THE UNIVERSITY OF CHICAGO

STRUCTURED VARIABILITY AND TEMPORAL HETEROGENEITY CONSTRAIN THE CORTICAL CODE FOR MOTION DIRECTION

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

COMMITTEE ON COMPUTATIONAL NEUROSCIENCE

BY
JOSEPH ALLEN LOMBARDO

CHICAGO, ILLINOIS
AUGUST 2018

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Firstly, I would like to thank my advisor, Stephanie Palmer, without whose support and guidance I would not have made it through the last six years. The intelligence, enthusiasm, and intellectual curiosity that Stephanie brought to lab every day has been a light and an inspiration.

I would like to thank the members of the Palmer lab, especially Nathan Buerkle Jared Salisbury, Audy Sederberg, and Eyal Nitzany, who created a great lab community and were always there for discussing ideas scientific or less so. Björn Beyer for his work and initial insights in the temporal heterogeneity project.

I would like to thank the members of the Osborne lab for the contribution to this work. Matt Macellaio and Bing Liu, for enormous help in physiological experiments, and Trishna Mukherjee for being an invaluable resource in the lab and for help working with the animals. And of course I would like to thank Leslie Osborne for her intellectual contributions to these works and for covering for the deficits in my writing ability.

I would like to thank Summer Thompson, who has been wonderfully patient and supportive.

I would like to thank my family for being incredibly supportive, even when that means letting me disappear into stress and work. They have done so much to give me the opportunities to get me where I am, and this work would not have been possible without them.

Finally, I would like to thank the other members of my thesis committee, Nicolas Brunel, Nicho Hatsopoulos, and especially my committee chair David Freedman. The time, feedback, and support has been an invaluable contribution to this work.

ABSTRACT

Neurons in cortical area MT respond to visual motion stimuli and drive downstream perception of motion and smooth pursuit eye movements. Each neuron responds to parametric changes in stimulus value, and the responses are pooled across a large population of neurons in a population code. The accuracy of such a code can depend subtly on specific features of the response properties of the neurons that make up the code. Trial-to-trial variability, correlated responses, response heterogeneity, and temporal dynamics can all affect how a neural population codes information. This work characterizes the variability of neurons in MT, and describes how the stimulus-dependent structure of this variability can impact the population code. A model of stimulus-dependent variability bridges this result with previous descriptions of variability in MT that lacked stimulus-dependence, and demonstrates how changes in attentional or behavioral state can alter the structure of this variability. This work also characterizes and models the diversity of temporal dynamics in MT neurons, and explores the contribution of temporal heterogeneity to population codes. A generative model of dynamic MT responses provides a foundation for further investigation of the population code.

CHAPTER 1

GENERAL INTRODUCTION

The goal of computational neuroscience, broadly, is to connect sensory inputs to behavioral outputs through an understanding of how populations of neurons process information. Neurons vary their firing rates in relation to varying stimulus features, and behavior is in turn correlated with that firing. The neural population code lies between these two end points. This work examines cortical area MT as one such neural population that clearly connects sensory inputs to behavioral outputs, and thus provides ample opportunity for exploring the neural code from both ends. The neurons of area MT respond to visual motion stimuli, and drive smooth pursuit eye movements. This introductory material provides background on the key smooth pursuit eye movements and how they are connected to cortical area MT. Subsequently some background is provided on the understanding and challenges of population coding in MT.

1.1 Area MT and its Role in the Visual Motion and Smooth Pursuit Pathways

Smooth pursuit is a voluntary eye movement common to primates that smoothly rotates the eye to keep the image of a small moving object stabilized on the retina. Humans have a fovea, a small central area of high acuity vision for which most of the visual processing in cortex is dedicated. In order to see a moving object clearly, the fovea must be stabilized on the retina by smooth pursuit eye movements. Thus smooth pursuit eye movements are an important physiological behavior for animal survival. Activities from a predator tracking prey to a human hitting a baseball depend on accurate smooth pursuit tracking of visual targets. Maintaining a stabilized image of a moving target on the retina requires a finely tuned mechanism that can act on very rapid time scales, taking visual motion information and finely controlling the muscles of the eye. This computation is thought to be done by

largely by cortical area MT. The following introduction examines the properties of smooth pursuit eye movements, the properties of neurons in MT, and connections between the two. This lays the foundation and justification for subsequent investigations of population coding in MT, which depend on the connection between visual motion stimuli and smooth eye movement behavior.

1.1.1 Properties of Smooth Pursuit

Smooth pursuit is unusual among eye movements in that while it is voluntary, it generally requires a small, smoothly moving target [114]. Anecdotally, individuals prompted to translate their eyes smoothly are unable to do so, instead making a series of small jumps at approximately 200ms intervals [61]. In the presence of a small moving target, pursuit involves a combination of smooth pursuit eye movements and small "catch up" saccades to keep the target centered on the fovea. Smooth pursuit is considered the differentiable part of pursuit, between the saccades.

The step-ramp stimulus has been an important tool for studying the properties of smooth pursuit [114]. The stimulus involves a fixation target, followed by a target step to a region eccentric of the foveal visual field immediately followed by smooth translation of the target. This stimulus is able to disentangle motion and position signals by stepping in one direction and ramping in the opposite direction. With the step-ramp stimulus, pursuit begins with a smooth translation in the direction of the target motion, followed by a small saccade to correct for the position error [114]. This suggests that smooth pursuit is responsive to the motion signal but not the position signal.

Additionally, with the step-ramp stimulus latency of pursuit eye movements could be disentangled from saccade latency, as it allowed pursuit to be initiated without a saccade. Smooth pursuit latency was found to be around 80-100 ms in macaques [62]. Pursuit latency is greater in humans than macaques, at approximately 125ms [117]. Pursuit initiation was recently found to be rather slow in marmosets, at about 200ms latency, although the relia-

bility of the study is called into question by systematic errors in their measurements of eye velocity [85]. The broad features of smooth pursuit appear to be common across primates, with variability in latency, gain, and accuracy.

Eye velocity during pursuit is not able to perfectly match target velocity, with a gain slightly less than one. The saturation of the pursuit system was for a long time reported to occur for targets moving at 30°/s, at which point pursuit gain would decrease, error would increase, and pursuit would require an excess of "catch up" saccades [41, 61, 154]. Systematic study, however, found pursuit to saturate at much higher velocities in humans, typically 90°/s [84]. The gain in the pursuit system is about 0.9 up until that saturation point, meaning eye velocity during pursuit is 90% of target velocity [84].

The first 20ms of pursuit match the direction of target motion, but pursuit speed is independent of target speed or position [62]. For step-ramp stimuli in the same direction but different speeds, the initial 20ms of smooth pursuit follow the same path, but diverge over the period from 20-80ms as the eye accelerates to match target velocity [62].

Retinal errors during smooth pursuit are unable to affect pursuit velocity until the after the subsequent 100 ms, as image motion and position information requires time to propagate through the pursuit system [62]. Lisberger and Westbrook used a step-ramp stimulus that, as soon as pursuit was initiated, would take into account the eye movements and artificially imposing retinal errors that could be corrected [62]. In this condition, the eye must continually use visual feedback to correct its velocity. The eye velocity trace for the first 100 ms of pursuit under this condition is the same as if the target were translating at a constant velocity, but after 100 ms the eye accelerates in attempt to correct the imposed error. Additionally, randomly changing the speed of the target during pursuit does not affect the speed of pursuit until 100 ms after the change in target speed. This strongly indicates that the first 100 ms of pursuit do not take into account the velocity of the eye, but instead rely on the estimate generate from image motion velocity during the first 80-100 ms prior to pursuit initiation. This so-called "open loop period" of pursuit is therefore purely sensory in

origin, allowing for comparisons of pursuit behavior to the underlying visual motion neural pathways.

The initiation of pursuit is inherently different from the maintenance of pursuit, since the role of smooth pursuit is to stabilize the image of a moving target on the retina. Initiation of pursuit requires motion of the image on the retina, but once pursuit is established, retinal slip is decreased and pursuit instead depends on small retinal errors about the fovea as well as estimates of eye velocity. Models of pursuit relying solely on visual feedback with a 100 ms delay can easily become unstable, leading to oscillations of eye velocity during pursuit [57, 61]. Instead, pursuit must rely a combination of efference copies of eye velocity commands and nonlinearities that take into account both retinal slip velocity and acceleration.

For periodic or predictable moving stimuli, the eye is able to make predictive pursuit movements that anticipate target motion [13, 23]. Predictive pursuit movements provide a useful tool for study because they show that at least some of the inputs to the pursuit system are not purely sensory and instead depend on memories of target velocity. This can help determine the neural pathways controlling pursuit.

Although smooth pursuit and saccadic eye movements operate through separate pathways, there is evidence of shared motion information between the two systems. Lesions in visual cortex that generate smooth pursuit deficits generate corresponding deficits in the compensation of saccadic movements to moving stimuli [33, 96, 128].

1.1.2 Role of MT in Pursuit

Lesions of MT help to reveal the role of MT as a source of visual motion information used to generate eye movements. A study by [96] found that small chemical lesions in MT using ibotenic acid create retinotopic deficits in pursuit, i.e. deficits in pursuing targets moving within the region of the visual field affected by the lesion. Using a step-ramp moving stimulus (see [114]), deficits in pursuit velocity were found only when the target was stepped into a localized region of the visual field. The retinotopic deficit corresponded to the center of the

receptive field at the site of the ibotenic acid injection. Subjects compensated for pursuit deficits velocity by making a saccade such that the image of the target was in an intact portion of the visual field, and pursuit continued normally. Saccades to stationary targets were unaffected by the lesions, but saccades to moving targets in the area of the visual field affected by the lesion did not compensate for the motion of the target, resulting in a greater error away from the direction of motion. This supports the idea that MT provides sensory estimates of target motion rather than motor commands. The motion signals would be used for both smooth pursuit and saccades.

The deficits produced by such focal chemical lesions disappeared rapidly, recovering to pre-lesion performance after a few days [114, 159]. The study by [159] also included a large unilateral lesion that included "all of MT and most, if not all, of MST," as well as "parts of V4, V4t, V3, possibly V3a, V2, a small portion of V1, possibly the saccade-related lateral portion of [LIP], and a portion of the lateral sulcus." Despite the extent of damage to visual cortex, motion information was still available to the pursuit system from outside the damaged the areas. Eye velocity was unable to match target velocity of 16°/s, but "even on the day of the injection when the pursuit deficit was at its maximum, the pursuit speed was 4.4°/s and not 0°/s." The large lesion generated severe retinotopic deficits (contralateral visual field) and more modest directional deficits (ipsilateral motion direction). Recovery from the lesion was incomplete after 7 months.

Electrical microstimulation studies further support the hypothesis that MT provides a sensory motion signals for both pursuit and saccades. One study tested the effects of microstimulation on smooth pursuit direction and velocity as well as saccade accuracy to moving and stationary targets [47]. For target motion in along the preferred direction of the stimulated cells, the stimulation increased the speed of pursuit. If the target motion were 180° in the opposite direction, stimulation had the effect of decreasing pursuit speed. For target motion in any other direction, the direction of pursuit was biased towards the preferred direction of the stimulated cells. Similarly for trials where saccades were mad to stationary targets,

eye position error was biased towards the preferred direction of the stimulated neurons as if compensating for target motion. However, a few trials had opposite effects for saccades and pursuit; this is still consistent with MT providing a shared motion signal for pursuit and saccades, but confirms that MT is likely not the only source of motion information for eye movements. In addition to affecting pursuit velocity and eliciting motion compensation in saccades, stimulation could also trigger pursuit for a stationary target after it was stepped into the receptive field, but had no effect in the absence of a target.

Microstimulation in the foveal representation of MT, MTf, generated ipsilateral accelerations during pursuit rather than acceleration toward some other preferred direction [60]. This may indicate some extraretinal representation of pursuit in MTf that is not present in the rest of MT.

1.1.3 Response Properties of MT

Cortical area MT is part of the visual pathway that plays an important role in visual motion processing. Neurons in MT respond selectively to localized motion of a stimulus within a receptive field. Most MT neurons are strongly directly tuned, responding most strongly to motion along a preferred direction, with response amplitude on average decaying by half for motion $\pm 40^\circ$ away from the preferred stimulus direction [3, 76]. Gaussian curves provide "an excellent fit" to the direction tuning of MT neurons [3]. In addition to selectivity to direction of motion, MT neurons also exhibited tuning for the speed of moving stimuli, the orientation of flashed, stationary stimuli, and the binocular disparity of the images from each retina [76, 77]. Preferred speeds of stimuli ranged from 2°/s to 256°/s, with an average near 32°/s [76]. The average speed tuning half-maximum width was two octaves from the preferred speed [76]. Receptive fields in MT were found to be in the range of 4°-25° in diameter, eccentric to the fovea, with an antagonistic surround in about half of the neurons [5, 38]. Receptive field center is in the contralateral hemi-field [59]. Response was not affected by color information [76]. The response properties of MT are consistent with the

hypothesis that MT mostly receives inputs from the magnocellular pathway, associated with motion and depth information, and little input from the parvocellular pathway, associated with color and form information [67].

Despite receiving its most significant projections from V1, which already exhibits significant selectivity to visual motion, area MT is distinguished in a number of ways. One of the most obvious differences is the significantly larger receptive fields of MT compared to V1 [42]. The larger receptive fields of MT coincide with sensitivity to motion of more complex patterns and objects, rather than simple local edge motion in V1. A study by Movshon et al. used plaid stimuli consisting of superimposed gratings, revealing the MT neurons responded to the the overall motion of the pattern whereas V1 responded to the component motion of the gratings [92]. Similarly, [79] claim that preferred speed in MT is constant across spatial and temporal stimulus frequencies, but in V1 preferred speed varies systematically with stimulus frequency. The strength of evidence supporting this claim in primates, however, is unclear.

Cells at the lateral edge of MT, with receptive fields that include the fovea, is often treated as a distinct region, MTf, separate from MT proper. The cells in MTf have much smaller receptive fields, typically less than 2° in diameter [59].

Careful study has shown responses in MT to be purely sensory in origin [97]. Such studies control retinal input by keeping the image of the target fixed on the retina to eliminate slip during pursuit or by transiently blanking out the target during pursuit to remove sensory input. Retinal slip of the background is eliminated by performing the experiment in total darkness and by turning on a light between trials to prevent dark adaptation, as a small amount of light is scattered from the target. MT does not respond when retinal slip is eliminated, and response is usually diminished when the target is removed [97] (although unpublished data from the Osborne lab shows some MT neurons responding transiently to target blanking, this is consistent with purely retinal inputs).

Response latencies for typical MT neurons (the time from onset of a moving stimulus to a

change in firing rate) have been reported at about 90 ms in anesthetized macaques [65, 113]. Latency is strongly correlated with target velocity, with faster speeds generally reducing latency [65]. Such latencies would not provide much time to transmit visual motion signals to the pursuit system for motor response, which has a latency of about 100 ms [62]. Other studies in anesthetized macaque have found average latencies as short as 40 ms, but may be attributable to differences in stimuli [10]. Recent data from awake macaques, collected by Macellaio and Liu in the Osborne lab, show a response latency of only 51 ms on average, which is more consistent with MT providing motion signals for pursuit initiation. This data also showed a systematic shift in latency as function of stimulus direction, with the shortest latencies for preferred direction, and increasing latencies for other directions.

Responses exhibit separate transient and sustained components, which may play different roles in motion processing [65, 146]. The sustained response is a roughly tonic firing in response to image motion on the retina. The transient response is a brief increase in firing rate in response to a change in stimulus velocity, usually of greater amplitude than the sustained rate [146]. A transient decrease in firing may occur if the change in velocity is moving away from the preferred direction.

1.1.4 Connections of MT

MT has inputs from a number of cortical regions in visual cortex, as well as some subcortical inputs. MT also projects to both cortical and subcortical targets. Area MT clearly plays a role in processing visual motion and in the initiation and maintenance of pursuit eye movements, but understanding the anatomical connections of gives a better understanding of the role MT plays in hierarchy of visual motion processing.

Cortical Inputs to MT

Several cortical visual areas have been found to contribute significant inputs to area MT. Retrograde tracing studies using injections of horseradish peroxidase in MT have found the

largest contribution of cortical projections appears to come from V1 directly or indirectly through V2 and V3 [78, 133, 132, 143]. The direct inputs to MT from V1 originate mostly in layer 4B, with a smaller portion coming from layer 6, or the border of layers 5 and 6 [78, 133, 143]. These layer 4B neurons are reported to be mostly large spiny stellate cells [133, 143], all though some studies have found a prevalence of large pyramidal cells [37, 153], with the reputed contributions of each cell type differing across studies. The discrepancy in the ratio of spiny stellate and pyramidal cells may be due to differences across primate species, the topographic region of MT studied, and methodological differences in the study of cell morphology. The layer 5/6 projection neurons are large and often referred to as solitary cells of Meynert [133, 143]. The direct and indirect V1 inputs to MT appear to be dominated by the magnocellular pathway from the retina and LGN, evidenced by anatomical as well as function inactivation studies. Maunsell et al. found that inactivating the magnocellular layers of LGN greatly reduced the responsiveness of MT, while inactivating the parvocellular layers had a much smaller effect [75]. Anatomical studies support this idea. The parallel magnocellular and parvocellular visual pathways project from distinct layers of the LGN and terminate in layers 4Cα and 4Cβ of V1, respectively [39]. Spiny stellate neurons, reputedly the strongest inputs to MT from layer 4B of V1, receive strong magnocellular input from 4Cα and no significant parvocellular input from 4cβ [158]. The pyramidal cells of 4B receive a mix of inputs from 4Cα and 4Cβ, but still receive a greater number of magnocellular inputs [158]. The indirect inputs to MT through V2 project from the thick cytochrome oxidase stripes [132], which in turn receive inputs from a population of cells in layer 4B of V1 distinct from the direct project neurons [134]. The indirect inputs projecting to V2 originate in layer 4B, comprising both stellate and pyramidal neurons, and suggest a predominantly magnocellular input with a smaller parvocellular component [68, 120, 134, 158]. Similarly, the indirect inputs through V3 originate in layer 4b of V1 and consist of small pyramidal as well as some stellate cells [20, 153], suggesting a predominantly magnocellular input to MT via V3 as well. Anatomical and functional studies suggest that the direct V1-MT inputs

are more significant than the indirect inputs. While retrograde tracing studies have found a larger number of MT input cells in V2 compared V1 [78], the projections from V1 have larger axons compared to V2 (3m to 1m) and the V1 projection axons have specialized beaded terminals with a greater number of multisynaptic boutons [6, 7, 118, 119]. Born and Bradley argue that these morphological features indicate a stronger input from V1, with fast and reliable connections [14]. Functional studies eliminating direct or indirect inputs to MT are inconclusive but insightful. Surgical removal of V1 weakens the activity in MT, but does not have a significant impact on the direction selectivity, tuning width, or binocularity of MT neurons [121]. Inactivation of V2 and V3 by cooling also reduces overall MT responsiveness, but disproportionately degrades binocular disparity tuning relative to direction tuning, suggesting different modalities for for the direct and indirect pathways [106]. Removal of superior colliculus in conjunction with V1 removal abolishes visual responsiveness in MT, providing evidence for significant subcortical pathways [122]. Other cortical inputs to MT include the ventral posterior area (VP), the posterior intraparietal area (PIP), but are thought to be less important based on the density of connections [78].

Subcortical Inputs to MT

Functional studies suggest that an alternative pathway to MT runs through the superior colliculus (SC) [122], although the details of this pathway are not entirely clear. One possibility is that input to MT from the SC is through the inferior pulvinar (PI); however, the sub-regions of the PI that receive input from the SC (PICM and PIP) do not have much overlap with the sub-regions with the densest projections to MT (PIM) [139]. Another strong possibility is a tecto-geniculo-cortical pathway to MT that bypasses MT. While most LGN neurons project to V1, direct connections from the intercalated konicellular layers of the LGN to MT have been found [135]. These konicellular regions also receive collicular input [139]. The lack of response in MT when both V1 and SC are removed suggests that if a konicellular pathway bypassing V1 exists, it does not receive direct retinal input [122].

Projections from MT

The densest cortical projections from MT are to its neighboring region in the STS, area MST [78, 149]. Other feedforward projections go to the ventral intraparietal area (VIP) and the frontal eye fields (FEF) [78, 149]. Cortical connections to MT tend to be reciprocal, including V1, V2, and V3 [78]. Subcortical projections of MT go to the superior colliculus, the pulvinar, the dorsal striatum, and the dorsolateral pontine nucleus [150].

1.1.5 Role of Other Visual Areas in Pursuit

Other areas of the brain besides MT are strongly implicated in pursuit eye movements. Most of these areas project to or receive inputs from MT. The pathway of visual motion information through the visual system must be understood in order to bridge sensation and action with population coding in MT.

V1

Primary visual cortex, or V1, is a key part of the pursuit pathway. Many V1 neurons respond to motion and are directionally selective and tune for binocular disparity [92, 111]. These response properties appear to be inherited by downstream area MT. Unilateral lesions of V1 generated deficits in pursuit behavior similar to those seen in MT lesions [96, 128]. Moving targets stepped into the contralateral hemi-field were unable to generate sufficient pursuit velocity to match the target, and saccades were unable to compensate for the motion of targets [128]. In this case, lesions ablated the entirety of V1, leading to temporary hemianopia, but permanent pursuit deficits. This is compared to MT and MST, where pursuit deficits were recovered quickly from small, focal lesions, but a lesion of the entirety of MT created permanent deficits [33, 96, 159]. While V1 lesions create serious pursuit deficits, it can be argued that motion information from V1 for smooth pursuit normally passes through MT, and that MT is a more specialized input for the pursuit system. MT

lesions are downstream of V1, and receive most of its inputs directly or indirectly from V1 [78], and MT lesions produce the same types of deficits without also losing visual perception in contralateral field for several months. Additionally, bilateral removal of V1 had drastically different effects than the unilateral ablation [160]. After several months of blindness, both saccadic and smooth pursuit eye movements returned. Smooth pursuit gain and saccade accuracy were normal, but the latencies of both eye movements increased and smooth pursuit tracking was more variable. Recover suggests that alternative pathways for pursuit exist, but the presence of V1 inhibits or overwhelms that pathway.

MST

The medial superior temporal area, MST, in the superior temporal sulcus of visual cortex, is adjacent to MT and receives the majority of the projections from MT [78]. Neurons in MST have receptive fields larger than those in MT [59]. Neurons in MST were often direction selective and responded to larger moving stimuli than MT neurons. Neurons in MST were also found to have extraretinal inputs related to pursuit movements [97]. This makes MST a candidate for integrating feedback from eye movements in the pursuit system. Microstimulation of MST during pursuit generates acceleration in the ipsilateral direction, but has little effect during fixation [60]. This is in contrast with extra-foveal MT, where microstimulation biases pursuit toward the preferred direction of the region stimulated, but similar to foveal MT [47, 60]. Similarly, lesions in MST (and MTf) led to directional pursuit deficits, i.e., an inability to match target speed for ipsilaterally moving targets, regardless of position in the receptive field [34, 33]. Retinotopic deficits similar to those observed in MT were also observed, with deficits in pursuit initiation and compensation for motion in saccades, for moving targets stepped in the contralateral hemi-field. Recovery from lesions was rapid, comparable to MT lesions.

Frontal Eye Field

A region of the frontal lobe known as the frontal eye field (FEF), in the fundus of the anterior bank of the arcuate sulcus, is the region most closely associated with eye movements in the frontal lobe. The FEF is divided into two subregions for smooth pursuit and saccades, and microstimulation of each subregion will evoke the respective eye movement [72, 141, 142]. Evoked smooth pursuit movements will tend toward the ipsilateral side [72]. The inputs to these subregions are thought to be segregated into mostly non-overlapping saccade and pursuit regions from the supplementary eye field (SEF), the lateral intraparietal area (LIP), the principal sulcus region (PSR), and MST [142]. Lesions of FEF cause deficits in smooth pursuit, but much less profound deficits to saccades [70, 72]. Smooth pursuit gain decreased to about 0.2, although [70] reported deficits in all directions, whereas [72] reported deficits in the ipsilateral motion direction. Lesions in FEF also eliminate the predictive component of smooth pursuit, which can occur for predictable or periodic moving stimuli [13, 72].

Pontine Nuclei

The pontine nuclei receive inputs from MT and MST, as well as LIP and FEF, SEF, and a few others [44, 45]. Neurons in the dorsolateral pontine nucleus (DLPN) respond variously to visual motion, smooth pursuit eye movements, or both [95]. Lesions of DLPN produce retinotopic and direction deficits in pursuit similar to those observed in MT and MST [80]. The dorsomedial pontine nucleus (DMPN) has neurons that respond to large-field motion, often with direction selectivity, and some neurons possess pursuit related activity [57]. The function observed in the pontine nuclei is consistent with transmitting motion and pursuit signals from cortex to the cerebellum.

