 74.1±0.11%, both $p < 0.05$, permutation test, data not shown). In contrast, pooling from a larger group of neurons failed to improve the decoding accuracy of decorrelated movies (individual vs. population, *K0*: 56.4±0.04% vs. 58.1±0.09%, $p = 0.58$).

To better characterize decoding accuracy, we computed mutual information (MI), which is a measure of the reduction in uncertainty about the presented movie by knowledge of a single-trial response (see Appendix 3A). At the level of both individual neurons and the entire population, we found a significant reduction in MI as spatial

correlations were removed from the stimulus ($p < 10^{-6}$, Friedman Test with Bonferroni correction, 16 mice, **Figure 3-15C**). These results suggest that population coding does not improve the representation of decorrelated movies and is consistent with the result that these movies evoke decorrelated responses. Thus, regardless of their phase information, neurons are not able to discriminate between movies with either *K0* or *K1* spectra.

Does pooling over specific ensembles improve decoding accuracy? We restricted our analysis to correlated stimuli because decorrelated stimuli failed to reliably recruit neural ensembles (**Figure 3-14C**). Strongly correlated neurons outperformed neurons that were either weakly correlated or randomly sampled from the population (*Orig.:* $p = 4.2 \times 10^{-2}$, *K1.5*: $p = 3.6 \times 10^{-2}$, *K2*: $p = 4.1 \times 10^{-2}$, rank-sum test relative to all neurons, 16 mice, **Figure 3-15D**, *left*). Additionally, pooling responses within the highly correlated ensembles resulted in a remarkably similar performance to the entire population. In contrast, decoding from just the weakly correlated clusters led to a significantly lower MI ($p = 4.32 \times 10^{-2}$, rank-sum test relative to all neurons, **Figure 3-15D**, *right*). Thus, despite containing fewer neurons, the high correlation clusters were as good as the entire population at discriminating between the different stimuli.

Was this improved performance due to the fact that the ensembles for *K1.5* and *K2* movies contained a larger fraction of neurons? To answer this, we compared the MI obtained from randomly sampling sub-populations of the same size as either the high or low correlation cluster (**Figure 3-15E**). For all stimuli, high correlation clusters outperformed randomly sampled sub-populations. Thus, these results, together with our clustering analysis, indicate that visual information is encoded with high fidelity within ensembles of strongly correlated neurons.

An ensemble-specific normalization model explains responses to noise movies

Finally, we sought a simple model that could explain responses to noise movies by taking into account the unique pattern of ensemble activation. To this end, we further extended the model introduced in **Figure 3-8Ci**. First, we obtained best-fit RF estimates

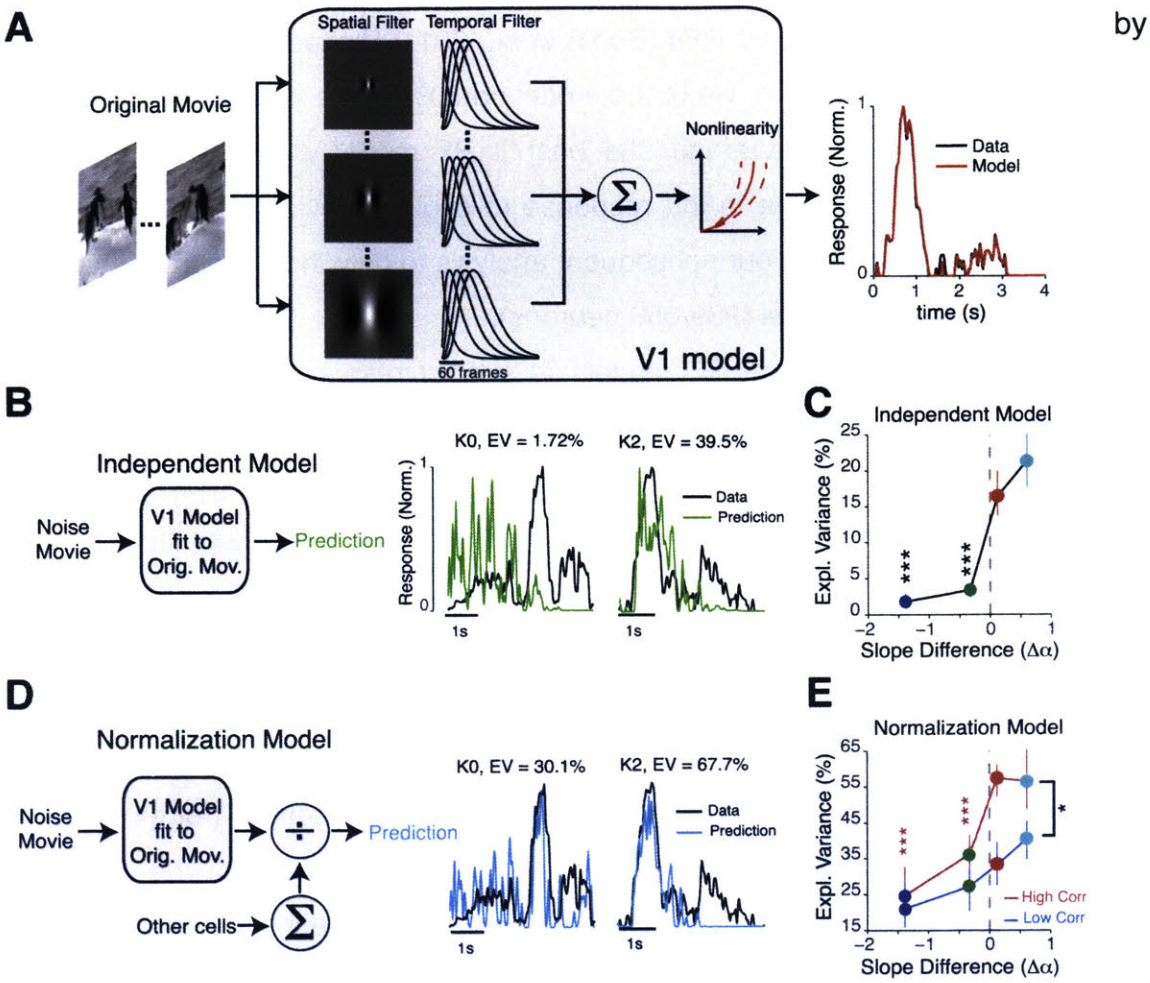

Figure 3-16. Linear-Nonlinear cascade model with normalization explains responses to correlated stimuli. (A) Schematic of the model used to determine the best receptive field estimate from responses to an original movie. **(B)** *Left,* Schematic describing independent model. This model assumes no interactions between neurons. *Right,* Representative examples showing the predicted output of this model (green) relative to the real data (black) from two cells. EV, explained variance. **(C)** Percent variance explained by the independent model. **(D)** *Left,* Schematic describing normalization model. This model assumes interactions between neurons via the divisive normalization rule. *Right,* Responses of the same cells as **(B)**, but predicted with the normalization model instead. (**E**) Percent variance explained by the normalization model for all stimulus conditions. Here, the high (dark red) or low (dark blue) correlation clusters were separately considered as normalization pools. Red stars indicate comparisons performed for data in the high correlation cluster (relative to *K2* movie). Black stars indicate comparisons between high and low correlation clusters. Data in (**C**) and (**E**) presented as median±95% CI from 239 neurons.

fitting LN models to the original movie response of each neuron (**Figure 3-16A**). Since we did not know the spatial locations or structures of RFs *a priori*, we had to estimate RFs by determining the least-squares best-fit model from a bank of linear filters, which

were analogous to mouse V1 RFs (Bonin et al., 2011; Ringach et al., 2002) (**Figure 3-16A**). To prevent over-fitting, we fit the model on 50% of the trials and used it to predict the remaining trials. On average, the best-fit RF model was able to explain up to 85.0±13.9% of the variance in the response to original movies (see **Figure 3-16A** for example fit). We restricted our subsequent analysis to only those neurons that could be well explained by this model (239/650 neurons).

Next, we used these RF models to predict responses to the various noise-masked movies. Specifically, we considered two alternative explanations. In the first, termed the independent model, we assumed that neurons acted independently, without interacting with other cells (**Figure 3-16B**). In the second, termed the normalization model, we applied the divisive normalization rule to pool activity from either the highly or weakly correlated clusters (**Figure 3-15D**). Specifically, with the second model, we tested the hypothesis that the highly correlated cluster of neurons formed a normalization pool (Carandini and Heeger, 2012). The independent model, which ignored population coding, poorly predicted noise movie responses (for example **Figure 3-16B**): this model could explain only 1.3-3.4% of the variance of the *K0* and 16.5-21.8% of the variance of *K2* movie responses ($p < 10^{-4}$, Friedman Test with Bonferroni correction, 239 neurons, **Figure 3-16C**). This result also affirmed our finding that SF tuning (or RF properties) alone are not able to predict responses to the noise movies (**Figure 3-8**). In contrast, the normalization model outperformed the independent model in predicting noise movie responses (for example **Figure 3-16D**). Considering only neurons in the high correlation cluster as the normalization pool, the model allowed us to explain up to 55.7% of the variance of the *K2* responses (**Figure 3-16E**). In comparison, only 39.7% of the variance of the *K2* responses could be explained by normalizing over neurons in the low correlation cluster ($p = 0.045$, rank-sum test between high and low correlation clusters, 239 neurons). The normalization model demonstrated substantially higher prediction accuracy of *K1.5* and *K2* movies when responses were normalized over neurons in the high correlation cluster compared to the independent model. Hence, a parsimonious explanation for the decreased reliability for either *K0* or *K1* movies is a failure of these stimuli to activate a significant ensemble of

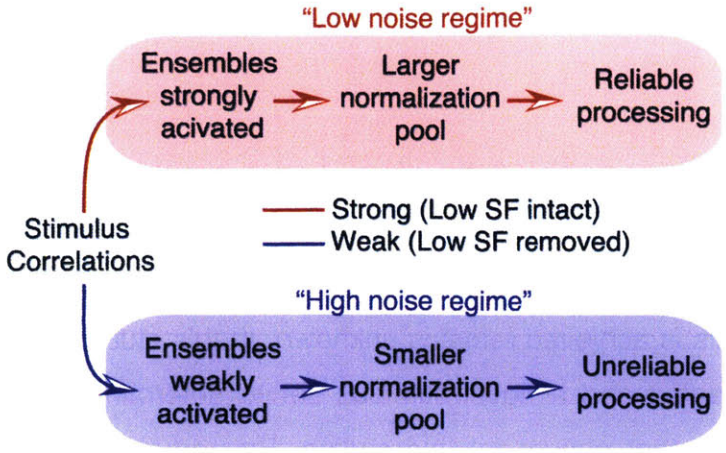

Figure 3-17. Schematic summarizing the main findings of this study. In the "low noise regime", strong stimulus spatial correlations dynamically alter inter-neuronal correlations to change the normalization pool, ensuring reliable processing. In the "high noise regime", weak stimulus correlations fail to activate ensembles, resulting in unreliable processing.

correlated neurons. However, we cannot exclude the possibility that stimuli with different spatial correlations change the spatiotemporal structure of RFs (Fournier et al., 2011).

In summary, our results support the idea that pooling responses from correlated ensembles of neurons decreases inter-trial variability. Notably, our work demonstrates that V1 has two distinct, stimulus-dependent coding regimes - a *low noise* regime, which is activated by movies which have strong spatial correlations and intact low SF information, and a *high noise* regime, which is activated by movies which lack these spatial correlations (**Figure 3-17**). In the *low noise* regime, inter-neuronal correlations reorganize to recruit more neurons, which in turn improves coding reliability and selectivity. In contrast, *K0* and *K1* movies, which lack the appropriate spectral content, fail to recruit these ensembles, resulting in a noisier code. Thus, the coding regime used by the visual cortex depends critically on the spectral statistics of the stimulus.

3.4 Discussion

The natural environment contains a vast amount of visual information, but not all of it is important to the behavior of an animal. As such, the visual system faces the challenging task of generating robust neural codes that parsimoniously convey this information. However, the mechanism by which this is achieved remains unknown. In this study, we used *in vivo* two-photon calcium imaging and a novel stimulus set to determine the statistical properties of natural scenes that ensure reliable coding. By perturbing the power spectrum in natural movies, we discovered that attenuating low SF information, which removed long-range correlations between pixels, resulted in an unreliable and unselective code. Surprisingly, the spatial structure of inter-neuronal correlations closely tracked stimulus correlations, leading to the activation of specific neural ensembles. Our study further revealed that only ensembles of correlated neurons reliably encoded the different movies. Taken together, we demonstrate that the unique spectral structure of natural scenes helps to improve coding fidelity by dynamically modulating inter-neuronal correlations in the visual cortex.

Power Spectrum and image discrimination

Natural scenes are widely used to probe coding in V1. However, unlike simple stimuli such as sinusoidal gratings, natural scenes are statistically more complex and are harder to parameterize (Olshausen and Field, 2005). Without parameterization, it is difficult to relate neural response properties, such as reliability, to a single stimulus attribute. Phase randomization is commonly used to relate the statistics of natural images to coding. For example, two recent studies showed that randomizing the phase of natural scenes decreased selectivity and reliability of V1 neurons without altering the overall firing rate (Froudarakis et al., 2014; Pecka et al., 2014). Phase randomization destroys the edge structure of images but keeps spatial correlations intact. However, because these stimuli are non-parametric, these studies were not able to attribute changes in reliability to specific aspects of the phase spectrum.

Although the power spectrum contains relatively low-level features, such as spatial correlations, it is believed to play an important role in rapid image discrimination and classification (Bar, 2004; McCotter et al., 2005; Oliva and Torralba, 2007). Specifically, it provides diagnostic information about the spatial organization of objects within the scene, which is used to learn the gist of an image (Oliva and Torralba, 2006). To date, very little information exists on how the power spectrum influences coding in V1. In this study, we reasoned that an image that is easily discriminated should be coded more efficiently. Thus, we hypothesized that the power spectrum could modulate neural reliability, which in turn would increase the fidelity of information processing and facilitate image discrimination.

Our novel noise-masking technique allowed us to test this hypothesis and revealed that the mouse visual system is highly sensitive to perturbations in the power spectrum. Importantly, our parametric approach permitted us to relate changes in neural coding to changes in the spectral slope. Subtle changes in the spectral slope strongly modulated the firing rates of V1 neurons in both anesthetized (**Figure 3-4**) and awake mice (**Figure 3-7**). These results are consistent with previous studies, which found no difference in rates when only the structure of spatial correlations was left intact (Froudarakis et al., 2014; Kayser et al., 2003). Similar results were also found in the visual cortex of primates (Freeman et al., 2013). Importantly, we found that the trial-to-trial reliability of responses was also strongly modulated by the low SF and TF content of the image. In particular, attenuating the low spatiotemporal content (*K0* or *K1* movies) was more detrimental to neural coding than attenuating the high SF content (*K1.5* or *K2* movies).

We asked whether these results could be explained just from the SF tuning properties of V1 neurons. Surprisingly, a simple linear-nonlinear model based on SF tuning poorly explained our data (**Figures 3-8 - 3-10**). These results indicate that the observed changes in firing rate and reliability were not the result of simple spatial filtering properties of neurons alone. Rather, the nonlinear interaction between diversely tuned neurons, via divisive normalization, better explained the responses (**Figure 3-16**).

Thus, our work provides evidence that V1 processes broadband stimuli by integrating over multiple SF channels, presumably by altering correlations between channels.

Implications for natural scene processing

We show that neurons in mouse V1 respond more reliably to movies with strong low SF power. The low spatiotemporal content in natural scenes is mainly large, coarse-grained objects, which can be used for basic scene recognition and motion discrimination (Bar, 2004). Our finding that mice respond similarly to movies that are much more strongly correlated than the original movie, implies that mice primarily use their vision to extract coarse grained information from their visual environments, presumably to guide navigation in low light conditions or to avoid aerial predators (see references within (Zoccolan, 2015)).

Interestingly, our results strongly support psychophysical studies, which show that the human visual system also uses information in the low SF bands to rapidly and accurately discriminate natural scenes (Gaspar and Rousselet, 2009; McCotter et al., 2005). We show that attenuating these bands, even to a small degree (~10 dB in the case of *K1* movies), strongly degrades processing in V1. Additionally, neurons were not able to discriminate movies with decorrelated power spectra, despite large differences in their phase information (**Figure 3-15**). While we cannot draw easy parallels between the visual systems of mice and humans, our data, together with Froudarakis et al. (2014), suggests that mouse V1 might also use spectral information in a similar manner. Specifically, stimuli that lack the appropriate spectral structure are processed less reliably, making them harder to discriminate. Further experiments, however, are required to determine how mice use spatial statistics to discriminate between different scenes (Vinken et al., 2014). In **Chapter 6**, we describe behavioral experiments aimed at answering this question.

Effects of stimulus statistics on population coding

Why are movies with strong spatiotemporal correlations more reliably processed in V1? Our analysis revealed that neighboring neurons were more strongly noise correlated

when stimulated with decorrelated movies. This suggests that shared variability between neurons is detrimental to the representation of weaker stimuli, such as *K0* movies (Zohary et al., 1994). One possible explanation to why correlated movies have lower noise correlations is that by co-activating distant neurons, these movies could increase inhibitory surround suppression (Adesnik et al., 2012; Snyder et al., 2014).

In fact, we demonstrate that increasing spatiotemporal correlations increase signal correlations between neurons. This is in contrast to the ideas of decorrelation that have been proposed for the early visual areas (Atick and Redlich, 1990), and suggest that V1 uses a different coding strategy. By pooling across neurons, V1 appears to multiplex information from multiple filters, allowing more filters to remove increase redundancy in the stimuli. This strategy could also be useful in overcome intrinsic variability in cortical neurons (Kara et al., 2000). Additionally, our finding that the size of neuronal ensembles recruited increase from K0 to K2 movies, suggests that the cortex is using fewer dimensions to encode stimuli with high spatiotemporal correlations. Similar low dimension coding strategies have been shown in the auditory cortex (Bathellier et al., 2012). In essence, we believe that by dynamically modulating inter-neuronal correlations, the cortex is optimizing its filter bank for more efficient coding.

The high spatial resolution provided by two-photon microscopy permitted us to investigate the spatial organization of inter-neuronal correlations. Notably, neuronal coupling patterns changed with stimulus correlations suggesting that scenes with different spectral statistics recruit distinct neural ensembles. These findings add to the growing body of evidence that visual processing is carried out by discrete clusters of functionally coupled neurons (Kampa et al., 2011; Miller et al., 2014; Okun et al., 2015). Remarkably, these strongly correlated ensembles were reliably and stably activated over multiple stimulus repetitions (**Figure 3-14**). Additionally, although these ensembles comprised ~30% of the population, they performed as well as the entire population in encoding various movies, supporting the notion of a sparse population code. Together, these results indicate that reading the activity of these clustered ensembles of neurons is sufficient to accurately discriminate different movies.

How these neural ensembles ensure reliable coding remains unknown. Our model proposes that these correlated ensembles could function as normalization pools. Divisive normalization is a canonical cortical computation (Carandini and Heeger, 2012) that is believed to be crucial for efficient coding (Schwartz and Simoncelli, 2001). One possibility is that normalization pools average responses from neurons and in doing so suppress cortical variability (Shadlen and Newsome, 1994). This has led to the notion that redundancy, through neuronal coupling, plays an important role in visual processing (Lin et al., 2015; Okun et al., 2015). Our work bolsters this idea by demonstrating that normalization pools can be dynamically recruited, in a stimulus-dependent manner, to ensure reliable and efficient coding.

Thus, our work establishes that stimulus correlations function in a manner similar to attention or neuromodulation to protect the neural code from intrinsic variability by increasing coupling between neurons (Moreno-Bote et al., 2014). In particular, by rapidly recruiting neural ensembles, scenes with naturalistic spatial correlations switch coding from a "high-noise" to a "low-noise" regime, where intrinsic variability is suppressed to permit reliable coding.

In conclusion, our study provides a novel insight into how the unique statistical features of the natural environment can modulate coding in V1. We provide strong evidence that spatial correlations are an important feature of natural scenes because of their role in shaping inter-neuronal correlations. While we have focused on V1, we believe that similar coding strategies could also operate in other sensory modalities (Bandyopadhyay et al., 2010; Hires et al., 2015), consistent with this being a general principle of cortical computations.

3.5 References

Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z.J., Scanziani, M., 2012. A neural circuit for spatial summation in visual cortex. Nature 490, 226–31. doi:10.1038/nature11526

Atick, J.J., Redlich, A.N., 1990. Towards a Theory of Early Visual Processing. Neural Comput. 2, 308–320. doi:10.1162/neco.1990.2.3.308

Averbeck, B.B., Latham, P.E., Pouget, A., 2006. Neural correlations, population coding and computation. Nat. Rev. Neurosci. 7, 358–66. doi:10.1038/nrn1888

Azouz, R., Gray, C.M., 1999. Cellular mechanisms contributing to response variability of cortical neurons in vivo. J. Neurosci. 19, 2209–2223.

Bandyopadhyay, S., Shamma, S.A., Kanold, P.O., 2010. Dichotomy of functional organization in the mouse auditory cortex. Nat. Neurosci. 13, 361–8. doi:10.1038/nn.2490

Bar, M., 2004. Visual objects in context. Nat. Rev. Neurosci. 5, 617–29. doi:10.1038/nrn1476

Barlow, H., 2001. The exploitation of regularities in the environment by the brain. Behav. Brain Sci. 24, 602–607.

Bathellier, B., Ushakova, L., Rumpel, S., 2012. Discrete neocortical dynamics predict behavioral categorization of sounds. Neuron 76, 435–449. doi:10.1016/j.neuron.2012.07.008

Baudot, P., Levy, M., Marre, O., Monier, C., Pananceau, M., Frégnac, Y., 2013. Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. Front. Neural Circuits 7, 206. doi:10.3389/fncir.2013.00206

Bayati, H., Davoudi, H., Fatemizadeh, E., 2008. A heuristic method for finding the optimal number of clusters with application in medical data. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf. 2008, 4684–7. doi:10.1109/IEMBS.2008.4650258

Bonin, V., Histed, M.H., Yurgenson, S., Reid, R.C., 2011. Local diversity and fine-scale organization of receptive fields in mouse visual cortex. J. Neurosci. 31, 18506–21.

doi:10.1523/JNEUROSCI.2974-11.2011

Bonin, V., Mante, V., Carandini, M., 2006. The statistical computation underlying contrast gain control. J Neurosci 26, 6346–6353. doi:10.1523/JNEUROSCI.0284-06.2006

Borst, A., Theunissen, F.E., 1999. Information theory and neural coding. Nat. Neurosci. 2, 947–57. doi:10.1038/14731

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–6.

Butts, D. a, Weng, C., Jin, J., Yeh, C.-I., Lesica, N. a, Alonso, J.-M., Stanley, G.B., 2007. Temporal precision in the neural code and the timescales of natural vision. Nature 449, 92–5. doi:10.1038/nature06105

Butts, D.A., Weng, C., Jin, J., Yeh, C.-I., Lesica, N.A., Alonso, J.-M., Stanley, G.B., 2007. Temporal precision in the neural code and the timescales of natural vision. Nature 449, 92–5. doi:10.1038/nature06105

Carandini, M., Heeger, D.J., 2012. Normalization as a canonical neural computation. Nat Rev Neurosci 13, 51–62. doi:10.1038/nrn3136

Carandini, M., Heeger, D.J., Movshon, J. a, 1997. Linearity and normalization in simple cells of the macaque primary visual cortex. J. Neurosci. 17, 8621–44.

Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., Looger, L.L., Svoboda, K., Kim, D.S., 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295–300. doi:10.1038/nature12354

Churchland, M.M., Yu, B.M., Cunningham, J.P., Sugrue, L.P., Cohen, M.R., Corrado, G.S., Newsome, W.T., Clark, A.M., Hosseini, P., Scott, B.B., Bradley, D.C., Smith, M.A., Kohn, A., Movshon, J.A., Armstrong, K.M., Moore, T., Chang, S.W., Snyder, L.H., Lisberger, S.G., Priebe, N.J., Finn, I.M., Ferster, D., Ryu, S.I., Santhanam, G., Sahani, M., Shenoy, K. V, 2010. Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat. Neurosci. 13, 369–78. doi:10.1038/nn.2501

Cohen, M.R., Kohn, A., 2011. Measuring and interpreting neuronal correlations. Nat. Neurosci.

14, 811–9. doi:10.1038/nn.2842

Cohen, M.R., Maunsell, J.H.R., 2009. Attention improves performance primarily by reducing interneuronal correlations. Nat. Neurosci. 12, 1594–600. doi:10.1038/nn.2439

Cossell, L., Iacaruso, M.F., Muir, D.R., Houlton, R., Sader, E.N., Ko, H., Hofer, S.B., Mrsic-Flogel, T.D., 2015. Functional organization of excitatory synaptic strength in primary visual cortex. Nature. doi:10.1038/nature14182

Dan, Y., Atick, J.J., Reid, R.C., 1996. Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. J. Neurosci. 16, 3351–62.

Desbordes, G., Jin, J., Weng, C., Lesica, N. a, Stanley, G.B., Alonso, J.-M., 2008. Timing precision in population coding of natural scenes in the early visual system. PLoS Biol. 6, e324. doi:10.1371/journal.pbio.0060324

Dong, D., Atick, J., 1995. Statistics of natural time-varying images. Netw. Comput. Neural Syst. 6, 345–358. doi:10.1088/0954-898X/6/3/003

Ecker, A.S., Berens, P., Tolias, A.S., Bethge, M., 2011. The effect of noise correlations in populations of diversely tuned neurons. J. Neurosci. 31, 14272–83. doi:10.1523/JNEUROSCI.2539-11.2011

El-Boustani, S., Sur, M., 2014. Response-dependent dynamics of cell-specific inhibition in cortical networks in vivo. Nat. Commun. 5, 5689. doi:10.1038/ncomms6689

Fournier, J., Monier, C., Pananceau, M., Frégnac, Y., 2011. Adaptation of the simple or complex nature of V1 receptive fields to visual statistics. Nat. Neurosci. 14, 1053–60. doi:10.1038/nn.2861

Freeman, J., Ziemba, C.M., Heeger, D.J., Simoncelli, E.P., Movshon, J.A., 2013. A functional and perceptual signature of the second visual area in primates. Nat. Neurosci. 16, 974–981. doi:10.1038/nn.3402

Froudarakis, E., Berens, P., Ecker, A.S., Cotton, R.J., Sinz, F.H., Yatsenko, D., Saggau, P., Bethge, M., Tolias, A.S., 2014. Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. Nat. Neurosci. doi:10.1038/nn.3707

Gao, E., DeAngelis, G.C., Burkhalter, A., 2010. Parallel input channels to mouse primary visual cortex. J. Neurosci. 30, 5912–26. doi:10.1523/JNEUROSCI.6456-09.2010

Garrett, M.E., Nauhaus, I., Marshel, J.H., Callaway, E.M., 2014. Topography and areal organization of mouse visual cortex. J. Neurosci. 34, 12587–600. doi:10.1523/JNEUROSCI.1124-14.2014

Gaspar, C.M., Rousselet, G.A., 2009. How do amplitude spectra influence rapid animal detection? Vision Res. 49, 3001–12. doi:10.1016/j.visres.2009.09.021

Goard, M., Dan, Y., 2009. Basal forebrain activation enhances cortical coding of natural scenes. Nat. Neurosci. 12, 1444–9. doi:10.1038/nn.2402

Haider, B., Häusser, M., Carandini, M., 2013. Inhibition dominates sensory responses in the awake cortex. Nature 493, 97–100. doi:10.1038/nature11665

Haider, B., Krause, M.R., Duque, A., Yu, Y., Touryan, J., Mazer, J. a, McCormick, D. a, 2010. Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. Neuron 65, 107–21. doi:10.1016/j.neuron.2009.12.005

Harris, K.D., Mrsic-Flogel, T.D., 2013. Cortical connectivity and sensory coding. Nature 503, 51–8. doi:10.1038/nature12654

Hires, S.A., Gutnisky, D.A., Yu, J., Connor, D.H.O., Svoboda, K., 2015. Low-noise encoding of active touch by layer 4 in the somatosensory cortex 1–18. doi:10.7554/eLife.06619

Hofer, S.B., Ko, H., Pichler, B., Vogelstein, J., Ros, H., Zeng, H., Lein, E., Lesica, N. a, Mrsic-Flogel, T.D., 2011. Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. Nat. Neurosci. 14, 1045–52. doi:10.1038/nn.2876

Hosoya, T., Baccus, S. a, Meister, M., 2005. Dynamic predictive coding by the retina. Nature 436, 71–7. doi:10.1038/nature03689

Kampa, B.M., Roth, M.M., Göbel, W., Helmchen, F., 2011. Representation of visual scenes by local neuronal populations in layer 2/3 of mouse visual cortex. Front. Neural Circuits 5, 18. doi:10.3389/fncir.2011.00018

Kara, P., Reinagel, P., Reid, R.C., 2000. Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. Neuron 27, 635–46.

Kayser, C., Einhäuser, W., König, P., 2003. Temporal correlations of orientations in natural scenes. Neurocomputing 52-54, 117–123. doi:10.1016/S0925-2312(02)00789-0

Kohn, A., Smith, M. a, 2005. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. J. Neurosci. 25, 3661–73. doi:10.1523/JNEUROSCI.5106-04.2005

Lesica, N. a, Jin, J., Weng, C., Yeh, C.-I., Butts, D. a, Stanley, G.B., Alonso, J.-M., 2007. Adaptation to stimulus contrast and correlations during natural visual stimulation. Neuron 55, 479–91. doi:10.1016/j.neuron.2007.07.013

Lin, I.-C., Okun, M., Carandini, M., Harris, K.D., 2015. The Nature of Shared Cortical Variability. Neuron 87, 644–56. doi:10.1016/j.neuron.2015.06.035

Mainen, Z.F., Sejnowski, T.J., 1995. Reliability of spike timing in neocortical neurons. Science 268, 1503–1506. doi:10.1126/science.7770778

McCotter, M., Gosselin, F., Sowden, P., Schyns, P., 2005. The use of visual information in natural scenes. Vis. cogn. 12, 938–953. doi:10.1080/13506280444000599

Miller, J.-E.K., Ayzenshtat, I., Carrillo-Reid, L., Yuste, R., 2014. Visual stimuli recruit intrinsically generated cortical ensembles. Proc. Natl. Acad. Sci. U. S. A. doi:10.1073/pnas.1406077111

Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., Pouget, A., 2014. Information-limiting correlations. Nat. Neurosci. 17, 1410–7. doi:10.1038/nn.3807

Movshon, J.A., Thompson, I.D., Tolhurst, D.J., 1978. Spatial summation in the receptive fields of simple cells in the cat's striate cortex. J. Physiol. 283, 53–77.

Niell, C.M., Stryker, M.P., 2008. Highly selective receptive fields in mouse visual cortex. J. Neurosci. 28, 7520–36. doi:10.1523/JNEUROSCI.0623-08.2008

Okun, M., Steinmetz, N.A., Cossell, L., Iacaruso, M.F., Ko, H., Barthó, P., Moore, T., Hofer, S.B., Mrsic-Flogel, T.D., Carandini, M., Harris, K.D., 2015. Diverse coupling of neurons to populations in sensory cortex. Nature. doi:10.1038/nature14273

Oliva, A., Torralba, A., 2007. The role of context in object recognition. Trends Cogn. Sci. 11, 520–7. doi:10.1016/j.tics.2007.09.009

Oliva, A., Torralba, A., 2006. Building the gist of a scene: the role of global image features in recognition. Prog. Brain Res. 155, 23–36. doi:10.1016/S0079-6123(06)55002-2

Olshausen, B.A., Field, D.J., 2005. How close are we to understanding v1? Neural Comput. 17, 1665–99. doi:10.1162/0899766054026639

Olshausen, B.A., Field, D.J., 2004. Sparse coding of sensory inputs. Curr. Opin. Neurobiol. 14, 481–7. doi:10.1016/j.conb.2004.07.007

Pecka, M., Han, Y., Sader, E., Mrsic-Flogel, T.D., 2014. Experience-Dependent Specialization of Receptive Field Surround for Selective Coding of Natural Scenes. Neuron. doi:10.1016/j.neuron.2014.09.010

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10, 437–42.

Renart, A., Machens, C.K., 2014. Variability in neural activity and behavior. Curr. Opin. Neurobiol. 25, 211–220. doi:10.1016/j.conb.2014.02.013

Ringach, D.L., Hawken, M.J., Shapley, R., 2002. Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. J. Vis. 2, 12–24. doi:10:1167/2.1.2

Ruderman, D., Bialek, W., 1994. Statistics of natural images: Scaling in the woods. Phys. Rev. Lett. 73, 814–817.

Sceniak, M.P., Hawken, M.J., Shapley, R., 2002. Contrast-dependent changes in spatial frequency tuning of macaque V1 neurons: effects of a changing receptive field size. J. Neurophysiol. 88, 1363–73.

Schwartz, O., Simoncelli, E.P., 2001. Natural signal statistics and sensory gain control. Nat. Neurosci. 4, 819–25. doi:10.1038/90526

Segal, I.Y., Giladi, C., Gedalin, M., Rucci, M., Ben-Tov, M., Kushinsky, Y., Mokeichev, A., Segev, R., 2015. Decorrelation of retinal response to natural scenes by fixational eye

movements. Proc. Natl. Acad. Sci. 112, 3110–3115. doi:10.1073/pnas.1412059112

Shadlen, M.N., Newsome, W.T., 1998. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. 18, 3870–96.

Shadlen, M.N., Newsome, W.T., 1994. Noise, neural codes and cortical organization. Curr. Opin. Neurobiol. 4, 569–79.

Simoncelli, E.P., Olshausen, B.A., 2001. Natural image statistics and neural representation. Annu. Rev. Neurosci. 24, 1193–216. doi:10.1146/annurev.neuro.24.1.1193

Singh, A., Lesica, N.A., 2010. Incremental mutual information: a new method for characterizing the strength and dynamics of connections in neuronal circuits. PLoS Comput. Biol. 6, e1001035. doi:10.1371/journal.pcbi.1001035

Snyder, A.C., Morais, M.J., Kohn, A., Smith, M.A., 2014. Correlations in V1 Are Reduced by Stimulation Outside the Receptive Field 34, 11222–11227. doi:10.1523/JNEUROSCI.0762-14.2014

Stevens, C.F., Zador, A.M., 1998. Input synchrony and the irregular firing of cortical neurons. Nat. Neurosci. 1, 210–217. doi:10.1038/659

Tiesinga, P., Fellous, J.-M., Sejnowski, T.J., 2008. Regulation of spike timing in visual cortical circuits. Nat. Rev. Neurosci. 9, 97–107. doi:10.1038/nrn2315

Tolhurst, D.J., Smyth, D., Thompson, I.D., 2009. The sparseness of neuronal responses in ferret primary visual cortex. J. Neurosci. 29, 2355–70. doi:10.1523/JNEUROSCI.3869-08.2009

Torralba, A., Oliva, A., 2003. Statistics of natural image categories. Network 14, 391–412.

van Hateren, J.H., Ruderman, D.L., 1998. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. Proc Biol Sci 265, 2315–2320. doi:10.1098/rspb.1998.0577

van Hateren, J.H., van der Schaaf, A., 1998. Independent component filters of natural images compared with simple cells in primary visual cortex. Proc. Biol. Sci. 265, 359–66. doi:10.1098/rspb.1998.0303

Vinje, W.E., Gallant, J.L., 2002. Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. J Neurosci 22, 2904–2915. doi:20026216

Vinken, K., Vermaercke, B., Op de Beeck, H.P., 2014. Visual categorization of natural movies by rats. J. Neurosci. 34, 10645–58. doi:10.1523/JNEUROSCI.3663-13.2014

Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R., Paninski, L., 2010. Fast nonnegative deconvolution for spike train inference from population calcium imaging. J. Neurophysiol. 104, 3691–704. doi:10.1152/jn.01073.2009

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Trans. Image Process. 13, 600–612. doi:10.1109/TIP.2003.819861

Wichmann, F.A., Drewes, J., Rosas, P., Gegenfurtner, K.R., 2010. Animal detection in natural scenes: critical features revisited. J Vis 10, 6 1–27. doi:10.1167/10.4.6

Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., Tanaka, J.W., 2010. Controlling low-level image properties: the SHINE toolbox. Behav. Res. Methods 42, 671–84. doi:10.3758/BRM.42.3.671

Willmore, B.D.B., Mazer, J. a, Gallant, J.L., 2011. Sparse coding in striate and extrastriate visual cortex. J. Neurophysiol. 105, 2907–19. doi:10.1152/jn.00594.2010

Wilson, N.R., Runyan, C. a, Wang, F.L., Sur, M., 2012. Division and subtraction by distinct cortical inhibitory networks in vivo. Nature 488, 343–8. doi:10.1038/nature11347

Wilson, N.R., Schummers, J., Runyan, C.A., Yan, S.X., Chen, R.E., Deng, Y., Sur, M., 2013. Two-way communication with neural networks in vivo using focused light. Nat. Protoc. 8, 1184–203. doi:10.1038/nprot.2013.063

Zoccolan, D., 2015. Invariant visual object recognition and shape processing in rats. Behav. Brain Res. 285, 10–33. doi:10.1016/j.bbr.2014.12.053

Zohary, E., Shadlen, M.N., Newsome, W.T., 1994. Correlated neuronal discharge rate and its implications for psychophysical performance. Nature 370, 140–3. doi:10.1038/370140a0

Appendix 3A – Experimental Procedures

Experiments

Animals and Surgery

Experiments were carried out under protocols approved by MIT's Animal Care and Use Committee and conformed to NIH guidelines. All data in this study were collected from adult (>8 weeks old), C57/BL6 (Charles River Laboratory) mice of either sex. Mice were anesthetized using isoflurane (3% induction, 1.5-2% during surgery). A custom-built metal head-post was attached to the skull using dental cement (C&B Metabond, Parkell) and a 3mm diameter craniotomy was performed over binocular V1 (approximately 2-3mm lateral and 0.5mm anterior to lambda). Care was taken not to rupture the dura mater. The core body temperature was maintained at 37.5°C using a heating blanket (Harvard Apparatus).

For anesthetized experiments, anesthesia was maintained with 0.5-0.8% isoflurane during imaging. A solution of Oregon Green Bapta-1 AM (OGB1, 1mM, Molecular Probes) and SR-101 (100μM, Molecular Probes) was pressure injected (10 PSI for 1 minute; Picospritzer) into the brain 180-250μm below the pial surface via a borosilicate glass pipette (5-7MΩ tip resistance) under visual guidance. After confirming successful expression of OGB1 (~20 minutes following injection), the craniotomy was covered with a 3mm-glass coverslip (Warner Instruments) and sealed with a silicone elastomer (Kwik-Sil, WPI).

For awake experiments, mice were first habituated for 5 days to head-fixation on a custom-built stage. Once habituated, the mice received a microinjection of 100-200nl of AAV1.Syn.GCaMP6f.WPRE.SV40 (University of Pennsylvania Vector Core, diluted to a titer of 10^{12} genomes ml^{-1}), following which a cranial window was implanted over the craniotomy and sealed, as described above. Mice were allowed to recover for 2-3 weeks to allow for adequate expression of the virus before imaging commenced. It has been previously shown that both OGB1 and GCaMP6f have similar response kinetics

(Chen et al., 2013), permitting us to make direct comparisons of the influence of brain state on coding.

Two-photon calcium imaging and analysis

Imaging was performed using a Prairie Ultima two-photon system (Prairie Technologies) driven by a Spectra Physics Mai-Tai laser, passed through a Deep-See module (Spectra Physics) and a high performance objective lens (25x Olympus XL Plan N objective, NA = 1.05). Cells were excited at 960nm for OGB1 and 920nm for GCaMP6f.

A custom-built MATLAB-based (MathWorks) software system was used to collect optimized raster scans at 50 frames/s and to perform offline data analysis, as described previously (Wilson et al., 2012; Wilson et al., 2013). Briefly, image segmentation algorithms were first used to identify cell bodies from a scanned image (see **Fig. 2A**). Next, a genetic algorithm was used to determine the shortest scan path between cells, and this line scan was run at 50Hz, ensuring a 90% dwell time inside the cells. This higher dwell time ensured calcium transients with slightly larger amplitudes (ΔF/F range: 20-50%). Frames with excessive brain movement were ignored.

Significantly visually responsive cells were determined from the fluorescence time changes (ΔF/F) by performing a one-tailed Student's t-test between visually evoked and spontaneous responses (gray screen collected for 2 minutes before start of experiment). Only cells with $p < 10^{-3}$ were classified as visually-responsive. Firing rates of these cells were then inferred using a fast nonnegative deconvolution algorithm (Vogelstein et al., 2010) using parameters that were previously verified in our lab (El-Boustani and Sur, 2014; Wilson et al., 2012). The Vogelstein algorithm infers the probability of spiking from calcium transients. To convert this probability into a firing rate (measured in events/s), we multiplied each probability by 50Hz, the frequency at which the calcium transients were sampled. Unless otherwise stated, all data analysis was performed using inferred firing rates.

Visual Stimuli

Creation of noise movies

We developed an algorithm that allowed us to create noise images with a user-defined spectral slope. To do so, we took advantage of the inverse-square law: $P \sim k^{-\alpha}$, which translates to a circle with radius α in two-dimensional Fourier space. Thus, we constructed all noise movies in the Fourier domain. We first defined a matrix of the same size as the original image (256x256 pixels) and then created a noise amplitude spectrum as a 2-D circle of radius α, with α taking values from 0 (*K0* movie) to $\sqrt{2}$ (*K2* movie). This was due to the squared-relationship between the amplitude spectrum and the power spectrum. To create the final noise image, we combined this noise amplitude spectrum with a random phase spectrum, where values were randomly sampled from the range (0-2π). The final noise images were visualized by computing its 2-D inverse Fourier transform. Each frame of the noise movie was created using a new random seed and, as a result, the raw noise movies had no temporal correlations between frames.

Noise-masking procedure

Figure 3-1A provides a schematic of the noise-masking procedure. First, each frame of a natural movie was decomposed into its Fourier components (phase and amplitude) via a 2-D Fast Fourier Transform implemented in MATLAB. Next, a noise image was created as described above. The phase spectrum of the original movie was then combined with the amplitude spectrum of the noise movie. The resulting image was then inverse Fourier transformed to yield a noise-masked movie frame. This procedure was repeated for all frames. We used a total of five different natural movies, each 4s in duration, from the van Hateren movie database (van Hateren and Ruderman, 1998).

Adjusting image statistics.

Gray scale values of each movie frame were discretized to 255 values and each frame was normalized to have equal mean luminance (mean of luminance histogram = 128, mean luminance = 77 cd/m^2, Luminance Range = 0.02-134 cd/m^2) and contrast (standard deviation of luminance histogram = 32). This normalization was carried out

using custom written MATLAB code together with the *lumMatch* function in the SHINE toolbox (Willenbockel et al., 2010). To minimize differences between the different original movies, we used the *sfMatch* function to ensure that all original movies had the same SF distribution and amplitude spectra ($\alpha = 1.3$, which was the average spectral slope of the selected movie database). We also used the *histMatch* function to match the histogram of an image with a target (set to the original movie) by remapping pixel values to how frequently they occur in the target histogram. This ensured that all pixel statistics, such as mean, contrast, kurtosis and skewness, were the same for both the noise movies and the original movies. Additionally, we normalized each power spectrum to have the same integral. That is all power spectra have the same total SF content. The noise-masking technique only reshuffles power into different bands, without changing the total SF content.

To slow down the movie from its original 60Hz frame rate, we updated every three frames creating an effective frame rate of 20Hz. The monitor refresh rate was fixed at 60Hz. All visual stimuli were displayed on a 23" gamma-corrected LCD monitor (Dell) covering a visual space of approximately 96x54 degrees^2. Stimuli were displayed using the *PsychoPhysics* Toolbox (Brainard, 1997; Pelli, 1997).

Image quality metrics.

We assessed the impact of perturbing spatial correlations on the quality of image using the structural similarity index (SSIM). The SSIM uses image structural information, such as mean, variance and covariance, to estimate dependencies between pixels (Wang et al., 2004). Specifically, we computed SSIM between images i and j using the following equation:

$$SSIM(i,j) = \frac{\left(2\mu_i\mu_j + (k_iL)^2\right)\left(2\sigma_{ij} + \left(k_jL\right)^2\right)}{\left(\mu_i^2 + \mu_j^2 + (k_iL)^2\right)\left(\sigma_i^2 + \sigma_j^2 + \left(k_jL\right)^2\right)}$$

where, $\mu_{i,j}$ and $\sigma_{i,j}$ is the mean and standard deviation of images i and j respectively, σ_{ij} is the covariance and L is the dynamic range of the image. Further details on the SSIM metric are provided in (Wang et al., 2004).

Data Analysis

Response similarity, sparseness and reliability analysis.

Let the response of a neuron to trial *i* of movie A be $f_{i,A}$; then the response similarity index (SI, **Figure 3-2**) between natural movie A and a noise movie B was calculated using the equation:

$$SI_{A-B} = \frac{1}{T^2}\sum_{i=1}^{T}\rho(f_{i,A}, f_{i,B})$$

where, $\rho(f_{i,A}, f_{i,B})$ is the Pearson correlation coefficient and T is the total number of trials. From this equation, SI is the average correlation of all possible pairwise combinations of single trial response vectors of two movies. Lifetime sparseness (selectivity, **Figure 3-2**) was computed using the equation:

$$S_A = \frac{N - \frac{\left(\sum_j \langle f_{i,A}\rangle_j\right)^2}{\sum_j \langle f_{i,A}\rangle_j^2}}{N-1}$$

where $\langle f_{i,A}\rangle_j$ is the trial-averaged response to frame j of movie A and N is the total number of movie frames.

Similarly, response reliability to movie A (R_A) was calculated using the equation:

$$R_A = \frac{2}{T^2 - T}\sum_{i=1}^{T}\sum_{j=i+1}^{T}\rho(f_{i,A}, f_{j,A})$$

Thus, the response reliability (**Figure 3-3**) is the average correlation of all pairwise combinations of trials for a single movie. Only neurons with significant responses on more than 5 trials were selected for this analysis.

To compute cluster activation reliability (**Figure 3-10C**), we used the same equation above but instead let $f_{i,A}$ be the population-averaged firing rate (i.e. averaged over all neurons in a population) of neurons on trial *i* of movie A.

Mapping neuron spatial frequency preferences.

