THE UNIVERSITY OF CHICAGO

MOTION TUNING TO DRIFTING GRATINGS AND WATER FLOW

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
AND
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BIOPHYSICAL SCIENCES

BY
VAUGHN SPURRIER

CHICAGO, ILLINOIS
DECEMBER 2019

This work is dedicated to the countless individuals who have supported me during my graduate career. I thank my family, friends, mentors and peers for the investment they have made in me. I hope to return a good dividend.

Revelation 5:13

TABLE OF CONTENTS

LIST OF FIGURES vi

ABSTRACT vii

1 GENERAL INTRODUCTION 1
1.1 Controllable, parametric stimuli and V1 response 1
1.2 Increased reliability and sparsity in V1 response to naturalistic stimuli 4
1.3 Experimental setups for interrogating vision 7

2 ENGINEERING A MICROSCOPE FOR IMAGING AWAKE MICE 10
2.1 A linear treadmill to measure ambulatory velocity 10
2.2 An infrared eye-tracking camera 12

3 MODELS FOR MOUSE VISION 17
3.1 Modeling at various levels of complexity 17
3.2 Normalization computations 19
3.3 Hierarchical modeling of the ventral stream 20

4 CLASSIC MODELS PREDICT MURINE PRIMARY VISUAL CORTICAL MOTION TUNING TO NATURALISTIC STIMULUS 23
4.1 Introduction 23
4.2 Methods 25
4.3 Results 31
4.4 Discussion 46

5 GENERAL DISCUSSION 48

REFERENCES 51

LIST OF FIGURES

2.1 Treadmill and eye-tracking illumination. 11

4.1 Stimuli characteristics. 32
4.2 Expectations informed by modeling. 33
4.3 Preferred direction of modeling. 34
4.4 Experimental design. 35
4.5 Visual responsiveness to G and W. 36
4.6 Tuning strength to G and W. 38
4.7 Neurons OT to G do not have a shared motion preference to W, while neurons DT to G have a shared motion preference to W. 40
4.8 Trial variance effect on confidence in preferred angles. 41
4.9 Characterization of water texture stimulus. 42
4.10 Difference in preferred angle similarity does not depend on natural motion. . . . 44
4.11 Trial variance explained and response similarity as a function of distance. 45

ABSTRACT

In murine primary visual cortex, simple cells extract edge orientation from visual inputs, while complex cells extract motion direction [45]. These properties are classically measured using synthetic psychophysical visual stimuli, typically drifting gratings. Natural movies contain complicated distributions of edge orientation and motion statistics, vastly differing from the edge and motion content of drifting gratings. It remains unclear the extent to which single neuron properties measured under synthetic stimuli extend to responses to natural movies. In this study, we use two-photon calcium imaging in 6 mice (both sexes) to measure the response to optimal frequency square wave gratings and water flow, each rotated into 8 directions. We find these stimuli recruit an overlapping neural population, recruited slightly stronger (signed rank test on average response across trials and directions, $p = 1.7e^{-5}$, r = 0.097, n = 981), but more reliably (rank sum test on variance explained by tuning curve: $p = 1.7e^{-8}$, r = 0.211, n = 225), to water than gratings. Water flow drives direction tuning in 50% more neurons than gratings (of n=2116, 18% direction tuned to G; 33% direction tuned to W). As predicted by a linear-nonlinear edge detector model and a motion energy model, edge extracting neurons have unrelated tuning across stimulus class (in water flow movies, edge orientation is not correlated with motion direction), while motion extracting neurons are tuned to similar directions across stimulus statistical context.

CHAPTER 1

GENERAL INTRODUCTION

Primary visual cortical neurons respond to oriented edges of luminance contrast. Primary visual cortical neurons respond to motion in the visual field. Visual signals arrive at the retina within a statistical context. Natural statistical contexts are distinct from synthetic contexts. It remains unknown how single-neuron responses to edges and motion are related across visual stimuli of differing statistics. The statistical context of a visual signal might affect the properties of a neuronal response. In this thesis, we explore the role statistical context plays in neuronal motion tuning - the preference a neuron displays in responding to a visual stimulus moving in a direction. We begin with a brief discussion of historical and modern visual cortical studies on which this work builds.

1.1 Controllable, parametric stimuli and V1 response

Synthetic statistical contexts have unlocked visual neuroscience concepts foundational to motion tuning. In the late 1950s to early 1960s, the progenitors of the field, Hubel and Wiesel, detailed the response of primary visual cortical neurons in the cat to oriented bars of light [44]. Previously it was shown that neurons in primary visual cortex respond to light hitting a particular region of the retina, called the retinotopic map. Hubel and Wiesel found that within the relevant area on the retinotopic map, a neuron responds to static or drifting light and dark luminance subregions, separated by an edge. To theoretically explain these responses, the field turned to a visual neuroscience concept, developed from similar results in retinal ganglion cells, that became the dominant concept over the next few decades - the linear receptive field. The linear receptive field, the idea that neurons employ a finite impulse response filter to their inputs over space and time, is so pervasive that it has become paramount throughout sensory neuroscience (auditory system - [2, 67, 113], somatosensory system - [37, 23]). The linear receptive field also has significant similarities to filters produced

in computational analysis of natural images, allowing fruitful collaboration between the visual neuroscience field and computational visual information processing [94, 5, 119].

One way to measure the linear receptive field of a neuron is to utilize another synthetic stimulus - white noise. With enough sampling, a neuronal response can be reverse-correlated with high statistical power with the stimulus frames preceding spikes, in a method called spike-triggered averaging [13, 85]. Many spatial filters from this analysis result in filters named for the Hungarian physicist, Denis Gabor. V1 neuronal Gabor filters are typically described as two-dimensional filters in space, most often with two or three subunits of positive or negative luminance (as compared to mean luminance blank gray), formed by the multiplication of a sine function with a Gaussian [49]. These filters result from independent component analysis of natural images [8], and result from learning a sparse code for natural images [77]. The temporal kernel that results from spike triggered averaging is often a monophasic or biphasic Gaussian [17]. Response properties that can be modeled with this space-time separable Gabor function in space and Gaussian in time are hallmark of a set of V1 neurons called simple cells.

A separate set of primary visual cortical neurons - complex cells - do not respond to synthetic stimuli in the same way as simple cells [45, 68]. The specific qualities of a complex cell are typically not well defined, but in general these cells are best described with multiple linear filters [99], and thus employ a space-time inseparable filter - these cells respond to motion, formally defined as oriented energy in the space-time domain [65, 33]. Complex cells employ Gabor-like filters on spatiotemporally oriented energy [1]. To obtain a complex cell filter, one need only sum and rectify the output of phase-offset space-time separable filters. This arithmetic computation results again in filters with subunits that respond to light and dark patches of light; the difference is that to activate the filter these patches need to change throughout time, most often through moving across space. Complex cell filters are thought to be the output of opponent summation of such filters moving along an axis, and are therefore directional, and respond to motion moving in one particular direction with

responding to motion along the axis in the opposite direction. Classically, the response of either a simple or complex neuron is the proportional weighting of the convolution of its filter with a visual stimulus.

In reality, the response is not so simple. A linear weighting of the stimulus does not produce the responses experimenters often measure. Linear models account for the response to white noise, but they fail to predict the relative strength of the response to gratings, including over-estimation of the bandwidth of neural tuning in orientation and frequency [112, 19], and over-estimating the response to non-preferred grating directions [4]. In simpler terms, neurons respond less to visual inputs of slight activation of their receptive field, and stronger to relatively strong activation of their receptive field, than would be expected from a linear weighting of the input. This behavior clearly evokes a threshold nonlinearity. Passing the output of their linear filter through a threshold nonlinearity (and then using this activation as the input to some probabilistic model, typically Poisson) results in better fits to spike rates observed in the visual system [79]. Various nonlinear computations are important for cortical processing and homeostasis, to deal with multi-faceted inputs and various contrasts [3, 55, 88].

Both simple and complex cells are excited by sequences of patches of light (or shadow) that match their spatiotemporal receptive field. In classical theories, the spatial or temporal context of the visual stimulus, excluding the classical receptive field in which the filtered pattern is found, should not affect the response of the neuron. However, experiments often find the classical neural response can be modulated by visual information outside of the classical receptive field. Neurons are not only matching their internal filters with the visual stimulus, they are also influenced by the visual content surrounding the filtered visual area, in the aptly named non-classical receptive field, the receptive field surround [123, 107]. Interconnections in the neural network gives rise to relationships between the response of a neuron and the responses of other neurons in the network [34]. Neurons in primary visual cortex are influenced by excitatory and inhibitory relationships with neurons that filter space

surrounding a neuron's classical receptive field [9, 102]. In experiments with impoverished visual stimuli, the surrounding visual context is controlled by the experimenter, and is often blank or contains a pattern identical to the classical receptive field.

1.2 Increased reliability and sparsity in V1 response to naturalistic stimuli

In nature the entire visual field is typically stimulated with complex patterns of vastly different statistical composition than classical synthetic stimuli, both within and outside the classical receptive field. Synthetic patterns are typically governed by a few generative rules, resulting in generally homogeneous scenes, often containing narrow frequency content in both space and time. Natural patterns, however, are exhibit power law spectra both spatial and temporal frequency content [115, 97, 106]. Natural scenes are complicated by foreground and backgrounds, and by objects of various sizes moving in them at various speeds at various depths. Dynamic visual scenes churn and swirl with complex patterns, following patterns that can be described statically but almost never entirely repeating themselves, lending a new meaning to the axiom of Heraclitus, the Greek philosopher in the sixth century: "You cannot step twice into the same river". In modern vernacular pertaining to vision: you cannot view the exact same scene in nature twice. The river metaphor is especially effective at illustrating this idea: water flowing down a stream always looks like water, with wavelets and objects floating along the surface, and underwater objects causing eddies and flows. However, the exact same image of these eddies, flows, and surface leaves or twigs will likely never occur twice if one is to stand at the edge of the stream for centuries. Natural scenes are largely stationary processes, meaning their first and higher-order statistics are often consistent over time. In other words, natural images are statistically redundant, but specific probabilistic instances of natural scenes are often unique. Changing the perspective of the viewer also complicates the picture: from the perspective of a single neuron, does an observer on the east

bank experience the same visual stimulus of an identical observer on the west, even if they are focused on the same visual location in the stream? What starts in the right visual field for one observer will start on the left of the other, so even this change of perspective reverses the retinotopic map. Synthetic stimuli often repeat and are often statistically homogeneous across space and time. Natural scenes are inhomogeneous and rarely replay.

Even though natural scene statistics are vastly different from synthetic scenes, neurons in natural stimulus experiments have been measured to use filters similar to the filters they employ under synthetic stimulation. Measurement of the linear receptive field to white noise stimuli is exceedingly tractable because white noise is uncorrelated, which provides fast sampling of the entire possible stimulus space. Natural scenes are correlated, resulting in slow statistical sampling of the possible input space [96]. However, formulation of a sufficiently rich stimulus set for measurement of the linear receptive field from spike-triggered averaging techniques is possible, at the cost of longer, but tractable, experiments of noise with natural correlation distributions (brown noise), or even natural images themselves [92, 117] utilizing information theoretic approaches or Bayesian likelihood estimation [81, 80]. Edge detection neurons are therefore expected to extract similar edges across statistical context, with the caveat that natural images likely do not present an optimal edge to their filter, so the nonlinear processing of the neuronal response is important. Less explored is how motion processing is affected by statistical context, although functional models describing motion tuning to synthetic stimulus data have no apparent weaknesses to processing natural inputs as well. Natural scenes can be used to measure receptive fields, and motion processing occurs on synthetic stimuli and natural stimuli alike.

The response to natural scenes, however, often defies expectations, and is not always well explained by these models [52, 128, 16]. As discussed previously, the edge content of natural scenes is varied and inhomogenous, and neurons are affected by their receptive field and the surround. Statistical context is therefore both varied within and across natural scenes, and important for the neuronal response. Although the response to natural scenes is not always

well explained by these models, numerous important discoveries have been made about the response to natural stimuli. The strong low frequency content of natural stimuli drives more reliable responses in primary visual cortex [118, 91]. Responses to natural movies are more sparse than responses to synthetic stimuli. Reliable and sparse responses are key components for efficient coding theoretical frameworks [75] which reduce redundant information, helping satisfy important metabolic constraints on neural activity [58, 6, 57]. The picture is muddied, however, when one considers various experimental conditions and considers dynamic stimuli - sparsity and reliability measurements can vary even across natural stimulation experiments, sometimes resulting in values similar to those measured under impoverished stimuli [123].

We have discussed at length the difference between synthetic and natural stimulus experiments. Within these broad categories there is no standard for determining what statistics a stimulus of a particular type should have. Even white noise experiments can vary the size of a pixel and the flickering rate of the stimulus. Vastly differing stimuli are referred to as "naturalistic". Should a natural stimulus include self-motions from an organism's body, head, and even eye movements? Should a natural scene be filtered for the spatial acuity of the model organism? The statistics of natural stimuli obey scaling laws, but does that mean Hollywood movies are fair natural scenes to show a salamander retina?

Synthetic movies have been utilized to uncover a vast amount of information about the neural response, but at the cost of exploring components of neural processing in isolation, because of the combinatorial explosion of including psychophysical stimuli for every possible neural response component (contrast, luminance, or speed gain control). Psychophysical stimuli are of course more useful for measuring these response properties than natural scenes, because natural scenes vary along these dimensions in an often uncontrolled manner [98], but it remains unclear whether the response of neurons to natural movies can be satisfactorily explained by models combining each of these response properties, or whether an entirely different filter set is employed by neurons driven by natural scenes.

1.3 Experimental setups for interrogating vision

Visual responses are measured in many experimental contexts. Various model organisms are of course utilized to explore visual processing systems. Some properties of visual anatomy and function are shared across species. Linear receptive fields have been documented in animals across taxonomic classes in analogous brain areas (iguana optic tectum - [109, 108]; pigeons - [36]). Some properties are specific to the evolutionary context of an organism. Macaques and cats display primary visual cortical columns that share orientation preference, while rodents do not [93]. Early visual neuroscience experiments relied on anesthetizing the model organism, and recent reports have found details of the visual response to be dependent on brain state [72, 82, 29].

Researchers employ various neural recording technologies to measure neural activity. The properties of a recording technology sometimes result in similarities to data measured with other methods, and sometimes they reveal biases, shortcomings, or strengths of another experimental approach. Fluorescent calcium reporter experiments have shown that historical studies using electrical probes were often biased towards measuring strongly responding neurons, leading to average descriptions of visual cortex biased towards these neurons and ignoring many others [78, 76]. Overcoming these limitations and unifying these various experimental conditions into a comprehensive theory of visual processing remains a challenge.

Here we present work from a relatively novel experimental model and modern technological design. In the last few decades the mouse model for vision has become increasing common [46]. Mice provide facile genetic manipulation within a mammalian visual system analogous to humans (gene expression systems - [60]; synaptic tracing - [129]; optogenetics - [131]), with V1 provided at an economical size of several millimeters across, located on the surface of the brain [125]. Long thought to be a poor model because of low spatial acuity relative to humans [84], the mouse is becoming a popular visual model. It has been documented that mice perform detailed visually guided tasks such as hunting crickets [43], implying the mouse mouse visual system can overcome its relatively low spatial acuity. Our mouse model is a

transgenic line that has a calcium reporter in pyramidal neurons throughout the brain. This calcium reporter, GCaMP6S [15], provides detailed spatially localized fluorescence signals at very high signal to noise, at the expense of slower fluorescence transients compared to fast versions of the same molecule. We excite this fluorescent molecule with light at half the excitation energy; therefore the reporter needs two photons of light to fluoresce. Two photon excitation occurs measurably at the focal point of a focused beam of this low energy excitation, giving excellent spatial resolution to the excitation. Blinking constellations of neural firing can be observed as this focused beam is scanned across the brain. This technique allows for simultaneous recording of many neurons (up to tens of thousands in recent reports) at temporal frequencies fast enough to sample average neuronal response to stimulus ensembles. In this work, we measured up to 300 neurons simultaneously at speeds of 30-40 Hz [100]. This experimental approach allows us to densely sample neurons in the field of view, recording high activity and low activity neurons alike.