Cerebellum

The flocculus of the cerebellum is closely connected with the motor circuits of the brainstem, as evidenced by eye movements evoked with flocculus stimulation on short time scales of 10ms [63]. Floccular stimulation elicits ipsilateral eye movements, which is consistent with the ipsilateral deficits and preferences observed in MTf, MST, and FEF lesions and stimulations [123]. Additionally, [64] found that neurons in the ipsilateral brainstem receive inhibition from the flocculus, and those neurons increased in firing rate for contralateral movement, consistent with disinhibition. Removal of the cerebellum results in total loss of pursuit eye movements, and floccular lesions result in partial loss of pursuit [155].

1.2 Cortical Population Coding

Sensory neurons in cortex are tuned to specific stimuli, but the accuracy of single neurons is insufficient to achieve levels of accuracy required for behavior. The responses of individual neurons are noisy, responding variably to repeated presentations of the same stimulus. In MT, single neurons can discriminate visual motion stimuli separated by about 30° apart, but the downstream perception and and smooth pursuit behavior indicate much finer discriminability [100]. The implied increase in accuracy is due to the pooling of response across a population of neurons. Averaging the response over many neurons can overcome the significant trial-to-trial variability of cortical neurons.

Different methods have been develop to to estimate the coding performance of a population. With a small number of neurons, it is possible to simply calculate the mutual information between the stimulus and the response of the neurons [103]. With a large population of neurons however, it is impossible to estimate the mutual information due to under-sampling of the distribution. Instead, Fisher information is commonly used to describe population coding performance [9, 36, 91, 131]. The Fisher information has the property of bounding the performance of an unbiased estimator (Cramér-Rao bound) and is related to the mutual

information between stimulus and a large population of tuned neurons [18, 27, 116].

The Cramér-Rao bound provides a nice comparison to behavioral data, because the performance of the behavioral output can be considered a result of an upstream unbiased estimator, and thus must be bounded upstream by Fisher Information of the population through which the information passes.

Vector averaging is another population coding method that is used to estimate a stimulus parameter from the response of the population [50, 51]. Each neuron in the population represents a vector in the stimulus space to which it is tuned, weighted by the firing rate of the response. The stimulus estimate is simply the normalized weighted sum of the vectors of the population.

Another method for estimating population coding performance is to use a linear classifier, with methods such as a support vector machine (SVM). This provides a comparison to discrimination tasks in behavior, such as two alternative forced choice tasks.

In addition to simple independent trial-to-trial variability, responses across neurons may be correlated. The effect of correlated noise on population coding depends on tuning curves of the population and structure of the correlations [131, 162, 36, 91]. In particular, small noise correlations have a large impact the Fisher information in populations of neurons with homogeneous tuning curves, but that effect is mitigated by having a diversity of tuning functions in the population [131, 162, 36]. It has been recently shown that the decrease in Fisher information from noise correlations is due to differential, or information-limiting, correlations [91]. Differential correlations are correlations in the direction perpendicular to the optimal decision boundary between adjacent stimuli, such that the effect is not distinguishable by a linear estimator.

The visual system can also operate in different states, such as shifts in attention, which alter the gain of the neural responses as well as the the variance and correlation in population [129, 148]. This adds another confounding factor for decoding a population, as a population may decode the same stimulus differently depending on context.

CHAPTER 2

STATE DEPENDENCE OF STIMULUS-INDUCED VARIABILITY TUNING IN MACAQUE MT

Behavioral states such as levels of arousal and attention modulate some properties of cortical responses (average firing rates, pairwise correlation) while leaving others largely intact (tuning bandwidth, preferred stimulus). Here we show that changes in state modulate the tuning of response variance-to-mean ratios in a fashion that is neither predicted by a Poisson spiking model nor changes in the mean firing rate, with a substantial effect on stimulus discriminability. We recorded motion-sensitive neurons in middle temporal cortex (MT) in two states: alert fixation and light, opioid anesthesia. Anesthesia tended to lower average spike counts, without decreasing trial-to-trial variability compared to the alert state. Under anesthesia, within-trial fluctuations in excitability were correlated over longer time scales compared to the alert state, creating supra-Poisson variance-to-mean ratios (Fano factors). In contrast, alert-state MT neurons have higher mean firing rates and largely sub-Poisson variability that is stimulus-dependent and cannot be explained by firing rate differences alone. The absence of such stimulus-induced variability tuning in the anesthetized state suggests different sources of variability between states. A simple model explains state-dependent shifts in the distribution of observed Fano factors via a suppression in the variance of gain fluctuations in the alert state and predicts the observed changes in stimulus discriminability by the cortical population.

2.1 Introduction

Sensory systems operate in many states (e.g. attentional states and stages of sleep) wherein the same anatomical network displays different scales of firing rates, variability, correlations, oscillation frequencies, and so forth while maintaining basic function [147, 81, 28, 35, 156, 29]. In some states, such as under light anesthesia, cortical networks encode roughly the same

information about a stimulus, but with different dynamics and firing rates [16, 25]. Each neuron's contribution to the accuracy of stimulus decoding depends on its tuning function, the smooth modulation of the firing rate in response to parametric changes in stimulus value, and a neuron's response reliability determines its impact on decoding precision [18, 130, 90]. Much is made of how a population code might be either robust or sensitive to neuronal variability, but an important aspect of variability is often left out of the discussion: how does neuronal variability, itself, depend on the stimulus? This question becomes particularly important in the context of how response reliability impacts information transmission over the brain's natural operating range[18, 100, 21, 53] and decoding [18, 71].

The impacts of variability on population decoding can be analytically derived for the case of an independent or correlated population of Poisson neurons [18, 161], where the mean and variance of the response are the same, yielding a Fano factor = 1. The Poisson model of spike generation replicates many of the features of cortical spike trains recorded under some conditions [130], but the Fano factor prediction is often violated in real sensory neurons [137, 54, 8, 22, 89], particularly when measuring responses over the $\sim$ 100 ms timescale of sensory estimation [100]. Decreases in Fano factor are observed at the onset of visual stimulation, alongside decreases in neuronal correlation, in MT and throughout cortex[100, 136, 28, 87, 26, 156]. Moreover, Fano factors can have their own stimulus tuning, which can impact stimulus encoding at the population level [109, 124, 107].

The response properties of neurons in the middle temporal cortical area (MT) have been particularly well described across a number of behavioral states. MT neurons respond selectively to visual motion and firing rates decrease with a Gaussian profile with angular distance from a preferred motion direction [76, 3]. Levels of arousal (anesthesia, alert behavior) and modulations of spatial attention affect the stimulus-averaged excitability of MT neurons, as elsewhere in the brain, but do not tend to shift preferred directions or tuning bandwidths [147, 126, 30, 81], much like changing contrast modulates the rate without changing tuning in primary visual cortex [144, 32, 2, 127]. Less well studied are the effects of behavioral

state on the variability of cortical responses, a critical measurement for assessing sensory discrimination. Increased attention tends to reduce variability in MT and V4 responses, particularly in narrow-spiking neurons [86, 28, 74]. Anesthetic effects may be analogous to a large reduction in attention, decreasing mean firing rates and increasing response variance [17, 156].

Here, we explore the state dependence of variability in neural responses and its implications for sensory discriminability. We record single unit responses to motion steps in MT in alert monkeys and under a light opioid anesthetic as a proxy for the range of natural brain states. Our goal is not to provide an exhaustive review of the myriad effects of anesthesias and other modulations of brain state on cortical variability, but rather to quantify several features of cortical responses under two particular brain states and ask whether a parsimonious model can explain these transitions. We find that scaled variability in spike count (Fano factor) is tuned for motion direction, but only in some network states. Alert responses display sub-Poisson variability that is inversely tuned to the stimulus, decreasing at the preferred direction of the cell. Anesthetized responses show flat, supra-Poisson tuning. We identify a simple model by which a single parameter accounts for changes in visually-driven spiking activity in both alert and anesthetized animals. Modulation of the size of gain fluctuations in the response can give rise to both qualitative regimes of Fano factor tuning. Finally, we show how changes in the tuning of the Fano factor enhance stimulus discriminability when animals are alert and actively engaged in visual behavior.

2.2 Results

In order to test the impact of network state on spiking statistics and stimulus encoding in cortical sensory neurons, we made extra-cellular recordings of isolated units from cortical area MT in both alert and anesthetized monkeys (Fig 2.1A). Some of the data from sufentanil-anesthetized experiments has been published elsewhere [100, 101, 103]. MT neurons respond robustly to motion steps of random dot patterns, with firing rates that often peak at over 100

spikes/s to preferred directions (Fig 2.1B) and Gaussian-shaped direction tuning functions (Fig 2.1C). We presented motion steps in 13 (anesthetized condition) or 24 (alert condition) directions and repeated each stimulus $\sim$ 100 times to estimate the distribution of spike counts (see Methods).

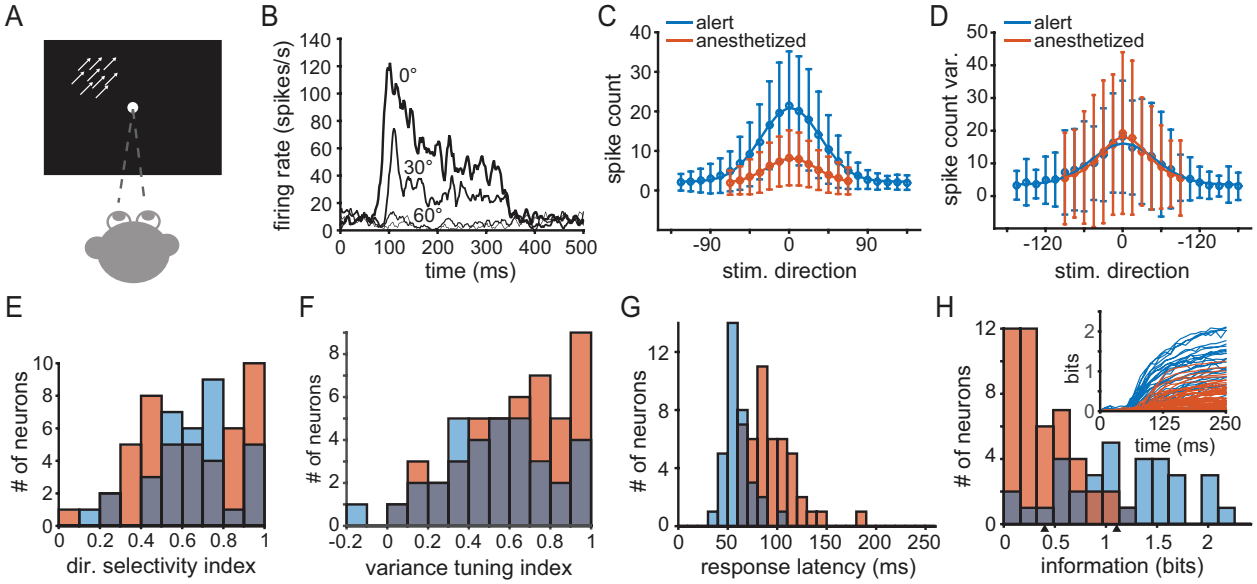

Figure 2.1: Response properties of MT neurons to constant motion stimuli. Blue indicates experiments on alert subjects and orange indicates experiments on anesthetized subjects throughout. (A) A cartoon of the experimental setup for measuring the response of MT neurons to steps of coherent random dot pattern motion. Monkeys maintained fixation while stimuli translated at constant speed in one of 13 or 24 directions behind a stationary aperture scaled to the excitatory receptive field. See Materials and Methods for detail. (B) Sample peri-stimulus time histogram (PSTH) from a representative unit showing the time course of firing rate for different motion directions, averaged over a 10 ms sliding window. Motion onset is at 0 ms and motion offset is at 250 ms. (C) Average spike count across neurons in the alert experiments (blue) and anesthetized (orange) by direction. Circles indicate mean, error bars indicate standard deviation, and the solid traces indicates a Gaussian best fit. Responses are aligned such that the preferred stimulus direction for each neuron is taken to be 0°. (D) Average spike count variance across populations of neurons in the same manner as (C). (E) Histogram of direction selectivity indexes, DI, across both populations. (F) Histogram of variance spike count tuning indexes. (G) Histogram of response latency distributions. (H) Distribution of single-unit mutual information (Shannon) values for alert (blue bars) and anesthetized (orange bars) states, based on spike count 250 ms after stimulus motion onset, Value are corrected for finite sample size (see Materials and Methods). Arrows below indicate average information for alert (1.11 bits) and anesthetized (0.40 bits). Inset: time course of the mutual information between the cumulative spike count and motion direction over time with respect to stimulus motion onset. Blue traces indicate alert-state units, orange traces anesthetized state.

2.2.1 Anesthetic state modulates mean and Fano factor in MT neurons

Our recordings from MT neurons in alert, behaving primates show a pronounced increase in their stimulus-evoked mean firing rates compared to neurons recorded in MT under light, opioid anesthesia (Fig 2.1C), which is consistent with comparisons between alert and anesthetized states in other systems (e.g. [25, 35, 156, 29, 4, 83, 58, 1, 48]). We quantified the mean and variance of the response by counting spikes within a 250 ms window starting from stimulus motion onset on each trial. The scale of fixational eye movements is small compared to MT neuron receptive field sizes, such that they do not alter the observed variability in motion-evoked responses in MT [11]. We focus on this 250 ms time interval during stimulus motion, but the main results we present are consistent across a range of behaviorally relevant time scales for motion estimation (50-500 ms). The variance is tuned to the direction of motion in recordings from both alert and anesthetized subjects, and displays a largely overlapping distribution of magnitudes in these two states (Fig 2.1D). The maintenance of roughly equal magnitude spike count variance with increasing mean firing rate in the alert condition implies a substantial change in the Fano factor (FF) in these two states.

In the alert state, we find that the variance, like the mean response, is tuned to the direction of motion. On average, variance peaks at the preferred direction (rotated to 0° in all figures) in both the alert (blue trace, Fig 2.1D) and anesthetized (orange trace, Fig 2.1D) populations and falls off with increasing angular separation from the preferred motion direction. But while firing rates are quite state-dependent, alert and anesthetized populations display a largely overlapping distribution of variance values, indicating a large change in Fano factor (FF). Consistent with past studies, the change in state does not affect the direction tuning of mean rate. We computed a direction selectivity index, DI, (see Equation 2.9, Materials and Methods) and found that the population distribution of values were statistically indistinguishable (two-tailed t-test, $p = 0.60$, Fig 2.1E). An analogous variance tuning index (see Equation 2.10) shows a similar tuning of the variance in both populations (two-tailed t-test,$p = 0.09$, Fig 2.1F). Response latencies, estimated as the first point when

the average response rose above baseline after motion onset and inspected manually for each neuron, decrease from 94 ± 24 ms (SD, n=46) under anesthesia (orange bars, Fig 2.1G) to 56 ± 13 ms (SD, n=34) in alert responses (blue bars). The latency difference is consistent with previously reported measurements [55, 125, 112, 113, 1]. The increase in latency under anesthesia contributes to the reduction in estimated firing rate (Fig 2.1C), but firing rates are lower under anesthesia even when estimated in time windows aligned to response onset.

Overall the impact of anesthesia is to lower signal (the mean rate) and maintain noise (count variance), suggesting that sensory information transmission is impaired with respect to alert behavior, though no studies have directly measured this effect. We confirm this intuition by information theoretic calculations. The mutual information between spike count and motion direction differed substantially between states (see Materials and Methods, Fig 2.1H). On average, MT units recorded in alert subjects encoded $1.11 \pm Y$ 0.57 (SD, n=34, blue bars, Fig 2.1H) about direction compared to 0.40 ± 0.32 bits (SD, n=46, orange bars, Fig 2.1G) in anesthetized subjects, a statistically significant difference ($p = 1.3 * 10^{-10}$, one-tailed t-test). The combined effect of the increase in response latency and reduction in information under anesthesia is that less information about motion direction is available over time. In Figure 2.1H inset we plot the time course of the mutual information between the cumulative spike count measured from motion onset and motion direction for each unit in both populations (see Materials and Methods). In both populations, stimulus information accumulates most rapidly with the first few spikes fired, but shorter latencies and higher overall firing rates in the alert state mean that more bits are available more quickly (blue versus orange traces, Fig 2.1H). Normalizing the mutual information by the response entropy reduced the difference between the two states but did not change the results. Changing the size of the time window over which counts are integrated does not recover the lost information under anesthesia and the difference in average coding capacities persists during stimulation. However, anesthesia does not abolish the capacity of MT to encode information about motion direction entirely. The network still functions even with substantially reduced firing rates.

The state-dependence of the Fano factor is illustrated in Figure 2.2. Very few neurons in either state display a Fano factor of 1 (dashed unity line in Fig 2.2A). The Fano factors (FF) of most units measured in the alert state fall below 1 (blue symbols), while most units in the anesthetized state display FFs above 1 (orange symbols). The distribution of measured FFs over the anesthetized data shows a significant rightward shift toward higher values compared to data from anesthetized units, from a mean of 1.02 ± 0.79 (SD, n=34) to a mean of 1.82 ± 0.84 (SD, n=46) (Fig 2.2B). Fano factors also show different dependencies on the count in the two states. Although values show a large degree of scatter, FF increases with spike count in anesthetized data (orange and black triangles, dashed line, Fig 2.2C). In the alert state, FFs tend to decrease slightly with increasing spike count (blue and black circles, solid line, Fig 2.2C) such that neurons with higher firing rates tend to be more precise. Quadratic polynomial fits of rate to Fano factor (black traces, Fig 2.2C) show these relationships, but low r^2 values (0.08 in alert, 0.02 in anesthetized) suggest that firing rate alone is not a great predictor of Fano factor. Overall, there is a marked shift in the scaling of variability with responsiveness in these two states. Alert state neurons have higher firing rates and higher rate precision relative to anesthetized responses, consistent with better direction discrimination during alert behavior.

2.2.2 In alert subjects, Fano factor is tuned to motion direction

All but one of the neurons showed directional tuning of the mean and 72 of 80 neurons showed directional tuning in the trial-to-trial variance (Fig 2.1C-F). However, the tuning of the variance does not simply follow the tuning of the mean, either in a one-to-one fashion, with a Fano factor of one, or with a constant factor not equal to one. In a linear regression between mean rate and variance, firing rate fails to explain much of the spike count variance measured in the alert experiments ($R^2 = 0.37$, see Materials and Methods for statistical tests used). A linear regression of spike count variance on firing rate explained more variance ($R^2 = 0.66$) in the anesthetized state.

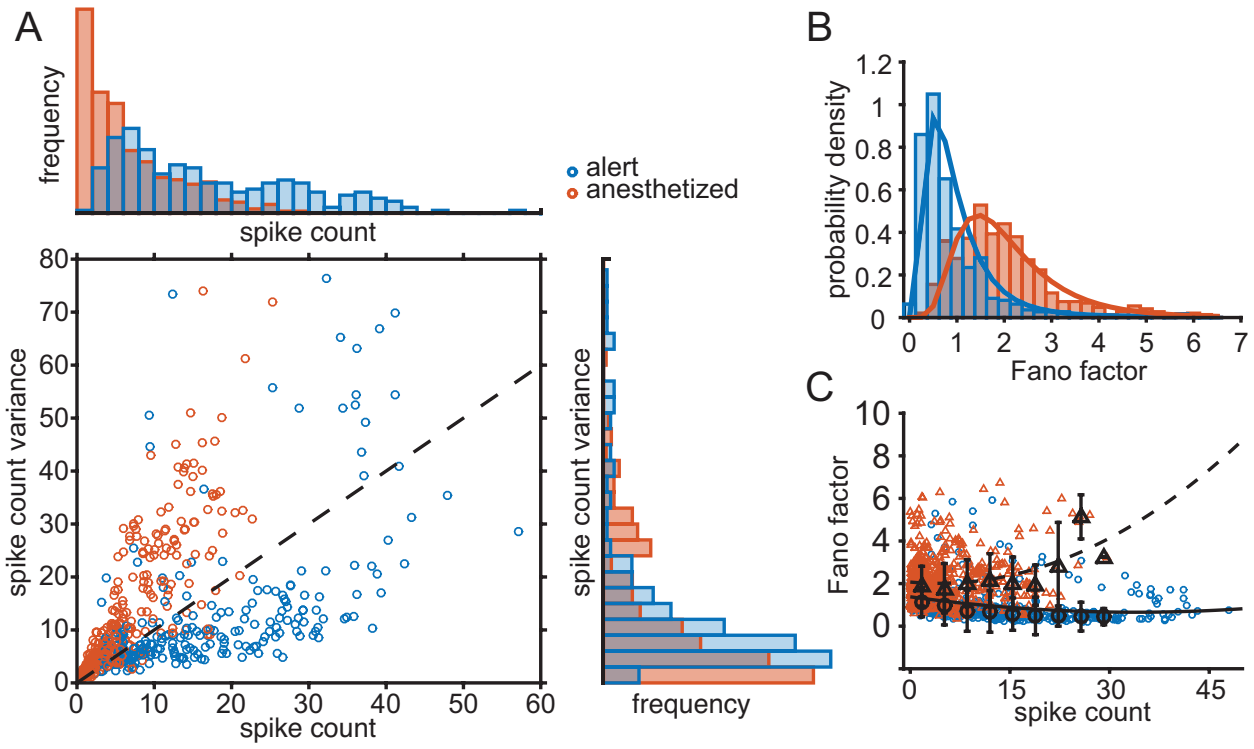

Figure 2.2: Fano factors differ significantly between MT responses recorded in alert versus anesthetized primates. (A) The relationship between spike count and spike count variance in alert (blue circles) and anesthetized (orange circles) experiments. Each circle indicates a single neuron under a single stimulus condition. Histogram above shows distribution of spike counts observed in both experiments. The histogram to the right shows the distribution spike count variance values in both experiments. Stimulus directions shown are within 45° of the preferred motion direction, in 15° increments. (B) Distribution of spike count Fano factors for alert (blue) and anesthetized (orange) experiments. Solid trace indicates the best-fit lognormal distribution for each case. (C) Distribution of Fano factors as a function of spike count for alert experiments (blue) and anesthetized experiments (orange). Black circles and triangles indicate the median Fano factor binned by spike count in alert and anesthetized experiments, respectively. Error bars show the standard deviation of Fano factors in each bin. The solid and dashed traces are the best-fit quadratic curve for alert and anesthetized states, respectively ($r^2 = 0.08$ alert, $r^2 = 0.02$ anesthetized).

Figures 2.3A and B show the qualitative difference in Fano factor tuning between three example units in the alert and anesthetized states, respectively. Although there are a range of FF tuning profiles measured within each state, the increased sharpness of FF tuning in the alert state is apparent. For each neuron recorded, we defined $\mathrm{FF}_{\mathrm{pref}}$, the Fano factor for the preferred stimulus direction of the neuron, and $\mathrm{FF}_{\mathrm{orth}}$, the Fano factor for the orthogonal stimulus directions (rightmost panel, Fig 2.3B). We used these terms to define a Fano factor

tuning index (FFTI) that captures the degree to which the FF depends on motion direction.

$$\mathrm{FFTI} = \frac{\mathrm{FF}_{\mathrm{orth}} - \mathrm{FF}_{\mathrm{pref}}}{\mathrm{FF}_{\mathrm{orth}} + \mathrm{FF}_{\mathrm{pref}}} \tag{2.1}$$

Equation 2.1 is analogous to standard methods for quantifying direction selectivity in first-order response statistics like the rate [110], and takes values from -1 to 1. Positive FFTI values indicate a decrease in Fano factor for the preferred stimulus direction relative to the off-preferred stimuli ("U-shaped" tuning); negative FFTI values indicate an increase in Fano factor for the preferred stimulus direction (Gaussian-like tuning). The three alert-state example units in Figure 2.3A display "U shaped" tuning functions (FFTI values > 0) whereas the anesthetized examples in Figure 2.3B either lack direction tuning (FFTI near 0, left and center panels) or show a more Gaussian-like profile (rightmost panel). We plot FF tuning functions for all units in Figure 2.3C (alert) and D (anesthetized) as gray lines. Although there is a noticeable diversity in FF tuning profiles for individual units, particularly in the anesthetized state, the population medians (green lines, Figs 2.3C-D) reflect a clear state-dependent shift in tuning. Fano factors in the alert state are clearly tuned, whereas in the anesthetized state, the median FF is flat across directions. These results mean FF as well as median FF and for a variety of integration windows, including aligning to response latency rather than stimulus onset (S1).

The distribution of FFTI values in the alert data is broader and more positive than in the anesthetized state (Fig 2.3E). Alert state MT units showed, on average, significantly greater FF tuning ($\langle \mathrm{FFTI} \rangle = 0.172, \mathrm{where} \langle \cdot \rangle$ indicates a mean over all units) compared to the anesthetized state ($\langle \mathrm{FFTI} \rangle = -0.012$) (two-tailed t-test, $p = 0.0016$). We chose FFTI ≥ 0.2 as a cutoff for a positively tuned Fano factor. In the alert state, 50% of the units had positively tuned Fano factors (17/34). In the anesthetized state, only 13% of the units had positively tuned Fano factors (6/45). One unit from the anesthetized group was excluded from the FFTI calculations as it did not emit any spikes to the orthogonal stimulus direction, and thus $\mathrm{FF}_{\mathrm{orth}}$ was not defined.

