

UNIVERSITY OF ROCHESTER

DOCTORAL THESIS

Neural Tuning in High-Dimensional Stimulus Spaces

Author:
Hayden SCOTT

Supervisor:
Adam SNYDER



*Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in the Department of*

Brain and Cognitive Sciences
School of Arts and Sciences
University of Rochester
Rochester, NY

August 21, 2023

Contents

Biographical Sketch	v
Acknowledgements	vii
Abstract	ix
1 Introduction	1
1.1 The Visual Heirarchy	2
1.2 The Neural Response Function	4
1.2.1 Tuning in One Dimension	4
1.2.2 nD Stimulus Spaces: Multiple Features	7
1.2.3 nD Response Spaces: Multiple Neurons	9
1.2.4 Next Steps	11
2 Neuronal Modulation Depends on Context and Feature Tuning	13
2.1 Introduction	13
2.2 Methods	15
2.3 Results	28
2.4 Discussion	42
2.5 Conclusion	46
3 Relating Neural Data to Stimulus Parameters in 3+ Dimensions	47
3.1 Simulated Neural Populations	47
3.1.1 Defining the Population Response Function	48
3.1.2 Deterministic Simulations	49

3.1.3	Poisson Variability Simulations	51
3.2	Optimization	53
3.2.1	Derivative-Based Optimization	54
3.2.2	Derivative-Free Optimization	55
4	Using Generative Models of Natural Images to Define Neural Tuning Manifolds	59
4.1	Introduction	59
4.2	Methods	61
4.3	Results	70
4.4	Discussion	81
4.5	Conclusion	84
5	General Discussion	85
Bibliography		87

Biographical Sketch

The author was born in Louisville, KY, USA. He attended The University of Kentucky and graduated cum laude with a Bachelor of Science in Neuroscience. He received a Masters degree in Brain and Cognitive Sciences from the University of Rochester in 2020. He began doctoral studies in Brain and Cognitive Sciences at the University of Rochester in 2020. He was awarded the NSF NRT-DESE: Data-Enabled Research into Human Behavior and its Cognitive and Neural Mechanisms Fellowship in 2018. He pursued his research in neural computations underlying feature based attention under the direction of Dr. Adam Snyder.

The following publications were a result of work conducted during doctoral study:

1. Scott, H., Wimmer, K., Pasternak, T., & Snyder, A. C. (2023). Altered task demands lead to a division of labour for sensory and cognitive processing in the middle temporal area. *European Journal of Neuroscience*, 1– 16. doi: [10.1111/ejne.14011](https://doi.org/10.1111/ejne.14011)
2. Scott, H., Murphey, A., Briggs, F., & Snyder, A. C. (2023). Using Generative Models of Natural Images to Define Neural Tuning Manifolds. In Prep.

Acknowledgements

A special thanks to everyone who helped me get here:

Advisor:

Adam Snyder

Collaborators:

Tania Pasternak

Farran Briggs

Klaus Wimmer

Allison Murphrey

Lab Members:

Ellie Sachse

Megan Conley

Chad Dekdebrun

Bikash Sahoo

Dennis Jung

Shizhao Liu

Jay Gonzolez-Ramirez

Other:

Ryan Burke

Rolo

Twizzler

Gremlin

UNIVERSITY OF ROCHESTER

Abstract

Hayden Scott

Brain and Cognitive Sciences

Doctor of Philosophy

Neural Tuning in High-Dimensional Stimulus Spaces

by Hayden SCOTT

One central goal of the field of neuroscience is to understand the neural code. How do neural systems process information, and what information is being conveyed across neural populations? Visual science in particular has been addressing these questions for decades, looking into the stimulus-response function of neurons throughout the visual hierarchy. In the early years, the goal was to record one neuron, manipulate one stimulus feature, and quantify what information is being processed. As the field moves towards more natural settings, such as neural populations and natural images, a new framework is needed to relate high-dimensional neural and stimulus data during behavioral tasks. In this thesis I will address the problem of complex and flexible feature tuning that aids behavioral goals using techniques pulled from neuroscience, information theory, and artificial intelligence.

List of Tables

2.1 Task effects summary. Significance of task effects (p value) separated by information clusters. Statistics for Tuning and entropy (HF) were a three-way anova between task, motion direction, and stimulus (S1 vs. S2). Delay ramping was compared using t-tests on the slopes (Passive minus active: (2.8 G-I).	45
--	----

List of Figures

1.1	<i>The Rhesus Macaque Visual Hierarchy.</i> Each area depicted here is distinct in cytoarchitecture and maintains a complete retinotopic map. Information flows from the Lateral Geniculate Nucleus (LGN) (not shown) to V1, then V2, V4, and finally the inferior temporal cortex (IT) in what we call the Ventral stream.	2
1.2	<i>Calculating a 1D Tuning Curve.</i> A.) The process for estimating a single neurons tuning curve to a single stimulus parameter θ . Discrete stimulus values were chosen (-180° by 45° to 180°) to span the stimulus space, and repeatedly presented (Gray dots). This provides the expected response to a given stimulus value (dotted line) for the neuron. B.) The estimated tuning and variability for the neuron in panel A.	5
1.3	<i>Tuning curves and Information.</i> A.) Von Mises tuning functions with different concentration parameters (κ). Color indicates highest information tuning functions for fine (Blue) and coarse (red) discrimination. B.) Mutual information (Firing Rate and orientation) as a function of κ . Blue line depicts 1° discrimination, Red line depicts the more realistic situation of 8 stimulus classes (0° to 315° in 45° steps).	6
1.4	<i>Gabors as points in a 2D Stimulus Space.</i> A.) examples of Gabor stimuli with various θ and spatial frequency values. B.) Positions of the Gabors depicted in panel A within a 2-Dimensional stimulus space.	7
1.5	<i>Tuning along multiple Feature Dimensions.</i> A neurons tuning curve, in 2 dimensions becomes a tuning surface. Gray plane represents the spatial frequency value at which the tuning for θ matches figure 1.2.	9

1.6 <i>Population Codes</i> A.) Tuning curves for 8 homogenous model neurons along the orientation feature dimension. Color at the bottom represents the values of θ . B.) A 2-Dimensional state space plotting the responses of neurons 1&2 from panel A. Color represents the value of θ that lead to each pair of neural responses.	10
2.1 Cross validation of clustering algorithm. We took 50% of trials for each neuron, calculated mutual information (MI), ran the clustering algorithm, and minimized AIC to pick the optimum number of clusters. After repeating this process 100 times, we got a distribution of models that supports 3 clusters as optimal for our data.	21
2.2 Relationship between Entropy Factor (HF) and Fano Factor (FF). We took the neural responses for our 8 directions of motion (S1), for each neuron in our sample, and plotted FF vs HF. Dotted lines are Poisson processes. Neural responses with a larger lambda parameter (average firing rate) result in a roughly linear relationship between the measures. When neural responses are low, FF overestimates the variability (darker blue dots are highly super-Poisson according to FF, but roughly Poisson according to HF)	22
2.3 Example and overview of tuning goodness of fit (R^2). A Gaussian fit to one example neuron during one bootstrap fold. The full distribution of R^2 for our sample of neurons.	25
2.4 Average Comparison Effect (CE) during the two windows of interest for neurons' preferred direction and 180° off preferred (anti-preferred).	27

- 2.7 *Motion Information Clustering.* A.) Time series for active minus passive motion information (MI) for each neuron. Rows correspond to neurons and color represents the difference in information. See Methods for the clustering algorithm. B.) Within-cluster averages from panel A. 34
- 2.8 *Effects of task demands differed depending on how informative a neuron was about task-relevant information.* A.) Average raw tuning curves for IDI neurons during the active (purple) and passive (green) tasks. Both curves contain the same neurons. Error bars are $\pm SEM$ across neurons. B-C.) Same as A for DDI and SDI groups respectively. D.) Average trial-to-trial variability during S1 quantified with entropy factor (HF) (See Methods). E-F.) Same as panel D for DDI neurons and SDI groups respectively. G.) Slope of a linear fit to IDI neurons' firing rates vs time during a 1s interval before S2 onset in the Active vs Passive tasks. There was a significantly higher ramping during the active task (**: $p = 5.66 \cdot 10^{-5}$, t-test). H-I.) Same as panel G for DDI (ns) and SDI (**: $p = 0.0098$) groups respectively. J.) Task effect of the width parameter of gaussian fits vs ΔMI during S1. For each cross validation fold, we fit a gaussian to the 50% of trials not used for the calculation of MI (i.e. the X and Y axes use different trials). We then averaged across folds, and subtracted across task to get a task effect of tuning width. We found no significant relationship between tuning width and ΔMI ($R^2 = 0.038, p = 0.552$). K.) Same as J but for the height parameter of the gaussian fits ($R^2 = 0.859, p = 3.505 \cdot 10^{-75}$). . . . 35

2.9 <i>DDI and SDI neurons have different comparison effects across the active and passive tasks.</i> A.) Time course of comparison effects (AUC between same and different trials within stimulus type) for the IDI neurons with significant $S > D$ in either task. B.) Same as panel A but for IDI neurons with $D > S$ signal. C-D.) Same as A-B but for DDI neurons. E-F.) Same as A-B but for SDI neurons. **: $p < 0.01$, ***: $p < 0.001$, n.s.: not significant t-test for the difference between tasks over the indicated interval. We did not find significant CEs or task effects during the S2 (gray shading) for any group of neurons, but we did see effects during the post-S2 interval for DDI neurons (C-D), and for SDI neurons with $S > D$ CEs (E).	38
2.10 <i>Proportion of neurons with significant comparison signals during active and passive tasks</i> A.) Proportion of neurons within each information group with significant $S > D$ signal during the S2. B.) Same as panel A for the post-S2 interval. C) Same as in panel A for the $D > S$ neurons. D) Same as in panel B for the $D > S$ neurons. (***: $p < 0.001$, χ^2 test). E.) Comparison Effects (absolute difference from 0.5 AUC) vs mutual information for each task. Neurons with the largest CE in the active task also conveyed the least direction information. Error bars = 95% binomial confidence intervals.	41
3.1 <i>The effect of tuning width on Canonical Correlations Analysis (CCA).</i> Top row.) Example tuning curves for three simulations with descending tuning widths. First column.) Canonical correlation variable pairs for a population of 3 neurons with wide tuning. Neurons were assumed to have the same tuning width but different preferred stimuli along the three dimensions. Second&Third column.) Same as first column but for medium and narrow tuning respectively.	50

3.2	<i>The effect of irrelevant stimulus dimensions on CCA.</i> Canonical correlation variable pairs for three models relating neural population space to stimulus space. Model 1.) Similar to models shown in figure 3.1, with all three stimulus dimensions being relevant to the neurons. Tuning for neurons was set at 50 a.u. across models. Model 2.) Same as model 1, except one stimulus parameter (consistent across neurons) was set to have no impact on the firing rate of the neurons. Model 3.) Same as model 2 but with multiple irrelevant stimulus dimensions.	51
3.3	<i>The effect of Poisson variability on CCA.</i> Same as figure 3.1 with the exception of added Poisson variability to the neural responses.	52
3.4	<i>Optimization Example.</i> This example demonstrates the simplest case of optimization. Some linear, deterministic function (blue line) is sampled repeatedly and the gradient is followed to find the minimum.	54
3.5	<i>Non-Convex Optimization Example.</i> Here I show how gradient-based optimization can fail in even simple 1D problems. When a function is non-convex (multiple minima/maxima), the local gradient can often be misleading about the global behavior of the function.	55
3.6	<i>Genetic Algorithm Flowchart.</i> The general process of how a genetic algorithm optimizes.	57
4.1	<i>Example GAN Stimuli.</i> Four example images generated by the 128 GAN latents (top row) and their pixel reconstruction after dimensionality reduction with PCA (bottom row).	71
4.2	<i>Individual neuron model fits</i> A.) AIC for GLMs using the GAN latents as predictors (x-axis) and the pixel predictors (y-axis) on neuron spike counts. B.) Relative change in AIC across models for individual neurons. Negative values indicate that the GAN latents are better predictors of neural data while positive values indicate that the pixel model was better. The latent model was better on average in both brain regions (V1: $t = -10.26$, $p = 1.54 \cdot 10^{-23}$, V4: $t = -6.02$, $p = 3.46 \cdot 10^{-9}$; t-test).	73

4.3 <i>Canonical Correlations in neural data</i> A.) Simulated stimulus data where points were randomly selected to span an arbitrary space. The red line indicates the axis chosen by CCA for panel D. B.) Neural state space of stimuli from panel A, where each axis is the firing rate of a neuron. Data was centered for visualization purposes as CCA also centers data. The purple line indicates the axis in neural space chosen by CCA for panel D. C.) Original correlation between neuron 1 (panel B) and latent 1 (panel A). D.) First canonical variable pair from the data in panels A and B. The X and Y axes are now linear combinations of the latent and neural dimensions respectively.	74
4.4 <i>Baseline-Corrected R values</i> The baseline-corrected r values for each canonical variable pair across brain region and algorithm. Baseline correction was calculated by subtracting out the distribution of r values calculated from randomly permuting predictors and calculating CCA 100 times. The number of significant CCA pairs was calculated using Rao's approximate F statistic.	75
4.5 <i>Linear Relationships in V1</i> A.) Top CCA pair for a PSO session in V1. Each point is one stimulus and color indicates the normalized L^2 norm of the population response vector to that image. B.) Same as panel A for a session optimized with the genetic algorithm. C.) Same as panels A&B for a session in which no optimization was performed, and instead random images were shown. D-F.) Same as panels A-C except the color now indicates the order of stimuli in the session. G-I.) The GAN latent dimension that CCA found was most linearly related to neural activity for each algorithm. These dimensions correspond to the X-axes in panels A-F.	76