Two-photon calcium imaging in the mouse is an excellent system to study visual processing of natural scenes. Fast and dense sampling of visual cortex allows for detection of neurons weakly responsive and tuned to our stimuli. Being terrestrial, the ensemble of natural scenes they are likely to see in the wild are not wildly different from the scenes other land mammals encounter. Because of their size, it is relatively easy to build a microscope for awake behaving mice, and surgically implanting a cranial window is tractable. Congenitally expressed mouse lines for our calcium reporter are commercially available. In this model, we will explore motion tuning to synthetic and naturalistic contexts. Because natural scenes are not monolithic, we choose to scale the complexity of our naturalistic stimulus by removing complicating features. Our naturalistic stimulus is flowing water, which does not have a foreground and a background, and does not contain objects. Our stimulus might be best described as a dynamic natural texture. We will compare the motion tuning properties of edge detecting and motion detecting neurons, identified by their response to synthetic stimuli, with these naturalistic water flow movies. In this manner, we hope to take a

step towards understanding the motion encoding of primary visual cortical neurons across statistical context, utilizing our stimulus that is a step along the spectrum from synthetic to natural.

CHAPTER 2

ENGINEERING A MICROSCOPE FOR IMAGING AWAKE MICE

2.1 A linear treadmill to measure ambulatory velocity

As previously discussed, the benefits of two photon microscopy and the mouse model for studying cortical processing of vision are myriad; however, there are also many challenges to recording the neural activity of an awake mouse. Awake mice prefer to be able to move rather than to be confined to a particular location. However, because our microscope has a fixed objective, our mice must be head-fixed. To allow our mice to ambulate in a fixed location during experiments, we designed and built a 3-D printed voluntary treadmill for mouse ambulation. The treadmill consists of two spoked wheels, turning about aluminum axles, with aluminum flanged ball bearings inserted into the treadmill walls (Fig 2.1). A thin strip of fabric was glued into a loop, and used as the treadmill band. In order to support the weight of the mouse, and also minimize friction, a treadmill bed of delrin was inserted into the walls between the wheels. The treadmill band was placed around the wheels and over the treadmill bed, allowing the mouse to push the band over the bed to ambulate in place.

Through testing we discovered mice do not prefer to ambulate on the treadmill alone, because on the treadmill they are essentially standing at the edge of a table. House mice prefer to ambulate close to walls and within enclosures, so they do not feel exposed to predators. For this reason, we engineered our treadmill with two small walls on either side of the location where the mouse ambulates without blocking the visual stimulus from the mouse, but giving the perception to the mouse that it is not fully exposed during imaging. Using a treadmill model with movable walls, we discovered that if the walls are too far away from the mouse, the mouse will freeze with fear as often as when the walls are not present. However, when the walls are too close, the mouse takes the opportunity to grab the walls with its back legs, in an attempt to wrest itself free from head-fixation. Optimal placement

from the mouse is around 5 cm on either side, so the mouse will remain centered on the treadmill and facing forward without reaching out to the side for the wall, but will remain calmer than if there were no treadmill walls altogether.

Our treadmill design is linear, which creates ample experimental space under the objective. Other treadmill designs, like floating foam balls or dinner-plate rotary designs, have a large footprint on the laser table. We learned through our experiments that the efficacy of clamping the mouse headbar to the table is directly related to the distance the clamping bar is to the mouse - the longer the clamping bar, the longer the lever arm the mouse has to fight against its constraints, leading to the possibility of motion artifacts in microscopy data. For this reason, a small treadmill footprint is preferred to larger footprints. Other treadmill designs, especially the foam ball floating on air treadmill design, encourage the mouse to ambulate by providing very little support when a mouse is stationary. In essence, on such balls the mouse is constantly working to correct its balance on the ball. Our treadmill is stable when the mouse is stationary, providing less distraction for the mouse on the rig.

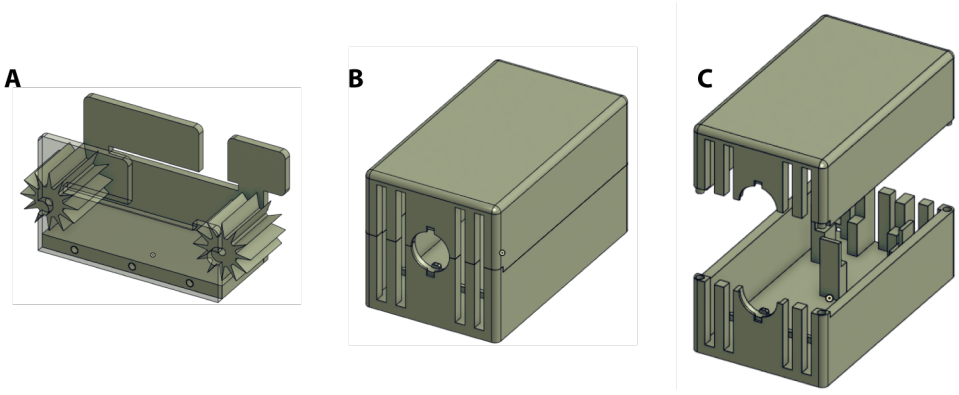

Figure 2.1: **Treadmill and eye-tracking illumination.** A) 3-D rendering of the plastic components of the mouse treadmill. B) 3-D rendering of the IR LED box. C) Exploded view of B).

The speed of the mouse was recorded using a rotary encoder (US Digital H5 Ball Bearing Optical Shaft Encoder) fixed to the axle of the wheel with a 3-D printed attachment. This encoder contains four important components: a light source, two light detectors, and a slotted wheel, with two circular rows of slits. The encoder operates by recording luminance changes

on the light detectors as the wheel is turned about the shaft. The two rows of slits are cut into quadrature phase, such that they oscillate blocking the light from the detectors in such a way as to indicate which detector has turned on first, and which has turned on second. The two square-waves created by these oscillating detectors allow the encoder to distinguish counter-clockwise from clockwise motion, as determined by the leading and lagging channels. The voltage output from this device was recorded, and because the diameter of the axle and the treadmill wheel is a fixed ratio, a simple transformation gave a measurement of the running speed, using the diameter of the treadmill wheel and the number of turns of the encoder per second. The encoder was designed to output 100 cycles per revolution of the encoder axle, and the wheel had a 5 cm diameter, leading to a translation of 0.05 cm per cycle of the encoder. Counting the cycles per unit time gave an estimate of the instantaneous velocity of the mouse.

Locomotion in the mouse is a proxy for brain state - essentially indicating the extent an awake mouse is "awake" (mouse brain oscillations can exhibit characteristics of slow-wave sleep during quiescent periods, [63]). Increased locomotion is associated with increased neuronal firing rates, but this increase in firing rate does not indicate a change in the tuning curve of neurons, providing evidence that this locomotion-induced firing rate increase is instantiated through a gain control mechanism [72]. Because the increased firing rate does not change the tuning properties of the neurons recorded, we chose to include running epochs with quiescent epochs in our data presented in Chapter 4. We did measure whether the results presented in that chapter differed significantly when running was removed, and we found that the general results were unchanged (data not shown).

2.2 An infrared eye-tracking camera

A second challenge to visual experiments with awake animals is the possibility that the animal in question moves its eyes - changing the area of visual angle each neuronal receptive field occupies. To track the eye movements of the mouse, we used an infrared-sensitive Prosilica

GX GigE camera to record eye position. Our eye-tracking methodology is similar to the methodology presented in [101]. With strong illumination of the eye, the pupil is easily distinguished from the iris, since the pupil is much darker in color than the iris. Strong illumination was provided by a 5 W (20mm Star Bead design, 1.5V, 1400mA) light-emitting diode (LED) in the infrared range at 940 nm, since mice lack a long wavelength opsin, and this light is therefore imperceptible to the mouse (no light-emitting diode emits with a delta function, but the emission spectrum of the LED was peaked sharply enough that humans, most of whom can perceive red light, can not perceive this IR LED either). LEDs are typically manufactured with minimal optics, and bare surface-mount LEDs provide diffuse light. It was therefore necessary to focus this light onto the mouse eye, in order to provide enough luminance to distinguish the pupil from the iris.

We accomplished this focusing with two methods. The first was to design a 3-D printed box to contain a high-power IR led, the current limiting resistor for the LED, an AC to DC power converter, and a cooling fan. Because the wattage through the current limiting resistor drove the current limiting resistor to high temperatures, the box was printed with Formlabs High Temperature Resin to mitigate warping from heat, and the box also contained a small brushless fan to cool the resistor. This box was outfitted with a commercially available plastic surface mount LED spot lens, which was fitted into the front opening of the box. This backscatter of this IR light onto the eye tracking camera provided enough contrast to be able to distinguish the pupil from the iris.

However, the singular spot from this box provided illumination to one side of the eye better than the other, making one side of the pupil very dark, but leaving the other with less contrast. The optimal location for placing the IR illumination was directly in front of the eye, but the visual stimulus is placed directly in front of the eye. We therefore designed a second illumination method. The second method consisted of replacing the LED in a commercially available handheld focusable flashlight with the same high power IR LEDs. With this method, we were able to fit two converted flashlights into the experimental rig

(with the current limiting resistor still housed in the heat-proof box) to provide light onto the eye from two vantage points. Filling the eye with light from two directions allowed complete filling of the eye, and high contrast all around the pupil. Two illumination sources instead of one allowed us to move the light sources to each side of the monitor presenting the visual stimulus.

Eye tracking must be performed while an experiment is underway. Our two-flashlight design was suited to fitting onto the experimental rig, and providing light to track the eye. However, the wavelength of light emitted by these LEDs happened to be identical to the wavelength of light emitted by the laser for two-photon illumination. While the illumination laser is on, intense laser light scatters through the brain of the mouse so that the entire eye is brightened at 940 nm during an experiment. To overcome this problem, we made use of the distribution of light emitted by our LED, and the fact that the laser has a very tight emission spectrum. Since the laser emits in a very narrow band, we were able to filter almost all the light from the laser with a Thorlabs longpass optical filter with a 950 nm cut-on wavelength. The LED provided enough power at the passed wavelengths for eye tracking, at the cost of moving further red on the absorbance spectrum of our IR camera - moving into a region of the camera chip absorbance spectrum with lower quantum efficiency. A slower shutter-speed on the camera made up for this loss in signal.

The final hurdle to eye tracking the awake mouse on our rig was that the ideal location for not only illumination but also the eye tracking camera is normal to the cornea - the exact location of the visual stimulus. In order to record normal to the cornea, but out of the field of view of the mouse, we recorded videos of the eye after the image was reflected off of a dichromatic mirror placed near the visual stimulus. This mirror passed visual light, and reflected IR light. We used again used two methods to fix this mirror in place. In the first method, the mirror was held in space by a custom-built brass wire mount, designed with a weaving pattern to be thin enough as to be almost imperceptible to the mouse, but strong enough to hold the weight of the mirror. The second method utilized again a brass wire

frame to hold the mirror, but instead of mounting this frame to the table, with frame was suspended aloft by thin clear nylon thread tied to the microscope and stimulus monitor. By focusing the camera onto the image formed in this mirror, we were able to place the camera at an angle away from the visual stimulus, while minimally affecting the visual field of the mouse.

We repurposed homebuilt LabView software from the Scanziani lab (UCSF) for recording the pupil position from this video data. Within a manually specified rectangular ROI around the eye, the video feed was low-pass thresholded to isolate the pupil. A circle was then fit to the pixels with luminance below this threshold. As an estimate of eye position, the pupil x-coordinate, y-coordinate, and diameter in pixels was recorded from this circular fit. Our eye tracking setup can record eye position in the mouse while neuronal fluorescence data is measured. However, because of the impracticalities associated with the technique, and for other reasons discussed below, we did not eye-track during the majority of the experiments discussed in this dissertation.

Because eye movements move the receptive field of a neuron in the visual field [64], reconstructing the neural receptive field from a visual stimulus benefits greatly from accounting for eye movements. However, in our hands typical eye movements in the mouse were small and infrequent. Even in the aforementioned study, in a foveate animal like the macaque, exemplar eye movements only covered 0.2°. In our studies, mice move their eyes at most 2.0°, and on average less than 1.0°per movement. Movements also occurred relatively infrequently, compared with foveate animals.. Our visual stimuli in this study are full field drifting textures and patterns, with similar statistics covering the entire visual stimulus. Because the statistics in one position on the visual stimulus are similar to the statistics elsewhere, the receptive field is not filtering vastly different statistical inputs no matter where it looks on the screen. However, this is less true for natural movies, as the statistical class in a natural movie is similar across the visual field, but the particular instantiation of natural statistics can change across the visual field. Nonetheless, because of the infrequency and small magnitude of eye

movements during experiments, we did not track eyes during the experiments presented in Chapter 4.

Eye tracking is useful for reconstructing the visual stimulus on the retina; similar to running, it is also useful for estimating the brain state of the mouse [86]. The diameter of the pupil indicates arousal levels, with a larger pupil diameter indicating increased arousal. Global brain state has varied effects on individual neuron activity. On average, arousal level is linked to increasing fire rates in V1, but some narrow-spiking neurons are suppressed by arousal (as measured by locomotion) in mouse V1 [72]. Suppressive effects of arousal have also been measured on average in other brain areas in the mouse [104]. When the effects of arousal as indicated by pupil diameter and running speeds are dissociated, it has been shown that these modalities of arousal measurement are associated with different synaptic efficacy changes in different neuron subtypes [86], and have differing consequences for functional flexibility in neural circuits [122].These results have inspired hypotheses that these changes are associated with shifts in behavioral modality between exploration (seeking out new resources to use) and exploitation (utilizing known sources of resources) tasks [39]. This framework lends an interesting perspective to mouse vision - if the ethological usage of vision in the mouse is predator detection, and if a mouse needs to forage to exploit its environment and is most vulnerable when its motion makes it easier for a predator to detect it, then an increase in the gain control of predator detecting circuits while running seems a possible target of positive evolutionary selective pressure. Exploration of this framework seems an especially pertinent area of study for naturalistic experiments. Nonetheless, it remains outside the scope of the work presented here.

CHAPTER 3

MODELS FOR MOUSE VISION

3.1 Modeling at various levels of complexity

This work makes extensive use of two canonical computations of mouse V1 - mouse visual cortical neurons extract luminance edges, typified by the simple cell linear receptive field and linear-nonlinear model of neural processing; mouse visual cortical neurons extract motion direction, exemplified by the motion energy model of neural processing. We present in Chapter 4 the essential components of these models, and here discuss some intricacies of each. Of course, there are many other canonical computations that cortex employs to accomplish its many functions, and in this chapter we also discuss some of these models which were not as essential to understanding the data presented in Chapter 4.

Computational models are rarely designed to capture all of the complexity of a process, and are often designed to abstract a complex process into digestible pieces [110]. Computation can be understood at various levels of abstraction, by choosing which details to include and which characteristics of the real-world process to match [61]. Rate based models, like the linear-nonlinear model and the motion energy model central to this work, are abstractions which deal in the realm of computation and do not discuss implementation of their algorithm. Various other models are more concerned with the biophysical plausibility of the computations they perform, and models of varying levels of complexity are employed. Each layer of the visual system employs many neuronal subtypes - the number of cell subtypes in the retina is ever increasing, and currently sits at over 30 [7]. To describe the complexity of the turtle visual system, Nenadic et al. (2003) employ a model with 744 compartments describing only 3 cells types, with an input layer of an additional 201 compartments, all to recapitulate the relatively simple (when compared with visual encoding) phenomenon of propagating waves in visual cortex. The number of cell subtypes and heterogeneity within a cell subtype results in an explosion of complexity if one is to simulate even a small number

of neurons in a biophysically plausible manner. Depending on the complexity of the neural behavior one seeks to model, it is generally possible to reduce the complexity of these models, to form tractable models that describe the most important components of the system but sacrifice matching some small percentage of the details [42]. In [10], the researchers describe a compartmental model that reduces the complexity of a 400 compartment model to 8-9 compartments, with a modest loss of fitting accuracy. Nonetheless, the question of how much complexity will be required to describe the cortical response in natural situations remains open.

As discussed in Chapter 1, Section 1.1, the linear receptive field is currently the linchpin of visual cortical processing, and enjoys a similar seat of significance in other sensory systems. The linear receptive field is so appealing because it supports a particular view of cortical processing that allows researchers to treat the brain as a feed-forward feature extractor, with simple features extracted at early points in the visual stream, and combined in complex representations in higher brain areas. This viewpoint is attractive because it allows theorists to bring the full weight of hierarchical information processing theories to the question of understanding cortex. Feed-forward networks are simple and are well-studied in the fields of computer science and information theory.