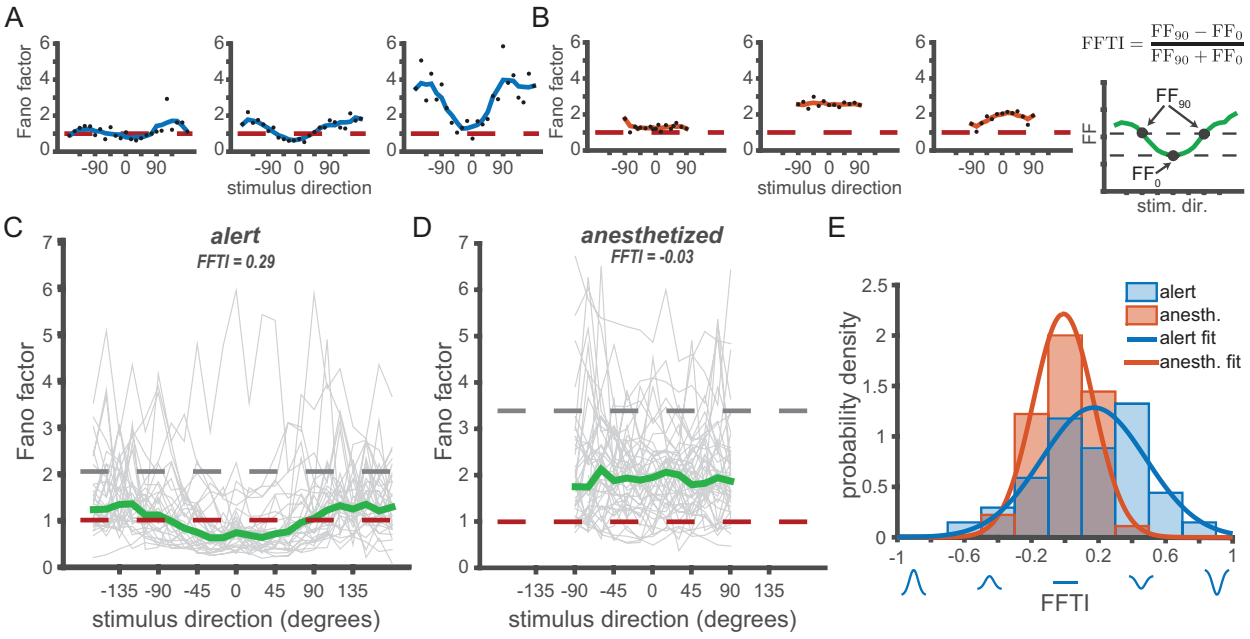

Figure 2.3: Fano factors show significant tuning to motion direction in the alert state. Fano factor across stimulus directions in example neurons from (A) alert and (B) anesthetized recordings in MT. Black dots indicate Fano factor by direction. Blue and orange traces show Fano factor values smoothed by a 30° moving window. for respective experimental conditions. Orange dashed lines indicate FF = 1. (C) Black dots show the Fano factors by stimulus direction for all neurons neurons in the alert experiments. Stimulus directions are aligned such that the preferred stimulus direction of each neuron is at 0°. The median Fano factor across the population (green) is significantly tuned by direction. Gray dashed line indicates median Fano factor in response to a stationary null stimulus. Red dashed line is at FF = 1 is as predicted for a Poisson process. (D) Same as (C) but for the anesthetized experiments. The median Fano factor is not tuned by stimulus direction. (E) Histogram distributions of Fano factor tuning indices (FFTI) from the alert (blue) and anesthetized (orange) experiments. Orange and blue traces are Gaussian best fit. The cartoon at top right shows how FFTI is calculated using the median trace from (C).

Overall, we observe a dependence of the Fano factor of the spike count on stimulus direction in the alert state. The tuning is U-shaped, with a dip at the preferred direction, and the occasional presence of side-peaks at near-orthogonal directions (see e.g. Fig 2.3A-B). The Fano factor tuning we observe in the alert state using high Fourier bandwidth dot pattern motion is similar in shape and amplitude to that observed by [107] using low Fourier bandwidth drifting sine-wave grating stimuli. Our data from anesthetized subjects showed no significant stimulus dependence of the Fano factor (repeated measures ANOVA, $p = 0.38$).

Firing rate changes do not explain differences in Fano factor tuning

The firing rates in our alert MT recordings are roughly twice as large as those in the anesthetized recordings (Fig 2.2A). Ideally, we would like to be able to exclude all effects changes of mean rate might have on observed Fano factors. One way to minimize the impact of rate differences is by analyzing subsets of the data samples that have matching firing rates, as in [26]. The mean-matching method excludes data from the groups being compared in order to match their spike count distributions. Mean-matching does not maintain cell identity, and thus can only describe population-wide rather than cell-specific tuning. We generated histograms of the spike counts for both states at the preferred stimulus direction and at the orthogonal directions. We then randomly excluded points from each dataset until the histograms of spike counts in the alert and anesthetized states matched at each stimulus direction (gray shaded bars, Fig 2.4A). We calculated the mean Fano Factor of this reduced data set for both stimulus conditions and both states. We calculated the population FFTI from the mean Fano factor at the preferred and orthogonal directions. This process was repeated one million times to average out effects of the particular random subset sampled on each draw.

The mean-matched analysis reveals that both the population-average shape of the FF tuning as well as the shift to higher FF in the anesthetized state are maintained (Fig 2.4B). In this figure, the mean-matched population Fano factors were fit to cosine tunings (Fig

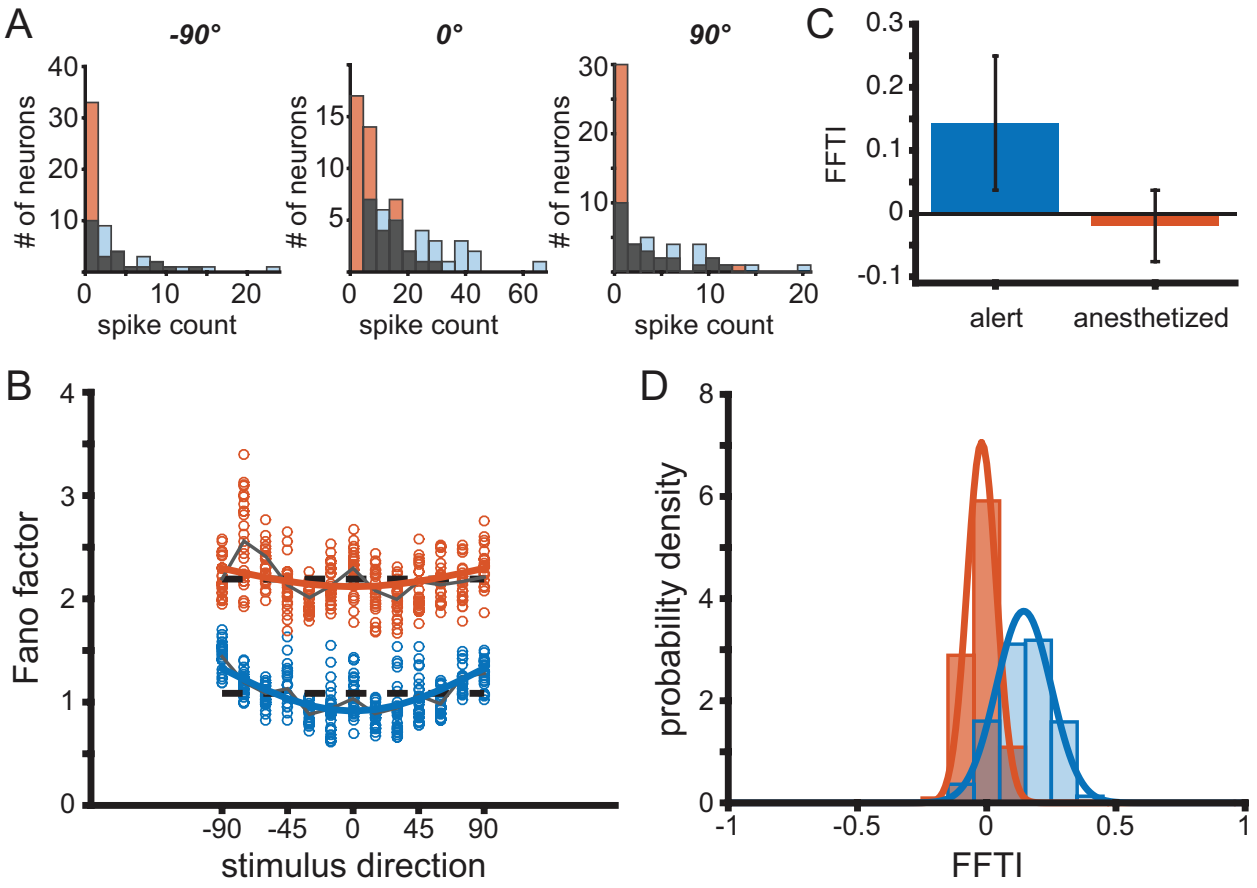

Figure 2.4: Mean-matched Fano factors preserve variability tuning. (A) Histograms of spike counts by neuron for a given stimulus direction (left: $-90°$; center: preferred direction; right: $+90°$) for alert (blue, throughout) and anesthetized (orange, throughout) recordings. The overlapping area is shown in gray. (B) Sample mean-matched Fano factors, with spike count distributions corresponding to the gray histograms in (A). Each circle is a sample mean. Dashed traces indicate the mean Fano factor across directions. Solid traces indicate the mean Fano factor across bootstrapped samples. (C) Fano factor tuning index (FFTI) of bootstrapped resamples for alert (FFTI $= 0.143$, $s.d. = 0.106$) and anesthetized states (FFTI $= -0.020$, $s.d. = 0.057$). (D) Distributions of FFTI for mean-matched data samples. Solid traces indicate a best-fit Gaussian curve.

2.4B). The best cosine fits corresponded to $\text{FFTI}_{\text{alert}} = 0.156$ and $\text{FFTI}_{\text{anesth}} = 0.039$ for the mean-matched populations, with significant tuning of the FF in the alert data. The state-dependent differences in Fano factor tuning for individual units within the sub-sampled population were also maintained after mean matching, with $\text{FFTI}_{\text{alert}} = 0.143$ (sd = 0.107) and $\text{FFTI}_{\text{anesth}} = -0.020$ (sd = 0.057) (Fig 2.4C). The distribution of FF tuning indices across sub-sampled data were also similar to the full data set, and recapitulated the shift to more positive FFTI in the alert state (Fig 2.4D).

An alternative possibility is that the state-dependent difference in FF stimulus tuning arises from differences in the scale of the FF between the two brain states. Fano factors in the anesthestized recordings were roughly 80% higher than in the alert experiments, which affects the normalization factor in the tuning index calculation. With this in mind, we can examine the raw change in Fano factor, where $\Delta\text{FF} = \text{FF}_{\text{orth}} - \text{FF}_{\text{pref}}$. In alert recordings we see $\Delta\text{FF} = 0.326$ (sd = 0.24) and for anesthetized, $\Delta\text{FF} = -0.09$ (sd = 0.26). Thus, the higher spike counts recorded in the alert state are not the source of the difference in Fano factor tuning.

It is notable that the qualitative relationship between Fano factor and firing rate differs in the alert and anesthetized recordings. Fano factor increases with spike count in the anesthetized state, and decreases with spike count in the alert state, over the same range of firing rates (Fig 2.2C). This suggests that changes in mean spike count alone cannot entirely explain the observed differences in measured Fano factors and their tuning.

Effect of tuning bandwidth on Fano factor tuning

While differences in mean spike count fail to explain the differences in observed Fano factor tuning between states, it is possible that differences in tuning bandwidths lead to the observed differences in Fano factor tuning. If mean tuning curves are narrower or steeper in the alert state, this could result in different Fano factor tuning curves when normalized by the same variance tuning curves. Conversely, if the mean tuning curves are the same in

both states, Fano factor tuning could arise from relatively broader or flatter variance tuning curves. We used an ANOVA to compare the widths of mean spike count tuning curves and spike count variance tuning curves in both states to the FFTI. There was no effect of tuning curve width or variance tuning width on FFTI, and no interaction with state, suggesting that FF tuning arises from systematic differences in the structure of the variance rather than differences in the first-order tuning properties.

Temporal correlations change with behavioral state

The temporal frequency of fluctuations in spiking differed substantially between the alert and anesthetized states. Under sufentanil anesthesia, neural excitability fluctuated more slowly such that deviations from the mean tended to accumulate during a trial. The temporal autocorrelation in spike count fluctuations displayed a higher peak with an exponential decay with a characteristic time constant of approximately 100 ms (Fig 2.5, orange curve). In contrast, alert-state correlations were weaker overall and no significant correlations in spiking were observed beyond 50 ms (blue curve, Fig 2.5). This difference in timescales contributes to the state-dependent difference in Fano factors. Longer, stronger temporal correlations in excitability create a larger count variance over a 250 ms time window, and therefore higher Fano factors [100]. The shorter timescales of correlation in the alert state create less variable counts, lowering the variance and the FF. To the observer, fluctuations in spike count are indistinguishable from fluctuations in response gain, so an alternative description of MT activity is that, on timescales longer than 50 ms, gain fluctuations are smaller in the alert state than in the anesthetized state. We explore a gain-based model of state dependent changes in network activity below.

Autocorrelation is not affected by stimulus presentation refresh rate

An analog oscilloscope was used for stimulus presentation in the anesthetized experiments while a 100 Hz CRT was used for alert experiments. It is conceivable that the disparity

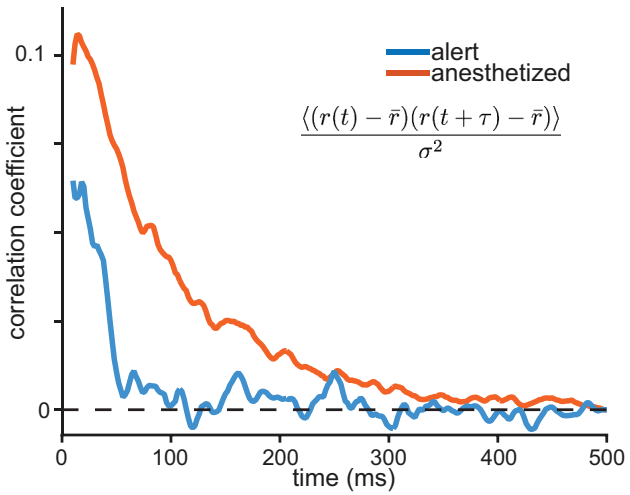

Figure 2.5: Behavioral state affects the spike-count autocorrelation function. Traces indicate the average temporal autocorrelation for spike count fluctuations in awake (blue) and anesthetized (orange) conditions. To calculate autocorrelation, spikes are binned in 2 ms windows and smoothed over a running average of 5 bins. Autocorrelation is normalized by variance and averaged over the population. The time course of autocorrelation is longer under anesthesia, decaying with an exponential time constant of 88 ms compared to 29 ms in the awake state.

in stimulus presentation methods could impact spike count variability. Specifically, there is concern that nature of the frame update of the CRT leads to phase-locking of the spikes and regular spiking. We examined the autocorrelation and power spectra of neurons from both experiments, and no periodic firing was found (Supplemental Fig 2.10A,B). Phase-locking to the stimulus would reveal itself as peaks in the power spectrum corresponding to the 100 Hz refresh rate of the CRT.

The power spectra of many neurons showed broad peaks in the power spectrum at approximately 300-600 Hz, but no peaks corresponding to stimulus refresh rates. This bump in the power spectrum in every instance could be explained by the distribution of interspike intervals (ISI) for the neurons (Supplemental Fig 2.10B,C). Neurons that exhibited a bump in the power spectrum had an abundance of very short ISI corresponding to pairs of spikes, or doublets, which occur in rapid succession about 1.5-3.5 ms apart. The timing of these doublets is particular to the individual neuron and consistently match peak in the power spectrum, where the frequency is the reciprocal of the mode of the ISI distribution. The

high frequency peak in the power spectrum and matching doublet ISI were observed in the anesthetized experiments. Therefore we deem it unlikely that the disparity in Fano factor distribution is significantly determined by refresh rates and stimulus phase-locking.

2.2.3 A simple model accounts for changes in Fano factor tuning

We created a simple model that could account for the observed state-dependent changes in FF tuning without relying on state-dependent differences in firing rates. The model is inspired by observations of decorrelation and enhanced variability in cortical responses in the alert compared to the anesthetized state. We reason that increased decorrelation with alertness could lead to lower relative variance in the neural response, and, hence, lower FF's. Models that include Fano factor tuning have been developed to explain the stimulus-dependent effects of neuronal response variability [107, 40, 163], while others have modeled variability as rate-dependent, or due to stimulus-independent gain fluctuations [46]. Our model combines contributions from rate-dependent variability as well as rate-independent gain-fluctuations to reproduce the diversity of Fano factor tunings observed in both the alert and anesthetized data.

A simple and common model of spike counts in cortical neurons is given by a Poisson process, or a Poisson mixture model [46]. A Poisson process is characterized by a Fano factor of 1. Including a multiplicative gain to describe the underlying rate of the Poisson process results in a super-Poisson spike count distribution, with a Fano factor that increases with firing rate. Under anesthesia, we found Fano factors larger than 1 that increased with mean spike count (Fig 2.2). These observations are consistent with a Poisson mixture model, and agree with previous experiments [130, 46]. Our recordings in the alert state, however, show a sub-Poisson spike count, with a Fano factor that decreases with increased firing rate. The inverse rate-dependence of Fano factor was previously documented by [35] in V1. Can a single model account for responses in more than one behavioral state? The model we use captures the difference in the rate-dependence of Fano factors between states as well as the

difference in Fano factor tuning.

We modeled stimulus averaged MT responses with Gaussian tuning functions $f(\theta)$, where θ is the stimulus direction. On any given trial, the mean rate is scaled by a multiplicative gain, g, yielding the underlying rate for the neuron on that trial

$$\mu = f(\theta) * g. \tag{2.2}$$

The gain g is taken to be a gamma-distributed variable with mean of 1. The spike count for a given trial is sampled from a Gaussian distribution with a mean value μ, and a variance μ^{α}, where α is an intrinsic property of the cell that determines how variance scales with firing rate. For a fixed value of g, a value of α less than 1 corresponds to a Fano factor less than 1, and a value of α greater than 1 corresponds to a Fano factor greater than 1.

If we allow g to fluctuate, the variance of the spike count x is given by

$$\mathrm{var}(x|\theta) = f(\theta)^{\alpha}\langle g^{\alpha}\rangle + f(\theta)^2 * \mathrm{var}(g). \tag{2.3}$$

Because $\langle g\rangle = 1$ and the variance of g is relatively small, we can approximate $\langle g^{\alpha}\rangle \approx 1$. Thus, we can approximate the spike count variance

$$\mathrm{var}(x|\theta) \approx f(\theta)^{\alpha} + f(\theta)^2 * \mathrm{var}(g). \tag{2.4}$$

For small values of var(g), *i.e.*, small gain fluctuations, the variance is dominated by the first term. This results in a U-shaped Fano factor tuning for $\alpha < 1$. As the gain fluctuations increase, the second term dominates the spike count variance, resulting in flat or Gaussian-shaped Fano factor tuning for the same value of α. This model suggests that a difference in the amplitude of gain fluctuations is sufficient to reverse or eliminate the observed Fano factor tuning.

The gain fluctuation parameter also explains the qualitative dependence of Fano factor

on increasing spike count observed in the anesthetized and alert states. For small gain fluctuations ($\mathrm{var}(g) \ll 1$) and $\alpha < 1$, the Fano factor for a given firing rate, r, is $r^{\alpha-1}$, which decreases as r increases. For stronger gain fluctuations, the Fano factor scales linearly with r, increasing with increasing rate.

To fit this model to our data, we first use as input the set of best-fit Gaussian tuning curves measured in each behavioral state. Next, we fit α and $\mathrm{var}(g)$ to reproduce the observed FFTI distribution in the each behavioral state (Fig 2.6A,B). The distribution of Fano factor tuning indices generated by the model (Fig 2.6A) were similar to those observed in the data (Fig 2.3D), indicating that the model can reproduce the observed data. The fitted values of the intrinsic variability parameter were $\alpha = 0.31$ in the alert condition and $\alpha = 0.74$ in the anesthetized condition. While these values are different, they both result in similarly U-shaped Fano factor tuning curves in the absence of large gain fluctuations. The model predicts a U-shaped Fano factor tuning for small gain fluctuations in both states, and inverted tuning for larger gain fluctuations. The gain variance parameters were $\mathrm{var}(g) = 0.0094$ in the alert condition and $\mathrm{var}(g) = 0.0732$ in the anesthetized condition, nearly an order of magnitude difference. It is this difference in gain variance that causes the qualitative difference in the Fano factor tuning between the two conditions.

To confirm that these results were not affected by differences in α arising from the fits to heterogeneous units in each behavioral state, we fit the same model using an optimized value of $\alpha = 0.66$ for every neuron in both behavioral states. The results were comparable to those shown, and the difference in gain variance between states was yet larger. This indicates that the observed FFTI distributions might be primarily modulated by changes in a single parameter, the variance in the gain distribution.

We also fit the model for the population-averaged mean-response tuning in each state, to test whether the differences in FFTI distributions could be wholly accounted for by changes in $\mathrm{var}(g)$, and not by changes in the distributions of mean tuning. The average alert and anesthetized tuning curves are show in Fig 2.6E, inset. The model parameters were then

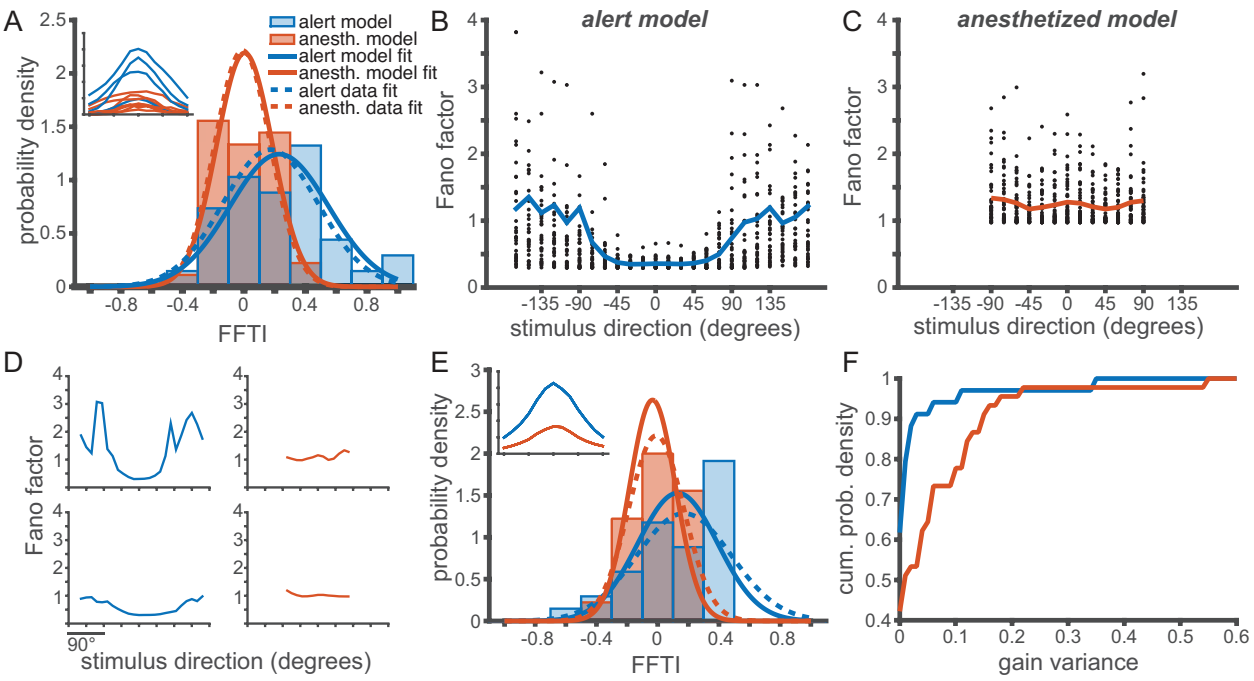

Figure 2.6: A single parameter, gain variance, can account for the observed changes in Fano factor tuning with behavioral state. (A-D) Model fitting captures differences in Fano factor tuning through changes in gain variance. The model is fit with the optimal α and var(g) for each population. (A) The distribution of FFTI for alert (blue, throughout) and anesthetized (orange, throughout) in the observed population (dashed trace) and the model values (solid trace). Inset: Sample tuning curves for the alert and anesthetized experiments. (B,C) The variance model predicts Fano factor direction tuning for the alert (B) and anesthetized (C) experiments. Compare model results to observed Fano factors in Fig 2.3C-D. Parameters are fit to match distribution of FFTI, but reasonably reproduce Fano factor tuning as well. (D) Sample Fano factor tunings generated by the best-fit models shown in (A-C). (E) Distribution of FFTI for alternate model fitting in which the spike count used for each neuron was replaced by the average tuning curve over all neurons recorded in each experimental condition. The gain variance was fit separately to match the observed FFTI distribution. The α parameter is identical for all neurons and was chosen to minimize mean-squared error in the fit. Inset: Mean tuning curves for the alert and anesthetized experiments. (F) The cumulative distribution of gain variance parameters from the model fit in (E), showing larger gain variance values in the anesthetized model.

fit, such that all neurons had the same value of α. The gain variance parameter var(g) was fit separately for each neuron to reproduce the observed distribution of FFTIs. This model again predicts a distribution of gain fluctuations that is much larger in the anesthetized condition than in the alert condition, and is centered on a significantly larger mean gain variance value (Fig 2.6F).