4.6 <i>Linear Relationships in V4</i> A.) Top CCA pair for a PSO session in V4. Each point is one stimulus and color indicates the normalized L^2 norm of the population response vector to that image. B.) Same as panel A for a session optimized with the genetic algorithm. C.) Same as panels A&B for a session in which no optimization was performed, and instead random images were shown. D-F.) Same as panels A-C except the color now indicates the order of stimuli in the session. G-I.) The GAN latent dimension that CCA found was most linearly related to neural activity for each algorithm. These dimensions correspond to the X-axes in panels A-F.	77
4.7 <i>Projections onto CCA1</i> The value of projections onto Canonical variable pair 1 for each stimulus shown. PSO tends to pick one side of the latent space that it knows has strong optimization evaluations, whereas the genetic algorithm explores more.	78
4.8 <i>DCA pair 1 in V1 across Algorithms</i> A.) Top DCA pair for a PSO session in V4. Each point is one stimulus and color indicates the normalized L^2 norm of the population response vector to that image. B.) Same as panel A for a session optimized with the genetic algorithm. C.) Same as panels A&B for a session in which no optimization was performed, and instead random images were shown. D-F.) Same as panels A-C except the color now indicates the order of stimuli in the session. G-I.) The GAN latent dimension that DCA found was most strongly related to neural activity for each algorithm. These dimensions correspond to the X-axes in panels A-F.	79

4.9 <i>DCA pair 1 in V4 across Algorithms</i>	A.) Top DCA pair for a PSO session in V4. Each point is one stimulus and color indicates the normalized L^2 norm of the population response vector to that image.	B.) Same as panel A for a session optimized with the genetic algorithm.	C.) Same as panels A&B for a session in which no optimization was performed, and instead random images were shown.	D-F.) Same as panels A-C except the color now indicates the order of stimuli in the session.	G-I.) The GAN latent dimension that DCA found was most strongly related to neural activity for each algorithm. These dimensions correspond to the X-axes in panels A-F.	80
4.10 <i>Relative Linearity by Brain region</i>	A.) Linearity of neural/stimulus relationships by DCA dimension pairs. Stimulus parameters used were the GAN latent variables. DCA dimension pairs are organized by descending distance covariance and relative linearity was calculated as described in Methods. Each line represents one session, and is colored according to brain area.	B.) Same as panel A for pixel predictors.	81			

List of Symbols

Information Theory:

H	entropy	<i>nats</i>
MI	mutual information	<i>bits</i>

Neuron Simulations:

$\vec{\theta}$	Parameter values for stimulus θ	a.u.
$f_n(\vec{\theta})$	Response for neuron n to stimulus $\vec{\theta}$	$\frac{\text{spk}}{\text{s}}$
R_{max}	Maximum response	$\frac{\text{spikes}}{\text{second}}$
τ_n	Tuning matrix for neuron n	a.u.
λ	Rate parameter	$\frac{\text{spikes}}{\text{second}}$
$D_n(\vec{\theta})$	Decay term for neuron n	a.u.
$\vec{\rho}_n$	Preferred stimulus for neuron n	a.u.

List of Abbreviations

GAN	Generative Adversarial Network
CNN	Convolutional Neural Network
MI	Mutual Information
HF	Entropy (H) Factor
RGC	Retinal Ganglion Cells
LGN	Lateral Geniculate Nucleus
IT	Inferior Temporal cortex
V1	Visual area 1
V4	Visual area 4
MT	Middle Temporal area
LPFC	Lateral Prefrontal Cortex
SF	Spatial Frequency
PCA	Principal Component Analysis
DVA	Degree of Visual Angle
CCA	Canonical Correlations Analysis
DCA	Distance Covariance Analysis
PSO	Particle Swarm Optimization
AUC	Area Under the Curve of receiver operating characteristic

Chapter 1

Introduction

Humans are highly visual animals. We rely on our sight to effectively navigate the world around us, which has lead to extensive research over the last few decades into the machinations of our brains that produce such vivid qualia. Most laymen would agree that light enters our eyes and this information is then processed by the brain, but the specific computations that occur have only begun to be elucidated with recent developments in artificial neural networks. These evolving methods for describing neural computations parallel more sophisticated electrophysiology techniques, such as high-yield neural recordings capable of recording hundreds-to-thousands of neurons simultaneously. This intersection of artificial and biological intelligence has brought about as many problems as it has solutions, including the curse of dimensionality, noise, and the black-box problem. I will focus first on a brief history of visual neuroscience, and introduce the idea of stimulus tuning and its' dependence on context (Chapter 2). Second I will expand beyond 1D tuning functions using computational models which introduce high-dimensional stimulus spaces, the curse of dimensionality, and optimization approaches to circumvent it (Chapter 3). Finally, I will expand these ideas through an experiment which leverages machine learning for adaptive stimulus selection (Chapter 4).

1.1 The Visual Hierarchy

The mammalian visual system is grossly organized hierarchically (Felleman and Essen, 1991; Barone et al., 2000; Batardiere, 2002). Light first enters the retina and activates photoreceptors (Field et al., 2010), then this information propagates through the Lateral Geniculate Nucleus (LGN) of the thalamus to V1. From V1, neurons send projections down the visual cortex posterior to anterior through areas V2, V3, V4 and into Inferior Temporal cortex (IT) (1.1) in what is called the ventral stream.

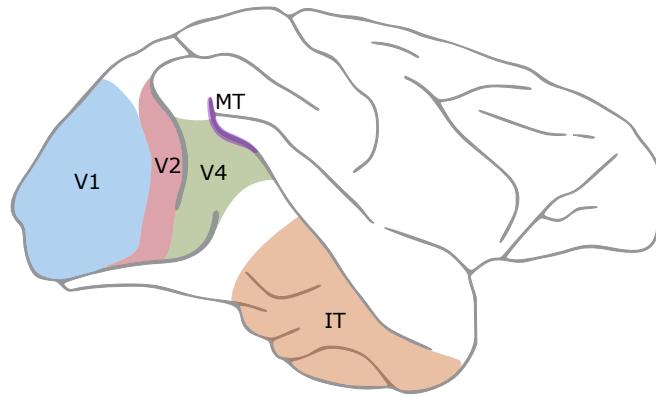


FIGURE 1.1: *The Rhesus Macaque Visual Hierarchy*. Each area depicted here is distinct in cytoarchitecture and maintains a complete retinotopic map. Information flows from the LGN (not shown) to V1, then V2, V4, and finally the inferior temporal cortex (IT) in what we call the Ventral stream.

It is known that each of these brain regions are unique and self-contained stages in hierarchical information processing. One reason they are thought to be unique stages is because each area has a full representation of the visual field (Retinotopic map) (Felleman, Xiao, and McClendon, 1997). It has also been shown that neurons in these brain areas respond differently to the same visual information (Mahon and Valois, 2001), and have different functional architecture (Yoshioka, Dow, and Vautin, 1996; Hubel and Wiesel, 1965). At each stage of the hierarchy, the visual information is processed in increasingly complex ways. Information in photoreceptors is quite simple: it consists of how many photons of light there are at each point in the visual field. The retinal ganglion cells (RGCs) then sum that information and describe the first spatial derivative of the photoreceptors, where the changes in the concentration of light are (i.e. contrast

edges; Wiesel, 1959). The simplicity of this intra-areal computation is contrasted by the more intricate transformations between brain areas. For example, Hubel and Wiesel demonstrated that neurons in V1 respond dynamically to the orientation of a bar of light (Hubel and Wiesel, 1959), and so their firing rate contains information about orientation. Knowledge about the neural activity reduces the uncertainty about the stimulus and vice-versa. This stimulus-response relationship is the prototypical sensory coding function called a tuning curve. Studies have shown that the amount of stimulus information represented in the tuning curve of a neuron is highest at the peak (the stimulus that causes the highest firing rate) and the peak of the first derivative (largest slope), depending on the noise present in the system (Butts and Goldman, 2006). Investigations into the neural code of the visual system therefore need to find stimulus dimensions that result in steep tuning curves and large peak responses from neurons. While bars of light are sufficient to investigate the neural code in early visual areas, these same stimuli are too impoverished to excite neurons in later areas. The low response rates are obstacles to describing the response function of neurons higher up in the visual hierarchy, like V4. The goal is still to find the domain in stimulus space that results in substantial, dynamic responses from a neuron, but the dimensionality of the stimulus space is astronomically larger. In the same manner that Hubel and Wiesel manipulated the width/orientation of a bar of light to excite V1 neurons, researchers have been developing novel strategies to make stimuli better suited for neurons in other visual areas (Hubel and Wiesel, 1959; Pasupathy and Connor, 2002; Ponce et al., 2019; Cowley et al., 2017). The resultant stimuli can be roughly categorized by the order of the underlying image statistics: first, second, and high-order. It is a precarious endeavor to match the complexity of a stimulus to the intricacy of a brain area's tuning preferences, but decades of work have guided a growing, cohesive understanding of stimulus-response functions. This work has lead to substantial improvements in sensory models of the brain as well as fundamental coding principles that generalize beyond the visual system.

1.2 The Neural Response Function

In its simplest form, one can consider visual neurons to be a biological implementation of some unknown function of a stimulus parameter (θ):

$$R_n = f(\vec{\theta}) \quad (1.1)$$

The neuron takes in inputs from dendrites, and produces an output in the form of action potentials (spikes). Computational neuroscientists then try to understand how the brain processes information by studying the input-output relationship of these functions. We begin with a single neuron, and a single variable input (θ), to build the foundations for high-dimensional cases later.

1.2.1 Tuning in One Dimension

The prototypical representation of a neurons input-output function is the tuning curve. Tuning curves are estimated by taking discrete samples along a single feature dimension (orientation for example) and plotting the neural responses against the stimulus parameter (Figure 1.2A). In this case, each dot represents one trial of the stimulus-response function, and by averaging across trials we get an estimate (dotted line) of how a neuron responds to that stimulus dimension (θ). This process provides the expected response of a neuron to each part of the stimulus space (Figure 1.2B) under the assumption that the tuning function is smooth where not sampled.

Importantly, this method also provides the reliability of the neural response function. While neurons are believed to respond in a deterministic manner, the presence of uncontrollable variables (biological noise: Faisal, Selen, and Wolpert, 2008, network oscillations: Fries, 2005, latent factors: Yu et al., 2009; Chen, Geisler, and Seidemann, 2006, etc.) makes most neurons appear to behave somewhat stochastically. That is, on any given trial, it may respond more or less than expected, which we call response variability. Tuning and variability both impact information signaling in unique ways, with different implications for the brain.

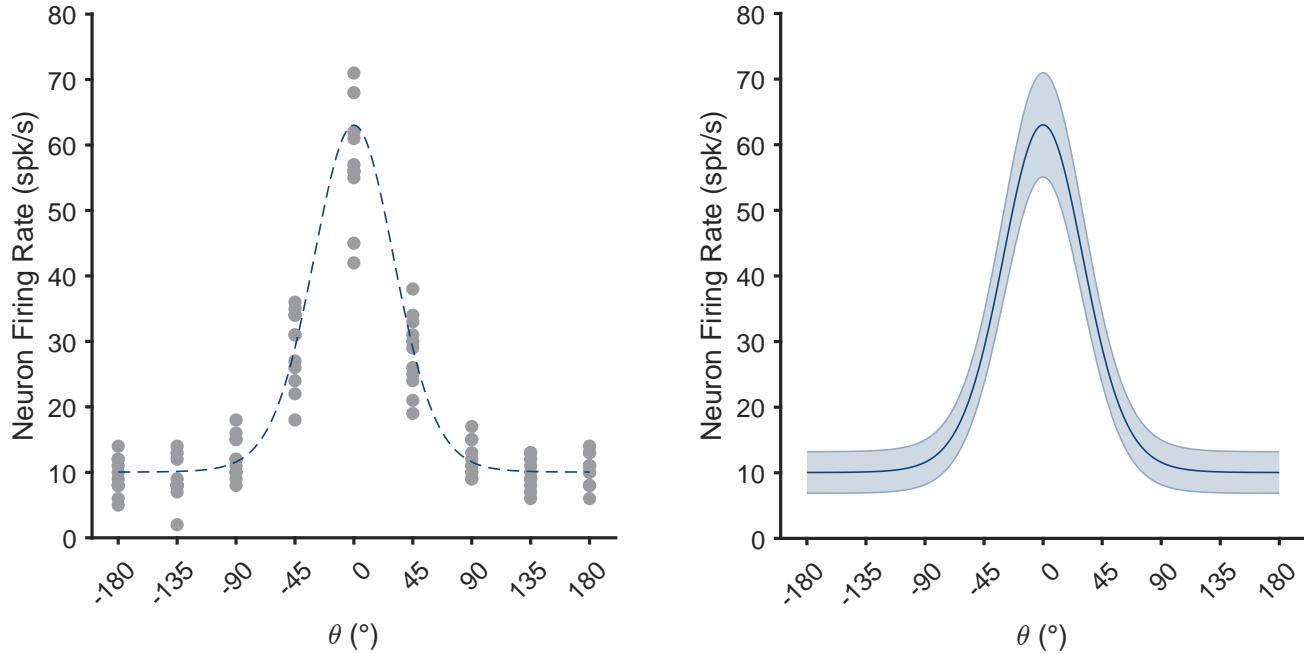


FIGURE 1.2: *Calculating a 1D Tuning Curve.* A.) The process for estimating a single neurons tuning curve to a single stimulus parameter θ . Discrete stimulus values were chosen (-180° by 45° to 180°) to span the stimulus space, and repeatedly presented (Gray dots). This provides the expected response to a given stimulus value (dotted line) for the neuron. B.) The estimated tuning and variability for the neuron in panel A.