In experiments where we mapped the spatial frequency tuning of neurons (**Figures 3-8 – 3-10**), we presented alternating blocks of noise-masked movies and sinusoidal gratings at full contrast. Here, we used gratings with 2Hz temporal frequency, at 8 different orientations (0-180°) and at 9 different spatial frequencies (0.01 to 0.32

cycles/degree). Each grating was presented for 3s and was flanked by 1s gray screens. We quantified spatial frequency responses by fitting the following empirical difference-of-Gaussians function to the spatial frequency tuning curves (Sceniak et al., 2002).

$$R(sf) = R_0 + R_E e^{-[(sf-\mu)^2/\sigma_E^2]} - R_I e^{-[(sf-\mu)^2/\sigma_I^2]}$$

Here R_0 is the spontaneous rate, measured from the 1s blank screen epochs. The parameters ($R_{E,I}$, μ and σ) were optimized to provide a least-square error fit of the data using the MATLAB function *lsqcurvefit* with the Levenberg-Marquardt algorithm. The quality of fit was assessed by calculating the adjusted-R^2 (coefficient of determination) value as follows: $R^2 = \frac{\sum_i (f-\hat{y})}{\sum_i (y-\hat{y})}$, where f is the fit and y is the raw data. Only neurons with fits better than 75% were selected for further analysis. The preferred spatial frequency and bandwidth was determined empirically from the fitted curves. We defined the preferred spatial frequency as the spatial frequency that elicits the maximal response (i.e. Preferred SF $= \mu$). The bandwidth (BW, measured in octaves) was computed as the log ratio of the SFs that elicited half the maximal response for the high-frequency cutoff to the low-frequency cutoff: $\mathrm{BW} = \log_2(SF_{high}/SF_{low})$.

Mapping receptive fields and measuring cortical magnification.

In experiments where we mapped the RF centers of neurons (**Figure 3-12A**), we presented sparse noise stimuli, which consisted of black and white squares (4x4 degrees each, 1 pixel corresponded to 0.8 degrees) on an isoluminant gray background (128 on a 256 gray scale). Each square was presented at a random location (7x12 grid) for 200ms followed by a 300ms blank period. The location of each black/white square was chosen from a pseudorandom distribution such that two consecutive squares were at least three nodes away from each other. Reponses to white squares was use to calculate the ON receptive field and responses to black squares were used to calculate the OFF receptive field following methods described elsewhere (Smith and Hausser, 2010). We computed the cortical magnification factor (CMF) in both the rostro-caudal (CMF_{R-C}) and medio-lateral (CMF_{M-L}) axes by performing a linear regression between RF position (in degrees) and neuron position (in μm) and computing the slope of the

best-fit regression line. The overall CMF was computed using the following equation: $CMF = CMF_{M-L} \times CMF_{R-C}$. For neural populations in which we were not able to calculate the CMF (**Figure 3-14B**), we used a CMF of $1x10^{-3}$ mm^2/deg^2 to scale between cortical space and visual space. This value is consistent with previously published reports (Garrett et al., 2014).

Analysis of signal and noise correlation between neurons.

Signal correlations between pairs of neurons were calculated as the Pearson correlation coefficient between trial-averaged responses binned at 200ms. To compute noise correlations, we first subtracted the trial average from responses in each trial and then computed the Pearson correlation coefficient (CC) between these mean-subtracted responses, again binned at 200ms (Singh and Lesica, 2010).

To compute signal and noise correlation decay functions (**Figure 3-14**), we first binned neurons according to the pairwise Euclidean distance between their centroids, determined from the imaging software. On average a majority of neurons were located within 20-180μm of each other, and only <1% were separated by a distance of 300μm. Next, we computed the median CC within each distance bin. We adjusted the bin width (range: 30-50μm) to ensure at least 5 neurons per bin. Data were then pooled from different experiments by first normalizing the CC to the first distance bin and then by computing a bootstrapped estimate of the median CC and its 95% confidence interval for each subsequent bin. To estimate the decay constant, we fit single exponential functions to each correlation decay function using a least-squares algorithm, and assessed fit using an adjusted-R^2 value. Only populations with $R^2 > 80\%$ were selected for further analysis.

Clustering analysis.

Clustering analysis (**Figure 3-13, 3-14**) was used to visually represent the correlation structure of the network. Clustering analysis was performed on all neurons in the imaged population and was repeated for each spectral condition. We first arranged either signal or noise correlation values in a matrix, where each element is the CC for a

pair of neurons. Next, an agglomerative hierarchical clustering algorithm was used to cluster this matrix by maximizing between-cluster variance and minimizing within-cluster variance. Inter-cluster linkages were formed using the Ward metric and within-cluster linkages were formed by minimizing the Euclidean distance between CC. This was achieved via a custom written code that called built-in MATLAB functions (*linkage, pdist* and *cluster*). The optimal number of clusters ($\hat{n}$, typically ranging between 2-4) were selected using the Calinski-Harabasz metric (Bayati et al., 2008), which minimizes within-cluster variance and maximized between-cluster variance according to the formula:

$$\hat{n} = \underset{n}{\operatorname{argmax}} \left(\frac{SSB/(n-1)}{SSW/(N-n)} \right)$$

where $SSB = \sum_i (x_i - \hat{x})^2$ and $SSW = \sum_{j=1}^{N} \sum_i (x_i - c\,(j))^2$ are the between- and within-cluster sum of squares respectively, x_i is the correlation coefficient of the i^{th} element and N is the maximum number of clusters, which was set to 4. The index i references all elements within a cluster while the index j references the clusters. Neurons belonging to each cluster were color-coded as shown in **Figure 3-13**.

Discrimination analysis.

We used a template-matching algorithm to determine if a single-trial response could be matched to one of five population templates (the five movies used in each spectral condition, **Figure 3-15**). This unsupervised classification was based on previous studies (Goard and Dan, 2009; Kampa et al., 2011). First, we created population templates for each movie by averaging (over trials) the responses of all neurons in a population. Next, individual trials for each neuron were assigned to a particular stimulus category by minimizing the Euclidean distance between the template and the single trial response. The classified stimulus identities were then compared to the stimulus presentation to compute the percentage of correctly classified trials for each neuron. To test the dependence of the number of neurons on classification, we used a Monte Carlo sampling technique (repeated 500 times) to pick *n* neurons (range: 1 to population size)

at random from the population with replacement. The single trial responses from these *n* neurons were compared to the template as described above. The decoding accuracy was quantified by computing mutual information (MI) using the equation:

$$MI = \sum_{i\in\{1,5\}} \sum_{j\in\{1,5\}} p_{ij} \log_2 \left(\frac{p_{ij}}{p_i p_j} \right)$$

where, p_{ij} is the probability of observing movie i given that the true label is j.

V1 Linear-Nonlinear-Normalization Model.

We first determined a best-fit RF estimate from neural responses to two original movies using a bank of linear and nonlinear filters (**Figure 3-15**). The spatial part of the linear filter bank were log-Gabor RFs, at 6 different orientations (0-180°) and ranged in size from 12-18° of visual angle (Bonin et al., 2011). Because we did not know the locations of the RFs *a priori*, we randomly picked 50 possible locations on the screen. The temporal part of the linear bank consisted of gamma functions with a range of temporal delays (140-200ms). Responses from each branch of this filter bank was summed and then passed through a point-wise exponential nonlinearity to provide a response estimate. We used a least squares method to determine the best-fit model. Only neurons with adjusted-R^2 greater than 80% were selected for further analysis.

In the independent model, we convolved each noise movie frame with the best-fit RF model for each neuron. In the normalization model, we modulated the output of each independent model (f) using the divisive normalization rule:

$$R = \frac{f^2}{\sigma^2 + \sum_j f_j^2}$$

where j indexes neurons in either the high or low correlation cluster. The adjusted-R^2 was computed between the predicted and actual response to assess the percentage of variance explained by each model.

Statistical Analysis.

All statistical analysis was performed using custom written scripts in MATLAB and R. No tests were conducted to determine sample size. Data from anesthetized experiments

came from 16 mice, yielding a total of 1006 neurons. For reliability analysis, we picked neurons (650 neurons, 16 mice) that had significant visually-evoked responses from 6 or more trials. These same 650 neurons were used for all correlation analysis. For SF tuning analysis, we collected data from 308 neurons for 4 mice. In awake experiments, we obtained data from 230 neurons in 4 mice.

Data were first tested for normality using the Shapiro-Wilk Test. All data presented in this paper are non-normally distributed, thus all statistical tests were conducted using non-parametric statistics. Our experiments involved testing the influence of different movies on the same population of neurons, thus all comparisons were performed using non-parametric repeated-measures ANOVA (Friedman Test) with Bonferroni's correction for multiple comparisons. For Bonferroni corrections, the significance value was set to 0.05. Post-hoc tests were performed using the two-tailed Wilcoxon rank-sum test. All other statistical tests that were performed are described in the text. The 95% confidence intervals (CI) of the median were computed by bootstrapping. In most figures, data are presented as median ± 95% CI, unless otherwise stated. Visually, non-overlapping 95% CI imply that the data is significant to at least $p<0.05$. Symbols used in figures: *$p<0.05$, **$p < 10^{-3}$, ***$p < 10^{-4}$, N.S., non-significant.

Author Contributions

Rajeev Rikhye and Mriganka Sur designed the study. Rajeev Rikhye performed all the experiments and analyzed all the data. Rajeev Rikhye and Mriganka Sur wrote the manuscript. The manuscript was reformatted and rewritten for this chapter.

Additional References

Bayati, H., Davoudi, H., Fatemizadeh, E., 2008. A heuristic method for finding the optimal number of clusters with application in medical data. Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf. 2008, 4684–7. doi:10.1109/IEMBS.2008.4650258

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–6.

Chen, T.-W., Wardill, T.J., Sun, Y., Pulver, S.R., Renninger, S.L., Baohan, A., Schreiter, E.R., Kerr, R.A., Orger, M.B., Jayaraman, V., Looger, L.L., Svoboda, K., Kim, D.S., 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature 499, 295–300. doi:10.1038/nature12354

Garrett, M.E., Nauhaus, I., Marshel, J.H., Callaway, E.M., 2014. Topography and areal organization of mouse visual cortex. J. Neurosci. 34, 12587–600. doi:10.1523/JNEUROSCI.1124-14.2014

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. 10, 437–42.

Sceniak, M.P., Hawken, M.J., Shapley, R., 2002. Contrast-dependent changes in spatial frequency tuning of macaque V1 neurons: effects of a changing receptive field size. J. Neurophysiol. 88, 1363–73.

Singh, A., Lesica, N.A., 2010. Incremental mutual information: a new method for characterizing the strength and dynamics of connections in neuronal circuits. PLoS Comput. Biol. 6, e1001035. doi:10.1371/journal.pcbi.1001035

van Hateren, J.H., Ruderman, D.L., 1998. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. Proc Biol Sci 265, 2315–2320. doi:10.1098/rspb.1998.0577

Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Trans. Image Process. 13, 600–612. doi:10.1109/TIP.2003.819861

Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., Tanaka, J.W., 2010. Controlling low-level image properties: the SHINE toolbox. Behav. Res. Methods 42, 671–84. doi:10.3758/BRM.42.3.671

This page intentionally left blank

Chapter 4

Which inhibitory mechanisms are responsible for reliable coding?[1]

Summary

Neurons in the primary visual cortex (V1) respond to repeated presentations of natural scenes with spike trains that are highly reliable from trial-to-trial. While it has been proposed that inhibition plays a major role in sculpting pyramidal cell reliability, the precise contribution of different interneuron subtypes remains unknown. In this study, we investigated the roles of Parvalbumin (PV) and Somatostatin (SST) expressing interneurons in modulating reliable coding of natural movies in awake, head-fixed mice. Specifically, we were interested in studying how interactions between SST and PV neurons sculpted the reliability of pyramidal neurons. To this end, we developed a new line of double transgenic mice, which allowed us genetic access to both SST and PV neurons. Dual color imaging revealed that these interneurons were temporally coupled in reliably processed movies. Optical activation of SST neurons improved the reliability of pyramidal neurons, and at the same time suppressed PV neurons. Optical inhibition of PV neurons also improved pyramidal cell reliability. Thus, these experiments reveal that the SST→PV circuit plays an important role in modulating reliability by dynamically routing inhibition between the soma and the dendrites. Together, the work presented in this chapter identifies a novel role of the SST→PV circuit in modulating the fidelity of neural coding in mouse V1.

Highlights

- Novel double transgenic mice allowed us to probe the SST→PV connection in awake, head-fixed mice.
- SST neurons respond reliably, and with delayed responses to natural scenes.
- Activation of SST neurons improves reliability, whereas activation of PV neurons decreased reliability.
- SST neurons suppress PV neurons and optical suppression of PV neurons also improved pyramidal cell reliability
- The SST→PV disinhibitory circuit is critical for ensuring reliable responses.
- VIP neuron activation, which suppresses SST neurons, can allow top-down signals to modulate coding fidelity in V1.

[1] This chapter is currently in preparation for submission

4.1 Introduction

Neurons transmit information through a series of precisely timed spikes. Surprisingly, *in vivo* recordings have shown that cortical neurons are highly unreliable, as repeated presentations of identical stimuli produce spike trains that vary greatly in both the number and timing of spikes between trials (Softky and Koch, 1993; Tolhurst et al., 1983). These observations have led to the idea that the cortex uses a rate code to transmit information (Shadlen and Newsome, 1998). At the same time, slice recordings have shown that neurons can produce highly reliable and low jitter spike trains when injected with stochastic currents (Mainen and Sejnowski, 1995; Stevens and Zador, 1998). Thus, despite intrinsic variability, neurons can transform noisy inputs into highly reliable outputs. More importantly, these studies also imply that the nature of inputs to cortical neurons can strongly modulate their variability (Renart and Machens, 2014).

In agreement with this notion, several recent studies have found that the reliability of visual cortical neurons is highly stimulus dependent (see **Chapter 2**). In particular, responses to natural scenes are sparser and have less trial-to-trial variability than responses to simpler stimuli (Baudot et al., 2013; Haider et al., 2010; Rikhye and Sur, 2015). This means that in the presence of statistical complex visual stimuli the number of spikes not explained (noise) by the stimulus is decreased. This low noise coding is beneficial as it enables neurons to transmit information rapidly using spike-timing based codes (Butts et al., 2007; Hires et al., 2015). In addition, minimizing variability in the earliest stage of cortical processing allows increased coding fidelity in higher stages of the cortical processing hierarchy (Moldakarimov et al., 2015). The impact that input statistics have on response variability means that the cortex can dynamically regulate variability to gate different stimuli (Sprague et al., 2015). However, the neural mechanisms that allow the cortex to achieve this remain unknown. In this chapter, I will focus on elucidating role that inhibitory interneurons play in ensuring reliable coding in mouse V1.

At the heart of all computations in the cortex are microcircuits of highly interconnected excitatory and inhibitory neurons. Thus, it is likely that inhibitory

interneurons play a critical role in sculpting the reliability of pyramidal neurons. Early work in the rat auditory cortex has shown that IPSPs are both balanced and temporally offset to EPSPs during epochs of reliable firing (Wehr and Zador, 2003). Following up this work, Haider and colleagues found that increased IPSP amplitudes correlated with increased reliability for full-field natural movie stimuli in cat visual cortex (Haider et al., 2010). These studies point to a simple mechanism where reliable spikes occur because large and temporally precise IPSPs cancel out noisy EPSPs. Unfortunately, it is still unclear which interneuron subtypes are responsible for modulating reliability.

Although interneurons comprise only 20% of the neurons in the cortex, they are highly diverse in morphology, biophysical properties and connectivity patterns (Kepecs and Fishell, 2014; Markram et al., 2004). The two major classes of cortical interneurons are Parvalbumin-expressing (PV) neurons and Somatostatin-expressing (SST) neurons, which make up 36% and 30% of cortical interneurons respectively (Pfeffer et al., 2013). SST neurons synapse on to the distal dendritic regions, whereas fast-spiking PV neurons target the soma and proximal dendrites. As a result of this domain specific inhibition, PV and SST neurons are likely to play specialized roles in controlling how pyramidal neurons transform synaptic inputs into spike trains. In particular, by inhibiting the soma, PV neurons provide strong divisive inhibition, which can alter both the threshold required to generate spikes and the temporal window over which integration can occur (Atallah et al., 2012; Wilson et al., 2012). On the other hand, by inhibiting branches of dendrites, SST neurons provide shunting or subtractive inhibition (Vu and Krasne, 1992)(Chance and Abbott, 2000) and can control the integration of synaptic inputs within the dendritic arbor (Gidon and Segev, 2012; Lovett-Barron et al., 2012). Additionally, SST neurons also inhibit PV neurons, forming a dis-inhibitory circuit (Pfeffer et al., 2013). While the role of this circuit is still ambiguous, it is possible that this circuit helps to route inhibition from soma to dendrites, which is known to play an important role in adjusting the temporal fidelity of spikes (Pouille and Scanziani, 2001; Pouille et al., 2013). Despite these proposed roles, it is still unclear how these neurons interact to sculpt the temporal reliability of spiking, especially in the context of natural scene processing.

Recent studies shed some light on this question. Using the light-sensitive proton pump archeorhodopsin (Arch) to inhibit PV and SST neurons in awake, passively-viewing mice, Zhu and colleagues found that suppressing PV neurons caused a modest decrease in the reliability of pyramidal cell responses to natural movies. In contrast, suppressing SST neurons did not have any effect on reliability or the signal-to-noise ratio. Based on these results, they concluded that feed-forward inhibition from PV neurons is the main effector of low-noise coding (Zhu et al., 2015). In contrast, another study showed that silencing SST neurons, with Arch, reduced spiking fidelity in mouse piriform cortex (Sturgill and Isaacson, 2015). These authors reasoned that, through subtractive inhibition, SST neurons help to produce sparser spike trains that are more reliable between trials.

These two studies have unfortunately led to conflicting results, and it is still unclear if PV or SST neurons contribute more to low noise coding in the cortex. In this chapter, I use novel imaging methods and optical manipulation to elucidate the circuit mechanisms of reliable coding. I show that SST neurons are crucial for modulating the reliability of coding by providing strong dendritic inhibition, and by modulating somatic inhibition via PV neurons. Together, this work shows that the cortex can use the SST→PV circuit to modulate the precision of pyramidal neuron spiking.

4.2 Hypotheses and Proposed Mechanism

To gain an intuition into how inhibition sculpts response reliability, we constructed a simple conductance-based leaky Integrate-and-Fire (LIF) neuron that received 800 excitatory and 200 inhibitory inputs (**Figure 4-1A**). Each input neuron was assumed to have Poisson spike statistics with variable rate (drawn from a Gaussian distribution). We assumed that both excitatory and inhibitory post synaptic potentials (EPSPs and IPSPs) behaved as alpha functions ($g_{e,i} = \alpha^2 t e^{-\alpha t}$) with identical time constants ($1/\alpha =$ 5ms) (Wehr and Zador, 2003). We changed the reliability of both the excitatory and inhibitory inputs by varying the rates of the input neurons. For example, for a perfectly reliable excitatory input, we set all excitatory-neurons to fire at the same Poisson rate. Indirectly, this allowed us to alter the variance of the EPSP and IPSPs on every trial. We computed reliability of our LIF neuron using the following formula:

$$Reliability = \frac{2}{T^2 - T}\sum_{i=1}^{T}\sum_{j=i+1}^{T}\rho\big(f(i), f(j)\big)$$

Where T is the number of trials and $f(i)$ is the response of the model neuron on trial i. Based on this formula response reliability is the average Pearson's correlation coefficient between every pairwise combination of trials. Thus, this metric quantifies the degree of response similarity between trials. We used the same formula to calculate reliability in the next section. We used 40 trials, and repeated each simulation 1000 times, each with a different random seed.

As expected, the LIF neuron produced highly reliable spikes when given reliable EPSPs and IPSP inputs (timing jitter = 0.5ms). Similarly, spikes were much more unreliable when both inputs were unreliable (timing jitter = 50ms, **Figure 4-1B**). We found that reliability decreased as the jitter in the timing of the IPSP input increased, even when EPSPs were highly reliable (**Figure 4-1C**). Surprisingly, increasing the reliability of IPSP barrages improved reliability in the model that had unreliable EPSPs (**Figure 4-1D**). These simulations, predict that reliable inhibition would serve to reduce membrane potential variance caused by noisy and randomly timed EPSP barrages.

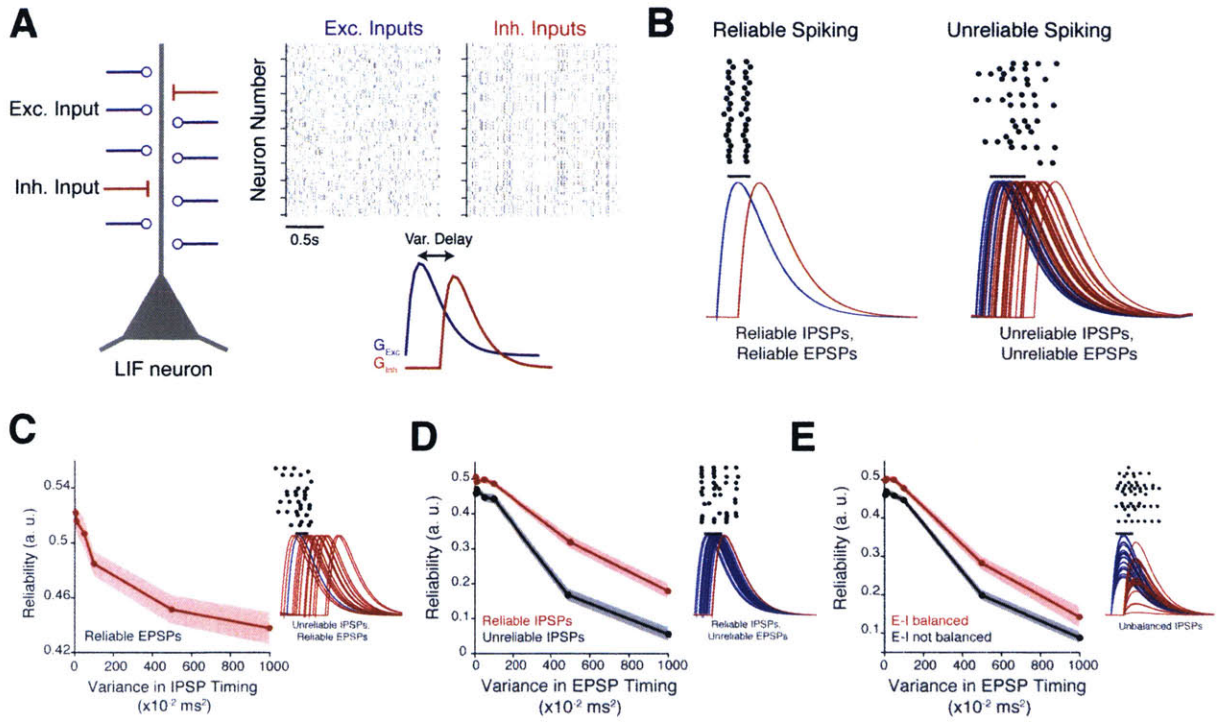

Figure 4-1. Model predicts that reliable and balanced inhibition is necessary for reliable coding. (A) Model schematic. Briefly, we simulated an integrate-and-fire neuron that received 800 excitatory (E) and 200 inhibitory (I) inputs, where were each modeled as Poisson spike generators. The E- and I- inputs were filtered through an alpha-function synapse to produce an EPSP and IPSP respectively. **(B)** Conditions required to produce reliable and unreliable spiking. **(C)** Increasing IPSP jitter with reliable EPSPs increases response unreliability. **(D)** Reliable inhibition (red) helps to mitigate unreliability caused by unreliable EPSP inputs. **(E)** Balanced inhibition also helps to mitigate unreliability caused by unreliable EPSP inputs. Data from 1000 simulations. Shaded areas, SEM.

Numerous studies have shown that excitation and inhibition (E-I) are balanced and highly correlated in cortical neurons (Isaacson and Scanziani, 2011; Okun and Lampl, 2008) [but see (Haider et al., 2013)]. Thus, we next examined how this balance influenced spiking reliability in our model neuron. We balanced excitation and inhibition by equalizing the integrals of both EPSPs and IPSPs on every trial. Conversely, randomly jittering the amplitude of the IPSP relative to the EPSP disrupted this balance. Our simulations revealed that reliability was higher in neurons with balanced inhibition than those with weak, and unbalanced inhibition (**Figure 4-1E**). This result recapitulated the findings of both Wehr and Zador (2003) and Haider et al. (2010).

Taken together, our modeling results predict that reliable spiking occurs as a consequence of three factors: (1) balanced, (2) reliable, and (3) temporally-offset

inhibition. In particular, our model shows that reliable inhibition can be used to quench variable EPSP barrages. We propose that temporally offset inhibition (relative to excitation) functions to create a small time window over which a neuron can integrate its inputs to create spikes (Haider and McCormick, 2009). Fixing the location of this window on every trial would lead to reliable spikes. In contrast, increasing the duration of this window would in turn allow more noisy EPSPs to be integrated, resulting in noisy spikes.

One caveat of this model is that we do not distinguish between IPSPs generated by PV or SST interneurons. Based on the known physiology and connectivity of these interneurons we propose the following mechanism. SST neurons, which primarily inhibit the dendrites of pyramidal neurons, would play a significant role in sculpting the dendritic integration of EPSPs. In contrast, PV neurons, which inhibit the soma, would alter the threshold for spiking by changing the gain of the summed membrane potential. Due to this, PV neurons are believed to provide a majority of the IPSP (Neske et al., 2015). Thus, we propose that reliable inhibition from SST neurons would help the neuron to filter out noisy EPSPs. At the same time, PV neurons would alter the threshold, so as to allow only reliably summed EPSPs to produce spikes. By inhibiting PV neurons, SST neurons can control both the amplitude and timing of the IPSPs. In doing so, we propose that SST neurons are crucial for setting the temporal window for reliable spiking via the SST→PV circuit.

Based on this proposed mechanism, we expect SST neurons to be reliable and to strongly modulate the activity of PV neurons. In the following section, we describe a set of experiments aimed at testing this proposed mechanism.

4.3 Results

SST neurons respond reliably and selectively to natural movies.

First, we sought to directly measure the dynamics of cortical inhibition in awake, head-fixed mice during natural scene stimulation. To do so, we performed cell-specific calcium imaging by injecting a floxed-variant of the genetically encoded calcium indicator (GECI) GCaMP6f into either PV-cre or SST-cre mice (**Figure 4-2A**). This method allowed us to restrict the expression of GCaMP6f to only these interneurons, and minimized contamination from neighboring pyramidal cells. Head-fixed and passively viewing mice were shown a pseudo-randomized sequence of five natural movies (4s each) that were interleaved with a 4s gray screen (mean luminance = 128). Each movie frame was adjusted to have a mean luminance of 128 and a mean contrast of 32, and each movie was repeated 80 times (see Appendix 4B for full experimental details). We restricted our analysis to neurons in L2/3 of primary visual cortex.

Both PV and SST neurons responded to natural movies with robust calcium transients (see examples in **Figure 4-2B**). We inferred response rates from these calcium transients using a temporal deconvolution method that is optimized for GCaMP signals (Theis et al., 2016). We found both interneuron subtypes were strongly activated by natural movies with an approximate 2-fold increase in response rate over spontaneous activity (**Figure 4-2C, D**). In particular, we noticed that the majority of PV neurons were maximally active within 0.89-1.02s of the movie onset. In contrast, SST neurons were activated much later during the movie (1.22-1.83s, **Figure 4-2E**). These results are consistent with previous studies that demonstrate that SST neurons are delayed relative to PV neurons when stimulated with drifting gratings (Ma et al., 2010). To gain a better insight into the activation dynamics of pyramidal neurons, we repeated these imaging experiments in PV-tdTomato or SST-tdTomato mice that expressed GCaMP6f in all neurons (see Appendix 4B). We identified excitatory (EXC) neurons as those cells, which did not have the tdTomato marker. We reasoned that since 80% of neurons in the cortex are excitatory (Markram et al., 2004), the large majority of tdTomato-negative

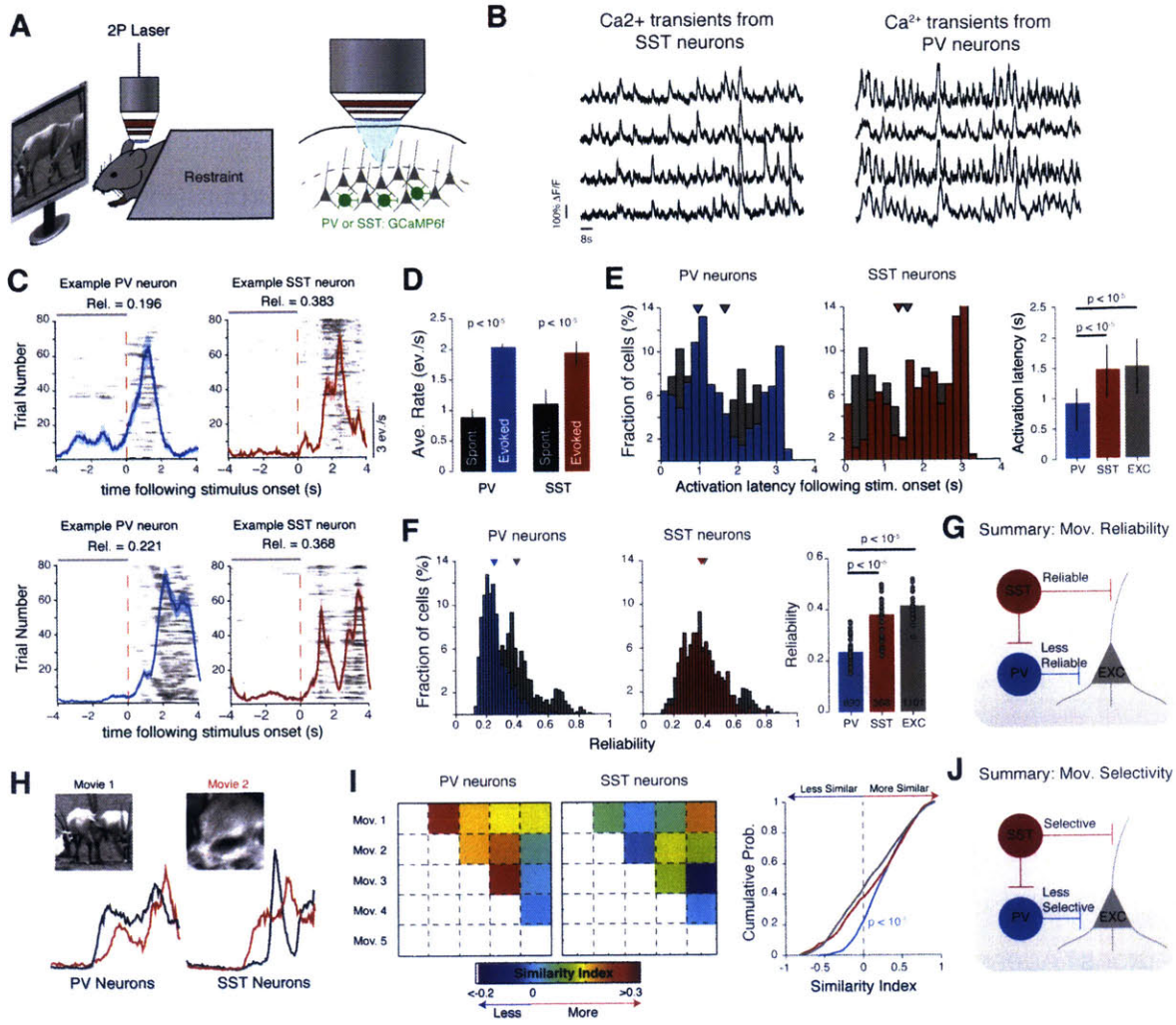

Figure 4-2. Dynamics of cortical inhibitory neurons in response to natural movies. (A) Cartoon illustrating experimental setup. **(B)** Example calcium transients from four PV and SST neurons in response to a natural movie. **(C)** Example raster plot depicting the trial-to-trial firing activity of two PV and two SST neurons to different natural movies. The neurons in each row were stimulated with the same natural movie. Solid line represents trial-averaged activity. **(D)** Average firing rate during gray screen (black bars) and stimulus-evoked epochs (colored bars) for both natural movies. P-value computed using Bonferroni-corrected rank-sum test. **(E)** PV neurons are activated earlier than SST or EXC neurons (gray bars). Left, Histogram showing fraction of neurons activated at different epochs (200ms time bin) during a movie. Triangles denote mean activation latency. Right, Bar plot quantifying median activation time (+/- 95% CI). **(F)** PV neurons are less reliable than SST or EXC neurons. Triangles again denote mean reliability. **(G)** Reliability summary. **(H)** Example responses of PV and SST neurons to two different natural movies. This example illustrates that PV neurons have similar responses for both movies, whereas SST neurons have different responses. **(I)** Matrix showing mean similarity index between all pairs of movies. Right, CDF of similarity index. (**G and J**) Schematic model summarizing findings in this figure. Data collected from: PV = 690 neurons (8 mice), SST = 368 neurons (8 mice) and EXC = 1,101 neurons (10 mice). P-values in **E**, **F** and **I** computed using Kruskal-Wallis ANOVA followed by post-hoc Bonferroni-corrected rank-sum test. Data in bar plots shown as median +/- 95% CI.

neurons would be pyramidal neurons. Expectedly, we also found that distinct EXC neurons responded at different epochs a particular movie, spanning the duration of the movie. However, most EXC neurons also responded within 1.27-1.83s of the onset of the movie, and had very similar activation time course as SST neurons (**Figure 4-1E**). Together, these results suggest that cortical inhibition is dynamic, with different interneurons recruited at different epochs during a natural movie.

Next, we quantified the reliability of both interneuron subtypes to natural movies. A reliably responding neuron would respond with highly stereotypical responses on each stimulus repetition. Thus, reliability provides a robust measure of both the trial-to-trial variability and the temporal precision of a neuronal response to a particular stimulus. Surprisingly, while PV neurons responded strongly to natural movies, their responses were quite variable between trials. In contrast, SST neurons responded precisely and stereotypically on every trial (see example in **Figure 4-2C**). As a result, SST neurons were more reliably recruited than PV neurons (**Figure 4-2F**). Additionally, we found no significant difference in reliability between SST and EXC neurons. One parsimonious explanation for this result is the possibility that these SST neurons inherit their visual response properties from other recurrently connected L2/3 pyramidal neurons (Adesnik et al., 2012)

We also noticed that PV neurons responded to different movies with very similar responses, whereas SST neurons showed a more diverse range of responses (**Figure 4-2H**). We quantified response similarity between movies by computing the Pearson's correlation coefficient between neuron responses for all pairs of movies. On average, both EXC and SST neurons had much lower response similarity than PV neurons (**Figure 4-2I**). Taken together, these results demonstrate that PV neurons provide strong, but unselective and unreliable inhibition on their targets (summarized in **Figure 4-2G, J**). In contrast, SST neurons provide reliable and stimulus-specific inhibition. Thus, SST neurons could potentially be the source of reliable inhibition that is necessary for reliable coding as predicted by our model.

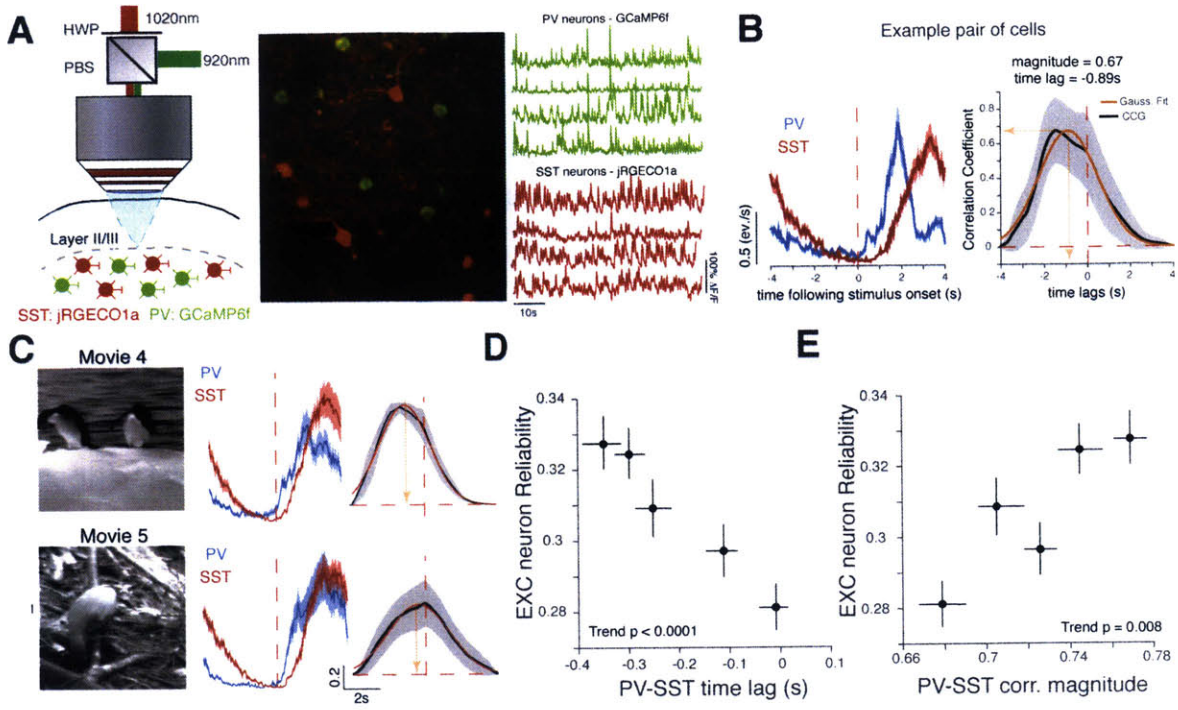

Figure 4-3. PV-SST networks demonstrate strong temporal correlations in reliably processed movies. (A) Left, Experimental setup. Briefly, a 1020nm laser and a 920nm laser were combined using a half wave plate (HWP) and a polarizing beam splitter (PBS) to optimally activate jRGECO1a expressing SST neurons and GCaMP6f expressing PV neurons. Right, Example FOV showing co-labeling of PV and SST neurons. Example calcium transients from neurons in this FOV are also shown. **(B)** Left, Trial-averaged responses from a pair of simultaneously recorded PV and SST neurons. Right, Cross-correlogram (CCG) of this pair of neurons. Orange line shows Gaussian fit. **(C)** Correlation between PV and SST neurons vary with different movies. Data averaged from one population (14 PV neurons and 8 SST neuron). **(D-E)** More reliable movies have longer time lags **(D)** and stronger correlation **(E)** between PV and SST neurons. Data in D-E from 2292 pairs (5 mice). P-values computed using F-test and quantify significance of the trend. Shaded error-bars indicate SEM. Data in D-E presented as median +/- 95% CI.

SST and PV neurons have a precise temporal relationship in response to natural movies

Our data show that PV and SST neurons have distinct dynamics in response to natural movies. Since, SST neurons also innervate PV neurons forming a di-synaptic disinhibitory loop (Pfeffer et al., 2013), we next sought to better understand the trial-to-trial co-activation dynamics of PV and SST neurons within the same neural population. To achieve this goal, we had to develop a method that allowed us to label both interneuron subtypes within the same field-of-view (FOV). To do so, we used an intersectional genetic strategy to engineer a new double transgenic mouse strain

(referred to as SXP mice) that expressed the recombinase enzymes Cre and Flpo exclusively in SST and PV neurons respectively (see Appendix 4B). These mice allowed us to express the red-shifted GECI (jRGECO1a, (Dana et al., 2016)) exclusively in SST neurons and GCaMP6f in PV neurons (**Figure 4-3A**). Using novel optical technology, we were able simultaneously scan the same FOV with two 2-photon laser beams - one tuned to 1020nm to excite jRGECO1a and the other tuned to 920nm to excite GCaMP6f. These wavelengths optimally activate each fluorophore with very little spectral overlap. As a consequence, our new two-color imaging approach allowed us to record from SST and PV neurons with little signal contamination between the two channels.

Our dual-color imaging revealed that SST neurons lagged behind PV neurons within the same population (see example in **Figure 4-3B**). This is consistent with our data in **Figure 4-2E**. We quantified the temporal relationship between PV and SST neurons by computing a cross-correlogram (CCG) between all PV-SST pairs on each trial. We then estimated a lag time as the mean of the Gaussian fit to the CCG. Across all recorded pairs, we observed an average time lag of -0.218s (CI: -0.652 to -0.08s) between PV and SST neurons. Interestingly, we found that the temporal relationship changed with different movies (**Figure 4-3C**). This is primarily due to different SST dynamics, as PV neurons responded with stereotypical responses to most movies. Also, this negative temporal relationship is indicative of an inhibitory connection between PV and SST neurons (Cottam et al., 2013; Pfeffer et al., 2013).

Similarly, we also noticed that EXC neurons responded more reliably to some movies than other movies, even though all movies had the same mean luminance and contrast. This is likely due to the fact that each movie had slightly different spatiotemporal content, which is known to influence coding reliability ((Rikhye and Sur, 2015), see **Section 3-2**). Plotting the PV-SST time lag against EXC reliability for different movies revealed that movies, which were more reliably processed, had more negative time lags than stimuli that were less reliably processed (**Figure 4-3D**). Similarly, movies that were more reliably processed by EXC neurons also evoked significantly stronger correlations between PV and SST neurons (**Figure 4-3E**). One

caveat, however, is that the EXC neuron data is not from the same FOV. Consequentially, the ordinate in **Figure 4-2E** is the maximum likelihood estimate (median +/- 95% CI) of the pyramidal cell reliability of that particular movie.

Taken together, these results suggest that reliably processed movies strongly co-activate both PV and SST neurons, with SST activity appearing later than PV activity. Specifically our data shows that inhibition dynamically routes between the soma and dendrites during the time course of a natural movie. At stimulus onset, PV neurons are most active, leading to stronger somatic inhibition. Shortly after this, SST neurons become active, which causes inhibition to be routed away from the soma and into the dendrites. Thus, this temporal interplay between PV and SST neurons is likely to be a crucial factor for reliable spiking (**Figure 4-S1**).

Optogenetic activation of SST and PV neurons bi-directionally alters pyramidal cell reliability

If this temporal relationship between PV and SST neurons is necessary for reliable coding, then disrupting it by decorrelating PV-SST interactions should change pyramidal cell reliability. To test this hypothesis, we conditionally expressed the light activated cation channel channelrhodopsin-2 (ChR2) in either PV or SST neurons and performed calcium imaging from pyramidal neurons expressing GCaMP6f (**Figure 4-4A**, see Appendix 4B). We used transgenic mice (PV-cre::Ai32 and SST-cre::Ai32) to minimize ChR2 expression variability (Madisen et al., 2012).

Given the heterogeneity of neural responses within a population, we had no way of knowing *a priori* when a pyramidal neuron would be reliably active during a particular movie. In fact, distinct EXC neuron subsets responded reliably throughout the duration of the movie (**Figure 4-4B**). To circumvent this problem, we developed a new activation strategy in which we briefly pulsed (4x20ms, 10ms inter-pulse interval) a 473nm (blue) laser at 22 different time points during a given movie (**Figure 4-4C**). To prevent

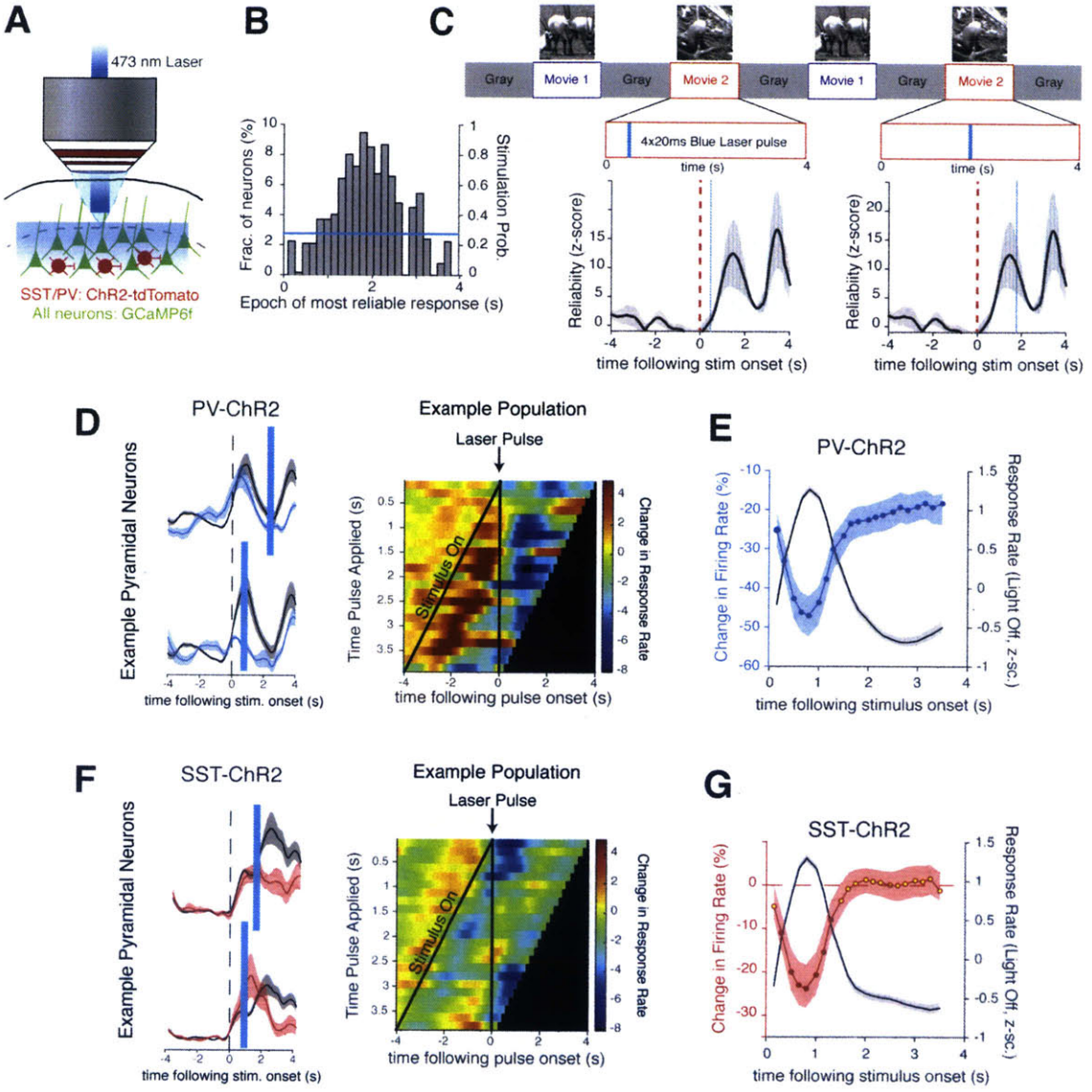

Figure 4-4. Random stimulation protocol to optically activate interneurons. (A) Cartoon of experimental set-up. **(B)** Distinct pyramidal neurons are active at different epochs during a natural movie. Histogram shows that reliable responses space the entire duration of a movie. **(C)** Cartoon describing random stimulation protocol. A brief laser pulse was applied at 22 different time points during a movie with pulses occurring at random. Bottom plot shows the timing of each pulse in relation to the reliability of a pyramidal neuron. **(D)** Pyramidal neurons are strongly suppressed following PV neuron activation. Heat map quantifies change in firing rate and is aligned to pulse onset. Each row is the average change a set of neurons within the same FOV. **(E)** Change in firing rate induced by PV neuron activation. For reference the original firing rate is plotted in black. To facilitate comparisons between movies and mice, the plot is aligned to have a maximum firing rate at 1s. **(F-G)** Pyramidal neurons are also suppressed by SST activation. Data from: PV-ChR2 = 7 mice (540 neurons), SST-ChR2 = 8 mice (622 neurons). Shaded areas in E and G denote 95%CI of median.

adaptation, we interleaved the *pulsed* movie with a *non-pulsed* movie and randomized the pulse pattern, such that each stimulation event occurred 8s or more apart. We repeated each pulse pattern 10 times, which allowed us to compute reliability. To quantify the change in reliability caused by laser stimulation, we randomly interleaved 20 *Light-Off* trials. Imaging from PV/SST neurons expressing both GCaMP6f and ChR2 confirmed that these brief pulses of blue light could robustly and reliably excite interneurons expressing ChR2 with at any time during a natural movie (**Figure 4-S2**). Thus, our stimulation technique allowed us to reliably increase inhibition independent of the ongoing visual stimulus. The main advantage of our technique is that it allows us to determine *post hoc* the effect of increasing inhibition when pyramidal neurons are either the most or the least reliable.