However, actual neural networks display a preponderance of feed-back recursive connections, and this interplay of feed-forward and feed-back interactions complicates the picture. These connections also occur both within brain layers, across brain layers, and even across brain regions, leading to multimodal crosstalk across sensory systems [26, 32]. As previously discussed, models informed by simple synthetic stimuli like white noise do not always explain responses to naturalistic conditions [89], even in the retina, the supposed simplest component of the visual stream [51]. These results seem to indicate that the feed-forward feature extracting viewpoint is incomplete at best. However, amazing progress has been made utilizing this framework, possibly best illustrated by successes in brain-machine interfaces that utilize simple feature based decoding strategies [121]. The receptive field remains an

important, but incomplete, and therefore controversial, model of cortical processing [30].

3.2 Normalization computations

The modeling landscape of motion extraction is complicated by the discovery that mouse brains (and other mammalian visual systems) employ motion extractors of two types - pattern motion extractors, and component motion extractors [70]. This separation arises from the fact that the motion of edges can only be discerned in the motion axis normal to the edge orientation, and any motion parallel to the edge orientation in lost, arising in many edge containing visual stimuli driving the same motion percept [69]. When presented with two drifting gratings at differing orientations, a stimulus called a drifting plaid, some neurons extract the motion of one edge and not the other, and some neurons extract the combined motion direction of the plaid [130]. Interestingly, the motion energy model can account for both of these types of motion computation, with the addition of another key canonical neural computation - divisive normalization.

Some neurons are tuned and respond maximally to visual components in an angular dependent manner. In simulation, if a population of linear-nonlinear neurons is shown a visual stimulus with a mixture of visual components, and all of the neurons respond maximally to detecting their preferred stimulus, the result is an epileptic cacophony of neural spiking. The brain overcomes this complication using neural circuits that suppress differently tuned neurons when a tuned neuron fires, so that the firing rate of an individual unit is normalized over the sum firing rate of the population [12]. This model was first proposed as the computation underlying directional selectivity of complex cells [116], and has since been shown to explain non-preferred direction suppression [99]. Motion opponency in the motion energy model of Chapter 4 could be accounted for with two units selective to motion in opposite directions, with divisive shunting inhibition between the pair. Although not explicit in the models proposed in Chapter 4, normalization is an important component of primary visual cortical computation, of both edge orientation and motion direction.

3.3 Hierarchical modeling of the ventral stream

As previously discussed, most visual cortical models treat the object identification pathway (the ventral stream) in the visual system as a hierarchical system of feature extractors. A hierarchical system of object identification is attractive from a theoretical perspective, as also previously discussed, but is also is suggested by to the serial anatomy of primary projections of visual cortical areas, and the fact that hierarchical visual processing provides a solution to an important problem in visual systems: in the problem of object identification, one would hope evolution has resulted a system that can recognize objects from different perspectives, so that a small change in the position of the eye relative to an object does not radically change the perception of that object. This feature of visual processing systems is called invariance - the ability of the cortex to identify an object across multiple scales, orientations, and contexts. Invariance is diametrically opposed to object perception selectivity - the ability to provide a precise characterization of an object, including its scale and orientation. These two features comprise a tradeoff in an inverse relationship, and it is theorized this tradeoff has played an important role in the evolution of mammalian vision [90, 24]. Invariance and selectivity tradeoffs can be optimized easily in a hierarchical object recognition system, in which the early visual areas are highly selective for particular visual features, and are combined in a hierarchical manner in order to create invariant visual responses in higher brain areas.

This hierarchical structure finds a cortical substrate in the various brain regions of the visual stream. In primates, primary visual cortex is thought to employ the selective filters necessary to implement selective responses to specific visual features, and this information is passed in various ways to higher brain areas for further processing [53]. By reducing a visual stimulus to the components that are necessary for maximal activation of a brain region, researchers have attempted to classify the computations that each brain region employs. Visual area V2 contains interwoven maps of both color-selective and disparity-selective cells that are broadly orientation tuned [95]. Area V4 contains orientation maps and stimulus

size maps that are thought to culminate in responses that are specific to stimulus form, giving optimal responses when the orientation, size, and spatial frequency of en edge match the selectivity of the specific functional population containing a particular neuron [22, 38]. Inferior temporal cortex is thought to have complex responses to visual stimuli that depend on many edges and shapes [**?**] converging into an object identity perception, with special attention given within particular domains to the identification of faces and expressions, with increasingly complex and specific response patterns [59, 21, 66].

The mouse visual ventral stream is similarly organized to the primate visual stream, with the important caveat that functional clustering in cortex seems to occur to a much lesser degree in the rodent than the primate [93]. The mouse corollary to V2 is referred to as the the lateromedial field (LM), and is still a primary target of mouse V1 [125], and it strongly projects to mouse temporal cortex [126]. Utilizing modern techniques, researchers have sub-categorized mouse visual cortex into smaller brain regions, and mapped their projections throughout the visual stream, finding it homologous to primates in most ways [127].

The classical roles of each visual area has been critiqued and revised many times, and current models seem to lean towards thinking of the process of object recognition of a particular object class as a task distributed across brain regions, because with the advent of functional imaging techniques, the measurement of correlations between brain activity and visual stimulus have been discovered in an increasingly distributed fashion [48]. The hierarchical perspective of object recognition is therefore being revised, as it becomes clear that feed-back connections from higher brain areas onto lower brain areas are critical components of the visual processing pathway. Primarily two roles for top-down connections in visual processing have been put forward. The first views feed-back as the higher brain areas as contributing task-dependent modulation (attentional, for example) of the feature sets employed at a particular visual processing level [132]. The second views feed-back as a predictive inference on future sensory information that is combined with the actual sensory input at a point in time, essentially instantiating a form of Bayesian hypothesis testing

between the perceptions of the higher brain areas and sensory inputs from the retina [56]. In both of these frameworks, and in cortical processing in general, backprojections are sure to be a key component of cortical processing. However they are not featured in the models presented in Chapter 4.

Diverse models of visual cortex have been put forth, and many have found great success in explaining visual cortical recordings. However, because of the immense complexity of the neural substrate, and the immense complexity of the task of visual processing, modern models do not account for a large portion of neural activity in primary visual cortex [78]. There remains many opportunities for further abstractions and models in order to comprehensively understand visual processing.

CHAPTER 4

CLASSIC MODELS PREDICT MURINE PRIMARY VISUAL CORTICAL MOTION TUNING TO NATURALISTIC STIMULUS

4.1 Introduction

Sensory information, including visual inputs, are collected by organisms so that they might be able to learn about their environment and react to it [54]. Visual inputs are useful for determining the content of surroundings (e.g. identifying predators, detecting prey, navigating complex environments, etc.). Visual tasks can often be broken up into two important components: identifying what is in the environment, and identify how the environment is changing, or moving, in order to interact with it [114]. From an ethological perspective, these two tasks are not mutually exclusive, because an environmental component can be identified using static signals and dynamic visual signals [14]. These two tasks are not mutually exclusive from a neural processing perspective because even motion processing units employ spatial filters [47]. Primary visual cortex employs neural circuits that use spatial and temporal signals to extract information about what is in the environment and how the environment is moving. These circuits perform both tasks in concert and to varying degrees. The output of these computations is represented in orientation and direction selective neurons [44]. Orientation selective neurons are primarily useful for classifying static edge components of visual scenes, while direction selective neurons primarily contribute information concerning motion in visual scenes [18]. It is plausible that neurons that lean towards extracting edges, and neurons that lean towards extracting motion, might differ in how well their motion tuning properties extend across stimulus class. Neural populations that are extracting what is in the scene would respond vastly different dependent upon the specific content of a stimulus, regardless of motion in the stimulus; neurons extracting motion might extract the

same direction of motion independent of what is actually moving.

A focus of neuroscience is to understand the brain in natural conditions, but to parameterize the natural visual world is a nontrivial task [103]. Vision researchers have often thus simplified natural scenes into parametric synthetic movies, and using these stimuli the field has made significant predictions and observations about visual cortical computation. In primary visual cortex, neurons integrate visual information in space and time, resulting in the idea of the classical receptive field [41, 45, 27, 71]. Neurons are also influenced by visual patterns surrounding the classical receptive field, resulting in important normative computations and inhibitive relationships with neurons with neighboring receptive fields [105, 124]. These and other discoveries build on two foundational computations, the touchstones of the visual neuroscience field: primary visual cortical neurons encode the orientation of luminance edges in a visual stimulus; primary visual cortical neurons encode the direction of motion of an object or texture in a visual stimulus.As discovered in Hubel and Wiesel's classic experiments, to illustrate these canonical computations, all that is necessary is to flash or drift a simple oriented edge. The spatial and temporal filters of such neurons responsive to these stimuli can even be uncovered using white noise stimuli [85]. Noise and bars of light is all that is necessary to explore vast regions of visual neuroscience space. Powerful theoretical ideas in information theory and visual filtering explain the response to noise and bars of light with great success [8, 120, 119].

Because the set of neuronal filters in response to noise and bars results in an optimal basis set for natural visual scenes [106], models of the neuronal response to noise and bars have persisted through many decades, even though models informed by response to impoverished stimuli often fail to explain neural response to naturalistic stimuli [31]. The relationship between activity driven by impoverished stimuli and rich stimuli remains unclear. Do neurons extract similar information from natural and synthetic scenes? Does the response to synthetic movies inform predictions of the response to natural scenes? With this study we hope to help point towards an answer.

Natural scenes are very complex and present an immense challenge to experimentalists in that comprehensive coverage of the entire space of natural scenes in an experiment is impossible. However, because natural scenes are not monolithic [87], it is possible to step through complexity of a naturalistic stimulus. Natural scenes often contain objects, a foreground and a background, and other complicating features that make their analysis difficult [97]. In this study, we decided to simplify the picture by choosing as our stimulus a natural texture with relatively uniform spatial and temporal statistics, with no objects, and no foreground/background distinction. Water flow provides a step between fully featured natural scenes and synthetic movies, by providing broad spatial frequency content and motion signals present in the natural world but often absent from synthetic stimuli. If our models informed by synthetic stimuli can predict response properties to this stimulus, we would therefore have a foothold to expect that our models of neuronal computation might be able to extend to natural conditions.

4.2 Methods

Animals and surgery

All procedures were approved by the Institutional Animal Care and Use Committee at the University of Chicago. Data was collected from C57BL/6J mice of either sex (n = 4 female; 2 male, 2-4 sessions per animal) expressing transgene Tg(Thy1-GCaMP6s)GP4.12Dkim (Jackson Laboratory) between ages P50–P200. After induction of anesthesia with isoflurane (induction at 4%, maintained at 1–1.5%), a 3 mm diameter cranial window was implanted above V1 by stereotaxic coordinates and cemented in place alongside a custom titanium headbar. Mice recovered for at least 8 days.

Intrinsic Signal Imaging

V1 was identified using intrinsic signal imaging. Under isoflurane anesthesia, mice were head-clamped under a CCD camera (Qimaging Retiga-SRV). A vertical white-bar stimulus

(100% contrast, 0.125 Hz) was repeatedly presented on an LED monitor (AOC G2460) located approximately 20 cm from the contralateral eye, while cortical reflectance under 625 nm illumination was collected. The retinotopic mapping of V1 was then estimated at each pixel from the phase of peak reflectance driven by increases in activity-dependent blood flow.

Two-Photon Calcium-Dependent Fluorescence Microscopy

Neural activity was recorded from awake, head-fixed mice on a linear treadmill. A rotary encoder attached to one axle measured running speed. Using galvanometer raster scanning, a layer 2/3 field of view roughly 800 μm in diameter in V1 was identified. Raster scans (1 Hz) were acquired for roughly 10 minutes, while high contrast full-field visual stimuli were presented. Neurons in the field of view (50-300) were automatically identified using custom image processing software. Activity from the identified neurons was captured at 25-33 Hz using Heuristically Optimal Path Scanning. Imaging was performed at 930 nm (Coherent; Chameleon Ultra) through a 20X 1.1NA Olympus objective, and fluorescence emission was collected with a GaAsP PMT (Hamamatsu; H10770A-40). The diameter of the field of view was estimated once by fitting circles to a single raster scan of 15 μm fluorescent microbeads. The true field of view size may vary up to 8% across datasets from realignment of laser beam path.

Visual presentation

Visual stimuli were presented on an ASUS VG248QE, 20 cm from the contralateral eye, covering 51°x90°of the visual field, and set to 60 Hz and 60 cd/m2. Using a luminance meter, the screen's gamma curve was measured, and generated stimuli accounted for this curve. Three luminance-matched visual stimuli were presented: 1) square-wave gratings (G), 95% contrast, 2 Hz, 0.04 cpd, 8 directions; 2) water flow (W); 3) water flow texture (T) generated to match spatial statistics of a still frame of the water movie. Each stimulus was rotated into 8 evenly spaced directions. In each experimental session, at least two of these stimulus types were presented in pseudo-random order. A mean-luminance blank epoch

was also presented in each experimental session. Stimulus epochs were presented for 3 s. Five seconds of mean-luminance grey screen was presented between stimulus epochs. Thirty presentations of each stimulus epoch was presented in blocks containing either 3 or 5 repeats of each stimulus.

Naturalistic visual stimuli

Our study utilized a novel naturalistic stimulus - water flowing downstream (W). The stimulus was made from a movie of water flowing across a dam, selected from the Chicago Motion Database, housed in the Palmer lab (www.chicagomotion.rcc.uchicago.edu). The original movie was collected at 60 fps and 512x512 resolution with an AVT GX1050 camera. This movie was cropped, rotated into 8 directions, and resampled to the resolution of our stimulus presentation monitor. A second naturalistic stimulus was constructed: a single frame of W was used to seed generation of a water texture stimulus (T), matching the joint statistics of complex wavelet coefficients [83]. This texture was generated with wavelet coefficients at 7 scales and 4 orientations, in a spatial neighborhood of 3x3, and 50 iterations of the generation algorithm. To control for the particular instance of texture generated, three textures were generated and randomly presented as the T stimulus. For each naturalistic stimulus, three-dimensional (spatiotemporal) amplitude spectrum were computed using a sliding 128x128x128 Hamming window with 50% overlap in each dimension. The direction of motion in W was calculated as the optic flow between adjacent stimulus frames [111]. To measure the types of motion in the W movie, local motion scores were measured based on the luminance correlations within an appropriate spatiotemporal template for each type of motion, and these scores were then combined across space [73]. To measure the number, orientation, and length of edges in W and T, Canny edge detection was performed [11] within simulated receptive fields, spaced evenly into nine locations spanning the visual stimulus (center, cardinal, and intercardinal locations). Edges shorter than three pixels were ignored. Edge orientation and length were calculated as the orientation and Euclidean distance from tip to tail. The standard deviation of the Gaussian filter during edge detection was swept

through 5°, 10°, 20°and 30°(to match the spatial size of receptive fields in mouse visual cortex, [71]. Major gradient threshold was swept through 0.2, 0.35, and 0.5, and minor gradient threshold was set to 40% of the major.

Linear-nonlinear modeling

In the same nine locations as edges were detected, the activation of a linear-nonlinear model was measured [79]. The spatial filter of the model was defined as a circular 2-D Gabor with width swept through 5°, 10°, 20°and 30°, oriented to respond to vertical edges. Both monophasic and biphasic temporal kernels were used. The monophasic temporal kernel was a Gaussian peaked at 100 ms, with a width of 90 ms. The biphasic temporal kernel was the difference of a Gaussian of 110 ms width with a Gaussian of 55 ms, offset by a 25 ms delay. This filter was convolved with the stimulus, and passed through threshold-linear nonlinearity, with threshold of 0.1 and gain of 5.0.

Motion energy modeling

In the same nine locations as edges were detected, the activation of a horizontal motion energy model was measured [1]. The spatial component of the filter was again a 2-D Gabor swept through the previous sizes, but with vertical width 10% of the horizontal (reducing the presence of a spatial edge in the filter). The model utilized two spatial filters, phase shifted by 90°, and two biphasic temporal filters defined by the difference of Gaussians. The faster temporal filter was the difference of Gaussians with width of 55 ms, offset by 30 ms, with second peak 90% the height of the first. The slower temporal filter was the difference of Gaussians with width of 80 ms, offset by 30 ms, with second peak 90% of the first. These spatial and temporal filters were combined to form space-time separable filters, then added to form four spatio-temporally oriented filters, two to the left and two to the right, normalized by the L-2 norm. These were convolved with the stimulus, the output was squared and averaged, and the directional energy was computed as the total rightward activation minus the total leftward activation.