A neurons tuning curve contains information. That is to say, knowledge of a neurons firing rate reduces uncertainty about the state of the world. Take the neuron from Figure 1.2 for example. If the neuron is firing at $40 \frac{\text{spk}}{\text{s}}$, it is pretty unlikely that the stimulus is either 0° or 180° , so it is conveying information about the stimulus to other neurons. The question then becomes, *how much* information a neuron is signaling. This quantity is directly related to three main components:

1. Response Range
2. Shape of tuning curve
3. Response Variability

The first point is addressed through natural upper and lower bounds. Neurons cannot have a negative firing rate, and neurons cannot fire arbitrarily fast due to biochemical constraints such as repolarization of chemical gradients and refractory periods (Kole

and Stuart, 2012). The final two points are intertwined due to quantization of response and precision of readout. Neurons cannot fire half a spike, so a single neuron cannot encode to arbitrary precision (e.g. 45° vs 45.000001° ; remember that firing rates are bounded), even with 0 response variability. This makes properly utilizing the available range essential, i.e. the shape of the tuning curve.

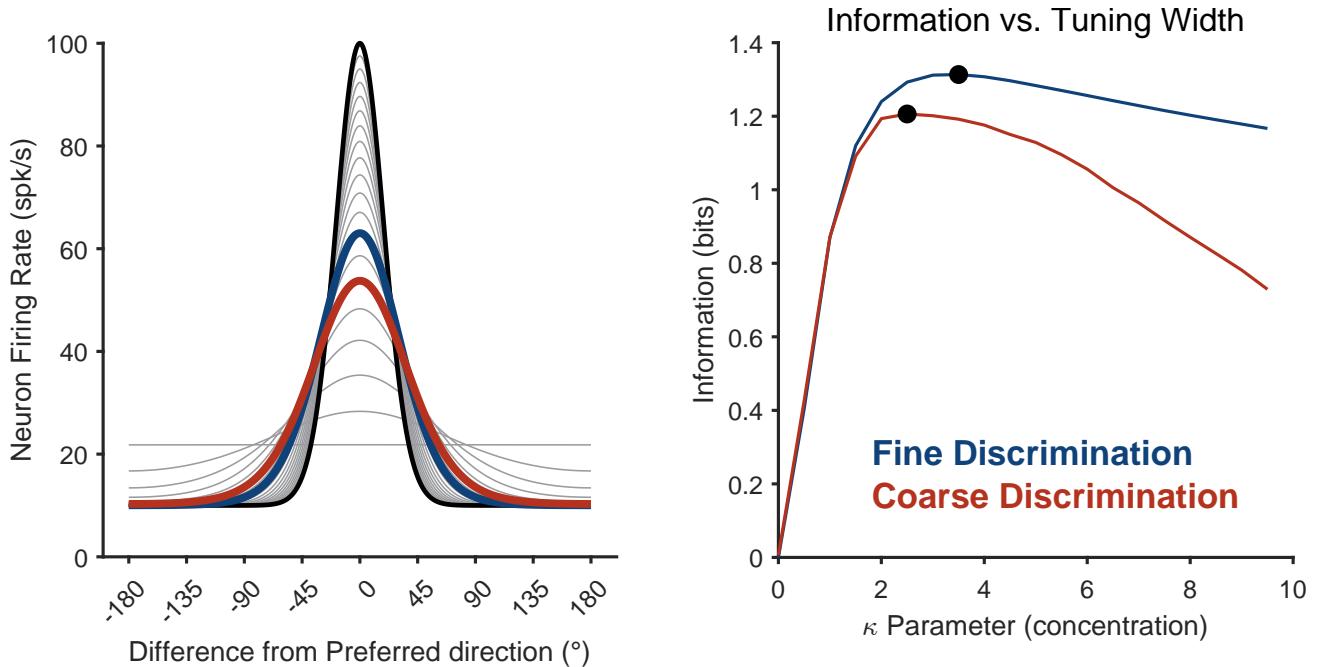


FIGURE 1.3: *Tuning curves and Information.* A.) Von Mises tuning functions with different concentration parameters (κ). Color indicates highest information tuning functions for fine (Blue) and coarse (red) discrimination. B.) Mutual information (Firing Rate and orientation) as a function of κ . Blue line depicts 1° discrimination, Red line depicts the more realistic situation of 8 stimulus classes (0° to 315° in 45° steps).

Previous work has demonstrated that the most informative parts of a tuning curve is where the slope is highest (Seriès, Latham, and Pouget, 2004). A small change in the stimulus results in the largest change in the response, making those sections the most sensitive to the stimulus dimension. Given this, one may fall into the trap of thinking larger slope means more information. Consider instead the black tuning curve in Figure 1.3A. If the task is to discriminate between $\theta = 45^\circ$ and $\theta = 90^\circ$, this is the worst stimulus-response function to use. In fact, I will show in chapter 2 that the shape of a tuning curve can vary with the demands on the system (Scott et al., 2023). In the case of

tuning width (κ parameter of a Von-Mises function), course discrimination (0° vs. 45°) necessitates a wider tuning function than fine discrimination (0° vs. 1° ; Figure 1.3B). This is because we are optimizing the tuning function for discrimination across the whole available range of θ , not just the high-slope regions.

1.2.2 nD Stimulus Spaces: Multiple Features

In the previous section I discussed the basics of tuning in 1 dimension. This explored how one neuron responds to changes in one stimulus parameter θ . This is an oversimplification of the true process, as no true stimulus can be described with only a single number. In this section I will expand the concept of a 1D tuning curve to an nD tuning surface spanning multiple feature dimensions.

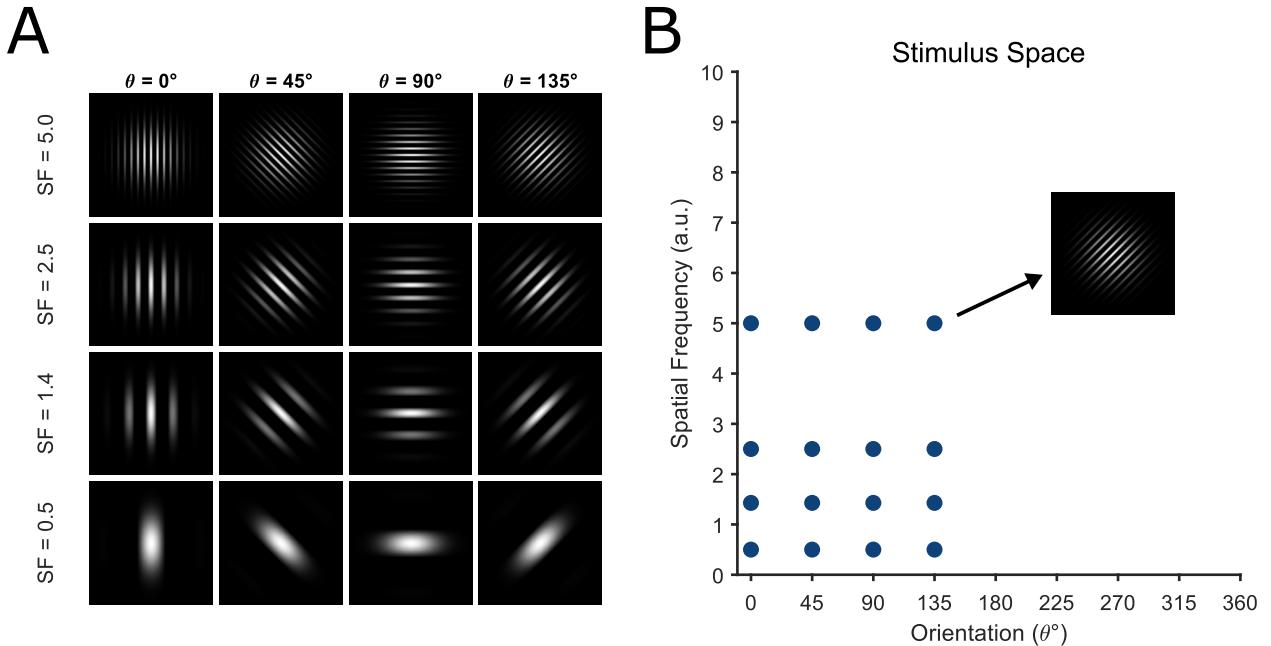


FIGURE 1.4: *Gabors as points in a 2D Stimulus Space.* A.) examples of Gabor stimuli with various θ and spatial frequency values. B.) Positions of the Gabors depicted in panel A within a 2-Dimensional stimulus space.

One of the most heavily-studied and simplest stimuli used in studies of visual neuroscience is the Gabor (Figure 1.4A). Gabors (or wavelets) are essentially oscillating light and dark patches at an orientation. These stimuli are often used in low-to-mid-level visual areas such as LGN (Mahon and Valois, 2001), V1 (Bredfeldt and Ringach, 2002;

Lennie, Krauskopf, and Sclar, 1990; Tanigawa, Lu, and Roe, 2010), V2 (Gegenfurtner, Kiper, and Suzanne, 1996; Liu et al., 2020) and Middle Temporal Cortex (MT) (Movshon and Newsome, 1996; Rust et al., 2006). Most such studies use the orientation dimension to produce tuning curves as described previously (Figure 1.2A), but many make necessary choices which limit them along non-studied feature dimensions (eg. spatial frequency, color, size, etc.). Let us begin with a 2D feature space by adding spatial frequency to orientation. Figure 1.4A demonstrates how independently varying the two feature dimensions results in a set of stimuli that a particular neuron may be tuned for. These are of course discrete steps representing two continuous feature dimensions (orientation and spatial frequency). This results in an important conceptualization, the stimulus space (1.4B). Here, every possible combination of orientation and spatial frequency are represented as points in a plane.

This is a natural progression from the 1-Dimensional orientation case shown before. All experiments investigating orientation tuning, necessarily limit stimuli along other dimensions. So while spatial frequency (for example) is kept constant in those experiments, neurons are likely still responding according to their n-dimensional tuning function. This results in a tuning surface (or tuning manifold in > 2 dimensions), such as in figure 1.5. Here we have the same model neuron from 1.2, where it's tuning before was just a slice (gray plane) through some 2D surface.

This idea of multidimensional tuning is more realistic, albeit more cumbersome. One of the biggest problems with studying neuron behavior this way is how intractable it is in 3+ dimensions. Take the 1D tuning curve in 1.2A. An experimenter must test the neural response function at several points to get an estimate of the tuning curve (in this case 8 orientations), because we do not know *a priori* which values along the stimulus dimension any one neuron is sensitive to. If, for example, we took the neuron from 1.2A and only sampled $\pm 90^\circ$, we would conclude that the neuron is not tuned to orientation. This is to say there is a minimum required sampling of each stimulus dimension we are interested in. Additionally, neurons often perform nonlinear operations (Carandini and Heeger, 2012) so we must sample each relevant dimension independently. This means

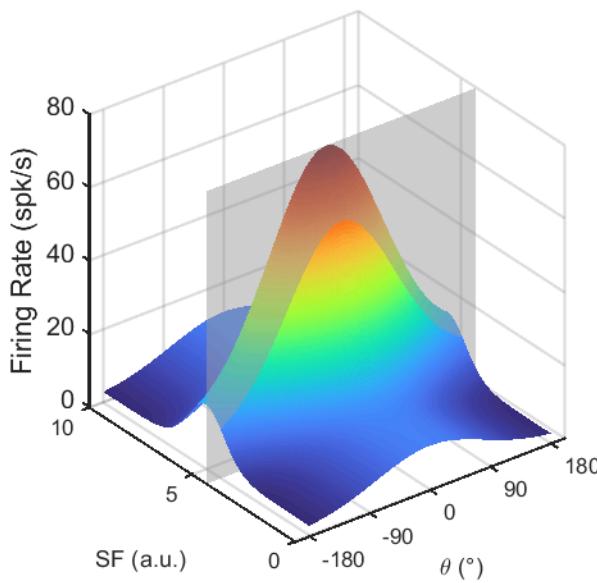


FIGURE 1.5: *Tuning along multiple Feature Dimensions.* A neurons tuning curve, in 2 dimensions becomes a tuning surface. Gray plane represents the spatial frequency value at which the tuning for θ matches figure 1.2.

to get a meaningful tuning curve for an N-dimensional feature space requires around 8^n individual stimuli (assuming 8 sampled values for each dimension) sampled with repetition (5-10 trials each). So for a Gabor stimulus with feature parameters 1.) Orientation, 2.) Spatial Frequency, 3.) Color, 4.) Contrast, and 5.) Size, with the minimum 5 trials each, an experimenter would have to show 163,840 stimuli in a single recording session. This is of course just for Gabors, one of the simplest visual stimuli. In more naturalistic settings, the number of feature dimensions could count in the hundreds. This problem motivates a more intelligent approach to searching for high dimensional feature tuning, which I will expand on in Chapter 3.