Expectedly, activating PV neurons resulted in a significant decrease in pyramidal neuron response rate (**Figure 4-4D**). To prevent stimulation laser artifacts, we restricted our analysis to a 600ms window following laser offset. Surprisingly, even though we used only brief pulses of light, we found that the inhibitory effects of PV activation lasted for up to 800-1000ms following laser offset (**Figure 4-S3**). We noticed that PV-mediated suppression was the strongest when pyramidal neurons were the most active (**Figure 4-4E**). This suggests that PV neurons do play an important role in preserving the balance between excitation and inhibition, as previously noted (Xue et al., 2014). Importantly, with the exception of the epoch of highest activity, activation of PV neurons resulted in an approximate 20% reduction in pyramidal neuron response. Similarly, activating SST neurons also caused a transient reduction in pyramidal neuron activity (**Figure 4-4F**). However the effect that SST activation had on pyramidal neurons was much weaker than PV activation (**Figure 4-4G**), possibly due to the disinhibitory effects of the SST→PV circuit. Together, these findings confirm the efficacy of our new stimulation paradigm in transiently altering the inhibitory tone of pyramidal neurons.

Next, we investigated the effect that PV/SST activation had on pyramidal cell reliability. PV neuron activation significantly decreased pyramidal cell reliability. In particular, the effect of PV activation was the strongest when the laser was applied close to stimulus onset ($p < 10^{-4}$, Permutation Test, **Figure 4-5A**). At stimulus onset,

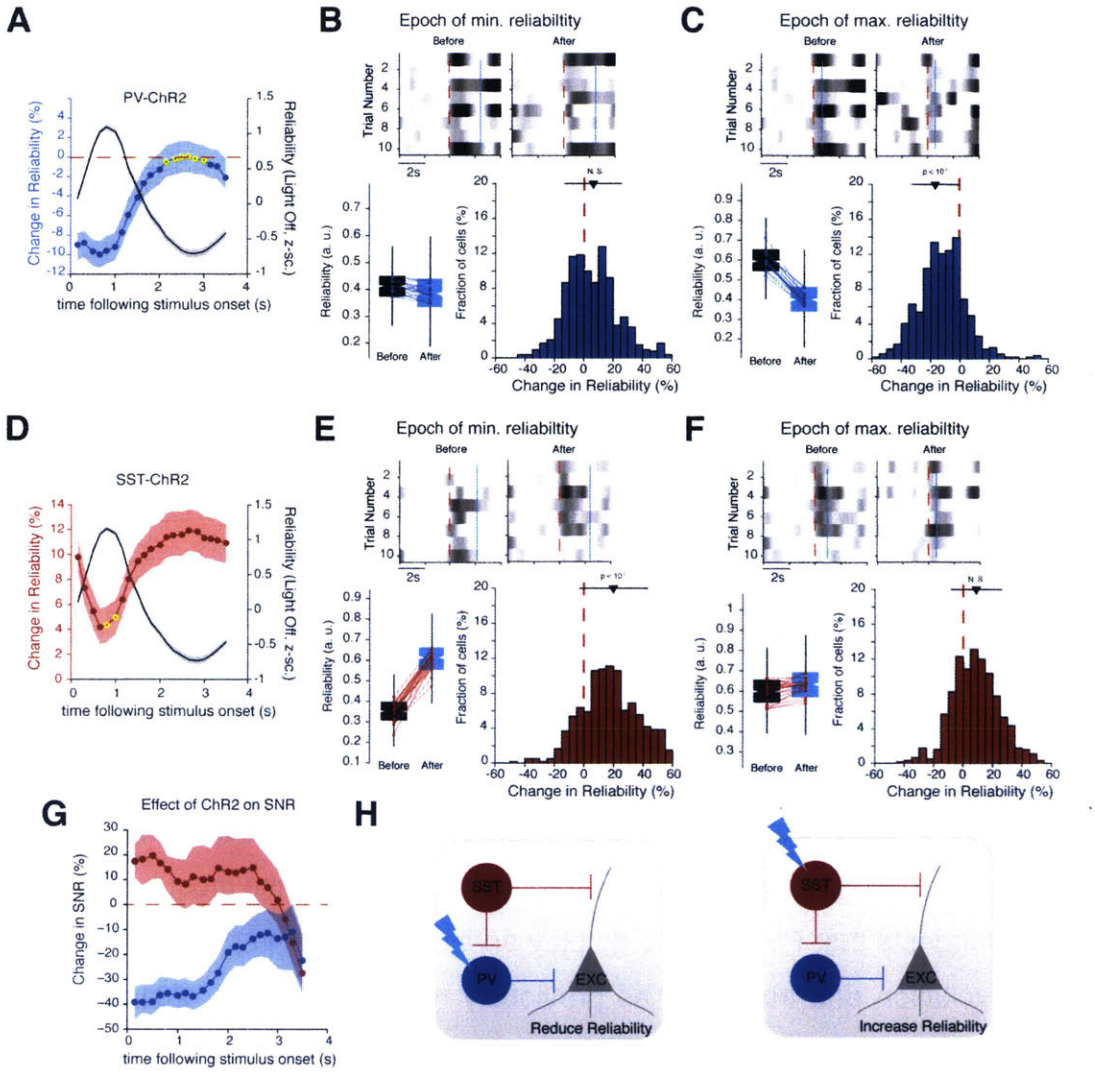

Figure 4-5. Optically activating SST and PV neurons bi-directionally alters response reliability. **(A)** Plot showing change in reliability for each laser activation time point. For comparison, reliability on the light-off trials is plotted in the background. To facilitate comparison between animals, plots are aligned such that the epoch of most reliable spiking occurs at 1s post stimulus onset. **(B)** Activating PV neurons when pyramidal neurons are least reliable does not significantly alter reliability. Raster plots from taken from a representative neuron. **(C)** Activating PV neurons when pyramidal neurons are most reliable significantly decreases reliability. (**D-F**) Same as (**A-C**) but shows that activating SST neurons improves reliability (E) when neurons are least reliable. Data taken from: PV-ChR2 = 7 mice (540 neurons) and SST-ChR2 = 8 mice (622 neurons). **(G)** Effect of ChR2 activation on signal-to-noise ratio. **(H)** Summary. P-values in A and D computed using Permutation test. Shaded area denotes 95% CI of the median. Yellow circles represent non-significant change. P-values in B, C, E, F computed using Bonferroni-corrected Wilcoxon rank-sum test.

feed-forward drive is the highest, due to the sudden transition from gray screen to movie. Thus, perturbing PV inhibition when feed-forward drive is the highest has the

most detrimental effect on the ability of the visual cortex to encode the incoming stimulus. Interestingly, activating PV neurons when pyramidal neurons are unreliable (and weakly responsive) only marginally reduced reliability (**Figure 4-5B**). In contrast, activating PV neurons when pyramidal neurons were the most reliable strongly decreased reliability (**Figure 4-5C**).

In stark contrast to PV neurons, activating SST neurons close to stimulus onset strongly increased reliability ($p < 10^{-4}$, Permutation Test, **Figure 4-5D**). As we show in **Figure 4-2** and **Figure 4-S1**, SST neurons activity correlates well with pyramidal neuron reliability, with SST neurons most active during epochs of reliable pyramidal firing. Thus, it is likely that weak reliability early in the movie could be due to weak SST activity. In agreement with this argument, we found that activating SST neurons when pyramidal neurons were weakly reliable strongly increased reliability (**Figure 4-5E**). Surprisingly, adding more SST inhibition when neurons are already at their most reliable state failed to further improve reliability (**Figure 4-5F**). This is likely to be due to a ceiling effect on reliability, as many other factors, such as arousal, can also influence reliability (Renart and Machens, 2014).

Were these changes in reliability simply due to a reduction in response rates of pyramidal neurons? We reasoned that if the variability changes were governed by a reduction in firing rate, then we would observe large reductions in reliability for those neurons that have experienced large reductions in response rate. Following SST activation, we found that, neurons that had the largest reduction in response rate displayed the smallest increase in reliability (**Figure 4-S4A**). We observed a similar trend following PV activation, suggesting a relationship between rate and reliability. In particular, in all mice, PV activation always reduced reliability, regardless of change of response rate. On the other hand, in a small subset of imaged populations (8/20), SST activation increased reliability without changing mean rate. In the rest, SST activation increased reliability and decreased mean rate. Interestingly, we observed that for the same level of suppression, SST activation increased pyramidal neuron reliability, whereas PV activation decreased pyramidal neuron reliability (**Figure 4-S4B**). Thus, our analysis demonstrates that, despite increased inhibition, SST activation can increase

the mean response rate of pyramidal neurons by increasing their reliability. Importantly, these results also indicate fundamental differences in the way SST neurons and PV neurons change the response properties of pyramidal neurons.

As further controls, we used mice, which expressed the red fluorescent protein tdTomato in either PV or SST neurons but no ChR2. In these mice, we found no significant change in reliability following laser activation, implying that the observed effect was indeed due to increased inhibition (**Figure 4-S5**). Additionally, we found that inferring firing rates via deconvolution did not influence our reliability calculation, as neurons that had reliable calcium transients also had reliable inferred rates (**Figure 4-S6A**). As a consequence, influence of PV or SST neuron activation on pyramidal cell reliability remained unchanged even when we computed reliability from raw fluorescence transients (**Figure 4-S6B**). Instead, deconvolution allowed us to extract a much wider range of reliability values. This is because deconvolution helps to remove the long tails in the fluorescence distribution caused by slow decay of the GECI. These long tails result in spuriously higher correlation coefficients, and thus give falsely higher reliability values. For this reason, all subsequent analysis was performed with inferred rates.

In addition to reliability, we also characterized the signal-to-noise ratio (SNR) of pyramidal neurons following laser activation. We defined signal as the fraction of variance in the response that can be explained by the stimulus, and noise as the fraction of the variance in the stimulus that cannot be explained by the stimulus. Following this definition, noise provides a measure of the non-stimulus driven portion of the response (see **Chapter 2**). A neuron with high noise and low signal (i.e. low SNR) would transmit information about the stimulus with much less fidelity than a high SNR neuron. Activating SST neurons early in the movie improved SNR ($p < 10^{-4}$, Permutation Test, **Figure 4-5G**). This is because activating SST neurons primarily reduces noise to improve reliability. In contrast, activating PV neurons early during the movie significantly reduced SNR, which was primarily due to a reduction in signal.

At the population level, we found that activating SST neurons also significantly reduced noise correlations between neurons without altering signal correlations (**Figure**

4-S7). Taken together, our optical manipulations demonstrate that SST and PV neurons have opposite effects on pyramidal cell reliability (**Figure 4-5H**). SST neurons are important regulators of trial-to-trial variability. In particular, adding reliable SST inhibition is sufficient to increase reliability and reduce noise between neurons. PV neurons, on the other hand, function mainly to control the level of the signal drive to a neuron, consistent with their role in preserving E-I balance by sacrificing temporal precision.

How does this change in pyramidal cell reliability influence coding efficiency, in particular the ability to discriminate between movies? To answer this question, we used a simple linear classifier to predict from single-trial responses which movie (pulsed movie vs. non-pulsed movie) had been presented (see **Chapter 3**). We repeated classification on both Laser-ON and Laser-OFF trials and assed discrimination accuracy as the fraction of correct classifications made by the classifier (**Figure 4-S8A**). Interestingly, we observed that activating SST neurons early during the movie improved discriminability, whereas activating PV neurons decreased discriminability (**Figure 4-S8B**). In contrast, activating both interneuron subtypes late in the movie (after 2.5s) did not significantly change discriminability compared to the Laser-OFF condition. This implies that modifying inhibition close to stimulus onset has the biggest impact on stimulus encoding. Additionally, these results also demonstrate that increasing neural reliability, by increasing SST inhibition, is beneficial to stimulus encoding. In **Chapter 6**, we explore how this improvement in encoding relates to improved stimulus perception.

SST neurons strongly inhibit PV neurons

Our data so far shows that movies that are reliably processed have strongly correlated, and temporally delayed, PV and SST activity. In addition, increasing SST reliability via optical activation increases pyramidal cell reliability. This suggests that it is the interaction between PV and SST neurons that is crucial for reliable coding. One likely mechanism is that, by suppressing PV neurons, SST neurons can alter the dynamics of PV-driven somatic inhibition. In doing so, SST neurons can alter the *window of opportunity* for reliable spiking in which only the most reliable inputs are integrated

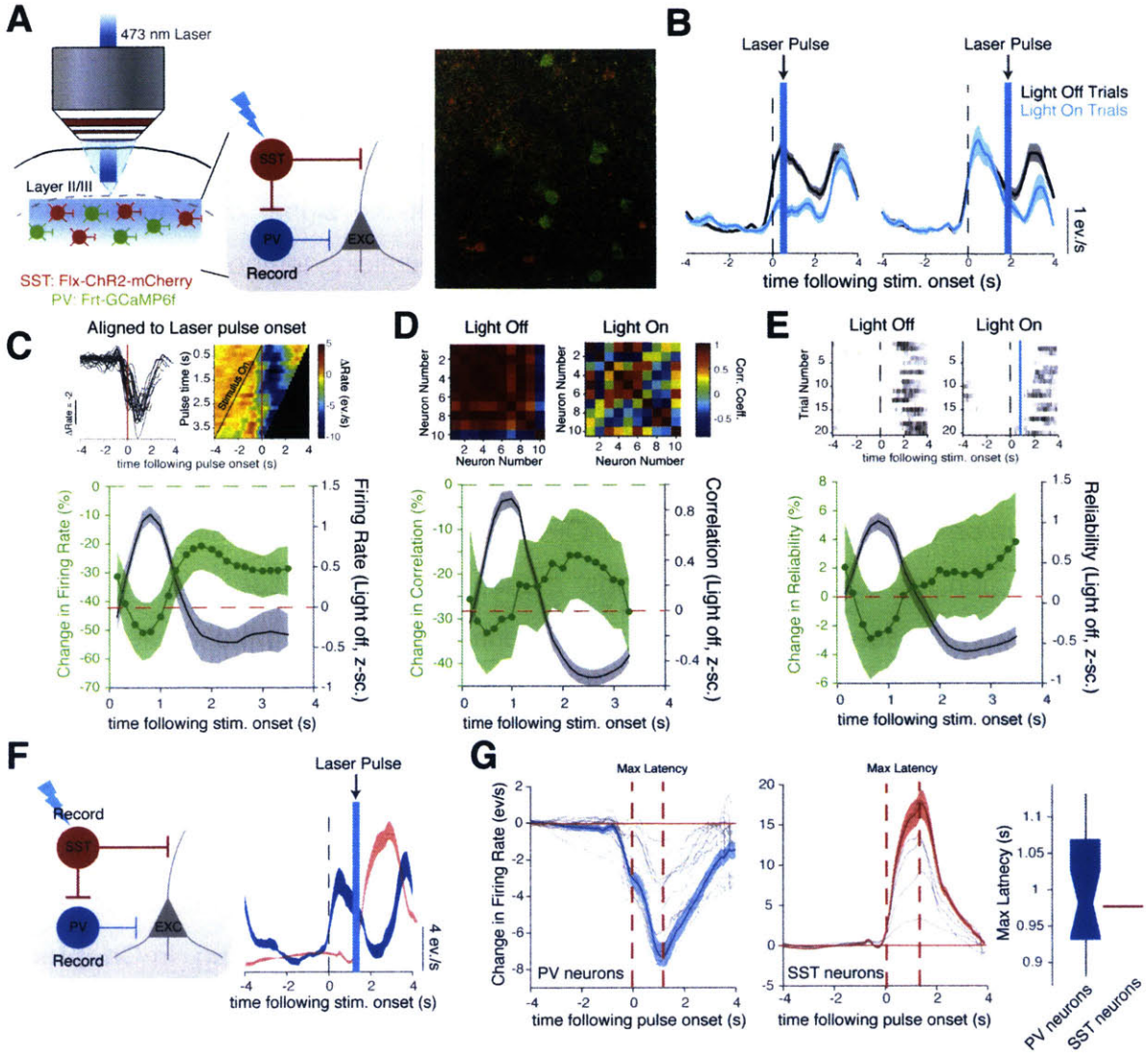

Figure 4-6. SST neurons strongly inhibit PV neurons. (A) Cartoon describing experimental set up. *Right,* Example FOV showing green PV neurons expressing GCaMP6f and red SST neurons expressing ChR2-mCherry. **(B)** Average firing rates of one PV neuron showing the effect of activating SST neurons at two different time points (indicated by arrows). Shaded area is SEM across trials. **(C)** *Top,* Activating SST neurons induces a strong suppression in PV neurons with short latency. Heat map shows suppression dynamics from one population of PV neurons (14 cells). *Bottom,* Quantification of change in firing rate caused by SST activation at different movie time points. **(D)** *Top,* Activating SST neurons strongly decorrelates PV neurons (10 cells). *Bottom,* Quantification of change in PV-PV correlation caused by SST activation. Change at all time points is significant ($p<10^{-4}$, permutation test). **(E)** *Top,* Activating SST neurons does not change the reliability of PV neurons. Bottom, Change in reliability is not significant ($p > 0.05$, permutation test). **(F)** Simultaneously recorded PV and SST neurons showing that laser activation rapidly increases SST activity and suppresses PV activity. **(G)** Quantification of activation/suppression latency following pulse onset. Box-whisker plot shows that the laser very reliably activates SST neurons. Data in C-E from 4 mice (212 PV neurons). Data in F-G from 3 mice (44 PV neurons, 19 SST neurons).

(Haider and McCormick, 2009). However, how SST neurons modulate the activity of PV neurons in awake mice, and under natural scene stimulation remains poorly understood. To test this hypothesis, we used SXP mice to conditionally expressed ChR2 in SST neurons and GCaMP6f in PV neurons (**Figure 4-6A**). We used the same stimulation protocol as described above to quantify the effect that SST activation had on PV neuron activity. We discovered that activating SST neurons at different movie time points rapidly and robustly suppressed PV neurons (see examples in **Figure 4-6B, Figure 4-6C**). SST-induced suppression was the strongest when PV neurons were most active. Importantly, these results demonstrate that, regardless of the on-going visual stimulus, activation of SST neurons was always followed by a strong, but transient, suppression of PV neurons (**Figure 4-5C**). Activating SST neurons also strongly decorrelated PV neurons (**Figure 4-6D**), suggesting that SST activation is sufficient to alter dynamics within the entire PV neuron sub-network. Surprisingly, however we did not find a change in PV neuron reliability following SST activation (**Figure 4-6E**). Thus, even though SST neurons inhibit the dendrites of PV neurons (Hioki et al., 2013), SST neurons are unable to improve the reliability of PV neurons. Together, these results suggest that SST neurons strongly inhibit PV neurons in the mouse visual cortex (Cottam et al., 2013).

Next, we sought to determine if the SST→PV circuit was always active or if it was active only under certain conditions. We reasoned that if SST neurons always inhibited PV neurons, then silencing SST neurons should cause a disinhibitory increase in PV activity. We tested this idea by conditionally expressing Arch in SST neurons (**Figure 4-S9A**). As expected, we found that optically inhibiting SST neurons always led to a strong disinhibitory increase in PV neuron activity, regardless of when the pulse was applied during a natural movie (**Figure 4-S9B-D**). Due to this, SST suppression caused an increase in PV neuron reliability **(Figure 4-S9D**). Thus, these loss-of-function experiments indicate that SST neurons modulate the activity of PV neurons throughout the duration of visual stimulation.

Interestingly, activation of PV neurons failed to suppress SST neurons (**Figure 4-S10**), suggesting that the PV neurons did not directly inhibit SST neurons. Also, we

found that the reduction in pyramidal neuron firing rate induced by PV activation is not sufficient to reduce SST activity during on-going visual stimulation.

To better understand how SST activation changed the dynamics between SST and PV neurons, we injected both floxed-ChR2 and jRGECO1a into SST neurons and FRT-GCaMP6f into PV neurons (see Appendix 4B). This allowed us to image the activities of PV and SST neurons while simultaneously activating SST neurons. Expectedly, SST neurons increased their firing rate, while PV neurons decreased their firing rate shortly after laser onset (see example pair of neurons in **Figure 4-6F**). We quantified the activation latency relative to pulse onset as the time of maximum suppression in the case of PV neurons and maximum activation in the case of SST neurons (**Figure 4-6G**). This analysis revealed that SST neurons were reliably activated and PV neurons were strongly suppressed with the same latency following laser onset. Thus, due to the SST→PV circuit, laser activation of SST neurons induces a time lag between SST and PV neurons (as illustrated in **Figure 4-6F**). This time lag could be reason why pyramidal neurons are more reliable following SST activation (see **Figures 4-3** and **4-6**).

Together, our results suggest that, through the SST→PV circuit, SST neurons are able to change the dynamics of PV-dependent somatic inhibition. Our loss-of-function experiments also suggest that the SST→PV circuit also allows SST neurons to continually modulate the activity of PV neurons. In particular, we propose that activation of SST neurons rapidly routes inhibition to the distal dendrites and lifts inhibition from the soma. Hence, it is possible that reliable coding is the consequence of increase dendritic inhibition and reduced somatic inhibition.

Optical suppressing PV neurons improves pyramidal cell reliability

Is suppressing PV neurons sufficient to increase the reliability of pyramidal neurons? To answer this question, we expressed the light activated proton pump archaeorhodopsin (Arch) selectively in PV neurons (PV-cre::Ai35 mice). We optically suppressed PV neurons with green light (532nm) at different times during a natural movie using the

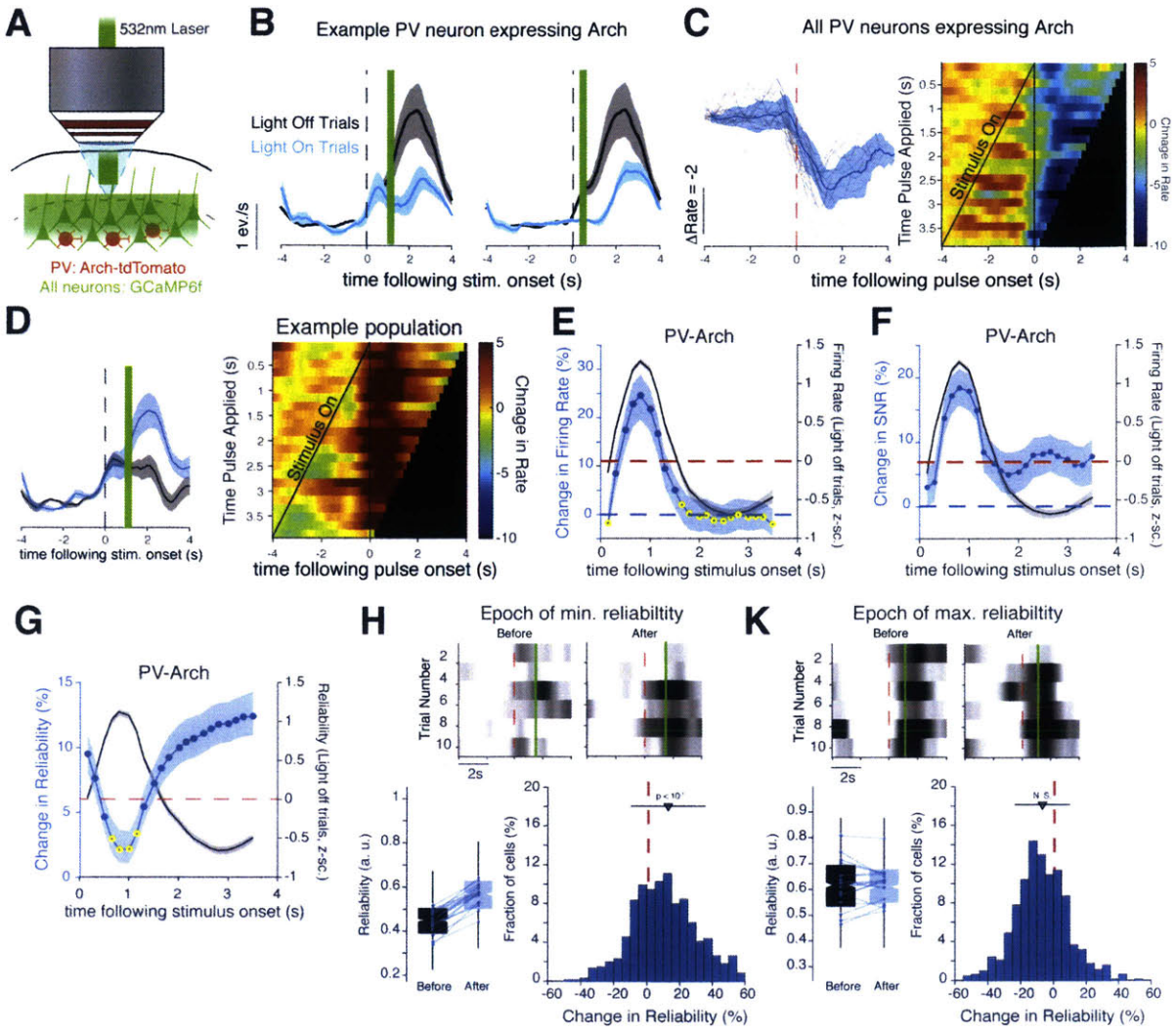

Figure 4-7. Suppression of PV neurons increases reliability, mimicking the effect of SST activation. **(A)** Cartoon describing experimental setup. **(B)** Average firing rate of PV neurons expressing Arch. Shaded areas denote SEM across trials. **(C)** Arch activation transiently suppresses PV neurons. Data from 3 mice (43 neurons). (**D**) Suppressing PV neurons transiently increases the firing rate of pyramidal neurons. Heat map averaged from one population (64 neurons). **(E-G)** Change in firing rate (**E**), change in SNR (**F**) and change in Reliability (**G**) following PV suppression. **(H)** Arch activation at epochs of least reliability significantly increases the reliability of pyramidal neurons. **(K)** Suppressing PV neurons when pyramidal neurons are at their most reliable does not change reliability. Data in D-K collected from 8 mice (634 neurons). Yellow circles in E-G indicate non-significant changes ($p > 0.05$, Permutation test). P-values in H-K computed using Bonferroni-corrected rank-sum test relative to light-off condition. Shaded areas in E-G denote 95% CI of the median.

same random stimulation protocol (**Figure 4-7A**). Expectedly, activating Arch with green light transiently suppressed PV neurons and was able to mimic the effect of SST activation (**Figure 4-7B, C**). Since SST neurons inhibit both pyramidal neurons and PV neurons, the effect of SST activation on reliability could be a combination of both

factors. Thus, by suppressing PV neurons directly, we hoped to dissociate the effect that reduced PV inhibition had on reliable coding from increased dendritic inhibition.

As expected, optically suppressing PV neurons strongly increased the response rate of pyramidal neurons, due to a transient lifting of somatic inhibition (**Figure 4-7D**). In particular, the effect of PV suppression on pyramidal firing rate was the largest only when pyramidal neurons were the most active (**Figure 4-7E**). This suggests that when neurons are weakly active, lifting somatic inhibition is insufficient to cause neurons to respond more. Interestingly, although SST neurons had a strong inhibitory effect on PV neurons, we did not observe a disinhibitory effect (i.e. increase in pyramidal neuron response rate) following SST activation (**Figure 4-4**). Together, these results suggest that the increase in dendritic inhibition induced by SST activation balances somatic disinhibition, resulting in a net suppression.

Additionally, we found that suppressing PV neurons improved the SNR, even when pyramidal neurons were weakly active (**Figure 4-7F**). This means that reducing somatic inhibition when neurons are weakly active decreases in trial-to-trial noise without changing the overall firing rate. Mice expressing tdTomato in PV neurons did not show a laser-induced change in firing rate or reliability (**Figure 4-S5**).

Similar to SST activation, we found that optically inhibiting PV activity increased the reliability of pyramidal neurons (**Figure 4-7G**). In particular, PV suppression most significantly increased pyramidal reliability when these neurons were weakly reliable (**Figure 4-7H**). Removing PV inhibition when pyramidal neurons were most reliable, however, did not alter reliability (**Figure 4-7H**). We observed similar effects following optical excitation of SST neurons; specifically, increasing SST inhibition increased reliability when pyramidal neurons were weakly reliable. Taken together, these results suggest that transiently suppressing PV neurons, either optically via Arch or by the SST→PV circuit, can improve the reliability of pyramidal cells. These results support the hypothesis that the suppressive action of SST on PV neurons, via the SST→PV disinhibitory circuit, is crucial for modulating the reliability of pyramidal neurons.

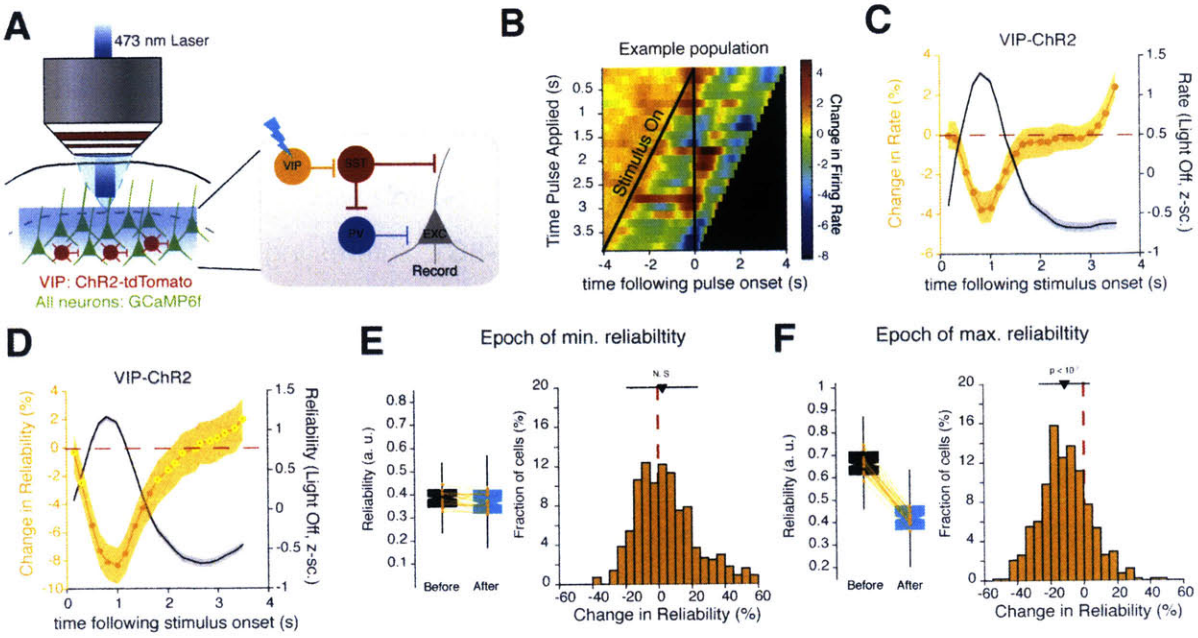

Figure 4-8. Activating VIP neurons decreases pyramidal cell reliability. (A) Cartoon describing experimental setup. **(B)** Heat map quantifying the change in pyramidal cell firing rate following pulse onset for each laser pulse. Data from one population of neurons. **(C)** A modest reduction in pyramidal cell firing following VIP activation. **(D)** VIP activation reduces pyramidal cell reliability only when pyramidal cells are at their most reliable. **(E-F)** Quantification of change in reliability following VIP activaiton at epoch of minimum **(E)** and maximum **(F)** reliability. Data shown in C-D obtained from 6 mice (582 neurons). Yellow circles indicate non-significant changes in reliability computed using permutation test. P-values in E and F computed using Bonferroni-corrected rank-sum test relative to light off trials. Shaded areas denote 95% CI of median.

VIP neurons can use the SST→PV circuit to reduce pyramidal cell reliability

Our data so far suggests that SST neurons can dynamically route inhibition away from the soma and into the dendrites to ensure reliable coding. How can this routing take place under normal conditions, without forcing SST neurons to respond reliably? Another important class of inhibitory interneurons in the visual cortex is the VIP interneurons. These interneurons function primarily to inhibit SST neurons, and are activated by top-down signals from other cortical or subcortical areas (Fu et al., 2014; Letzkus et al., 2011; Pi et al., 2013). Thus, by suppressing SST neurons, VIP neurons could reduce pyramidal cell reliability by routing inhibition back to the soma.

To test this hypothesis, we virally expressed ChR2 in VIP neurons and measured the effect that VIP activation had on pyramidal cell responses (**Figure 4-8A**). Surprisingly, we found that activating VIP neurons only had a modest effect on

pyramidal neuron response rates (**Figure 4-8B**). This could possibly be due to the fact that VIP neurons do not synapse directly onto pyramidal cells, and thus only indirectly influence pyramidal neurons through either PV or SST neurons. In particular, we noticed a small reduction in response rate (average = -4.2%) following VIP activation when pyramidal neurons were most active (**Figure 4-8C**). Interestingly, increasing VIP inhibition during the epoch of maximum reliability strongly reduced pyramidal neuron reliability (**Figure 4-8D, F**). As we have established, SST neurons are most active when pyramidal neurons are at their most reliable (**Figure 4-S1**). In contrast, SST neurons are weakly active when pyramidal neurons are unreliable. Thus, these results demonstrate that increasing VIP inhibition when SST neurons are most active decreases pyramidal reliability. In contrast, VIP neurons are ineffective at altering pyramidal reliability when both SST neurons and pyramidal neurons are weakly active (**Figure 4-8E**). This implies that VIP→SST inhibition is most effective when SST neurons are strongly active. Together, these results further underscore the importance of SST neurons in modulating pyramidal neuron reliability. In particular, VIP neurons could make the neural code less reliable by lifting SST→PV inhibition.

Collectively, our results show that both SST and PV neurons work cooperatively together, via the SST→PV disinhibitory circuit, to modulate the reliability of pyramidal neuron responses to natural movies in visual cortex.

4.4 Discussion

In this study, we investigated the inhibitory mechanisms responsible for reliable coding in mouse V1. Our work has revealed distinct contributions of PV- and SST-expressing interneurons to modulating trial-to-trial variability of pyramidal neurons to repeated presentation of natural movies. Specifically, we demonstrate using a new imaging methodology that the activities of SST and PV neurons are tightly correlated during epochs of reliable pyramidal neuron firing. Optical activation of SST neurons resulted in a strong, but transient, suppression of PV neurons, confirming the existence of a strong SST→PV inhibitory circuit. Furthermore, we found that either optical activating SST neurons or suppressing PV neurons increased pyramidal neuron reliability. Together, our work reveals that the importance of the inhibitory SST→PV circuit in modulating the spiking variability of pyramidal neurons.

The role of SST neurons in reliable coding

Cortical neurons are highly unreliable to sensory stimulation, despite receiving reliable inputs from the earlier visual areas (Kara et al., 2000). Thus, cortical unreliability must therefore be the consequence of variability in synaptic transmission between cortical neurons. Previous studies have demonstrated that cortical neurons fire reliably for some stimuli because these stochastic EPSP barrages are quenched by strong and reliable IPSPs (Haider et al., 2010; Wehr and Zador, 2003). We further extend these studies by demonstrating that SST neurons fire reliably to natural movies and are most active when pyramidal neurons are most active. These results suggest that SST neurons could be the source of these reliable IPSPs. In particular, we propose that SST neurons are responsible for transforming noisy inputs into reliable events.

Given that SST neurons inhibit distal dendrites, they are ideally poised to filter out noisy synaptic inputs. Several studies have shown that shunting inhibition can alter synaptic integration (Gidon and Segev, 2012; Jadi et al., 2012). Interestingly, SST neurons have been shown to provide compartmentalized inhibition onto specific

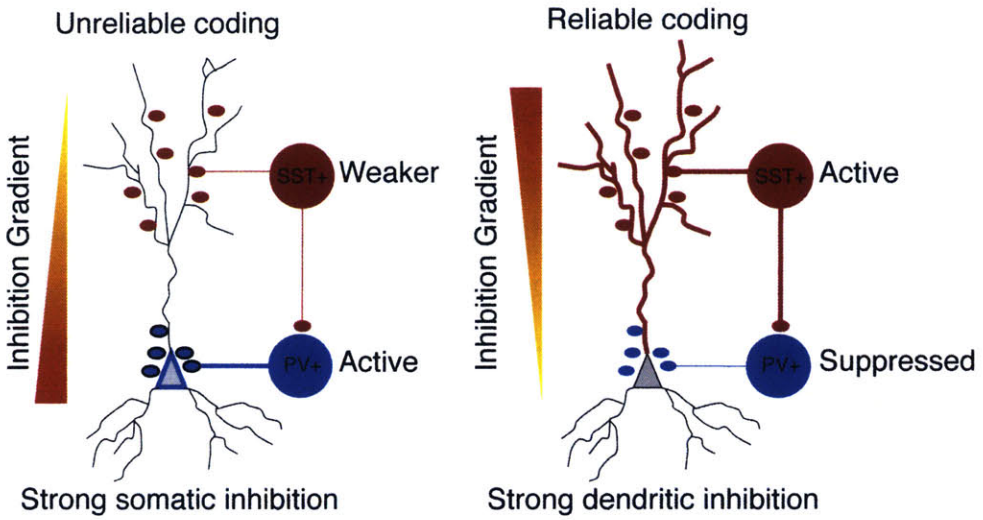

Figure 4-9. The main findings of this chapter. Specifically, by suppressing PV neurons, SST neurons can dynamically route inhibition from the soma into the dendrites. In doing so, SST neurons function to filter out unreliable inputs.

branches of the dendritic arbor (Bloss et al., 2016; Chiu et al., 2013). In doing so, SST neurons can regulate excitatory synaptic integration in a highly branch specific manner (Branco and Häusser, 2011; Lovett-Barron et al., 2012). Thus, we propose that SST neurons alter pyramidal neuron reliability by tuning the threshold for synaptic integration, which would allow only the most reliable and correlated inputs to be integrated to produce spikes (Huang et al., 2016). However, to test this hypothesis, further work is required to determine the strength and distribution of synaptic responses during epochs of reliable spiking.

Reliable SST neuron firing could also be the reason why full-field natural movies are more reliably processed (Haider et al., 2010; Ozeki et al., 2009). A recent study found that SST neurons primarily receive inputs from other L2/3 pyramidal neurons (Adesnik et al., 2012). Thus, in response to full-field natural movies, these recurrent connections can reliably recruit the SST network to provide surround suppression. This increased dendritic inhibition from SST neurons can change pyramidal neuron input-output functions to increase the threshold for spiking. As a result, pyramidal neurons produce sparser and more reliable responses.

A recent study found that suppressing SST neurons with Arch did not change reliability, where as suppressing PV neurons decreased reliability (Zhu et al., 2015).

This is in stark contrast to our results, which conclusively show that SST neurons are crucial for reliable coding. One possible explanation for this result is that Zhu and colleagues suppressed these inhibitory neurons for a long period of time, whereas we used a more subtle perturbation (4s vs. 20ms). Disrupting PV neurons for the entire duration of the stimulus is likely to disrupt the E-I balance of the neuron and alter other network properties (El-Boustani and Sur, 2014; Seybold et al., 2015). It is possible that the network also reduced the strength of excitatory drive following the chronic suppression of PV inhibition to preserve E-I balance. This reduced drive would in turn cause SST neurons to fire less, and would allow noisy inputs to be filtered through. Another explanation for this discrepancy is that Zhu used anesthetized mice. Anesthesia is known to decrease cortical inhibition and is especially harmful to SST neurons (Haider et al., 2013). Because SST neurons are already firing weakly, inhibiting with Arch would likely yield no effect on reliability. These discrepancies underscore the importance of using small perturbations to study the dynamics of cortical inhibition.

Inhibitory circuit between SST and PV neurons

Importantly, our work shows that SST neurons and PV neurons are tightly coupled via the SST→PV inhibitory circuit. Our work confirms pervious studies that have shown that SST neurons provide stronger inhibition onto PV neurons than onto pyramidal neurons (Cottam et al., 2013; Xu et al., 2013). Additionally, our work provides a more mechanistic insight into the previously unknown function of this inhibitory circuit. Our dual color imaging method allowed us to show that the interactions between PV and SST neurons change dynamically from movie to movie. Specifically, movies that are more reliably processed have stronger temporal correlations between PV and SST neurons, than movies that are less reliably processed. This interaction between PV and SST neurons implies that inhibition is dynamically routed between the soma and dendrites during the time course of the movie (**Figure 4-9**). We propose in this study that the SST→PV circuit serves a critical role in modulating the temporal fidelity of pyramidal neurons in the visual cortex. Specifically, given that PV neurons control the strength of feed forward inhibition (Pouille et al., 2009; Xue et al., 2014), they can alter

the temporal window over which spiking can take place (Gabernet et al., 2005; Haider and McCormick, 2009). Through the SST→PV circuit, SST neurons are able to dynamically control both dendritic integration and the duration of the spiking window at the soma. By transiently suppressing PV neurons, we propose that SST neurons allow only the strongest and most reliable inputs can be integrated at the soma to produce reliable spikes. Thus, we propose that the primary function of the SST→PV circuit is to improve the temporal fidelity of spiking by reducing noise and increasing signal (i.e. increasing SNR). This mechanism could allow feedback signals about network state or stimulus statistics to act via SST neurons to adaptively change the spiking threshold at the soma. A recent computational study demonstrated that adaptive thresholds are crucial for increasing the temporal fidelity of spiking in the presence of stochastic inputs (Fontaine et al., 2014; Huang et al., 2016). Thus, our experiments support this model and imply that this adaptive threshold could be achieved through the interactions between SST and PV neurons.

Dynamic modulation of coding reliability via the SST→PV circuit

It is a well-established fact that variability in neural responses influence how computations are implemented in neural networks (Renart and Machens, 2014). Specifically, noise in neural codes are amplified by subsequent processing stages (Moldakarimov et al., 2015). Thus, the most efficient way to encode behaviorally relevant stimuli is to reduce response variability directly in the first stage of processing. Indeed studies have shown that behaviorally relevant stimuli are more reliably encoded than other stimuli (Hires et al., 2015). Our results demonstrate the cortex can use the SST→PV circuit to modulate coding efficiency.

In addition to local recurrent inputs, SST neurons also receive strong top-down projections from higher cortical areas (Zhang et al., 2014), such as the cingulate cortex, and also cholinergic signals from the basal forebrain (Chen et al., 2015; Polack et al., 2013). A recent study found that activating cingulate projections in V1 increased neuronal gain and improved detection to low contrast stimuli (Zhang et al., 2014). Similarly, activating cholinergic axons in V1 also improved response reliability to natural

movies and improved perceptual thresholds (Pinto et al., 2013). These non-sensory inputs could act through SST neurons, and the SST→PV circuit, to improve the fidelity of visual processing (Egger et al., 2015). This provides a powerful mechanism for arousal or attention to change coding in V1. It would be interesting to record from these interneurons during a visual discrimination task to determine when the SST→PV circuit is more active.

Additionally, VIP neurons can use the SST→PV circuit to modulate coding reliability in V1. Unlike PV or SST neurons, VIP neurons rarely synapse onto pyramidal neurons. Instead, they strongly inhibit SST and, thus specialize in disinhibitory control (Pfeffer et al., 2013; Pi et al., 2013). We show that activating VIP neurons reduced pyramidal neuron reliability. We believe that this is because activating VIP neurons suppresses SST neurons. As we have demonstrated, optically suppressing SST neurons increases PV neuron activity. Thus, VIP neurons could work to shift inhibition away from the dendrites and back to the soma. Several recent papers have found that VIP neurons are strongly activated by punishment and motor outcomes (such as licking and locomotion) (Fu et al., 2014; Letzkus et al., 2015; Pinto and Dan, 2015). These findings suggest that VIP neurons might carry contextual information from other cortical areas. Thus by suppressing SST neurons, VIP neurons are likely to alter the gain of pyramidal neurons, while sacrificing temporal precision. Our results are also at odds with a recent finding that found that neurons are more reliable when then pupil is actively dilating, which also coincides with increase VIP neuron activity and reduced SST activity (Reimer et al., 2014). A more detailed dissection of the VIP→SST circuitry using the *cre-flp* system is required to better reconcile these results.

In conclusion, our study shows that the responsibility of reliable coding does not lie exclusively with one interneuron subtype. Instead, it is the mutual dynamics of SST and PV neurons, via the SST→PV circuit, that ensure that complex visual stimuli are reliably processed. Top-down projections from higher cortical areas could take advantage of this unique circuit to enhance the temporal fidelity of behaviorally relevant stimuli.