Data acquisition and pre-processing

Data acquisition synchronization was performed at 2 kHz using an NI PCI-6259. Three data channels were simultaneously collected: 1) the voltage trace from a photodiode placed onto the monitor that tracked stimulus presentation, 2) a TTL pulse coinciding with the start of each frame of imaging, 3) the voltage trace output of the encoder measuring running speed. For each neuron, baseline fluorescence was estimated from raw fluorescence by thresholding to eliminate spike-induced fluorescence transients and smoothing with a 4th-order, 81-point Savitzky-Golay filter. Fluorescence time-series were then normalized to percent change from baseline ($\delta F/F_0$) using this time-varying baseline. The mean normalized fluorescence was calculated for each 3 s stimulus epoch. This value was considered the neuron's trial response.

Response properties and tuning classification

Neurons were classified as visually responsive to a stimulus type if the mean response to any direction was significantly greater than the mean response of the neuron to the interleaved blank stimulus epochs by Dunnett-corrected one-way ANOVA (alpha = 0.05). Responsive neurons were then tested for statistically significant orientation or direction tuning to G according to the trial vectors in orientation or direction spaces [62], and for significant tuning to W by bootstrapping across directions. For each responsive, tuned neuron the mean response at each direction was considered the tuning response (curve). To calculate the orientation tuning curve, the responses of the neuron to opposite directions were concatenated, and the mean response at each orientation was considered the tuning curve. Tuning curve fits were not performed. The tuning strength was defined as the normalized length of the vector of the sum of the mean responses in either orientation space:

$$L_{ori} = \frac{\sum_k R(\Theta_k) exp(2i\Theta_k)}{\sum_k R(\Theta_k)} \tag{4.1}$$

where $R(\Theta_k)$ is the response to angle Θ_k, or in direction space:

$$L_{dir} = \frac{\sum_k R(\Theta_k) exp(i\Theta_k)}{\sum_k R(\Theta_k)} \tag{4.2}$$

The tuning axis or direction was the angle of this vector, formalized here in direction space:

$$\Theta = \tan^{-1} \frac{\sum_k R_y(\Theta_k)}{\sum_k R_x(\Theta_k)} \tag{4.3}$$

where R_y is the response y-component, and R_x is the response x-component.

Statistical Methods

Neuronal fluorescence response similarity was measured using partial correlation, controlling for the population mean response:

$$r_{x,y|z} = \frac{r_{x,y} - r_{x,z} r_{y,z}}{\sqrt{1 - r_{x,z}^2} \sqrt{1 - r_{y,z}^2}} \tag{4.4}$$

where x(t) is the fluorescence trace of the first neuron, y(t) is the fluorescence trace of the second neuron, z(t) is the population mean fluorescence trace, and r denotes the Pearson correlation coefficient between respective fluorescence traces. Tuning curve similarity was measured using Pearson's correlation coefficient. Both parametric (two-sample t-test) and non-parametric (Wilcoxon signed rank, rank sum test) hypothesis testing was performed, with p-values listed inline. Effect size for comparing means was calculated as Cohen's d. Effect size, r, for comparing medians was calculated from the appropriate hypothesis test z-value:

$$r = \frac{z}{\sqrt{N}} \tag{4.5}$$

4.3 Results

Water flow movies are an exemplary naturalistic stimulus.

It remains unclear the extent to which neurons in visual cortex respond similarly to visual stimuli which contain directional flow of differing statistical composition. In order to test the degree to which neurons respond similarly across statistical categories of visual stimulus we measured neuronal responses to drifting gratings and to a novel naturalistic stimulus: water flowing downstream. Typical of dynamic natural scenes [97, 25], water flow contains a broad spatial and temporal frequency distribution and various types of motion signals (Fig. 4.1), including Fourier, non-Fourier and glider motion [73]. Primary visual cortex is thought to predominantly encode two features of visual stimuli: edge orientation, and directional motion energy. We therefore measured the edge length and orientation in the water movie within spatially-distributed simulated receptive fields (Fig. 4.2). We measured the edge orientation and length distributions to again be broad, indicative of the broad spatial frequency distribution and fast mixing times of swirling water.We also measured the direction of optic flow in the water movie, finding a distribution of motion direction with mean 10.9°below the horizontal, which we round in the following analyses to 10°.

Water flow creates varied motion tuning curves to edge extracting models.

To compare motion directional tuning across stimulus conditions, we rotated the water flow movie into 8 evenly distributed directions. How should we expect an edge extracting neuron to respond to this water stimulus? To start towards an answer to this question, we measured the activation of a linear-nonlinear (LN) model by the rotated water flow stimulus (see Methods). The water flow contains inhomogenous edge content across space and time, which drives various directional responses to water flow, with the strongest responding direction dependent upon the analyzed region of the movie and the parameters of the model filter. This LN modeling result suggests the preferred direction of a neuron direction tuned to water and orientation tuned to gratings should not be consistent across stimuli (Fig. 4.3).

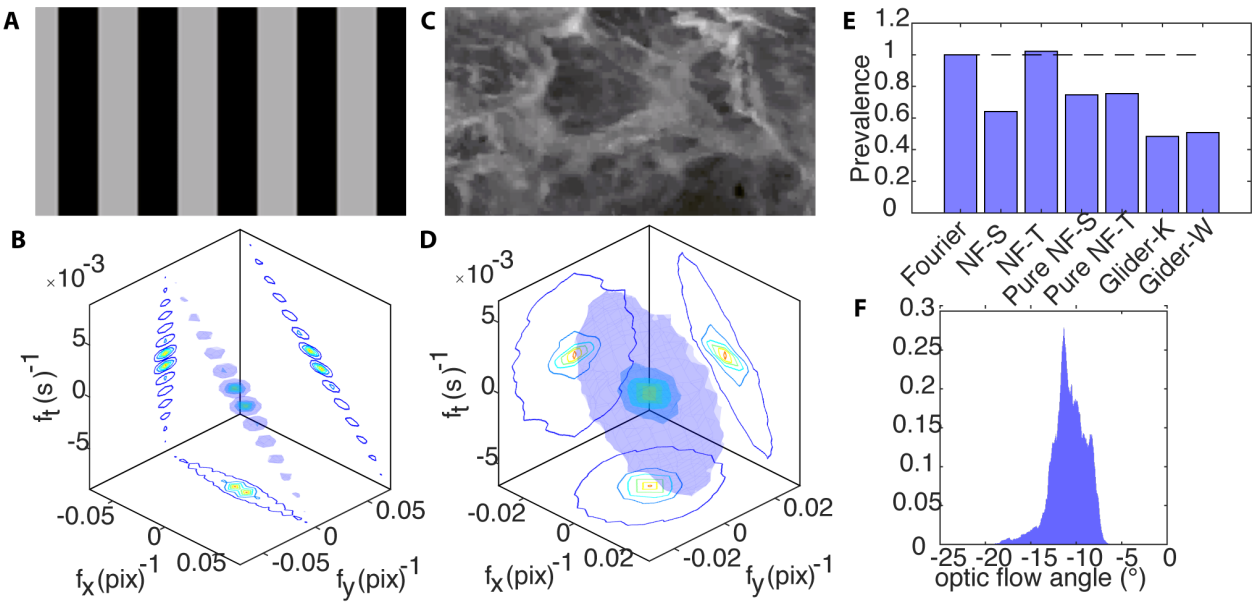

Figure 4.1: **Stimuli characteristics.** A) Still frame of the grating (G) 0°stimulus. B) Spatiotemporal energy spectrum of drifting G, with ripples in frequency characteristic of square wave gratings, and oriented motion energy in x-t plane. C) Still frame of the water (W) 0°stimulus. D) Spatiotemporal energy spectrum of W, with broad spatial frequency distributions and oriented x-t energy. E) Types of motion in the W stimulus, normalized to the prevalence score of the Fourier (standard translational) motion. F) Optic flow angle of W, with mean -10.9°.

Water flow creates consistent motion tuning curves to motion detecting models.

How should we expect a motion energy extracting neuron to respond to this stimulus? We measured the activation of an opponent motion energy (ME) model by the water stimulus (Fig. 4.3). The water flow contains strong directional motion energy across space and time, which drives consistent directional responses to the water flow, independent of the analyzed region or the parameters of the model filter. This ME modeling result gives the opposite expectation as the LN model: the preferred direction of a neuron tuned to water and gratings should agree across stimulus conditions, and the response of motion direction detecting neurons should be correlated across motion directions.

Murine visual cortical neurons similarly respond to distinct visual stimuli classes.

To test these predictions, we needed to identify neurons that were responsive and tuned

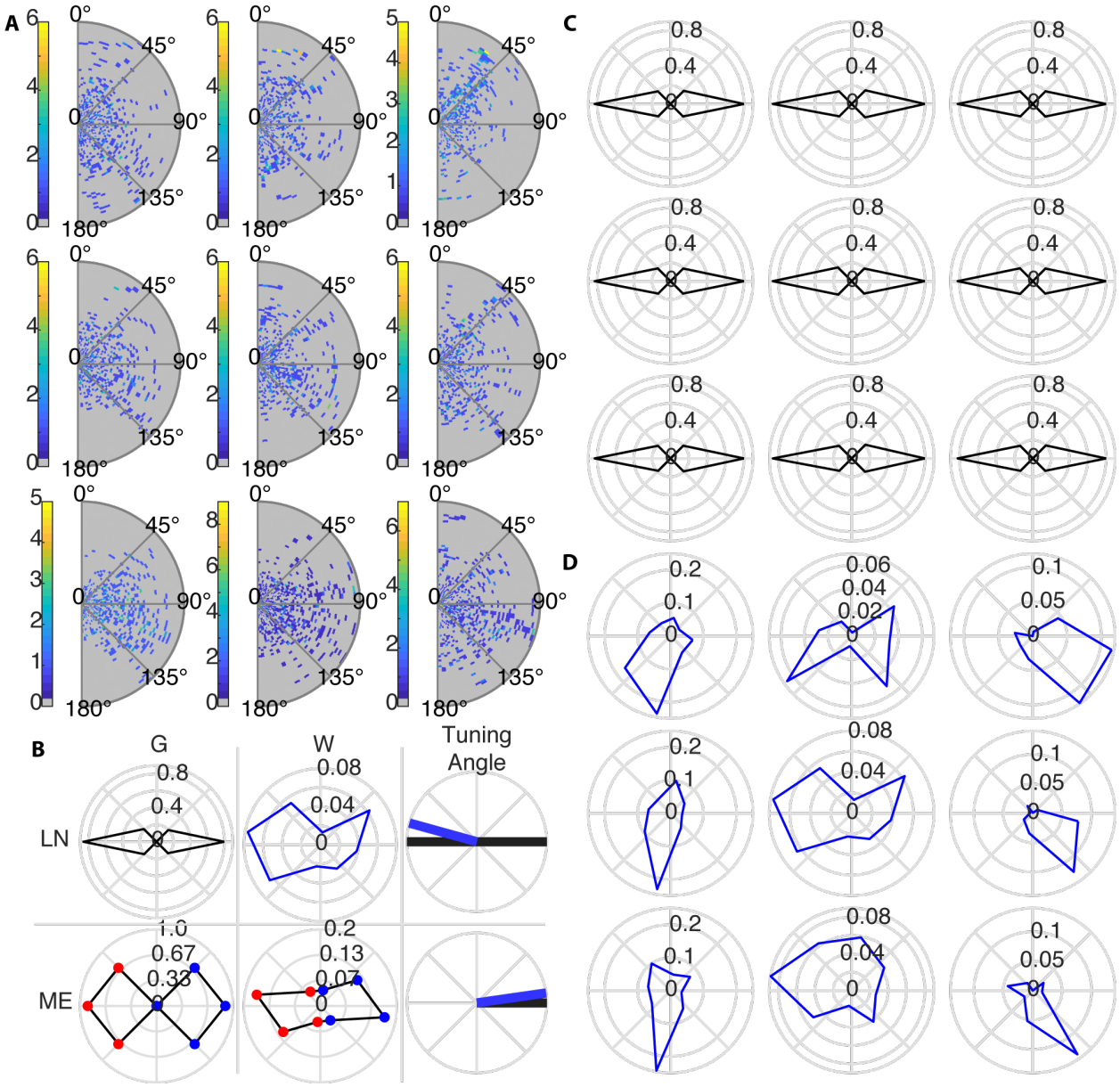

Figure 4.2: **Expectations informed by modeling.** A) Distributions of edge orientation and length, detected by canny edge detection, across 9 spatial locations, to the W 0°stimulus. Distributions are broad, and depend on spatial location. B) Example LN and ME preferred angle analysis. Top-left: response of LN model to G (black in subsequent figures). Top-center: response of LN model to W (blue in subsequent figures). Top-right: preferred axis and direction from response curves in top-left and top-center. Bottom-left, -center, and -right: Same as top, for ME model (blue - positive opponent energy, right-ward signal, red - negative opponent energy, left-ward signal). C) Tuning curves of the LN model with a 20°spatial filter to G, across the 9 spatial locations in A). D) Same as C), for response to W.

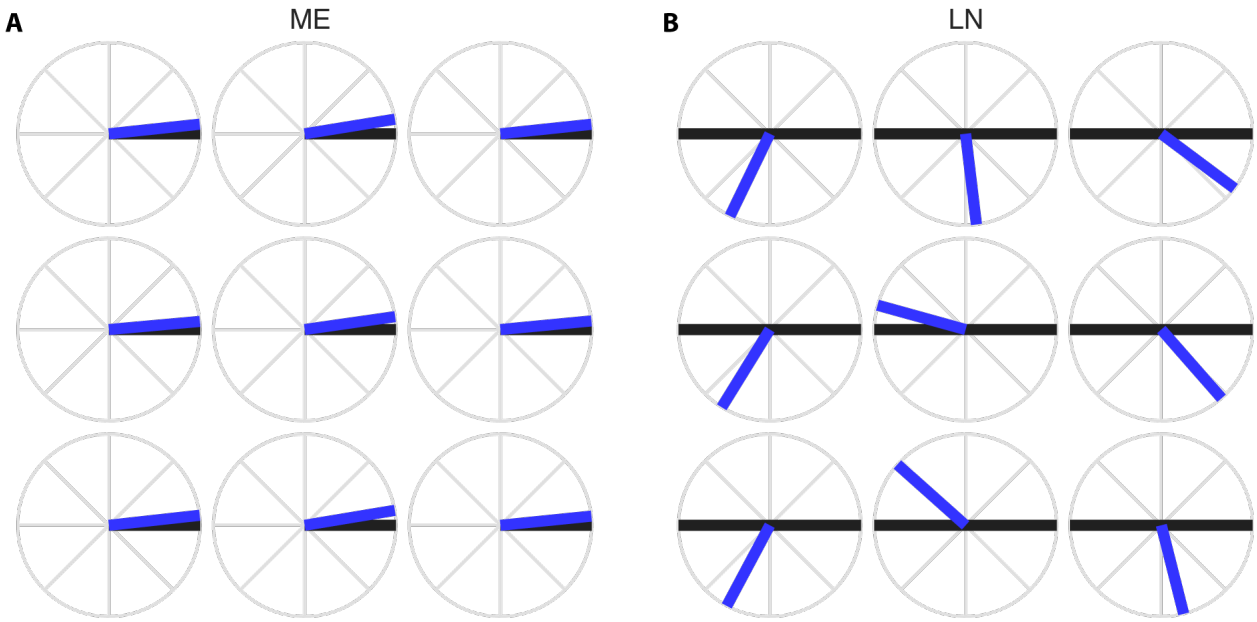

Figure 4.3: **Motion tuning of modeling.** A) Comparison of extracted motion direction by ME model across space for W and G movies, indicating strong alignment. B) Comparison of extracted motion axis from G and motion direction from W by LN model, indicating uncorrelated motion preference.

to both stimulus conditions. We therefore imaged local cortical populations (<800μm diameter imaging plane), of L2/3 excitatory neurons constitutively expressing GCaMP6S (see Methods) in mouse V1 during presentation of the classes of visual stimuli containing directional flow in eight directions (Fig. 4.4). Again, the classes of visual stimuli were drifting square-wave gratings (G), and water (W) flow. As mentioned previously, the spatial frequency distribution and the distribution of motion signals of natural movies are both broad distributions. The gratings, however, are narrow in both spatial distribution and motion. To compare the responses of individual neurons across these stimulus classes, we restrict our analysis to populations of neurons were shown both conditions in shuffled order within an experiment (n = 2116).