In both models, whether the tuning curves are heterogeneous or identical, the proportion of neurons with high values of gain variance is much greater in the anesthetized condition.

2.2.4 Impact on information transmission

It seems reasonable to assume that higher Fano factor in the anesthetized state will result in lower rates of information transmission and stimulus discriminability from these neural populations. Conversely, the reduction in relative variability at the preferred direction for most cells recorded from alert, behaving subjects seems to imply enhanced information transmission and decoding. To test this intuition, we simulated responses of MT neuron populations and varied the stimulus-dependence of the noise. We quantified population encoding performance via the Fisher information (FI) metric, which determines the bound on the performance of an unbiased estimator reading out the population code.

We first tested a homogeneous population with identically shaped tuning curves (Fig 2.7A). The first-order response properties of the simulated populations were matched to the population-averaged statistics of MT neuron responses measured in alert macaques in the first 150 ms following the onset of stimulus motion. We varied this averaging window systematically to determine how it affects the resulting decoding performance of the model. The mean tuning curve was fit to a von Mises function as in [35, 163] and rotated such that the preferred directions evenly tiled all directions.

In order to separate the effects of the magnitude of the spike-count variability from the stimulus-dependent part of the variability, we imposed various Fano factors tunings on the populations while keeping the average Fano factor across all directions constant (FF=1).

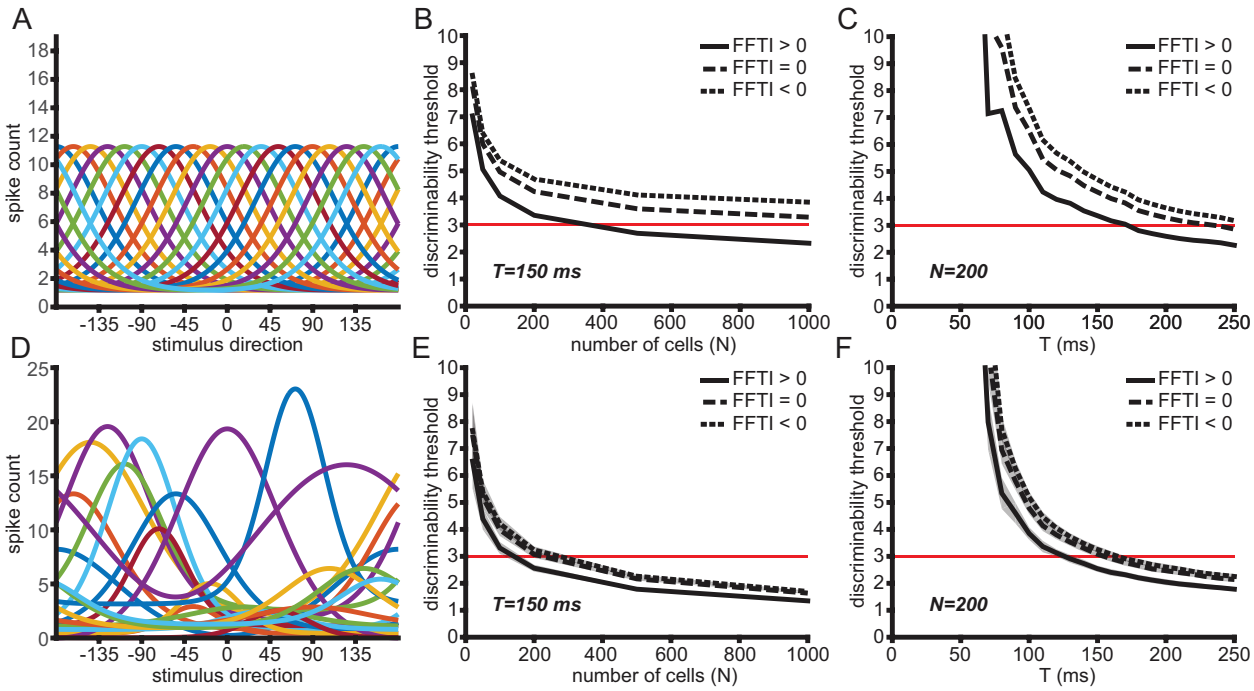

Figure 2.7: Fano factor tuning and heterogeneity both contribute to lower discriminability thresholds in MT populations. (A) Model population with 20 homogeneous tuning curves. (B) Cramer-Rao bound in homogeneous population models of different sizes with short-range correlations ($c_{max} = 0.1$, $\langle c \rangle = 0.0438$). Tuning curves are fit to the average cumulative response up to 150 ms after motion onset. Black traces show performance of models with varying stimulus-dependent variance. The solid trace is FFTI > 0, the dashed trace is FFTI $= 0$, and the dotted trace is FFTI < 0. The red line indicates the stimulus discriminability threshold for smooth pursuit behavior in macaques 125 ms after pursuit initiation. (C) Same as in (B) but for populations of 200 neurons with first-order statistics matched to average response at time t after motion onset. (D) Sample population of 20 heterogeneous tuning curves drawn from measured tuning curves in recorded neurons. (E,F) Same as in (B,C) but for heterogeneous populations. The shaded areas show the standard deviation of the Cramer-Rao bound.

The Fano factor tunings were modeled as von Mises functions and imposed directly, rather than using the spike count model described above, so that the average Fano factor could be held constant. The positive (U-shaped) and negative (inverted-U) tuning of the stimulus-dependence of the Fano factor are symmetric by reflections over the line FF=1.

Up to this point we have treated neurons within the cortical population as independent. To test whether correlations between neurons would affect model predictions, we tested correlated population responses. We synthesized a population of correlated MT neurons using parameters consistent with the experimental literature [12, 51]. In our simulations, pair-wise correlation levels, c, peaked at 0.1 and fell off with the angular distance between preferred directions, d, according to a von Mises function

$$c(d) = c_{\max} \frac{e^{\kappa(\cos(d)+1)} - 1}{e^{2\kappa} - 1} \tag{2.5}$$

with width parameter, $\kappa = 1$, corresponding to a half-width at half-max = 64°. For the results shown in Figure 2.7, the maximum pairwise correlation $c_{\max}$ was chosen to be 0.1, with an average pairwise correlation of $\langle c \rangle = 0.0438$.

In the homogeneous population, the population with the positive FFTI performed better than a population with either a flat Fano factor tuning or a negative FFTI (Fig 2.7B). Performance was quantified by the Cramer-Rao bound, given by $\frac{1}{\sqrt{J}}$, where J is the Fisher information. Because MT coding performance is a constraint on the precision of downstream smooth pursuit behavior, we can compare the decoding performance of the population to the discrimination thresholds measured in behavioral experiments [47, 15, 100, 102, 101, 138, 50, 93]. To align with behavioral measurements of motion discrimination in primates, we estimated the size of the model population needed to reach a Cramer-Rao bound of 3° [101]. The model population with positive FFTI crossed this threshold with 361 neurons (Fig 6B). The population in which the variability was structured to give a constant Fano factor (FFTI = 0) required 3764 neurons to reach same level of discrimination, and the FFTI < 0

population appeared to asymptote before reaching Cramer-Rao bound of 3°.

We also looked at the time course of decoding performance by finding the average tuning curves after cumulative intervals of time after motion onset (Fig 2.7C). We assumed a population size 200 neurons and $c_{max} = 0.1$. We estimated the time after motion onset when the Cramer-Rao bound for these populations would cross the 3° threshold. Pursuit has a directional precision of $2 - 3°$ within 250ms of motion onset [102, 101]. We asked whether our model populations could reach that direction precision within the behaviorally relevant timescale. The population with FFTI > 0 crossed the threshold after 171 ms, while the FFTI $= 0$ population crossed the threshold in 235 ms. This shows that a population with stimulus-dependent variability tuning can reach the same level of stimulus discrimination in less time. The FFTI < 0 population did not reach 3° threshold within a 250 ms analysis window.

Real cortical populations are heterogeneous, and that heterogeneity is expected to improve overall coding [131, 103]. We introduced heterogeneity into our simulations by sampling with replacement from the measured tuning curves in each behavioral state and randomly reassigning preferred directions uniformly (Fig 2.7D-F). The tuning curves were again fit to von Mises functions that matched the cumulative spike counts assessed at a range of intervals following stimulus motion onset and the preferred directions were spaced evenly to tile the space of stimulus directions. We found that adding this heterogeneity did improve coding performance, as assessed by the Cramer-Rao bound. Fewer neurons were required to reach behavioral performance, and this bound could be reached more quickly after motion onset in the heterogeneous populations (Fig 2.7E-F). Short-range correlations and stimulus-dependent variability were imposed in the same manner as with the homogeneous population.

We found that the effect of stimulus-dependent variability was smaller in the heterogeneous population model as compared to the homogeneous one, but the same qualitative trend was observed. The heterogeneous population sampled at 150ms after motion onset and with FFTI > 0 required only 141 neurons to reach the Cramer-Rao bound of 3°, while

the population with constant Fano Factor required 224 neurons (Fig 2.7E). The population with FFTI < 0 required 273 neurons to reach the same level of stimulus discriminability (Fig 2.7E).

Estimating the time course of stimulus discriminability in populations of 200 heterogeneous neurons, the FFTI > 0 population reached the 3° discriminability threshold in 125 ms while the FFTI $= 0$ population reached the same threshold in 154 ms and the FFTI < 0 population reached the threshold in 164 ms (Fig 2.7F). Once again, U-shaped tuning of the Fano factor, as observed experimentally, allows the threshold on behaviorally relevant direction discrimination levels to be reached with fewer neurons in a shorter amount of time after motion onset.

Accounting for information-limiting correlations

The Fisher information of neural populations may also be limited by small information-limiting correlations that cannot easily be measured [91]. Downstream behavioral performance can be used to set bounds potential strength of such correlations and their potential effects can be taken into account. We can add information-limiting correlations to our covariance, giving us

$$\mathbf{\Sigma}_\epsilon(\theta) = \mathbf{\Sigma_0}(\theta) + \epsilon \mathbf{f'}(\theta)^T \mathbf{f'}(\theta). \tag{2.6}$$

Here, $\mathbf{\Sigma_0}$ is the initial covariance matrix, and $\mathbf{\Sigma}_\epsilon$ is the covariance with information limiting correlations. If the Fisher information associated with the covariance $\mathbf{\Sigma_0}$ is J_0, the Fisher information associated with $\mathbf{\Sigma}_\epsilon$ is

$$J_\epsilon = \frac{J_0}{1 + \epsilon J_0}. \tag{2.7}$$

We quantify performance by the square root of the Cramer-Rao bound, given by

$$\frac{1}{\sqrt{J_\epsilon}} = \sqrt{\frac{1}{J_0} + \epsilon}, \tag{2.8}$$

which represents the lower bound on the standard deviation of an unbiased estimator. As J_0 increases with population size, the decoder performance is bounded by $\sqrt{\epsilon}$. Note that while information-limiting correlations may affect the absolute coding performance of a population, they do not affect the relative performance of models that differ only in their covariance structures $\mathbf{\Sigma_0}$, as long as the bound is not saturated ($J_0 \not\gg \epsilon$). While different models will saturate to the same bound, the question of interest is how quickly each model approaches that asymptotic bound in terms of population size and temporal integration.

For an estimator to achieve a 3° level of precision comparable to behavioral smooth pursuit discrimination, then information-limited correlations are bounded by $\epsilon \leq 9$. We therefore introduced the strongest reasonable information-limiting correlations into our population, for $\epsilon = 4$ and $\epsilon = 9$, corresponding to asymptotic limits of 2° and 3°. These additional correlations decrease the decoding performance in each model, but do not change the qualitative results, as the FFTI > 0 population still reaches the asymptote with fewer neurons and less integration time than populations FFTI ≤ 0 (S1).

2.3 Discussion

The brain functions across a wide range of network states that encompass levels of arousal and attention, yet little is known about how behavioral state affects sensory discrimination. Much of our historical understanding of the nature of visual coding arises, of necessity, from experiments under anesthesia, mimicking a stage of sleep [52, 76, 16]. Those observations continue to inform modern neuroscience because the response characteristics of cortical neurons, such as tuning curves, remain consistent across network states [147, 81, 30, 4]. But there are clear state-dependent differences in sensitivity, background firing rates and pairwise correlation structure with alertness and attention [147, 28, 35, 156]. Attention, for example,

decorrelates local cortical populations [28] and increases firing rates for preferred stimuli [147, 74], enhancing signal to noise ratio in the cortical network and improving stimulus detection and discrimination [130, 21]. The fine spatial [87] and temporal [43, 151, 28, 69] scales over which attention can operate suggests a localized control mechanism. Here we show that a simple model operating at the level of individual neurons can reproduce many features of the state-dependent variance-to-mean changes we observe in sensory cortex.

We have used a combination of physiological data analysis from cortical area MT alongside modeling to characterize the statistics of cortical responses under two candidate behavioral states: attentional alertness required for maintaining fixation during a visual experiment, and a quiescent state induced by the opioid anesthestic agent, sufentanil. A characteristic effect of opioid anesthesia is to increase cortical wave activity, low frequency spatiotemporally structured activity modulations (e.g. [16, 145]). While the scale of wave activity is increased, it is not observed to be unnaturally structured. Similarities in the correlations between functionally connected brain areas in the quiet awake and lightly anesthestized states suggests that general activity patterns under anesthesia can mimic network states during active behavior [88, 49, 152].

At the single unit level, we find that sufentanil increases the timescale over which fluctuations in spiking are correlated, increasing spike count variance on time scales relevant for visual motion estimates. Overall, these correlations lower the precision of MT responses and degrade stimulus discriminability compared to the alert state. Models of MT activity that can describe the shift in variance and mean firing rates between the alert and anesthetized states have an additional constraint. We find, as did [156, 107], that response variance is stimulus-dependent and has its own tuning function that is similar, but not identical, to that of the mean count. A measure of response precision, the Fano factor, acquires stimulus tuning in the alert state and becomes stimulus-independent under anesthesia. Significant tuning of the Fano factor during alert behavior may enhance stimulus readout. Lower variance at and around the preferred direction of each neuron in a population leads, unsurprisingly, to

a finer discrimination threshold, as estimated via the Fisher information. Fewer cells are needed to achieve the same level of direction discriminability, and stimulus information is available earlier in populations that have this kind of Fano factor tuning.

The fact that the response variance is not tied to the mean firing rate violates the usual assumption that cortical spiking has Poisson statistics. A Poisson process yokes mean and variance together to maintain a Fano factor of 1, or values slightly less than 1 when refractoriness is revealed at high firing rates [86]. Neural deviations from idealized Poisson behavior are well documented [12, 19, 140, 54, 73] and these non-Poisson effects are known to be important for accurate modeling of neural response [56, 105]. Deviations from Poisson behavior are particularly acute in our data and represent a strong constraint on a feasible model of cortical activity that can generalize to different behavioral states.

We show that a model that incorporates a state-dependent shift in gain variance alone can reproduce the changes in the variance level and variance tuning observed in MT. Other recordings and models of attentional effects on neuronal firing, particuarly in area MT, focus on tuning-dependent shifts in the gain of mean responses [115, 99, 98]. Here, we focus on the effects of variance in the gain on the stimulus tuning of the Fano factor. Using the cortical gain model proposed by [46], we are able to show that a change in the gain variance model can not only explain overall shifts in the Fano factor, but also reproduces tuning of the Fano factor in the alert state, when gain variance is low. This suggests that a higher gain variance state underlies the observed flat tuning of the FF under anesthesia. This aligns with results that suggest that anesthesia corresponds to a more synchronous mode of brain coupling [16, 24, 25]. These results point to a simple physiological mechanism that achieves the shift in response statistics with alertness. A recent study in rat V1 shows effects of anesthetic state on Fano factors and FF tuning to the stimulus period [156], adding general support to our observation that noise is suppressed in the alert state.

The active suppression of gain fluctuations during wakefulness may be directed specifically at those neurons encoding stimulus variables in an active task (here, fixation). This

suggests a simple knob that attentional modulation can turn to drive more reliable decoding of the stimulus. Reducing gain fluctuations may require processes, such as activation of inhibitory networks, that are metabolically costly and, thus, are only engaged when needed to maximize sensory discrimination during active behavior.

2.4 Materials and Methods

2.4.1 Ethics statement

All animal care, behavioral, and surgical procedures were performed in compliance with National Institute of Health guidelines and were approved in advance by the *Institutional Animal Care and Use Committee* of the University of California, San Francisco and The University of Chicago. All procedures complied with guidelines for animal welfare in accordance with the recommendations of the Weatherall report, "The use of non-human primates in research."

2.4.2 Experimental methods

We made extracellular single-unit microelectrode recordings of MT neurons in both alert, behaving and anesthetized monkeys. All procedures were performed in compliance with *Institutional Animal Care and Use Committee* guidelines. The "anesthetized motion-step experiments" consisted of recordings from four adult male macaques (*Macaca fasicularis*). Animals were implanted with a head-restraint and a craniotomy was performed under isoflurane anesthesia using sterile technique. During the anesthetized experiments, anesthesia was maintained with continuous infusion of the opioid sufentanil, and paralysis was induced with vecuronium bromide to minimize eye movements, and midazolam was administered periodically. Pain responses were monitored at 15 minute intervals and additional analgesics were used if necessary. Unit recordings were made using tungsten-in-glass microelectrodes. Visual stimuli comprised randomly drawn patterns of white dots that moved coherently within

a stationary aperture against the dark screen of analog oscilloscopes (models 1304A and 1321B, P4 Phosphor; Hewlett-Packard, Palo Alto, CA). The size and position of the stimulus aperture was chosen to maximally excite each isolated unit, as was motion speed. The direction of stimulus motion was pseudo-randomly chosen from a set of at least 13 directions that spanned $\pm 90^\circ$ around the preferred direction including 15° increments. A stationary random dot "null" stimulus was interleaved with the motion stimuli in 36 of 46 recorded neurons. For these 36 neurons, stimuli were repeated 51-223 times. The remaining 10 neurons had 20-30 stimulus repetitions. In all experiments, dot textures appeared and remained stationary for 256 ms, translated for 256 ms with constant direction and speed, and were again stationary for 256 ms. When the motion of a dot carried it outside of the aperture, it was randomly positioned along the leading aperture edge to maintain dot number. Trials were separated by a brief pause of 1-2s. Spike waveforms were sampled at 10kHz and isolation was aided by a window discriminator. Spike times were determined by threshold crossings. Data from three of the four monkeys have been previously published [100, 103].

Similar "alert motion-step" experiments were performed with two adult male monkeys (*Macaca mulatta*) that maintained fixation during visual stimulus presentation. These methods have been described in more detail elsewhere [66, 94]. Animals were pair-housed when possible and had daily access to enrichment activities and play areas. Pre-study instrumentation with a head stabilization post, an eye coil, and a recording chamber was performed under isoflurane anesthesia using sterile surgical technique and post-operative analgesia with buprenorphine. Animals were trained over a period of time to acclimate them to the laboratory environment. Daily experiments involved seating the monkey in a plastic "chair" in front of a visual display in a dimly lit room. The task required fixation within 2° throughout the 2-3 second trial to obtain a juice reward. Eye position was monitored via a surgically implanted scleral coil. We employed similar visual stimuli to the anesthestized experiments. We presented bright random dot stimuli against the dark screen of a CRT display set to 1024x768 resolution and a 100 Hz frame rate (Sony GWFM-FW9011). Recordings were made with an

various 3 quartz-platinum/tungsten single microelectrodes (TREC, Germany). We sampled the voltage waveforms from the array at 30 kHz (Plexon Omniplex) and stored them for offline analysis. As in the anesthetized experiments, we performed online analyses to map the direction and speed tuning, and the size and location of each unit's excitatory receptive field. We confirmed unit isolation through principal component analysis of spike waveforms along with inspection of interspike interval distributions. Motion stimuli were comprised of 24 directions, evenly spread between the preferred direction of the isolated single unit and $\pm 180^\circ$. A stationary random dot "null" stimulus was interleaved with the motion stimuli.

2.4.3 Analytical methods

We computed the spike count in the 250 ms following motion onset, while the dot textures were translating. We computed the Fano factor of the spike count (variance divided by mean) as a function of stimulus direction using a repeated-measures ANOVA on the null hypothesis of a constant Fano factor across directions. Stimulus direction relative to the preferred direction of the isolated single unit was found to have an effect on Fano factor in the alert experiments ($p = 7.4 \times 10^{-7}$) but not in the anesthetized experiments ($p = 0.42$). Because there was a narrower range of stimulus directions in the anesthetized experiments, we limited both data sets to stimulus directions within $\pm 90^\circ$ of the preferred direction.

Tuning widths were obtained by fitting a Gaussian curve to the spike counts as a function of direction for each neuron. The tuning width was taken to be the standard deviation of the Gaussian of best-fit. Direction selectivity was computed as a direction index (DI), where

$$\mathrm{DI} = \frac{r_{\mathrm{pref}} - r_{\mathrm{orth}}}{r_{\mathrm{pref}} + r_{\mathrm{orth}}}. \tag{2.9}$$

Here, r_{pref} is the mean response to the preferred stimulus direction and r_{orth} is the mean

response to the orthogonal directions. Similarly, a variance tuning index was calculated as

$$\mathrm{VTI} = \frac{\sigma^2_{\mathrm{pref}} - \sigma^2_{\mathrm{orth}}}{\sigma^2_{\mathrm{pref}} + \sigma^2_{\mathrm{orth}}}. \tag{2.10}$$

FFTI was defined analogously.

$$\mathrm{FFTI} = \frac{\mathrm{FF}_{\mathrm{orth}} - \mathrm{FF}_{\mathrm{pref}}}{\mathrm{FF}_{\mathrm{orth}} + \mathrm{FF}_{\mathrm{pref}}}. \tag{2.11}$$

The procedure for mean-matched Fano factor is adapted from [26]. The distribution of mean spike counts was computed for each condition. The distribution of spike counts for each stimulus direction was approximated by binning values in 15 evenly spaced bins that spanned the range of responses. For each bin in which one condition had more data points than the other, data points were randomly discarded from that condition until both distributions matched. The Fano factor was calculated for each neuron in that stimulus condition. The mean Fano factor for each stimulus condition was used to calculate FFTI for each subject condition. This resampling procedure was repeated one million times to generate a distribution of Fano factor tunings. The resampled Fano factor values were fit to cosine curves by a least squares fitting procedure.

The Fisher information was calculated for simulated neuronal populations using

$$J(\theta) \approx J_{\mathrm{mean}}(\theta) = \vec{f}'(\theta)^T \Sigma^{-1}(\theta) \vec{f}'(\theta), \tag{2.12}$$

where the derivative is taken with respect to motion direction θ, and $\vec{f}$ is a vector containing the tuning curves for each neuron in the population, and we ignore the contributions to the Fisher information from derivatives of the covariance matrix [36, 163]. The covariance matrix between neurons is given by $\Sigma(\theta)$. The diagonal elements contain each neuron's variance, as imposed by the Fano factor tuning curves we model in the three qualitative regimes (U-shaped tuning of the variance relative to the mean, flat, and inverted-U). The off-diagonal

elements of $\Sigma(\theta)$ are imposed by our correlation model, which is a von Mises distribution, as given by Eq 2.5. For the homogeneous population, each neuron has the same tuning curve, given by a von Mises distribution

$$f(\theta) = b + A\frac{e^{\kappa(cos(\theta-\theta_{\text{pref}})+1)} - 1}{e^{2\kappa} - 1} \tag{2.13}$$

fit to the population averaged tuning in each behavioral state, where b and A are background firing rates and peak firing rates, respectively; θ_{pref} is the preferred direction of the neuron; θ is the stimulus direction; and κ is the width of the tuning curve. Preferred directions in the model population are distributed evenly across all recorded directions. For the heterogeneous population model, tuning curves are again modeled with von Mises distributions fit to the diversity of tunings measured in each behavioral state. Populations models are built up by sampling with replacement from the recorded, fit tuning curves.

Mutual information was calculated for each neuron from cumulative spike counts across 13 stimulus directions, in 15 degree increments $\pm 90^\circ$ around the preferred direction. The mutual information, I, between spike count k and stimulus direction θ is given by

$$I(k;\theta) = \sum_{\theta} p(\theta) \sum_{k} p(k|\theta) \log_2 \frac{p(k|\theta)}{p(k)}. \tag{2.14}$$

Sampling bias in information estimates were corrected via bootstrap resampling at fractions of the data between $95-50\%$; finite size effects were estimated using the method of quadratic extrapolation [104].