4.5 References

Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z.J., Scanziani, M., 2012. A neural circuit for spatial summation in visual cortex. Nature 490, 226–31. doi:10.1038/nature11526

Alitto, H.J., Dan, Y., 2013. Cell-type-specific modulation of neocortical activity by basal forebrain input. Front. Syst. Neurosci. 6, 79. doi:10.3389/fnsys.2012.00079

Atallah, B. V, Bruns, W., Carandini, M., Scanziani, M., 2012. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. Neuron 73, 159–70. doi:10.1016/j.neuron.2011.12.013

Baudot, P., Levy, M., Marre, O., Monier, C., Pananceau, M., Frégnac, Y., 2013. Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. Front. Neural Circuits 7, 206. doi:10.3389/fncir.2013.00206

Bloss, E.B., Cembrowski, M.S., Karsh, B., Colonell, J., Fetter, R.D., Spruston, N., 2016. Structured Dendritic Inhibition Supports Branch-Selective Integration in CA1 Pyramidal Cells. Neuron 89, 1016–30. doi:10.1016/j.neuron.2016.01.029

Branco, T., Häusser, M., 2011. Synaptic integration gradients in single cortical pyramidal cell dendrites. Neuron 69, 885–92. doi:10.1016/j.neuron.2011.02.006

Butts, D. a, Weng, C., Jin, J., Yeh, C.-I., Lesica, N. a, Alonso, J.-M., Stanley, G.B., 2007. Temporal precision in the neural code and the timescales of natural vision. Nature 449, 92–5. doi:10.1038/nature06105

Chance, F.S., Abbott, L.F., 2000. Divisive inhibition in recurrent networks. Network 11, 119–29.

Chen, N., Sugihara, H., Sur, M., 2015. An acetylcholine-activated microcircuit drives temporal dynamics of cortical activity. Nat Neurosci 18, 892–902. doi:10.1038/nn.4002\rhttp://www.nature.com/neuro/journal/v18/n6/abs/nn.4002.html#supplementary-information

Chiu, C.Q., Lur, G., Morse, T.M., Carnevale, N.T., Ellis-Davies, G.C.R., Higley, M.J., 2013. Compartmentalization of GABAergic inhibition by dendritic spines. Science 340, 759–62. doi:10.1126/science.1234274

Cottam, J.C.H., Smith, S.L., Häusser, M., 2013. Target-specific effects of somatostatin-

expressing interneurons on neocortical visual processing. J. Neurosci. 33, 19567–78. doi:10.1523/JNEUROSCI.2624-13.2013

Dana, H., Mohar, B., Sun, Y., Narayan, S., Gordus, A., Hasseman, J.P., Tsegaye, G., Holt, G.T., Hu, A., Walpita, D., Patel, R., Macklin, J.J., Bargmann, C.I., Ahrens, M.B., Schreiter, E.R., Jayaraman, V., Looger, L.L., Svoboda, K., Kim, D.S., 2016. Sensitive red protein calcium indicators for imaging neural activity. Elife 5. doi:10.7554/eLife.12727

Egger, R., Schmitt, A.C., Wallace, D.J., Sakmann, B., Oberlaender, M., Kerr, J.N.D., 2015. Robustness of sensory-evoked excitation is increased by inhibitory inputs to distal apical tuft dendrites. Proc. Natl. Acad. Sci. U. S. A. 112, 14072–7. doi:10.1073/pnas.1518773112

El-Boustani, S., Sur, M., 2014. Response-dependent dynamics of cell-specific inhibition in cortical networks in vivo. Nat. Commun. 5, 5689. doi:10.1038/ncomms6689

Fu, Y., Tucciarone, J.M., Espinosa, J.S., Sheng, N., Darcy, D.P., Nicoll, R.A., Huang, Z.J., Stryker, M.P., 2014. A cortical circuit for gain control by behavioral state. Cell 156, 1139–52. doi:10.1016/j.cell.2014.01.050

Gabernet, L., Jadhav, S.P., Feldman, D.E., Carandini, M., Scanziani, M., 2005. Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. Neuron 48, 315–27. doi:10.1016/j.neuron.2005.09.022

Gidon, A., Segev, I., 2012. Principles Governing the Operation of Synaptic Inhibition in Dendrites. Neuron 75, 330–341. doi:10.1016/j.neuron.2012.05.015

Haider, B., Häusser, M., Carandini, M., 2013. Inhibition dominates sensory responses in the awake cortex. Nature 493, 97–100. doi:10.1038/nature11665

Haider, B., Krause, M.R., Duque, A., Yu, Y., Touryan, J., Mazer, J. a, McCormick, D. a, 2010. Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. Neuron 65, 107–21. doi:10.1016/j.neuron.2009.12.005

Haider, B., McCormick, D. a, 2009. Rapid neocortical dynamics: cellular and network mechanisms. Neuron 62, 171–89. doi:10.1016/j.neuron.2009.04.008

Hioki, H., Okamoto, S., Konno, M., Kameda, H., Sohn, J., Kuramoto, E., Fujiyama, F., Kaneko, T., 2013. Cell type-specific inhibitory inputs to dendritic and somatic compartments of parvalbumin-expressing neocortical interneuron. J. Neurosci. 33, 544–55.

doi:10.1523/JNEUROSCI.2255-12.2013

Hires, S.A., Gutnisky, D.A., Yu, J., Connor, D.H.O., Svoboda, K., 2015. Low-noise encoding of active touch by layer 4 in the somatosensory cortex 1–18. doi:10.7554/eLife.06619

Isaacson, J.S., Scanziani, M., 2011. How inhibition shapes cortical activity. Neuron 72, 231–43. doi:10.1016/j.neuron.2011.09.027

Jadi, M., Polsky, A., Schiller, J., Mel, B.W., 2012. Location-dependent effects of inhibition on local spiking in pyramidal neuron dendrites. PLoS Comput. Biol. 8, e1002550. doi:10.1371/journal.pcbi.1002550

Kepecs, A., Fishell, G., 2014. Interneuron cell types are fit to function. Nature 505, 318–26. doi:10.1038/nature12983

Lee, S.-H., Kwan, A.C., Zhang, S., Phoumthipphavong, V., Flannery, J.G., Masmanidis, S.C., Taniguchi, H., Huang, Z.J., Zhang, F., Boyden, E.S., Deisseroth, K., Dan, Y., 2012. Activation of specific interneurons improves V1 feature selectivity and visual perception. Nature 488, 379–83. doi:10.1038/nature11312

Letzkus, J.J., Wolff, S.B.E., Lu, A., 2015. Review Disinhibition , a Circuit Mechanism for Associative Learning and Memory. doi:10.1016/j.neuron.2015.09.024

Letzkus, J.J., Wolff, S.B.E., Meyer, E.M.M., Tovote, P., Courtin, J., Herry, C., Lüthi, A., 2011. A disinhibitory microcircuit for associative fear learning in the auditory cortex. Nature 480, 331–5. doi:10.1038/nature10674

Lovett-Barron, M., Turi, G.F., Kaifosh, P., Lee, P.H., Bolze, F., Sun, X.-H., Nicoud, J.-F., Zemelman, B. V, Sternson, S.M., Losonczy, A., 2012. Regulation of neuronal input transformations by tunable dendritic inhibition. Nat. Neurosci. 15, 423–430. doi:10.1038/nn.3024

Madisen, L., Mao, T., Koch, H., Zhuo, J., Berenyi, A., Fujisawa, S., Hsu, Y.-W.A., Garcia, A.J., Gu, X., Zanella, S., Kidney, J., Gu, H., Mao, Y., Hooks, B.M., Boyden, E.S., Buzsáki, G., Ramirez, J.M., Jones, A.R., Svoboda, K., Han, X., Turner, E.E., Zeng, H., 2012. A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. Nat. Neurosci. 15, 793–802. doi:10.1038/nn.3078

Mainen, Z.F., Sejnowski, T.J., 1995. Reliability of spike timing in neocortical neurons. Science 268, 1503–1506. doi:10.1126/science.7770778

Markram, H., Toledo-Rodriguez, M., Wang, Y., Gupta, A., Silberberg, G., Wu, C., 2004. Interneurons of the neocortical inhibitory system. Nat. Rev. Neurosci. 5, 793–807. doi:10.1038/nrn1519

Moldakarimov, S., Bazhenov, M., Sejnowski, T.J., 2015. Feedback stabilizes propagation of synchronous spiking in cortical neural networks. Proc. Natl. Acad. Sci. U. S. A. 112, 2545–2550. doi:10.1073/pnas.1500643112

Neske, G.T., Patrick, S.L., Connors, B.W., 2015. Contributions of Diverse Excitatory and Inhibitory Neurons to Recurrent Network Activity in Cerebral Cortex. J. Neurosci. 35, 1089–1105. doi:10.1523/JNEUROSCI.2279-14.2015

Okun, M., Lampl, I., 2008. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nat. Neurosci. 11, 535–7. doi:10.1038/nn.2105

Ozeki, H., Finn, I.M., Schaffer, E.S., Miller, K.D., Ferster, D., 2009. Inhibitory stabilization of the cortical network underlies visual surround suppression. Neuron 62, 578–92. doi:10.1016/j.neuron.2009.03.028

Pfeffer, C.K., Xue, M., He, M., Huang, Z.J., Scanziani, M., 2013. Inhibition of inhibition in visual cortex: the logic of connections between molecularly distinct interneurons. Nat. Neurosci. 1–12. doi:10.1038/nn.3446

Pi, H.-J., Hangya, B., Kvitsiani, D., Sanders, J.I., Huang, Z.J., Kepecs, A., 2013. Cortical interneurons that specialize in disinhibitory control. Nature 503, 521–4. doi:10.1038/nature12676

Pinto, L., Dan, Y., 2015. Cell-Type-Specific Activity in Prefrontal Cortex during Goal-Directed Behavior. Neuron 87, 437–450. doi:10.1016/j.neuron.2015.06.021

Pinto, L., Goard, M.J., Estandian, D., Xu, M., Kwan, A.C., Lee, S.-H., Harrison, T.C., Feng, G., Dan, Y., 2013. Fast modulation of visual perception by basal forebrain cholinergic neurons. Nat. Neurosci. 16, 1857–63. doi:10.1038/nn.3552

Polack, P.P.-O., Friedman, J., Golshani, P., 2013. Cellular mechanisms of brain state-dependent gain modulation in visual cortex. Nat. Neurosci. 16, 1–11. doi:10.1038/nn.3464

Pouille, F., Marin-Burgin, A., Adesnik, H., Atallah, B. V, Scanziani, M., 2009. Input normalization by global feedforward inhibition expands cortical dynamic range. Nat. Neurosci. 12, 1577–1585. doi:10.1038/nn.2441

Pouille, F., Scanziani, M., 2001. Enforcement of temporal fidelity in pyramidal cells by somatic feed-forward inhibition. Science 293, 1159–1163. doi:10.1126/science.1060342

Pouille, F., Watkinson, O., Scanziani, M., Trevelyan, A.J., 2013. The contribution of synaptic location to inhibitory gain control in pyramidal cells. Physiol. Rep. 1, n/a–n/a. doi:10.1002/phy2.67

Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D., Denfield, G.H., Tolias, A.S., 2014. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. Neuron 84, 355–62. doi:10.1016/j.neuron.2014.09.033

Renart, A., Machens, C.K., 2014. Variability in neural activity and behavior. Curr. Opin. Neurobiol. 25, 211–220. doi:10.1016/j.conb.2014.02.013

Rikhye, R. V, Sur, M., 2015. Spatial Correlations in Natural Scenes Modulate Response Reliability in Mouse Visual Cortex. J. Neurosci. 35, 14661–80. doi:10.1523/JNEUROSCI.1660-15.2015

Seybold, B.A., Phillips, E.A.K., Schreiner, C.E., Hasenstaub, A.R., 2015. Inhibitory Actions Unified by Network Integration. Neuron 87, 1181–92. doi:10.1016/j.neuron.2015.09.013

Shadlen, M.N., Newsome, W.T., 1998. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. 18, 3870–96.

Singh, A., Lesica, N.A., 2010. Incremental mutual information: a new method for characterizing the strength and dynamics of connections in neuronal circuits. PLoS Comput. Biol. 6, e1001035. doi:10.1371/journal.pcbi.1001035

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

Sprague, T.C., Saproo, S., Serences, J.T., 2015. Visual attention mitigates information loss in small- and large-scale neural codes. Trends Cogn. Sci. doi:10.1016/j.tics.2015.02.005

Stevens, C.F., Zador, A.M., 1998. Input synchrony and the irregular firing of cortical neurons. Nat. Neurosci. 1, 210–217. doi:10.1038/659

Sturgill, J.F., Isaacson, J.S., 2015. Somatostatin cells regulate sensory response fidelity via subtractive inhibition in olfactory cortex. Nat. Neurosci. 18, 531–5. doi:10.1038/nn.3971

Theis, L., Berens, P., Froudarakis, E., Reimer, J., Román Rosón, M., Baden, T., Euler, T., Tolias, A.S., Bethge, M., 2016. Benchmarking Spike Rate Inference in Population Calcium

Imaging. Neuron 90, 471–82. doi:10.1016/j.neuron.2016.04.014

Tolhurst, D.J., Movshon, J.A., Dean, A.F., 1983. The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res. 23, 775–85.

Vu, E.T., Krasne, F.B., 1992. Evidence for a computational distinction between proximal and distal neuronal inhibition. Science 255, 1710–2.

Wall, N.R., De La Parra, M., Sorokin, J.M., Taniguchi, H., Huang, Z.J., Callaway, E.M., 2016. Brain-Wide Maps of Synaptic Input to Cortical Interneurons. J. Neurosci. 36, 4000–9. doi:10.1523/JNEUROSCI.3967-15.2016

Wehr, M., Zador, A.M., 2003. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. Nature 426, 442–446. doi:10.1038/nature02116

Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., Tanaka, J.W., 2010. Controlling low-level image properties: the SHINE toolbox. Behav. Res. Methods 42, 671–84. doi:10.3758/BRM.42.3.671

Wilson, N.R., Runyan, C. a, Wang, F.L., Sur, M., 2012. Division and subtraction by distinct cortical inhibitory networks in vivo. Nature 488, 343–8. doi:10.1038/nature11347

Xu, H., Jeong, H.-Y., Tremblay, R., Rudy, B., 2013. Neocortical Somatostatin-Expressing GABAergic Interneurons Disinhibit the Thalamorecipient Layer 4. Neuron 77, 155–167. doi:10.1016/j.neuron.2012.11.004

Xue, M., Atallah, B. V, Scanziani, M., 2014. Equalizing excitation-inhibition ratios across visual cortical neurons. Nature 511, 596–600. doi:10.1038/nature13321

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.C., Jenvay, S., Miyamichi, K., Luo, L., Dan, Y., 2014. Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science (80-.). 345, 660–665. doi:10.1126/science.1254126

Zhu, Y., Qiao, W., Liu, K., Zhong, H., Yao, H., 2015. Control of response reliability by parvalbumin-expressing interneurons in visual cortex. Nat. Commun. 6, 1–11. doi:10.1038/ncomms7802

Appendix 4A– Supplementary Figures

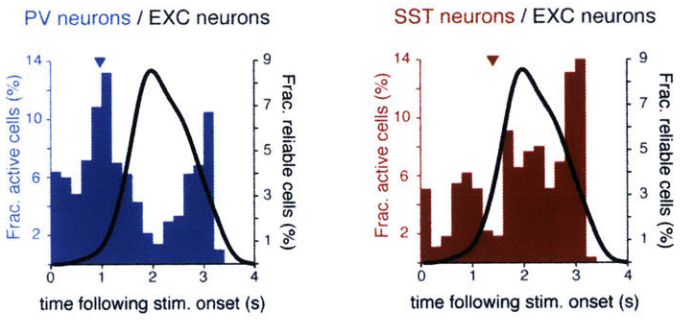

Figure 4-S1. SST neurons are most active when pyramidal neurons are most reliable. Distribution of active PV neurons (left) and SST neurons (right) at different movie epochs. The black line shows the distribution of reliably activated EXC neurons. PV neurons are the least active and SST neurons are the most active when pyramidal neurons are the most reliable. Correlation analysis reveals a significant positive correlation between SST neuron activation and EXC neuron reliability ($p < 0.001$) and a significant negative correlation with PV neuron activation ($p < 0.01$).

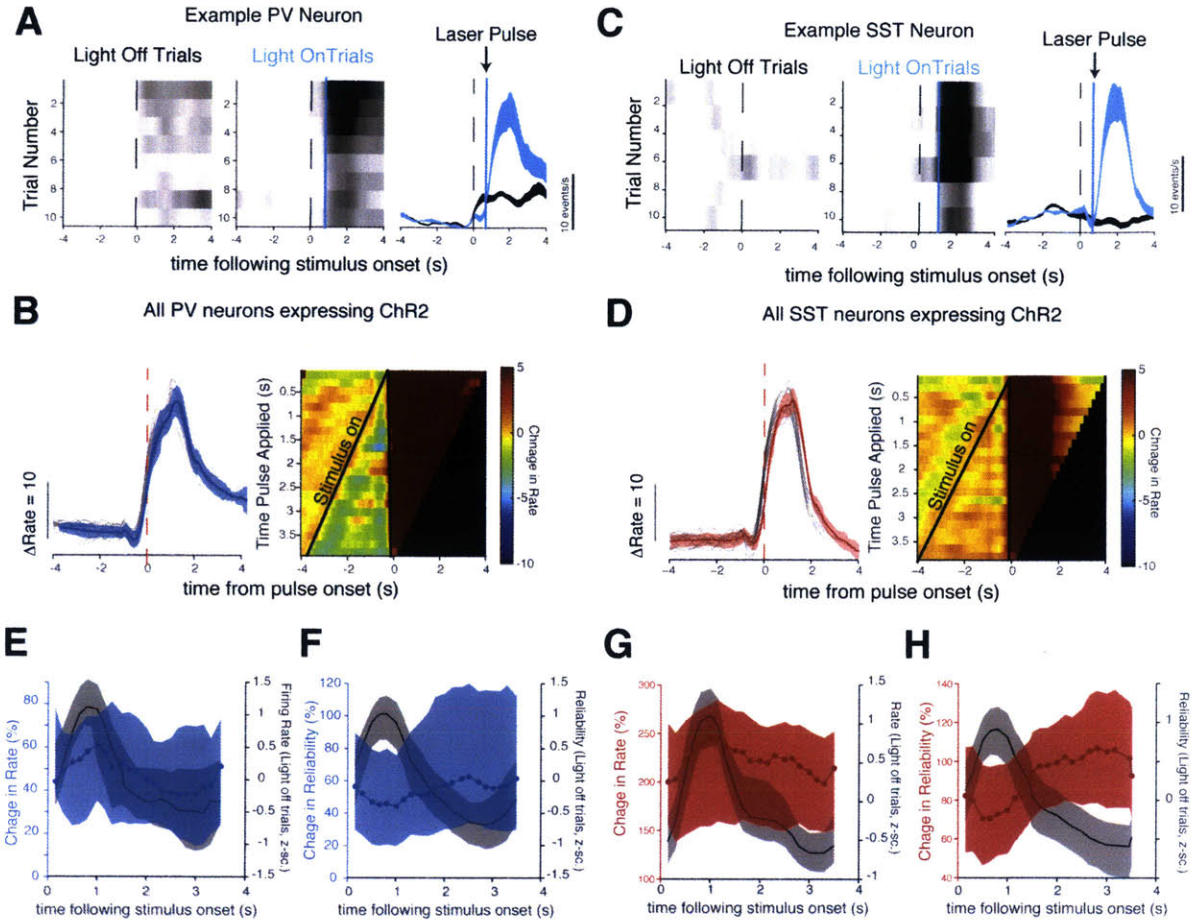

Figure 4-S2. Laser activation reliably activates both PV and SST neurons. (A) Example raster plots of a PV neuron expressing both GCaMP6f and ChR2. Shaded area, SEM over trials. **(B)** Firing rate trace aligned to pulse onset. Heat map shows average over all PV neurons. **(C, D)** Same as (A, B) but for SST neurons instead. This shows that both PV and SST neurons are also reliably and strongly activated following laser activation. **(E, G)** Change in firing rate following each laser pulse. **(F, H)** Change in reliability caused by each laser pulse. Data from: PV-cre = 3 mice (45 neurons), SST-cre = 3 mice (52 neurons).

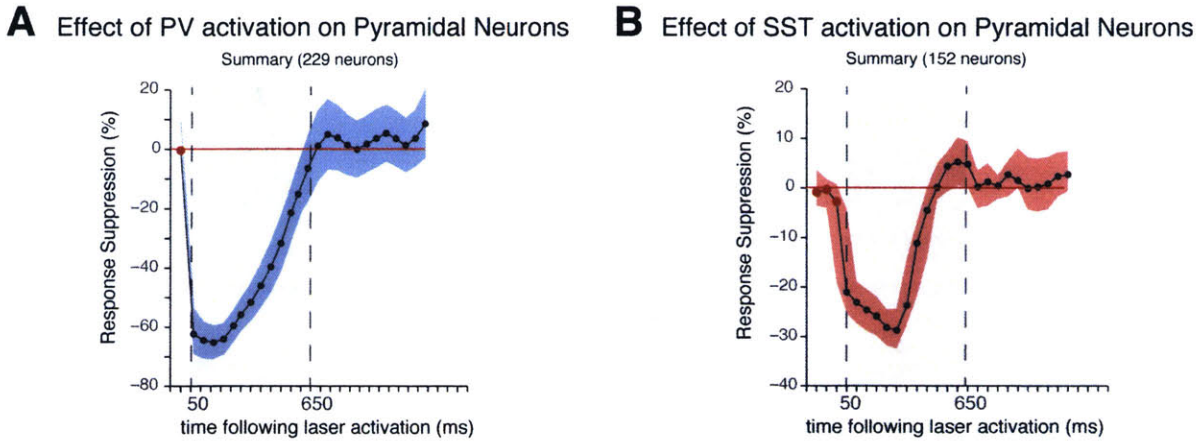

Figure 4-S3. Dynamics of laser induced suppression in pyramidal neurons. Percent change in pyramidal cell response following laser activation of PV neurons **(A)** and SST neurons **(B)**. All data analysis was limited to a 600ms window indicated by the gray box. During this period, the laser maximally suppresses pyramidal neurons.

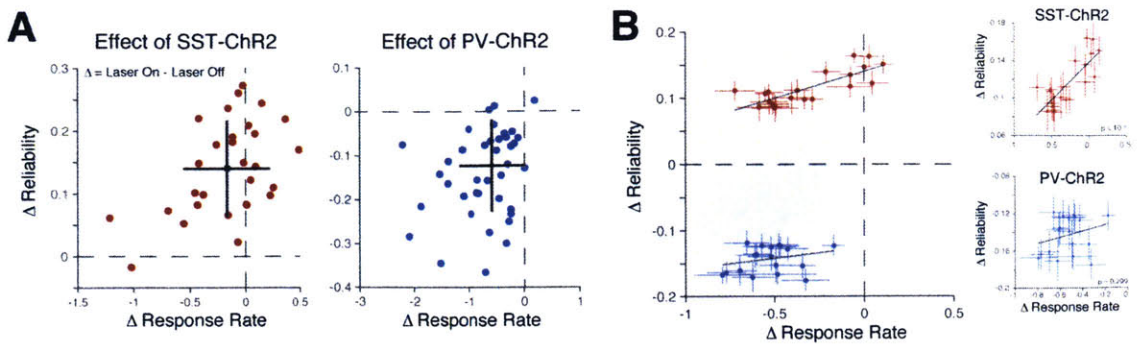

Figure 4-S4. SST neuron activation increases reliability despite decreasing mean response rate. (A) Scatter plot of change in reliability versus change in response rate for two example pyramidal neuron populations. Both these populations show an almost linear relationship between response rate and reliability. In both cases, following SST (left) or PV activation (right), neurons with the largest reduction in response rate also had the largest change in reliability. This change in reliability is positive following SST activation (indicating increase) and negative following PV activation. The black crosses represent mean +/- SEM. **(B)** Scatter plot showing relationship between mean firing rate and mean response rate for all imaged populations (SST: 20, PV: 22). Crosses again represent mean +/- SEM for each population. This plot shows that for the same reduction in response rate, SST activation increases in pyramidal reliability, whereas PV activation decreases reliability (indicated by gray box). Regression analysis (right panel) indicates that PV activation-induced change in reliability is independent of the change in response rate ($F(16) = 1.15$, $p = 0.299$, F-test). This means that, regardless of the effect on response rate, PV neuron activation always reduces pyramidal neuron reliability. In contrast, SST activation-induced change in reliability is strongly correlated with the change in response rate ($F(18) = 31.2$, $p = 2.66 \times 10^{-5}$). Thus, SST activation improves the reliability of pyramidal neurons, which in turn leads to an increase in mean response rate, possibly due to the dual effect that SST neurons have on somatic and dendritic inhibition.

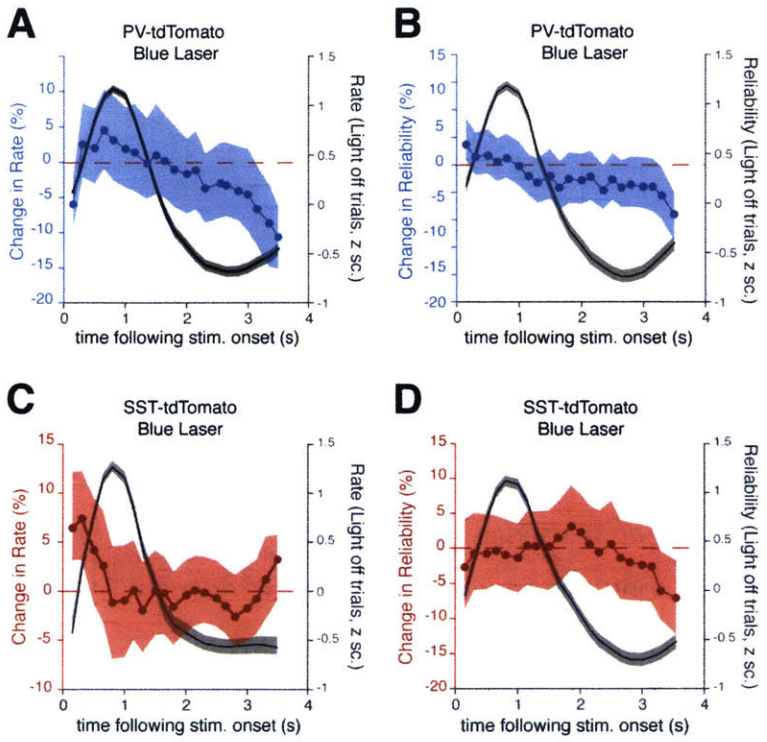

Figure 4-S5. Photostimulation of non-ChR2 expressing neurons does not cause change in firing rate or reliability. (A-B) Photo-stimulation (blue light) applied to PV-cre::Ai14 mice (PV neurons expressing tdTomato) does not change pyramidal neuron response rate **(A)** or reliability **(B)**. **(C-D)** Similarly, photo-stimulation applied to SST-cre::Ai14 mice (SST neurons expressing tdTomato) does not change pyramidal neuron firing rate **(C)** or reliability **(D)**. Thus, the effects of inhibitory neuron activation on mean response rate and reliability that we have observed are due to an increase in inhibition, and are not due to an artifact induced by photo-stimulation. Data in both conditions collected from 3 mice each (SST-tdTomato = 367 neurons, PV-tdTomato = 280 neurons). Shaded areas, 95% CI of the median.

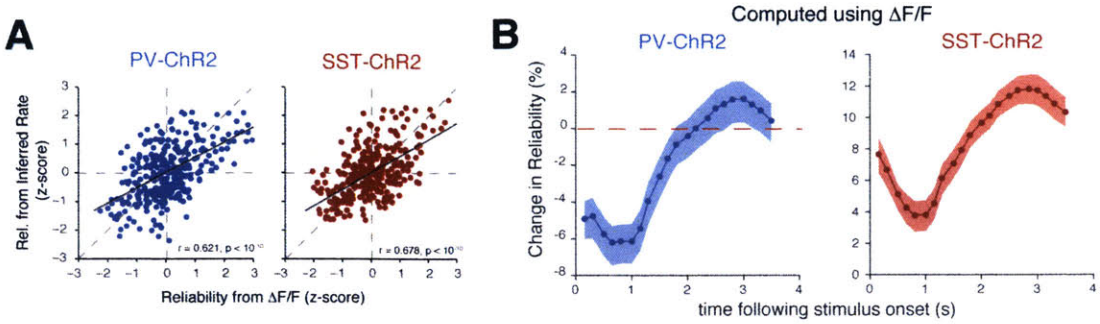

Figure 4-S6. Deconvolution does not affect results. (A) Reliability computed from inferred firing rates is similar to reliability measured from raw fluorescence changes (DF/F). Scatter plot shows that neurons with reliable DF/F will also have reliable inferred firing rates. **(B)** Change in reliability measured using DF/F without deconvolution.

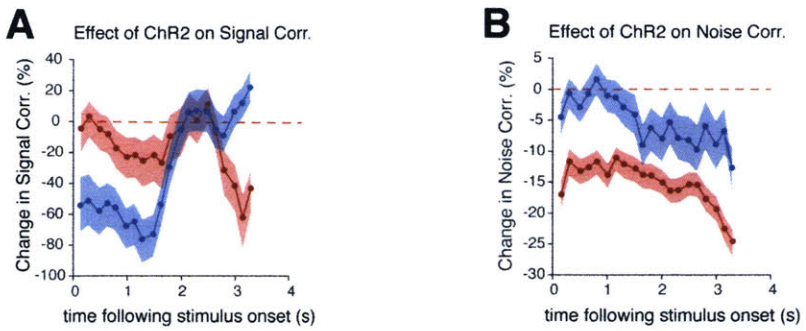

Figure 4-S7. Effect of ChR2 activation on inter-neuronal correlations. (A) Activating PV neurons strongly decreases signal correlations primarily due to a decrease in mean firing rate. In contrast, activating SST neurons does not change signal correlation. **(B)** Activating SST neurons strongly reduces noise correlations regardless when the laser pulse was applied during the movie.

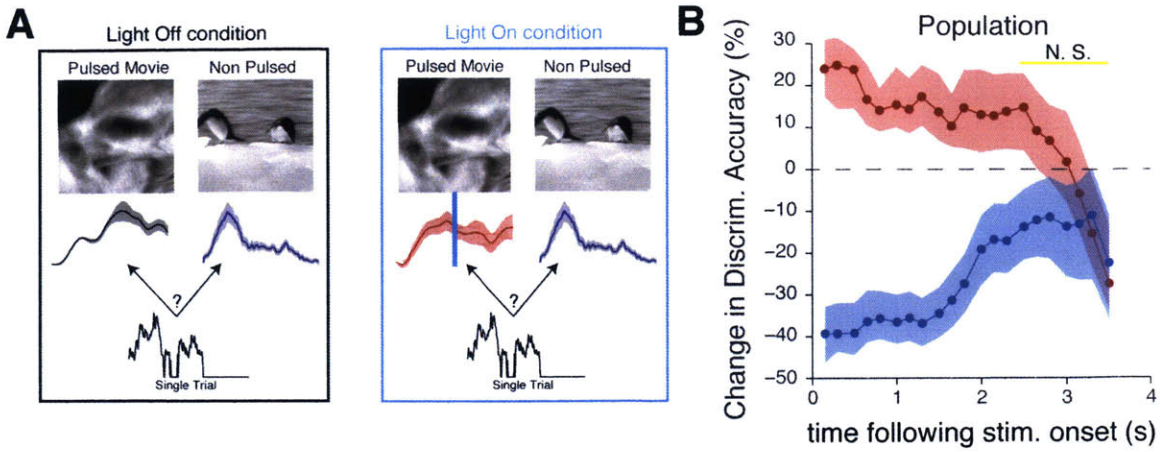

Figure 4-S8. Activating SST neurons improves stimulus discriminability. (A) Example showing template matching decoding method. Briefly, a single trial response was classified into one of two movies – Pulsed movie and Non-Pulsed movie – based on its Euclidean distance to a population template. Decoding accuracy was compared between light-on and light-off conditions. **(B)** Activating SST neurons early on in the movie improves discriminability, whereas activating PV neurons decreases discriminability. Discrimination Accuracy is computed as the number of correct classifications made by the classifier. Data used in this figure same as **Figure 4-5**. Shaded error-bars denote 95% CI of the median.

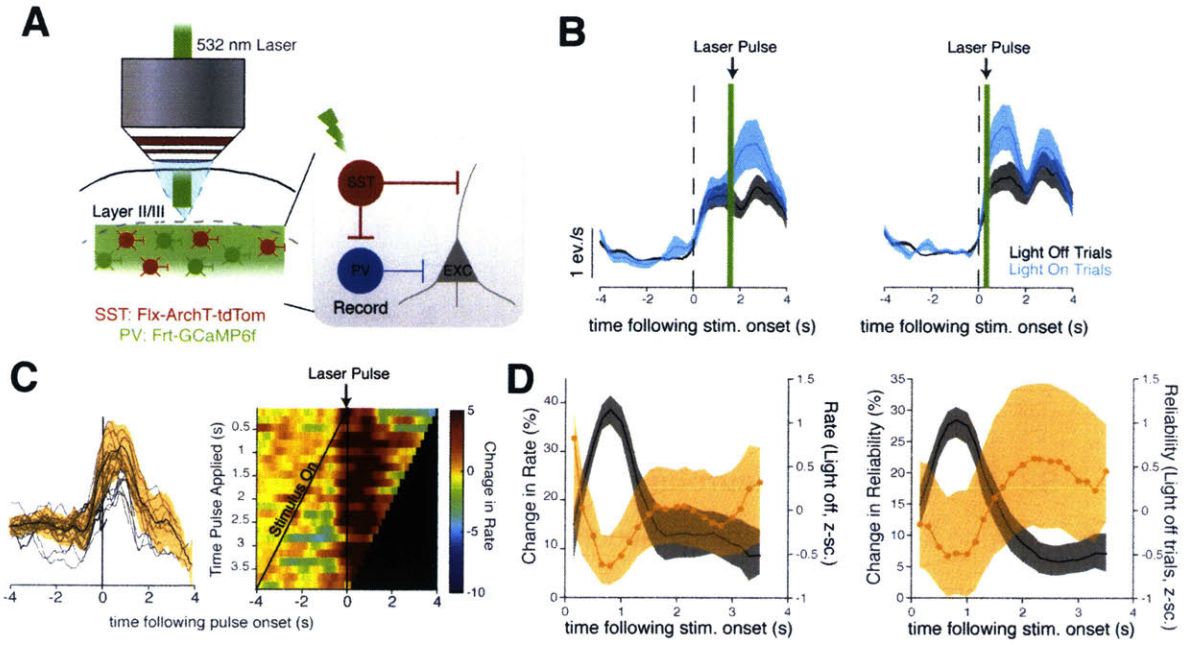

Figure 4-S9. SST inactivation disinhibits PV neurons. (A) Cartoon describing experimental setup. **(B)** Effect of SST inactivation at different times during a movie on a PV neuron. Shaded area, SEM over trials. **(C)** Change in response rate aligned to pulse onset. Heat map shows average change in response rate of all PV neurons (112 neurons) following each laser pulse. There is always an increase in PV neuron activity following SST inhibition. **(D)** Quantification of change in response rate (left) and change in reliability (right) of PV neurons following SST inactivation. We observe a significant increase in both activity ($P < 0.001$) and reliability ($P < 0.001$). Data from 3 mice (112 neurons). Shaded area in **(D)** is 95% CI of the median. P-values computed using Permutation test.

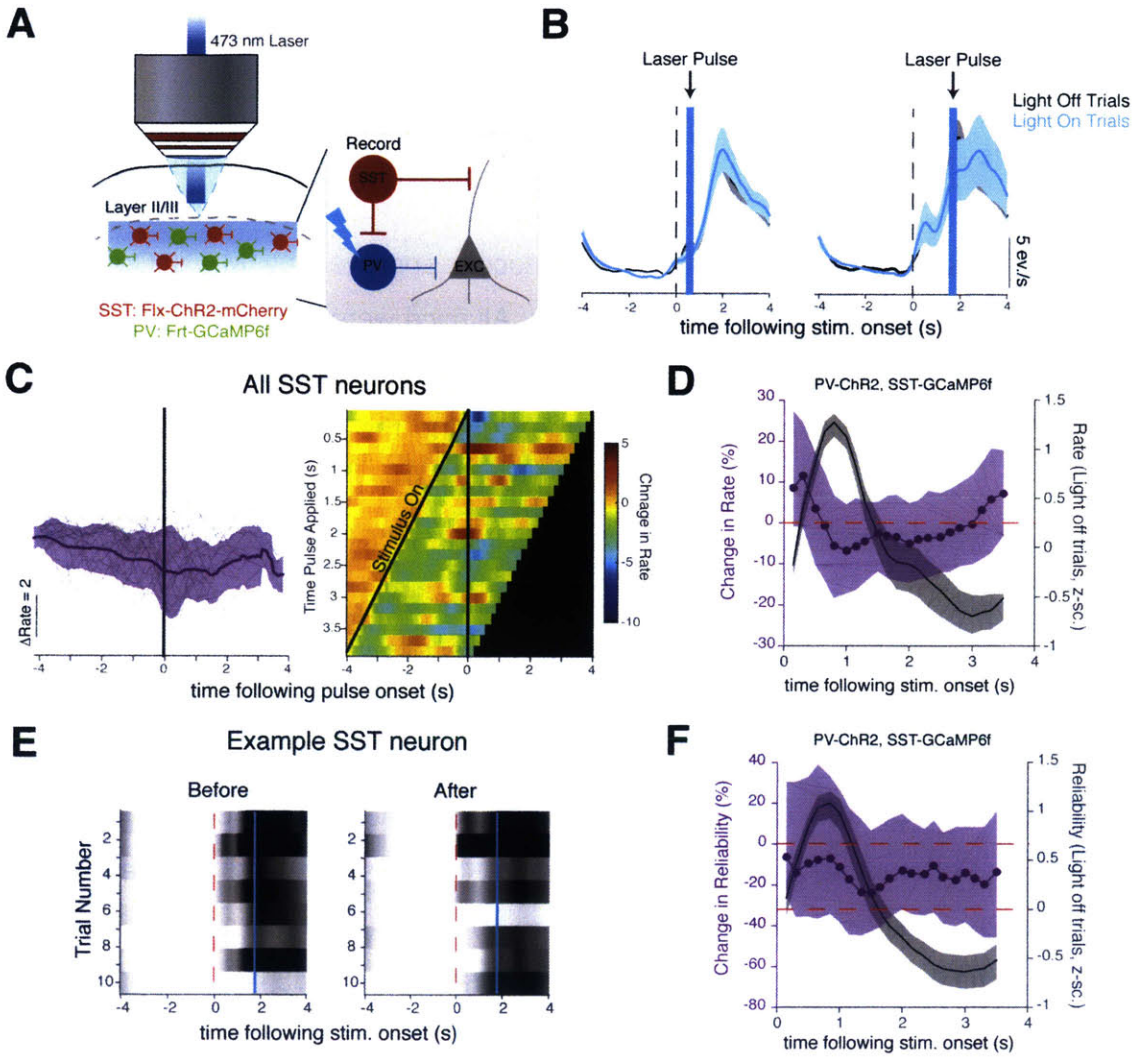

Figure 4-S10. PV activation does not affect SST neurons. (A) Cartoon describing experimental setup. To determine the existence of a PV→SST connection, we expressed ChR2 in PV-flp neurons GCaMP6f into SST-cre neurons. **(B)** Example SST neuron showing no effect following PV-ChR2 photo-stimulation. Shaded area, SEM over trials. **(C)** Change in response rate aligned to pulse onset. Heat map shows average change in response rate of all SST neurons for each laser pulse. **(D)** Activating PV neurons does not significantly change mean response rate of SST neurons. P = 0.125. **(E)** Raster plot showing no effect of PV-ChR2 photo-stimulation on SST reliability. **(F)** Activating PV neurons does not change reliability of SST neurons. P = 0.762. Data from 3 mice (72 neurons). Shaded area in D, F is 95% CI of median. P-values computed using Permutation test.

Appendix 4B – Experimental Procedures

Experiments

Animals

All experiments were carried out under protocols approved by MIT's Committee on Animal Care and conformed to NIH guidelines. All mice were maintained on a C57BL6/J background. Mice were housed in a standard 12 hour light-dark cycle with *ad libitium* food and water. Experiments were performed during the light cycle. Only mice older than 8 weeks old were used for this study. Mice of either sex were used. Three main mouse lines were used in this study: Pvalb-IRES-Cre (PV-Cre, Jax: 008069), Sst-IRES-Cre (Sst, Jax: 013044) and Pval-2A-FlpO-D (PV-Flp, Jax: 022730).

For cell specific imaging (**Figure 4-2**), we obtained pyramidal neurons from tdTomato negative neurons in SST-tdTomato or PV-tdTomato animals. We obtained these mice by crossing homozygous PV-Cre and SST-Cre mice with homozygous Ai14 mice (Jax: 007908). We used the same mice for optogenetics control experiments (Figure 4-X). The number of mice used for Figure 4-2 is as follows: SST-Cre (Imaging) = 11 mice (328 neurons), PV-Cre (Imaging) = 10 mice (690 neurons), SST-tdTomato (SST-Cre x Ai14) = 5 mice (541 neurons) and PV-tdTomato (SST-Cre x Ai14) = 5 mice (560 neurons).

For photostimulation experiments (**Figure 4-4, 5, 7**) we crossed homozygous PV-Cre or SST-Cre mice with homozygous Ai32 mice (to express ChR2, Jax:012569) or Ai35 mice (to express Arch, Jax: 012735). All pups from these crosses were genotyped at P28 to verify the correct expression of ChR2 or Arch. Thus, these mice transgenically expressed either ChR2 or Arch in all SST/PV neurons (Madisen et al., 2011). We used these SST-cre::Ai32 mice and PV-cre::Ai32 mice for behavioral experiments (Figure 4-11). In these mice, we co-injected a 1:1 mixture of AAV1-Syn-GCaMP6f.WPRE.SV40 (University of Pennsylvania Vector Core) and AAV1-CAG-flex-TdTomato (UNC Vector core). The tdTomato injection to helped us to better visualize the ChR2-expressing neurons. For positive control experiments aimed at testing the effectiveness of ChR2 at activating PV an SST interneurons, we co-injected a 1:1 mixture of AAV1-EF1-dflox-

hChR2(H134R)-mCherry.WPRE(flex-ChR2)and AAV1-Syn-Flex-GCaMP6f.WPRE.SV40 (flex-GCaMP6f, both from Penn vector core) into PV-cre and SST-cre mice. Similarly to test the effectiveness of Arch at suppressing we co-injected AAV1-CAG-FLEX-Arch-GFP (UNC vector core) with flex-GCaMP6f in PV-cre mice (**Figure 4-7**).

To create SST-Cre::PV-Flp (SXP) mice (**Figure 4-3, 6**), we crossed male homozygous Sst-IRES-Cre mice with female homozygous Pvalb-2A-FlpO-D mice. Pups were genotyped at P28 to verify the expression of Cre in SST neurons and Flp in PV neurons. For dual color imaging experiments (**Figure 4-3**), we co-injected a 1:1 mixture of AAV1-Syn-Flex-NES-jRGECO1a.WPRE.SV40 (flex-jRGeco1a, Penn vector core) and AAV1-syn-fdio-GCaMP6f.WPRE.SV40 (frt-GCaMP6f) into these SXP mice. To examine the SST→PV disinhibitory circuit (**Figure 4-6**), we co-expressed flex-ChR2 in SST neurons and frt-GCaMP6f in PV neurons.

Surgery

Adult mice (between 6-10 weeks old) were an anesthetized with 1–2% isoflurane (vol/vol) and a circular piece of scalp was removed (Gooley et al., 2015). After cleaning and drying the skull using a razor blade and sterile cotton swabs, a custom-built head-post was implanted to the exposed skull with cyanoacrylate glue (Loctite) and cemented with dental acrylic mixed with black paint (C&B Metabond). A craniotomy (3 mm in diameter) was made over the left V1 (2.5 mm lateral and 0.5 mm anterior to lambda). Care was taken not to damage the dura during the craniotomy. Viruses (described above) were then injected were injected using a beveled pipette (20-30-µm diameter tip Drummond Scientific) backfilled with mineral oil at a speed of 50nl/min at 5-6 injections sites. The 100-150nl of virus was injected per injection site. Notably, in experiments when we co-injected flex-ChR2 and flex-GCaMP6f we only injected 50nl at two sites. We found that this helped to minimize over expression of GCaMP6f. After each injection, pipettes were left in the brain for an additional 5-10 minutes (depending on injection volume) to prevent backflow and to ensure proper virus spread. Following virus injections, a chronic imaging window was placed in the craniotomy. The imaging window consisted of an inner 3mm glass window and an outer 5mm glass window (Warner

Scientific), which were glued together using UV curing glue (Norland Optical). Once mice recovered from anesthesia, they were returned to their home cage (singly housed). Mice were provided with analgesia (meloxicam, 0.1 mg per kg of body weight) subcutaneously three days post-surgery. Imaging experiments typically started 14 to 21 days post-surgery. Mice with limited optical access due to bone growth or virus infection were excluded.

Two-photon calcium imaging

Imaging was performed using a Prairie Ultima two-photon system (Prairie Technologies) driven by two Spectra Physics Mai-Tai lasers, both passed through a Deep-See modules (Spectra Physics). Imaging was performed using high performance objective lens (25x Olympus XL Plan N objective, NA = 1.05). In most experiments (except dual color imaging, see below), we tuned the laser to 965nm to enable us to optimally visualize both GCaMP6f and tdTomato fluorescence. To separate red and green fluoresence, we used a 565nm dichroic filter, a 520/40nm green filter and a 600/50nm red filter (Chroma). We used a removable curtain made from blackout material (Thorlabs) to isolate the visual display from the microscope.

In dual color imaging experiments, we tuned one laser to 920nm to excite GCaMP6f and another laser to 1020nm (limit of the laser) to excite jRGECO1a. Both laser beams were combined using a half wave plate and a polarizing beam splitter (Thorlabs) before being focused onto a pair of galvanometer mirrors. In doing so, the pair of mirrors scanned two laser beams over the same FOV. This allowed us to image from jRGeco1a-expressing SST and GCaMP6f-expressing PV neurons.

In all experiments, images were acquired using ScanImage (Vidrio Technologies) at 20 Hz, 512 x 100 pixels (250 x 250 μm, 2x optical zoom). Images were collected at a depth of 180-280 μm below the pial surface. Typically, 5-8 non-overlapping FOV (each an independent neural population) was collected for each mouse. Images with photostimulation laser artifacts were discarded from analysis (usually 1-2 frames) and spline interpolation was used to smooth over these blanked frames. Images were corrected for brain movements using an ImageJ plugin that maximized the cross-

correlation coefficient between frames. Neuronal ROIs were manually segmented in ImageJ (Cell Magic wand tool) and fluorescence time series for each cell was computed by averaging each ROI. This data was then imported into Matlab for further analysis.