1.2.3 nD Response Spaces: Multiple Neurons

In the last section I built on the 1D neural response function by introducing the idea of multiple input dimensions: multiple features. In this section I will bring it back to 1 feature dimension, and discuss the implications of multiple response dimensions: multiple neurons.

Take again the example of neurons in V1 that are tuned for the orientation of a Gabor. This time we show a population of 8 neurons with equally-spaced preferences for θ (Figure 1.6A). We can assume for the moment that each neuron independently contributes information about the value of θ with its' firing rate. A brain area or decoder with knowledge of all 8 neurons firing rates can accurately estimate θ . For example, if neuron 1 has a low firing rate, and neuron 2 has a high firing rate, θ is most likely to be around 0° .

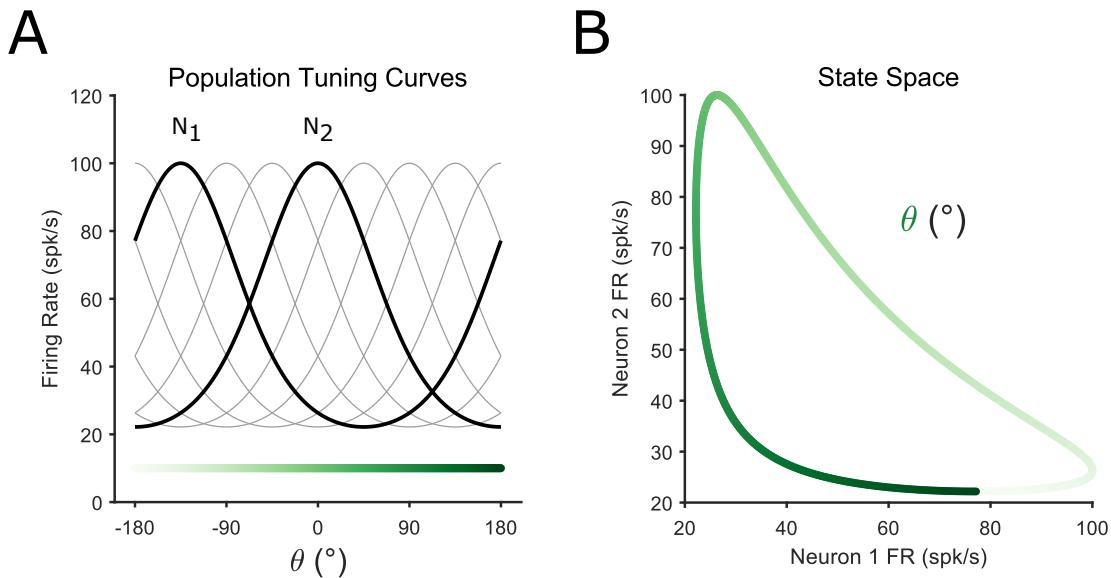


FIGURE 1.6: *Population Codes* A.) Tuning curves for 8 homogenous model neurons along the orientation feature dimension. Color at the bottom represents the values of θ . B.) A 2-Dimensional state space plotting the responses of neurons 1&2 from panel A. Color represents the value of θ that lead to each pair of neural responses.

This brings up another important concept that parallels the stimulus space from figure 1.4B, the Neural state space (Paninski et al., 2010; Cross et al., 2021). A state space is a representation of neural data in which each dimension (or axis) is the firing rate of one neuron (Figure 1.6B). We can then plot experimental variables (eg. stimulus features) according to the neural activity they elicited. In figure 1.6B, θ is represented by color, and plotted according to the relative responses of neurons 1 and 2 from panel A. This plot also represents the simplest case of a 1-Dimensional response manifold (Kriegeskorte and Wei, 2021; Chung, Lee, and Sompolinsky, 2018).

An important point to make here, as discussed in Kriegeskorte and Wei (2021), is that tuning curves (Figure 1.6A) determine the shape of the neural manifold (Figure 1.6B), but not the other way around. Different sets of tuning curves are capable of generating the same manifold geometry.

For example, if neuron 1&2 from figure 1.6A had the same tuning curve, adding neuron 2 provides no extra information. The fact that their preferences are offset is where the extra information comes from. This does not remain true when one considers response variability from before, as two identical neurons can provide multiple samples simultaneously (e.g. one neuron with two trials, or two identical neurons with one trial).

Likewise, another major issue when expanding intuitions about tuning beyond a single neuron, are shared neural variability. When neurons co-vary together within stimulus condition, we call these noise correlations (Cohen and Maunsell, 2009; Ruff and Cohen, 2016; Snyder et al., 2014; Moreno-Bote et al., 2014). Noise correlations provide their own set of drawbacks and complications which I briefly touch on in chapter 3.

1.2.4 Next Steps

The bulk of this thesis resides in the final section of this introduction, comparing high dimensional neural and feature spaces. As discussed previously, many studies have looked into tuning of neural populations to single feature variables, and more still investigated individual neural tuning in multiple feature dimensions. Quite few have attempted to investigate both simultaneously, due again to intractability of the problem. In chapter 2, I first discuss how even the simple case of neural tuning to one feature dimension is complicated by task demands and behavioral context. Then in chapter 3 I expand on neural optimization techniques to elucidate tuning for multiple features from a population of neurons. I will also expand on how to interpret and analyzed such rich data using techniques such as Canonical Correlations Analysis (CCA) and Distance Covariance Analysis (DCA) on simulated neural populations with known tuning parameters. Finally, in chapter 4, I apply all previous theory and techniques

in a series of experiments aimed at deriving simultaneous understanding about both high-dimensional spaces.

Chapter 2

Neuronal Modulation Depends on Context and Feature Tuning

2.1 Introduction

The MT area of monkey cerebral cortex is a well-studied extrastriate visual area known to have robust selectivity for visual motion features (Britten et al., 1992; Born and Bradley, 2005; Pasternak and Tadin, 2020). Motion signals represented by the activity of neurons in this area are likely to contribute to memory-guided comparisons of visual motion direction (Lui and Pasternak, 2011; Bisley, Zaksas, and Pasternak, 2001; Katz et al., 2016). Lesions of this area impair the ability to perform direction discrimination during motion comparison tasks (Rudolph and Pasternak, 1999; Bisley and Pasternak, 2000); whereas electrical stimulation of this area can bias the behavior of an animal as though a particular direction of motion was seen that was not present in the stimulus (Bisley, Zaksas, and Pasternak, 2001; Salzman et al., 1992; Salzman, Britten, and Newsome, 1990). These findings indicate a key role for MT in representing behaviorally relevant visual motion information and demonstrate that MT has an important role in task-related perceptual decisions. One open question, however, is how behavioral goals impact the neural activity in area MT during sensory (e.g., encoding/decoding of stimulus) and non-sensory (e.g., reward contingencies, decision making, etc.) periods of the task.

It is likely that during memory-guided comparisons of motion direction, area MT acts in concert with another brain area, the Lateral Prefrontal Cortex (LPFC). LPFC and

MT are reciprocally connected (Ungerleider and Desimone, 1986; Barbas, 1988; Petrides and Pandya, 2006), and both are involved in memory-guided motion comparison tasks (Zaksas and Pasternak, 2006). Neurons in LPFC contain direction information during motion discrimination tasks, but this selectivity is largely attenuated when animals do not have to make perceptual decisions, or when that information is not relevant to the task (Hussar and Pasternak, 2009; Hussar and Pasternak, 2012; Hussar and Pasternak, 2013). Previous research has also found reduced trial-to-trial variability in LPFC with increased task demands (Hussar and Pasternak, 2010), suggesting LPFC neurons may be recruited to support task performance when needed. Because of the interconnected structural and functional relationship between LPFC and MT, such changes in LPFC activity with task demands may be reflected in MT dynamics and altered processing of stimulus information, but this hypothesis remains untested.

To investigate this question, we measured how activity in MT was affected by different demands encountered during an active task that required monkeys to report memory-guided comparisons of motion direction, and a passive, identically structured task involving the same stimuli that did not require the animals to report their decisions. The active task consisted of two random dot motion stimuli separated by a delay, where monkeys were rewarded for correctly reporting whether or not the direction of motion in the second stimulus ("S2") matched the first ("S1"). During the passive task, signaled by a distinct fixation target, the trial structure and sensory conditions were the same, but required no perceptual decision, and the reward was delivered on each trial. Thus, despite identical sensory conditions, the task demands during active and passive tasks were different; the active task can be parsed as a sequence of: stimulus encoding (S1), memory maintenance (delay period), recall and comparison (S2), and finally perceptual report (S2/post-S2). We hypothesized that the sensory responses to motion in S1 and S2 during the active task would be reflected in sharper tuning, reduced trial-to-trial variability, and/or modulation gain, in line with prior findings about selective attention to motion direction (Ponce-Alvarez et al., 2013; Cohen and Newsome, 2008; Arandia-Romero et al., 2016).

We found that the effects of task demands depended on the amount of task-relevant stimulus information (e.g. direction selectivity) neurons conveyed. Specifically, we found three main modulation profiles for stimulus processing in the active task compared to the passive task: neurons that showed increased direction information (“IDI” neurons), neurons that showed decreased direction information (“DDI” neurons), and neurons with with largely the same amount of direction information during the two tasks (“SDI” neurons). These profiles likely reflect neural heterogeneity and should therefore be taken as soft boundaries more-so than distinct groupings. We found that during the active task, IDI neurons displayed tuning gain and a reduction in trial-to-trial variability, but showed similar comparison signals during the two tasks. In contrast, the task effects for DDI neurons were reflected in weaker tuning for direction, but not in trial-to-trial variability. The task effects on direction tuning and response variability of the SDI neurons were weak. Instead, these neurons, similarly to the DDI neurons, showed task dependent comparison signals. These results suggest a division of labor in area MT, grading sensory processing and cognitive effects of the task across subpopulations of neurons.

2.2 Methods

Resource Availability

Lead Contact

Requests for further information should be directed to the lead contact Adam Snyder (adam.snyder@rochester.edu).

Materials Availability

This study did not generate any new reagents.

Data and Code Availability

Data and code used for this paper are made available by request through the lead contact.

Subjects

Subjects used in this study were three adult male rhesus macaques 201, 202, and 317. All training, surgery, and experimental procedures were performed in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and were approved by the University of Rochester Committee for Animal Research. Surgery was performed using aseptic technique with isofluorane general anesthesia and perioperative opiate analgesics. An initial surgery was performed to implant a surgical steel head holder restraint embedded in bone cement secured to the cranium with ceramic bone screws. After behavioral training, a subsequent surgery was performed to implant a PEAK canula (19 mm diameter; Crist Instruments, Hagerstown, MD) over MT.

Data Acquisition

Recording chambers enclosed the craniotomies and contained 1mm-spaced CILUX grids (Crist Instruments). Thirty-two-channel S-probes (Plexon, Dallas Texas) were lowered through custom-made steel guide tubes, which were themselves inserted just low enough to penetrate the dura. We then drove the electrodes using a NAN electrode drive (NAN Instruments, Nof Hagalil, Israel). Recordings were done by coordinating a Plexon (Dallas, Texas) Multichannel Acquisition Processor and the data acquisition system TEMPO (Reflective Computing).

Receptive Field Mapping

Prior to each recording session, receptive fields were mapped in order to identify where the stimuli should be displayed. The initial manual phase consisted of moving a patch of dot motion around on the stimulus display with a computer mouse while listening to spiking activity via an audio monitor. Afterwards, we presented small dot motion

patches (1°) in different directions on the vertices of a grid spanning the likely RF area identified in the manual step. Stimuli for each session were placed such that they covered the aggregate receptive field area for the recording (Figure 2.5C).

Stimulus Presentation

The stimuli and the behavioral tasks were similar to those used in previous studies (Hussar and Pasternak, 2012; Wimmer et al., 2016). Stimuli were presented at 75 Hz monitor refresh rate on a 19-inch Iiyama Vision Master Pro-513 monitor with 1,152 by 870 pixel resolution. The stimuli consisted of random dots placed in a circular aperture fit to maximize the overlap of the recorded receptive fields. The dot density was $3 \frac{\text{dots}}{\text{deg}^2}$ and each dot was 0.03° Degrees of Visual Angle (DVA) in diameter with a luminance of $15 \frac{\text{cd}}{\text{m}^2}$ shown on a dark background of $0.1 \frac{\text{cd}}{\text{m}^2}$. The stimuli moved in one of 8 directions ($0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$), and their sizes ranged from 2° DVA to 15° DVA in diameter depending on eccentricity. Monkeys were required to maintain fixation within 1.5° of a central dot throughout the trial until the response interval, and eye position was monitored with an ISCAN infrared eye-tracking device.