2.5 Supplemental Materials

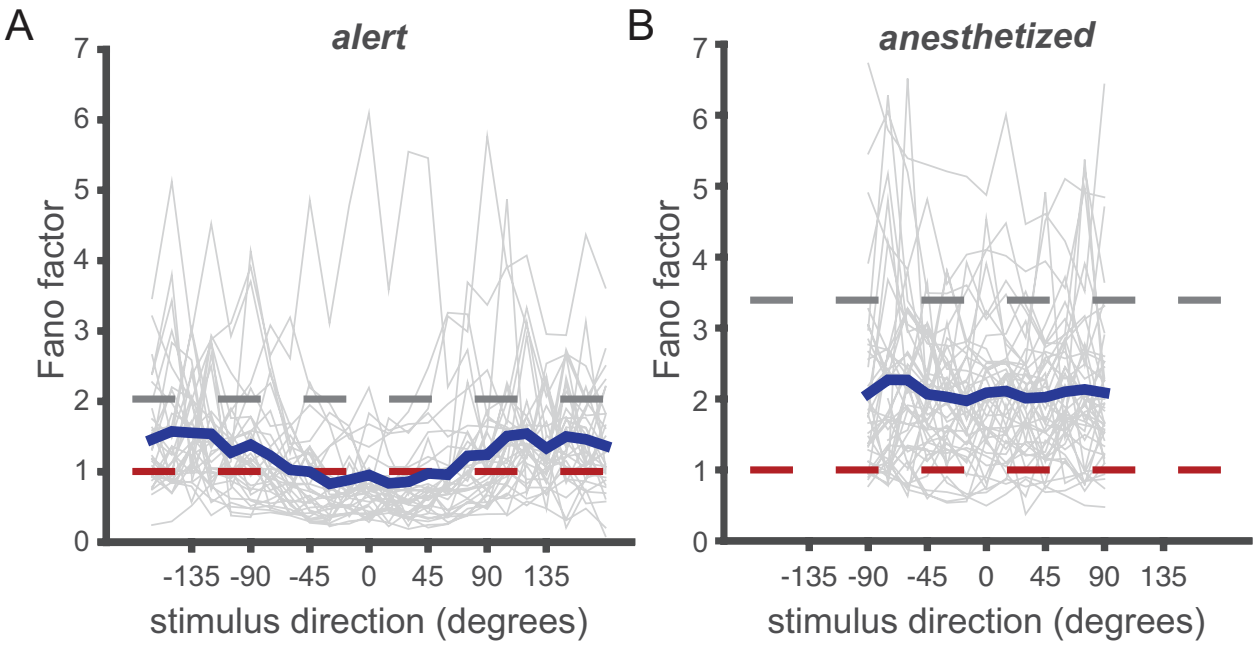

Figure 2.8: Mean Fano factor by stimulus direction exhibits the same state-dependence of stimulus-induced variability as median. Same as Fig 2.3C-D but with the mean Fano factor for each stimulus direction shown in blue rather than the median.

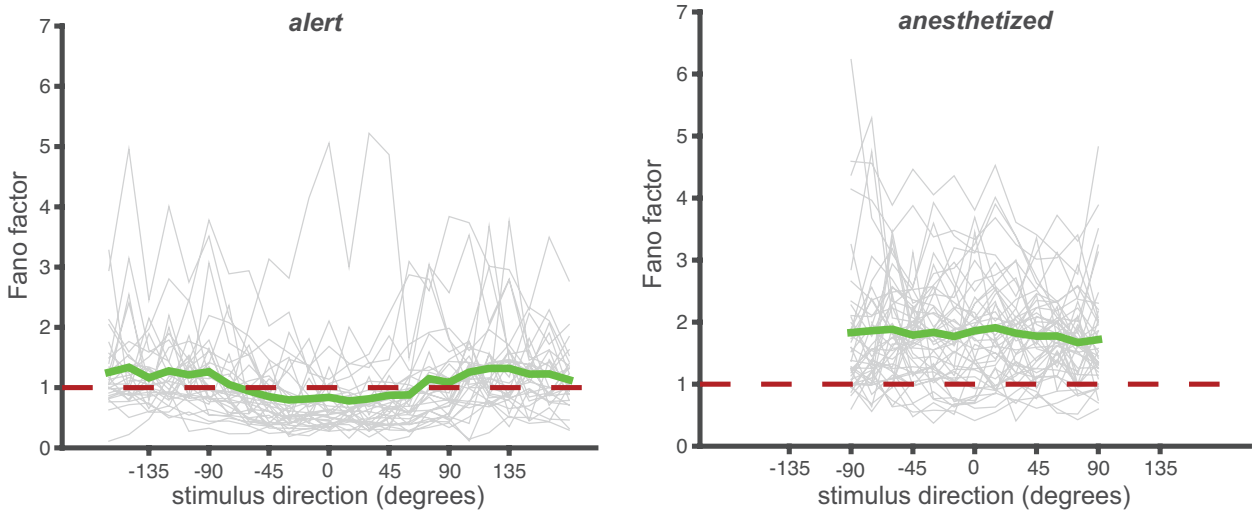

Figure 2.9: Aligning spike count windows by response onset does not affect stimulus-dependent Fano factor tuning. Figure is as in Fig 2.3C-D but spike count windows are aligned to response onset rather than stimulus motion onset. Neurons in the alert state tend to have shorter latencies than those in the anesthetized state, which may bias the spike counts and spike count variance and give rise to different Fano-factor tunings. Instead we find the effect is the same when aligned by response latency.

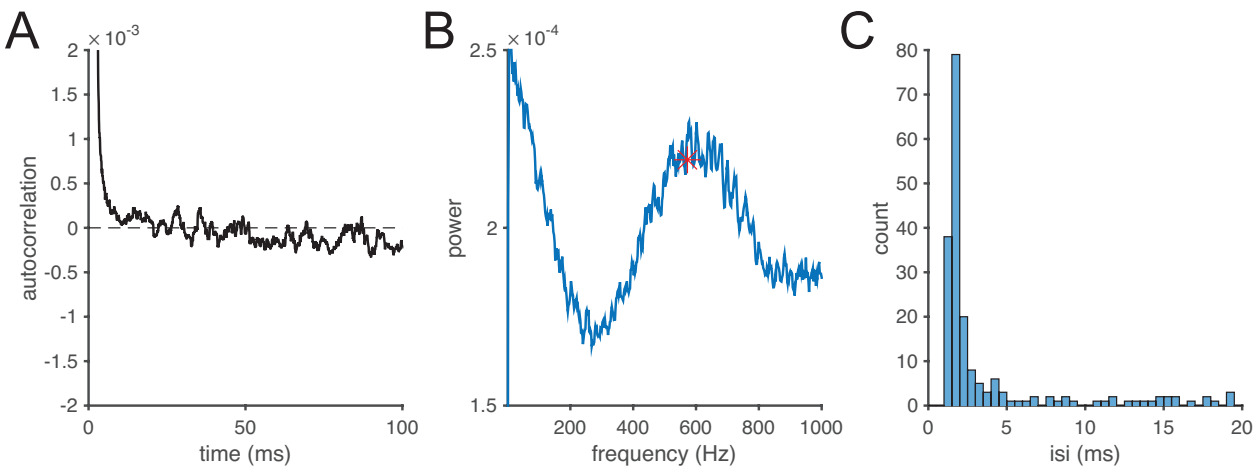

Figure 2.10: Neuronal response is not locked to refresh rate of stimulus presentation. It is conceivable that the refresh rate of the CRT (100 Hz) produces regular spiking that affects the spike count variability. Because the two experiments in different states use different stimulus presentation mechanisms, the power spectrum was analyzed to determine if the CRT had a significant impact on spike timing. (A) The autocorrelation for a representative neuron in alert state does not reveal obvious phase locking. (B) The power spectrum of the same neuron shows a peak around 600 Hz. The red star indicates the frequency corresponding to the mode of the distribution of inter-spike intervals (ISI). (C) A histogram of the distribution of short ISI for the same neuron in 0.5 ms bins peaks in the 1.5-2.0 ms range. These short ISI corresponds to doublet spikes firing in rapid succession. An ISI of 1.75 ms corresponds to 571 Hz, shown as red star in (B). For all neurons that exhibited a peak in the power spectrum, it was in the 300-600 ms range and corresponded to the ISI of doublet spikes for that neuron.

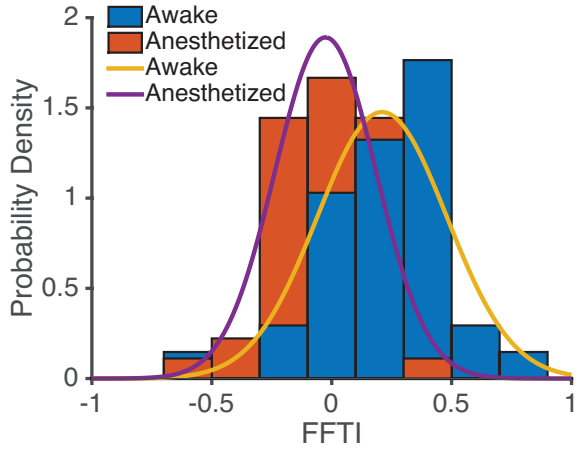

Figure 2.11: Aligning spike count windows by response onset preserves distributions of Fano factor tunings. Figure is as in Fig 2.3E but spike count windows are aligned to response onset rather than stimulus motion onset. Yellow and purple traces are Gaussian best fits to FFTI distributions in alert and anesthetized states, respectively.

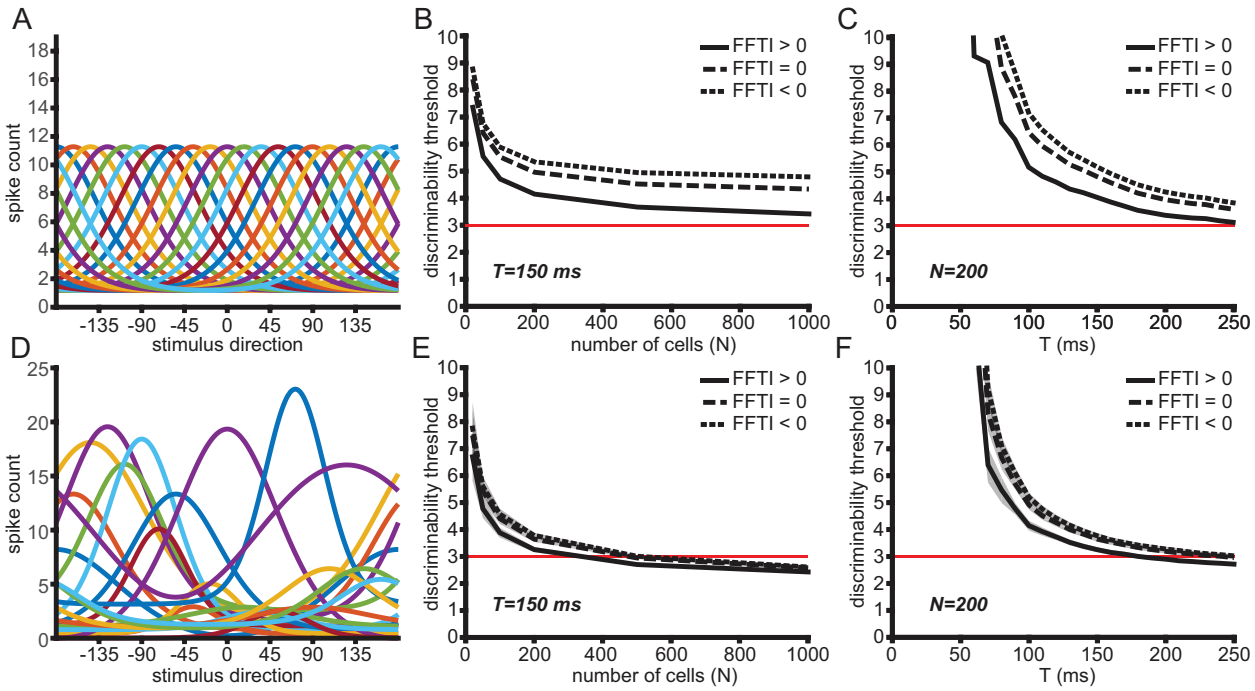

Figure 2.12: Information-limiting correlations do not affect relative decoding performance between different variance models. Same as Fig 2.7 but with the addition of information-limiting correlations where $\epsilon = 4$. Inputs to MT contain only finite information, and as such a reasonable decoding model must have limited information. This is accomplished by adding differential correlations, *i.e.*, noise correlations along the direction of the signal correlations. The differential correlations decrease decoding performance, but in a population where the information bound is not fully saturated, there is still an advantage to stimulus-dependent variability tuning.

CHAPTER 3

EFFECTS OF TEMPORAL HETEROGENEITY IN CORTICAL POPULATION CODES

Stimulus selectivity in cortical neurons is often described by a simple tuning curve that varies firing rate with stimulus features. Such a description necessarily ignores the temporal dynamics of cortical responses, with neurons varying their firing rates over time. Cortical area MT is seen to exhibit complex temporal dynamics in response to constant motion stimuli, and such temporal dynamics can improve coding performance over stationary responses. We develop a low dimensional description of temporal heterogeneity in MT and explore how the space of response dynamics impacts coding performance. A generative model of MT response dynamics allows us to probe which features of heterogeneity that are important for population coding and how heterogeneity shapes the way neural populations process information.

3.1 Introduction

Neurons in MT respond selectively to visual motion, with a tuning function that is typically described by a described by a Gaussian tuning profile that varies with stimulus direction, centered on the preferred stimulus direction of the neuron [76, 3]. On short time scales, however, the responses of MT neurons exhibit diverse temporal dynamics [31, 101]. While these temporal dynamics have some stereotyped features, such as a broadly tuned transient response to motion onset followed by a sustained response, there is considerable diversity in the specifics of these dynamics [65, 146, 101]. The heterogeneity of responses across neurons over time may improve coding performance in cortical populations [103, 131, 31, 82]. This work aims to better understand how temporal dynamics and heterogeneity fits into the tradeoff between population size, integration time, and decodability.

Many earlier attempts at understanding coding in MT treated the response as a Gaussian

tuning with a static firing rate and integration of spikes over relatively long intervals, on the order of hundreds of milliseconds (see review [108]). The timescales relevant to behavior, however, are short, and firing rates are not constant on these time scales [101]. It is not obvious how such temporal dynamics would affect coding performance. Previous work has used a principal components analysis to decompose temporal responses in V1 to suggest that temporally dynamic responses can improve information transmission over a simple rate code [82]. This work uses similar dimensionality reduction techniques to characterize the the space of temporal dynamics in MT and model the heterogeneity of the population. With a model of temporal dynamics, we provide a method to probe how the degree of diversity within the population affects coding performance in regards to population size and integration window.

3.2 Methods

3.2.1 Experimental Methods

Experimental methods and data are the same as described in chapter 2. Single-unit recordings were made with tungsten electrodes in adult macaques. Half the data ($n = 46$) was collected in sufentanil-anesthetized and paralyzed macaques and half ($n = 34$) were recorded in awake, head-fixed and fixating macaques. Random dot stimuli were presented to the receptive fields of the recorded neurons. Dot patterns translated coherently behind an aperture with size and speed of stimulus optimized for the neuron. The stimulus motion epoch lasted for 250-256 ms, with 250-256 ms periods of stationary stimulus on either side. Stimulus motion direction was randomly drawn from a range of $\pm 180^\circ$ or $\pm 90^\circ$ around the preferred direction of the neuron, in 15° increments.

3.2.2 Dimensionality Reduction

The peri-stimulus time histograms (PSTH) are calculated for each stimulus direction in each neuron by averaging spike counts in 2 ms time bins across trials, beginning at stimulus onset.

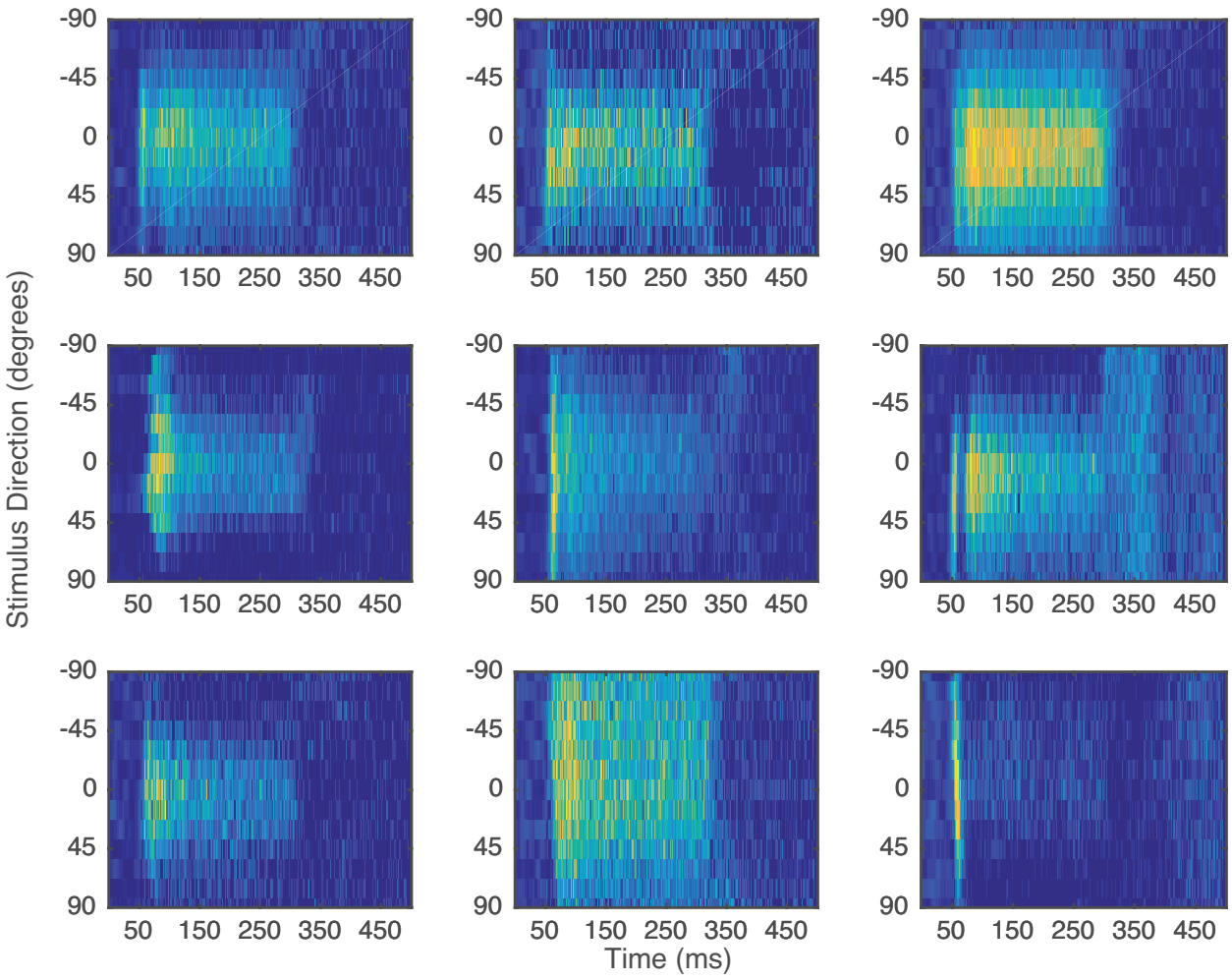

Figure 3.1: MT populations exhibit diverse temporal dynamics across neurons and stimuli. Heat maps represent the PSTH of 9 sample neurons across 13 evenly spaced stimulus directions from -90 to 90 degrees around the preferred direction of the neuron. Color maps to the probability of spiking in a 2 ms bin. Time window shown is 500 ms starting from stimulus motion onset.

Response latencies for the preferred stimulus direction are determined by the time bin at which the cumulative spike count distribution diverges from the response to the stationary stimulus. For most neurons this was done by a Kolmogorov-Smirnoff test ($p < 0.05$), but a minority of neurons required a manual correction. The PSTH for 9 sample neurons are shown in Fig 3.1, demonstrating the heterogeneity of temporal dynamics across the population and across stimuli within a single neuron.

Before taking the principal components, the PSTH were normalized by firing rate. To do this, the PSTH for each neuron across all stimulus directions was divided by a rate normalization factor defined by the average PSTH of the preferred direction for that neuron. Latencies were normalized by shifting the entire PSTH for each neuron so that each neuron began its response to the preferred latency on the same time bin as the neuron with the shortest latency, which was 48 ms. The missing bins at the end of the shifted responses were backfilled to match the PSTH length with the average response to the stationary stimulus for that neuron. The rate normalization factors and latency shifts are saved to be used in the generative model to match the distribution of rates and latencies. The distribution of rate normalization factors was fit to an exponential distribution with a rate parameter of $\lambda = 0.08615$. The fit passed a χ^2 goodness-of-fit test ($p = 0.37$). The distribution of latency shifts was fit to a negative binomial, with a mean of 34.7 ms and parameters $r = 1.95$ and $p = 0.101$ (χ^2 goodness-of-fit test $p = 0.37$).

Principal components analysis (PCA) was then performed on the normalized neurons. This is simply the eigen-decomposition of the covariance of the mean-subtracted PSTH. The variance explained by each principal component (PC) was calculated by arranging the eigenvalues of the principal components in descending order and normalizing by the sum of the eigenvalues. The first 12 PCs accounted for 90% of variance explained, and the first 15 PCs were used for most models in this document, unless otherwise stated.

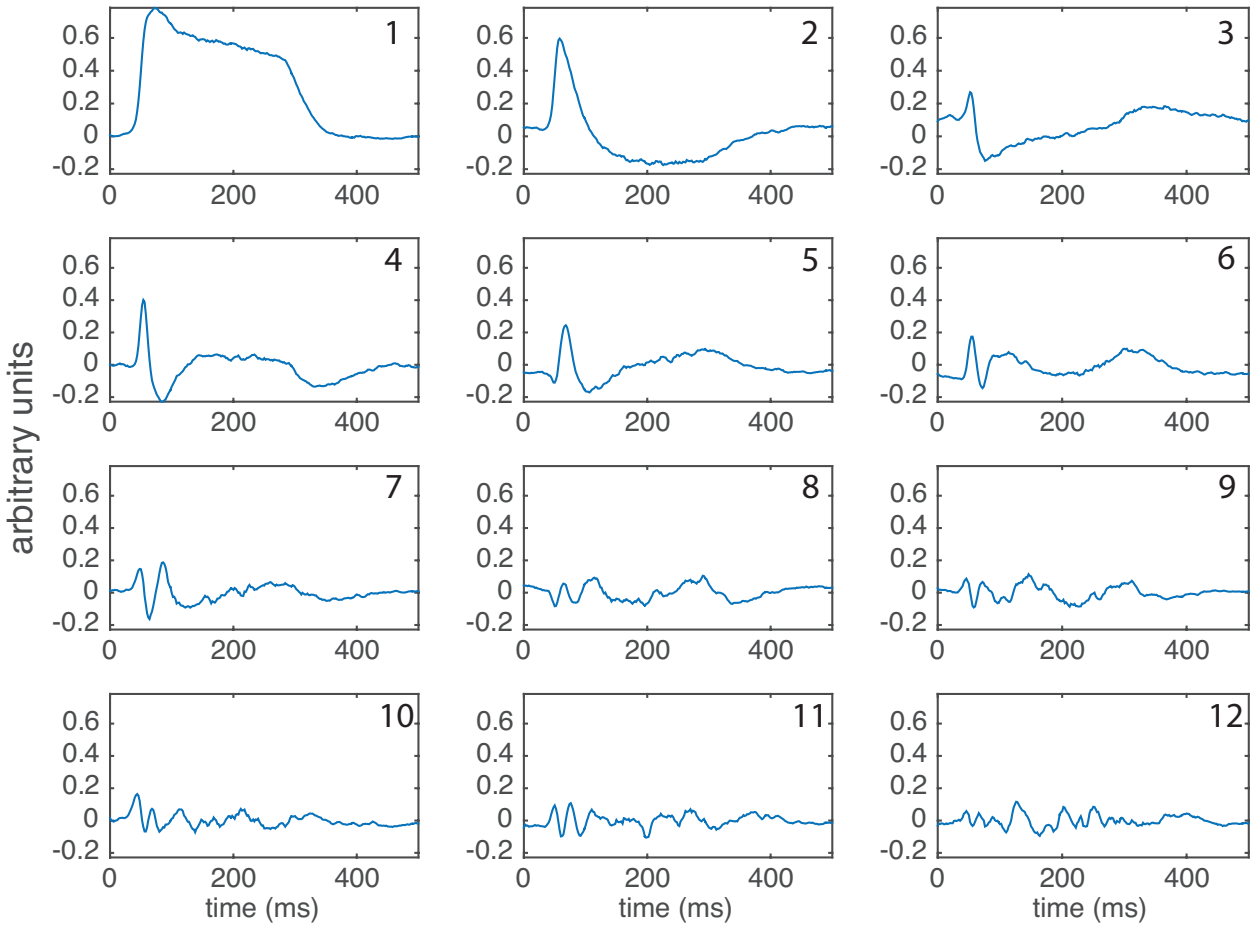

Figure 3.2: PSTH can be decomposed into principal components. Each of 12 plots shows the vector representing the corresponding principal component. The number of the principal component is shown in the top right of each plot. Principal components are scaled by the square root of the corresponding eigenvalue, in order to show the relative impact of each component.