We first asked whether V1 neurons would respond to broad- or narrow-band stimuli uniquely, or whether they would respond to both. The chances a neuron responds to G and W is higher than expected from the product of the chance of response to each individual stimulus class. Of neurons shown both G and W, 67% were visually responsive to either G

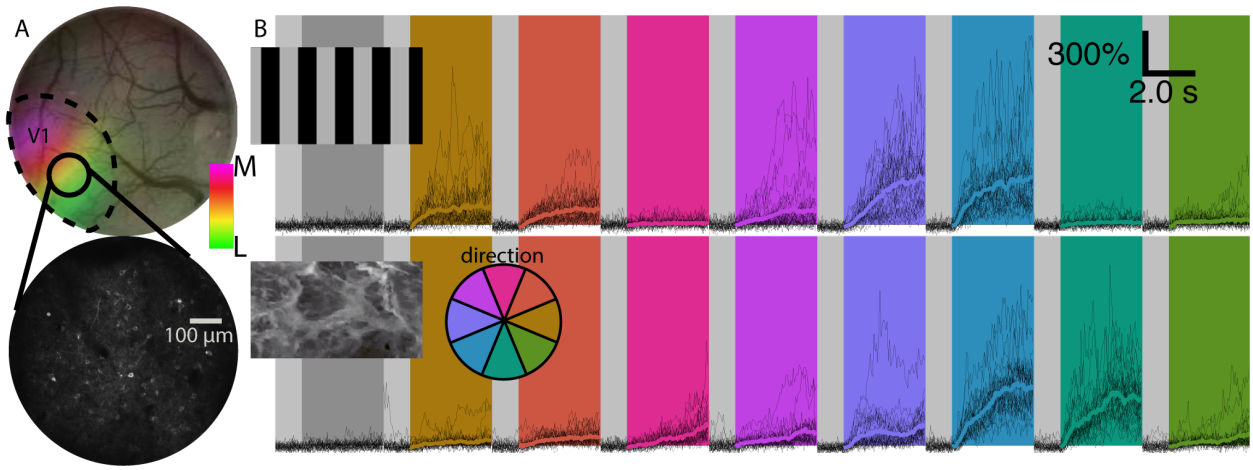

Figure 4.4: **Experimental design.** A) Top: Craniotomy over primary visual cortex, V1 retinotopy mapped by intrinsic signal imaging (colormap - visual azimuth, lateral to medial). Bottom: Two-photon field of view, L2/3, from region in Top. B) Fluorescence responses of one exemplary neuron to W (top) and G (bottom). Gray region indicates last 1s of mean luminance blank epoch before stimulus presentation.

or W, 54% were responsive to G, 60% were responsive to W, and 46% were responsive to both G and W (n = 2116, Fig. 4.5). If the mechanism for responding to G and W was independent, approximately a third (32%) of the neurons would be expected to be visually responsive to both classes of stimuli. About half (46%) of all neurons were responsive to G and W, and only 21% of neurons were responsive to either G or W alone. These results indicate that both water and gratings recruit an overlapping cortical population, even across statistical context.

Taking the neurons (981) that were statistically significantly visually responsive to both G and W, we then asked whether these neurons show increased responsiveness to one stimuli class over the other. We compared the mean response of each neuron across all directions and all trials between conditions. The median response averaged across directions and trials to W was higher than the median response to G, with small effect size (W to G signed rank test, p = $1.7e^{-5}$, r = 0.097, n = 981). To control for differences in expression level of GCaMP6S and to control for the fact that not every neuron imaged was perfectly bisected by the focal plane, we both Z-scored and unity-scaled the data for normalized comparisons.

In the normalized comparisons, we found that differences in responsiveness to W and G were again significant and small (Z-score, signed rank tests: W to G, p = $1.4e^{-9}$, d = 0.388; Unity-scaled, signed rank tests: W to G, p = $3.3e^{-8}$, d = 0.214).

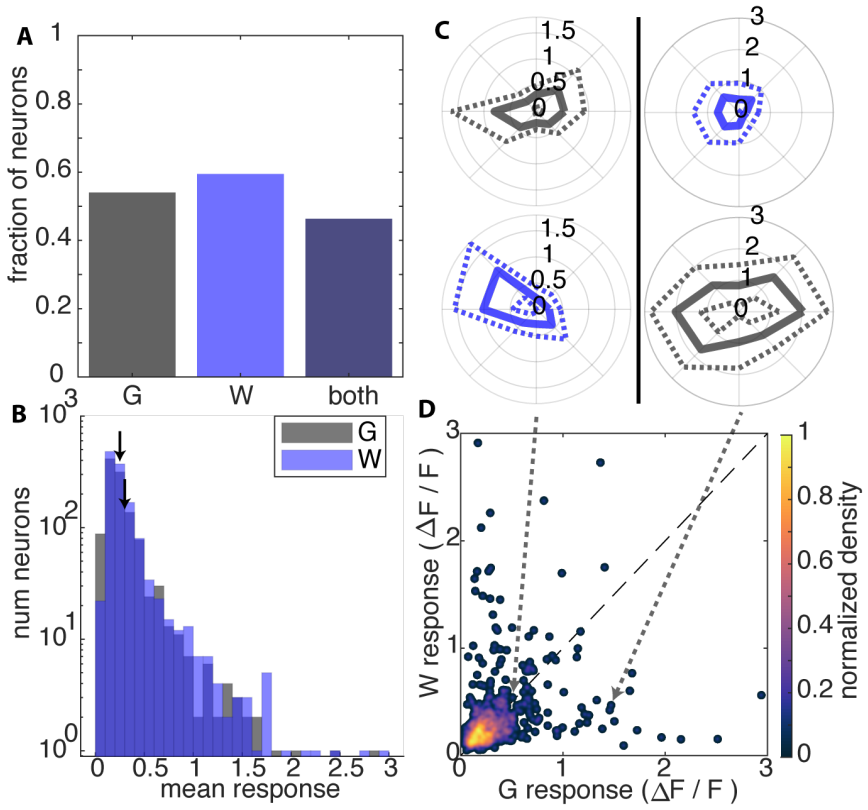

Figure 4.5: **Visual responsiveness to G and W.** A) Barplot of fraction of neurons responsive to G, W, and both G and W. B) Histogram of mean response over directions and trials for neurons responsive to G or W. C) Example tuning curves of neurons responsive to both G and W. D) Heatmap of density of strength of response of neurons responsive to both W and G. Dashed line is unity.

Naturalistic stimuli recruit more directional, asymmetric responses.

We further investigated the tuning properties of murine visual cortex to W and G by comparing direction tuning and orientation tuning. In general, both edge extracting neurons and motion extracting neurons exhibit a preferred axis of motion. The symmetry of the edges in the G movie along a motion axis results in symmetric motion responses, called orientation tuning, detected by averaging responses along an axis. Direction tuned neurons, however, display a motion preference for one direction along this axis over the other. The statistical

tests to detect orientation and direction tuning are not mutually exclusive, and are not a complete subset of one another. Both tests detect motion along an axis, but in different ways. We therefore compared orientation and direction tuning separately, but included in our comparison of tuning strength neurons that were exclusively orientation or direction tuned, and neurons that were both orientation and direction tuned, in the comparisons of each tuning type.

The distribution of orientation of edges in the W stimulus does not give reason to expect symmetric activation of neurons along a motion axis. Nevertheless we did measure the tuning of neurons to the W stimulus along an axis. We refer to these neurons as bi-directionally tuned neurons, because they are likely not extracting the same oriented luminance edge in both directions of their preferred motion axis. We measured 10% of neurons to be exclusively orientation tuned to the W stimulus. Unlike neurons orientation tuned to G, 54% of these neurons were measured to have asymmetric bi-directional tuning, with a stronger response in a direction less than 180°from the maximum direction than the direction opposite the maximal response (ignoring the two directions adjacent to the maximum). Because of the relatively small number of neurons in this uniquely orientation tuned to W group, and the asymmetry in their responses, complicating measurement of a preferred motion axis, we give our attention exclusively in the following analyses to neurons that exhibited significant direction tuning to the W stimulus.

Of neurons shown G and W (n = 2116, 6 mice, 18 sessions), and including both direction and orientation tuning to G, 54% of neurons were tuned to either stimulus class, 41% were tuned to G, 33% were tuned to W, and 20% were tuned to both (Fig. 4.6). The product of the fractions of neurons tuned to each individual stimulus was again lower than the measured fraction of neurons tuned to both stimuli (in the case of G and W, 14% compared to 20%).

Of neurons shown both G and W, 34% were orientation tuned to G, and 18% were direction tuned to G. Of these same neurons, 33% were direction tuned to W, and 9% were direction tuned to both. Doubly-direction-tuned neurons were weakly statistically

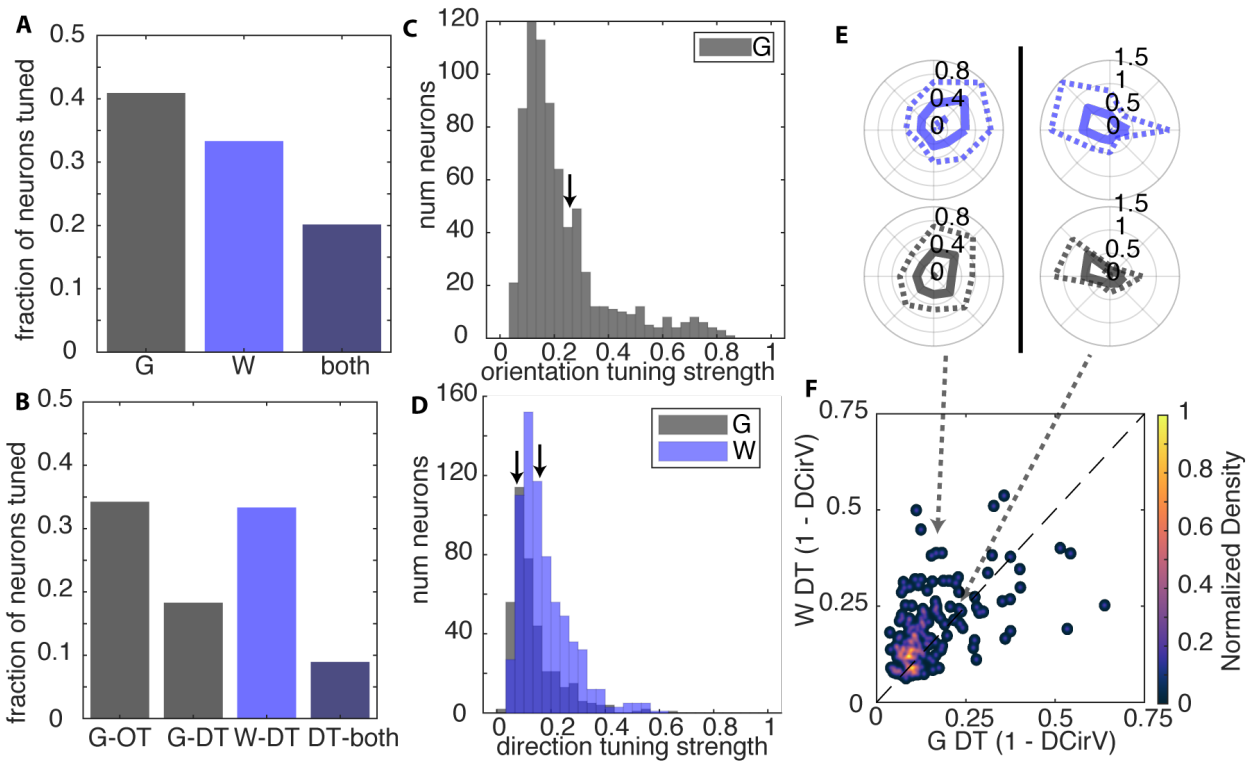

Figure 4.6: **Tuning strength to G and W.** A) Barplot of fraction of neurons tuned to G, W, or both, including both OT and DT to G. B) Barplot of fractions of neurons OT or DT to G or W. C) Histogram of OT strength to G. D) Histogram of DT strength to G and W. E) Exemplary neurons DT to both G and W. F) Heatmap of density of DT strength of neurons DT to G and W. Dashed line is unity.

significantly stronger direction tuned to W over G (signed rank test: p = $7.6e^{-6}$, r = 0.211, n = 225). These results show that more neurons are direction tuned to W than G, the directional tuning strength of neurons tuned to both W and G is slightly shifted towards W, and there is a significant fraction of neurons that encode angular motion information about both W and G.

Neurons orientation tuned to G have uncorrelated preferred angles when direction tuned to W.

We next explored whether directional tuning preference to W is related to axial motion preference to G (Fig. 4.7). Because confidence in the angular preference of a neuron increases with increasing tuning strength, we present these analyses as a function of tuning strength

threshold. By chance, the preferred motion axis to G would be on average 45°different from the preferred direction to W. The preferred angle difference was at every tuning strength threshold very close to chance, and every tuning value was indistinguishable from uniformly distributed (μ = 47.9°at threshold of 0.0 tuning strength, μ = 46.6°at threshold of 2.5, Kolmogorov-Smirnov test: p = $7.4e^{-2}$, p = $4.3e^{-2}$, p = $1.5e^{-1}$, p = $7.4e^{-}1$, p = $7.3e^{-1}$, p = $6.3e^{-1}$). Preferred angle is a first order comparison of tuning preference. We next explored a higher-order comparison: signal correlation in tuning curve as a function of tuning strength. Tuning curves of neurons orientation tuned to G and direction tuned to W were uncorrelated as a function of tuning strength. These results indicate that for neurons orientation tuned to G and direction tuned to W, the preferred direction to W is unrelated to the preferred motion axis to G.

Doubly direction tuned neurons exhibit a similar preferred angle across stimulus conditions.

The same analysis was performed for neurons direction tuned to G and direction tuned to W. By chance, two direction tuning curves would have a preferred angle difference of 90°. When comparing the preferred direction of neurons direction tuned to both G and W, the average preferred angle difference is less than chance, decreases with tuning strength of the neuron, and is statistically significantly different from uniformly distributed at every threshold value (mew = 72.3°at tuning threshold of 0.0; mew = 52.6°at threshold of 2.5, Kolmogorov-Smirnov test: p = $4.2e^{-5}$, p = $9.2e^{-6}$, p = $1.0e^{-4}$, p = $4.1e^{-3}$, p = $2.1e^{-2}$, p = $1.3e^{-2}$). The correlation of tuning curves of neurons direction tuned to both G and W increased as a function of tuning threshold, and at a tuning threshold of 0.25 is much greater than the correlation of tuning curves of neurons orientation tuned to G and direction tuned to W (rank sum test: p = $4.4e^{-2}$, r = $5.6e^{-1}$).

Confidence analysis around preferred angle corroborates preferred angle comparison differences.