Task Design

Trials started once subjects fixated on a central dot. They then had to maintain fixation for 1 second before S1 onset. Both S1 and S2 lasted 500ms. S1 was followed by a 1.5 second delay period, and then the second stimulus S2 was presented. Subjects then had to wait an additional 1 second after S2 offset before either a choice was required (active task), or they were simply rewarded (passive task). For the active task, the central fixation point was extinguished at the beginning of the response interval, and two identical choice targets were presented 5° DVA to the right and left of fixation, and the animal indicated its judgement as to whether S1 and S2 were the same direction or not by making a saccade to the right or left choice target, respectively. In most recording sessions, we collected data for both the active and the passive tasks. We always collected data for the active task before the passive task, because animals would have been unmotivated

to perform the more difficult active task if they had performed the easier passive task first.

Data Analysis

Spike sorting was done offline with Plexon sorter (Version 3.3.2) by using Principal Component Analysis (PCA) on the spike waveforms and manually clustering. A neuron was included for analyses if there were at least ten trials in each direction condition, and if the stimuli covered at least 30% of its 50% isointensity response field. A total of 1264 neurons across both tasks passed these criteria (201 : n = 263, 202 : n = 454, 317 : n = 547), with 928 neurons recorded in the active task and 336 neurons recorded in the passive task. Of those neurons, 254 were confidently matched across task and included for further analysis. Neurons were considered matched across task if they were recorded on the same channel and their waveforms were highly correlated ($r^2 > 0.99$). All subsequent analyses were done using custom routines for Matlab 2020a (The MathWorks, Inc, Natick, MA).

Mutual Information

Mutual information (MI) quantifies information (in bits) shared between two variables, i.e.: how much the uncertainty about one variable (motion direction condition, in our case) may be reduced with knowledge about the second variable (spike counts). MI provides a robust metric of direction selectivity that accounts for changes in tuning curves, variability, and absolute firing rates. We calculated MI using:

$$MI(x, y) = \sum_x \sum_y p(x, y) \log_2 \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (2.1)$$

Where $p(x, y)$ is the joint probability distribution between number of spikes and direction condition, and $p(x)p(y)$ is the product of marginal probability distributions. This provides how many bits of information a neuron's firing rate contains about direction and vice versa. For $p(x)$, we used the probability of observing a spike count in a

bin of width Δx sp/s, where the bin width was determined by “Scott’s rule” (SCOTT, 1979):

$$\Delta x = \frac{3.5n^{-\frac{1}{3}}}{n} \sum_i^n (x_i - \bar{x}) \quad (2.2)$$

We calculated the mutual information between spike count and S1 direction for each non-overlapping 50ms bin from the start of each trial up to the end of the delay, and then the MI between spike count and S2 direction through the end of the trial. In order to correct for spurious effects that come from differences in trial-counts, we baseline-corrected MI by calculating it many times with shuffled conditions, and subtracted this out (Hatsopoulos et al., 1998).

$$I(x; y) = MI(X; Y) - MI_{\text{shuffled}}(X; Y) \quad (2.3)$$

Classification of neurons into information Enhanced, Suppressed, and Consistent subgroups

We calculated baseline-corrected mutual information (MI) between the firing rate of a neuron and motion direction shown for each time bin. From trial start up to S2 onset MI was calculated according to S1 direction. From S2 onset through the end of the trial, it was calculated between firing rate and S2 direction.

We then took the MI time series for a neuron in active, minus passive, in order to get a task-effect of motion information over time. Note that this was only possible for neurons that were successfully matched across tasks (n=254). The clustering algorithm contains 4 distinct steps.

1. Calculate the Pearson distance (i.e., one minus the Pearson correlation divided by 2) between each pair of neurons’ task-effect time series. This was done on a window during S1. We did this in order to disentangle the sensory and cognitive components, since clustering on S2 would create a circularity for later analyses.

2. Reduce the dimensionality with multidimensional scaling to five dimensions (to mitigate the “curse of dimensionality” for clustering with high-dimensional data).
3. Model selection to pick the appropriate number of clusters. This consisted of fitting multiple Gaussian mixture models to the data, each with different numbers of components (using the MATLAB ‘gmdistribution’ class). We then selected the appropriate number of clusters by selecting the model with the lowest Akaike Information Criterion (AIC), which was three clusters.
4. K-means clustering using the appropriate number of components (3) determined in step 3.

In this way, neurons were partitioned into groups based on how motion information signaling changed across the active and passive tasks. This allows us to then see how other task-related effects may depend on the type of information modulation a neuron receives.

Resampling Procedure for Clustering Robustness

In order to confirm the validity of our clustering approach, we used a resampling procedure with 100 iterations. We calculated mutual information for each neuron, using random samples of 50% of trials (on each iteration, we used the remaining half of trials to analyze changes in tuning curves; see section [Tuning Curves](#)). We then ran the clustering algorithm on each iteration, and calculated the most parsimonious model each time using AIC. The majority of iterations resulted in 3 clusters of enhanced, suppressed, and consistent direction information, while no iterations resulted in a single cluster being the best model (Supplemental Figure 2.1).

Entropy Factor

Entropy Factor (HF) quantifies the entropy of a neuron’s spike count (across trials) relative to a Poisson process with the same rate parameter, λ (Rajdl, Lansky, and Kostal, 2017). It is similar to Fano factor, in that a value of 1 indicates equivalence to a Poisson

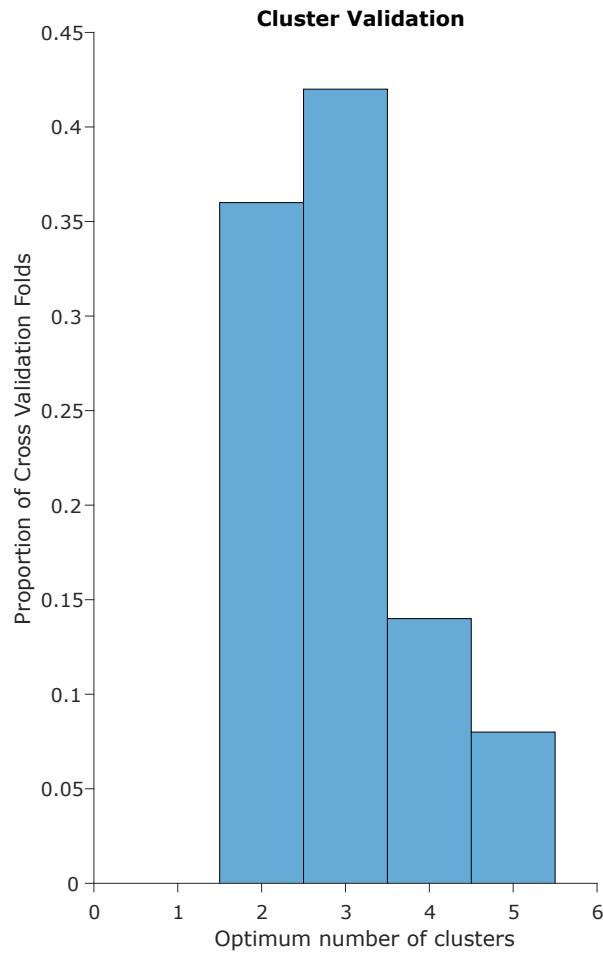


FIGURE 2.1: Cross validation of clustering algorithm. We took 50% of trials for each neuron, calculated MI, ran the clustering algorithm, and minimized AIC to pick the optimum number of clusters. After repeating this process 100 times, we got a distribution of models that supports 3 clusters as optimal for our data.

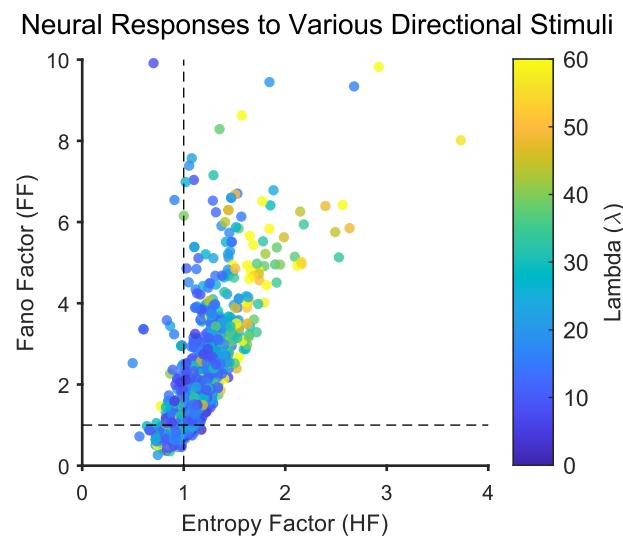


FIGURE 2.2: Relationship between Entropy Factor (HF) and Fano Factor (FF). We took the neural responses for our 8 directions of motion (S1), for each neuron in our sample, and plotted FF vs HF. Dotted lines are Poisson processes. Neural responses with a larger lambda parameter (average firing rate) result in a roughly linear relationship between the measures. When neural responses are low, FF overestimates the variability (darker blue dots are highly super-Poisson according to FF, but roughly Poisson according to HF)

process, while values greater or less than 1 indicate distributions that are supra-poisson or sub-poisson, respectively. However, unlike Fano factor, which compares the variance of a distribution of spike counts to its mean, entropy factor takes into account the entire distribution of spike counts. This provides a more reliable metric than Fano Factor, particularly when there are a low number of observations (Supplemental Figure 2.2). For each 50ms time bin, t , we calculated the entropy $H_{Observed}(\vec{x}_t)$ of the distribution of spike counts (\vec{x}_t) across trials with the same stimulus direction and divided it by the entropy of a simulated Poisson distribution with the same intensity and number of trials.

$$HF(\vec{x}_t) = \frac{H_{Observed}(\vec{x}_t)}{H_{Poisson}(\hat{x})} \quad (2.4)$$

Here, $H_{Observed}(\vec{x}_t)$ is the entropy of the distribution of spike counts across trials $i \in [1, n]$ for a neuron at time t , defined by:

$$H_{Observed}(\vec{x}_t) = - \sum_{i=1}^n p(x_t^i) \log(p(x_t^i)) \quad (2.5)$$

where $H_{Poisson}(\hat{x})$ is the entropy of a simulated Poisson process with rate parameter λ :

$$\lambda = \bar{x}_t \quad (2.6)$$

The entropy of a Poisson process was calculated by doing 200 simulations of spike counts for n trials pulled from a Poisson distribution with rate parameter λ_t , and the same unbiased bins used for the data (eq. 2.2). By then averaging the observed entropies of the simulations, we get an estimate of how variable, on average, we would expect the matched Poisson process to be. This normalization accounts for bias in sample entropy due to differences in firing rate or trial count.

Tuning Curves

Tuning curves were calculated as the average firing rate during the stimulus window for each of the 8 directions. For analyzing tuning curves, we used the half of trials that

was not used for clustering based on mutual information from each resampling iteration (see section [Resampling Procedure for Clustering Robustness](#)), and averaged results across iterations. For testing the effect of stimulus (S1 vs. S2) on tuning and HF, we used a three-way ANOVA with factors of ‘direction’ (eight levels corresponding to each motion direction), ‘stimulus’ (two levels: S1 vs. S2) and ‘task’ (two levels: active vs. passive).

To test whether changes in tuning functions could be explained through changes in amplitude, width, preferred direction, or baseline shifts, we fit Gaussian functions to the responses of individual neurons using the trust-regions algorithm implemented by the Matlab ‘`cfit`’ object class. The fitted function was

$$f(\theta) = ae^{-\frac{(\theta-b)^2}{2c^2}} + d \quad (2.7)$$

where a is the amplitude (in spikes/s), b is preferred direction (in degrees), c is the width parameter in degrees —related to the full width at half height (FWHH) by $FWHH = 2\sqrt{2\ln 2}c$, d is the baseline response (in spikes/s), and θ is the motion direction of a stimulus. Because motion direction has a periodic domain ($[0, 360]$ degrees) whereas the Gaussian function does not, we “centered” the responses of each neuron so its preferred direction of motion was around zero prior to fitting, then we applied the inverse shift to the estimated b parameter. We imposed the following constraints: $a \in [0, 1.2 \cdot \max(\vec{x})]$, $b \in [-45, 45]$, $c \in [1, 90]$, $d \in [0, \text{median}(\vec{x})]$, where \vec{x} is the set of observed responses (average firing rate during stimulus window) to be fit for a given neuron. The starting point for fitting was $a = \text{range}(\vec{x})$, $b = 0$, $c = 40$, $d = \min(\vec{x})$.