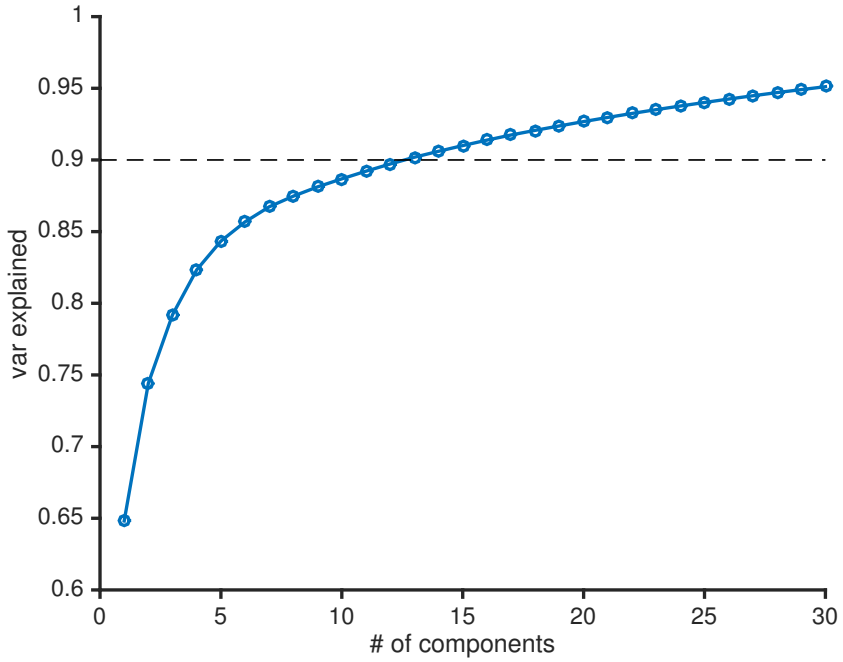

Figure 3.3: Twelve principal components explain 90% of variance in the temporal dynamics of PSTH. The cumulative variance explained is shown for the first 30 principal components, ordered by descending eigenvalues. The dashed trace indicates the 90% variance explained threshold.

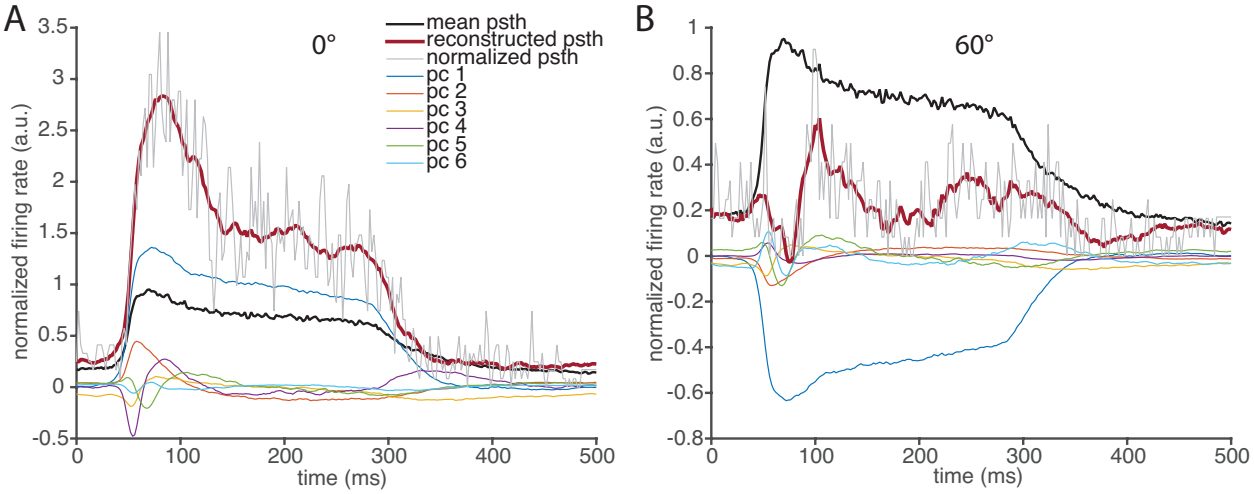

Figure 3.4: Decomposition of sample neuron PSTH into principal components. (A) The decomposition of the PSTH of the preferred stimulus direction for a sample neuron. The thicker black trace shows the mean normalized PSTH across all neurons and stimuli, which is not included in the principal components. The thick red trace is the PSTH reconstructed from the first 15 PCs. The thin gray trace is the true, normalized PSTH. The first 6 weighted PCs are also shown. (B) Same as (A) but for the stimulus 60° away from the preferred stimulus direction. The temporal dynamics can vary greatly across stimulus directions.

3.2.3 Generative Model

To generate synthetic PSTH, we first start with the response to the preferred stimulus. The distribution of PC weights for each neuron was modeled as a multivariate Gaussian with the observed mean and covariance. We then make n draws from the distribution to generate weights for n preferred-direction responses.

From here, we use a chain of conditional distributions to generate realistic PSTH. We make the assumption that the weights for PCs across two adjacent stimulus directions ($\pm 15^\circ$) are multivariate Normally distributed. We find the covariance of the weights for the preferred stimulus direction and the adjacent stimulus directions, $\pm 15^\circ$ (Fig 3.6A). The covariance matrix $\mathbf{\Sigma}$ can be decomposed as

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Sigma}_{22} & \mathbf{\Sigma}_{21} \\ \mathbf{\Sigma}_{12} & \mathbf{\Sigma}_{11} \end{pmatrix}, \tag{3.1}$$

where $\mathbf{\Sigma}_{11}$ is the covariance of weights at the preferred stimulus direction and $\mathbf{\Sigma}_{22}$ is the covariance of weights for the stimulus direction $\pm 15^\circ$ away from preferred. The other parts, $\mathbf{\Sigma}_{12} = \mathbf{\Sigma}_{21}^T$ are the terms of the covariance matrix between the two stimulus directions. Similarly the mean weights for each stimulus are given by μ_1 for the preferred stimulus and μ_2 for the adjacent directions.

Given a set of weights $\mathbf{w}_1$ for the preferred stimulus direction, the weights for the adjacent directions will be $\mathbf{w}_2 \sim \mathcal{N}(\bar{\mu}, \bar{\mathbf{\Sigma}})$, where

$$\bar{\mu} = \mu_2 + \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}(\mathbf{w}_1 - \mu_1) \tag{3.2}$$

and

$$\bar{\mathbf{\Sigma}} = \mathbf{\Sigma}_{22} - \mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12}. \tag{3.3}$$

With the conditional distribution, we draw weights $\mathbf{w}_2$ and repeat this process for the

next stimulus direction, in this case $\pm 30^\circ$. The weights $\mathbf{w}_2$, means μ_2, and covariance $\mathbf{\Sigma}_{22}$, become the new $\mathbf{w}_1$, μ_1, and $\mathbf{\Sigma}_{11}$. The corresponding means and covariance for $\pm 30^\circ$ becomes the new μ_2 and $\mathbf{\Sigma}_{22}$, $\bar{\mu}$ and $\bar{\mathbf{\Sigma}}$ are calculated, and the next set of weights $\mathbf{w}_2$ are drawn. This continues in 15° increments all the way to $\pm 180^\circ$.

Once weights are drawn, normalized PSTH are generated by multiplying the weights by the corresponding PCs and adding the together along with the mean normalized PSTH. Each synthetic PSTH is scaled by the mean rate for the preferred direction, as before, so that they can be scaled to match the true population distribution. Random rate normalization factors are drawn from the exponential distribution fit to the population rates and multiplied to scale each PSTH.

Shifting latencies requires replacing missing values with a background rate, which is determined by the mean PSTH value in all time bins up to the minimum response latency across all stimulus directions. Latencies are shifted by drawing random shifts from the negative binomial distribution fit to the observed latencies, and each PSTH is shifted accordingly. All time bins up to new response latency are replaced by the background rate.

3.2.4 Limiting Model Heterogeneity

There are several methods for limiting the diversity of temporal dynamics in the model population for testing the effects of such diversity on coding performance.

Clustering Weights

The above model is extended to generate a population of PSTH that does not make use of the full heterogeneity of temporal dynamics and instead is clustered around a narrower trajectory through PC space. To accomplish this, a multivariate Gaussian distribution is fit to the weights of PCs as before, but only a single set of weights is $\bar{\mathbf{w}}$ is drawn, which becomes the center around which subsequent weights are drawn. Weights are then generated from the distribution $\mathbf{w}_1 \sim \mathcal{N}(\bar{\mathbf{w}}, \epsilon\mathbf{\Sigma}_{11})$, where $0 < \epsilon < 1$ is a scale factor that reduces the

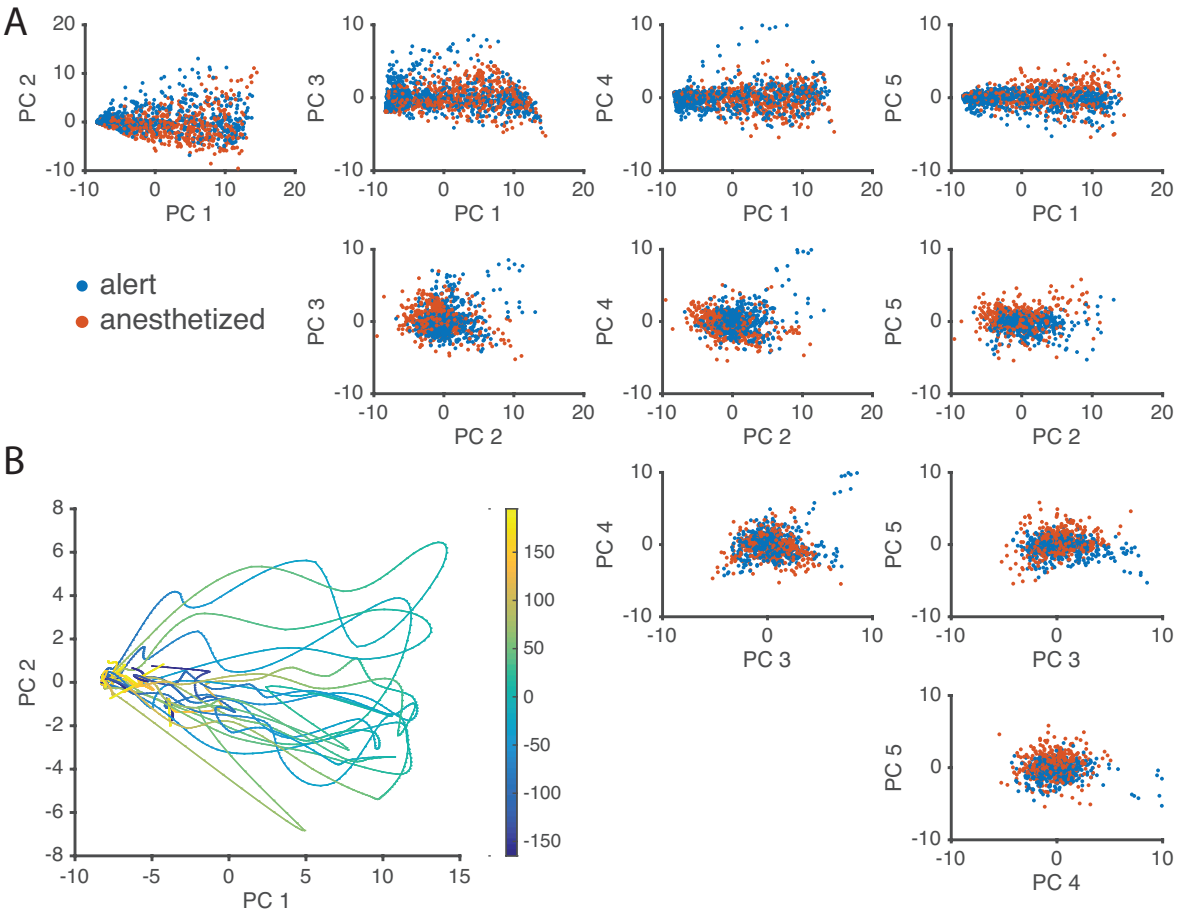

Figure 3.5: Peri-stimulus time histograms of MT neurons do not cluster in PC space. (A) Projections of PSTH onto pairs of principal components. Data points are separated by experiments recorded in an alert state (blue) or anesthetized state (orange). (B) Trajectories of 10 sample neurons through PC space, colored by stimulus direction. Trajectories correspond to the PSTH in PC space as the stimulus motion direction is rotated through 360 degrees (colorbar). The response of each neuron is considered to be the response distribution across all stimulus directions, and is therefore a manifold through PC space rather than a single point.

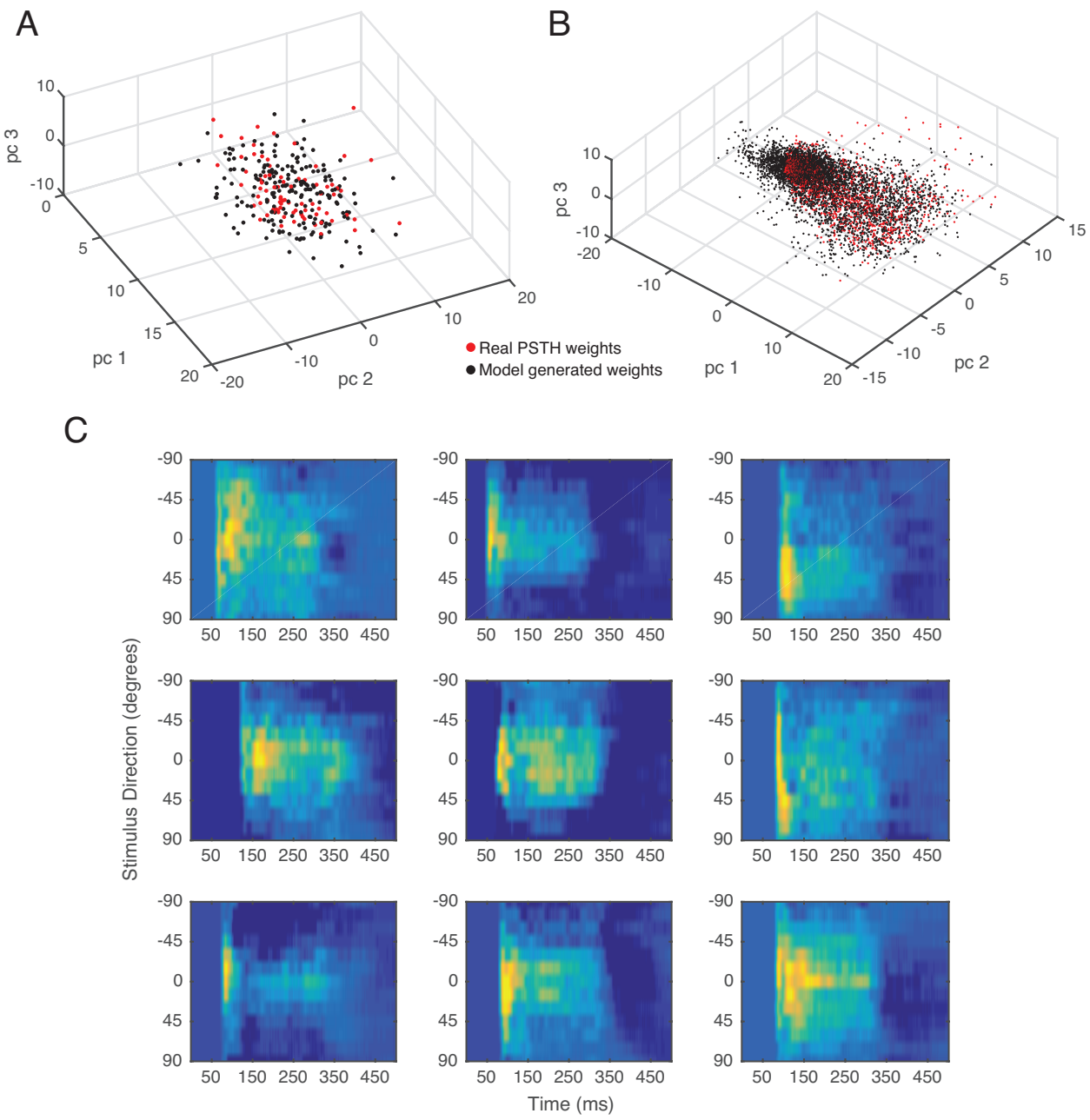

Figure 3.6: Generative model produces PSTH that span the space of temporal heterogeneity in observed populations (A) The weights of the first 3 PCs for the preferred stimulus direction for the real PSTH (red) and 200 model-generated weights (black). Generated weights are drawn from a multivariate Gaussian distribution with the same mean and variance as observed weights. (B) Same as (A) but including weights for all stimulus directions, sequentially generated. (C) Heat maps represent 9 sample model-generated PSTH across 13 evenly spaced stimulus directions from -90 to 90 degrees around the preferred direction of the neuron. Color maps to the probability of spiking in a 2 ms bin. Time window shown is 500 ms starting from stimulus motion onset. The generative model replicates temporal diversity of real neurons.

covariance of the distribution. From the weights $\mathbf{w}_1$, the conditional distribution is drawn as before, but the next set of weights are drawn from the distribution $\mathbf{w}_2 \sim \mathcal{N}(\bar{\mu}, \epsilon\bar{\mathbf{\Sigma}})$, again scaling the covariance to limit the space of weights.

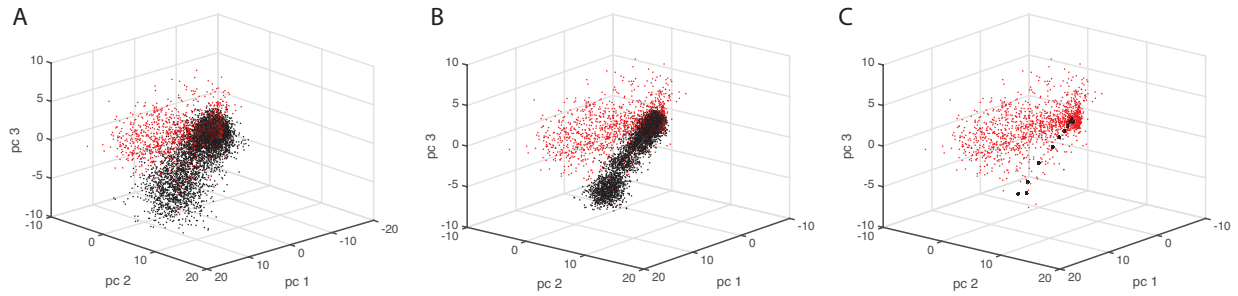

Figure 3.7: Generating model responses with clustered temporal dynamics. (A-C) The weights of the first 3 PCs for the preferred stimulus direction for the real PSTH (red) and 200 model-generated weights (black) for varying amounts of temporal heterogeneity in the population: (A) $\epsilon = 0.5$, (B) $\epsilon = 0.1$, (C) $\epsilon = 0$. In each case, the same seed was for drawing the initial weights $\bar{\mathbf{w}}$ around which the samples are drawn. This allows a more direct comparison of the effects of heterogeneity, as the temporal dynamics are clustered around the same point in PC space in each case.

Uniform Latency

Varying response latencies provide a source of time information that can be used to for decoding stimulus direction more accurately. This source of temporal diversity can be limited by drawing the same latency shift for each neuron, aligning the response latency of each neuron with the mean response latency of the population.

Uniform Dynamics

The generative model produces realistic PSTH for which the the temporal dynamics vary by stimulus direction. This diversity of temporal dynamics can be limited by only drawing weights for the preferred-stimulus direction, and using a scaled version of that PSTH for every other stimulus direction. The scaling of the PSTH was determined by randomly drawing tuning curves from the observed population of MT neurons, so that the synthetically

generated tuning curves have tuning curves that match the experimental data.

Static Rate

The temporal dynamics can be removed by replacing the PSTH, beginning with response onset, by the average firing rate. This preserves the tuning curve over the window without any additional temporal information.

3.2.5 Simulating MT Populations

A simulated population of n MT neurons is created by drawing n random PSTH from the model. The preferred stimulus direction of each neuron is drawn uniformly on $\pm 180°$. The response to a given stimulus direction is found by 2-dimensional spline interpolation of the model PSTH. The PSTH is smoothed by a sliding 10 ms averaging window prior to interpolation for more consistent results. The interpolation increases time resolution to 1 ms bins. Spike trains are simulated as a Bernoulli process, with the probability of a spike occurring in a given time bin proportional to the amplitude of the PSTH. Responses are uncorrelated across neurons.

3.2.6 SVM Decoder

For a simulated population of size n neurons, 200 spike trains were generated for each neuron for each of two stimulus directions, θ_0 and θ_1. For each spike train, spike counts were calculated at test time bins t. Spike counts were either calculated as a cumulative spike count beginning at stimulus motion onset, or as a sliding window of 30 ms terminating on time t. A support vector machine (SVM) classifier with a linear kernel was trained on the spike counts to distinguish stimuli θ_0 and θ_1. A k-fold cross-validation was used with $k = 10$. The SVM trained at time t was used to predict the stimulus direction at other time points in 10 ms increments.

Decoding performance for each model was tested with a range of population sizes ($n \in \{20, 50, 100, 200, 500\}$). Time points t were chosen to be 10 ms increments from 10 ms to 200 ms after stimulus onset. The difference in the two tested stimulus directions, $\delta\theta = \theta_1 - \theta_0$, was chosen so that $\delta\theta \in \{2°, 4°, 8°, 12°, 15°, 30°\}$. For each test condition, at least 10 different populations were tested by generating a new set of PSTH.

3.3 Results

3.3.1 Dimensionality Reduction

Recorded MT neurons exhibit a diversity of temporal dynamics across neurons and stimuli (Fig 3.1). Parsing the impact of this temporal heterogeneity on coding is difficult because it requires a principled model for comparison. We use dimensionality reduction to develop a model that can help answer the questions of how diverse is the space of temporal dynamics in MT responses, and how diverse do the dynamics need to be to achieve the same level of coding performance.

One source of heterogeneity, response latency, was accounted for by fitting a negative binomial distribution to observed latencies. Mean response latency in MT was found to be about 83 ms (see Methods).

Another source of heterogeneity, while not temporal in nature, is the average firing rate of the neurons. However, normalizing preferred-stimulus firing rates isolates the temporal dynamics so they are comparable across the population. Without normalization, most of the variance across the population is due to differences in firing rate rather than temporal dynamics.

After normalization of PSTH, PCA is used to find the dimensions of greatest variability in the space of temporal heterogeneity (Fig 3.2). While principal components are not necessarily interpretable, the results of PCA in this case do appear to have to have qualitative significance. The first PC is similar to the average PSTH across all neurons and stimuli, most

likely accounting for shifts in firing rate across stimuli and for varying transient/sustained ratios in the PSTH. The second PC increases the transient/sustained ratio by increasing the firing rate at response onset and suppressing the response during the sustained period. The next several PCs shift or sharpen the transient response.

The first 12 PCs explain 90% of the variance in the temporal dynamics of the population, which allows for a considerable dimensionality reduction while preserving most of the temporal dynamics (Fig 3.3). Adding more PCs will explain more of the variance, but 90% was chosen to be sufficient as some of that variance is due to high frequency noise in the original data due to under-sampling of the PSTH in a finite number of trials.

The weights of the PCs in the population did reveal any clustering of temporal dynamics into different classes of neurons, and instead evenly filled the space of responses (Fig 3.5A). While some responses could be categorized by eye as having features such as "transient-only" or "broad transient and narrow sustained," the responses seemed to continuously tile the space. The weights of PCs from the two separate data sets, anesthetized experiments and awake experiments, were not separable in PC space and clustering algorithms such as t-SNE and k-means clustering were not able to discriminate the two data sets.

The response of each neuron is heterogeneous across stimulus directions, and is therefore described by the set of responses as stimulus direction is varied. The response is then a manifold through PC space rather than a single point, creating a path through PC space as the stimulus parameter is varied (Fig 3.5B). The responses of an MT neuron to different stimulus directions are correlated but heterogeneous, and the characterization of these manifolds is necessary for a full description of the temporal dynamics of MT populations. Again, these manifolds or trajectories through PC space did not reveal clusters.

3.3.2 Generative Model

Given the PCs and weights, it is possible to generate PSTH from the distribution of responses observed in real neurons. This requires an estimate of the probability density function of

PSTH, which is a manifold in a high-dimensional space (Fig 3.5B). A generative model requires and estimate of the distribution of manifolds through this high dimensional space. For a 500 ms PSTH in 2 ms bins, with 24 stimulus directions, the dimensionality of the PSTH is 6000, which is too many dimensions to allow an estimate of the probability density function. Using PCA and only considering the first 15 principal components (92% of variance explained), the dimensionality is reduced to 360. This is still too high-dimensional, as there are more dimensions than the number of recorded neurons in the data set and the probability density cannot be estimated.