The preferred direction or motion axis of a neuron is subject to the uncertainty in the tuning curve, which is measured by the uncertainty of trial-by-trial variance around the mean

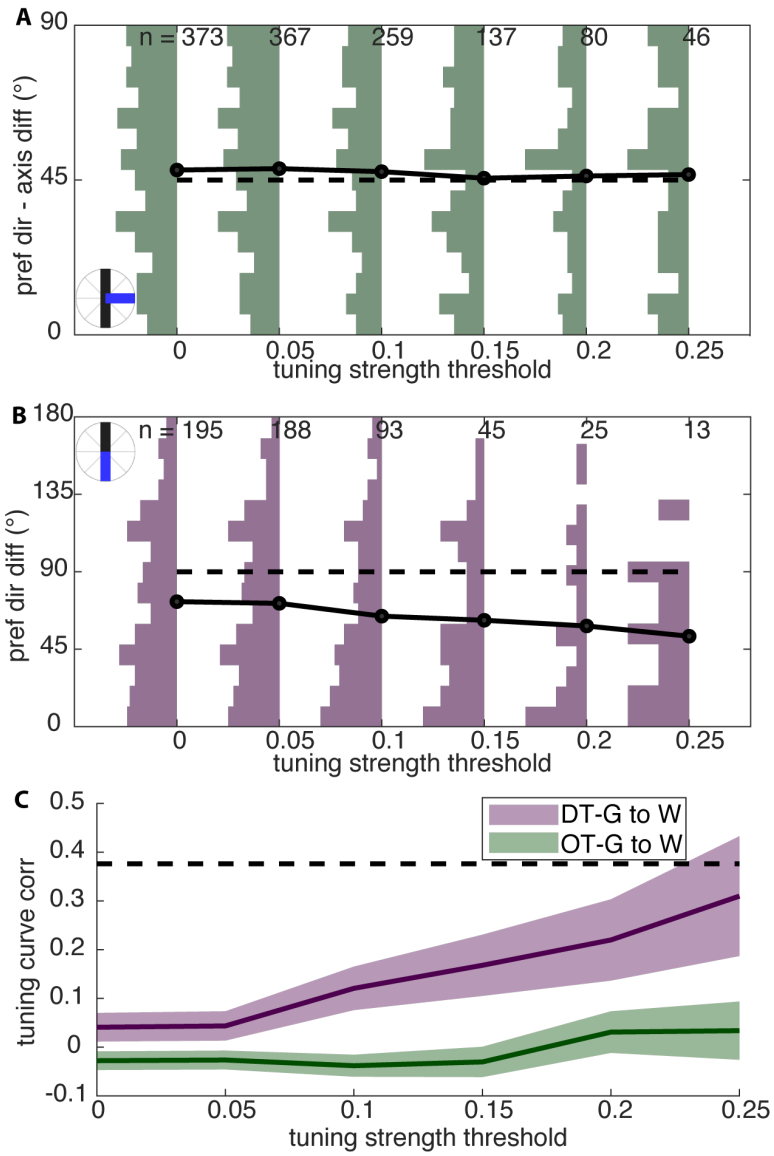

Figure 4.7: **Neurons OT to G do not have a shared motion preference to W, while neurons DT to G have a shared motion preference to W.** A) Angle between preferred axis to G and preferred direction to W, for neurons OT to G and DT to W, as a function of threshold on tuning strength. Neurons with tuning values to both stimuli conditions greater than or equal to each threshold value are included. Inset fills at each point are PDF of angular difference. Dashed line is value of randomly distributed angular preferences. Circular inset is an illustration of maximal angular difference, 90°. B) Same as A), for neurons direction tuned to G and W. C) Tuning curve (signal) correlation of neurons tuned to both G and W (purple - DT to G; green - OT to G). Shaded fill is SEM. Dashed line is one standard deviation of the distribution of tuning curve correlation over all neurons.

response to a stimulus epoch. To measure the influence of this variance on the preferred angle differences detailed above, we propagated this error through our measurements of preferred

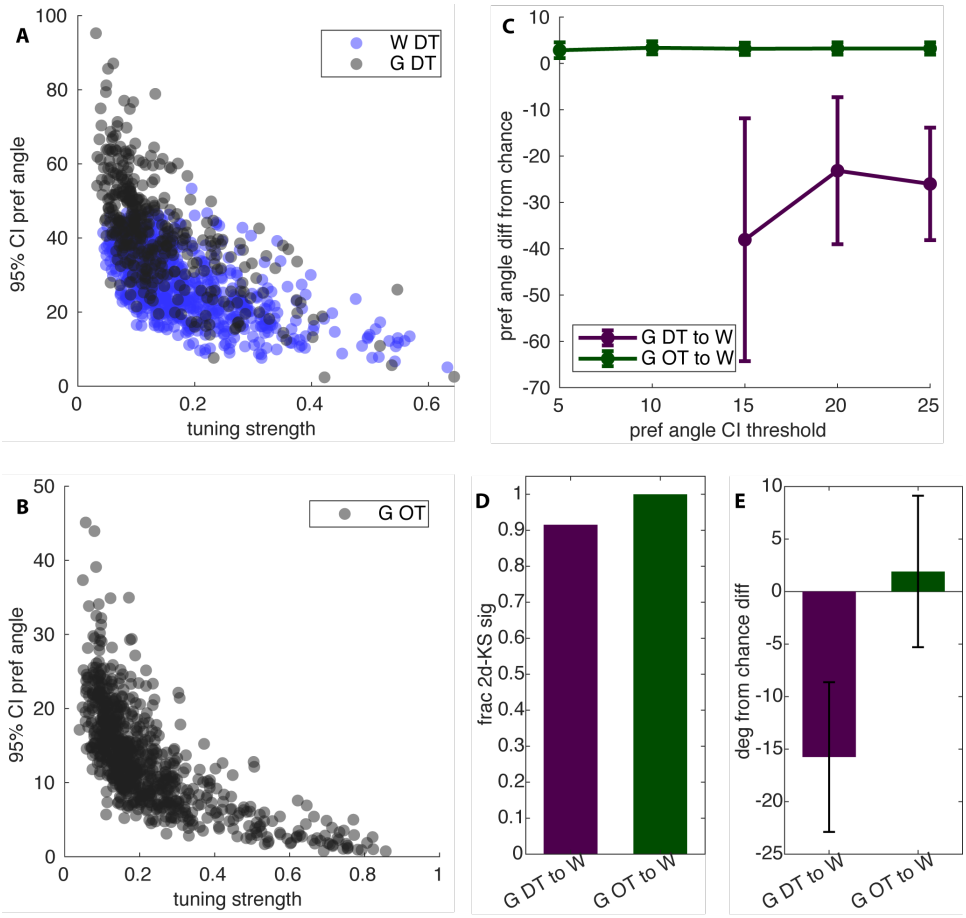

Figure 4.8: **Trial variance effect on confidence in preferred angles.** A) 95% confidence interval of preferred motion direction to W and G as a function of tuning strength. B) 95% confidence interval of preferred motion axis to G as a function of tuning strength. C) Preferred angle difference of G direction tuned neurons (purple) and G orientation tuned neurons (green) to W tuning, as a function of preferred angle 95% confidence interval threshold. Neurons with low uncertainty in preferred angle exhibit the same preferred angle relationship as seen in previous analysis: neurons that extract directional motion from both water and gratings with low directional uncertainty have similar directional preferences across stimuli conditions, while neurons that extract motion axis from gratings and motion direction from water have uncorrelated motion preference across stimuli. D) Fraction of neurons the pass a two-dimensional K-S test on trial response clouds. E) Mean difference from chance preferred angle agreement for neurons from D). Error bars are 95% confidence interval.

direction or preferred axis, to obtain a 95% confidence interval around the preferred angle. The width of the confidence interval of the preferred direction and the preferred axis of neurons to both G and W varied widely for neurons with low tuning strength, and decreased with tuning strength (Fig. 4.8). The confidence interval for the preferred direction to G was

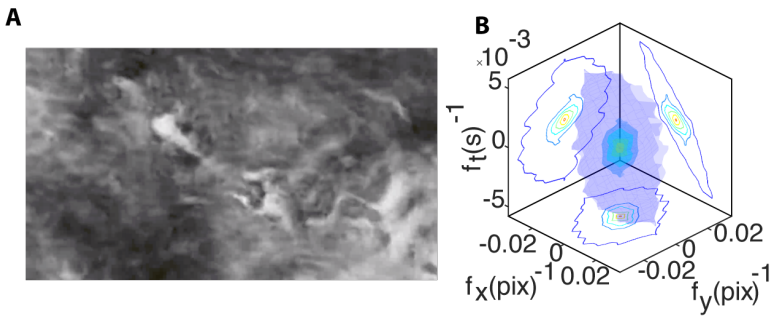

Figure 4.9: **Characterization of water texture stimulus.** A) Still frame of the water texture (T) 0°stimulus. B) Spatiotemporal energy spectrum of T, with broad spatial frequency distributions and oriented x-t energy, similar to W spectrum.

especially variable. Some neurons exhibited a confidence interval in excess of 50°. A large confidence interval around the preferred angle indicates low resolution in measuring angular agreement or disagreement across stimulus conditions. We therefore repeated the preferred angle difference measurement, thresholding by confidence interval around the preferred angle across both conditions instead of tuning strength. We display the preferred angle difference from chance levels (a flat distribution of preferred angle difference between an axis and a direction has a mean of 45°, while a flat distribution of preferred angle difference between two directions is 90°) The comparison of G direction tuned and W tuned neurons again showed a large deviation below chance levels, indicating agreement in preferred direction, while the comparison of G orientation tuned and W tuned neurons remained close to 0°, indicating chance agreement in preferred angles (t-test: $p = 2.2e^{-2}$, $p = 7.3e^{-3}$, $p = 5.2e^{-5}$). The measurement of the response to a stimulus epoch lives in a two-dimensional space, with the angle of the stimulus epoch motion as one dimension, and the mean response to the stimulus as the second. The responses of a neuron to all the stimulus epochs belonging to a stimulus condition is a cloud of such points. To compare the discriminability of these clouds, we performed a two-dimensional Kolmogorov–Smirnov test on the clouds for the response of neurons from each preferred angle comparison (while controlling for a direction or orientation comparison by first folding responses into the correct response space). This statistical significance test combines information from both the uncertainty in the preferred

angles and the uncertainty in the tuning strength. Over 90% of the neurons direction tuned to G and tuned to W have discriminable response clouds across stimulus conditions, and all the neurons orientation tuned to G and tuned to W have discriminable response clouds across stimulus conditions. When we compare the mean preferred angle difference to chance levels for only these discriminable neurons, the same preferred angle comparison result holds (t-test: $p = 1.8e^{-5}$).

Natural motion is not necessary for tuning preference results, as shown by response to water drifting textures.

To control for the swirling, non-standard motions in the water movie, we also constructed a stimulus with the spatial statistics of the water, but the temporal statistics of the gratings - water texture (T), generated from the spatial statistics of a frame of the water stimulus, but uniformly drifting like the gratings (Fig. 4.9). If this stimulus drives the same trend in tuning angular responses, then the result is present in both stimuli containing natural spatial statistics, and does not depend on natural motion signals.

The comparison of motion direction preference of G to T was equivalent to the comparison of G to W (Fig. 4.10). Edge-extracting neurons orientation tuned to G and direction tuned to T displayed a preferred orientation to G unrelated to their preferred direction to W, with angular difference close to chance ($\mu = 47.0°$at tuning threshold of 0.0; $\mu = 37.7°$at threshold of 2.5, Kolmogorov-Smirnov test to uniform distribution: $p = 2.3e^{-1}$, $p = 2.6e^{-1}$ $p = 3.0e^{-1}$, $p = 4.6e^{-1}$, $p = 6.7e^{-1}$, $p = 3.1e^{-1}$). Motion-extracting neurons direction tuned to G and direction tuned to T displayed a preferred direction difference lower than chance ($\mu = 69.9°$at tuning threshold of 0.0; $\mu = 61.4°$at threshold of 2.5, Kolmogorov-Smirnov test: $p = 2.1e^{-6}$, $p = 2.1e^{-6}$, $p = 2.7e^{-5}$, $p = 7.4e^{-2}$, $p = 1.1e^{-1}$, $p = 1.3e^{-2}$). The T stimulus, with natural spatial statistics but uniformly drifting motion statistics, drives similar preferred angle comparisons to G as W, with uncorrelated tuning preferences to neurons orientation tuned to G, but similar tuning preferences to neurons direction tuned to G. Therefore, natural spatial statistics, and not natural motion statistics, drive this result.

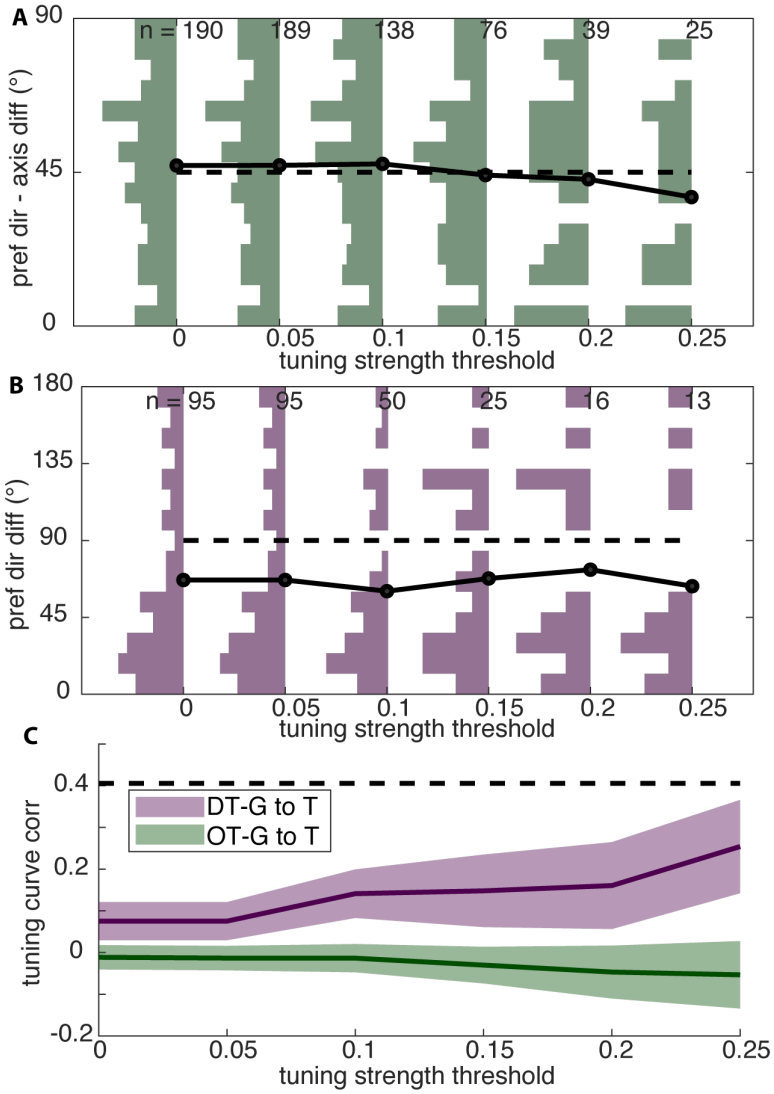

Figure 4.10: **Difference in preferred angle similarity does not depend on natural motion.** A) Angle between preferred axis to G and preferred direction to T, for neurons OT to G and DT to T, as a function of threshold on tuning strength. Inset fills at each point are PDF of angular difference. Dashed line is value of randomly distributed angular preferences. B) Same as A), for neurons direction tuned to G and T. C) Tuning curve (signal) correlation of neurons tuned to both G and W (purple - DT to G; green - OT to G). Shaded fill is SEM. Dashed line is one standard deviation of the distribution of tuning curve correlation over all neurons.

Tuning curves to W explain more trial variance than tuning curves to G.

We also compared the responses to our water flow stimulus to previous reports concerning the cortical response to natural movies. It has been reported that neurons respond more reliably to natural movies than synthetic stimuli [91]. We therefore explored how much of

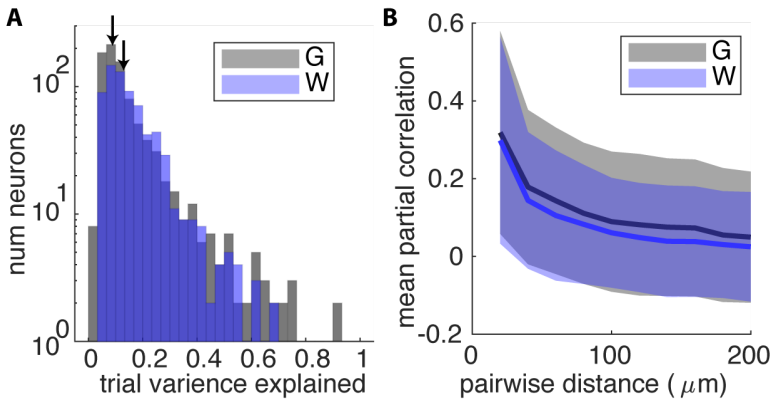

Figure 4.11: **Trial variance explained and response similarity as a function of distance.** A) Histogram of the variance explained by neurons tuned to G or W, with increased trial variance explained by W. B) Partial correlation of pairs of neurons tuned to the same stimulus condition, as a function of Euclidean distance between neurons.

the variance across trials is explained by the trial average (Fig. 4.11). The tuning curve explains 14.7% of the variance of neurons tuned to G, and 15.9% of the variance to W, a small but statistically significant difference (rank sum test: p = $1.7e^{-8}$, r = 0.211, n = 225). We measure more trial average variance explained by the tuning curve to W, indicating more reliable responses to the naturalistic stimulus, as has been previously reported.

The spatial distribution of tuning is similar across stimulus types.