To test for differences in fitted parameters between tasks, we performed paired t-tests with $\alpha = 0.05$, Bonferroni-corrected for the three groups of neurons tested (i.e., DDI, IDI, and SDI). We report the results of this analysis using all neurons, but we also tested subsamples of neurons restricted to have R^2 values for their fits greater than 0.2 (68% of neurons) or 0.5 (34% of neurons) and found that the pattern of results was similar to the results when testing the full sample, and the conclusions were not affected. Note that we used individual trials rather than within-condition averages for fitting

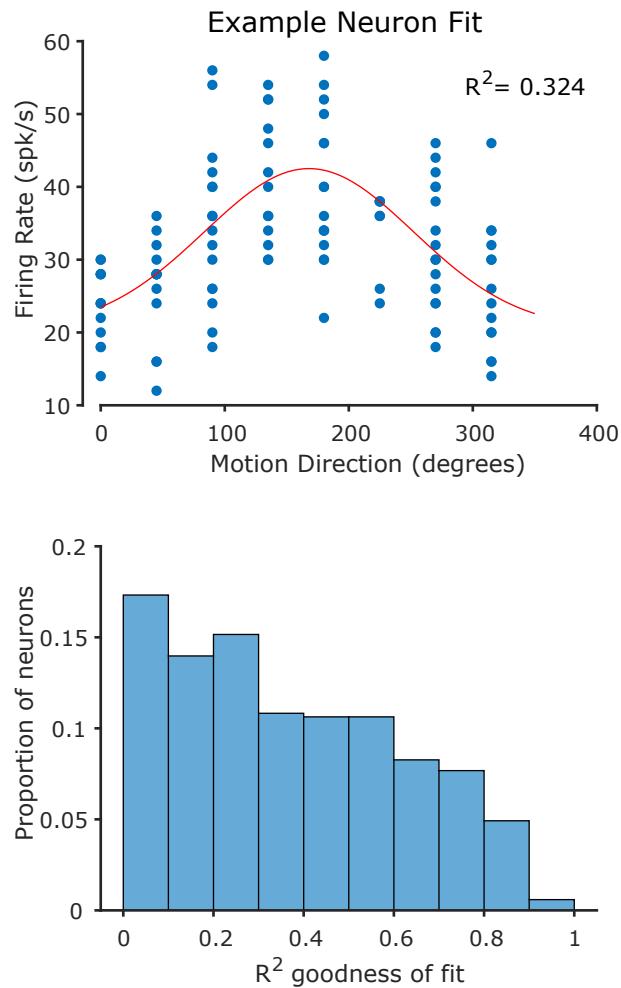


FIGURE 2.3: Example and overview of tuning goodness of fit (R^2). A Gaussian fit to one example neuron during one bootstrap fold. The full distribution of R^2 for our sample of neurons.

tuning functions, so some of the variance unexplained by the fitted model includes trial-to-trial variability.

For reports directly comparing MI to tuning parameters, we used the above procedure on the half of trials not used to calculate MI across all 100 cross-validation folds. This allowed us to avoid any circularity in our comparison of tuning and mutual information.

Comparison Effects

Comparison effects (CE) were calculated as the area under the curve (AUC) of the receiver operating characteristic. For each neuron, trials were separated by direction condition (during S2), and then by trial type (same direction or different direction trials depending on whether S1 matched S2). CE was calculated on firing rates in 100 ms bins, sliding by 10 ms steps, as the AUC between ‘same’ trials and ‘different’ trials within direction. AUC values were then averaged across their preferred direction and $\pm 45^\circ$ off preferred within each time bin, in order to reduce noise in the estimate (see Supplemental Figure 2.4 for a comparison of CEs to preferred versus anti-preferred directions). This resulted in a time course of AUC values for each neuron where values greater than 0.5 indicates a neuron that responds more robustly to identical visual stimuli on trials where the directions of S1 and S2 matched ($S > D$ neurons), while values less than 0.5 indicated the opposite ($D > S$ neurons).

Statistical Significance of CE

For testing task effects on the population average CE (Figure 2.9), we first averaged for each neuron within each of two time windows of interest: (1) the period when S2 was presented, and (2) the period from S2 offset to the onset of the choice targets. We then tested for a difference between tasks using a paired t-test, with each neuron as a degree of freedom ($\alpha = 0.01$, two-tailed).

Statistical significance of CE for individual neurons (Figure ??) was determined using a bootstrapping procedure. Trial labels ('same' or 'different' trial) were randomly

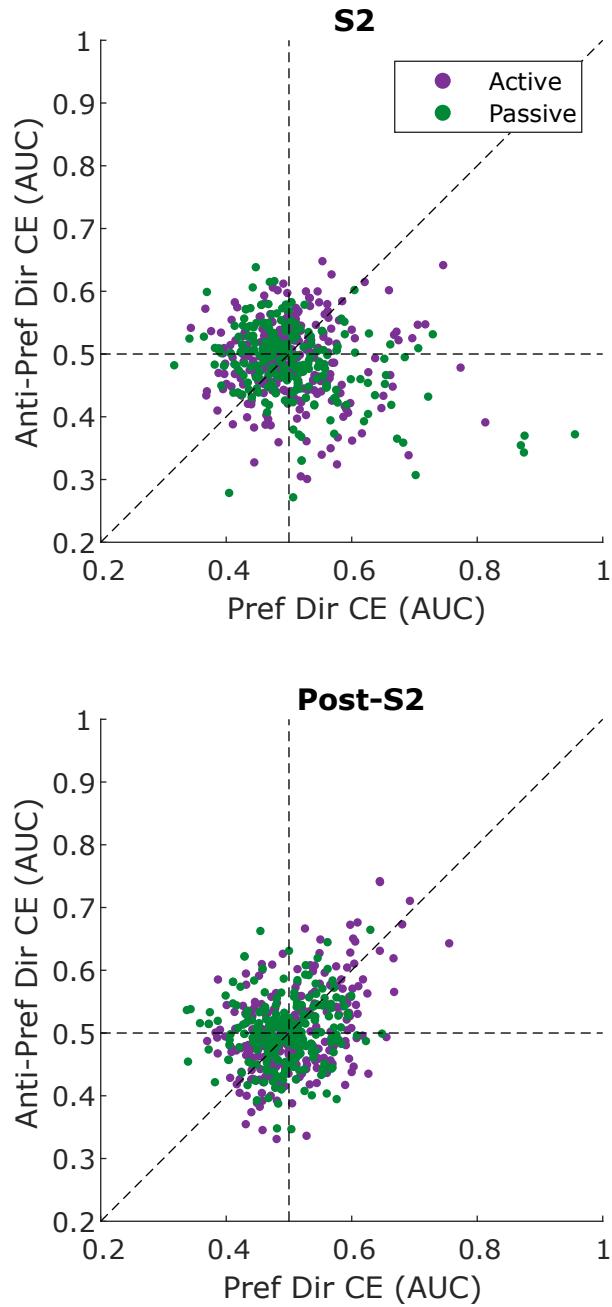


FIGURE 2.4: Average Comparison Effect (CE) during the two windows of interest for neurons' preferred direction and 180° off preferred (anti-preferred).

permuted, and AUC values were calculated 5000 times per neuron following the procedure described above (section *Comparison Effects*), and then averaged within each time period of interest (S2, and post-S2). A neuron was considered significantly CE signaling (Figure ??) if its AUC was significantly different from the bootstrap distribution ($\alpha = 0.01$, two-tailed). Observed CE values greater than the bootstrap distribution indicate same-preferring ($S > D$) and lesser than the bootstrap distribution indicate different-preferring ($D > S$).

2.3 Results

Three adult male Rhesus macaques (*Macaca mulatta*) performed delayed visual motion comparison tasks while we recorded neural activity in MT (Figure 2.5A). In the active task, monkeys indicated whether a second stimulus ("S2") moved in the same or an orthogonal direction to a first stimulus ("S1") by making a saccade to one of two choice dots. All three monkeys performed above chance in the active task (201, 10 sessions: $77.79\% \pm 1.08\%$ accuracy, 202, 28 sessions: $81.63\% \pm 0.95\%$, 317, 16 sessions: $85.77\% \pm 1.36\%$; *mean* \pm *Std*). The passive task, except for a different fixation point shape, was identical up to the choice window, at which time no choice target stimuli were shown, and no further action was required to receive reward (Figure 2.5A). Note that during both tasks, the timing of salient events, i.e., the onset time of S1 and S2 during both tasks was highly predictable, since on each trial the periods preceding S1 (1 s fixation period) and S2 (1.5 s delay) were always the same. Similarly, on each trial there was always a 1 s period separating the offset of S2 and the decision/reward (active task) or the reward (passive task). We were interested in how neurons' signaling for motion direction changed from the passive to the active task, and how that related to changes in neurons' signaling for cognitive processes. The cognitive signals we considered were anticipatory ramping of firing rates and comparison effects. We were unable to consider another cognitive signal of potential interest, choice probability (Britten et al., 1996), because error trials were relatively infrequent in the active task, and no choices were made in the passive task.

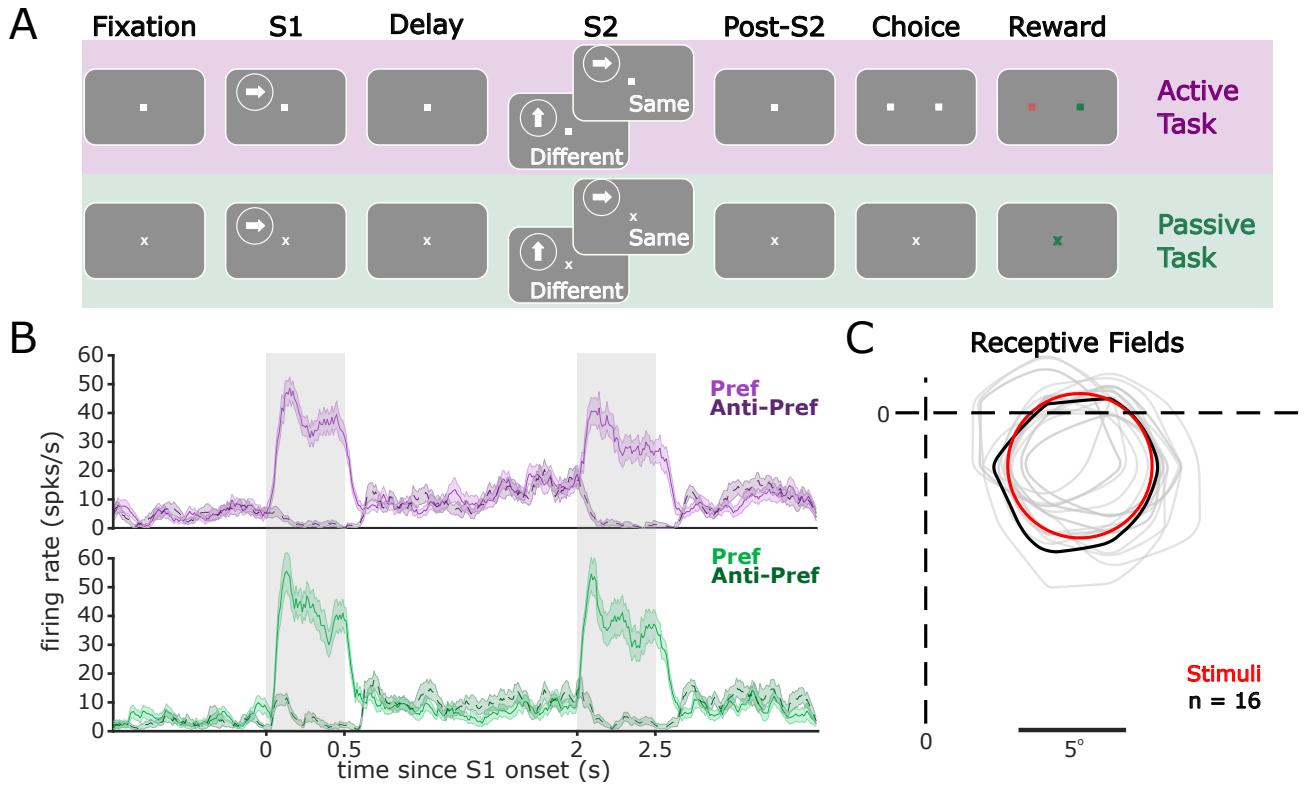


FIGURE 2.5 *Task design and task modulation of MT firing rates*. A.) Trials consisted of a pre-stimulus 1 second fixation period, a first stimulus (S1), followed by a 1.5 second delay, then a second stimulus (S2) and a post-S2 fixation period. After the 1 second post-S2 fixation period subjects were either rewarded (passive task) or had to report a decision with a saccade (active task). During the active task, correct choices were rewarded with juice while incorrect choices resulted in a 3 s time-out signaled by a tone and no reward. During S1, stimuli moved in one of 8 directions ($0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$), followed by the S2 that moved either in the same direction, or 90° off of S1 (rotated left or right). B.) PSTHs of an example neuron recorded in both tasks. Solid line represents the neuron's response to its preferred direction while dashed lines indicate 180° away from preferred ("anti-preferred"). C.) Receptive fields (grey) of each simultaneously recorded neuron from one example session, along with the location and size of the stimuli (red). Black curve corresponds to example neuron in B. Contours represent iso-intensity contours at 50% of the peak response.