Optical activation and inactivation

A 473nm (blue) laser and a 532nm laser (both from Optoengine) we were used to activate ChR2 and Arch respectively. Both lasers were coupled to a 0.12NA optical fiber (Thorlabs) and these fibers were launched into the uncaging beam path of the two-photon microscope. The uncaging beam path was co-aligned with the imaging path such that the single wavelength laser illuminated the same FOV as the two-photon laser. In this way, we were able to provide focused activation (or inactivation) of neurons within the same FOV. These single wavelength lasers were triggered using a TTL pulse generated by the visual stimulus computer (see description below). Laser power at the tip of the objective was 1.1-1.5mW for ChR2 and 2-2.5mW for Arch experiments respectively. Laser power was measured before the start of each experiment. Using single cell imaging, we determined that these laser powers were sufficient to reliably drive activation (or suppression) of PV and SST neurons (see supplementary figure). We typically used less power to drive PV neurons (1.2mW on average) and more power to drive SST neurons (1.2mW on average). These power values are consistent with previously published reports using similar mice (Madisen et al., 2012; Seybold et al., 2015).

In all experiments, we used a stimulus-triggered random stimulation protocol to activate/inactivate cells. Each stimulation pattern consisted of four 10ms pulses of laser with a 10ms interpulse interval (i.e. 70ms total duration per stimulation epoch). In Arch experiments, we used 4x20ms laser pulses with a 10ms inter-pulse interval. This pulse pattern was applied at 22 different frames during a natural movie, with the first pulse occurring at stimulus onset (i.e. triggered by frame one). Also, each pulse pattern was repeated 10 times. This resulted in a total of 220 laser-on events and 10 laser-off events (used for controls). The frame numbers that triggered the pulses were the same for each experiment but were pseudo-randomized before the start of each experiment. This

helped to minimize spurious network activity caused by rhythmic photo-stimulation. Also, we interleaved the "pulsed" movie with a "non-pulsed" movie, during which no laser was applied. As a consequence, the network was allowed at least 8s to recover before the next laser pulse was applied.

Visual stimulus

All stimuli were presented on a 7-inch 1080p LCD computer monitor (Xenarc) placed 3-inches in front of the ipsilateral eye. This computer monitor covered a visual space of approximately 50x70 degrees. The monitor was gamma corrected before all experiments. All visual stimuli were presented using Psychtoolbox-3 with custom written Matlab (Mathworks) scripts. First, we performed hand-mapping experiments to confirm that the neurons within the FOV had receptive fields on the screen. Specifically, we displayed a 0.04cpd square wave grating in 30-degree diameter patches in the screen. The location that evoked the strongest response (at least 40% of neurons with $p < 0.001$ Students t-test over baseline) was assumed the location of the receptive field of that FOV. Where possible, the screen was adjusted to center the receptive field relative to the screen. However, FOVs with receptive fields too close to the edges of the screen after adjustment were discarded.

In experiments where we measured the reliability of inhibitory interneurons (**Figure 4-2**), the visual stimulus ensemble comprised 5 natural movies repeated 80 times. Movies were selected from the Van Hateren natural movie database (Rikhye and Sur, 2015). Each movie was presented for 4s and were interleaved with 4s gray screen (mean luminance = 128). Each movie frame was adjusted to have a mean luminance of 128 and a mean contrast of 32 using the SHINE toolbox (Willenbockel et al., 2010). This stimulus ensemble lasted a total of 3200s (53.3 mins.).

Before commencing photostimulation experiments, we used the same stimulus ensemble, but with 20 repeats of each movie (total duration: 800s), to determine the two movies that evoked the most reliable responses. We selected these two movies as the "pulsed" and "non-pulsed" movies respectively for the random stimulation protocol. Each

movie was also presented for 4s with a 4s interleaved blank screen. Thus, the total stimulus duration was 4000s (66.6 mins.).

Data Analysis

Visually responsive neurons and deconvolution

All data analysis was performed using custom written scripts in Matlab. Significantly visually responsive cells were determined from the fluorescence time changes (ΔF/F) by performing a one-tailed Student's t-test between visually evoked (4s movie on) and spontaneous responses (4s gray screen in between movies). To obtain a better estimate of the spontaneous activity we also collected 120s of activity to an isoluminance gray screen before the start of the experiment. Only cells with $p < 0.001$ were classified as visually-responsive. Firing rates of these cells were then inferred using an optimized deconvolution algorithm for GCaMP6f (Theis et al., 2016). This algorithm inferred the probability of spiking from calcium transients. To convert this probability into a firing rate (measured in events/s), we multiplied each probability by 20Hz, the frequency at which the calcium transients were sampled. Unless otherwise stated, all data analysis was performed using inferred firing rates.

Change in firing rate and change in reliability following photostimulation

In all photostimulation experiments, analysis was restricted to 500ms time window following laser activation (Figure 4-S2). Within this time window, we determined the change in firing rate using the following formula

$$\Delta Rate(p) = \frac{Rate_{Laser\ On}(p) - Rate_{Laser\ off}(p)}{Rate_{Laser\ off}(p)}$$

where p is the pulse number. Since each neuron responded at different time points of the movie, averaging this across population of neurons would obscure any changes in the firing rate (or reliability). Thus, we aligned the firing rate trace of each neuron obtained on the Laser Off trials (control condition) to have a maximum rate at 1s into the movie. This index was then used to align the firing rate traces on the laser on trials.

These aligned traces were then used to calculate $\Delta Rate(p)$ according to the equation above.

Response reliability to natural movies was calculated using the equation:

$$Rel = \frac{2}{T^2 - T}\sum_{i=1}^{T}\sum_{j=i+1}^{T} Corr(f_{i,A}, f_{j,A})$$

where $i \in [1, T]$ indexes the trial number and $f_{i,A}$ is the PSTH on the i^{th} trial for movie A. Thus, from this equation the response reliability is the average correlation of all pairwise combinations of trials for a single movie (Rikhye and Sur, 2015). In photo-stimulation equations, the change in reliability induced by laser activation was calculated using the formula:

$$\Delta Rel(p) = \frac{Rel_{Laser\ On}(p) - Rel_{Laser\ off}(p)}{Rel_{Laser\ off}(p)}$$

Similar to the firing rate, we aligned the reliability on the Laser Off trials such that each neuron was maximally reliable at 1s. The same time index was then used to align reliability on the Laser On trials. To determine the epoch of maximum and minimum reliability, found the time index corresponding the maximum and minimum reliability on the laser off trials. Note, for this we used unaligned traces.

Signal-to-Noise ratio

The signal-to-noise ratio in both Laser On and Laser Off conditions was computed using the following equation

$$SNR = \frac{Var_i(E_t[f_i])}{Var_i(f_i)}$$

where $Var_i(f_i)$ is the variance of the firing rate across trials and $Var_i(E_t[f_i])$ is the across trial variance of the mean firing rate. Since $Var_i(E_t[f_i])$ is the stimulus-dependent variance and $Var_i(f_i)$ is the total variance of the neuron, SNR provides a measure of the fraction of the variance of the response that cannot be explained by the stimulus (i.e. noise). When all the variance of the response is explainable by the stimulus then SNR will be 1. Thus, the lower the SNR, the higher will be the fraction of the unexplainable variance.

Signal and Noise Correlation

Signal correlations and noise correlations between pairs of neurons were calculated using the following equation (Singh and Lesica, 2010)

$$CC\left(r_{ni}(t), r_{nj}(t)\right) = \frac{E\left[r_{ni}(t)r_{nj}(t)\right] - E\left[r_{nj}(t)\right]E\left[r_{ni}(t)\right]}{\sqrt{Var\left(r_{nj}(t)\right)}\sqrt{Var\left(r_{ni}(t)\right)}}$$

In the case of signal correlations (SC), $r_{nj}(t)$ is the trial averaged response of neuron nj binned at 250ms, whereas in the case of noise correlations (NC) we first subtracted the trial average response from each trial such that $\widehat{r_{nj}(t)} = r_{nj}(t) - E\left[r_{nj}(t)\right]$. Change in correlations was computed relative to Light Off trials using the same equations as reliability or firing rate.

Statistical Analysis.

All statistical analysis was performed using custom written scripts in MATLAB. No tests were conducted to determine sample size. Data were first tested for normality using the Shapiro-Wilk Test. All data presented in this paper are non-normally distributed, thus all statistical tests were conducted using non-parametric statistics. Our experiments involved testing the influence of laser activation on the same population of neurons, thus all comparisons were performed using non-parametric repeated-measures ANOVA (Friedman Test) with Bonferroni's correction for multiple comparisons. For Bonferroni corrections, the significance value was set to 0.05. Post-hoc tests were performed using the two-tailed Wilcoxon rank-sum test relative to the Light Off condition. To determine if the change in reliability was significant, we performed Permutation test (corrected for family-wise error rate) where we resampled with replacement (10,000 permutations) from the change distribution and tested if the sampled distribution was significantly different from 0 using a one-tailed test. In most figures, data are presented as median±95% CI, unless otherwise stated. All p-values are labeled in the figure legends.

Author Contributions

Rajeev Rikhye and Mriganka Sur designed the study. Rajeev Rikhye performed all the experiments and analyzed all the data. Rachael Neve designed the FRT-GCaMP6f and FRT-ChR2-mCherry constructs. Murat Yildrim designed and implemented the dual color imaging system. Rajeev Rikhye wrote the chapter, with input from Mriganka Sur. The manuscript will be a reformatted version of this chapter.

Chapter 5

What happens when interneurons go awry[1]?

Summary

Rett syndrome is a severe neurodevelopmental disorder caused by a loss-of-function mutation of the X-linked gene *MECP2*. Several studies have now established that many of the deficits associated with Rett Syndrome are the consequence of improper synaptic function, specifically caused by a dysfunction of cortical interneurons. Although it is accepted that inhibitory interneurons play an important role in sculpting neural computations, very little is known about how their dysfunction contribute to the pathophysiology of Rett Syndrome. The aim of our study was to characterize the effect of deleting MeCP2 from PV and SST neurons on visual processing. Deletion of MeCP2 from both cell types reduced their response magnitudes and orientation selectivity. PV-specific deletion was able to recapitulate many of the effects seen in MeCP2-null mice, by reducing the response magnitudes and increasing trial-to-trial variability of pyramidal neurons to both gratings and natural movies. In contrast, deletion of MeCP2 from SST neurons only altered the correlation structure of the pyramidal neuron network. Additionally, we found that administration of recombinant IGF-1 was able to ameliorate many of these coding defects in mice with PV-specific MeCP2 deletion. Thus, loss of MeCP2 from specific interneuron subtypes contributes crucially to the circuit-wide deficits of Rett Syndrome, suggesting that these neurons have a pivotal role in the functional deficits that characterize the disorder.

Highlights

- Deletion of MeCP2 from the forebrain reduces excitatory and inhibitory synaptic conductances in pyramidal neurons.
- MeCP2 deletion from PV and SST neurons alters the response properties of these neurons and of target pyramidal neurons.
- PV-specific deletion impairs the ability of pyramidal neurons to reliably and efficiently process natural scenes, whereas SST-specific deletion only changes neuronal correlations.
- PV specific loss of MeCP2 is sufficient to recapitulate effects of global loss of MeCP2, and these deficits can be reversed with IGF-1.

[1] The findings of this chapter are currently under review in PNAS. I would like to thank Drs. Abhishek Banerjee and Vincent Breton-Provencher for contributing crucial data for this chapter.

5.1 Introduction

In the previous chapter, I demonstrated that interneurons, namely PV and SST-INs, serve critical roles in sculpting the temporal precision of pyramidal cell responses to natural scenes. Importantly, this work, together with many others in the field (see **Chapter 2**), highlights the importance of intact inhibitory microcircuits in cortical information processing. What happens to neural computations when these interneurons go awry, for example during neurodevelopmental or psychiatric diseases? Numerous disorders are characterized by a disruption of the tightly coupled balance between cortical excitation and inhibition (Rubenstein and Merzenich, 2003; Sahin and Sur, 2015). Hence, a crucial first step towards understanding the pathophysiology of these diseases is elucidating the effect of compromising cortical inhibition. In this Chapter, we investigate the effect of specifically deleting the gene *MECP2* from cortical interneurons, which has been implicated in Rett Syndrome, on visual processing.

Autism spectrum disorders (ASDs) are typically caused by loss-of-function mutations of numerous genes, which in turn leads to severe developmental defects in brain connectivity (Zoghbi and Bear, 2012). However, a small fraction of ASDs is monogenic in nature, such as Rett Syndrome. These monogenic disorders are advantageous as they allow investigators to study the effect of single gene mutations on neural network dynamics, and allow for a better dissection of the circuitry underlying the pathophysiology of these diseases. For this reason, we have chosen to study Rett Syndrome, which is caused by a sporadic loss-of-function mutation of the gene *MECP2* located on the X-chromosome (Chahrour and Zoghbi, 2007). This gene encodes the methyl-CpG binding protein 2 (MeCP2), which is a critical regulator of brain development and adult neural function (Guy et al., 2011; McGraw et al., 2011). Since MECP2 is an X-linked gene, Rett Syndrome primarily affects girls. Affected girls develop normally during the first year, but quickly regress losing language and acquired motor skills. Notably, patients with Rett Syndrome demonstrate ataxia, have respiratory dysrhythmias, seizures and severe cognitive defects (Chahrour and Zoghbi, 2007).

Mouse models, with conditional knock-out of *MECP2* from either pyramidal (McGraw et al., 2011) or inhibitory interneurons (Chao et al., 2010) seem to recapitulate many of these cognitive and behavioral deficits. These deficits are primarily due to changes in synaptic function and cell morphology (Banerjee et al., 2012; Jiang et al., 2013).

While there are no apparent deficits in sensory processing in Rett Syndrome patients, it is possible that improper coding of sensory information could exacerbate many of the observed behavioral defects. Indeed, fMRI studies have demonstrated unreliable processing of visual stimuli, which is characterized by weak and noisy responses, in human patients with ASDs (Dinstein et al., 2012). In mouse models, several studies have noted change in visual acuity (Durand et al., 2012) and abnormal experience-dependent plasticity in the visual cortex (He et al., 2014). Additionally, a recent study identified critical deficits in auditory evoked potentials (ERPs) in the MeCP2-null mice (Goffin et al., 2014). Specifically, deleting MeCP2 from GABAergic neurons decreased ERP power at within all frequency bands, and reduced the reliability of evoking ERPs. Together, these results, along with several others (Chao et al., 2010; He et al., 2014) suggest that deleting MeCP2 in cortical neurons, especially INs, has a strong deleterious effect on neuronal computations.

It is well appreciated that at the heart of all cortical computations lie microcircuits of highly interconnected excitatory and inhibitory interneurons. Three major classes of INs mediate inhibition in the cortex: parvalbumin-expressing (PV), somatostain-expressing (SST) and vasoactive intestinal peptide expressing (VIP) (Tasic et al., 2016). Each IN subtype target different domains on pyramidal neurons, and consequentially performs distinct computations (Kubota, 2014; Womelsdorf et al., 2014). For example, PV neurons provide strong divisive inhibition onto the somatic compartment of pyramidal neurons, and play a role in gain normalization (Wilson et al., 2012). In contrast, SST neurons inhibit the distal dendrites of pyramidal neurons, and thus gate the integration of synaptic inputs ((Lovett-Barron et al., 2012), see also **Chapters 2** and **4**). These IN subtypes also express MeCP2 and a recent paper found that deleting MeCP2 from PV- or SST-INs results in distinct behavioral deficits. Specifically, mice lacking MeCP2 in PV-INs developed severe cognitive and motor deficits, reminiscent of

Rett Syndrome, while mice lacking MeCP2 in SST-INs developed seizures (Ito-Ishida et al., 2015). Thus, the pathophysiology of Rett syndrome is modular – loss of MeCP2 from different neuron subtypes lead to distinct deficits. However, it remains unexplored how the response properties of these INs or the circuit they are embedded in changes in Rett Syndrome.

In this study, we investigated the effect of deleting MeCP2 from INs on neuronal processing in mouse visual cortex. Using whole-cell electrophysiology, we noticed that V1 neurons in MeCP2-null male mice had weaker, and unbalanced, excitatory and inhibitory currents. This implied that MeCP2 deletion alters the magnitude of synaptic drive onto pyramidal neurons. To elucidate the contribution of different IN subtypes to this weakened drive, we specifically knocked-out MeCP2 from PV- and SST-INs. Deleting MeCP2 in both cell types resulted in weaker response magnitudes and orientation selectivity. Next, we assayed the effect that this IN-specific MeCP2 deletion had on visual processing using high-speed calcium imaging using both sinusoidal gratings and natural movies. Knocking-out MeCP2 from PV neurons (*PV-cKO*) had the most deleterious effect, resulting in weaker, less reliable responses to both grating and movies, and weaker orientation selectivity. At the network level, *PV-cKO* mice had strongly decorrelated responses. These changes were comparable to deleting MeCP2 from all neurons (*MeCP2-gKO*), suggesting that perturbing the normal function of PV neurons strongly influences sensory processing. Interestingly, deleting MeCP2 from SST neurons (*SST-cKO*) had slightly different effects. In contrast to *PV-cKO*, *SST-cKO* mice did not have weaker, less reliable or unturned responses. However, neurons in *SST-cKO* mice were more strongly correlated, and thus were unselective to different movies. This highlights the role of SST neurons in organizing the spatial activity of pyramidal neurons in V1. Finally, we found that administering IGF1, a drug known to ameliorate many of the synaptic and behavioral deficits in Rett Syndrome (Castro et al., 2014), was also able to rescue many of these coding deficits in *PV-cKO* mice. Together, our findings indicate that a reduction of inhibition, primarily through PV neurons, underlies many of the cortical circuit deficits of Rett Syndrome.

5.2 Predictions and Hypothesis

Cortical interneurons play a critical role in sculpting neural computations. Dysfunctions in interneurons have been implicated in several psychiatric and neurodevelopmental disorders (Cohen et al., 2015; Lewis, 2012). Often, these dysfunctions manifest as reduced inhibition, which in turn results in a circuit-wide disruption of the balance between excitatory and inhibitory currents (Benes and Berretta, 2001). Due to technical limitations, very few studies have characterized how dysfunctions in inhibitory interneurons alter cortical computations, such as the processing of complex visual stimuli. In this chapter, we have used two-photon calcium imaging along with cell-specific deletion of MeCP2 to answer this question. First, we used the DREADD (designer receptors exclusively activated by designer drugs, (Urban and Roth, 2015)) to either chronically increase or decrease inhibition from PV or SST neurons. This method allowed us to gain insight into how chronically, but reversibly, perturbing inhibition alters visual processing in V1.

Chronically perturbing inhibition reduces reliability to natural scenes

We first sought to determine the effect of chronically perturbing inhibition from PV and SST neurons on pyramidal cell reliability. To do so, we used a *chemo-genetic* approach to either chronically silence (hM_4Di, Gi-DREADD) or excite (hM_3Dq, Gq-DREADD) these interneurons following administration of the drug Clozapine-N-Oxide (CNO, **Figure 5-1A**). We used a fast, genetically encoded calcium indicator (GECI) GCaMP6f and high-speed two-photon imaging to monitor pyramidal neuron response in layer 2/3 of V1 in awake, head-fixed mice with high spatial and temporal resolution. Specifically, we scanned a 250x250μm patch of cortex at 20 frames/second. Our stimulus set comprised 5 different natural movies, each 4s in duration and each repeated 40 times. The order of the movies was randomized to prevent adaptation. We used a temporal deconvolution algorithm to infer firing rates from the observed calcium transients. For further experimental details see Appendix 5A.

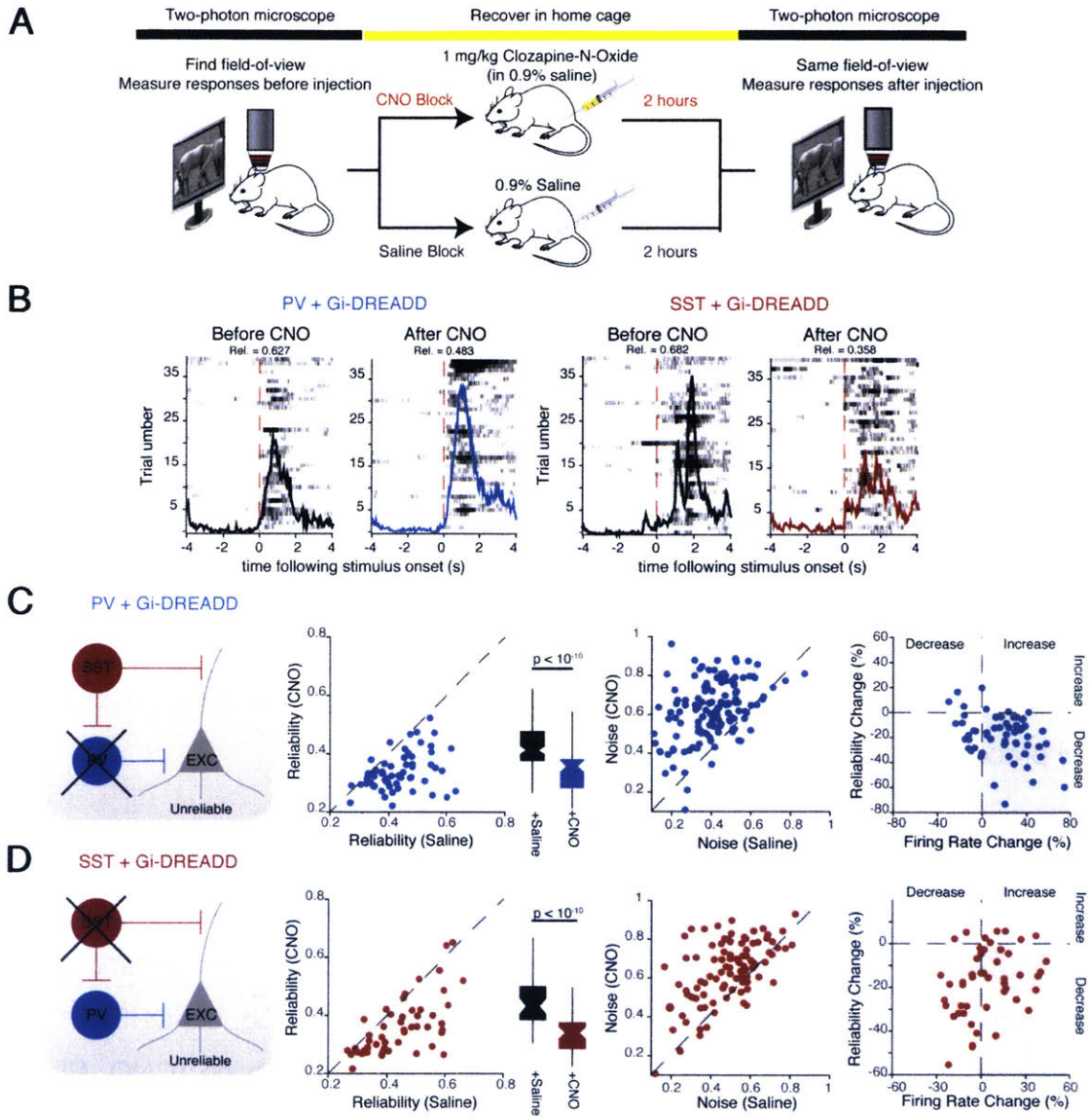

Figure 5-1. Intact inhibition is required for reliable coding. (A) Schematic illustrating experimental protocol. **(B)** Example raster plots from two pyramidal neurons showing the effects of inactivating PV neurons (left) and SST neurons (right) on movie-evoked response. Solid lines show trial-averaged firing rate. A clear increase in non-stimulus evoked spikes can be seen in both conditions. (C-D) Quantification of change in reliability, noise and firing rate when PV (C) and SST (D) neurons are silenced with Gi-DREADD. Data obtained from 3 mice each. P-values computed using Wilcoxon signed-rank test relative to saline controls.

A schematic of our experimental set-up is shown in **Figure 5-1A**. Briefly, we imaged the same population of neurons both before and 2-3 hours after administering either CNO or saline (control). A previous study (Kuhlman et al., 2013) determined that

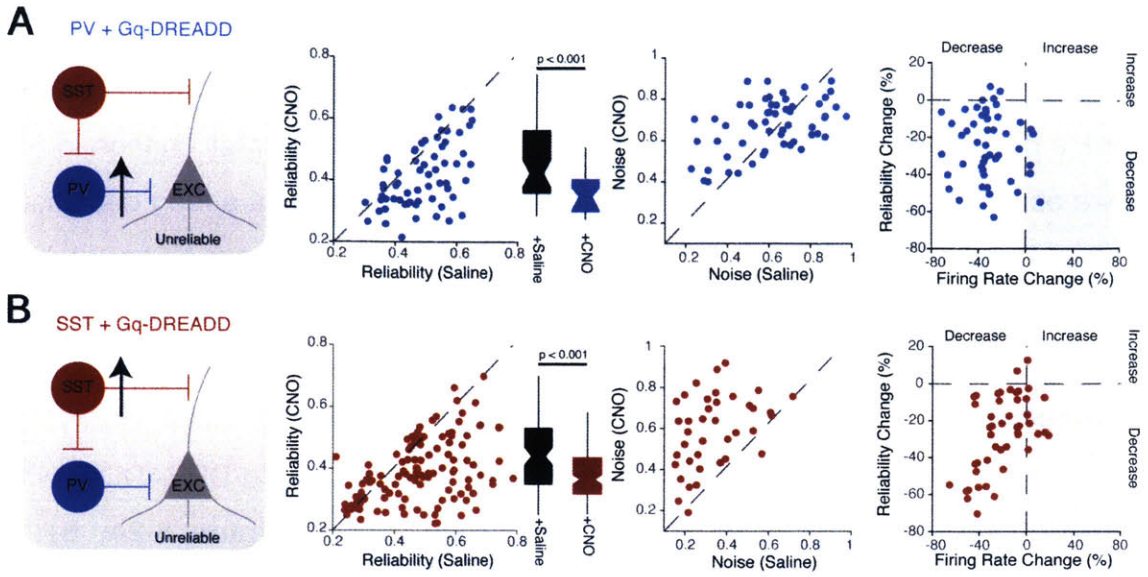

Figure 5-2. Chronically increasing the activity of interneurons decreases pyramidal neuron reliability. Exciting PV **(A)** and SST **(B)** neurons with Gq-DREADD receptors decreases reliability, increase noise and increases firing rate (compared to saline controls). Data from 3 mice each.

a 2-3 hour window was sufficient for CNO to bind to DREADD receptors. To minimize stress, we allowed mice to recover in their home cage during this interval. As a control, we injected saline instead of CNO, and compared responses from the same population of cells. We computed response reliability as the average Pearson's correlation coefficient between every pairwise combination of trials (see Appendix 5A).

Silencing both PV and SST neurons, with Gi-DREADDs, significantly reduced response reliability within L2/3 pyramidal neurons, but had different effects on the trial-averaged firing rate. Specifically, suppressing PV inhibition caused a large increase in pyramidal neuron response rate, which was primarily due to an increase in firing during non-stimulus-evoked epochs (see example in **Figure 5-1B**). To better quantify this, we measured noise in the response, which we defined as the fraction of the trial-to-trial variance not explained by the stimulus. Based on this metric, a noisy response would have large non-stimulus evoked activity. As shown in **Figure 5-1C**, silencing PV neurons significantly increased noise, which in turn led to a decrease in trial-to-trial reliability.

In contrast, silencing SST neurons only weakly changed the firing rate of pyramidal neurons (**Figure 5-1D**). We suspect this was due to the fact that SST neurons inhibit PV neurons in addition to pyramidal neurons. Despite this small change in firing rate, we observed a strong increase in noise, and consequentially a strong decrease in reliability. Thus, whereas silencing PV increased both signal (average firing rate) and noise in their target neurons, silencing SST primarily increased noise. This also reaffirms the central role of PV neurons in providing feed-forward inhibition.

What effect does non-specifically increasing inhibition have on reliability? We answered this question by repeating these experiments with Gq-DREADDs, which allowed us to chronically increase interneuron activity instead (**Figure 5-2A, B**). We were surprised to find that exciting both PV and SST neurons also decreased pyramidal neuron reliability and increased noise. As expected, we found that increasing inhibition from both PV and SST decreased pyramidal cell firing rate. The fact that hyperactive SST neurons decreased firing rates suggests either very strong dendritic inhibition or some compensatory decrease in PV inhibition as well.

Taken together, these results suggest that the temporal precision of spiking activity in V1 depends critically on a set inhibition value. Perturbing inhibition, either by increasing or decreasing the activity of interneurons, chronically prevents pyramidal neurons from responding reliably to natural scenes.

Chronically perturbing inhibition alters correlations between neurons

Our results so far show that perturbing cortical inhibition, by either non-specifically increasing or decreasing PV and SST neuron activity, strongly increases trial-to-trial variability (noise). Several theoretical and experimental studies have now established that inter-trial variability strongly influences the ability of the cortex to encode stimuli (Abbott and Dayan, 1999; Averbeck et al., 2006). In particular, correlated noise between neurons is believed to limit the capacity of neuronal networks to store information (Moreno-Bote et al., 2014; Zohary et al., 1994). Thus, we next sought to understand how chronically perturbing cortical inhibition influenced the magnitude of correlations between pyramidal cells.

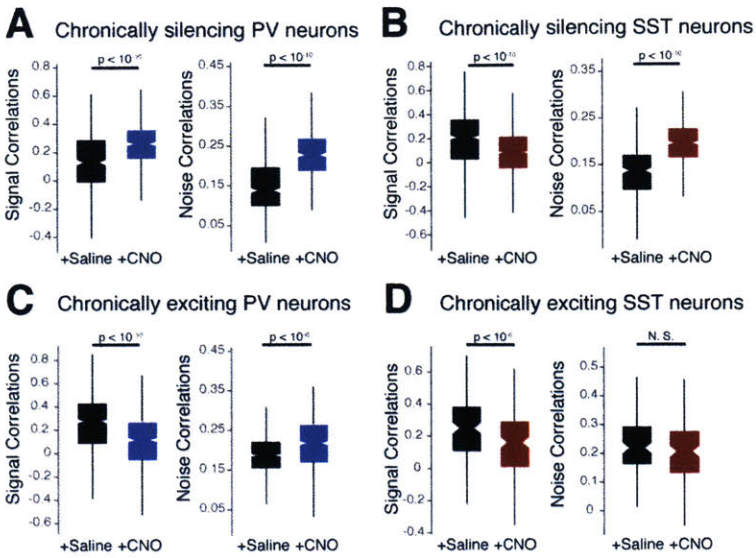

Figure 5-3. Chronically perturbing interneurons alters both signal and noise correlations between neurons. **(A-B)** Box-whisker plots quantifying the effect that silencing PV **(A)** and SST **(B)** neurons with Gi-DREADDs has on signal and noise correlations. **(C-D)** Analysis of correlations when PV **(C)** and SST **(D)** neurons are chronically silenced with Gq-DREADDs instead. Data from 3 mice each. P-values computed using Friedman's test with Bonferroni's correction for multiple comparisons.

We first computed signal correlations between neurons. In simple terms, signal correlations (SC) provide a measure of similarity between the trial-averaged responses of a pair of neurons. Neurons with high SC tend to respond together, and which could arise from shared receptive fields or common sources of inputs. We discovered that silencing PV inhibition with Gi-DREADD led to a significant increase in SC between neurons (**Figure 5-3A**). This stems from the fact that silencing PV inhibition non-specifically increases pyramidal neuron firing rates (**Figure 5-1C**), which would in turn result in strong stimulus-locked activity. Correspondingly, we found that increasing PV inhibition via Gq-DREADD reduced SC between neurons (**Figure 5-3C**). This is again consistent with the reduced pyramidal activity caused by increased PV inhibition. In stark contrast, we discovered that bi-directionally perturbing SST neurons lowered SC (**Figure 5-3B, D**). SST neurons are known to participate in recurrent processing in cortical Layer 2/3, resulting in surround suppression (Adesnik et al., 2012). In doing so, it is likely that SST neurons coordinate the activation of ensembles of cortical neurons (Stefanelli et al., 2016). Thus, despite increasing the firing rates of pyramidal neurons, exciting or silencing SST neurons can decorrelate the network by altering the activation of ensembles (Chen et al., 2015).

We next computed noise correlations (NC), which provides a measure of how much trial-to-trial variability co-fluctuates between a pair of neurons. By definition, noise

correlations capture dependencies between neurons that are not locked to the external stimulus, such as shared connections (Singh and Lesica, 2010). Expectedly, we found that silencing both PV and SST neurons increased NC between pyramidal neurons (**Figure 5-3A, B**). This is again because silencing interneurons increases both firing rates and trial-to-trial variability. Similarly, exciting PV neurons also increased NC (**Figure 5-3C**), suggesting that chronically perturbing PV neurons increases correlated noise between neurons due to an increase in inter-trial variability. We were, however, surprised to find that chronically exciting SST neurons did not significantly alter NC (**Figure 5-3D**), despite strongly increasing trial-to-trial variability. A parsimonious explanation is that the increased SST activity does not change the magnitude of NC, but could change the structure of NC between pyramidal neuron ensembles.

In sum, our correlation analysis shows that perturbing PV neurons not only increases trial-to-trial variability but also increases correlated variability between neurons. In contrast, perturbing SST neurons tends to decorrelate pyramidal neurons. Overall, our results show that silencing cortical inhibition has a more deleterious effect than chronically increasing cortical inhibition. We hypothesize that this is because several compensatory mechanisms, such as increase in feed-forward drive, can compensate for increased inhibition to preserve cortical E-I balance.

Chronically perturbing inhibition alters the ability to discriminate between natural scenes

We next sought to determine how this chronic perturbation of inhibition influenced the ability of the cortex to discriminate between different stimuli. To this end, we computed the Mahalanobis distance between neural population vectors. Briefly, we formed population vectors using the trial averaged activity of neurons, binned at 500ms. Assuming that we recorded from *N* neurons, then this population vector would span a *N*-dimensional subspace, as depicted in **Figure 5-4A**. Since not all dimensions are informative, we first performed principle components analysis to reduce the data set to the three most informative dimensions. On average, these dimensions were able to explain approximately 70% of the variance in the data (see Appendix 5A for details).

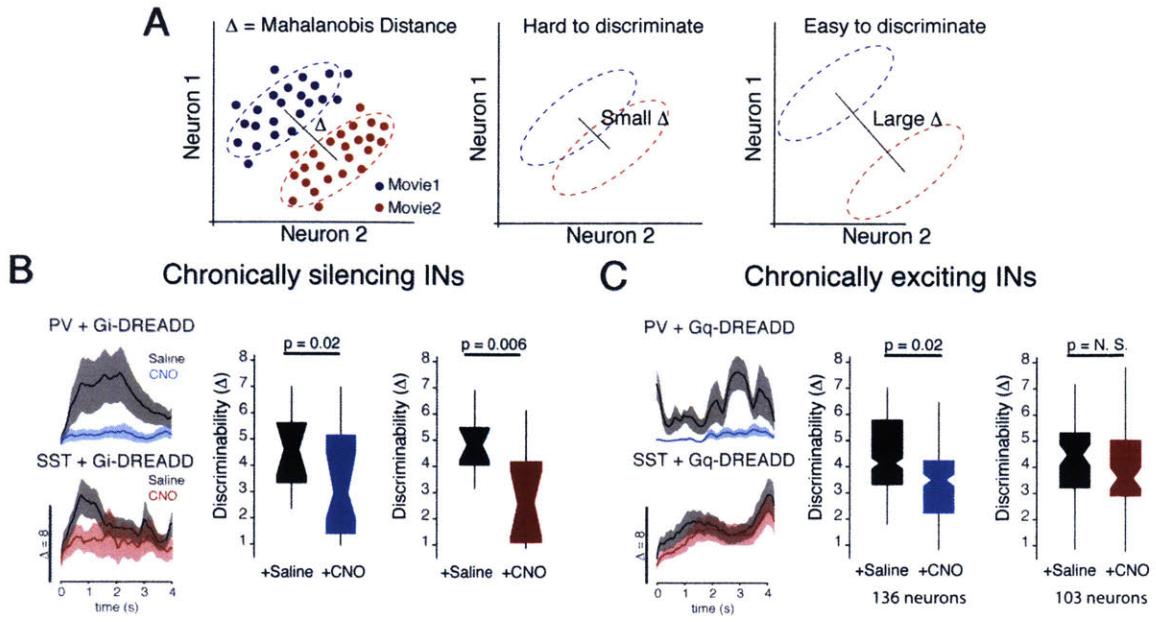

Figure 5-4. Chronically perturbing inhibition disrupts the ability of the cortex to discriminate between different movies. (A) Schematic illustrating the concept of a Mahalanobis distance. Ellipses denote the 95% confidence interval that contains the population vector. For simplification, only a 2-dimensional population vector is shown. **(B-C)** Quantification of the effect that either chronically silencing **(B)** or exciting **(C)** inhibition has on the Mahalanobis distance (or Disciminability) between different natural movies. Time series plots show Mahalobis distance within each time bin following stimulus onset. Shaded area indicates SEM. Data collected from 3 mice in each condition. P-values computed using Friedman's test with Bonferroni correction.

Next, we computed the Mahalanobis distance between these principle components (PCs) for each movie combination (total 10 movie combinations). The Mahalanobis distance measures the degree of overlap between two representations by taking into account both the distance between two distributions and the scatter within each distribution. A large Mahalanobis distance indicates small overlap between two population vectors, suggesting independent coding of both movies. In contrast, a small Mahalanobis distance indicates a large overlap and non-independent coding. Ideally, the cortex would maximize the Mahalanobis distance between its population vectors for different stimuli in order to maximize the number of stimuli it can encode.

We discovered that chronically silencing both PV and SST-INs with Gi-DREADDs significantly reduced the Mahalanobis distance, and consequentially the discriminability between different movies (**Figure 5-4B**). We noticed that the Mahalanobis distance

between population vectors varied dynamically over time following stimulus onset. This suggests that movie information is dynamically encoded through the duration of the movie. This dynamics was, however, lost when PV and SST-INs were silenced. We were surprised to find that silencing SST neurons had a stronger effect on stimulus discriminability than when we silenced PV neurons. This speaks to the importance of SST neurons in regulating the temporal precision of spiking and in coordinating interneuronal correlations (see **Chapter 4**).

Similarly, chronically increasing PV inhibition using Gq-DREADDs also reduced discriminability (**Figure 5-4C**). Surprisingly, increasing SST inhibition did not significantly change discriminability, even though this perturbation had comparable effects on reliability and noise. A parsimonious explanation is that PV inhibition is very finely tuned – small chronic perturbations, which increase or decrease PV activity, pushes the visual cortex into a pathological regime. In contrast, SST inhibition is much more dynamic and robust to perturbations

Summary

Together, these results demonstrate that perturbing cortical inhibition; by non-specifically either increasing or decreasing PV and SST neuron activity, strongly decreases the temporal precision of pyramidal neurons, and consequentially, reduces the ability of the cortex to encode different natural movies. We believe that this effect is due to a disruption of the E-I balance in the visual cortex. Specifically, without intact inhibition, neurons also fire more spikes during the stimulus, increasing noise. This also means that individual spikes convey less information about the stimulus, which is not metabolically efficient (Vinje and Gallant, 2002). Thus, the precision of spike trains in V1 is very sensitive to small perturbations in inhibition.

Importantly, our DREADD experiments show that the cortex is poised to reliably and efficiently process natural scenes. In particular cortical inhibition is finely tuned and diseases, which disrupt this balance, will likely have a very detrimental effect on stimulus encoding. We found that PV neurons, which play an important role in modulating the gain of pyramidal neurons, are most sensitive to perturbations. Our

results support the notion that PV neurons are critical for maintaining the E-I balance in the cortex (Xue et al., 2014). In particular, given their dense connectivity with neighboring pyramidal neurons (Fino and Yuste, 2011) and strong thalamic inputs (Wall et al., 2016), PV neurons function to balance feed-forward excitation and prevent neural activity from saturating. In agreement with this notion, PV inhibition is chronically reduced in Schizophrenia and other neuropsychiatric diseases, which are characterized by altered cortical oscillations (Lewis, 2012; Rubenstein and Merzenich, 2003). Thus, based on these results, we hypothesize that deleting MeCP2 from PV neurons should have detrimental effects on neural coding in V1.

Similarly, SST neurons are also susceptible to perturbations. In particular, inhibiting SST neurons is much more detrimental to network computations than chronically exciting them. Interestingly, a recent study showed that Layer-5 SST neurons are hyperactive in mouse models of Amyotrophic Lateral Sclerosis (ALS) and Dementia (Zhang et al., 2016). This study claimed that hyperactive SST neurons lead to excito-toxic effects on M1 pyramidal neurons due the SST→PV disinhibitory circuit. Our results agree with this finding because, both activating SST and inhibiting PV neurons caused reduced reliability. Thus, hyperactive or hypoactive SST neurons prevent pyramidal neurons from reliably conveying information. We were surprised to find that hyperactive SST neurons did not alter stimulus discriminability. This could be due to the fact chronically exciting SST neurons can lead to compensatory changes in PV and pyramidal activity. It is worth investigating the effect that hyper/hypo-active SST neurons have on PV neurons (see **Chapter 4**). Thus, based on these findings, we hypothesize that deleting MeCP2 from SST neurons would have a more subtle effect on neural coding. Specifically, we postulate that perturbing SST neurons would decrease reliability, would alter inter-neuronal correlations, but would not change neuronal selectivity.

5.3 Results

MeCP2-KO mice have altered E-I ratios in visual cortex[2]

Numerous studies have now shown that precisely-timed barrages of excitatory and inhibitory inputs into pyramidal neurons underlie spike-based computations in the cortex (Okun and Lampl, 2008; Wehr and Zador, 2003). In particular, inhibitory inputs are often temporally delayed with respect to excitation, creating a "window of opportunity" during which the neuron integrates these inputs and responds with a series of spikes (reviewed by (Haider and McCormick, 2009), see also **Chapter 4**). Additionally, in numerous cortical areas, inhibition is tightly correlated with excitation (Isaacson and Scanziani, 2011), giving rise to the notion of E-I balance. Not surprisingly, disrupting this balance strongly influences the ability of the cortex to perform several computations. Thus, we first sought to understand how this E-I balance was altered in male MeCP2-null mice (referred to henceforth as *MeCP2-gKO*).

To directly probe alterations in excitation and inhibition in pyramidal neurons after MeCP2 deletion, we used *in vivo* whole-cell patch-clamp recordings to measure both excitatory (Ge) and inhibitory (Gi) synaptic conductance from putative excitatory Layer 2/3 neurons in V1 of adult mice (postnatal age approximately 45days, **Figure 5-5A;** see Appendix 5B for experimental details). We measured post-synaptic currents, in voltage-clamp mode, in response to full-field drifting sinusoidal gratings. In order to isolate visually-evoked excitatory and inhibitory postsynaptic currents (EPSCs and IPSCs), we held neurons at the reversal potential for excitatory (-70 mV) and inhibitory currents (+20 mV) respectively and restricted our analysis to the first 500ms of stimulus presentation (**Figure 5-5B**). Expectedly, in wild-type mice, we found both excitatory and inhibitory synaptic conductance to be strongly correlated during in response to the visual stimulus (**Figure 5-5C**).

[2] Data in this section was collected and analyzed by Dr. Vincent Breton-Provencher

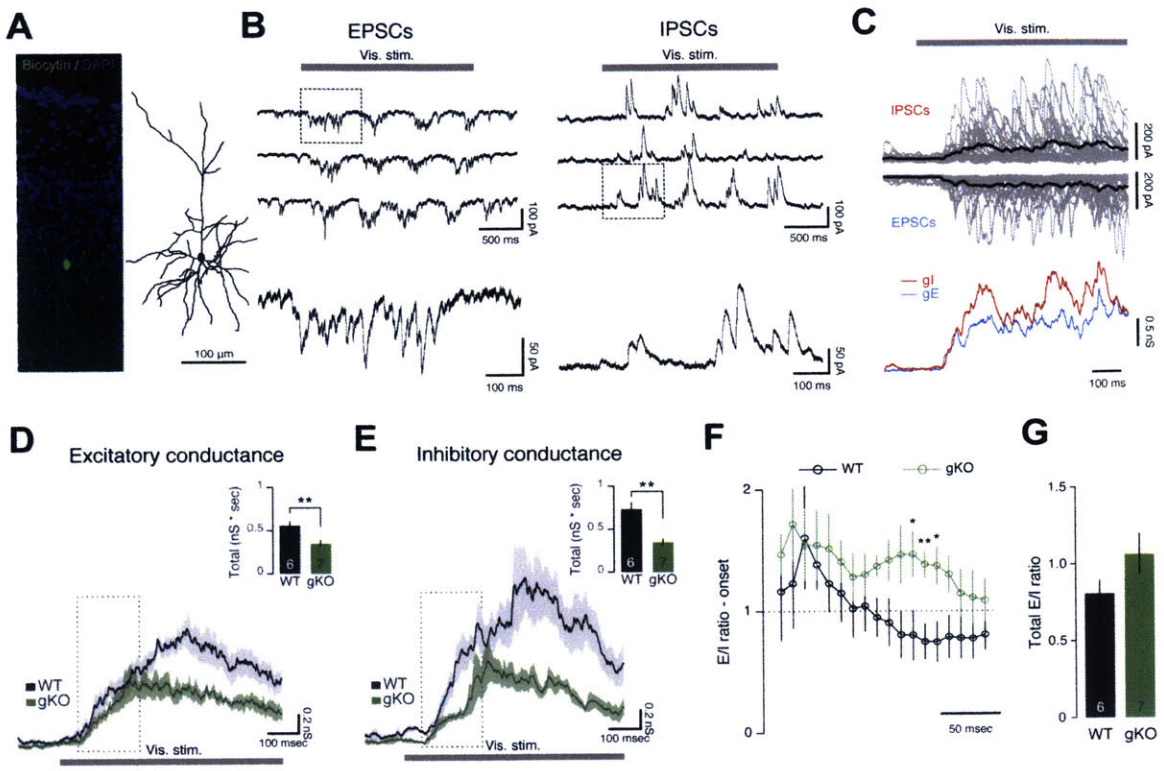

Figure 5-5. MeCP2 deletion leads to reduction of visually-evoked synaptic conductances and spike responses in pyramidal neurons. (A) Confocal image (*left*) and reconstruction (*right*) of a layer 2/3 pyramidal neuron in a WT mouse filled with biocytin during whole-cell *in vivo* patch-clamp recording in V1. **(B)** Example traces of visually-evoked excitatory and inhibitory post-synaptic currents (EPSCs and IPSCs), at a holding potential of -70 mV (*left*) and +20 mV (*right*). Drifting grating stimulation was presented at the neuron's preferred orientation at a temporal frequency of 2 Hz. The traces in the boxed areas are expanded at bottom. **(C)** (*Top*) Averaged postsynaptic currents in a WT neuron during the first 500 ms after the onset of visual stimuli. Gray traces represent single trials, black traces represent the average. (*Bottom*) Corresponding synaptic conductance derived from the voltage-clamp recordings. **(D)** Population averaged traces of the excitatory conductance for Wild-type (WT) (n = 6 neurons from 5 mice) and MeCP2 gKO mice (n = 7 neurons from 4 mice). (Inset) Mean ± SEM of the total excitatory conductance. **(E)** Population averaged traces, and (*inset*) mean ± SEM of the total inhibitory conductance. **(F)** Timing of E/I ratio after the onset of the visual response as highlighted by the boxes in **D** and **E**. Mean E/I ± SEM ratio for each time point was calculated over 10 ms intervals. **(G)** Summed E/I ratio (0-500 ms post-stimulus onset) for WT and gKO cells.