The distribution of tuning preference in rodent primary visual cortex layer 2/3 has been called "salt-and-pepper" because it does not exhibit columnar organization of tuning preference to the extent of other mammals [74]. However, the distribution of tuning preference to grating stimuli does cluster within a small radius [50, 93]. We thus measured whether this property is present under naturalistic stimuli (Fig. 4.11). Considering pairs of neurons tuned with moderate strength (tuning value >= 0.15) to one stimulus in a field-of-view, we measured the partial correlation of these neurons as a function of distance. For both G and W, the correlation within small distances was larger than the correlation at all distances (ranksum test to 20 micron bins to 200 microns, G: p = $2.7e^{-27}$, $2.7e^{-14}$, $3.1e^{-10}$, $5.5e^{-5}$, $2.1e^{-2}$, $3.4e^{-2}$, $2.5e^{-1}$, $2.0e^{-1}$, $2.2e^{-1}$, $1.8e^{-1}$; W: p = $2.2e^{-23}$, $5.5e^{-9}$, $3.5e^{-6}$, $1.5e^{-3}$, $8.1e^{-4}$, $2.3e^{-1}$, $7.4e^{-1}$, $1.8e^{-1}$, $4.7e^{-1}$, $7.2e^{-1}$). These binned values did not show significant differences

tested by Wilcoxon rank sum test across stimulus conditions, within the same distance bin. There is therefore a small but significant increase in tuning curve correlation within small distances for W as well as for G.

4.4 Discussion

Summary and comparison with previous findings

We describe in this study a comparison of responses of murine primary visual cortical neurons to a classic synthetic visual stimulus and to a novel naturalistic stimulus: drifting square wave gratings and water flow. An overlapping population of neurons is driven by each stimulus, and an overlapping population of neurons prefers flow in particular angles driven by each stimulus. Drifting gratings and water flow drives different motion preferences in neurons that extract edge orientation; however, these same stimuli drive similar motion preference in neurons that extract motion from gratings and water.

Because this study utilized a novel naturalistic stimulus, and was performed in relatively contemporary experimental settings, direct comparison of this result to literature is difficult, because to our knowledge this is the first study to directly compare motion tuning direction across a synthetic and naturalistic stimulus. A related study can be found in [40], detailing differences in response of cat striate cortex to visual texture and static and drifting bars. Our study corroborates the previous result of dissimilar preferred directions of drifting texture and an optimally oriented bar stimulus for complex cells. Spatial frequency distributions with strong low frequency power, such as those found in natural images, are known to drive reliable responses in mouse V1 [91]. Our study reproduces this result by measuring increased trial variance explained by the tuning curve to W over G. Neural codes in mouse V1 are sparser under stimulation with natural scenes versus phase-scrambled natural scenes lacking high-order correlations present in the original movie [35]. We do not measure such a difference between our comparison of G and W; however the difference in the previous study was found to be dependent on brain state, and was not present during quiet wakefulness, the

brain state corresponding to much of our data. We do not measure increased sparseness to our natural movies; however Vinje and Gallant 2000 describes that even though naturalistic stimuli produce increasingly sparse activity through activation of the non-classical receptive field surround, the sparseness of the neural response to naturalistic stimuli is statistically indistinguishable from sparseness to drifting gratings, leading the researchers to conclude that sparseness might be a property of the response to oriented spatiotemporal energy, rather than natural spatial and temporal frequency.

Future directions

This study utilized a naturalistic stimulus that is simpler than most natural scenes, because of the lack of foreground, background, and objects, but is more natural than purely synthetic stimuli because it contains natural distributions of spatial and temporal frequency. An increasing number of studies [28] have taken to this approach of simplifying natural movies, or complicating synthetic stimuli, to take steps between purely synthetic or purely natural movies. This approach has already lead to strong research programs, because it incorporates facility of parameterizing the information in the stimuli while including components that provide the capability of the results to extend to understanding of natural movies. This work foreshadows future work taking further steps along the continuum of simple and complex stimuli.

This study details a difference in motion preference agreement across natural and synthetic stimuli of neurons that extract edges or extract motion. This difference predicts a divergence in encoding fidelity of the identity of a stimulus or the motion of a stimulus. These discrepancies are often thought of in terms of the divergence of the cortical ventral and dorsal visual processing streams, but both types of information are bound to be present in primary visual cortex at least in rodents, because rodent V1 is located upstream of the stream bifurcation.

CHAPTER 5

GENERAL DISCUSSION

Drifting square wave gratings and water flow engage a similar proportion of primary visual cortex. This engaged proportion overlaps significantly, allowing direct comparison of response and tuning properties to both stimuli. Edge detecting neurons respond to gratings along a motion axis, but respond to water flow in an unrelated direction, because the orientations of edges in the water flow stimulus are not related to the direction of motion. Motion detecting neurons respond to gratings in one direction, and respond to water along a similar direction, because these neurons extract motion along the same direction in both statistical contexts. In this fashion, we describe response to a naturalistic stimulus that can be explained from the response to a synthetic stimulus. We have given an experimental foothold for extending models informed by single neuron responses to a synthetic stimulus to a naturalistic one. Our work provides numerous paths to push further into understanding the single neuron response across stimuli.

In this work we have provided a perspective on tuning that is rarely appreciated: tuning is a continuum along various dimensions. Recently, due to dense sampling of neurons using calcium imaging, it has become clear that tuning at least lives along an axis from weak to strong, and that there are many more neurons that respond with weak tuning than was originally expected from experiments using electrical probes, fishing for tuned neurons. In this work, we also present another tuning axis: neurons have varying degrees of space-time separable and space-time inseparable tuning. A neuron that responds with significant orientation and direction preferences might display a receptive field that has some space-time orientation, but also has a distinct edge detecting spatial filter. Further study is necessary to probe this relationship, but it at least hints at the prospect of unifying these two computations towards a single model of single neuron tuning properties in visual cortex.

These neurons of course reside in a network, and downstream areas are likely not recruited by a single neuron. Information is transmitted through a brain area by readout of populations

of neurons. The informational content of directional motion of population responses to gratings and water has not been explored. The mostly uni-directional responses of neurons to water flow might result in better motion decoding from this stimulus, versus the motion axis ambiguity produced by symmetrical orientation tuned responses. This dichotomy also suggests a possible decoding dichotomy in visual cortex along decoding content of scenes and decoding motion in scenes. Edge detecting neurons might be useful for detecting what is in a scene, and motion detecting neurons might be useful for detecting the motion of what is in a scene. Since, as previously discussed, neurons encode edges and motion concurrently along a spectrum, there might even be a synergistic relationship of stimulus decoding using both edges and motion to decode visual stimuli. We could begin to address these questions by decoding the direction of motion from our original experiments, but further experimental exploration would be required for a deep answer to these inquiries.

The work presented here has largely concerned neuronal trial-averaged response, but the visual system operates in real time, without averaging over repeated trials. The single neuron response to stimuli is not well explained by the average response, and further work is required to understand where variability arises and how the brain overcomes it. Typically this variability is attributed to noisy sensory channels, and adding some probabilistic noise into models (Gaussian or Poisson) reduces the variance explained by the mean to levels similar to those in data. However, work in the MacLean lab [20] has shown that this variability is predictable from correlated activity of neurons in the network within the same layer. This variability therefore arises from predictable causes, and could be folded into a unifying model of primary visual cortical response. Brain state, neurochemical background, undulating waves of neural activity and other factors combine with the visual stimulus to produce the response of neurons we record. A unified biophysical theory of neural response taking into account these complications might be able to more accurately describe the trial-by-trial response of neurons in sensory cortical areas.

The Palmer group has produced numerous intriguing visual stimuli that are also along the

spectrum from synthetic to natural, that would be useful for illuminating the relationship of response of neurons to natural and synthetic stimuli. Focusing primarily on natural motion, the group has extracted the motion vector fields from natural scenes, which can be used to apply natural motion statistics to synthetic scenes. These movies would allow an experimenter to explore whether the brain is optimized for the motion statistics it is likely to encounter in the natural world. By dropping tracers and textures into these flow fields, these movies also allow exploration of the relative importance of natural spatial statistics to natural motion statistics in neuronal processing.

Visual neuroscience aims to explain the activity of neurons in natural conditions. The field at large is still seeking a comprehensive theory to explain activity in natural conditions. The results presented here take a step towards that aim.

REFERENCES

[1] Edward H. Adelson and James R. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(2):284, February 1985.

[2] A. M. H. J. Aertsen and P. I. M. Johannesma. The Spectro-Temporal Receptive Field: A functional characteristic of auditory neurons. *Biological Cybernetics*, 42(2):133–143, 1981.

[3] D G Albrecht and D B Hamilton. Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology*, 48(1):217–237, July 1982.

[4] Duane G. Albrecht and Wilson S. Geisler. Motion selectivity and the contrast-response function of simple cells in the visual cortex. *Visual Neuroscience*, 7(6):531–546, December 1991.

[5] Joseph J. Atick and A. Norman Redlich. Towards a Theory of Early Visual Processing. *Neural Computation*, 2(3):308–320, September 1990.

[6] David Attwell and Simon B. Laughlin. An Energy Budget for Signaling in the Grey Matter of the Brain. *Journal of Cerebral Blood Flow & Metabolism*, 21(10):1133–1145, October 2001.

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

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

[9] Colin Blakemore and ElisabethA. Tobin. Lateral inhibition between orientation detectors in the cat's visual cortex. *Experimental Brain Research*, 15(4), September 1972.

[10] Paul C. Bush and Terrence J. Sejnowski. Reduced compartmental models of neocortical pyramidal cells. *Journal of Neuroscience Methods*, 46(2):159–166, February 1993.

[11] J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, November 1986.

[12] M Carandini and DJ Heeger. Summation and division by neurons in primate visual cortex. *Science*, 264(5163):1333, May 1994.

[13] E J Chichilnisky. A simple white noise analysis of neuronal light responses. *Network: Computation in Neural Systems*, page 15, 2000.

[14] Thomas W Cronin, Sönke Johnsen, N. Justin Marshall, and Eric Warrant. *Visual Ecology*. 2017. OCLC: 1034268373.

[15] Hod Dana, Tsai-Wen Chen, Amy Hu, Brenda C. Shields, Caiying Guo, Loren L. Looger, Douglas S. Kim, and Karel Svoboda. Thy1-GCaMP6 Transgenic Mice for Neuronal Population Imaging In Vivo. *PLOS ONE*, 9(9):e108697, September 2014.

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

[17] Russell L De Valois, Nicolas P Cottaris, Luke E Mahon, Sylvia D Elfar, and J. Anthony Wilson. Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. *Vision Research*, 40(27):3685–3702, December 2000.

[18] Russell L. De Valois, E. William Yund, and Norva Hepler. The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, 22(5):531–544, January 1982.

[19] G. C. DeAngelis, I. Ohzawa, and R. D. Freeman. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. II. Linearity of temporal and spatial summation. *Journal of Neurophysiology*, 69(4):1118–1135, April 1993.

[20] Joseph B. Dechery and Jason N. MacLean. Functional triplet motifs underlie accurate predictions of single-trial responses in populations of tuned and untuned V1 neurons. *PLOS Computational Biology*, 14(5):e1006153, May 2018.

[21] Mohammad-Reza A. Dehaqani, Abdol-Hossein Vahabie, Roozbeh Kiani, Majid Nili Ahmadabadi, Babak Nadjar Araabi, and Hossein Esteky. Temporal dynamics of visual category representation in the macaque inferior temporal cortex. *Journal of Neurophysiology*, 116(2):587–601, May 2016.

[22] R. Desimone and S. J. Schein. Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *Journal of Neurophysiology*, 57(3):835–868, March 1987.

[23] James J. DiCarlo, Kenneth O. Johnson, and Steven S. Hsiao. Structure of Receptive Fields in Area 3b of Primary Somatosensory Cortex in the Alert Monkey. *Journal of Neuroscience*, 18(7):2626–2645, April 1998.

[24] James J. DiCarlo, Davide Zoccolan, and Nicole C. Rust. How Does the Brain Solve Visual Object Recognition? *Neuron*, 73(3):415–434, February 2012.

[25] Dawei W. Dong and Joseph J. Atick. Statistics of natural time-varying images. *Network: Computation in Neural Systems*, 6(3):345–358, January 1995.

[26] U. C. Drager and D. H. Hubel. Responses to visual stimulation and relationship between visual, auditory, and somatosensory inputs in mouse superior colliculus. *Journal of Neurophysiology*, 38(3):690–713, May 1975.

[27] Ursula C. Dräger. Receptive fields of single cells and topography in mouse visual cortex. *Journal of Comparative Neurology*, 160(3):269–289, 1975.

[28] Luciano Dyballa, Mahmood S. Hoseini, Maria C. Dadarlat, Steven W. Zucker, and Michael P. Stryker. Flow stimuli reveal ecologically appropriate responses in mouse visual cortex. *Proceedings of the National Academy of Sciences*, 115(44):11304–11309, October 2018.

[29] Alexander S. Ecker, Philipp Berens, R. James Cotton, Manivannan Subramaniyan, George H. Denfield, Cathryn R. Cadwell, Stelios M. Smirnakis, Matthias Bethge, and Andreas S. Tolias. State Dependence of Noise Correlations in Macaque Primary Visual Cortex. *Neuron*, 82(1):235–248, April 2014.

[30] Adrienne Fairhall. The receptive field is dead. Long live the receptive field? *Current Opinion in Neurobiology*, 25:ix–xii, April 2014.

[31] Gidon Felsen and Yang Dan. A natural approach to studying vision. *Nature Neuroscience*, 8(12):1643–1646, December 2005.

[32] Eva M. Finney, Ione Fine, and Karen R. Dobkins. Visual stimuli activate auditory cortex in the deaf. *Nature Neuroscience*, 4(12):1171–1173, December 2001.

[33] David J. Fleet, Hermann Wagner, and David J. Heeger. Neural encoding of binocular disparity: Energy models, position shifts and phase shifts. *Vision Research*, 36(12):1839–1857, June 1996.

[34] Tobe C. B Freeman, Séverine Durand, Daniel C Kiper, and Matteo Carandini. Suppression without Inhibition in Visual Cortex. *Neuron*, 35(4):759–771, August 2002.

[35] Emmanouil Froudarakis, Philipp Berens, Alexander S. Ecker, R. James Cotton, Fabian H. Sinz, Dimitri Yatsenko, Peter Saggau, Matthias Bethge, and Andreas S. Tolias. Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. *Nature Neuroscience*, 17(6):851–857, June 2014.

[36] Y.-X. Fu, Hong-Feng Gao, M.-W. Guo, and S.-R. Wang. Receptive field properties of visual neurons in the avian nucleus lentiformis mesencephali. *Experimental Brain Research*, 118(2):279–285, January 1998.

[37] E. P. Gardner and R. M. Costanzo. Spatial integration of multiple-point stimuli in primary somatosensory cortical receptive fields of alert monkeys. *Journal of Neurophysiology*, 43(2):420–443, February 1980.

[38] Geoffrey M. Ghose and Daniel Y. Ts'O. Form Processing Modules in Primate Area V4. *Journal of Neurophysiology*, 77(4):2191–2196, April 1997.

[39] Mark S. Gilzenrat, Sander Nieuwenhuis, Marieke Jepma, and Jonathan D. Cohen. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, affective & behavioral neuroscience*, 10(2):252–269, May 2010.

[40] P. Hammond and D. M. MacKay. Differential responsiveness of simple and complex cells in cat striate cortex to visual texture. *Experimental Brain Research*, 30(2):275–296, November 1977.

[41] H. K. Hartline. The receptive fields of optic nerve fibers. *American Journal of Physiology-Legacy Content*, 130(4):690–699, September 1940.

[42] Andreas V. M. Herz, Tim Gollisch, Christian K. Machens, and Dieter Jaeger. Modeling Single-Neuron Dynamics and Computations: A Balance of Detail and Abstraction. *Science*, 314(5796):80–85, October 2006.

[43] Jennifer L. Hoy, Iryna Yavorska, Michael Wehr, and Cristopher M. Niell. Vision Drives Accurate Approach Behavior during Prey Capture in Laboratory Mice. *Current Biology*, 26(22):3046–3052, November 2016.

[44] D. H. Hubel and T. N. Wiesel. Receptive fields of single neurones in the cat's striate cortex. *The Journal of Physiology*, 148(3):574–591, October 1959.

[45] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, 160(1):106–154.2, January 1962.