To overcome the dimensionality and estimate the distribution of PC weights across stimuli for MT neurons, we treated the weights for each stimulus direction as conditioned only on the weights for the adjacent stimulus direction rather than on all stimulus directions (see Methods). The weights for the preferred-stimulus direction was model as a multivariate gaussian (Fig 3.6), which is only 15-dimensional, a manageable size for our data set. Then the weights for both the 0° stimulus and the $\pm 15^\circ$ stimuli are considered jointly as a 30-dimensional multivariate Gaussian. Given weights for the 0° stimulus, the distribution of weights for $\pm 15^\circ$ can be considered as a conditional multivariate Gaussian (see Methods). This method of chained conditional distributions provides a reasonably good estimate of the distribution of weights while keeping dimensionality sufficiently low. The generative model produces PSTH that capture the temporal dynamics of real MT neurons (Fig 3.6).

For testing the impact of heterogeneity on coding, it is useful to be able to compare this model to a model with less heterogeneity or different temporal dynamics. To do this, the covariance of chained conditional distributions was scaled by a factor ϵ so weights would be limited to a narrower region of response space (Fig 3.7, see Methods). By varying the scale factor ϵ, the model still generates realistic PSTH but can limit its sampling to a smaller subspace of responses, limiting the heterogeneity.

3.3.3 Population Decoder

Model performance was tested with a support vector machine (SVM) with a linear kernel. The performance of the model can be compared to behavioral smooth pursuit data, which provides a way to place bounds on decoder performance for the MT population. Since the MT population is thought to be driving smooth pursuit, the direction discriminability threshold inferred from smooth pursuit should be comparable to the discriminability threshold of a population decoder such as the SVM. An estimate of signal-to-noise ratio in smooth pursuit provides an estimate a 3° discriminability threshold 125 ms after pursuit onset, or roughly just over 200 ms after stimulus onset given pursuit latencies in macaques [101]. The inferred discriminability is comparable to a 69% performance on a two-alternative forced choice discrimination task, so we use 69% as a threshold for discriminability in the SVM decoder. The discriminability threshold inferred from the performance of the decoder for a given population size and integration time window is compared to the behavioral data to infer properties of the population.

We first used a simple population model with the full temporal dynamics of real data with an arbitrary number of randomly generated neurons to understand the tradeoff between integration time and population size (Fig 3.8). This model indicates that after 80 ms, about the time of pursuit initiation, a population of 200 neurons can discriminate stimuli about 12° apart, and a population of 500 neurons can already distinguish stimuli separated by about 7° (Fig 3.8A). As time progress, decoder performance increases. A population of only 20 neurons achieves the level of discriminability at 160 ms of a 500 neuron population at 80 ms, twice as long 3.8C). After 200 ms, when behavioral data suggests that the neural populations driving pursuit must have a discriminability of about 3°, we find a population of between 200 to 500 neurons is required to achieve this level of performance. Introducing correlated firing across the population may change these values, but this provides a starting ground for estimating the size of a neural population required to estimate stimulus motion direction.

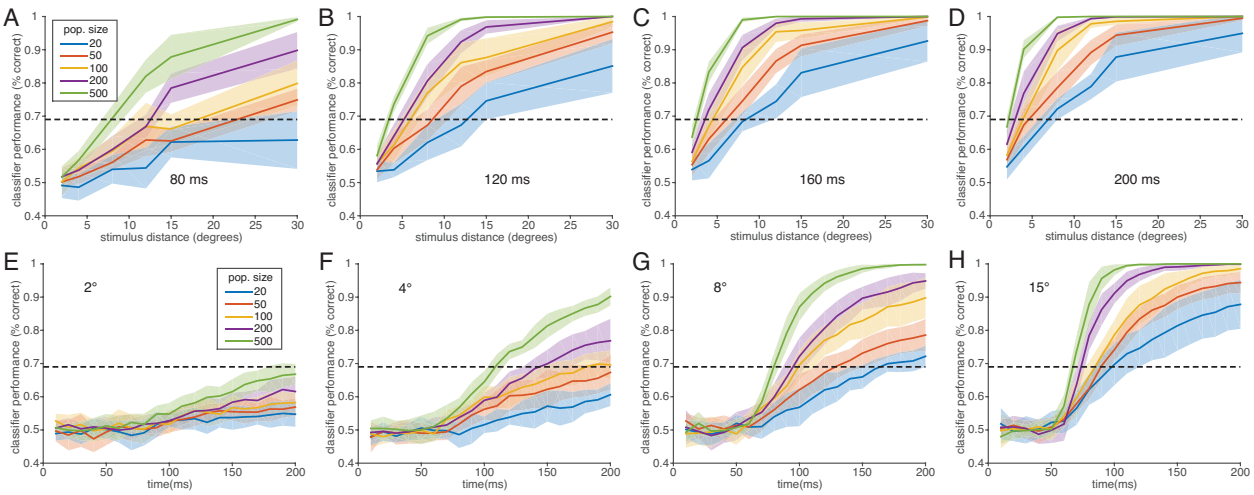

Figure 3.8: Decoder performance demonstrates tradeoff between population size and integration time. (A-D) SVM decoder performance as a function of the angular difference in stimuli being distinguished. Each trace is for a different simulated population size as shown in the legend of (A). Shaded area represents standard deviation of decoder performance. Decoder performance is shown for time points (A) 80 ms, (B) 120 ms, (C) 160 ms, and (D) 200 ms after stimulus motion onset. The dashed traces indicate a decoder performance of 69% and represents the threshold of discriminability as estimated from behavioral data [101]. (E-H) SVM decoder performance as a function of time since stimulus motion onset. Colors are population sizes as in (A-D). Performance is shown for distinguish stimuli separated by (A) 2°, (B) 4°, (C) 8°, and (D) 15°.

Dynamic response improves decoding compared to static rate

We compare the model with the full temporal heterogeneity of the population to a model in which the firing rate has been averaged out over time, producing a static response (Fig 3.9A-B). The decoder for the temporally heterogeneous model performs consistently better than the static model (Fig 3.9C). This model was only trained on one time point (160 ms) that was used to predict stimulus direction at all time points. The optimal decoder for the static tuning curve is the same at all time points, because the firing rate is the same at all points in time. Therefore, decoder performance steadily increases in time as information is accumulated until discriminability is saturated, regardless of at which time point the decoder is trained.

In contrast, for the temporally dynamic and heterogeneous response, the time at which the decoder is trained has an impact on decoder performance. Close to 160 ms, where the

decoder was optimized, the dynamic model performs better than the static model for all stimulus pairs. However, at time points further away from the training time point for the decoder, the response approaches that of the static rate.

Limiting Heterogeneity

The generative model allows for producing PSTH that do not span the full space of temporal heterogeneity, but instead cluster around a subspace. Testing the impact of this limited heterogeneity can demonstrate the impact of of heterogeneity on decoding performance. Figure 3.10 shows the impact of temporal heterogeneity on SVM decoding performance. The limited heterogeneity model uses $\epsilon = 0.1$ to restrict the response dynamics to a subspace of the full distribution. Since decoding performance may be affected by the specific subspace of dynamics chosen, multiple samplings of the distribution were made to average out the effects.

Decoding performance shows an increase in decoding performance for a more diverse population. This suggests that there is a benefit to having a more temporally heterogeneous population. For $\epsilon = 0.5$ the decoding performance was not significantly different from the model that spanned the full space of responses (not shown).

3.4 Discussion

These results demonstrate that the heterogeneous temporal dynamics of responses in MT increases decoding performance, and an accurate representation of temporal dynamics may be necessary for understanding population codes in cortex. We provide methods for characterizing the temporal heterogeneity of a population and for a generative model of population responses. These methods may be used in other areas of cortex for comparison, examining the dimensionality of responses and clustering of the response space. While we found 12 PCs account for most of the variance in the population, with responses continuously spanning

this space, that may not be the case at different levels of processing in cortex.

This model of heterogeneity, combined with behavioral outputs, can improve our understanding of coding populations. Pending analyses impose correlated spike trains while maintaining the temporal dynamics of the response as in [103]. Previously results suggest that the diversity of responses, across the population and across time, have a bigger impact on coding performance in the presence of correlations [103, 131, 157].

Other pending analyses will examine the window of integration for a cortical population. Analyses shown here use an expanding integration window. It is notable that in neurons that respond statically to a stimulus, decodability does not change in time unless the integration window expands. Additionally, after response onset of a statically firing neuron, it provides no additional information about the timing of motion onset. We can use our model to examine the impact of integration time on a temporally diverse population, and how a decoding population might optimize its decoding window.

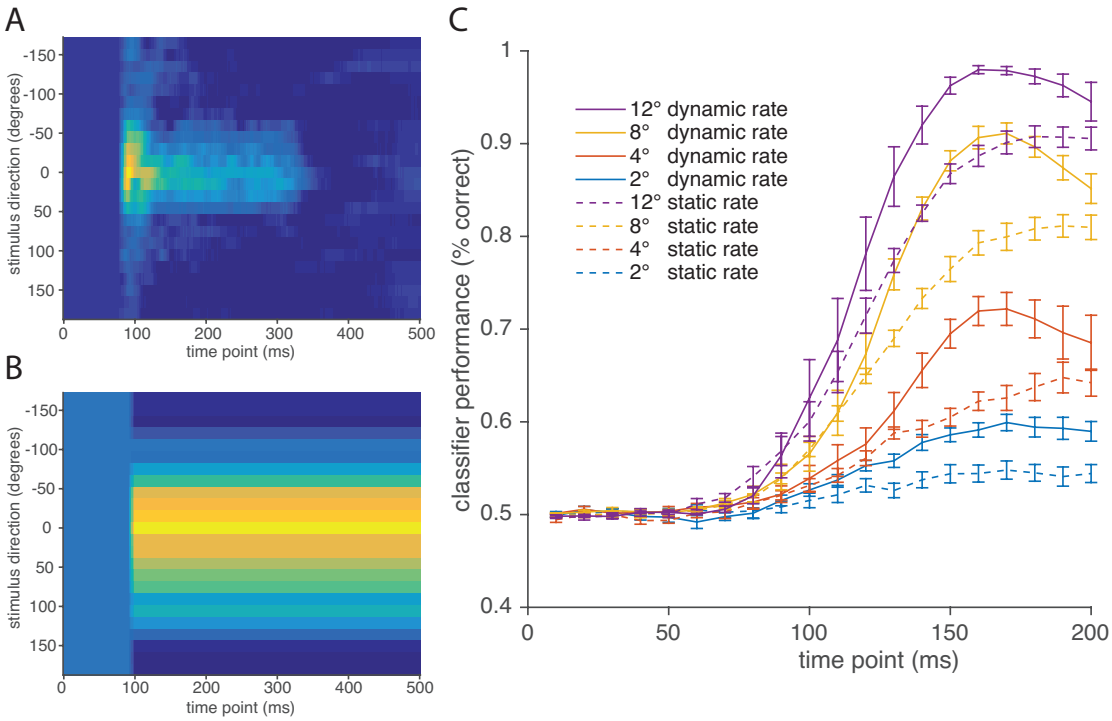

Figure 3.9: Temporally dynamic response profiles improve population decoder performance. (A) Sample model-generated PSTH response profile with realistic temporal dynamics. (B) Sample model-generated PSTH with static response profile. The PSTH without temporal dynamics is generated the same was as the model with temporal dynamics and the rate is simply averaged over time beginning at response onset. (C) Model decoder performance distinguishing motion stimuli separated by 2° (blue), 4° (red), 8° (yellow), or 12° (purple). The decoder is trained on data 160 ms after stimulus onset and tested on data in 10 ms increments from 10 ms to 200 ms after stimulus onset. For each stimulus pair, the model with temporal dynamics (solid traces) outperformed the static rate model (dashed traces). The model with temporal dynamics performs better around the training time of the decoder, but performance is comparable for time points far away from 160 ms because the temporal information is averaged out by the downstream integrator.

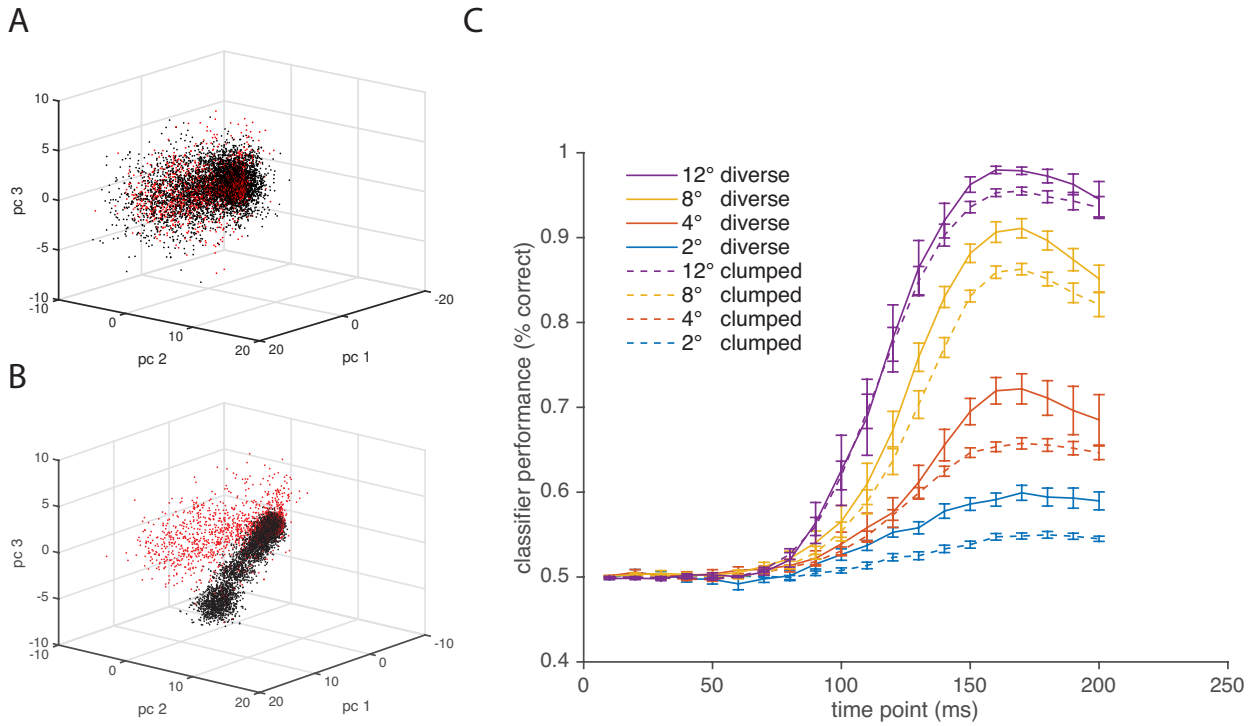

Figure 3.10: Diversity of temporal dynamics improve population decoder performance. (A) Distribution of scores of first 3 PCs for model-generated PSTH with full heterogeneity of temporal dynamics. (B) Distribution of scores of first 3 PCs for model-generated PSTH with limited heterogeneity of temporal dynamics, clumped together (scaling factor $\epsilon = 0.1$. The weights of only one population with limited heterogeneity is shown, but multiple populations were generated clustered around different trajectories through PC space. (C) Model decoder performance distinguishing motion stimuli separated by 2° (blue), 4° (red), 8° (yellow), or 12° (purple). The decoder is trained on data 160 ms after stimulus onset and tested on data in 10 ms increments from 10 ms to 200 ms after stimulus onset. For each stimulus pair, the model with realistic temporal heterogeneity (solid traces) outperformed the model with limited heterogeneity (dashed traces), particularly for more similar stimulus directions. Error bars show the standard error of the mean.

CHAPTER 4

DISCUSSION

One of the largest looming problems in neuroscience is the question of how information is coded and transmitted in the brain. What is the neural code?

The roles of different brain regions can be teased apart at a systems level through various techniques. Physical connections between brain areas can be traced. Lesions can reveal where brain regions are connected along a behavioral pathway. The feature selectivity of a brain region can be determined by reverse correlation of stimuli or behavior to neural responses. These systems level descriptions of inputs and outputs can go a long way toward answering the *what* questions. What information is encoded in a brain region? What is that information used for downstream? What stimulus drove this response? These questions can be answered while the particulars of the neural code remain something of a black box.

Cortical area MT is well described at the systems level. The areas that input to MT and the areas that receive its projections are known. The role of area MT in driving eye movements and perceiving visual motion is well known and characterized. The correlation of MT neuron responses with the presentation of simple stimuli reveal the selectivity for visual motion direction and speed. What is not known is precisely *how* information is manipulated and transmitted through this system.

In order to understand how a brain area performs its computations, the distribution of its responses must be known. However, the complexity of the problem is so great that we need to make certain simplifying assumptions in our models. One cannot know *a priori* which simplifying assumptions are justified, which details of the population response are critical for an efficient neural code. Features such as spike-count variability, temporal dynamics, inter-neuronal correlations, and state-dependent changes can all affect the population code.

We find, along with others [156, 107], that neurons in area MT exhibit spike count variability that is is tuned to the stimulus identity and is not described by a simple Poisson process. However, we find that this changes with behavioral state, and a Poisson or gamma-

Poisson assumption holds for MT populations in animals under Sufentanil anesthesia. Spike-count variability is therefore both stimulus-dependent and state-dependent, changing with the attentional state of the subject. We demonstrate how these state-dependent changes in the population code can impact the efficiency and accuracy of the code. A deeper understanding of the neural code must account for how changes in the stimulus and the attentional state affect the responses of the neural population as well their downstream effects.

We also examined the temporal dynamics of the responses in MT. While models of responses are often simplified to tuning functions that are static in time, the responses of cortical neurons are heterogeneous through time. Responses are often simplified to static tuning curves because it is difficult to concisely describe the complex and diverse temporal dynamics of cortical neurons. However, this simplification throws away a possibly large source of information. We used dimensionality reduction techniques to developed a method for characterizing the temporal dynamics of MT neurons in a simple, low-dimensional way. Being able to accurately model the temporal dynamics of a neural population opens up the possibility of testing different features of temporal dynamics and population heterogeneity in information coding. We find that a rich, diverse, and dynamic neural response consistently improves coding over simpler alternatives.

REFERENCES

[1] Ida EJ Aasebø, Mikkel E Lepperød, Maria Stavrinou, Sandra Nøkkevangen, Gaute Einevoll, Torkel Hafting, and Marianne Fyhn. Temporal processing in the visual cortex of the awake and anesthetized rat. *eNeuro*, 4(4):ENEURO–0059, 2017.

[2] Duane G Albrecht and David B Hamilton. Striate cortex of monkey and cat: contrast response function. *Journal of neurophysiology*, 48(1):217–237, 1982.

[3] T. D. Albright. Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, 52(6):1106–1130, December 1984.

[4] Henry J Alitto, Daniel L Rathbun, W Martin Usrey, et al. A comparison of visual responses in the lateral geniculate nucleus of alert and anaesthetized macaque monkeys. *The Journal of physiology*, 589(1):87–99, 2011.

[5] J Allman, F Miezin, and E McGuinness. Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception*, 14(2):105–126, 1985.

[6] John C. Anderson, Tom Binzegger, Kevan A. C. Martin, and K. S. Rockland. The Connection from Cortical Area V1 to V5: A Light and Electron Microscopic Study. *The Journal of Neuroscience*, 18(24):10525–10540, December 1998.

[7] John C. Anderson and Kevan A. C. Martin. Connection from cortical area V2 to MT in macaque monkey. *The Journal of Comparative Neurology*, 443(1):56–70, January 2002.

[8] Ehsan Arabzadeh, Stefano Panzeri, and Mathew E Diamond. Whisker vibration information carried by rat barrel cortex neurons. *Journal of Neuroscience*, 24(26):6011–6020, 2004.

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

[10] Wyeth Bair, James R. Cavanaugh, Matthew A. Smith, and J. Anthony Movshon. The Timing of Response Onset and Offset in Macaque Visual Neurons. *The Journal of Neuroscience*, 22(8):3189–3205, April 2002.

[11] Wyeth Bair and Lawrence P O'keefe. The influence of fixational eye movements on the response of neurons in area mt of the macaque. *Visual neuroscience*, 15(4):779–786, 1998.

[12] Wyeth Bair, Ehud Zohary, and William T Newsome. Correlated firing in macaque visual area mt: time scales and relationship to behavior. *Journal of Neuroscience*, 21(5):1676–1697, 2001.

[13] G R Barnes and P T Asselman. The mechanism of prediction in human smooth pursuit eye movements. *The Journal of Physiology*, 439(1):439–461, July 1991.

[14] Richard T. Born and David C. Bradley. Structure and Function of Visual Area Mt. *Annual Review of Neuroscience*, 28(1):157–189, 2005.

[15] RT Born, JM Groh, R Zhao, and SJ Lukasewycz. Segregation of object and background motion in visual area mt: effects of microstimulation on eye movements. *Neuron*, 26(3):725–734, 2000.

[16] Emery N Brown, Ralph Lydic, and Nicholas D Schiff. General anesthesia, sleep, and coma. *New England Journal of Medicine*, 363(27):2638–2650, 2010.

[17] Emery N Brown, Patrick L Purdon, and Christa J Van Dort. General anesthesia and altered states of arousal: a systems neuroscience analysis. *Annual review of neuroscience*, 34:601–628, 2011.

[18] Nicolas Brunel and Jean-Pierre Nadal. Mutual information, fisher information, and population coding. *Neural computation*, 10(7):1731–1757, 1998.

[19] Giedrius T Buracas, Anthony M Zador, Michael R DeWeese, and Thomas D Albright. Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron*, 20(5):959–969, 1998.

[20] A. Burkhalter, D. J. Felleman, W. T. Newsome, and D. C. Van Essen. Anatomical and physiological asymmetries related to visual areas V3 and VP in macaque extrastriate cortex. *Vision Research*, 26(1):63–80, 1986.

[21] Daniel A Butts and Mark S Goldman. Tuning curves, neuronal variability, and sensory coding. *PLoS biology*, 4(4):e92, 2006.

[22] Matteo Carandini. Amplification of trial-to-trial response variability by neurons in visual cortex. *PLoS biology*, 2(9):e264, 2004.

[23] J. R. Carl and R. S. Gellman. Human smooth pursuit: stimulus-dependent responses. *Journal of Neurophysiology*, 57(5):1446–1463, May 1987.

[24] Miguel Castelo-Branco, Sergio Neuenschwander, and Wolf Singer. Synchronization of visual responses between the cortex, lateral geniculate nucleus, and retina in the anesthetized cat. *Journal of neuroscience*, 18(16):6395–6410, 1998.

[25] Zhe Chen, Sujith Vijayan, ShiNung Ching, Greg Hale, Francisco J Flores, Matthew A Wilson, and Emery N Brown. Assessing neuronal interactions of cell assemblies during general anesthesia. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 4175–4178. IEEE, 2011.

[26] Mark M. Churchland, Byron M. Yu, John P. Cunningham, Leo P. Sugrue, Marlene R. Cohen, Greg S. Corrado, William T. Newsome, Andrew M. Clark, Paymon Hosseini, Benjamin B. Scott, David C. Bradley, Matthew A. Smith, Adam Kohn, J. Anthony

Movshon, Katherine M. Armstrong, Tirin Moore, Steve W. Chang, Lawrence H. Snyder, Stephen G. Lisberger, Nicholas J. Priebe, Ian M. Finn, David Ferster, Stephen I. Ryu, Gopal Santhanam, Maneesh Sahani, and Krishna V. Shenoy. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nature Neuroscience*, 13(3):369–378, March 2010.

[27] Bertrand S Clarke and Andrew R Barron. Information-theoretic asymptotics of bayes methods. *IEEE Transactions on Information Theory*, 36(3):453–471, 1990.

[28] Marlene R Cohen and John HR Maunsell. Attention improves performance primarily by reducing interneuronal correlations. *Nature neuroscience*, 12(12):1594–1600, 2009.

[29] Christine M Constantinople and Randy M Bruno. Effects and mechanisms of wakefulness on local cortical networks. *Neuron*, 69(6):1061–1068, 2011.

[30] Erik P Cook and John HR Maunsell. Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey. *Journal of Neuroscience*, 22(5):1994–2004, 2002.

[31] Erik P Cook and John HR Maunsell. Dynamics of neuronal responses in macaque mt and vip during motion detection. *Nature neuroscience*, 5(10):985–994, 2002.

[32] AF Dean. The variability of discharge of simple cells in the cat striate cortex. *Experimental Brain Research*, 44(4):437–440, 1981.

[33] M. R. Dursteler and R. H. Wurtz. Pursuit and optokinetic deficits following chemical lesions of cortical areas MT and MST. *Journal of Neurophysiology*, 60(3):940–965, September 1988.

[34] M. R. Dursteler, R. H. Wurtz, and W. T. Newsome. Directional pursuit deficits following lesions of the foveal representation within the superior temporal sulcus of the macaque monkey. *Journal of Neurophysiology*, 57(5):1262–1287, May 1987.