When compared to age-matched wild-type mice, we found that both Ge and Gi were significantly weaker in *MeCP2-gKO* mice (**Figure 5-5D, E**). Measurements of E-I ratio during the onset of the visual response revealed that, E-I balance was less dynamic in *MeCP2-gKO* mice. Specifically, inhibition remained at a significantly lower value than excitation throughout the stimulus onset epoch (**Figure 5-5F**). This is in stark contrast to WT mice, where inhibition surpassed excitation soon after the stimulus

onset. These results suggest that inhibitory currents dominate responses in WT mice, and these currents are strongly reduced when MeCP2 is deleted. Due to this lowered Gi, the total E-I ratio was marginally higher in *MeCP2-gKO* mice than in WT mice (**Figure 5-5G**).

Deleting MeCP2 from interneurons reduces their visual response properties[3]

Our results so far show that deleting MeCP2 from neurons alters both magnitude, and balance, of both excitatory and inhibitory currents. Thus, we next asked if this alteration influenced the ability of V1 pyramidal neurons to process visual stimuli. To do so, we preformed two-photon guided cell-attached recordings from visually identified neurons in binocular V1 in anesthetized mice, in response to full-field drifting gratings (see Appendix 5B).

We first made targeted recordings from pyramidal neurons in *CaMKII-cKO* mice, where MeCP2 was selectively deleted from CaMKII-expressing excitatory neurons. Compared to neurons in WT animals, neurons in *CaMKII-cKO* mice had broader orientation tuning curves (**Figure 5-6A**). In particular, we observed a significant reduction in peak response (i.e. response measured at the preferred orientation) and a similar reduction in the orientation selectivity index (OSI). We also observed similar reduction in both peak response and OSI in neurons in *MeCP2-gKO* mice (**Figure 5-6B, C**). These results indicate that deleting MeCP2 reduces the selectivity of pyramidal neurons to oriented edges, which is a primary computation performed by V1. These deficits could be due to the reduction in excitatory and inhibitory drive to these neurons, as we have shown through whole-cell recordings (**Figure 5-5**).

How are the visual response properties of different inhibitory interneuron (IN) subtypes altered by MeCP2 deletion? To answer this question, we deleted MeCP2 specifically from either PV or SST neurons using Cre/loxP recombination (to generate *PV-cKO*, and *SST-cKO* mice, see Appendix 5B for experimental details), and measured cell-specific responses with using cell-attached recordings, as described above. We

[3] Data in this section was collected by Dr. Abhishek Banerjee and Dr. Caroline Runyan, but was analyzed by me.

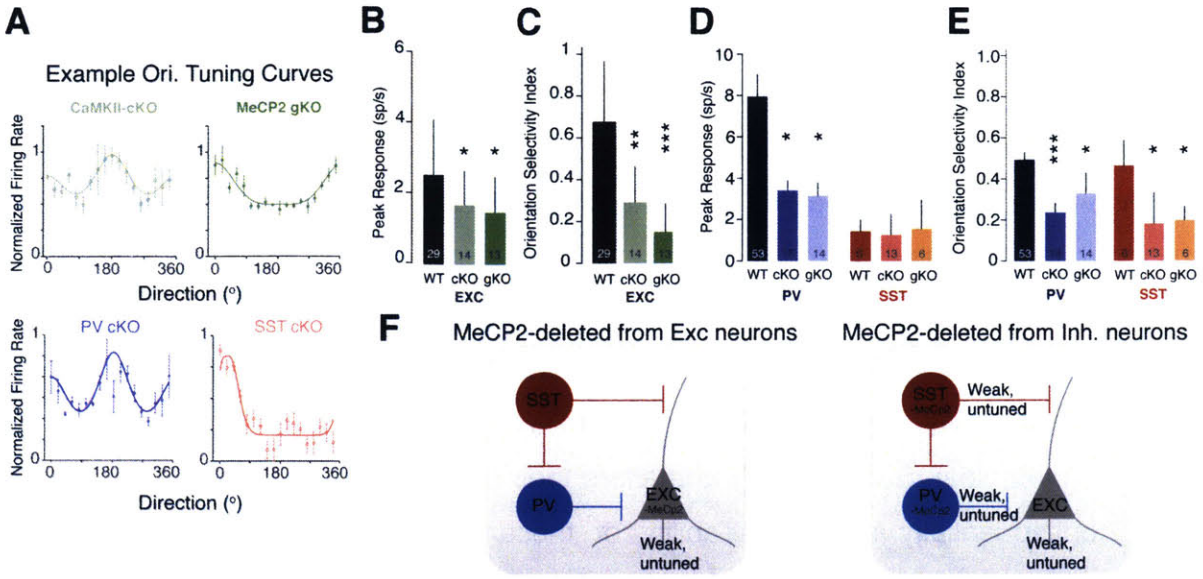

Figure 5-6. Deletion of MeCP2 alters response features and selectivity of interneuron subtypes. (A) Example orientation-selective responses of pyramidal neurons in *CaMKII-cKO* and *MeCP2-gKO* mice (top row), and PV and SST neurons in conditional knock-out mice (bottom row). (B-E) Quantification of peak response (spikes/s) at preferred orientation and orientation selectivity index from pyramidal neurons (EXC, **B-C**) and interneurons (**D-E**). Number of cells recorded is marked in the histogram bars. **(F)** Model summarizing the main results. All values are expressed as mean ± SEM. *p<0.05, ***p<0.0001, Wilcoxon rank-sum test.

specifically focused on PV and SST neurons because of their well-established role in sculpting cortical computations through feed-forward and feed-back inhibition (Isaacson and Scanziani, 2011). Interestingly, a recent study showed that several key behavioral deficits could be recapitulated by deleting MeCP2 specifically from PV or SST neurons alone (Ito-Ishida et al., 2015), underscoring the importance of these interneurons in the pathophysiology of Rett Syndrome.

Since the PV locus is turned on late during postnatal development (Gonchar et al., 2007), we found that MeCP2 deletion from PV neurons was only complete by postnatal day 60. Thus, we restricted our analysis to mice 60 days and older for both *PV-* and *SST-gKO* mice. We discovered that PV neurons lacking MeCP2 had weaker firing rates at preferred orientation and also had weaker orientation selectivity (**Figure 5-6D, E**). Notably, PV neurons, identified on the basis of their spike waveform, recorded in *MeCP2-gKO* mice, also responded poorly to drifting gratings. Similarly, SST neurons in *SST-gKO* mice also had significantly reduced firing rates at peak orientation (**Figure 5-**

6D, E). Interestingly, we found no change in the OSI of MeCP2-deficient SST neurons. This suggests that the receptive field properties of this neuronal subtype are not influenced by MeCP2 deletion. In contrast, deleting MeCP2 has a much stronger influence the response properties of PV neurons, which could stem from a reduction of excitatory drive to these neurons.

Taken together, our electrophysiological results reveal that deleting MeCP2 strongly alters the magnitude and timing of excitatory and inhibitory currents in cortical neurons. Additionally, deleting MeCP2 reduces visually evoked responses in all neuron subtypes (pyramidal, SST and PV) and alters their selectivity for oriented edges, which would consequentially affect the ability of these neurons to parse visual scenes (summarized in **Figure 5-6F**).

MeCP2 deletion from Interneurons leads to specific defects in visual processing

We were next interested in understanding how this change in interneuron responses influences V1 pyramidal neuron responses and the computations performed by these neurons. To study large populations of putative pyramidal L2/3 neurons *in vivo* at single cell resolution, we performed high-speed Ca^{2+} imaging after loading cells with the synthetic calcium indicator dye OGB1-AM (**Figure 5-7A**; see Appendix 5B for experimental details). As expected, pyramidal cells in both *PV-cKO* and *SST-cKO* mice responded to full-field sinusoidal drifting gratings with robust Ca^{2+} transients (**Figure 5-7B**). We inferred spiking rates from these Ca^{2+} responses using a temporal deconvolution algorithm (Vogelstein et al., 2010) and used these inferred firing rates to obtain orientation-tuning curves (**Figure 5-7C**). Interestingly, both the average firing rate (**Figure 5-7D**) and signal-to-noise ratio measured at the preferred orientation (**Figure 5-7E**), were reduced in pyramidal neurons in the *PV-cKO*, *SST-cKO* and *MeCP2-gKO* mice compared to WT and floxed-MeCP2 control mice. This suggests that deleting MeCP2 specifically from interneurons causes a reduction in visually evoked activity of pyramidal neurons and a commensurate increase in trial-to-trial variability.

As expected from previous studies in mouse V1, we found a wide range of OSI values in both WT and MeCP2-deficient mice (**Figure 5-7F**). Pyramidal cells in *PV-cKO*

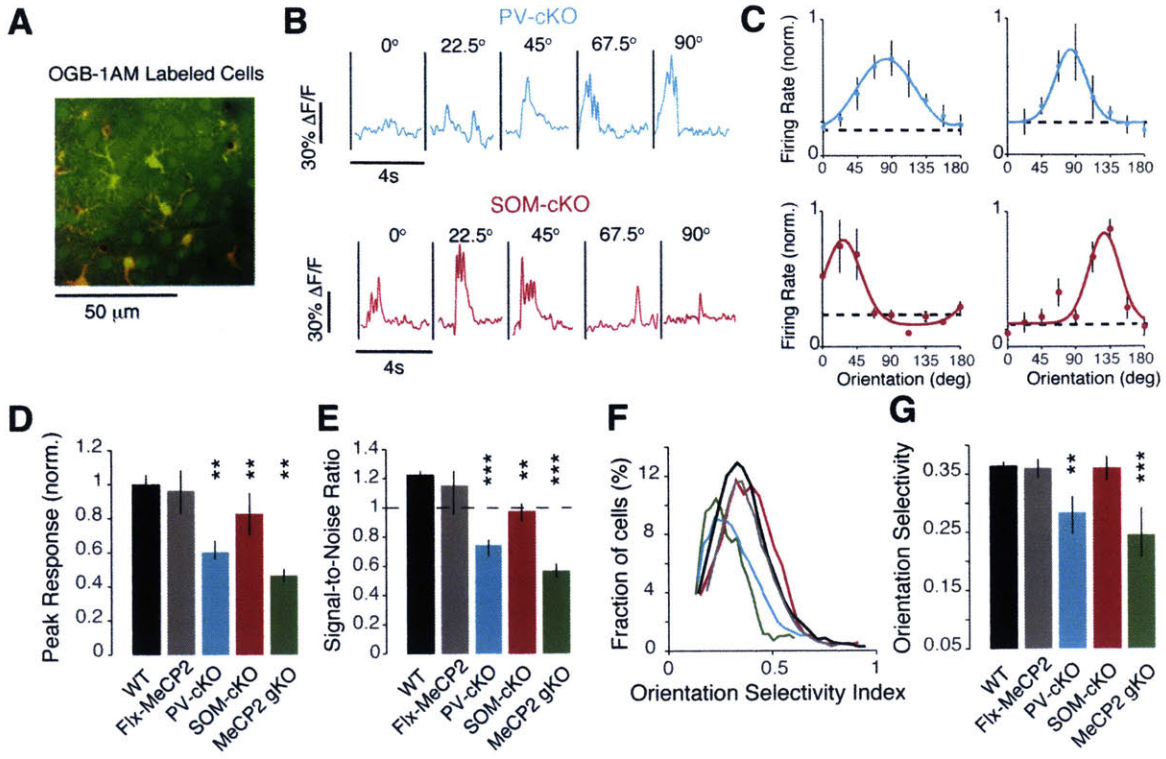

Figure 5-7. Deleting MeCP2 results in weak and unselective responses to sinusoidal gratings. (A) Example of an OGB1-AM labeled population of neurons. Neurons were identified by OGB1 expression (green) while astrocytes were identified by SR-101 expression (red). Astrocytes were discarded from further analysis. **(B-C)** Representative Ca^{2+} responses (**B**) and orientation tuning curves (**C**) from two neurons from PV-cKO (*top*) and SOM-cKO (*bottom*) mice. The Ca^{2+} responses correspond to the leftmost tuning curves in (**C**). Error bars represent SEM across 10 repeats of each orientation. **(D-E)** Normalized peak response (**D**) and signal-to-noise ratio (**E**) at preferred orientation for each of the 5 experimental conditions. **(F-G)** Histogram (**F**) and normalized orientation selectivity index (OSI, **G**) values for the 5 different experimental conditions. Bar colors denote experimental conditions as labeled in (**D**). Error-bars denote SEM; *p<0.05; **p<0.01; ***p<0.001, Kruskal-Wallis ANOVA followed by post-hoc Bonferroni corrected rank-sum tests relative to WT.

and *MeCP2-gKO* mice had significantly lower OSI values (**Figure 5-7G**), but surprisingly, cells in *SST-cKO* mice showed no change in OSI relative to WT or floxed-MeCP2 mice. Thus, even though MeCP2-deleted SST-INs neurons show weak orientation selectivity themselves (**Figure 5-7E**), this did not affect the orientation selectivity of their target pyramidal cells. Furthermore, OSI values in *MeCP2-gKO* mice were only marginally lower than *PV-cKO* mice ($p = 0.049$, Bonferroni-corrected rank-sum test).

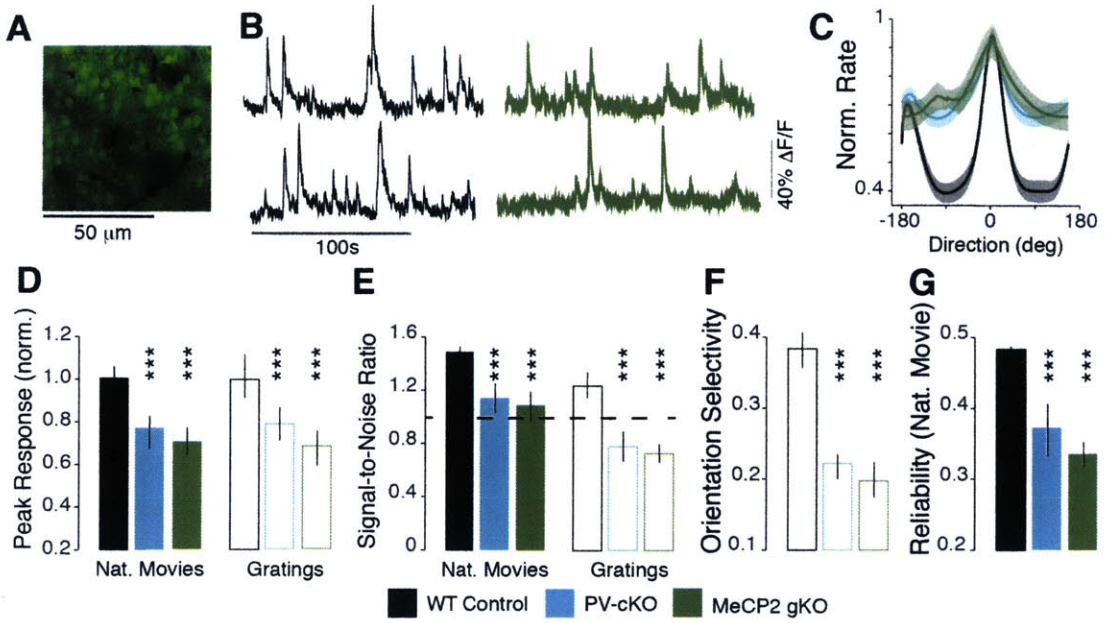

Figure 5-8. Comparison of the response properties of pyramidal neurons in awake, head-fixed mice. (**A**) Example population of neurons labelled with OGB1. (**B**) Example calcium transients from 2 neurons from WT control animal (left) and 2 neurons from a MeCP2 gKO animal (right). (**C**) Population-averaged tuning curves from WT (n = 53 neurons, 3 mice), PV-cKO (n = 25 neurons, 2 mice) and MeCP2 gKO (n = 28 neurons, 2 mice) conditions. These tuning curves were centered at 180°; shaded area represents SEM. (**D**) Quantification of normalized firing rate in response to natural movies (filled bars) and gratings (open bars) from WT (n = 156 neurons), PV-cKO (n =79 neurons) and MeCP2 gKO conditions (n = 99 neurons). (**E**) Quantification of signal-to-noise ratio (SNR) for both natural movies and gratings. (**F**) Quantification of mean orientation selectivity measured from drifting gratings. (**G**) Quantification of mean reliability measured from natural movies. All data in (**D-G**) are represented as mean +/- SEM. ***p<0.0001, Kruskal-Wallis ANOVA with Bonferroni Correction with post-hoc rank-sum tests relative to WT.

Given that anesthesia can impede the function of cortical interneurons (Haider et al., 2013), we repeated these experiments in awake, head-fixed, passively-viewing mice (**Figure 5-8**). In these mice, we confirmed our observation that *PV-cKO* mice also had reduced pyramidal neuron firing rates and orientation selectivity. Together, these results suggest that the poor orientation selectivity, which is perhaps due to the weak spatial acuity in MeCP2-null mice (Durand et al., 2012) can be recapitulated by just knocking out MeCP2 from PV neurons.

How does deleting MeCP2 from these interneuron subtypes influence neural computations under more naturalistic conditions? To answer this question, we presented natural movies, which are spatially and temporally more complex than

sinusoidal drifting gratings. Our previous work (Rikhye and Sur, 2015), together with several other groups (Baudot et al., 2013) established that, because natural movies are broadband in both spatial frequency and orientations, they would elicit complex interactions between neurons (see **Chapter 2**). In agreement with this, we found that neurons in WT mice responded strongly and synchronously to natural movies (**Figure 5-9A**). Surprisingly, neurons in *PV-cKO* and *MeCP2-gKO* mice responded weakly and with greater variability, and hence with reduced signal-to-noise ratio, than neurons in WT and floxed-MeCP2 animals (**Figure 5-9A-D**). We found no significant reduction in average firing rate or signal-to-noise ratio for neurons in *SST-cKO* mice.

Numerous studies have shown that natural scenes are processed in V1 with sparse and highly reliable codes (Haider et al., 2010; Pecka et al., 2014; Rikhye and Sur, 2015; Vinje and Gallant, 2000). A highly reliable and sparsely responding neuron neuron responds with very similar responses on every stimulus repetition, and thus can convey information with high fidelity, while conserving metabolic energy. An unreliable response, on the other hand, would be contaminated by intrinsic variability (i.e. noise) and would not be able to convey information adequately (Sprague et al., 2015). Thus, to assess how well natural scenes are encoded within MeCP2 deficient mice we measured both the reliability and the sparseness of neuronal responses. As described previously, we computed response reliability as the average correlation coefficient over all pair-wise combinations of trials (see Appendix 5B for equation). We discovered that neurons in *PV-cKO* and *MeCP2-gKO* mice, but not *SST-cKO*, responded unreliably to natural movies (**Figure 5-9E**). This is commensurate with the fact that neurons in both *PV-cKO* and *MeCP2-gKO* mice have weaker and more variable responses.

Sparsely responding neurons not only convey information with small metabolic cost (Vinje and Gallant, 2002) but also are selective to only few stimulus features (Pecka et al., 2014). Thus, we used temporal sparseness (see Appendix A) to quantify the selectivity of neurons in normal and MeCP2 deficient mice to different natural movies. We were surprised to find that neurons in all conditions had significantly less selective responses than neurons in WT mice (**Figure 5-9F**). Surprisingly, our results show that, when MeCP2 is deleted from SST neurons, pyramidal cells have unchanged

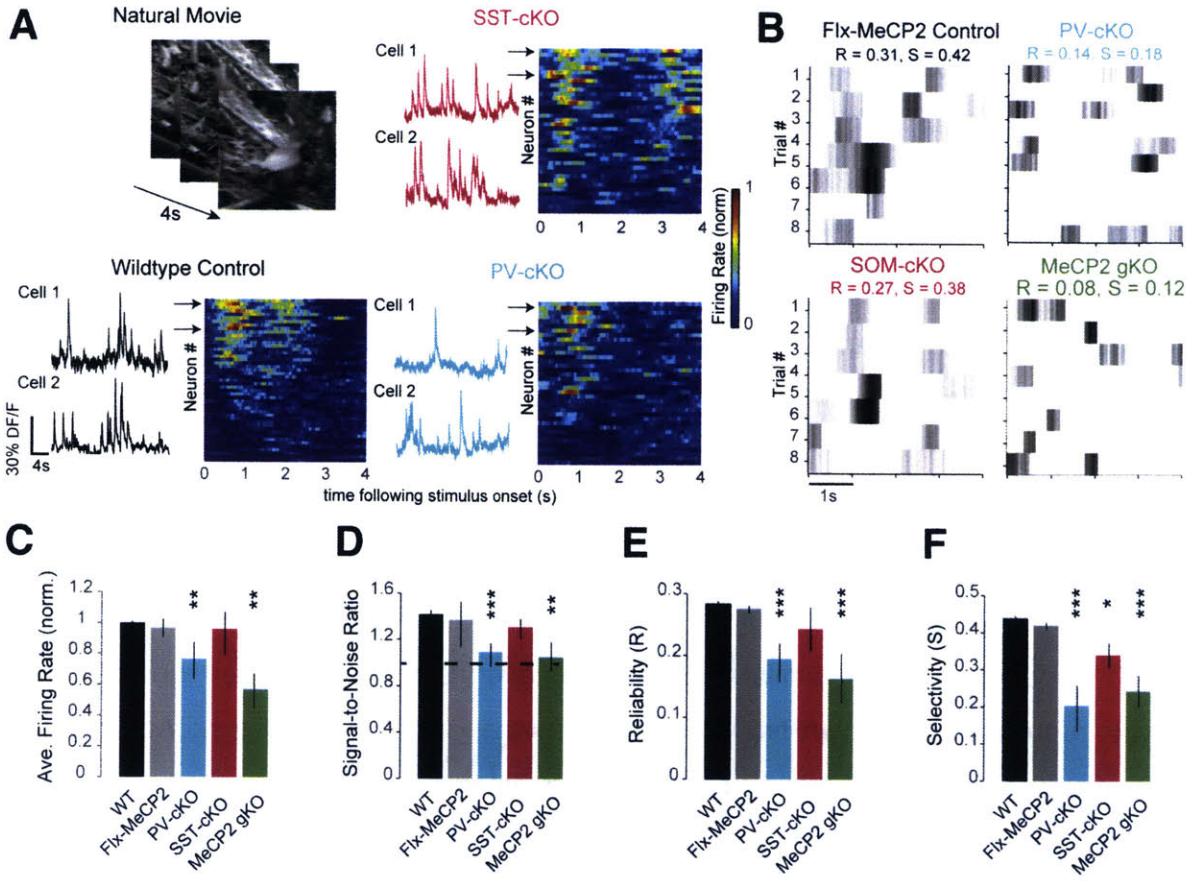

Figure 5-9. Impaired processing of natural movies in MeCP2-deficient mice. (A) Example frames from a natural movie lasting 4s (top, left). Example Ca^{2+} responses from two simultaneously recorded neurons from wild type (WT) control (bottom, left), PV-cKO (bottom, right) and SOM-cKO (top, right) mice. The heat maps show the normalized firing rates from simultaneously recorded populations of neurons. **(B)** Example raster plot (trials vs. time) of a representative neuron from flx-MeCP2 control, PV-cKO, SOM-cKO and MeCP2 gKO mice. Reliability (R) and Selectivity (S) values for each neuron are labeled. (**C-D**) Mean Firing Rate (**C**) and Signal-to-Noise ratio (**D**) of neurons in the 5 experimental conditions to natural movies. **(E-F)** Quantification of Reliability (**E**) and Selectivity (**F**) for each of the 5 experimental conditions. Error bars denote SEM. *p<0.05, **p<0.01, ***p<0.001, Kruskal-Wallis ANOVA followed by post-hoc Bonferroni corrected rank-sum tests relative to WT.

orientation selectivity, and respond reliably to natural movies, albeit with unselective responses. A parsimonious explanation is that deleting MeCP2 from SST neurons does not alter the receptive field structure of pyramidal cells, but changes the organization of the network in which these neurons are embedded in.

To probe these network deficits further, we asked how deleting MeCP2 altered the interactions between neurons. Specifically, we measured signal and Noise

correlations between pairs of recorded neurons. Signal correlations between neuronal responses arise from dependencies between neurons sharing similar receptive field properties and thus provide a measure of similarity in feed-forward input (Ko et al., 2011). Noise correlations, on the other hand, capture dependencies between neurons that are not locked to the sensory stimulus and thus are related to shared network properties including neuronal coupling (Cohen and Kohn, 2011). Deleting MeCP2 from PV, but not SST neurons, resulted in a strong reduction in signal correlation (**Figure 5-10A**). We took advantage of the high spatial resolution of two-photon calcium imaging to analyze how neuronal correlations varied with distance between neurons. Expectedly, signal correlations decayed rapidly in *PV-cKO* mice (**Figure 5-10B, C**), suggesting that deleting MeCP2 from PV neurons strongly decorrelated the network. This result is congruent with several experimental and computational findings that neuronal signal correlations can be modulated by feed-forward PV-dependent inhibition (Litwin-Kumar et al., 2016; Middleton et al., 2012). In contrast, signal correlations persisted over a much longer distance in *SST-cKO* mice, implying that more neurons were co-active, albeit weakly. This is consistent with recent studies, which showed that, by integrating mainly local L2/3 inputs; SST-INs integrate to play a crucial role in spatial summation (Adesnik et al., 2012). This result also reconciles our findings in **Figure 5-9**, as long-range correlations imply redundant network activation, which in turn leads to unselective responses. Not surprisingly, the spatial organization of signal correlations in *MeCP2-gKO* mice closely resembled the *PV-cKO* condition, underscoring the importance of intact feed-forward inhibition in cortical computations.

Noise correlations were also reduced in *PV-cKO* and *MeCP2-gKO* animals but increased in *SST-cKO* animals (**Figure 5-10D**). Across the imaged population, noise correlations decreased almost exponentially between neurons in all conditions, but again persisted over a longer distance in *SST-cKO* mice (**Figure 5-10E, F**). Taken together our results reveal that deficits in *MeCP2-gKO* mice were nearly completely recapitulated in *PV-cKO* mice, indicating that deletion of MeCP2 from PV-INs is sufficient for the circuit-wide deficits of global MeCP2 deletion. Also, our work reveals that deleting MeCP2 exclusively from PV neurons decorrelates the V1 neural network,

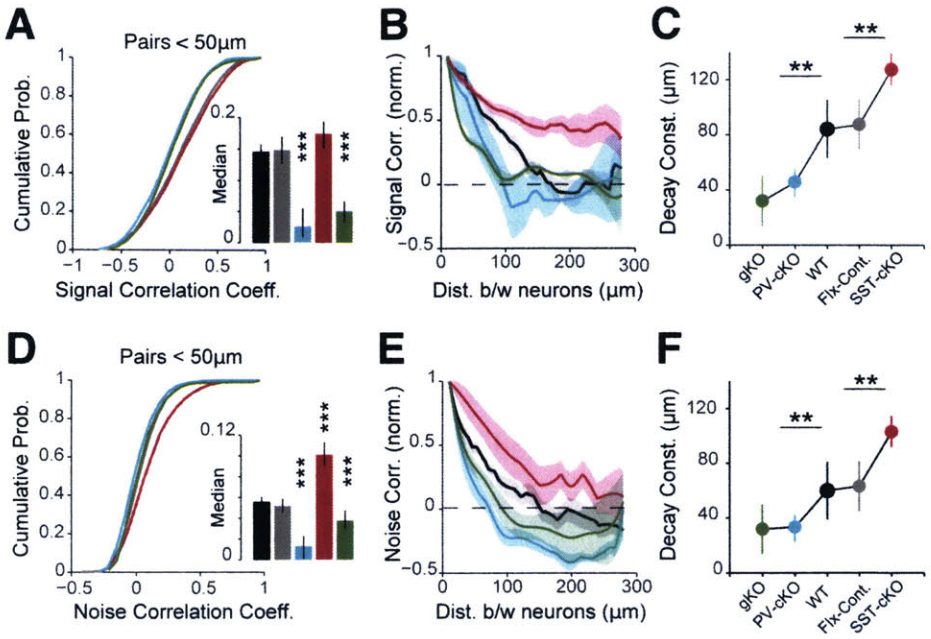

Figure 5-10. Deleting MeCP2 from interneurons alters correlations between neurons. (A) Cumulative distribution of Signal Correlation Coefficient between pairs of neurons closer than 50 μm apart. Inset shows bootstrapped estimate of median signal correlation coefficient. Error-bars denote 95% confidence interval of the median. **(B)** Signal Correlation Coefficient as a function of distance between neurons. Shaded area denotes 95% confidence interval of the median. **(C)** Correlation decay coefficient obtained by fitting single exponentials to the curves in (H). **p<0.01, Wilcoxon rank-sum test relative to WT control. **(D-F)** Same as (**A-C**) respectively but for Noise Correlation Coefficient. Error bars denote SEM. *p<0.05, **p<0.01, ***p<0.001, Kruskal-Wallis ANOVA followed by post-hoc Bonferroni corrected rank-sum tests relative to WT.

making it more noise-dominated, and consequentially limits the ability of neurons to efficiently process visual information. In contrast, deleting MeCP2 from SST neurons increases redundancy between neurons, which also limits the efficiency of visual processing. Thus, perturbing the finely tuned inhibitory microcircuit by deleting MeCP2 has detrimental effects on visual processing. Given the canonical nature of neural computations, we will not be surprised if similar deficits are found in other sensory areas (e.g. auditory cortex (Goffin et al., 2014).

Treatment with rhIGF1 ameliorates the effect of MeCP2 deletion

It is well established that loss of MeCP2 leads to immature synapses in the brain, which is the consequence of a down-regulation of many key molecular signaling pathways (Castro et al., 2014; Tropea et al., 2009). For example, it is known that MeCP2 reduces the expression of brain-derived neurotrophic factor (BDNF), which in turn affects

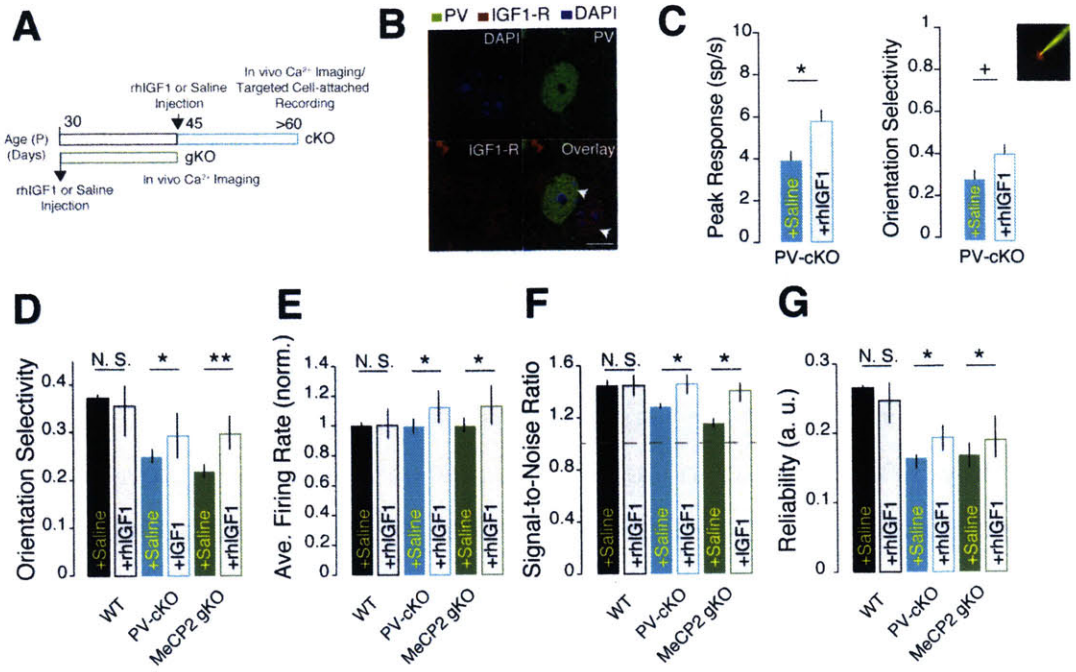

Figure 5-11. rhIGF1 treatment restores response rate, selectivity and reliability in MeCP2-deleted mice. (A) Schematic showing timeline of rhIGF1 administration and experiments. MeCP2 gKO animals were treated (by I.P injection) for two weeks (starting at P30), and recordings were performed ~P45; treatment for PV-cKO animals was started at P45 and Ca^{2+} imaging and electrophysiological recordings were done ~P60. Control animals received saline injections. **(B)** Immunohistochemical labeling showing the presence of IGF1 receptor (IGF1-R, in red) in PV interneurons (in green) and excitatory neurons in L2/3 of mouse V1. Arrows in the overlay show a PV interneuron and a neighboring excitatory neuron, both expressing IGF1-R. Scale bar = 10 μm. (**C**) Two-photon-guided cell-attached recording (*inset*) from PV+ cells in saline (control, n = 10 neurons from 4 mice) and rhIGF1-treated animals (n = 10 neurons from 5 mice) in response to drifting gratings. Values normalized to saline controls. **(D)** Ca^{2+} imaging of responses, showing normalized orientation selectivity index (OSI) computed from responses to drifting gratings. **(E-G)** Mean firing rate (**E**), signal-to-noise ratio (**F**) and response reliability **(G)** in response to natural movies. All values represented as mean ± SEM.). $^{+}p=0.06$; *p<0.05, **p<0.001, Wilcoxon rank-sum test relative to saline control.

signaling pathways (ERK and AKT/mTOR), which are critical for brain development. Thus, these BDNF-mediated pathways are prime targets novel therapeutic strategies aimed at ameliorating the symptoms of Rett Syndrome. Indeed, previous work from our lab has shown that daily treatment with *insulinlike growth factor-1* (rhIGF1) is able to ameliorate a wide range of social and behavioral deficits in both male and female MeCP2-null mice (Castro et al., 2014). We thus sought to investigate the effect that rhIGF1 treatment had on improving the circuit deficits that we observed when we deleted MeCP2 from PV neurons. Note, we have focused mainly on *PV-cKO* mice due

to the fact that deleting MeCP2 just from PV-INs is sufficient to recapitulate the defects seen in *MeCP2-gKO* mice.

Does rhIGF1 treatment improve responses in PV neurons? To answer this question, we performed two-photon imaging guided loose-patch recordings from identified PV-INs in *PV-cKO* mice. Specifically, we performed recordings under two different conditions – a saline-treated condition (control) and a rhIGF1-treated condition (**Figure 5-11A)**, see Appendix 5B for experimental procedures). We found that systemically administering rhIGF1 improved the firing rates of MeCP2-lacking PV-INs by approximately 50% compared to the control condition. This large improvement in firing rate was also accompanied by a smaller increase in OSI (**Figure 5-11C**). We believe that this improvement in PV-IN could be either due to direct activation of IGF1 receptors on PV neurons (**Figure 5-11B**) or through an indirect circuit effect.

Next, we performed two-photon calcium imaging experiments from putative excitatory neurons to determine the effect of rhIGF1 treatment on visual processing. In response to sinusoidal gratings, we found an increase in OSI in both *PV-cKO* and *MeCP2-gKO* mice following treatment (**Figure 5-11D**). In response to natural movies, we also found that treatment rescued deficits in average firing rate, signal-to-noise ratio and response reliability (**Figures 5-11E-G**). Thus, these results demonstrate that rhIGF1 treatment enhances the sensory-driven responses of neurons in both *PV-cKO* and *MeCP2-gKO* mice. Importantly, our data shows that the deficits caused by MeCP2 deletion in circuits performing sensory computations can be specifically ameliorated by rhIGF1 treatment. This amelioration in cortical computations could explain some of the behavioral improvement seen in rhIGF1 treated MeCP2-null mice.

5.4 Discussion

The aim of this study was to investigate the effect that deleting MeCP2 from cortical interneurons had on visual processing in mouse visual cortex. Using a combination of techniques, we show three major results. **(1)** Global deletion of MeCP2 from all neurons results in a reduction of visually driven excitatory and inhibitory currents in V1 pyramidal neurons and alterations in their relative timing. This reduction in magnitude leads to an altered E-I balance. **(2)** PV- and SST-INs lacking MeCP2 have reduced responses and weaker orientation selectivity. As a result of this, pyramidal neurons in these mice also have aberrant responses to both gratings and natural movies. In particular, when we probed visual processing with natural scenes, we discovered that deleting MeCP2 from PV-INs caused reduced responses, increased trial-to-trial variability and reduced selectivity. At the network level, neurons in *PV-cKO* were strongly decorrelated. In contrast, deletion of MeCP2 from SST neurons had more subtle effects. Pyramidal neurons in these mice were more strongly correlated, with inter-neuronal correlations persisting over longer distances. This altered correlation structure made neurons in *SST-cKO* mice less selective for different natural movies. **(3)** Deletion of MeCP2 from PV-INs alone recapitulates effects of global MeCP2 deletion on pyramidal neurons. This underscores the importance of PV-INs in regulating computations in the visual cortex. **(4)** Treatment of mutant mice with rhIGF1 restores PV and pyramidal neuron responses. These results demonstrate that reduction of both inhibition and excitation, importantly via MeCP2 effects on PV-INs, contributes to the deficits of Rett Syndrome.

The role of inhibition in Rett Syndrome pathophysiology

Loss of MeCP2 from a subset of forebrain GABAergic neurons recapitulates diverse and prominent features of RTT (Chao et al., 2010) and the behavioral symptoms of MeCP2 loss can be explained by PV- or SOM-specific MeCP2 deletion (Ito-Ishida et al., 2015). Furthermore, interneuron-specific re-expression of MeCP2 can ameliorate some of the deficits seen in RTT (Goffin et al., 2014). Our findings show that excitatory conductance is reduced concurrently with inhibitory conductance upon MeCP2 deletion. Reduced

excitation and inhibition within adult cortical circuits *in vivo* is consistent with previous findings in slices demonstrating reduced excitatory glutamatergic synapse number and weaker synaptic connections or drive in pyramidal neurons (Dani and Nelson, 2009; Dani et al., 2005; Wood et al., 2009). Anatomical measurements have suggested increased PV-IN puncta and enhanced perisomatic innervation in MeCP2 mutant mice (Durand et al., 2012; Krishnan et al., 2015; Lazarus et al., 2015). However, at the same time, studies have also noted reduced GABA and GAD65 levels in visual cortex (Durand et al., 2012). Together, these contradictory findings suggest that, while PV→Pyramidal connectivity has increased, these synapses do not function normally, which consequentially results in a reduced inhibitory tone. Our functional measurements of inhibitory conductance in adult, male *MeCP2-gKO* mice, reveal reduced inhibition onto pyramidal neurons. Also, our targeted cell-specific recordings of INs show that PV-driven inhibition is weaker in both global KO and PV-specific KO. Interestingly, we found no reduction in SST-IN responses, suggesting that much of the inhibition missing in Rett Syndrome is through PV neurons. Importantly, our measurements of reduced visually driven excitatory conductance demonstrate that reduced visual responses of pyramidal neurons or reduced propagated activity in slices in MeCP2 mutant mice are due not to increased inhibition but to reduced feed-forward excitatory drive. This decrease in feed-forward drive can explain why pyramidal neurons are weakly, and unreliably, driven by visual stimuli in in both *PV-cKO* and *MeCP2-gKO* mice. It also could explain the increased propensity for seizures observed in Rett patients and mouse models (W. Zhang et al., 2014). SST-INs on the other hand may function as integrators of top-down inputs that modulate visual processing (Hangya et al., 2014), or in modulating recurrent inhibition (Adesnik et al., 2012). Thus, altering feed-forward excitation would not influence SST neurons as much as PV neurons.

How feed-forward drive onto PV neurons changes during Rett syndrome is still unknown. It has been shown that deleting MeCP2 leads effects including immature membrane and synaptic properties in PV-INs (He et al., 2014). However no study has yet shown how E- and I- conductance within PV-INs change following MeCP2 deletion.

This could be a potential therapeutic target, as restoring the normal function of PV neurons could compensate for the decrease of feed-forward excitation.

Unreliable processing of visual information in Rett Syndrome

A neuron's responses to visual stimuli directly reflect the circuits within which the neuron is embedded, and responses of V1 neurons in mutant mice thus provide sensitive assays of circuit dysfunction. In particular, the fidelity with which neurons spikes provides an important measure of how well these neurons transmit information. It has been well established that reliable spiking is the consequence of the integration of precisely timed excitatory and inhibitory synaptic currents (Isaacson and Scanziani, 2011; Wehr and Zador, 2003). We found that loss of MeCP2 significantly alters the timing of both excitatory and inhibitory conductance on pyramidal neurons. Also, these synaptic barrages were highly unreliable in *MeCP2-gKO* mice (data not shown). Together, these factors cause pyramidal neurons in MeCP2 null mice to respond to both to both simple (oriented gratings) and complex (natural movies) visual stimuli with high trial-to-trial variability. Similarly, deleting MeCP2 from PV-INs alone causes a significant increase in trial-to-trial variability, and reduced reliability. These results are consistent with our prediction made in **Section 5.2** – chronically perturbing inhibition reduces response reliability.

Surprisingly, in contradiction to this prediction, the effects of deleting MeCP2 from SST-INs on pyramidal neuron reliability are less severe. This is likely due to the fact that deleting MeCP2 from SST-INs did not significantly alter the response properties of SST neurons. This implies that SST-INs are only weakly perturbed in Rett Syndrome. In **Chapter 4**, we established that SST neurons played a major role in modulating pyramidal cell reliability, through its effect on PV neurons. Specifically, our analysis showed that activating SST neurons increase dendritic inhibition, and at the same time, reduced somatic inhibition via the SST→PV disinhibitory circuit. This routing of inhibition towards the dendrites allows only the strongest and most reliable inputs to be integrated by the neuron, leading to reliable spikes. Since knocking-out MeCP2 from SST neurons does not alter SST neuron responses, we believe that dendritic inhibition should still be

intact. Thus, the circuit we have described in **Chapter 4** should still be able to allow reliable spiking to occur. Future experiments can be aimed at studying dendritic integration in before and after MeCP2 deletion.

Why are SST neurons not affected by MeCP2 deletion? Studies have shown that SST neurons gain most of their excitatory inputs from local L2/3 pyramidal neurons (Adesnik et al., 2012) and from higher cortical areas, such as the cingulate cortex (S. Zhang et al., 2014). Additionally, SST neurons also receive cholinergic inputs from the basal forebrain (Wall et al., 2016) and inhibitory inputs from local VIP-INs (Fu et al., 2014). Contrary to common belief, SST neurons also get a small fraction of thalamic inputs (Wall et al., 2016), poising them to mediate both feed-forward and feed-back inhibition. Thus, it is likely that deleting MeCP2 from SST neurons only the strength of their feed-forward excitatory inputs, which could explain why SST-neurons in *SST-cKO* mice have weaker orientation selectivity. Compensatory inputs from higher cortical areas or stronger cholinergic tone could explain why SST neurons do not reduce their firing rates following MeCP2 deletion. Together, the weak effect of SST-specific deletion, underscores further the notion that the circuit deficits seen in Rett Syndrome are due to a loss of feed-forward drive.

Altered population coding of natural movies following MeCP2 deletion

The accuracy of information encoding in the cortex depends not only on the response properties of individual neurons but also on the structure and magnitude of correlations between neurons (Zohary et al., 1994). We demonstrate that deleting MeCP2 from PV neurons, or globally from all cells, significantly decorrelates the network by reducing the magnitude of both signal and noise correlations. Using Gi-DREADDS, we found that silencing PV neurons increase signal and noise correlations, which was primarily due to a non-specific increase in the firing rate of pyramidal cells. In contrast, deleting MeCP2 reduces the overall feedforward drive into L2/3, which in turn leads to weak, noisy and decorrelated firing. In essence, neurons in both *PV-cKO* and *MeCP2-gKO* mice are weakly driven by visual stimuli and can be said to be noise dominated. Surprisingly, we found that deleting MeCP2 from SST neurons significantly increased the magnitude of

noise correlations. This means that pyramidal neurons in these mice shared a stronger "common signal", which could be shared noise. This result suggests that SST neurons work to remove common-mode noise between neurons and could function to coordinate activity within neural ensembles (see **Chapter 4**). In agreement with this notion, we found that signal correlations persisted over a much longer range in *SST-cKO* mice. When read in light of our results in **Chapter 3**, this result implies that coding in *SST-cKO* mice is more redundant (i.e. more neurons are performing the same computation), which could explain why selectivity for both oriented edges and different natural movies is reduced in these mice.

It is well established that Rett Syndrome is not a deficit in sensory processing, as patients with Rett Syndrome can still see normally (Ref?). Despite this, others and we have found sever deficits in the coding properties of neurons in the visual (Durand et al., 2012; He et al., 2014) and auditory cortices (Goffin et al., 2014) of mutant mice. Additionally, fMRI studies have found that visual stimuli are also unreliably processed in patients with ASDs (Dinstein et al., 2012). Together, these studies suggest that Rett Syndrome patients should have sever deficits in the ability to accurately process sensory information. Ultimately, the goal of sensory processing is extract and encode salient features such that higher areas can make decisions and actions based on these features. In this study, we demonstrate that, a strong reduction feed-forward excitation and inhibition in MeCP2-null mice, which in turn leads to noisy and unreliable responses in V1. It is likely that these noisy responses propagate up the visual hierarchy resulting in corrupted representations of the visual scene (Moldakarimov et al., 2015). This corrupted sensory information, along with altered coding in the motor and prefrontal cortices, can lead to many of the behavioral and cognitive deficits seen in Rett patients. Thus, a decrease in sensory processing efficiency may represent a fundamental physiological signature of neural processing in autism spectrum disorders. Future fMRI studies should be targeted at investigating how the flow of visual information is altered in patients with the disease.

rhIGF1 improves cell-specific and circuit-dependent responses

Insulin-like growth factor (IGF) is known to activate key signaling pathways and prime the development and maturation of V1 neurons in mouse visual cortex (Tropea et al., 2006). Levels of IGF1 are severely reduced in both mouse models (Castro et al., 2014) and patients with Rett Syndrome (Khwaja et al., 2014), presumably due to reduced mRNA levels (Mellios et al., 2014). Previous work has shown that rhIGF1 administration enhances multiple components of the PI3K/AKT and ERK signaling pathways, increases PSD95 at excitatory synapses, and enhances excitatory synaptic transmission, and ameliorates behavioral deficits in adult MeCP2 null mice (Castro et al., 2014). Our results demonstrate that rhIGF1 treatment in PV-cKO mice improves not only improves PV-IN responses, but also improves circuit-level deficits. In particular, responses to both gratings and natural movies became stronger and less noisy with rhIGF1 treatment. Thus, taken together, these results show that IGF1 maybe a powerful method to enhance excitatory synaptic transmission following MeCP2 deletion.