Neurons were recorded between 2° and 25° eccentricity, with receptive field sizes proportional to distance from the fovea. A linear fit to receptive field size vs. eccentricity resulted in an intercept of 2.37° and a slope of 0.44, comparable with previous results from this lab (Bisley et al., 2004) and other studies in MT (Albright and Desimone, 1987). Neurons were included for analysis if: 1.) Their 50% isointensity response field (see [Methods](#)) was at least 30% covered by the stimuli, 2.) the subject did at least 10 trials in each direction condition, and 3.) they were confidently recorded across both tasks (see [Methods](#)). This resulted in 254 well-defined neurons included for analysis.

We were interested in the impact of cognitive demands on MT activity. During the active task, such demands varied over the course of a trial, including periods where stimuli were present (i.e., during S1 and S2) and also periods where stimuli were absent (i.e., pre-S1, the delay, and post-S2). For example, during the 1 second fixation period prior to S1, the animal is likely to prepare cognitive resources in anticipation of the stimulus, the appearance of which was highly predictable. Then, during S1, the animal must process and store task-relevant information (motion direction). During the delay, this information must then be maintained in working memory and the appearance of S2 may be anticipated. During S2, task-relevant information must again be encoded, the stored information about S1 retrieved, and a comparison judgement made. These dynamic task demands may result in varying effects on MT activity across these different stages. We first describe the effect of task demands on signaling motion direction by MT activity during S1 and S2. Second, we will describe the effects of task demands on MT activity when stimuli were absent. Third, we will describe effects of task demands on MT activity during the comparison stage of the task, i.e., during the S2 and the post-S2 period preceding the decision/reward (active task) and the reward (passive task).

Task effects in MT depend on neural tuning

We hypothesized that the accuracy of encoding of motion direction would improve when that information was task-relevant (i.e., in the active task compared to the passive task). To test this, we measured the mutual information (MI) between neural firing

rates and the motion direction of S1 and S2. MI quantifies how much information about one signal is provided by another signal (and vice-versa) and takes into account both the tuning of neural responses with respect to a stimulus as well as trial-to-trial response variability (Hatsopoulos et al., 1998; Quiroga and Panzeri, 2009).

Task-dependent selectivity would be consistent with the role of MT as a mid-level sensory area. That is, switching from passive viewing to a task requiring judgment of motion direction might have a particularly strong effect on neurons strongly signaling motion direction compared to neurons that did not signal that information so strongly. To test this hypothesis, we calculated the task effect of motion information for each neuron by subtracting the MI during the passive task from the active task (Figure 2.6 for examples). This approach provided a metric which quantifies, at each time point, how each neuron signals motion direction information during the two tasks (Figure 2.6D). We then calculated the Pearson distance between each neuron, which provided a dissimilarity matrix for our population of neurons. Neurons close together in this space have similar task effects, whereas neurons far apart are different (Figure 2.6E). We then reduced the dimensionality to D=5 with multidimensional scaling to minimize distortion of the relative Pearson distances and clustered neurons using a Gaussian Mixture model. We fit models with 1 to 5 clusters using 50% trial hold-out cross-validation. We then took the median Akaike Information Criterion (AIC) across folds and found that three clusters of task effects represent the most parsimonious model (1 Cluster: $AIC = 6.655 \cdot 10^3$, 2 Clusters: $AIC = 6.626 \cdot 10^3$, 3 Clusters: $AIC = 6.624 \cdot 10^3$, 4 Clusters: $AIC = 6.630 \cdot 10^3$, 5 Clusters: $AIC = 6.638 \cdot 10^3$; smaller values indicated better models). Importantly, the model with a single cluster was never the model with the best AIC on any cross-validation fold, suggesting that the presence of cluster structure in the data is robust. Nevertheless, we caution that cluster identities are best conceived as “soft” boundaries, and interpret our results as reflecting heterogenous variation across the population. We label the task effects corresponding to these clusters as: Increased Direction Information (IDI), Decreased Direction Information (DDI), and Similar Direction Information (SDI) groups (Figure 2.7A&B), referring to the average task effect on

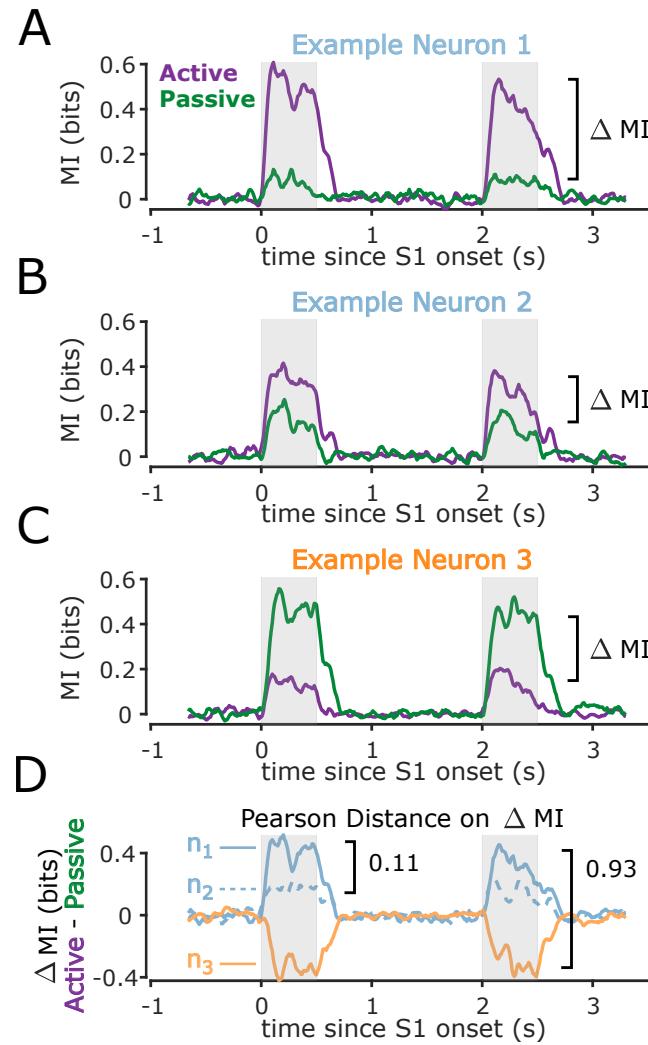


FIGURE 2.6 Example Mutual Information Time-Courses. A.) Mutual information between motion direction and firing rate for an example neuron in the active and passive tasks. B-C.) Same as A for two more example neurons. D.) Task effect on motion information (i.e., MI during active minus MI during passive) during the course of the trial for the three example neurons. Pearson Distance was calculated on these time courses for each pair of neurons. Distance values close to 0 mean similar task effects, whereas distance values close to 1 indicate opposite task effects. We then used the set of pair-wise Pearson distances to cluster neurons based on their task effects for MI.

direction information for neurons within each cluster.

Remarkably, during the active task a substantial group of neurons displayed a marked reduction in MI (DDI neurons), indicative of a detriment in direction coding, while the IDI neurons followed the expected pattern of improved direction signaling. These results suggest that heightened task demands result in task-relevant information becoming disproportionately represented in subgroups of neurons.

Differential Effect of Task Demands on Tuning for Motion Direction

We found that mutual information between neuron firing rates and motion direction changed when motion direction was task-relevant (active task) compared to when it was not (passive task). Since both the tuning for motion direction and trial-to-trial variability contribute to MI, the observed changes in MI could be consistent with a change in tuning, a change in trial-to-trial variability, or a mixture of both. To better characterize the nature of these effects, we separately analyzed tuning and trial-to-trial variability for IDI, DDI, and SDI neuron groups.

We tested for differences in responses to motion for each group by calculating the average firing rate for each neuron, and compared effects of task, direction condition, and stimulus epoch with 3-way ANOVAs. The main effect of ‘stimulus epoch’ (S1 vs S2) was negligible for each group (IDI: $p = 0.238$; DDI: $p = 0.187$; SDI: $p = 0.531$), so for visualization purposes we plotted S1, which unlike S2 does not involve concurrent comparison processes (Figure 2.8A-C). As expected, all three groups were significantly tuned to motion direction (3-way ANOVA: IDI: $F = 96.09, p = 5.47 \cdot 10^{-126}$; DDI: $F = 107.77, p = 1.49 \cdot 10^{-140}$; SDI: $F = 49.72, p = 2.70 \cdot 10^{-66}$). We found main effects of task in all three groups (IDI: $F = 15.27, p = 9.54 \cdot 10^{-5}$; DDI: $F = 18.28, p = 1.97 \cdot 10^{-5}$; SDI: $F = 7.96, p = 0.005$), indicating slight changes in overall firing rates with task condition (IDI, active: 27.36 ± 0.52 sp/s mean \pm SEM, passive: 24.96 ± 0.44 sp/s; DDI, active: 25.47 ± 0.41 sp/s, passive: 27.82 ± 0.46 sp/s; SDI, active: 23.60 ± 0.43 sp/s, passive: 22.12 ± 0.37 sp/s). We also found interaction effects between task and stimulus direction for IDI and DDI neurons, indicative of task-dependence changes in direction tuning.

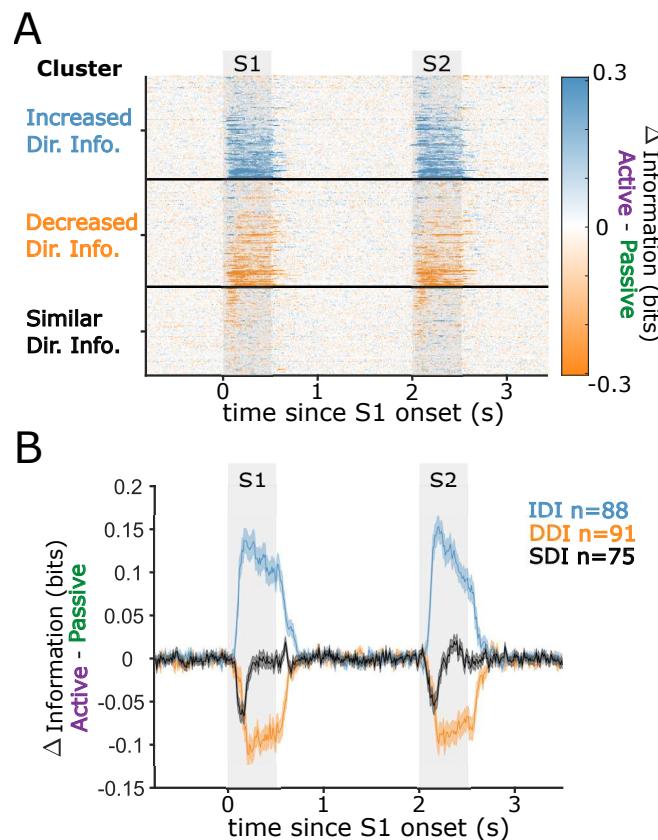


FIGURE 2.7 Motion Information Clustering. A.) Time series for active minus passive motion information (MI) for each neuron. Rows correspond to neurons and color represents the difference in information. See [Methods](#) for the clustering algorithm. B.) Within-cluster averages from panel A.