[35] 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.

[36] Alexander S Ecker, Philipp Berens, Andreas S Tolias, and Matthias Bethge. The effect of noise correlations in populations of diversely tuned neurons. *Journal of Neuroscience*, 31(40):14272–14283, 2011.

[37] G. N. Elston and M. G. Rosa. The occipitoparietal pathway of the macaque monkey: comparison of pyramidal cell morphology in layer III of functionally related cortical visual areas. *Cerebral Cortex*, 7(5):432–452, July 1997.

[38] D. J. Felleman and J. H. Kaas. Receptive-field properties of neurons in middle temporal visual area (MT) of owl monkeys. *Journal of Neurophysiology*, 52(3):488–513, September 1984.

[39] D. J. Felleman and D. C. Van Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex (New York, N.Y.: 1991)*, 1(1):1–47, February 1991.

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

[41] A. F. Fuchs. Saccadic and smooth pursuit eye movements in the monkey. *The Journal of Physiology*, 191(3):609–631, August 1967.

[42] R. Gattass and C. G. Gross. Visual topography of striate projection zone (MT) in posterior superior temporal sulcus of the macaque. *Journal of Neurophysiology*, 46(3):621–638, September 1981.

[43] Geoffrey M Ghose and John HR Maunsell. Attentional modulation in visual cortex depends on task timing. *Nature*, 419(6907):616–620, 2002.

[44] M. Glickstein, J. G. May, and B. E. Mercier. Corticopontine projection in the macaque: the distribution of labelled cortical cells after large injections of horseradish peroxidase in the pontine nuclei. *The Journal of Comparative Neurology*, 235(3):343–359, May 1985.

[45] Mitchell Glickstein, Janet Lee Cohen, Bryan Dixon, Alan Gibson, Mark Hollins, Eilene Labossiere, and Farrel Robinson. Corticopontine visual projections in macaque monkeys. *The Journal of Comparative Neurology*, 190(2):209–229, March 1980.

[46] Robbe L. T. Goris, J. Anthony Movshon, and Eero P. Simoncelli. Partitioning neuronal variability. *Nature Neuroscience*, 17(6):858–865, June 2014.

[47] Jennifer M Groh, Richard T Born, and William T Newsome. How is a sensory map read out? effects of microstimulation in visual area mt on saccades and smooth pursuit eye movements. *Journal of Neuroscience*, 17(11):4312–4330, 1997.

[48] Moshe Gur, Alexander Beylin, and D Max Snodderly. Response variability of neurons in primary visual cortex (v1) of alert monkeys. *Journal of Neuroscience*, 17(8):2914–2920, 1997.

[49] Kenneth D Harris and Alexander Thiele. Cortical state and attention. *Nature reviews neuroscience*, 12(9):509–523, 2011.

[50] Sonja S Hohl, Kris S Chaisanguanthum, and Stephen G Lisberger. Sensory population decoding for visually guided movements. *Neuron*, 79(1):167–179, 2013.

[51] Xin Huang and Stephen G Lisberger. Noise correlations in cortical area mt and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. *Journal of Neurophysiology*, 101(6):3012–3030, 2009.

[52] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106–154, 1962.

[53] Krešimir Josić, Eric Shea-Brown, Brent Doiron, and Jaime de la Rocha. Stimulus-Dependent Correlations and Population Codes. *Neural Computation*, 21(10):2774–2804, July 2009.

[54] Prakash Kara, Pamela Reinagel, and R Clay Reid. Low response variability in simultaneously recorded retinal, thalamic, and cortical neurons. *Neuron*, 27(3):635–646, 2000.

[55] K Kawano, M Shidara, Y Watanabe, and S Yamane. Neural activity in cortical area mst of alert monkey during ocular following responses. *Journal of neurophysiology*, 71(6):2305–2324, 1994.

[56] Justin Keat, Pamela Reinagel, R Clay Reid, and Markus Meister. Predicting every spike: a model for the responses of visual neurons. *Neuron*, 30(3):803–817, 2001.

[57] E. L. Keller and S. J. Heinen. Generation of smooth-pursuit eye movements: neuronal mechanisms and pathways. *Neuroscience Research*, 11(2):79–107, July 1991.

[58] Michael A Kisley and George L Gerstein. Trial-to-trial variability and state-dependent modulation of auditory-evoked responses in cortex. *Journal of Neuroscience*, 19(23):10451–10460, 1999.

[59] H. Komatsu and R. H. Wurtz. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *Journal of Neurophysiology*, 60(2):580–603, August 1988.

[60] H. Komatsu and R. H. Wurtz. Modulation of pursuit eye movements by stimulation of cortical areas MT and MST. *Journal of Neurophysiology*, 62(1):31–47, July 1989.

[61] S G Lisberger, E J Morris, and L Tychsen. Visual Motion Processing and Sensory-Motor Integration for Smooth Pursuit Eye Movements. *Annual Review of Neuroscience*, 10(1):97–129, 1987.

[62] S. G. Lisberger and L. E. Westbrook. Properties of visual inputs that initiate horizontal smooth pursuit eye movements in monkeys. *The Journal of Neuroscience*, 5(6):1662–1673, June 1985.

[63] SG Lisberger. Neural basis for motor learning in the vestibuloocular reflex of primates. iii. computational and behavioral analysis of the sites of learning. *Journal of Neurophysiology*, 72(2):974–998, 1994.

[64] SG Lisberger and TA Pavelko. Functional properties of brainstem cells inhibited from the cerebellar flocculus in monkey. In *Soc. Neurosci. Abstr*, number 290.7 in 10, 1984.

[65] Stephen G. Lisberger and J. Anthony Movshon. Visual Motion Analysis for Pursuit Eye Movements in Area MT of Macaque Monkeys. *The Journal of Neuroscience*, 19(6):2224–2246, March 1999.

[66] Bing Liu, Matthew V. Macellaio, and Leslie C. Osborne. Efficient sensory cortical coding optimizes pursuit eye movements. *Nature Communications*, 7, 2016.

[67] M. Livingstone and D. Hubel. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, 240(4853):740–749, May 1988.

[68] M. S. Livingstone and D. H. Hubel. Connections between layer 4b of area 17 and the thick cytochrome oxidase stripes of area 18 in the squirrel monkey. *The Journal of Neuroscience*, 7(11):3371–3377, November 1987.

[69] Steven J Luck, Leonardo Chelazzi, Steven A Hillyard, and Robert Desimone. Neural mechanisms of spatial selective attention in areas v1, v2, and v4 of macaque visual cortex. *Journal of neurophysiology*, 77(1):24–42, 1997.

[70] J. C. Lynch. Frontal eye field lesions in monkeys disrupt visual pursuit. *Experimental Brain Research*, 68(2):437–441, October 1987.

[71] Wei Ji Ma, Jeffrey M Beck, Peter E Latham, and Alexandre Pouget. Bayesian inference with probabilistic population codes. *Nature neuroscience*, 9(11):1432–1438, 2006.

[72] Martha G. MacAvoy, Jacqueline P. Gottlieb, and Charles J. Bruce. Smooth-Pursuit Eye Movement Representation in the Primate Frontal Eye Field. *Cerebral Cortex*, 1(1):95–102, January 1991.

[73] Gaby Maimon and John A Assad. Beyond poisson: increased spike-time regularity across primate parietal cortex. *Neuron*, 62(3):426–440, 2009.

[74] Julio C Martinez-Trujillo and Stefan Treue. Feature-based attention increases the selectivity of population responses in primate visual cortex. *Current Biology*, 14(9):744–751, 2004.

[75] J. H. Maunsell, T. A. Nealey, and D. D. DePriest. Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *The Journal of Neuroscience*, 10(10):3323–3334, October 1990.

[76] J. H. Maunsell and D. C. van Essen. Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, 49(5):1127–1147, May 1983.

[77] J. H. Maunsell and D. C. van Essen. Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *Journal of Neurophysiology*, 49(5):1148–1167, May 1983.

[78] J. H. Maunsell and DC van Essen. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of Neuroscience*, 3(12):2563–2586, December 1983.

[79] J H R Maunsell and W T Newsome. Visual Processing in Monkey Extrastriate Cortex. *Annual Review of Neuroscience*, 10(1):363–401, 1987.

[80] J. G. May, E. L. Keller, and D. A. Suzuki. Smooth-pursuit eye movement deficits with chemical lesions in the dorsolateral pontine nucleus of the monkey. *Journal of Neurophysiology*, 59(3):952–977, March 1988.

[81] Carrie J McAdams and John HR Maunsell. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area v4. *Journal of Neuroscience*, 19(1):431–441, 1999.

[82] John W McClurkin, Lance M Optican, Barry J Richmond, and Timothy J Gawne. Concurrent processing and complexity of temporally encoded neuronal messages in visual perception. *Science*, 253(5020):675–677, 1991.

[83] Matthew J McGinley, Martin Vinck, Jacob Reimer, Renata Batista-Brito, Edward Zagha, Cathryn R Cadwell, Andreas S Tolias, Jessica A Cardin, and David A McCormick. Waking state: rapid variations modulate neural and behavioral responses. *Neuron*, 87(6):1143–1161, 2015.

[84] Craig H. Meyer, Adrian G. Lasker, and David A. Robinson. The upper limit of human smooth pursuit velocity. *Vision Research*, 25(4):561–563, 1985.

[85] Jude F. Mitchell, Nicholas J. Priebe, and Cory T. Miller. Motion dependence of smooth pursuit eye movements in the marmoset. *Journal of Neurophysiology*, page jn.00197.2015, April 2015.

[86] Jude F Mitchell, Kristy A Sundberg, and John H Reynolds. Differential attention-dependent response modulation across cell classes in macaque visual area v4. *Neuron*, 55(1):131–141, 2007.

[87] Jude F Mitchell, Kristy A Sundberg, and John H Reynolds. Spatial attention decorrelates intrinsic activity fluctuations in macaque area v4. *Neuron*, 63(6):879–888, 2009.

[88] Majid H Mohajerani, Allen W Chan, Mostafa Mohsenvand, Jeffrey LeDue, Rui Liu, David A McVea, Jamie D Boyd, Yu Tian Wang, Mark Reimers, and Timothy H Murphy. Spontaneous cortical activity alternates between motifs defined by regional axonal projections. *Nature neuroscience*, 16(10):1426–1435, 2013.

[89] Nathan Montgomery and Michael Wehr. Auditory cortical neurons convey maximal stimulus-specific information at their best frequency. *Journal of Neuroscience*, 30(40):13362–13366, 2010.

[90] Jorrit S. Montijn, Martin Vinck, and Cyriel M. A. Pennartz. Population coding in mouse visual cortex: response reliability and dissociability of stimulus tuning and noise correlation. *Frontiers in Computational Neuroscience*, 8:1–15, June 2014.

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

[92] J. A. Movshon, E. H. Adelson, M. S. Gizzi, and W. T. Newsome. The analysis of moving visual patterns. *Pattern Recognition Mechanisms*, 1986.

[93] Trishna Mukherjee, Matthew Battifarano, Claudio Simoncini, and Leslie C Osborne. Shared sensory estimates for human motion perception and pursuit eye movements. *Journal of Neuroscience*, 35(22):8515–8530, 2015.

[94] Trishna Mukherjee, Bing Liu, Claudio Simoncini, and Leslie C Osborne. Spatiotemporal filter for visual motion integration from pursuit eye movements in humans and monkeys. *Journal of Neuroscience*, 37(6):1394–1412, 2017.

[95] M. J. Mustari, A. F. Fuchs, and J. Wallman. Response properties of dorsolateral pontine units during smooth pursuit in the rhesus macaque. *Journal of Neurophysiology*, 60(2):664–686, August 1988.

[96] W. T. Newsome, R. H. Wurtz, M. R. Dursteler, and A. Mikami. Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *The Journal of Neuroscience*, 5(3):825–840, March 1985.

[97] W. T. Newsome, R. H. Wurtz, and H. Komatsu. Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *Journal of Neurophysiology*, 60(2):604–620, August 1988.

[98] Amy M Ni and John HR Maunsell. Spatially tuned normalization explains attention modulation variance within neurons. *Journal of neurophysiology*, 118(3):1903–1913, 2017.

[99] Amy M Ni, Supratim Ray, and John HR Maunsell. Tuned normalization explains the size of attention modulations. *Neuron*, 73(4):803–813, 2012.

[100] Leslie C. Osborne, William Bialek, and Stephen G. Lisberger. Time Course of Information about Motion Direction in Visual Area MT of Macaque Monkeys. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(13):3210–3222, March 2004.

[101] Leslie C Osborne, Sonja S Hohl, William Bialek, and Stephen G Lisberger. Time course of precision in smooth-pursuit eye movements of monkeys. *Journal of Neuroscience*, 27(11):2987–2998, 2007.

[102] Leslie C Osborne, Stephen G Lisberger, and William Bialek. A sensory source for motor variation. *Nature*, 437(7057):412–416, 2005.

[103] Leslie C. Osborne, Stephanie E. Palmer, Stephen G. Lisberger, and William Bialek. The Neural basis for combinatorial coding in a cortical population responseeural basis for combinatorial coding in a cortical population response. *Journal of Neuroscience*, 28(50):13522–13531, December 2008.

[104] Stefano Panzeri, Riccardo Senatore, Marcelo A Montemurro, and Rasmus S Petersen. Correcting for the sampling bias problem in spike train information measures. *Journal of neurophysiology*, 98(3):1064–1072, 2007.

[105] Jonathan W Pillow and Eero P Simoncelli. Biases in white noise analysis due to non-poisson spike generation. *Neurocomputing*, 52:109–115, 2003.

[106] Carlos R. Ponce, Stephen G. Lomber, and Richard T. Born. Integrating motion and depth via parallel pathways. *Nature Neuroscience*, 11(2):216–223, February 2008.

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

[108] Alexandre Pouget, Peter Dayan, and Richard Zemel. Information processing with population codes. *Nature Reviews Neuroscience*, 1(2):125, 2000.

[109] Nicholas J Priebe and David Ferster. Mechanisms of neuronal computation in mammalian visual cortex. *Neuron*, 75(2):194–208, 2012.

[110] Nicholas J. Priebe, Stephen G. Lisberger, and J. Anthony Movshon. Tuning for Spatiotemporal Frequency and Speed in Directionally Selective Neurons of Macaque Striate Cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(11):2941–2950, March 2006.

[111] Simon J. D. Prince, Andrew D. Pointon, Bruce G. Cumming, and Andrew J. Parker. The Precision of Single Neuron Responses in Cortical Area V1 during Stereoscopic Depth Judgments. *The Journal of Neuroscience*, 20(9):3387–3400, May 2000.

[112] SE Raiguel, D-K Xiao, VL Marcar, and GA Orban. Response latency of macaque area mt/v5 neurons and its relationship to stimulus parameters. *Journal of Neurophysiology*, 82(4):1944–1956, 1999.

[113] Steven E Raiguel, Lieven Lagae, Balàzs Gulyàs, and Guy A Orban. Response latencies of visual cells in macaque areas v1, v2 and v5. *Brain research*, 493(1):155–159, 1989.

[114] C. Rashbass. The relationship between saccadic and smooth tracking eye movements. *The Journal of Physiology*, 159(2):326–338, December 1961.

[115] John H Reynolds and David J Heeger. The normalization model of attention. *Neuron*, 61(2):168–185, 2009.

[116] Jorma J Rissanen. Fisher information and stochastic complexity. *IEEE transactions on information theory*, 42(1):40–47, 1996.

[117] D A Robinson. The mechanics of human smooth pursuit eye movement. *The Journal of Physiology*, 180(3):569–591, October 1965.

[118] K. S. Rockland. Bistratified distribution of terminal arbors of individual axons projecting from area V1 to middle temporal area (MT) in the macaque monkey. *Visual Neuroscience*, 3(2):155–170, August 1989.

[119] K. S. Rockland. Morphology of individual axons projecting from area V2 to MT in the macaque. *The Journal of Comparative Neurology*, 355(1):15–26, April 1995.

[120] Kathleen S. Rockland. Laminar Distribution of Neurons Projecting from Area V1 to V2 in Macaque and Squirrel Monkeys. *Cerebral Cortex*, 2(1):38–47, January 1992.

[121] H. R. Rodman, C. G. Gross, and T. D. Albright. Afferent basis of visual response properties in area MT of the macaque. I. Effects of striate cortex removal. *The Journal of Neuroscience*, 9(6):2033–2050, June 1989.

[122] H. R. Rodman, C. G. Gross, and T. D. Albright. Afferent basis of visual response properties in area MT of the macaque. II. Effects of superior colliculus removal. *The Journal of Neuroscience*, 10(4):1154–1164, April 1990.

[123] SAMUEL Ron and David A Robinson. Eye movements evoked by cerebellar stimulation in the alert monkey. *Journal of Neurophysiology*, 36(6):1004–1022, 1973.

[124] Srivatsun Sadagopan and David Ferster. Feedforward origins of response variability underlying contrast invariant orientation tuning in cat visual cortex. *Neuron*, 74(5):911–923, 2012.

[125] Matthew T Schmolesky, Youngchang Wang, Doug P Hanes, Kirk G Thompson, Stefan Leutgeb, Jeffrey D Schall, and Audie G Leventhal. Signal timing across the macaque visual system. *Journal of neurophysiology*, 79(6):3272–3278, 1998.

[126] Joseph W Schumacher, David M Schneider, and Sarah MN Woolley. Anesthetic state modulates excitability but not spectral tuning or neural discrimination in single auditory midbrain neurons. *Journal of Neurophysiology*, 106(2):500–514, 2011.

[127] G Sclar and RD Freeman. Orientation selectivity in the cat's striate cortex is invariant with stimulus contrast. *Experimental brain research*, 46(3):457–461, 1982.

[128] M. A. Segraves, M. E. Goldberg, S. Y. Deng, C. J. Bruce, L. G. Ungerleider, and M. Mishkin. The role of striate cortex in the guidance of eye movements in the monkey. *The Journal of Neuroscience*, 7(10):3040–3058, October 1987.

[129] Eyal Seidemann and William T Newsome. Effect of spatial attention on the responses of area mt neurons. *Journal of neurophysiology*, 81(4):1783–1794, 1999.

[130] Michael N. Shadlen and William T. Newsome. The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding. *The Journal of Neuroscience*, 18(10):3870–3896, May 1998.

[131] Maoz Shamir and Haim Sompolinsky. Implications of Neuronal Diversity on Population Coding. *Neural Comput.*, 18(8):1951–1986, August 2006.

[132] S. Shipp and S. Zeki. The Organization of Connections between Areas V5 and V1 in Macaque Monkey Visual Cortex. *The European Journal of Neuroscience*, 1(4):309–332, 1989.

[133] S. Shipp and S. Zeki. The Organization of Connections between Areas V5 and V2 in Macaque Monkey Visual Cortex. *European Journal of Neuroscience*, 1(4):333–354, July 1989.

[134] Lawrence C. Sincich and Jonathan C. Horton. Independent Projection Streams from Macaque Striate Cortex to the Second Visual Area and Middle Temporal Area. *The Journal of Neuroscience*, 23(13):5684–5692, July 2003.

[135] Lawrence C. Sincich, Ken F. Park, Melville J. Wohlgemuth, and Jonathan C. Horton. Bypassing V1: a direct geniculate input to area MT. *Nature Neuroscience*, 7(10):1123–1128, October 2004.

[136] Matthew A Smith and Adam Kohn. Spatial and temporal scales of neuronal correlation in primary visual cortex. *Journal of Neuroscience*, 28(48):12591–12603, 2008.

[137] William R Softky and Christof Koch. The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *Journal of Neuroscience*, 13(1):334–350, 1993.

[138] Greg J Stephens, Leslie C Osborne, and William Bialek. Searching for simplicity in the analysis of neurons and behavior. *Proceedings of the National Academy of Sciences*, 108(Supplement 3):15565–15571, 2011.

[139] Iwona Stepniewska, Hui-Xin Qi, and Jon H. Kaas. Do superior colliculus projection zones in the inferior pulvinar project to MT in primates? *European Journal of Neuroscience*, 11(2):469–480, February 1999.

[140] Charles F Stevens and Anthony M Zador. Input synchrony and the irregular firing of cortical neurons. *Nature neuroscience*, 1(3):210–217, 1998.

[141] J. R. Tian and J. C. Lynch. Corticocortical input to the smooth and saccadic eye movement subregions of the frontal eye field in Cebus monkeys. *Journal of Neurophysiology*, 76(4):2754–2771, October 1996.

[142] J. R. Tian and J. C. Lynch. Functionally defined smooth and saccadic eye movement subregions in the frontal eye field of Cebus monkeys. *Journal of Neurophysiology*, 76(4):2740–2753, October 1996.

[143] J. Tigges, M. Tigges, S. Anschel, N. A. Cross, W. D. Letbetter, and R. L. McBride. Areal and laminar distribution of neurons interconnecting the central visual cortical areas 17, 18, 19, and MT in squirrel monkey (Saimiri). *The Journal of Comparative Neurology*, 202(4):539–560, November 1981.

[144] DJ Tolhurst. Separate channels for the analysis of the shape and the movement of a moving visual stimulus. *The Journal of Physiology*, 231(3):385–402, 1973.

[145] Rory G Townsend, Selina S Solomon, Spencer C Chen, Alexander NJ Pietersen, Paul R Martin, Samuel G Solomon, and Pulin Gong. Emergence of complex wave patterns in primate cerebral cortex. *Journal of Neuroscience*, 35(11):4657–4662, 2015.

[146] Andreas Traschütz, Andreas K. Kreiter, and Detlef Wegener. Transient activity in monkey area MT represents speed changes and is correlated with human behavioral performance. *Journal of Neurophysiology*, 113(3):890–903, February 2015.

[147] Stefan Treue and John HR Maunsell. Effects of attention on the processing of motion in macaque middle temporal and medial superior temporal visual cortical areas. *Journal of Neuroscience*, 19(17):7591–7602, 1999.

[148] Stefan Treue and Julio C Martinez Trujillo. Feature-based attention influences motion processing gain in macaque visual cortex. *Nature*, 399(6736):575, 1999.

[149] Leslie G. Ungerleider and Robert Desimone. Cortical connections of visual area MT in the macaque. *The Journal of Comparative Neurology*, 248(2):190–222, June 1986.

[150] Leslie G. Ungerleider, Robert Desimone, Thelma W. Galkin, and Mortimer Mishkin. Subcortical projections of area MT in the macaque. *The Journal of Comparative Neurology*, 223(3):368–386, March 1984.

[151] Freek van Ede, Floris de Lange, Ole Jensen, and Eric Maris. Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha-and beta-band oscillations. *Journal of Neuroscience*, 31(6):2016–2024, 2011.

[152] Justin L Vincent, Gaurav H Patel, Michael D Fox, Abraham Z Snyder, Justin T Baker, David C Van Essen, John M Zempel, Lawrence H Snyder, Maurizio Corbetta, and Marcus E Raichle. Intrinsic functional architecture in the anaesthetized monkey brain. *Nature*, 447(7140):83–86, 2007.

[153] D. M. Vogt Weisenhorn, R. B. Ilung, and W. B. Spatz. Morphology and connections of neurons in area 17 projecting to the extrastriate areas mt and 19dm and to the superior colliculus in the monkey Callithrix jacchus. *The Journal of Comparative Neurology*, 362(2):233–255, November 1995.

[154] G. Westheimer. Eye movement responses to a horizontally moving visual stimulus. *A.M.A. Archives of Ophthalmology*, 52(6):932–941, December 1954.

[155] GERALD WESTHEIMER and SIDNEY M BLAIR. Oculomotor defects in cerebellectomized monkeys. *Investigative Ophthalmology & Visual Science*, 12(8):618–621, 1973.

[156] Benjamin White, Larry F Abbott, and József Fiser. Suppression of cortical neural variability is stimulus-and state-dependent. *Journal of neurophysiology*, 108(9):2383–2392, 2012.

[157] Adrien Wohrer and Christian K Machens. On the number of neurons and time scale of integration underlying the formation of percepts in the brain. *PLoS computational biology*, 11(3):e1004082, 2015.

[158] N. Harumi Yabuta, Atomu Sawatari, and Edward M. Callaway. Two Functional Channels from Primary Visual Cortex to Dorsal Visual Cortical Areas. *Science*, 292(5515):297–300, April 2001.

[159] D. S. Yamasaki and R. H. Wurtz. Recovery of function after lesions in the superior temporal sulcus in the monkey. *Journal of Neurophysiology*, 66(3):651–673, September 1991.

[160] D. S. Zee, R. J. Tusa, S. J. Herdman, P. H. Butler, and G. Gucer. Effects of occipital lobectomy upon eye movements in primate. *Journal of Neurophysiology*, 58(4):883–907, October 1987.

[161] Kechen Zhang and Terrence J Sejnowski. Neuronal tuning: To sharpen or broaden? *Neural Computation*, 11(1):75–84, 1999.

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

[163] Joel Zylberberg, Jon Cafaro, Maxwell H. Turner, Eric Shea-Brown, and Fred Rieke. Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code. *Neuron*, 89(2):369–383, January 2016.