In conclusion, by specifically deleting MeCP2 from interneuron subtypes, we have discovered that most of the circuit deficits associated with Rett Syndrome is due to a chronic reduction in feed-forward excitatory drive and a concomitant reduction in inhibition from PV neurons. Our work underscores the importance of PV neurons in maintaining the balance between feed-forward excitation and inhibition within the cortex. Importantly, we have found that IGF1, which increases the strength of excitatory drive, can ameliorate many of the deficits seen in Rett Syndrome. Thus, our work has brought us closer towards understanding the pathophysiology underlying Rett Syndrome. A finer dissection of the mechanisms involved in Rett will allow us to design better therapeutic strategies to ameliorate the deficits associated with the disease.

5.5 References

Abbott, L.F., Dayan, P., 1999. The effect of correlated variability on the accuracy of a population code. Neural Comput. 11, 91–101.

Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z.J., Scanziani, M., 2012. A neural circuit for spatial summation in visual cortex. Nature 490, 226–31. doi:10.1038/nature11526

Averbeck, B.B., Latham, P.E., Pouget, A., 2006. Neural correlations, population coding and computation. Nat. Rev. Neurosci. 7, 358–66. doi:10.1038/nrn1888

Banerjee, A., Castro, J., Sur, M., 2012. Rett syndrome: genes, synapses, circuits, and therapeutics. Front. psychiatry 3, 34. doi:10.3389/fpsyt.2012.00034

Baudot, P., Levy, M., Marre, O., Monier, C., Pananceau, M., Frégnac, Y., 2013. Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. Front. Neural Circuits 7, 206. doi:10.3389/fncir.2013.00206

Benes, F.M., Berretta, S., 2001. GABAergic interneurons: implications for understanding schizophrenia and bipolar disorder. Neuropsychopharmacology 25, 1–27. doi:10.1016/S0893-133X(01)00225-1

Castro, J., Garcia, R.I., Kwok, S., Banerjee, A., Petravicz, J., Woodson, J., Mellios, N., Tropea, D., Sur, M., 2014. Functional recovery with recombinant human IGF1 treatment in a mouse model of Rett Syndrome. Proc. Natl. Acad. Sci. U. S. A. 111, 9941–6. doi:10.1073/pnas.1311685111

Chahrour, M., Zoghbi, H.Y., 2007. The story of Rett syndrome: from clinic to neurobiology. Neuron 56, 422–37. doi:10.1016/j.neuron.2007.10.001

Chao, H.-T., Chen, H., Samaco, R.C., Xue, M., Chahrour, M., Yoo, J., Neul, J.L., Gong, S., Lu, H.-C., Heintz, N., Ekker, M., Rubenstein, J.L.R., Noebels, J.L., Rosenmund, C., Zoghbi, H.Y., 2010. Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. Nature 468, 263–9. doi:10.1038/nature09582

Cohen, M.R., Kohn, A., 2011. Measuring and interpreting neuronal correlations. Nat Neurosci 14, 811–819. doi:10.1038/nn.2842

Cohen, S.M., Tsien, R.W., Goff, D.C., Halassa, M.M., 2015. The impact of NMDA receptor hypofunction on GABAergic neurons in the pathophysiology of schizophrenia. Schizophr. Res. 167, 98–107. doi:10.1016/j.schres.2014.12.026

Dani, V.S., Chang, Q., Maffei, A., Turrigiano, G.G., Jaenisch, R., Nelson, S.B., 2005. Reduced

cortical activity due to a shift in the balance between excitation and inhibition in a mouse model of Rett syndrome. Proc. Natl. Acad. Sci. U. S. A. 102, 12560–5. doi:10.1073/pnas.0506071102

Dani, V.S., Nelson, S.B., 2009. Intact long-term potentiation but reduced connectivity between neocortical layer 5 pyramidal neurons in a mouse model of Rett syndrome. J. Neurosci. 29, 11263–70. doi:10.1523/JNEUROSCI.1019-09.2009

Dinstein, I., Heeger, D.J., Lorenzi, L., Minshew, N.J., Malach, R., Behrmann, M., 2012. Unreliable evoked responses in autism. Neuron 75, 981–91. doi:10.1016/j.neuron.2012.07.026

Durand, S., Patrizi, A., Quast, K.B., Hachigian, L., Pavlyuk, R., Saxena, A., Carninci, P., Hensch, T.K., Fagiolini, M., 2012. NMDA receptor regulation prevents regression of visual cortical function in the absence of Mecp2. Neuron 76, 1078–90. doi:10.1016/j.neuron.2012.12.004

Fino, E., Yuste, R., 2011. Dense inhibitory connectivity in neocortex. Neuron 69, 1188–203. doi:10.1016/j.neuron.2011.02.025

Fu, Y., Tucciarone, J.M., Espinosa, J.S., Sheng, N., Darcy, D.P., Nicoll, R.A., Huang, Z.J., Stryker, M.P., 2014. A cortical circuit for gain control by behavioral state. Cell 156, 1139–52. doi:10.1016/j.cell.2014.01.050

Goffin, D., Brodkin, E.S., Blendy, J.A., Siegel, S.J., Zhou, Z., 2014. Cellular origins of auditory event-related potential deficits in Rett syndrome. Nat. Neurosci. 17, 804–6. doi:10.1038/nn.3710

Gonchar, Y., Wang, Q., Burkhalter, A., 2007. Multiple distinct subtypes of GABAergic neurons in mouse visual cortex identified by triple immunostaining. Front. Neuroanat. 1, 3. doi:10.3389/neuro.05.003.2007

Guy, J., Cheval, H., Selfridge, J., Bird, A., 2011. The role of MeCP2 in the brain. Annu. Rev. Cell Dev. Biol. 27, 631–52. doi:10.1146/annurev-cellbio-092910-154121

Haider, B., Krause, M.R., Duque, A., Yu, Y., Touryan, J., Mazer, J.A., McCormick, D.A., 2010. Synaptic and network mechanisms of sparse and reliable visual cortical activity during nonclassical receptive field stimulation. Neuron 65, 107–121. doi:10.1016/j.neuron.2009.12.005

Haider, B., McCormick, D. a, 2009. Rapid neocortical dynamics: cellular and network mechanisms. Neuron 62, 171–89. doi:10.1016/j.neuron.2009.04.008

Hangya, B., Pi, H.-J., Kvitsiani, D., Ranade, S.P., Kepecs, A., 2014. From circuit motifs to

computations: mapping the behavioral repertoire of cortical interneurons. Curr. Opin. Neurobiol. 26, 117–24. doi:10.1016/j.conb.2014.01.007

He, L., Liu, N., Cheng, T., Chen, X., Li, Y., Shu, Y., Qiu, Z., Zhang, X., 2014. Conditional deletion of Mecp2 in parvalbumin-expressing GABAergic cells results in the absence of critical period plasticity. Nat. Commun. 5, 5036. doi:10.1038/ncomms6036

Isaacson, J.S., Scanziani, M., 2011. How inhibition shapes cortical activity. Neuron 72, 231–43. doi:10.1016/j.neuron.2011.09.027

Ito-Ishida, A., Ure, K., Chen, H., Swann, J.W., Zoghbi, H.Y., 2015. Loss of MeCP2 in Parvalbumin-and Somatostatin-Expressing Neurons in Mice Leads to Distinct Rett Syndrome-like Phenotypes. Neuron 88, 651–8. doi:10.1016/j.neuron.2015.10.029

Jiang, M., Ash, R.T., Baker, S.A., Suter, B., Ferguson, A., Park, J., Rudy, J., Torsky, S.P., Chao, H.-T., Zoghbi, H.Y., Smirnakis, S.M., 2013. Dendritic arborization and spine dynamics are abnormal in the mouse model of MECP2 duplication syndrome. J. Neurosci. 33, 19518–33. doi:10.1523/JNEUROSCI.1745-13.2013

Khwaja, O.S., Ho, E., Barnes, K. V, O'Leary, H.M., Pereira, L.M., Finkelstein, Y., Nelson, C.A., Vogel-Farley, V., DeGregorio, G., Holm, I.A., Khatwa, U., Kapur, K., Alexander, M.E., Finnegan, D.M., Cantwell, N.G., Walco, A.C., Rappaport, L., Gregas, M., Fichorova, R.N., Shannon, M.W., Sur, M., Kaufmann, W.E., 2014. Safety, pharmacokinetics, and preliminary assessment of efficacy of mecasermin (recombinant human IGF-1) for the treatment of Rett syndrome. Proc. Natl. Acad. Sci. U. S. A. 111, 4596–601. doi:10.1073/pnas.1311141111

Ko, H., Hofer, S.B., Pichler, B., Buchanan, K.A., Sjostrom, P.J., Mrsic-Flogel, T.D., 2011. Functional specificity of local synaptic connections in neocortical networks. Nature 473, 87–91. doi:10.1038/nature09880

Krishnan, K., Wang, B.-S., Lu, J., Wang, L., Maffei, A., Cang, J., Huang, Z.J., 2015. MeCP2 regulates the timing of critical period plasticity that shapes functional connectivity in primary visual cortex. Proc. Natl. Acad. Sci. U. S. A. 112, E4782–91. doi:10.1073/pnas.1506499112

Kubota, Y., 2014. Untangling GABAergic wiring in the cortical microcircuit. Curr. Opin. Neurobiol. 26, 7–14. doi:10.1016/j.conb.2013.10.003

Kuhlman, S.J., Olivas, N.D., Tring, E., Ikrar, T., Xu, X., Trachtenberg, J.T., 2013. A disinhibitory microcircuit initiates critical-period plasticity in the visual cortex. Nature 501, 543–6. doi:10.1038/nature12485

Lazarus, M.S., Krishnan, K., Huang, Z.J., 2015. GAD67 deficiency in parvalbumin interneurons

produces deficits in inhibitory transmission and network disinhibition in mouse prefrontal cortex. Cereb. Cortex 25, 1290–6. doi:10.1093/cercor/bht322

Lewis, D.A., 2012. Cortical circuit dysfunction and cognitive deficits in schizophrenia--implications for preemptive interventions. Eur. J. Neurosci. 35, 1871–8. doi:10.1111/j.1460-9568.2012.08156.x

Litwin-Kumar, A., Rosenbaum, R., Doiron, B., 2016. Inhibitory stabilization and visual coding in cortical circuits with multiple interneuron subtypes. J. Neurophysiol. 115, 1399–409. doi:10.1152/jn.00732.2015

Lovett-Barron, M., Turi, G.F., Kaifosh, P., Lee, P.H., Bolze, F., Sun, X.-H., Nicoud, J.-F., Zemelman, B. V, Sternson, S.M., Losonczy, A., 2012. Regulation of neuronal input transformations by tunable dendritic inhibition. Nat. Neurosci. 15, 423–430. doi:10.1038/nn.3024

McGraw, C.M., Samaco, R.C., Zoghbi, H.Y., 2011. Adult neural function requires MeCP2. Science 333, 186. doi:10.1126/science.1206593

Mellios, N., Woodson, J., Garcia, R.I., Crawford, B., Sharma, J., Sheridan, S.D., Haggarty, S.J., Sur, M., 2014. β2-Adrenergic receptor agonist ameliorates phenotypes and corrects microRNA-mediated IGF1 deficits in a mouse model of Rett syndrome. Proc. Natl. Acad. Sci. U. S. A. 111, 9947–52. doi:10.1073/pnas.1309426111

Middleton, J.W., Omar, C., Doiron, B., Simons, D.J., 2012. Neural correlation is stimulus modulated by feedforward inhibitory circuitry. J. Neurosci. 32, 506–18. doi:10.1523/JNEUROSCI.3474-11.2012

Moldakarimov, S., Bazhenov, M., Sejnowski, T.J., 2015. Feedback stabilizes propagation of synchronous spiking in cortical neural networks. Proc. Natl. Acad. Sci. U. S. A. 112, 2545–2550. doi:10.1073/pnas.1500643112

Moreno-Bote, R., Beck, J., Kanitscheider, I., Pitkow, X., Latham, P., Pouget, A., 2014. Information-limiting correlations. Nat. Neurosci. 17, 1410–7. doi:10.1038/nn.3807

Okun, M., Lampl, I., 2008. Instantaneous correlation of excitation and inhibition during ongoing and sensory-evoked activities. Nat. Neurosci. 11, 535–7. doi:10.1038/nn.2105

Pecka, M., Han, Y., Sader, E., Mrsic-Flogel, T.D., 2014. Experience-Dependent Specialization of Receptive Field Surround for Selective Coding of Natural Scenes. Neuron. doi:10.1016/j.neuron.2014.09.010

Rikhye, R. V, Sur, M., 2015. Spatial Correlations in Natural Scenes Modulate Response Reliability in Mouse Visual Cortex. J. Neurosci. 35, 14661–80.

doi:10.1523/JNEUROSCI.1660-15.2015

Rubenstein, J.L.R., Merzenich, M.M., 2003. Model of autism: increased ratio of excitation/inhibition in key neural systems. Genes. Brain. Behav. 2, 255–67.

Sahin, M., Sur, M., 2015. Genes, circuits, and precision therapies for autism and related neurodevelopmental disorders. Science 350. doi:10.1126/science.aab3897

Singh, A., Lesica, N.A., 2010. Incremental mutual information: a new method for characterizing the strength and dynamics of connections in neuronal circuits. PLoS Comput. Biol. 6, e1001035. doi:10.1371/journal.pcbi.1001035

Sprague, T.C., Saproo, S., Serences, J.T., 2015. Visual attention mitigates information loss in small- and large-scale neural codes. Trends Cogn. Sci. doi:10.1016/j.tics.2015.02.005

Tasic, B., Menon, V., Nguyen, T.N., Kim, T.K., Jarsky, T., Yao, Z., Levi, B., Gray, L.T., Sorensen, S.A., Dolbeare, T., Bertagnolli, D., Goldy, J., Shapovalova, N., Parry, S., Lee, C., Smith, K., Bernard, A., Madisen, L., Sunkin, S.M., Hawrylycz, M., Koch, C., Zeng, H., 2016. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat. Neurosci. 19, 335–46. doi:10.1038/nn.4216

Tropea, D., Giacometti, E., Wilson, N.R., Beard, C., McCurry, C., Fu, D.D., Flannery, R., Jaenisch, R., Sur, M., 2009. Partial reversal of Rett Syndrome-like symptoms in MeCP2 mutant mice. Proc. Natl. Acad. Sci. U. S. A. 106, 2029–34. doi:10.1073/pnas.0812394106

Urban, D.J., Roth, B.L., 2015. DREADDs (designer receptors exclusively activated by designer drugs): chemogenetic tools with therapeutic utility. Annu. Rev. Pharmacol. Toxicol. 55, 399–417. doi:10.1146/annurev-pharmtox-010814-124803

Vinje, W.E., Gallant, J.L., 2002. Natural stimulation of the nonclassical receptive field increases information transmission efficiency in V1. J Neurosci 22, 2904–2915. doi:20026216

Vinje, W.E., Gallant, J.L., 2000. Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. Science (80-.). 287, 1273–1276. doi:10.1126/science.287.5456.1273

Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R., Paninski, L., 2010. Fast nonnegative deconvolution for spike train inference from population calcium imaging. J. Neurophysiol. 104, 3691–704. doi:10.1152/jn.01073.2009

Wall, N.R., De La Parra, M., Sorokin, J.M., Taniguchi, H., Huang, Z.J., Callaway, E.M., 2016. Brain-Wide Maps of Synaptic Input to Cortical Interneurons. J. Neurosci. 36, 4000–9. doi:10.1523/JNEUROSCI.3967-15.2016

Wehr, M., Zador, A.M., 2003. Balanced inhibition underlies tuning and sharpens spike timing in

auditory cortex. Nature 426, 442–446. doi:10.1038/nature02116

Wilson, N.R., Runyan, C.A., Wang, F.L., Sur, M., 2012. Division and subtraction by distinct cortical inhibitory networks in vivo. Nature 488, 343–348. doi:10.1038/nature11347

Womelsdorf, T., Valiante, T.A., Sahin, N.T., Miller, K.J., Tiesinga, P., 2014. Dynamic circuit motifs underlying rhythmic gain control, gating and integration. Nat. Neurosci. 17, 1031–9. doi:10.1038/nn.3764

Wood, L., Gray, N.W., Zhou, Z., Greenberg, M.E., Shepherd, G.M.G., 2009. Synaptic circuit abnormalities of motor-frontal layer 2/3 pyramidal neurons in an RNA interference model of methyl-CpG-binding protein 2 deficiency. J. Neurosci. 29, 12440–8. doi:10.1523/JNEUROSCI.3321-09.2009

Xue, M., Atallah, B. V, Scanziani, M., 2014. Equalizing excitation-inhibition ratios across visual cortical neurons. Nature 511, 596–600. doi:10.1038/nature13321

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.-C., Jenvay, S., Miyamichi, K., Luo, L., Dan, Y., 2014. Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science 345, 660–5. doi:10.1126/science.1254126

Zhang, W., Peterson, M., Beyer, B., Frankel, W.N., Zhang, Z., 2014. Loss of MeCP2 from forebrain excitatory neurons leads to cortical hyperexcitation and seizures. J. Neurosci. 34, 2754–63. doi:10.1523/JNEUROSCI.4900-12.2014

Zhang, W., Zhang, L., Liang, B., Schroeder, D., Zhang, Z.-W., Cox, G.A., Li, Y., Lin, D.-T., 2016. Hyperactive somatostatin interneurons contribute to excitotoxicity in neurodegenerative disorders. Nat. Neurosci. 19, 557–9. doi:10.1038/nn.4257

Zoghbi, H.Y., Bear, M.F., 2012. Synaptic dysfunction in neurodevelopmental disorders associated with autism and intellectual disabilities. Cold Spring Harb. Perspect. Biol. 4. doi:10.1101/cshperspect.a009886

Zohary, E., Shadlen, M.N., Newsome, W.T., 1994. Correlated neuronal discharge rate and its implications for psychophysical performance. Nature 370, 140–3. doi:10.1038/370140a0

Appendix 5A – Experimental Procedures for DREADD experiments

Experiments

Animals

All experiments were carried out under protocols approved by MIT's Committee on Animal Care and conformed to NIH guidelines. Two mouse lines were used in these experiments: PV-IRES-Cre (Jax: 008069) and SST-IRES-Cre (Jax: 013044). All mice were maintained on a C57BL6/J background. Mice were housed in a standard 12 hour light-dark cycle with *ad libitium* food and water. Experiments were performed during the light cycle. Only mice between 6-8 weeks old were used for this study.

The number of mice and neurons used are as follows. PV + Gi-DREADD = 3 mice (78 neurons), SST + Gi-DREADD = 3 mice (82 neurons), PV + Gq-DREADD = 3 mice (102 neurons), SST + Gq-DREADD = 3 mice (98 neurons).

Surgery and Virus Injection

Mice were anesthetized using isoflurane (3% induction, 1.5-2% during surgery). A custom-built metal head-post was attached to the skull using dental cement (C&B Metabond, Parkell) and a 3mm diameter craniotomy was performed over monocular V1 (approximately 2-3mm lateral and 0.5mm anterior to lambda). Care was taken not to rupture the dura mater. The core body temperature was maintained at 37.5°C using a heating blanket (Harvard Apparatus). To deliver virus particles, we performed microinjection of AAV1.Syn.GCaMP6f.WPRE.SV40 (University of Pennsylvania Vector Core) and either AAV1.hSyn.DIO.hM3d(Gq)-mCherry or AAV1.hSyn.DIO.hM4d(Gi)-mCherry (UNC Vector Core), using a beveled (30μm tip diameter) micropipette. Both viruses (GCaMP6f and DREADD) were mixed at a 50-50% ratio before performing the injection. Typically, to allow for optimal labeling, we performed 5-6 microinjections of 100nl each at different, non-overlapping locations. Injections were performed at a rate of 50nl/min and the pipette was allowed to remain in the cortex for approximately 5 minutes post-injection. Following microinjection, the craniotomy was sealed with a 5mm

cranial window using Loctite cyanoacrylate glue. More Metabond was applied between the cranial window and head-post. For more details see (Goldey et al., 2014)

Clozapine-N-Oxide administration

Clozapine-N-Oxide (CNO, Sigma-Aldrich) was dissolved in 0.9% sterile saline to an effective concentration of 1mg/kg. Mice received an intraperitoneal injection of either CNO or 0.9% saline (control) after one imaging session and were allowed to recover in their home cage for approximately 2.5 hours before imaging commenced again.

Two-photon calcium imaging

Mice were allowed to recover up to 14 days post-surgery to allow for adequate expression of the virus before imaging commenced. Mice were restrained under the two-photon microscope in a custom built apparatus. Typically, we allowed mice to habituate to head fixation for 1-2 sessions before commencing imaging experiements. We found that this helped to minimize motion artifacts.

Imaging was performed using a Prairie Ultima two-photon system (Bruker) with a Spectra Physics Mai-Tai eHP laser coupled with a DeepSee module. The excitation laser was tuned to 940nm and imaging was performed through an Olympus XL Plan N Objective (25x, NA = 1.05). Images were collected at 20 frames/s using ScanImage 3.8 (Vidrio Technologies) and raster scans covered a cortical area of 250x250μm^2. This allowed us to sample from approximately 100 neurons per field of view.

Images were corrected for motion artifacts offline using standard ImageJ (NIH) plugins (CV Template Matching). Image stacks were segmented manually with regions of interest (ROIs) only drawn over neurons that had GCaMP6f excluded from the nucleus, as filled neurons are known to be unhealthy (Ref). Neurons expressing red puncta, which is indicative of DREADD expression, were also not analyzed. The same cortical field of view was found over several days using blood vessel landmarks and average intensity projection maps. Care was taken to ensure that the z-plane was almost exactly aligned. To ensure that the same cells were include in the "before" and "after" conditions, the same ROI masks were used to segment images collected after

Saline or CNO administration. Cells in the after condition that moved relative to the original ROIs were not analyzed.

Visual stimulation

Movies were presented on a 7-inch 1080p TFT LCD monitor (DBPower) using custom software written in PsychToolbox-3 on a Windows 7 computer (Dell Precision) with a GeForce 8800 GTS 640MB graphics card (PNY). The monitor was placed approximately 4 inches from the contralateral eye and covered a visual space of approximately 70 degrees x 90 degrees. Movies were displayed on the entire monitor (full-field) and were not vignetted. Prior to starting experiments, we used a hand mapping technique to verify that the neural population being measured had receptive fields on the screen. Briefly, the stimulus used for hand mapping was a sequence of square-wave gratings drifting at 2 cycles/s at 8 different orientations (2s per orientation) vignetted to an annulus of diameter of 30 degrees. Neural populations with receptive fields not located on the screen or at the very edge of the screen were not used. Often, the screen location was adjusted such that the "population receptive field" was close to the center of the screen.

A total of 5 natural movies were selected from the van Hateren natural movie database (Rikhye and Sur, 2015). Gray scale values of each movie were discretized to 255 values and each frame was normalized to have equal mean luminance (mean of histogram = 128) and contrast (std. deviation of histogram = 32). Each movie was adjusted to have the same total SF distribution (Rikhye and Sur, 2015). To slow down the movie from its original 60Hz frame rate, we updated every three frames creating an effective frame rate of close to 30Hz. Movies were presented for 4s and were flanked by 4s gray screens (mean luminance = 128). Each movie was repeated 80 times in order to calculate response reliability.

Data Analysis

Significantly visually responsive cells were determined from the fluorescence time changes ($\Delta F/F$) by performing a one-tailed Student's t-test between visually evoked and

spontaneous responses (gray screen collected for 2 minutes before start of experiment). Only cells with $p < 10^{-3}$ were classified as visually-responsive. Firing rates of these cells were then inferred using an optimized spike rate inference algorithm (Vogelstein et al., 2010) using parameters that were previously verified in our lab. The Vogelstein algorithm infers the probability of spiking from calcium transients. To convert this probability into a firing rate (measured in events/s), we multiplied each probability by 20Hz, the frequency at which the calcium transients were sampled. Unless otherwise stated, all data analysis was performed using inferred firing rates.

Response reliability analysis.

Let the response of a neuron to trial *i* of movie A be $f_{i,A}$ then the response reliability to movie A (R_A) was calculated using the following equation, where CC is the Pearson's correlation coefficient.

$$R_A = \frac{2}{T^2 - T} \sum_{i=1}^{T} \sum_{j=i+1}^{T} CC(f_{i,A}, f_{j,A})$$

Thus, the response reliability is the average correlation of all pairwise combinations of trials for a single movie. Only neurons with significant responses on more than 10 trials were selected for this analysis.

Analysis of signal and noise correlation between neurons.

Signal correlations between pairs of neurons were calculated as the Pearson correlation coefficient between trial-averaged responses binned at 200ms. To compute noise correlations, we first subtracted the trial average from responses in each trial and then computed the Pearson correlation coefficient (CC) between these mean-subtracted responses, again binned at 200ms (Singh and Lesica, 2010).

Between-movie discriminability analysis

We performed discrimination analysis by quantifying the degree of overlap between the population codes of two different movies. To do so, we first used trial-averaged responses of all neurons to form a population vector

$$\boldsymbol{F}(t) = [f_1(t), f_2(t), \ldots, f_N(t)]$$

where N is the total number of neurons in the recorded population and $f_1(t)$ is the firing rate of a neuron binned at 200ms. This population vector spans a N - dimensional subspace. However, not all the dimensions are significant. Thus, to reduce the dimensionality of the population vector, we performed PCA (Cunningham and Yu, 2014) and kept only the three dimensions that explained more than 80% of the variance in the population vector. We projected the population vector onto the subspace spanned by these three dimensions. Next, we computed the discriminability between the vectors for different movies using the Mahalanobis distance metric

$$\Delta(\boldsymbol{x}) = \sqrt{(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})}$$

where $\boldsymbol{x}$ is the projection of the population vector on the PC subspace and $\boldsymbol{\Sigma}$ is the covariance matrix between the different population vectors. Using this analysis, a high Mahalanobis distance indicates low overlap between population codes and thus an easily discriminable pair of movies. We repeated this analysis for all time points (binned at 200ms) and for all pairs of movies (10 in total).

Statistical analysis

All statistical analysis was performed using custom written routines in MATLAB. The Shapiro-Wilk test was first used to determine if the data was normally distributed. All data presented in this paper are non-normally distributed. Since we were making comparisons between the same set of neurons exposed to either Saline or CNO, we performed repeated measures ANOVA (Friedman Test) and corrected for multiple comparisons using Bonferroni's method.

Appendix 5B – Experimental Procedures for Section 5.3

Experiments

Animals

All experiments were carried out under protocols approved by MIT's Committee on Animal Care and conformed to NIH guidelines. *Mecp2*$^{-/y}$ hemizygous KO mice and wild type littermates were obtained by breeding *Mecp2*$^{+/-}$ heterozygous females on a C57BL/6 background with wild type male mice on the same background. Neuronal subtype-specific deletion of MeCP2 was achieved by crossing cell-type specific Cre-driver lines (Pvalb-IRES-Cre and Sst-IRES-Cre) with homozygous female mice carrying two functional loxP sites flanking exons 3-4 of MeCP2 on the X- chromosome (B6.129P2-Mecp2^{tm1Bird}/J or flox-MeCP2, Jax: 007177). Male mice were used in all experiments. Mice were grouped with their wild type siblings and housed at 24°C and variable humidity in a 12 h dark-light cycle. All mice were on a C57BL6/J background.

To specifically delete MeCP2 in PV-INs (*PV-cKO*), we breed homozygous female flox-MeCP2 mice with PV-IRES-Cre knock-in homozygous male mice. Similarly, *SST-cKO* mice were obtained by breeding homozygous flox-MeCP2 female with SOM-IRES-Cre (Jax: 013044) to achieve specific deletion of MeCP2 in SOM+ interneurons. Excitatory neuron-specific MeCP2 deletion was achieved by crossing homozygous flox-MeCP2 females with homozygous Camk2a-cre males (Jax: 005359). Male flox-MeCP2 animals were used as controls. We also crossed PV- and SST-cre male mice with *Mecp2*$^{+/-}$ heterozygous females, took F1 animals and genotyped for -cre+ and MeCP2 allele. We used this to compare effects of MeCP2 deletion on PV and SST response properties in a global KO (gKO) background compared to cell-specific KO (cKO) mice.

For **Figure 5-5**, the number of mice and neurons used is: Wild-type (WT) = 5 mice (6 neurons) and *MeCP2-gKO* = 4 mice (7 neurons). **Figure 5-6**: WT = 12 mice (29 neurons), *MeCP2-gKO* = 6 mice (14 neurons) and *CaMKII-cKO* = 6 mice (13 neurons). PV-WT = 15 mice (53 neurons), *PV-cKO* = 7 mice (18 neurons), *PV-gKO* = 5 mice (14 neurons), *SST-WT* = 5 mice (6 neurons), *SST-cKO* = 5 mice (13 neurons), *SST-gKO* =

3 mice (6 neurons). **Figure 5-8**: WT = 6 mice (327 neurons), flx-MeCP2 = 6 mice (270 neurons), *PV-cKO* = 4 mice (146 neurons), *SST-cKO* = 4 mice (240 neurons) and *MeCP2-gKO* = 4 mice (163 neurons). **Figure 5-9**: WT = 8 mice (515 neurons), flx-MeCP2 = 6 mice (380 neurons), *PV-cKO* = 4 mice (146 neurons), *SST-cKO* = 4 mice (240 neurons) and *MeCP2 gKO* = 4 mice (163 neurons). Figure 5-10: WT (saline) 4 mice (189 neurons), WT (rhIGF1) = 4 mice (157 neurons), *PV-cKO* (saline) = 4 mice (294 neurons), *PV-cKO* (rhIGF1) = 4 mice (207 neurons), *MeCP2-gKO* (saline) = 3 mice (277 neurons), *MeCP2 gKO* (rhIGF1) = 3 mice (125 neurons).

Virus Injection

Three- to four-week-old mice were briefly anesthetized with isoflurane (3%) in oxygen in an anesthesia chamber and taken to injection platform. The cranium was secured with ear bars and anesthesia was maintained during the surgery with 1.5-2.0% isoflurane. L-shaped incision was made in the skin, and surface of the cranium was cleaned using absorbent swabs (Sugi; Kettenbach, GmbH). Injection coordinates were 3 mm lateral (from Bregma) and 1 mm anterior from the posterior suture. The skull was thinned along a 1 mm line at the rostral edge of V1 using a dremel drill with occasional cooling with saline, and the remaining dura was punctured using a glass micropipette filled with the virus suspended in mineral oil.

To study cellular response properties, red fluorescent protein (RFP) was expressed in PV+ and SOM+ neurons in V1 of PV/SOM-Cre mice by infection of an lox-STOP-lox (LSL) RFP construct packaged into adeno-associated virus.

Several injections (3-4) were made at neighbouring sites, at a depth of 200-250 µm. A volume of 200 nl of virus was injected at 75 nl/min rate at each site. After each injection, the pipette was held in place for 3-5 minutes prior to retraction to prevent leakage. Skin was sutured using a synthetic, monofilament, non-absorbable suture (Prolene 7.0, Ethicon).

Surgery

Two to three weeks after virus injection, mice were used for electrophysiology experiments. Mice were anesthetized using urethane (1.5 mg/g, i.p.) and chlorprothixene (10 mg/kg, i.p.). The eyes were protected with ophthalmic ointment during the surgery and moistened afterward with saline. The skin was excised, cleaned with sugi and a metal head plate was attached to the skull using superglue and dental acrylic, and a 2 mm x 2 mm craniotomy was performed over the primary visual cortex (V1, area 17). The exposed area was then covered with a thin layer of 1.5% agarose in artificial cerebrospinal fluid (aCSF; 140 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 0.01 mM EDTA, 10 mM HEPES, 10 mM glucose; pH 7.4). Head plates were screwed onto a moveable platform and mice were transferred to the microscope. During the recording, additional anesthetic was injected, as needed. Body temperature was maintained throughout the surgery and experiment at 37.5°C with a heating blanket (Harvard Apparatus) and supplementary heating pads. Saline (0.9% NaCl) was injected multiple times during the whole experiment to hydrate the mouse. The recording phase of electrophysiology experiments typically lasted for about 4-6h.

***In vivo* whole-cell electrophysiology**

A small craniotomy (> 0.5 mm diameter) was performed over V1. The dura was removed using fine forceps making sure that the brain remained accessible. The brain was covered with aCSF containing in mM: 140 NaCl, 5 KCl, 2 $CaCl_2$, 1 $MgCl_2$, 0.01 EDTA, 10 HEPES and 10 Glucose. 4-6 MΩ patch pipettes were filled with a Cesium-based intracellular solution containing in mM: 125 Cs-Methanesulfonate, 2 KCl, 10 HEPES, 4 Mg-ATP, 0.3 Na_2-GTP, 10 EGTA, 5 TEA-Cl, 8 Na_2-Phosphocreatine, 2 QX-314; pH 7.3; 290-295 mOsm. Access resistance was monitored throughout the recording period by applying a small -10 mV square pulse at the beginning of each visual stimulus trial. Cells with an access resistance exceeding 50 MΩ, or with access resistance varying more than 10% during the entire recording period, were discarded. The average access resistance was 38 ± 5 MΩ, and 37 ± 3 MΩ in WT and MeCP2 gKO conditions respectively. Excitatory and inhibitory conductances were isolated using the formula:

$$\Delta I = G_e(V - E_e) + Gi(V - E_i)$$

where ΔI is the current recorded at any time point after subtracting the average baseline current ($I - I_0$); G_e and G_i are the excitatory and inhibitory synaptic conductance; V is the holding voltage, and E_e and E_i are the reversal potentials for excitatory and inhibitory currents (+20 mV and -70 mV respectively). The clamping voltage was corrected for the effective series resistance for each cell using: $V = V_h - R_s I$, where V_h is the applied holding voltage, and R_s is the effective series resistance.

Blind patch recording was performed to target layer 2/3 neurons of V1. The position of each cell was monitored using the reading on the micromanipulator (Sutter Instrument). On average, cells from WT and MeCP2 were recorded at a depth of 220 ± 26 μm and 189 ± 11 μm respectively. For some of the recordings, 0.5 % biocytin (wt/wt) was added to the intracellular solution to reconstruct the neuronal morphology. Animals were deeply anesthetized at the end of the experiment and transcardially perfused with 0.9% NaCl followed by 4% PFA. 100 μm coronal sections of V1 were cut using a vibratome (Leica VT1200s). Slices were treated with 1% (wt/wt) Triton in PBS for 4h, and stained overnight with PBS+ 0.1% Triton + 3% BSA + 1/200 Alexa-488-conjugated streptavidin (Lifescience). Neurons were imaged using a confocal microscope (Zeiss) and traced using the neurite tracer in ImageJ.

In vivo two-photon-guided cell-attached recordings

Borosilicate pipettes (outer diameter = 1.5mm, inner diameter = 1.17mm, WPI) were pulled using a Sutter P-2000 laser puller (Sutter Instruments). Tip diameter was around 1 μm while the resistance was between 3-7 MΩ. Recordings were made using custom software (Network Prism, Sur Lab) written in Matlab (MathWorks) controlling a MultiClamp 700B amplifier (Axon). A glass pipette electrode was inserted into the brain at an angle of 20-35 deg and an Ag/AgCl ground electrode pellet (Warner Instruments) was positioned in the same solution as the brain and the objective. For visualization, pipettes were filled with Alexa Fluor 488 (100 μM, Molecular Probes). The pipette was first targeted to the injection site using a 10X lens, and then targeted to individual cells using a 25x lens via simultaneous two-photon imaging at 770 or 920 nm. Cells were

targeted either by moving the pipette into the 3D vicinity and then commencing diagonal advance, or cells were patched blindly by advancing diagonally through the cortex. While approaching a cell, either zero pressure or light positive pressure was applied. Cell proximity was detected through deflections in resistance observed in voltage-clamp during a rapidly time-varying 5 mV command voltage pulse. Once resistance had increased by 5-10 MΩ, slight negative pressure was applied and the pipette was advanced more slowly until resistance increased to 20-80 MΩ. At that point, the amplifier was switched to current-clamp, and spikes were recorded with zero injected current, under a Bessel filter of 4 KHz and an AC filter of 300 Hz. All recorded cells were located in the superficial cortical layer (between 150-330 µm below the pial surface).

In vivo two-photon Ca^{2+} imaging

Craniotomies were performed over binocular V1 as described above. The synthetic calcium indicator OGB-1AM (Invitrogen) was dissolved in 20% Pluronic F-127/DMSO (Invitrogen) for 25 minutes to a concentration of 10 mM. This solution was then diluted in sterile Calcium-free PBS (Invitrogen; pH 7.4) to achieve a final concentration of 1 mM. In addition 100 uM of AlexaFluor 488 (Invitrogen) was added to the solution to visualize the pipette during the injection. Before injection, this solution was filtered with a 0.45 µm centrifugal filter. Following the craniotomy, the dye solution was pressure injected 180-200 µm below the pial surface using a Picospritzer. Typically, we used 10 PSI for 1 minute and repeated the injection at 3-5 slightly overlapping regions. Following this, a small 4mm diameter glass coverslip (Warner Instruments) was implanted directly over the exposed cortex and sealed with a silicone elastomer (Kwik-Sil, WPI). This prevented large brain movements and allowed stable recording.

Imaging was performed using a Prairie Ultima two-photon system (Prairie Technologies) with a Spectra Physics Mai-Tai eHP laser coupled with a DeepSee module. The excitation laser was tuned to 960nm and imaging was performed through an Olympus XL Plan N Objective (25x, NA = 1.05). We used custom written Matlab scripts (MathWorks) to automatically detect neurons, to define an optimized scan path and to collect raster scans at approximately 50 frames per second (Wilson et al., 2013).

Visual stimulation

To assess the orientation tuning of neurons, we presented oriented gratings on a 23" 1080p LCD monitor (Dell) using custom software written in PsychToolbox-3 on a Windows 7 computer (Dell Precision) with a GeForce 8800 GTS 640MB graphics card (PNY). Gratings were optimized for cellular responsiveness using a contrast of 100%, spatial frequency of 0.03 cycles/degree, and a temporal frequency of 1-3 Hz. Gratings were then presented by stepping the orientation from 0-360 degrees in steps of 20 degrees, with each grating presentation being preceded for 4 seconds "off" followed by 4 seconds "on", for a total presentation duration of 144 seconds.

Natural movies were selected from the van Hateren natural movie database. Gray scale values of each movie were discretized to 255 values and each frame was normalized to have equal mean luminance (mean of histogram = 128) and contrast (std. deviation of histogram = 32). To slow down the movie from its original 60Hz frame rate, we updated every three frames creating an effective frame rate of close to 30Hz. Movies were presented for 4s and were flanked by 2s gray screens (mean luminance = 128). Each movie was repeated 10 times in order to calculate response reliability.

Systemic administration of rhIGF1

Animals were weighed and injected intra-peritoneally once daily with either vehicle (saline) or recombinant human insulin-like growth factor-1 (rhIGF1, 70 amino acid; 2.5 mg/kg; Peprotech, NJ) dissolved in saline with 0.01% BSA (weight/volume). Animals were kept in individual home cages and injected daily for 10-14 days, starting on P45.

Data analysis

Data was acquired directly into Matlab (The MathWorks, MA) via custom software Spike detection was performed via analysis routines that used manually defined thresholds followed by spike shape template matching for further verification. Every spike was tagged and displayed on screen in a graphical user interface whereupon it was manually reviewed for false positives and negatives by an experimenter. Spike times in

response to every stimulus were then grouped into "on" or "off" periods based on their timing relative to visual stimulation, and "on" spikes for each stimulus were decremented by the number of "off" spikes observed during an equal time period. For orientation experiments, # spikes per stimulus = (# spikes "on") - (# spikes "off") because "on" and "off" periods were of the same duration.

For calcium imaging experiments, we used a temporal deconvolution algorithm (Vogelstein et al., 2010) to estimate neural firing rates from the recorded calcium transients. Following this, visually responsive cells were determined by performing a one-tailed Student's *t-test* between visually evoked responses (4s stimulus) and spontaneous responses (4s gray screen). Only cells with $p < 0.001$ were classified as visually responsive.

Cells were selected for analysis based on Gaussian fits to their responses, and analyzed using the Orientation Selectivity Index. Gaussian fits were computed for the tuning curves using a sum of two Gaussians (cell data) with peaks 180° apart and five parameters: preferred orientation θp, tuning width σ (shared by the two directions), baseline response $R0$, response at the preferred orientation Rp, and response at the null orientation Rn. Optimal parameters were determined with least-squares fits using Matlab's *lsqnonlin* routine. Goodness-of-fit was determined from the R^2 value ($R^2 = 1 - \frac{\sum_i (y-f)^2}{\sum_i (y-\hat{y})^2}$). Only cells with $R^2 > 65\%$ and with significant visually evoked responses ($p<0.001$) were selected for further analysis. The peak response was taken to be the maximum response at preferred orientation. For quantitative analyses, the Orientation Selectivity Index (OSI) was computed by taking the vector average of responses to all orientations, according to the formula:

$$OSI = \frac{\sqrt{(\sum_i R(\theta_i)\sin(2\theta_i))^2 + (\sum_i R(\theta_i)\cos(2\theta_i))^2}}{\sum_i R(\theta_i)}$$

Reliability and Sparseness analysis

Lifetime sparseness provides a measure of the selectivity of a neural response to a visual scene. Cells with a sparseness of 0 imply dense and unselective responses,

while cells with a sparseness of 1 selectively encode only one feature (Willmore et al., 2011). Sparseness was calculated using the equation shown below, where R_i is the response to the i[th] movie frame (averaged over trials) and N is the number of frames:

$$Sparseness = \frac{N - \frac{(\sum_i R_i)^2}{\sum_i R_i^2}}{N - 1}$$

Response reliability measures spiking precision between trials. A highly reliable neuron responds with stereotypical responses on each repetition of the stimulus. As a result, a highly reliable neuron would have a between-trial Pearson's correlation coefficient of 1.

$$Reliability = \frac{2}{T^2 - T} \sum_{i=1}^{T} \sum_{j=1+1}^{T} \rho(R_i, R_j)$$

Following the equation above, reliability was calculated as the average Pearson's correlation of all pairwise combinations of trials (T) for a single movie. Only neurons with significant responses on more than 10 trials were selected for this analysis.

Statistical analysis

All statistical analysis was performed using custom written routines in MATLAB. The Shapiro-Wilk test was first used to determine if the data was normally distributed. All data presented in this paper are non-normally distributed. First, non-parametric one-way ANOVA (Kruskal-Wallis test) was performed to determine statistical significance between experimental conditions. Further comparisons were made relative to WT data using Bonferroni-corrected rank-sum tests. In all statistical tests we set $\alpha = 0.05$. No statistical tests were performed to determine sample size.

Author Contributions

Abhishek Banerjee and Mriganka Sur designed the study. Abhishek Banerjee and Caroline Runyan performed in vivo cell attached electrophysiology, and Rajeev Rikhye analyzed the data. Vincent Breton-Provencher performed in vivo whole cell electrophysiology and analyzed the data. Rajeev Rikhye performed all calcium imaging experiments and analyzed the data. Abhishek Banerjee, Rajeev Rikhye and Mriganka

Sur wrote the manuscript. The manuscript was reformatted and rewritten for this chapter.

Additional References

Cunningham, J.P., Yu, B.M., 2014. Dimensionality reduction for large-scale neural recordings. Nat. Neurosci. 17, 1500–9. doi:10.1038/nn.3776

Goldey, G.J., Roumis, D.K., Glickfeld, L.L., Kerlin, A.M., Reid, R.C., Bonin, V., Schafer, D.P., Andermann, M.L., 2014. Removable cranial windows for long-term imaging in awake mice. Nat. Protoc. 9, 2515–38. doi:10.1038/nprot.2014.165

Singh, A., Lesica, N.A., 2010. Incremental mutual information: a new method for characterizing the strength and dynamics of connections in neuronal circuits. PLoS Comput. Biol. 6, e1001035. doi:10.1371/journal.pcbi.1001035

Vogelstein, J.T., Packer, A.M., Machado, T.A., Sippy, T., Babadi, B., Yuste, R., Paninski, L., 2010. Fast nonnegative deconvolution for spike train inference from population calcium imaging. J. Neurophysiol. 104, 3691–704. doi:10.1152/jn.01073.2009

Willmore, B.D., Mazer, J.A., Gallant, J.L., 2011. Sparse coding in striate and extrastriate visual cortex. J Neurophysiol 105, 2907–2919. doi:10.1152/jn.00594.2010

Chapter 6

Does reliable coding lead to improved visual perception?

Summary

The efficient coding hypothesis asserts that the visual system functions to maximize information about the stimulus by reducing variability and redundancy within the neural code. Numerous studies have reported behavioral uncertainty is the consequence of neuronal variability in sensory areas. Thus, by this token, stimuli that are more reliably processed should be easier to detect. To test this hypothesis, we trained head-fixed mice to discriminate between two natural movies to earn a water reward. First, we sought to elucidate the scene statistics used by mice to discriminate between the movies. By perturbing the spectral statistics of both movies, we found that mice relied on information contained in both the amplitude and the phase spectrum to make their decisions. This is analogous to higher mammals, suggesting that the principles of image recognition are conserved across species. Next, we sought to determine if stimuli that are more reliably processed are better perceived. To this end, we trained mice to categorize movies that were more spectrally similar to a target movie. Activating SST neurons, which made neuronal responses more reliable, improved the accuracy of categorization. In contrast, activating PV neurons decreased reliability and decreased accuracy. Together, these results show that: (1) mice use similar strategies as higher mammals to discriminate natural scenes and (2) reliably processed stimuli are more accurately perceived. Finally, I discuss these results and relate them to the results from other chapters in this thesis.

Highlights

- Mice can rapidly learn how to perform a simple natural movie discrimination task (Go/No Go) and a more complex categorization task.
- Mice can effectively use phase information to discriminate natural movies that have the same amplitude spectrum.
- Improving the reliability of the neural code by activating SST neurons improves the accuracy of movie discrimination.