[46] Andrew D. Huberman and Cristopher M. Niell. What can mice tell us about how vision works? *Trends in Neurosciences*, 34(9):464–473, September 2011.

[47] Uwe J. Ilg and Guillaume S. Masson, editors. *Dynamics of Visual Motion Processing: Neuronal, Behavioral, and Computational Approaches*. Springer, New York, 2010. OCLC: ocn401157138.

[48] Alumit Ishai, Leslie G. Ungerleider, Alex Martin, Jennifer L. Schouten, and James V. Haxby. Distributed representation of objects in the human ventral visual pathway. *Proceedings of the National Academy of Sciences*, 96(16):9379–9384, August 1999.

[49] Joni-Kristian Kamarainen. Gabor features in image analysis. In *2012 3rd International Conference on Image Processing Theory, Tools and Applications (IPTA)*, pages 13–14, Istanbul, Turkey, October 2012. IEEE.

[50] Björn M. Kampa, Morgane M. Roth, Werner Göbel, and Fritjof Helmchen. Representation of visual scenes by local neuronal populations in layer 2/3 of mouse visual cortex. *Frontiers in Neural Circuits*, 5, December 2011.

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

[52] Christoph Kayser, Rodrigo F. Salazar, and Peter König. Responses to Natural Scenes in Cat V1. *Journal of Neurophysiology*, 90(3):1910–1920, September 2003.

[53] E. Kobatake and K. Tanaka. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, 71(3):856–867, March 1994.

[54] Michael F. Land and Dan-Eric Nilsson. *Animal Eyes*. Oxford Animal Biology Series. Oxford University Press, Oxford ; New York, 2nd ed edition, 2012. OCLC: ocn761379932.

[55] Joonyeol Lee and John H. R. Maunsell. A Normalization Model of Attentional Modulation of Single Unit Responses. *PLOS ONE*, 4(2):e4651, February 2009.

[56] Tai Sing Lee and David Mumford. Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A*, 20(7):1434, July 2003.

[57] Peter Lennie. The Cost of Cortical Computation. *Current Biology*, 13(6):493–497, March 2003.

[58] William B Levy and Robert A. Baxter. Energy Efficient Neural Codes. *Neural Computation*, 8(3):531–543, April 1996.

[59] Nikos K. Logothetis, Jon Pauls, and Tomaso Poggio. Shape representation in the inferior temporal cortex of monkeys. *Current Biology*, 5(5):552–563, May 1995.

[60] Linda Madisen, Theresa A. Zwingman, Susan M. Sunkin, Seung Wook Oh, Hatim A. Zariwala, Hong Gu, Lydia L. Ng, Richard D. Palmiter, Michael J. Hawrylycz, Allan R. Jones, Ed S. Lein, and Hongkui Zeng. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neuroscience*, 13(1):133–140, January 2010.

[61] D. Marr and T. Poggio. From Understanding Computation to Understanding Neural Circuitry. May 1976.

[62] Mark Mazurek, Marisa Kager, and Stephen D. Van Hooser. Robust quantification of orientation selectivity and direction selectivity. *Frontiers in Neural Circuits*, 8, August 2014.

[63] David A McCormick, Matthew J McGinley, and David B Salkoff. Brain state dependent activity in the cortex and thalamus. *Current Opinion in Neurobiology*, 31:133–140, April 2015.

[64] James M. McFarland, Adrian G. Bondy, Bruce G. Cumming, and Daniel A. Butts. High-resolution eye tracking using V1 neuron activity. *Nature Communications*, 5:4605, September 2014.

[65] Judith McLean and Larry A. Palmer. Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. *Vision Research*, 29(6):675–679, January 1989.

[66] Naohisa Miyakawa, Kei Majima, Hirohito Sawahata, Keisuke Kawasaki, Takeshi Matsuo, Naoki Kotake, Takafumi Suzuki, Yukiyasu Kamitani, and Isao Hasegawa. Heterogeneous Redistribution of Facial Subcategory Information Within and Outside the Face-Selective Domain in Primate Inferior Temporal Cortex. *Cerebral Cortex*, 28(4):1416–1431, April 2018.

[67] Aage R. Møller. Frequency selectivity of phase-locking of complex sounds in the auditory nerve of the rat. *Hearing Research*, 11(3):267–284, September 1983.

[68] J A Movshon, I D Thompson, and D J Tolhurst. Receptive field organization of complex cells in the cat's striate cortex. *The Journal of Physiology*, 283:79–99, October 1978.

[69] J. Anthony Movshon, Edward H. Adelson, Martin S. Gizzi, and William T. Newsome. The Analysis of Moving Visual Patterns. In Carlos Chagas, Ricardo Gattass, and Charles Gross, editors, *Pattern Recognition Mechanisms*, volume 11, pages 117–151. Springer Berlin Heidelberg, Berlin, Heidelberg, 1985.

[70] Dylan Richard Muir, Morgane Roth, Fritjof Helmchen, and Bjoern Kampa. Model-based analysis of pattern motion processing in mouse primary visual cortex. *Frontiers in Neural Circuits*, 9, 2015.

[71] Cristopher M. Niell and Michael P. Stryker. Highly Selective Receptive Fields in Mouse Visual Cortex. *The Journal of Neuroscience*, 28(30):7520–7536, July 2008.

[72] Cristopher M. Niell and Michael P. Stryker. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron*, 65(4):472–479, February 2010.

[73] Eyal I. Nitzany and Jonathan D. Victor. The statistics of local motion signals in naturalistic movies. *Journal of Vision*, 14(4):10–10, April 2014.

[74] Kenichi Ohki, Sooyoung Chung, Yeang H. Ch'ng, Prakash Kara, and R. Clay Reid. Functional imaging with cellular resolution reveals precise micro-architecture in visual cortex. *Nature*, 433(7026):597–603, February 2005.

[75] B Olshausen and D Field. Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14(4):481–487, August 2004.

[76] Bruno A Olshausen. What is the other 85% of V1 doing? In *23 Problems in Systems Neuroscience. T.J. Sejnowski, L. van Hemmen, Eds.*, page 29. Oxford University Press, 2005.

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

[78] Bruno A. Olshausen and David J. Field. How Close Are We to Understanding V1? *Neural Computation*, 17(8):1665–1699, August 2005.

[79] Srdjan Ostojic and Nicolas Brunel. From Spiking Neuron Models to Linear-Nonlinear Models. *PLoS Computational Biology*, 7(1):e1001056, January 2011.

[80] Il Memming Park and Jonathan W. Pillow. Bayesian Spike-Triggered Covariance Analysis. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 24*, pages 1692–1700. Curran Associates, Inc., 2011.

[81] Jonathan W. Pillow and Eero P. Simoncelli. Dimensionality reduction in neural models: An information-theoretic generalization of spike-triggered average and covariance analysis. *Journal of Vision*, 6(4):9–9, February 2006.

[82] Pierre-Olivier Polack, Jonathan Friedman, and Peyman Golshani. Cellular mechanisms of brain state–dependent gain modulation in visual cortex. *Nature Neuroscience*, 16(9):1331–1339, September 2013.

[83] Javier Portilla and Eero P. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *International journal of computer vision*, 40(1):49–70, 2000.

[84] G. T. Prusky and R. M. Douglas. Characterization of mouse cortical spatial vision. *Vision Research*, 44(28):3411–3418, December 2004.

[85] R. C. Reid, J. D. Victor, and R. M. Shapley. The use of m-sequences in the analysis of visual neurons: Linear receptive field properties. *Visual Neuroscience*, 14(6):1015–1027, November 1997.

[86] Jacob Reimer, Emmanouil Froudarakis, Cathryn R. Cadwell, Dimitri Yatsenko, George H. Denfield, and Andreas S. Tolias. Pupil Fluctuations Track Fast Switching of Cortical States during Quiet Wakefulness. *Neuron*, 84(2):355–362, October 2014.

[87] Pamela Reinagel and Anthony M Zador. Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, 10:1–10, 1999.

[88] John H. Reynolds and David J. Heeger. The Normalization Model of Attention. *Neuron*, 61(2):168–185, January 2009.

[89] Fred Rieke and Michael E. Rudd. The Challenges Natural Images Pose for Visual Adaptation. *Neuron*, 64(5):605–616, December 2009.

[90] Maximilian Riesenhuber and Tomaso Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–1025, November 1999.

[91] Rajeev V. Rikhye and Mriganka Sur. Spatial Correlations in Natural Scenes Modulate Response Reliability in Mouse Visual Cortex. *Journal of Neuroscience*, 35(43):14661–14680, October 2015.

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

[93] Dario L. Ringach, Patrick J. Mineault, Elaine Tring, Nicholas D. Olivas, Pablo Garcia-Junco-Clemente, and Joshua T. Trachtenberg. Spatial clustering of tuning in mouse primary visual cortex. *Nature Communications*, 7, August 2016.

[94] John Robson. Receptive fields: Spatial and intensive representation of the visual image. In *Handbook of Perception*. Academic Press, New York, NY, 1975.

[95] A. W. Roe and D. Y. Ts'o. Visual topography in primate V2: Multiple representation across functional stripes. *Journal of Neuroscience*, 15(5):3689–3715, May 1995.

[96] Daniel L. Ruderman. Origins of scaling in natural images. *Vision Research*, 37(23):3385–3398, December 1997.

[97] Daniel L. Ruderman and William Bialek. Statistics of natural images: Scaling in the woods. *Physical Review Letters*, 73(6):814–817, August 1994.

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

[99] Nicole C. Rust, Odelia Schwartz, J. Anthony Movshon, and Eero P. Simoncelli. Spatiotemporal Elements of Macaque V1 Receptive Fields. *Neuron*, 46(6):945–956, June 2005.

[100] Alexander J. Sadovsky, Peter B. Kruskal, Joseph M. Kimmel, Jared Ostmeyer, Florian B. Neubauer, and Jason N. MacLean. Heuristically optimal path scanning for high-speed multiphoton circuit imaging. *Journal of Neurophysiology*, 106(3):1591–1598, June 2011.

[101] Tomoya Sakatani and Tadashi Isa. PC-based high-speed video-oculography for measuring rapid eye movements in mice. *Neuroscience Research*, 49(1):123–131, May 2004.

[102] Frank Sengpiel, Roland J Baddeley, Tobe C. B Freeman, Richard Harrad, and Colin Blakemore. Different mechanisms underlie three inhibitory phenomena in cat area 17. *Vision Research*, 38(14):2067–2080, July 1998.

[103] Tatyana Sharpee, Nicole C. Rust, and William Bialek. Analyzing Neural Responses to Natural Signals: Maximally Informative Dimensions. *Neural Computation*, 16(2):223–250, February 2004.

[104] Daisuke Shimaoka, Kenneth D. Harris, and Matteo Carandini. Effects of Arousal on Mouse Sensory Cortex Depend on Modality. *Cell Reports*, 22(12):3160–3167, March 2018.

[105] A. M. Sillito. Inhibitory processes underlying the directional specificity of simple, complex and hypercomplex cells in the cat's visual cortex. *The Journal of Physiology*, 271(3):699–720, 1977.

[106] Eero P Simoncelli and Bruno A Olshausen. Natural Image Statistics and Neural Representation. *Annual Review of Neuroscience*, 24(1):1193–1216, 2001.

[107] Lothar Spillmann, Birgitta Dresp-Langley, and Chia-huei Tseng. Beyond the classical receptive field: The effect of contextual stimuli. *Journal of Vision*, 15(9):7–7, July 2015.

[108] B. E. Stein and N. S. Gaither. Receptive-field properties in reptilian optic tectum: Some comparisons with mammals. *Journal of Neurophysiology*, 50(1):102–124, July 1983.

[109] Barry E. Stein and Neal S. Gaither. Sensory representation in reptilian optic tectum: Some comparisons with mammals. *Journal of Comparative Neurology*, 202(1):69–87, 1981.

[110] Anja K. Sturm and Peter König. Mechanisms to synchronize neuronal activity. *Biological Cybernetics*, 84(3):153–172, February 2001.

[111] Deqing Sun, Stefan Roth, and Michael J. Black. Secrets of optical flow estimation and their principles. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2432–2439, 2010.

[112] Y. Tadmor and D. J. Tolhurst. The effect of threshold on the relationship between the receptive-field profile and the spatial-frequency tuning cure in simple cells of the cat's striate cortex. *Visual Neuroscience*, 3(5):445–454, November 1989.

[113] Ivar L. Thorson, Jean Liénard, and Stephen V. David. The Essential Complexity of Auditory Receptive Fields. *PLOS Computational Biology*, 11(12):e1004628, December 2015.

[114] Michelle P. S. To, Iain D. Gilchrist, and David J. Tolhurst. Perception of differences in naturalistic dynamic scenes, and a V1-based model. *Journal of Vision*, 15(1):19–19, January 2015.

[115] D. J. Tolhurst, Y. Tadmor, and Tang Chao. Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, 12(2):229–232, 1991.

[116] V. Torre and T. Poggio. A synaptic mechanism possibly underlying directional selectivity to motion. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 202(1148):409–416, July 1978.

[117] Jon Touryan, Gidon Felsen, and Yang Dan. Spatial Structure of Complex Cell Receptive Fields Measured with Natural Images. *Neuron*, 45(5):781–791, March 2005.

[118] J. H. van Hateren. Real and optimal neural images in early vision. *Nature*, 360(6399):68–70, November 1992.

[119] J H van Hateren and D L Ruderman. Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. *Proceedings of the Royal Society B: Biological Sciences*, 265(1412):2315–2320, December 1998.

[120] J H van Hateren and A van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1394):359–366, March 1998.

[121] Meel Velliste, Sagi Perel, M. Chance Spalding, Andrew S. Whitford, and Andrew B. Schwartz. Cortical control of a prosthetic arm for self-feeding. *Nature*, 453(7198):1098–1101, June 2008.

[122] Martin Vinck, Renata Batista-Brito, Ulf Knoblich, and Jessica A. Cardin. Arousal and Locomotion Make Distinct Contributions to Cortical Activity Patterns and Visual Encoding. *Neuron*, 86(3):740–754, May 2015.

[123] William E. Vinje and Jack L. Gallant. Natural Stimulation of the Nonclassical Receptive Field Increases Information Transmission Efficiency in V1. *The Journal of Neuroscience*, 22(7):2904–2915, January 2002.

[124] Martin J Wainwright, Odelia Schwartz, and Eero P Simoncelli. Natural Image Statistics and Divisive Normalization: Modeling Nonlinearities and Adaptation in Cortical Neurons. In *Statistical Theories of the Brain*, page 22. MIT Press, Cambridge, Mass, 2001. In Press.

[125] Quanxin Wang and Andreas Burkhalter. Area map of mouse visual cortex. *Journal of Comparative Neurology*, 502(3):339–357, 2007.

[126] Quanxin Wang, Enquan Gao, and Andreas Burkhalter. Gateways of Ventral and Dorsal Streams in Mouse Visual Cortex. *Journal of Neuroscience*, 31(5):1905–1918, February 2011.

[127] Quanxin Wang, Olaf Sporns, and Andreas Burkhalter. Network Analysis of Corticocortical Connections Reveals Ventral and Dorsal Processing Streams in Mouse Visual Cortex. *Journal of Neuroscience*, 32(13):4386–4399, March 2012.

[128] Michael Weliky, József Fiser, Ruskin H Hunt, and David N Wagner. Coding of Natural Scenes in Primary Visual Cortex. *Neuron*, 37(4):703–718, February 2003.

[129] Ian R. Wickersham, David C. Lyon, Richard J. O. Barnard, Takuma Mori, Stefan Finke, Karl-Klaus Conzelmann, John A. T. Young, and Edward M. Callaway. Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons. *Neuron*, 53(5):639–647, March 2007.

[130] S. M. Zeki. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. https://physoc.onlinelibrary.wiley.com/doi/abs/10.1113/jphysiol.1974.sp010452, February 1974.

[131] Feng Zhang, Alexander M. Aravanis, Antoine Adamantidis, Luis de Lecea, and Karl Deisseroth. Circuit-breakers: Optical technologies for probing neural signals and systems. *Nature Reviews Neuroscience*, 8(8):577–581, August 2007.

[132] Siyu Zhang, Min Xu, Tsukasa Kamigaki, Johnny Phong Hoang Do, Wei-Cheng Chang, Sean Jenvay, Kazunari Miyamichi, Liqun Luo, and Yang Dan. Long-range and local circuits for top-down modulation of visual cortex processing. *Science*, 345(6197):660–665, August 2014.