6.1 Introduction

This thesis has focused on understanding the neural mechanisms of reliable coding. So far, I have shown that the trial-to-trial variability of neurons in V1 varies depending on the spatiotemporal statistics of the stimulus. Also, I have found that the cortex uses inhibitory microcircuits, especially the SST→PV circuit to modulate reliability, by reducing noise in the dendritic input and maximizing spiking output. Additionally, I demonstrate that diseases that chronically disrupt cortical inhibition, such as Rett Syndrome, are characterized by unreliable visual processing, which in turn could explain cognitive deficits observed in these diseases. In this chapter I will focus on understanding how neuronal variability influences visual perception.

The pioneering work of Britten and colleagues showed that neural activity and behavioral choice covary from trial-to-trial even if the stimulus remained constant (Britten et al., 1996). This work implied that the unreliability in the processing of sensory stimuli could be the origin of behavioral variability. Since then, numerous groups have found a strong relationship between neuronal activity in sensory areas and perceptual decision (reviewed by (Nienborg et al., 2012)). This relationship, termed choice probability (CP), depends strongly on the structure of noise correlations between ensembles of neurons (Cumming and Nienborg, 2016; Haefner et al., 2013). Specifically, an influential model proposes high CPs require noise correlations to be stronger for similarly tuned neurons than neurons with different tuning (Shadlen et al., 1996). Supporting this model, several studies have shown that perceptual learning changes the sign of noise correlations in the cortex (Gu et al., 2011; Jeanne et al., 2013). Together, these studies argue that top-down signals, or other neural mechanisms, that alter the structure of noise correlations can help to reduce behavioral variability. For example, in mice, cholinergic signals from the basal forebrain have been shown to reduce noise correlations and improve coding reliability in V1 (Goard and Dan, 2009), which in turn improves the ability of mice to perceive lower contrast visual stimuli (Pinto et al., 2013).

These studies demonstrate that improved reliability can contribute to improved behavioral performance. Studies have shown that natural scenes are more reliably processed than simpler stimuli (reviewed in **Chapter 2**). This observation correlates well with the fact that humans are more efficient at classifying natural scenes compared with simplistic unnatural stimuli (Li et al., 2002) Thus, a central role of efficient processing is to improve detection of complex visual stimuli. However, no study has yet linked the cellular and circuit mechanisms of efficient coding to improved visual perception. Mice are a promising tool to answer this question (Huberman and Niell, 2011). While it has been shown that mice are able to use vision for complex tasks, such as navigation, no study has yet elucidated which statistical features mice use for scene discrimination.

In this chapter, we developed a head-fixed natural movie discrimination task to better understand natural scene perception in mice. We found that water-restricted mice were able to rapidly learn to discriminate between natural scenes to earn a water reward. By perturbing the amplitude spectrum of natural scenes, we discovered that mice relied on phase information to discriminate the target movie from the non-target movie. This finding agrees with studies in humans and monkeys, and it appears that the principles of image recognition are highly conserved across species. Finally, we trained mice on an image categorization task, where mice had to identify movies that had more *target-like* features. We generated distractor movies by blending the phase spectrum of the target and non-target movies at different fractions. Activating SST neurons at the onset of the movie, improved the behavioral performance of mice. In particular, mice were able to more accurately detect *target-like* movies and withhold licking to *non-target-like* movies when SST neurons were activated. Calcium imaging revealed that this improvement in behavioral performance correlated with an increase in coding reliability. Thus, the work presented in this chapter shows that improved coding fidelity does result in better perception of complex natural scenes. Importantly, this work also reinforces the notion that mice are a promising model system to study the neuronal mechanisms underlying natural scene perception.

6.2 Results

Training mice to discriminate between different natural movies

To better understand the relationship between reliable coding and visual perception; we trained mice to discriminate a target movie from a non-target movie in a simple Go/NoGo paradigm[1]. Specifically, water-restricted mice had to report the presence of a target movie by licking a lick-spout to gain a water reward. At the same time, they had to withhold licking when they saw a non-target movie (**Figure 6-1A**). During the training stage, we presented mice with a fixed sequence of target and non-target movies. Specifically, in one block, the target appeared after a run of four non-target movies. We reasoned that minimizing the uncertainty of stimulus appearance would help mice learn the task faster.

We assessed the performance of mice by quantifying the number of correct responses made by the mouse. A correct response is both if the mouse correctly licks to the target movie (Hit) and correctly rejects (CR) a non-target movie (**Figure 6-1B**). From an ethological perspective, the goal of the mouse is to consume a sufficient amount of water. An impulsive mouse would lick to every visual stimulus, which would result in a low CR and an artificially high Hit rate. This mouse would expend more energy, as it would only get rewarded 20% of the time. Alternatively, a mouse that is proficient at the task would lick only when the target movie appears, and would have a high Hit rate and a high CR rate. We considered mice proficient only if they achieved a Hit > 70% and CR > 70% on three consecutive training sessions.

Once mice were proficient, we made the task more difficult by introducing a stochastic trial structure (**Figure 6-1C**). Specifically we randomized the appearance of both target and non-target movies such that target movies appeared with a probability of 20%. As a result, mice were unable to predict the next movie from the trial history. We reasoned that to solve this task mice have to actively use the identity of the visual stimulus to make the correct decision. On average, mice took between 14-21 days to

[1] This work is done in collaboration with Dr. Ming Hu.

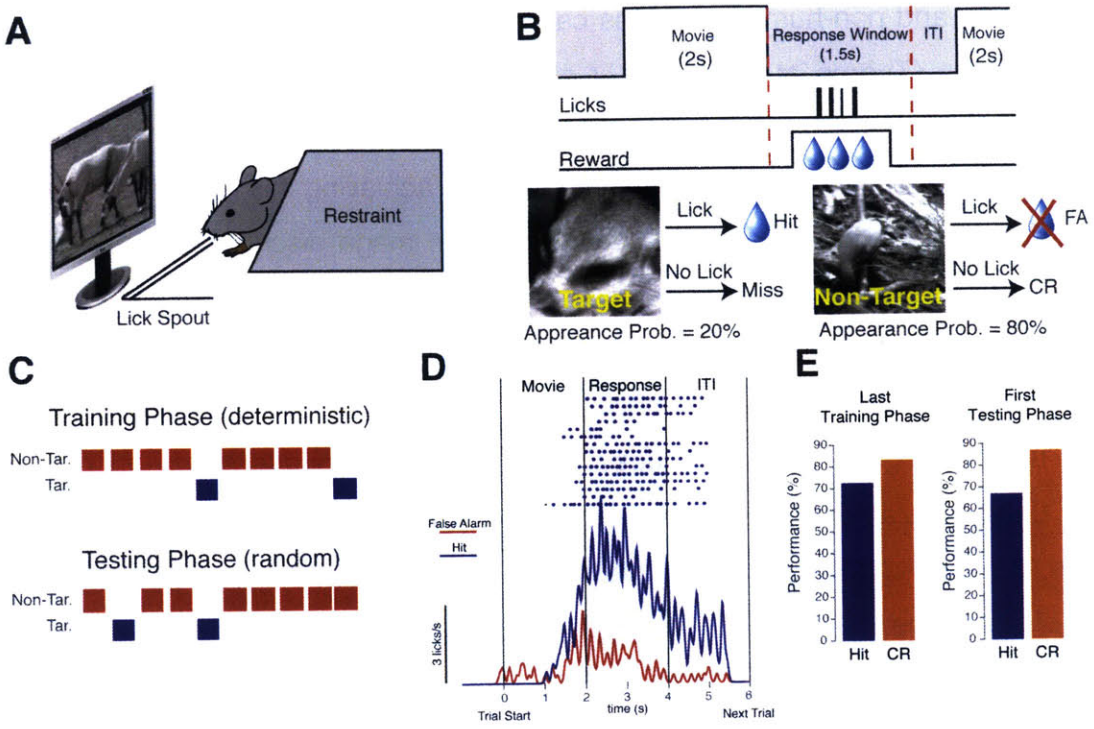

Figure 6-1. Natural movie discrimination task. (A) Cartoon illustrating the experimental set up. Briefly, headfixed mice had to report the presence of a target movie by licking a lick-spout to receive a water reward. **(B)** Structure of the task. A movie was presented for 2s and following the movie mice had a 1.5s response window in which to make their choice. If mice licked to the target movie they received a 10ul water reward. **(C)** Cartoon illustrating the sequence of movies shown in the different phases of training. In the training phase, movies were presented in a fixed sequence, while in the testing phase, the target movie appeared at random. In all cases, the target movie only appeared 20% of the time. **(D)** Quantification of licking behavior of one mouse. Top shows a lick raster plot on 15 target movie trials. Solid lines show trial-averaged response. (E) Performance of one mouse during the last training phase, and first testing phase. CR, correct reject,

become proficient at the task (9/11 mice). We observed that once proficient at the task, mice began licking approximately 1s into the target movie and quickly increased their rate of licking in once the movie ended (example from one mouse, **Figure 6-1D**). Additionally, proficient mice were able to maintain a high hit rate even when we switched to the much harder stochastic testing phase task (Training: Hit = 72.5%, CR = 83.1%, Testing: Hit = 67.1%, CR = 86.9%, **Figure 6-1E**). Since no other cues were used in this task, this performance indicates that head-fixed mice are able to rapidly learn to discriminate between different movies.

Mice use phase differences to discriminate between natural movies

Both humans and non-human primates can rapidly and effortlessly categorize complex natural scenes (Thorpe et al., 1996, Li et al., 2002). Understanding the mechanisms behind this ability to rapidly recognize and categorize complex scenes remains a central goal in systems neuroscience. A critical first step in achieving this goal is elucidating the natural scene statistics the visual system uses for image discrimination. A commonly held theory is that the visual system relies on phase information to discriminate between different scenes (Wichmann et al., 2010, see references in **Chapter 2**). The Fourier phase spectrum contains salient image features, such as edges and contours, which allows us to recognize shapes. Indeed, randomizing the phase spectrum renders an image unrecognizable (Wichman et al., 2005). Another theory posits that the visual system uses low-level spectral information contained in the amplitude spectrum to form a gist of an image (Bar, 2004). This spectral categorization hypothesis claims that the visual system uses gist to rapidly categorize scenes into different categories (Oliva and Torralba, 2006). In this section, we sought to determine if these ideas apply to mice. Establishing visual recognition analogs in mice is important, as it will allow us to leverage the genetic tractability of mice to dissect the circuits underlying visual perception, as we have done in this thesis.

We examined whether mice also use subtle differences in amplitude or phase spectra to discriminate between the target and non-target movies (**Figure 6-2A**). We first sought to determine which statistics mice used to recognize natural movies. We first trained mice (n = 1 at time of writing) to discriminate between the original target and non-target movies using the deterministic task structure. Next, using noise masking (see **Chapter 3**), we replaced the amplitude spectrum of a target movie with a *K1* noise mask but left its phase spectrum intact. Thus, in this version of the task, we tested the ability of mice to discriminate between an original non-target and a *K1* target (**Figure 6-2B**).

If mice learned to recognize the target movie using its phase structure, then replacing the target with a *K1* version should not impact performance, as only the amplitude spectrum has changed. Instead, if mice learned to recognize the target movie with its amplitude structure, then there should be a marked decrease in performance, as the *K1* target should be perceived as an entirely new movie. Also, if mice were using the

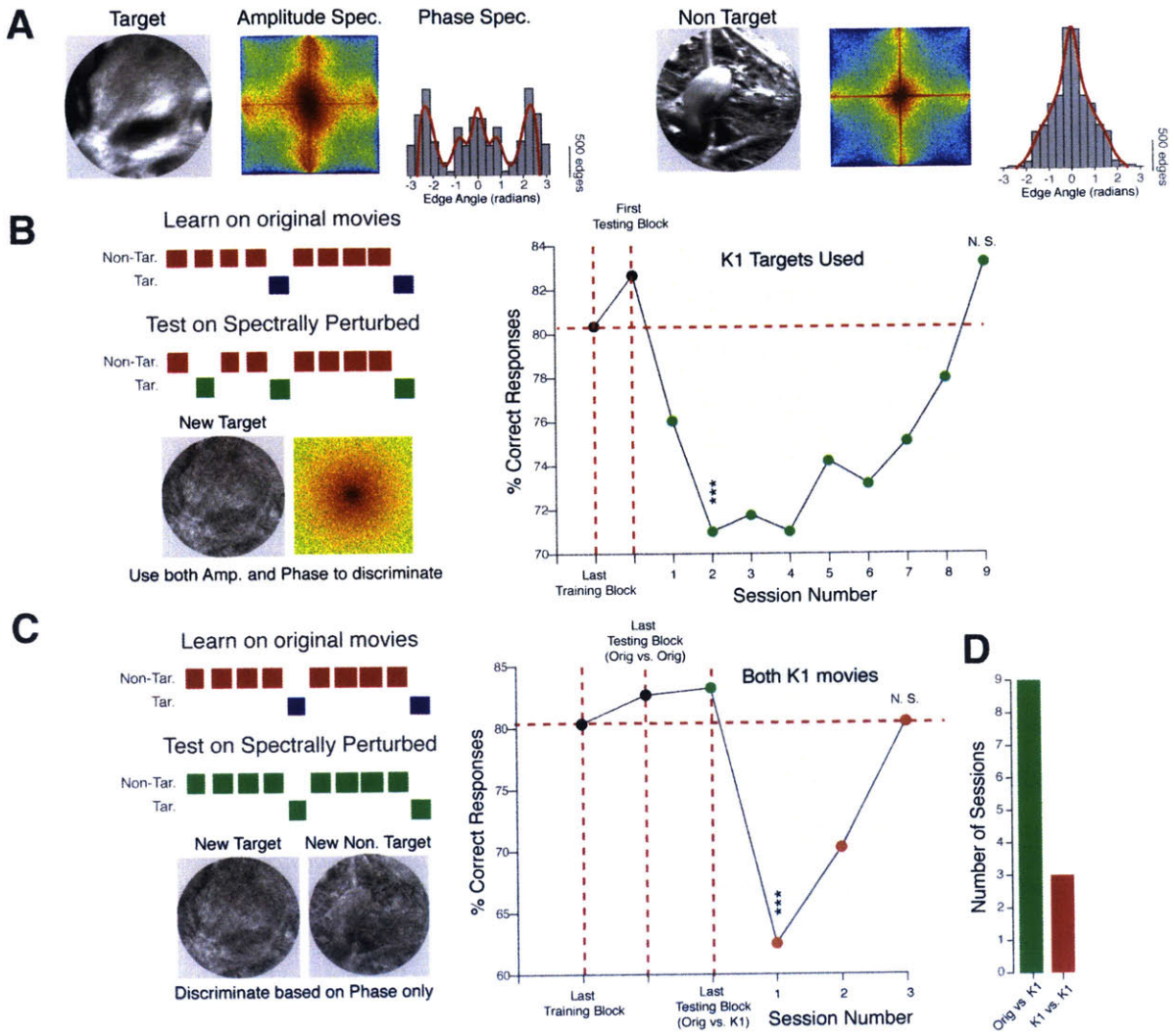

Figure 6-2. Mice use amplitude and phase information to discriminate between different natural scenes. (A) Quantification of Amplitude spectrum and Phase spectrum of the Target and non-Target movies. The target movie has a broader distribution of edges, whereas the non-target movie is biased to vertical edges. **(B)** Image recognition task. Mice take a long time to adjust to a perturbation of the amplitude spectrum of the target movie. (C) Mice use can rapidly learn how to use phase information to discriminate between scenes. (D) Comparison of learning time for the two different tasks. Data from 1 mouse.

"if not target, then lick" strategy, then this perturbation should also not affect performance as the non-target movie is unchanged.

We noticed a significant decrease in the percentage of correct responses made by the mouse in the first session using the *K1* target movie (82% vs. 72%). This was due to an increase in the number of False Alarms and a decrease in the number of Hits made by the mouse, suggesting that the mouse perceived the K1 target movie as a new

stimulus and, consequentially, was confused by the switch. Notably, it took this mouse 9 sessions (consecutive days) for performance to return back its original expert level (**Figure 6-2B**). When compared to the overall learning time of 16 sessions from naïve state, this long learning time suggests that the mouse had to relearn the association between the *K1* target movie and the water reward. This implies that the mouse perceived the *K1* target movie as a new stimulus even though the phase content remained unchanged. Taken together, these preliminary results suggest that mice do not only rely on phase information when discriminating between different images. Instead, mice appear to rely on spectral information contained in the amplitude spectrum to recognize different scenes. This might serve an ethological role as broad, and blurry shapes, which contain little phase information, are perceived as predators to mice and robustly activate fear responses (Yilmaz and Meister, 2013).

Having established that mice rely on amplitude information to recognize different scenes, we next sought to determine which statistics mice used to discriminate between different natural movies. To do so, we replaced the amplitude spectra of both the target and non-target movies with *K1* noise masks (**Figure 6-2C**). Since we equalized the amplitude spectra for both movies, mice were forced to use only information contained in the phase spectra to discriminate between the two movies. Notably, as shown in **Figure 6-2A**, the target movie has more broadly distributed edges than the non-target movie. Thus, we sought to determine if mice could use differences in the edge distributions to discriminate between the movies.

Although performance dropped in the first session (86% vs. 63%), the mouse was able to reach expert levels (80%) only after two sessions. Thus, this mouse was successfully able to learn to withhold licking to a *K1* non-target movie, which had the same amplitude information as the target movie. This result suggests that mice can use phase information to discriminate movies, when differences in amplitude information is uninformative. Importantly this result also suggests that, even though mice have broader orientation selectivity and randomly organized orientation maps compared to cats or monkeys (see **Chapter 2**), their visual systems are capable of using subtle differences in phase information to make goal-directed actions. More experiments are planned to

concretely elucidate how mice use amplitude and phase information to discriminate between different movies.

Activating SST neurons improves discriminability of complex movies

Our results in **Chapter 4** show that reliable coding in visual cortex is modulated by the interplay between SST and PV neurons. Specifically, we propose that the cortex uses the SST→PV inhibitory circuit to maximize SNR, and improve coding fidelity. How does this improved coding fidelity, and improved reliability, relate to better visual perception? As we have discussed in **Chapter 2**, visual processing is hierarchical, with unique stimulus transformations taking place in distinct brain areas. Given the multi-area nature of visual processing, reliable spiking in V1 has to propagate to numerous higher visual and decision-making areas in the cortex before the mouse is able to make a decision. A recent computational study showed that feed-forward processing amplifies neuronal variability (Moldakarimov et al., 2015). This means that unreliability in the V1 code can propagate up the visual hierarchy and prevent the formation of goal-directed motor plans (Britten et al., 1992, Nienborg et al., 2012). Thus, refining the neural code at the earliest stage would be beneficial to visual perception.

To test this hypothesis, we trained mice expressing ChR2 either in SST or PV neurons (SST-cre::Ai32 and PV-cre::Ai32) to discriminate between two natural movies, as shown in **Figure 6-1**. Once mice became proficient in at associating the target movie with a water reward, we introduced a new set of visual stimuli, in which we contaminated the phase spectrum of the target movie with different fractions of the non-target movie (**Figure 6-3B**). In particular, we blended the phase spectrum of the target movie with the phase spectrum of the non-target movie at different fractions (see Appendix 6A). We also equalized the amplitude spectra of all movies to be the mean spectrum of the target and non-target movies. We established that mice use phase information to discriminate between the target and non-target movies (**Figure 6-2**). Thus, by blending phase information, we created an ensemble of stimuli with increasing uncertainty levels. Specifically, since the target movie is most behaviorally relevant, its phase

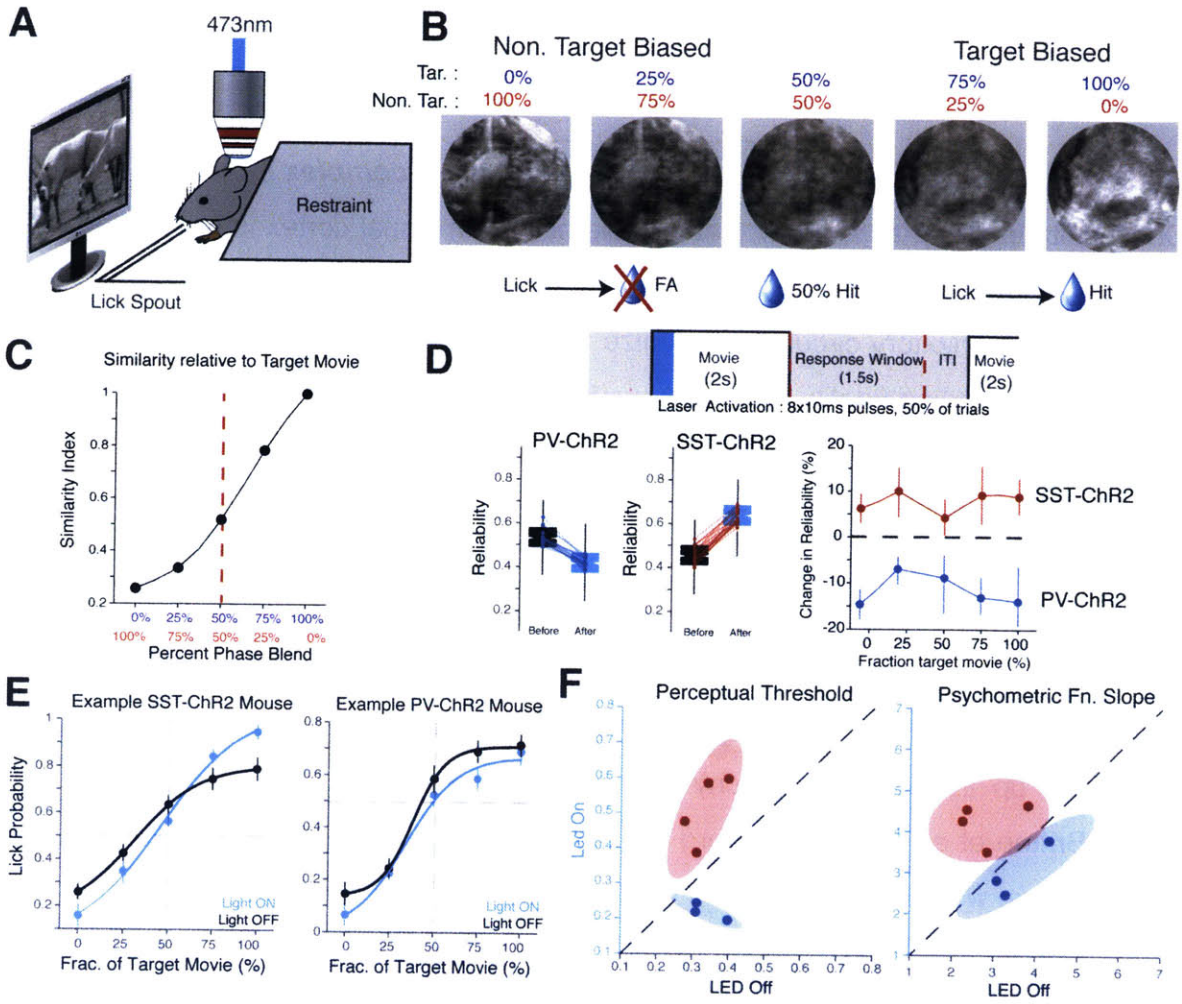

Figure 6-3. Activating SST neurons improves movie discriminability. (A) Cartoon of experimental setup. **(B)** Example frames from the stimulus used in this experiment. These movies were created by mixing the phase spectrum of the target movie with that of the non-target movie at different fractions. Non-target-biased movies are not rewarded, while target-biased movies are rewarded. **(C)** Quantification of the similarity of these movies relative to the target movie (SSIM, see Appendix 3A). Psychometric functions should follow this trend. **(D)** *Top,* trial structure showing the timing of the laser activation (blue) in relation to stimulus onset. *Bottom,* Box-whisker plots showing that activation of PV neurons at onset decreases reliability, while activation of SST neurons increases reliability. Data same as Figure 4-5. *Right,* Change in pyramidal cell reliability induced by the same stimulation paradigm in passively viewing mice. SST activation improves reliability of all blended stimuli. Data from 3 mice each. **(E)** Example psychometric curves. Error bars computed across sessions, 12 sessions for both mice. **(F)** Scatter plots showing the effect of SST and PV activation on Perceptual threshold and psychometric function slope. Shaded area represent 95% CI. Data from 4 SST and 3 PV mice.

spectrum can be considered to a *signal*, which the mouse must extract. At the same time, mice have to ignore the non-target movie, so its phase spectrum can be considered to be *noise*. Thus, in this paradigm, mice now have to identify and categorize movies with more *signal* than *noise*. Notably, these phase-blended movies

had a monotonically increasing similarity index relative to the original movie (**Figure 6-3C**). Thus, we expect the psychometric functions to closely match this SSIM plot.

To determine the effect that SST and PV neuron activation had on the ability of mice to extract *signal* from *noise* in the visual stimulus, we applied a blue (473 nm) laser at random on 50% of the trials, coincident with the onset of the stimulus (**Figure 6-3D**). Our calcium imaging experiments revealed that activating PV and SST neurons at stimulus onset had the largest effect on reliability, with PV activation suppressing reliability and SST activation increasing reliability (**Figure 4-5** and **Figure 6-3D**). Repeating these experiments using phase-blended movies instead also confirmed these results. In particular, activating SST neurons improved the reliability ($p<0.001$, permutation test) for all phase-blended movies, whereas activating PV neurons decreased reliability ($p<0.001$, permutation test, **Figure 6-3D**). Based on these experiments, we expect that, following SST activation, V1 neurons should be better at processing all stimuli. As a result, the mouse should lick more to the target movies and less to the non-target movies, which would in turn sharpen the psychometric curve (increased slope and higher threshold). In contrast, when PV neurons are activated, V1 should be worse at processing all stimuli. This would in turn result in a decrease in lick rates for all movies and a scaling down of the psychometric function (no change in slope and decrease in threshold).

Mice were able to learn this task within 10-12 days (5/7 mice). As expected, mice licked more to the target-biased movies and less to the non-target-biased movies, resulting in psychometric functions as shown in **Figure 6-3E**. Interestingly, most mice showed a bias to lick more to the 50-50 blended movies even though this movie was only 52% similar to the target movie. In agreement with our hypothesis, activating SST neurons resulted in a sharper and more step-like psychometric curve (see example in **Figure 6-3E,** data from 520 light-off and 510 light-on trials, $p = 0.041$, ANOVA for PV-ChR2 and $p = 0.044$ for SST-ChR2). In contrast, activating PV neurons scaled down responses to the target stimuli. We defined the perceptual threshold as the point of inflexion of the psychometric function. Physically, the perceptual threshold indicates the phase fraction that is at the threshold for discrimination. Interestingly, we found that

activating SST neurons increased the perceptual threshold and increased the slope of the psychometric function ($p < 0.001$, two-way ANOVA (slope x threshold), relative to PV neurons, **Figure 6-3F**). This means that mice are better able to identify movies with more *signal* than *noise* following SST activation. In stark contrast, activating PV neurons did not significantly alter either the threshold or the slope.

Together, these results suggest that SST activation not only improves the reliability of the neural code, but also improves the ability of mice to efficiently discriminate ambiguous visual stimuli.

6.3 Discussion

In summary, our behavioral results have revealed several interesting insights into visual processing in mice. First, head-fixed mice are able to learn how to discriminate between complex natural scenes within a short period of time, suggesting that the mouse visual system is also adapted to efficiently process information from the natural environment. Also, by perturbing the spectral statistics of movie, we discovered that mice rely on information contained in both the amplitude and the phase spectrum to recognize and discriminate between different scenes. Importantly, this result demonstrates that mice might also use similar strategies as higher mammals (cats, non-human primates, etc.) to categorize different scenes. Finally, by combining behavioral testing with optical manipulation, we found that improving the reliability of the neuronal code in V1, via SST activation, does improve the ability of mice to discriminate between complex natural movies. Thus, further studies can build on our behavioral paradigm to probe more complex circuitry underlying visual processing.

Role of amplitude and phase spectrum in image discrimination

In **Chapter 3** of this thesis, we investigated the role of the amplitude spectrum in reliable coding. The amplitude spectrum contains information about the distribution of spatial frequencies within an image. We found that neurons in mouse V1 respond more reliably to movies with strong low SF power. The low SF content in natural scenes is mainly large, coarse-grained objects, which can be used for basic scene recognition and motion discrimination (Bar, 2004). Thus, our finding that mice respond robustly and reliably to strongly correlated movies suggests that mice primarily use their vision to extract coarse grained information from their visual environments, presumably to guide navigation in low light conditions or to avoid aerial predators (see references within (Zoccolan, 2015)).

Interestingly, our behavioral results (**Figure 6-2**) agree with psychophysical studies, which show that the human visual system also uses information in the low SF bands to rapidly and accurately discriminate between different natural scenes (Gaspar

and Rousselet, 2009; McCotter et al., 2005). These suggest that the algorithms for scene discrimination are highly conserved between species. In particular, because mice use their vision in low-light conditions, they have to be able to discriminate between distant landmarks and have to rapidly identify threatening stimuli with limited structural information. In these conditions, low SF information would be dominant. Thus, the visual system of mice has to be optimized to detect changes in low SF content.

We also demonstrate that attenuating these bands, even to a small degree (~10 dB in the case of *K1* movies), strongly degrades processing in V1 by increasing trial to trial variability (see **Chapter 3**). This reduction in coding fidelity could explain why behavioral performance significantly drops when the amplitude spectrum of the target movie is replaced with *K1* noise. In particular, this result demonstrates that a stimulus-induced change in neuronal variability can also introduce large behavioral uncertainty (Britten et al., 1996; Renart and Machens, 2014). Additionally, *K1* movies elicit higher noise correlations and activate smaller neural ensembles than the original movie. This restructuring of interneuronal correlations could also explain the decrease in behavioral performance. Surprisingly, mice take approximately 9 sessions to re-associate the *K1* target movie with a water reward. It is possible during this time frame; learning related signals could reduce noise correlations. Further experiments are required to determine how the neural network changes during perceptual learning.

Additionally, we have discovered that, when both target and non-target movies have the same amplitude spectrum, mice rely on phase information for discrimination (**Figure 6-2**). If mice used only the amplitude spectrum, then performance should have also decreased. In contrast, mice rapidly learned how to discriminate *K1* target movies from *K1* non-target movies. The importance of phase information in image discrimination has been well documented (Wichmann and Gegenfurtner, 2010). This is mainly due to the fact that the phase spectrum contains information about edges and other salient features. Importantly, this result suggests that, even though mice lack organized orientation maps, they are able to detect subtle changes in edge structure within complex natural scenes.

Together, these results suggest that mice use information contained within the amplitude and phase spectra to discriminate between different movies. It is likely that mice use this information is by making a frequency template by comparing spatial frequency content at different screen positions. Similar strategies have been observed in the rat and monkeys (Ohayon et al., 2012; Vermaercke and Op de Beeck, 2012). However, more experiments using better-designed stimuli are required to elucidate how mice use visual information for scene discrimination.

Reliable coding and visual perception

A central aim of this thesis is to uncover the neural mechanisms responsible for reliable coding. In **Chapters 4** and **5**, we demonstrated that reliable, and intact inhibition, is required for reliable coding. In particular, we discovered an inhibitory microcircuit that can be used by the cortex to improve coding fidelity. In this chapter we sought to determine the impact that improved reliability had on visual perception. Indeed, we found that activating SST neurons improved the reliability of the neural code, reduced noise correlations and also improved the ability of mice to accurately categorize phase-blended movies. This result strongly suggests that reliable coding is in fact a hallmark of more efficient visual processing, Specifically, by minimizing noise (i.e. spikes not explained by the stimulus), SST neurons can improve the fidelity of coding for all stimuli. This in turn improves stimulus discriminability and results in better perceptual performance. Studies in piriform cortex have shown SST activation can also improve odor discrimination (Sturgill and Isaacson, 2015). Together, these results suggest that improving coding fidelity even in the earliest stage of cortical processing can improve behavioral performance. Interestingly, this means that neural codes are uncorrupted as they pass from one cortical stage to another. Further experiments should be directed at studying how the reliability of visual information changes as it propagates up the visual hierarchy.

In stark contrast to SST neurons, activating PV neurons reduced Hit (lick) rates to all stimuli. This suggests that mice were more uncertain on the identity of the stimulus on trials when PV neurons were optically activated. This decrease in behavioral

performance correlates with decreased reliability. Surprisingly, this is result disagrees with a recent study that found that activating PV neurons improved orientation discriminability in head-fixed mice (Lee et al., 2012). Notably, this study activated PV neurons for the entire duration of the stimulus, wheras we only increase PV inhibition at stimulus onset. This long duration PV activation has been shown to improve neuronal reliability (Zhu et al., 2015), possibly due to different network-dependent mechanisms (Seybold et al., 2015). Thus, despite differences in the result, both studies show that increasing coding reliability in V1 leads to improved perceptual performance.

In **Chapter 4**, we established that SST neurons are crucial mediators of reliable coding. Our experiments revealed that, by strongly suppressing distal dendrites and PV neurons, SST neurons control both synaptic integration in the dendrites and temporal integration in the soma. This mechanism controls the temporal precision of spiking at the soma primarily by minimizing noise in the inputs. Interestingly, learning suppresses SST neuron activity (Makino and Komiyama, 2015) and reduces their axonal arborization (Chen et al., 2015) within the cortex. In addition, suppressing SST neurons during learning stabilizes dendritic spines allowing enhanced synaptic transmission (Chen et al., 2015). When taken together with our results, these findings suggest that neuronal variability is higher during learning. Indeed, studies have shown that variability during learning allows the network to explore all possible states and select the best possible outcome (Fiete and Seung, 2006; Rokni et al., 2007). Thus, once the mouse learns to perform the task, SST neurons should become more active to minimize variability within the neural ensembles. This would ensure that the same neural repertoire is reliably activated from trial-to-trial.

We propose that both stimulus features and top-down signals from other cortical areas can recruit these inhibitory mechanisms to dynamically modulate reliability. In particular, as we have shown, perturbing features of the stimulus, such as the amplitude spectrum, can introduce unreliability in the neural code and can also increase behavioral variability. Top-down signals carrying arousal and attention signals can also modulate reliability. For example, a recent study showed that cholinergic modulation of V1 improved perceptual thresholds by increasing firing rates and decreasing correlated

noise (Pinto et al., 2013). In addition, top down glutamatergic projections from the cingulate cortex also improve the detection of low contrast stimuli (Zhang et al., 2014). It is possible these signals work through the SST→PV circuit to improve reliability, however this still remains unknown. Future experiments should be aimed at studying how top-down signals alter SST neuron activity during learning and once the mouse is proficient at the task.

6.4 Conclusion

The overall aim of this thesis is to understand the neural mechanisms underlying reliable coding. In **Chapter 3**, I investigated how stimulus statistics modulated neural responses in V1. I found that low SF information contained in the amplitude spectrum is necessary for reliable responses. This is because spatial correlations in the stimulus altered the structure of neuronal correlations. Specifically, movies with strong correlations recruited ensembles of neurons that worked cooperatively to reduce noise correlations. Thus, the work presented in this chapter revealed that the conditions required for reliable coding are: (1) stimuli with strong spatiotemporal correlations and (2) strong activation of correlated ensembles of neurons.

In **Chapter 4**, I investigated the inhibitory mechanisms that allow neurons to transform stochastic inputs into reliable outputs. To do so, I developed a novel line of transgenic mice that allowed us to study the interactions between SST and PV neurons within the same field of view. Our experiments revealed that the SST→PV circuit is responsible for modulating coding reliability by changing the signal-to-noise ratio (SNR). Specifically, by inhibiting dendrites, SST neurons can prevent noisy synaptic inputs from being integrated. At the same time, by inhibiting PV neurons, SST neurons can modulate temporal window over which spiking occur. In doing so, the SST→PV circuit allows SST neurons to regulate noise in the input and signal in the output. These novel results show that the cortex can use the SST→PV circuit dynamically modulate coding reliability by altering the dynamics of SST neurons. Specifically, we found that activating VIP neurons can reduce the reliability of visually evoked responses by suppressing PV neurons. In **Chapter 5**, we show that chronically perturbing cortical inhibition, either with DREADD or by deleting the MeCP2 gene, strongly reduces reliability and increases correlated noise. Together, the work presented in this chapter shows that: (1) reliable inhibition from SST neurons and (2) the SST→PV circuit are crucial components of the reliable coding circuit.

Finally in **Chapter 6**, I show that reliable coding is indeed necessary for accurate visual perception. Increasing the reliability of the neural code, by activating SST

neurons, improved the ability of mice to categorize complex visual stimuli. These results show that increased reliability is indeed a hallmark of efficient coding as it allows mice to better perceive different visual stimuli.

Since 1961, Barlow's efficient coding hypothesis has asserted that sensory systems have evolved to maximize information transmitted to the brain from the natural environment. Efficient coding is achieved by reducing information-limiting variability and redundancy in the neuronal code. By leveraging the genetic tractability of mice, the work done in this thesis has revealed several intricate neural mechanisms required for reducing neuronal variability, and the impact this has on visual perception. More importantly, the tools that have been developed in this thesis can be used to further elucidate the cells and circuits underlying more complex visual perception, such as object recognition.

6.5 References

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

Chen, S.X., Kim, A.N., Peters, A.J., Komiyama, T., 2015. Subtype-specific plasticity of inhibitory circuits in motor cortex during motor learning. Nat. Neurosci. 18, 1109–15. doi:10.1038/nn.4049

Cumming, B.G., Nienborg, H., 2016. Feedforward and feedback sources of choice probability in neural population responses. Curr. Opin. Neurobiol. 37, 126–32. doi:10.1016/j.conb.2016.01.009

Fiete, I.R., Seung, H.S., 2006. Gradient learning in spiking neural networks by dynamic perturbation of conductances. Phys. Rev. Lett. 97, 048104. doi:10.1103/PhysRevLett.97.048104

Gaspar, C.M., Rousselet, G.A., 2009. How do amplitude spectra influence rapid animal detection? Vision Res. 49, 3001–12. doi:10.1016/j.visres.2009.09.021

Goard, M., Dan, Y., 2009. Basal forebrain activation enhances cortical coding of natural scenes. Nat. Neurosci. 12, 1444–9. doi:10.1038/nn.2402

Gu, Y., Liu, S., Fetsch, C.R., Yang, Y., Fok, S., Sunkara, A., DeAngelis, G.C., Angelaki, D.E., 2011. Perceptual learning reduces interneuronal correlations in macaque visual cortex. Neuron 71, 750–61. doi:10.1016/j.neuron.2011.06.015

Guo, Z. V, Hires, S.A., Li, N., O'Connor, D.H., Komiyama, T., Ophir, E., Huber, D., Bonardi, C., Morandell, K., Gutnisky, D., Peron, S., Xu, N., Cox, J., Svoboda, K., 2014. Procedures for behavioral experiments in head-fixed mice. PLoS One 9, e88678. doi:10.1371/journal.pone.0088678

Haefner, R.M., Gerwinn, S., Macke, J.H., Bethge, M., 2013. Inferring decoding strategies from choice probabilities in the presence of correlated variability. Nat. Neurosci. 16, 235–42. doi:10.1038/nn.3309

Huberman, A.D., Niell, C.M., 2011. What can mice tell us about how vision works? Trends Neurosci. 34, 464–73. doi:10.1016/j.tins.2011.07.002

Jeanne, J.M., Sharpee, T.O., Gentner, T.Q., 2013. Associative learning enhances population coding by inverting interneuronal correlation patterns. Neuron 78, 352–63. doi:10.1016/j.neuron.2013.02.023

Lee, S.-H., Kwan, A.C., Zhang, S., Phoumthipphavong, V., Flannery, J.G., Masmanidis, S.C., Taniguchi, H., Huang, Z.J., Zhang, F., Boyden, E.S., Deisseroth, K., Dan, Y., 2012. Activation of specific interneurons improves V1 feature selectivity and visual perception. Nature 488, 379–83. doi:10.1038/nature11312

Li, F.F., VanRullen, R., Koch, C., Perona, P., 2002. Rapid natural scene categorization in the near absence of attention. Proc. Natl. Acad. Sci. U. S. A. 99, 9596–601. doi:10.1073/pnas.092277599

Makino, H., Komiyama, T., 2015. Learning enhances the relative impact of top-down processing in the visual cortex. Nat. Neurosci. 18, 1116–1122. doi:10.1038/nn.4061

McCotter, M., Gosselin, F., Sowden, P., Schyns, P., 2005. The use of visual information in natural scenes. Vis. cogn. 12, 938–953. doi:10.1080/13506280444000599

Nienborg, H., Cohen, M.R., Cumming, B.G., 2012. Decision-related activity in sensory neurons: correlations among neurons and with behavior. Annu. Rev. Neurosci. 35, 463–83. doi:10.1146/annurev-neuro-062111-150403

Ohayon, S., Freiwald, W.A., Tsao, D.Y., 2012. What makes a cell face selective? The importance of contrast. Neuron 74, 567–81. doi:10.1016/j.neuron.2012.03.024

Pinto, L., Goard, M.J., Estandian, D., Xu, M., Kwan, A.C., Lee, S.-H., Harrison, T.C., Feng, G., Dan, Y., 2013. Fast modulation of visual perception by basal forebrain cholinergic neurons. Nat. Neurosci. 16, 1857–63. doi:10.1038/nn.3552

Renart, A., Machens, C.K., 2014. Variability in neural activity and behavior. Curr. Opin. Neurobiol. 25, 211–220. doi:10.1016/j.conb.2014.02.013

Rokni, U., Richardson, A.G., Bizzi, E., Seung, H.S., 2007. Motor learning with unstable neural representations. Neuron 54, 653–66. doi:10.1016/j.neuron.2007.04.030

Seybold, B.A., Phillips, E.A.K., Schreiner, C.E., Hasenstaub, A.R., 2015. Inhibitory Actions Unified by Network Integration. Neuron 87, 1181–92. doi:10.1016/j.neuron.2015.09.013

Sturgill, J.F., Isaacson, J.S., 2015. Somatostatin cells regulate sensory response fidelity via subtractive inhibition in olfactory cortex. Nat. Neurosci. 18, 531–5. doi:10.1038/nn.3971

Vermaercke, B., Op de Beeck, H.P., 2012. A multivariate approach reveals the behavioral templates underlying visual discrimination in rats. Curr. Biol. 22, 50–5. doi:10.1016/j.cub.2011.11.041

Wichmann, F.A., Gegenfurtner, K.R., 2010. Animal detection in natural scenes : Critical features revisited 10, 1–27. doi:10.1167/10.4.6.Introduction

Willenbockel, V., Sadr, J., Fiset, D., Horne, G.O., Gosselin, F., Tanaka, J.W., 2010. Controlling low-level image properties: the SHINE toolbox. Behav. Res. Methods 42, 671–84. doi:10.3758/BRM.42.3.671

Zhang, S., Xu, M., Kamigaki, T., Hoang Do, J.P., Chang, W.C., Jenvay, S., Miyamichi, K., Luo, L., Dan, Y., 2014. Selective attention. Long-range and local circuits for top-down modulation of visual cortex processing. Science (80-.). 345, 660–665. doi:10.1126/science.1254126

Zhu, Y., Qiao, W., Liu, K., Zhong, H., Yao, H., 2015. Control of response reliability by parvalbumin-expressing interneurons in visual cortex. Nat. Commun. 6, 1–11. doi:10.1038/ncomms7802

Zoccolan, D., 2015. Invariant visual object recognition and shape processing in rats. Behav. Brain Res. 285, 10–33. doi:10.1016/j.bbr.2014.12.053

Appendix 6A – Experimental procedures

Behavior

Water restriction and training schedule

As described in Appendix 4A, PV-cre::Ai32 and SST-cre::Ai32 mice were implanted with a cranial window and a headpost. Mice were allowed to recover from surgery for 1-2 weeks before beginning water restriction. Specifically mice were placed on a water restriction schedule in which they received a minimum daily amount (40µL water per gram, daily mouse weight) (Guo et al., 2014). Weight, behavior and general condition, were monitored for signs of dehydration by veterinary staff.

Once stable body weight was reached, mice began training on the natural movie discrimination task. Typically mice were trained daily (7 days a week) in one-hour sessions. On the first two-three training days, mice were head-fixed and given water reward in order to acclimatize to the both head fixation and drinking from a waterspout. We trained mice in stages. In the first stage (Stage 0), mice were taught how to associate a water reward with the target movie. Specifically a 5µL water reward was given to mice 2s into the target movie. Once mice achieved a stable hit rate (HR) of 90% over three consecutive sessions, mice were graduated to Stage 1. In Stage1, the non-target movie was introduced (80% of trials) and mice were trained to withhold licking. False alarms were punished with a 200ms white noise burst. In this stage, mice were auto-rewarded on every target movie trial. Once correct reject rate (CR) reached 80% mice were graduated to the next stage. In Stages 2-3 the probability of auto-reward was gradually reduced from 50% (Stage 2) to 25% (Stage3). Mice were promoted to Stage 4 only once they were able to achieve a CR and HR of 80%. In Stage 4, target and non-target movies were played in a deterministic sequence but no auto-reward was given. Finally in Stage 5, mice were trained on the randomized sequence. Once mice reached a HR and CR > 70% on the randomized sequence they were tested on the Target discrimination version of the task. Typically, this training paradigm lasted for 3-4 weeks with mice completing 400-600 trials per day.

Target categorization task

In this version of the task, three more movies were introduced and the mice had to determine which movies from this ensemble had more target-like features. Specifically, to create these movies we blended the phase spectrum of the non-target movie with the phase spectrum of the target movie according the formula

$$Phase_{new} = \alpha Phase_{Tar} + (1 - \alpha) Phase_{Non-Tar}$$

with $\alpha = [0,0.25,0.5,0.75,1]$. Also we replaced the amplitude spectrum of all movies with the mean amplitude spectrum of the non-target and target movie.

$$Amp_{new} = 0.5(Amp_{Tar} + Amp_{Non-Tar})$$

All movies were corrected to have the same mean luminance and contrast using the SHINE toolbox (Willenbockel et al., 2010). Movies with $\alpha = [0.75,1]$ were treated as target movies and presented with a probability of 0.4, whereas movies with $\alpha = [0,0.25,0.5]$ were treated as non-target movies and were presented with a probability 0f 0.6. Non-target movies were not rewarded except for $\alpha = 0.5$, which was rewarded on 50% of the trials. On average mice completed 600-800 trials per day.

Once mice were proficient at this task, we optogenetically activated ChR2 using a fiber-coupled 470nm LED source on 50% of the trials. The average power at the tip of the fiber was 8mW, and the power spectral density was similar to what we used for physiology experiments.

To determine the effect that laser activation had on task performance, we compared psychometric functions on Light-Off trials with Light-On trials. We quantified parameters like threshold and slope, by fitting sigmoidal functions to the psychometric functions on each session. Only trials with fit quality >85% were kept (12/18 session).

The End

This page intentionally left blank