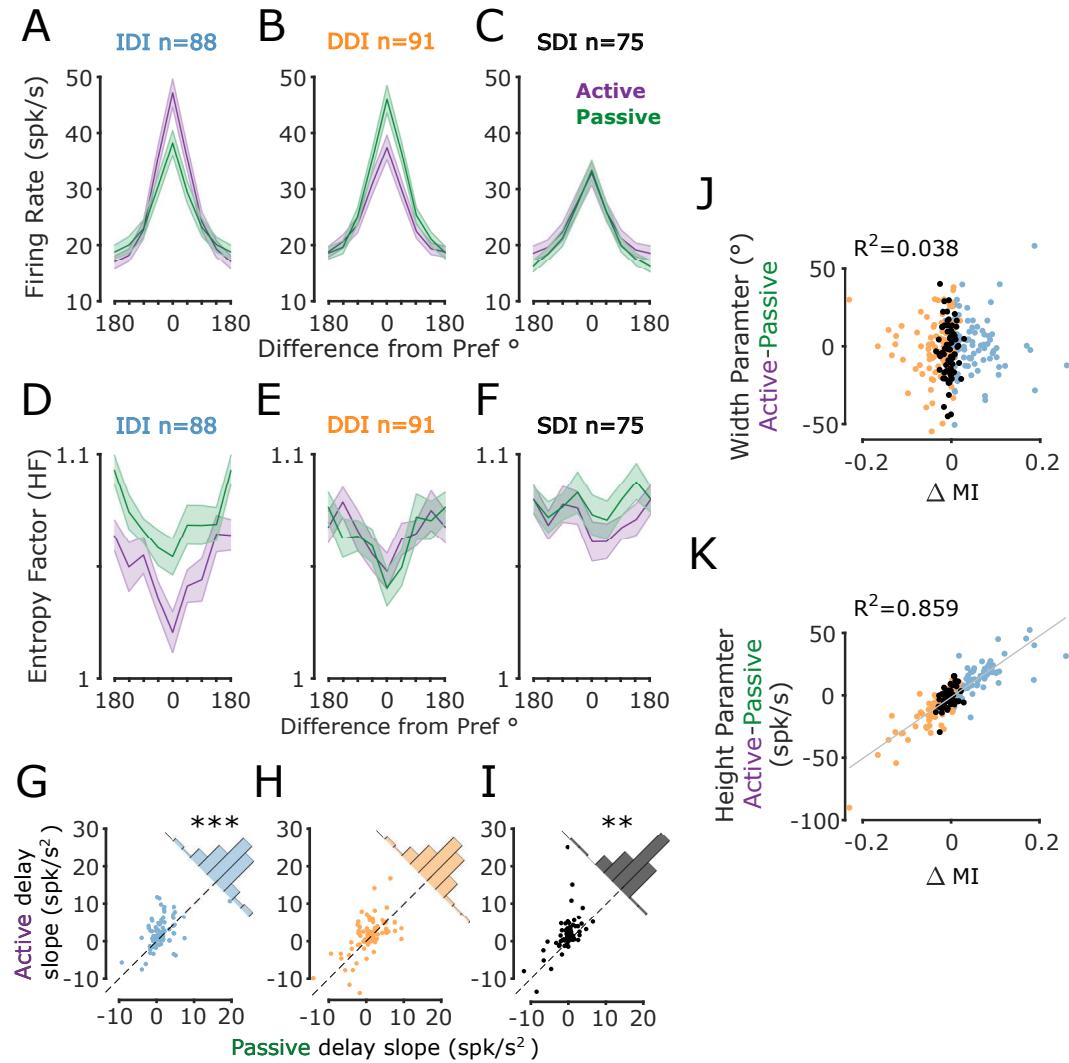


FIGURE 2.8 *Effects of task demands differed depending on how informative a neuron was about task-relevant information.* A.) Average raw tuning curves for IDI neurons during the active (purple) and passive (green) tasks. Both curves contain the same neurons. Error bars are $\pm SEM$ across neurons. B-C.) Same as A for DDI and SDI groups respectively. D.) Average trial-to-trial variability during S1 quantified with entropy factor (HF) (See Methods). E-F.) Same as panel D for DDI neurons and SDI groups respectively. G.) Slope of a linear fit to IDI neurons' firing rates vs time during a 1s interval before S2 onset in the Active vs Passive tasks. There was a significantly higher ramping during the active task (**: $p = 5.66 \cdot 10^{-5}$, t-test). H-I.) Same as panel G for DDI (ns) and SDI (**: $p = 0.0098$) groups respectively. J.) Task effect of the width parameter of gaussian fits vs ΔMI during S1. For each cross validation fold, we fit a gaussian to the 50% of trials not used for the calculation of MI (i.e. the X and Y axes use different trials). We then averaged across folds, and subtracted across task to get a task effect of tuning width. We found no significant relationship between tuning width and ΔMI ($R^2 = 0.038$, $p = 0.552$). K.) Same as J but for the height parameter of the gaussian fits ($R^2 = 0.859$, $p = 3.505 \cdot 10^{-75}$).

Firing rates around the preferred direction became enhanced in IDI neurons in the active task ($F = 5.25, p = 5.76 \cdot 10^{-6}$: Figure 2.8A,K). In contrast, in the DDI group, firing rates around the preferred direction were reduced in the active task compared to the passive condition ($F = 4.24, p = 1.16 \cdot 10^{-4}$ Figure 2.8B,K). The SDI group had no such interaction effect ($F = 0.24, p = 0.974$). Taken together, these results imply that direction tuning differs between tasks for IDI and DDI neurons, while task demands do not alter direction tuning of SDI groups.

Next, we sought to quantify how responses changed with task condition in terms of their direction tuning curves. Direction tuning curves can be conceived as bell-shaped functions with four parameters: amplitude, width, shift (preferred direction) and offset (baseline firing rate). To test whether the changes in tuning that we observed could be attributed to changes in one or more of these parameters, we fit Gaussian functions to the responses of each neuron as a function of stimulus direction, and then tested for an effect of task ('active' vs. 'passive') for each parameter with paired t-tests for each group of neurons (IDI, DDI, and SDI). For the IDI neurons, we found the amplitude parameter was significantly greater in the active task than in the passive task (change in amplitude for S1: 11.66 ± 1.34 sp/s [mean \pm SEM], $86.09 \pm 13.54\%$, $t_{87} = -8.73, p = 2 \cdot 10^{-13}$. For S2: 11.01 ± 1.55 sp/s, $98.11 \pm 17.03\%$, $t_{86} = -7.19, p = 2 \cdot 10^{-10}$). All other parameters were not significantly different (all p's > 0.07). In contrast, the DDI neurons showed lesser tuning curve amplitudes in the active task than in the passive task (change in amplitude for S1: -11.52 ± 1.66 sp/s, $28.46 \pm 4.05\%$, $t_{89} = 6.99, p = 5 \cdot 10^{-10}$. For S2: -9.95 ± 1.44 sp/s, $28.58 \pm 4.01\%$, $t_{89} = 6.90, p = 7 \cdot 10^{-10}$). All other parameters, including tuning width, were not significantly different (all p's > 0.08). The SDI group showed slightly greater offset (baseline) values in the active task than in the passive task during S2 ($p = 0.03$), consistent with the main effect of 'task' for the ANOVA results reported above, but this was not robust to correction for multiple comparisons for the three groups of neurons tested. We further tested the relationship between tuning and ΔMI by considering the full distribution, void of any assumptions about clusters. We again calculated tuning curves for each neuron, using only the 50% of trials

not used for the clustering analysis. This resulted in a distribution of parameters for each neuron across cross-validation folds, which we then averaged and plotted against ΔMI (Figure 2.8J-K). Again, only the amplitude (height) parameter could explain the differences in direction information ($R^2 = 0.859, p = 3.505 \cdot 10^{-75}$). Taken together these results show that during the active task, neurons displayed a continuous, approximately multiplicative, modulation of the amplitude of their tuning functions. These task-based modulations likely contribute to the patterns of IDI, DDI, and SDI neurons observed previously.

Trial-to-Trial Response Variability is Modulated by Task Demands

In addition to tuning, MI also takes into consideration trial-to-trial response variability. We reasoned that the task effects on MI may be partially due to altered variability of neural responses. To test this, we calculated the entropy of neural responses and normalized measurements relative to those of a simulated Poisson process matched for intensity (entropy factor (HF)).

We found that during the active task, IDI and SDI neurons showed an overall reduction in HF (3-way ANOVA: IDI: $F = 60.24, p = 1.18 \cdot 10^{-14}$; DDI: $F = 0.02, p = 0.879$; SDI: $F = 10.67, p = 0.001$ Figure 2.8D-F) compared to the activity recorded during the passive task. Additionally, all three classes of neurons had changes in trial-to-trial variability that was a function of motion direction (3-way ANOVA: IDI: $F = 6.64, p = 8.14 \cdot 10^{-8}$; DDI: $F = 6.97, p = 2.98 \cdot 10^{-8}$; SDI: $F = 2.07, p = 0.044$; Figure 2.8D-F). All three groups of neurons showed the lowest HF for their preferred direction and the variability increased progressively for less preferred directions. The reduction in variability and its dependence on motion direction was strongest in the IDI group, consistent with the improvement to both tuning and MI during the active task. While the DDI group also displayed an effect of motion direction on HF, unlike the IDI group, it did not show dependence on task demands. SDI neurons showed only a weak reduction in HF during the active task and a very modest dependence on motion direction.

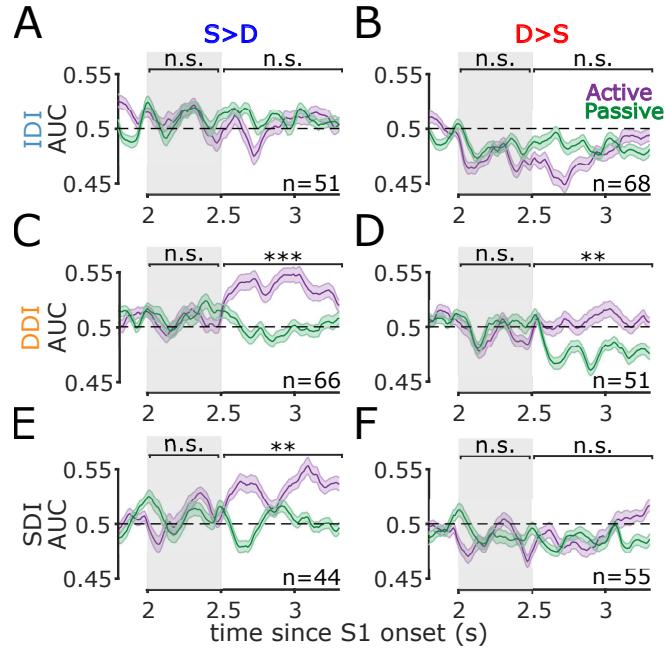


FIGURE 2.9 DDI and SDI neurons have different comparison effects across the active and passive tasks. A.) Time course of comparison effects (AUC between same and different trials within stimulus type) for the IDI neurons with significant $S > D$ in either task. B.) Same as panel A but for IDI neurons with $D > S$ signal. C-D.) Same as A-B but for DDI neurons. E-F.) Same as A-B but for SDI neurons. **: $p < 0.01$, ***: $p < 0.001$, n.s.: not significant t-test for the difference between tasks over the indicated interval. We did not find significant CEs or task effects during the S2 (gray shading) for any group of neurons, but we did see effects during the post-S2 interval for DDI neurons (C-D), and for SDI neurons with $S > D$ CEs (E).

Task Demands Modulate Activity in MT During Non-stimulus Periods

We next turned our attention to the effects of varying task demands on MT activity during periods when motion stimuli were absent (pre-S1, during the delay between S1 and S2, and post-S2). Because no stimuli were present, effects during these intervals would be a likely reflection of cognitive processes (e.g., anticipation of S1 or maintenance of memoranda during the delay). Since the duration of the pre-S1 period was highly predictable (1000ms), a gradual increase (“ramping”) of firing rate during this period, is likely to reflect anticipation (see example neuron in Figure 2.5B). We quantified pre-S1 ramping as the slope of a linear fit to FR vs. time during a one second window before S1 onset. We found significant pre-S1 ramping only during the active task and only for the IDI group (T-test Passive minus Active slopes: $p = 1.02 \cdot 10^{-4}$).

Next, we examined the pattern of activity during the delay separating S1 and S2.

Because the duration of this delay was also highly predictable (always 1.5 s) animals could anticipate the appearance of S2. Similar to the earlier report (Bisley et al., 2004), we found significant ramping in the delay activity during the active task and this ramping was most pronounced for the IDI neurons (Paired sample T-Test on delay slopes: $p = 5.66 \cdot 10^{-5}$) and SDI neurons ($p = 0.0098$; Figure 2.8G-I). This increased ramping in the active task may reflect additional task demands related to the preparation for S2, the comparison stage of the task.

Effects of Task Demands on Comparison Signals

After S2, our active task required subjects to make a saccade to the right if the direction of the preceding S1 was the same (“same trial”), and a saccade to the left if the direction of the preceding S1 was different (“different trial”). It has previously been observed that MT neurons’ responses to physically identical S2 stimuli differed whether they came from “same” or “different” trials, which we termed a “comparison effect” (CE) (Lui and Pasternak, 2011; Zaksas and Pasternak, 2006). Here, we examined whether the activity during S2 showed CEs and asked: (1) are CEs different in the active versus the passive task (when no comparison is needed); and (2) if these effects differ across the three neuron groups (IDI, DDI, SDI)?

We calculated CE as the area under the receiver operating characteristics curve (AUC) within direction condition during S2 and the post-S2 interval, between “same” and “different” trials (Lui and Pasternak, 2011). This metric quantifies the separation of the firing rate distributions for “same” and “different” trials, where a value of $AUC = 0.5$ means the two distributions are indistinguishable, values $0 \leq AUC < 0.5$ means the neuron fires more strongly to identical stimuli on “different” trials compared to “same” trials, and values $0.5 < AUC \leq 1$ means the neuron fires more strongly on “same” trials compared to “different” trials. On average across the population, we did not find significant CEs during S2 (Figure 2.9, Figure ??A,C), but we did find CEs during the post-S2 interval for some groups of neurons (Figure 2.9, Figure ??B,D). We next asked how these CEs were affected by changing task demands. We expected that task effects on cognitive

signals may be independent from task effects on sensory processing, a result that would indicate a division of labor across MT neurons for sensory and cognitive processes: the effect of increasing task demands for a given MT neuron would be to either increase direction signaling, or increase comparison effects, but not both. Thus, neurons with stronger motion direction signals in the active task would not also increase comparison effects in the active task; whereas neurons that did not increase signaling of motion direction in the active task would be more likely to show increased comparison effects.

In line with this prediction, we found that DDI and SDI groups showed significantly stronger $S > D$ CEs during the post-S2 period in the active task, the period immediately preceding the monkeys' report (Figure 2.9C,E; DDI: Kolmogorov-Smirnov test: $p = 4.46 \cdot 10^{-4}$, SDI: Kolmogorov-Smirnov test: $p = 0.0019$ on average CE in window) and a higher proportion of neurons with significant CEs (Figure ??B; DDI: $\chi^2 = 11.14$, $p = 8.43 \cdot 10^{-4}$, SDI: $\chi^2 = 14.23$, $p = 1.61 \cdot 10^{-4}$). Interestingly, the DDI group showed weaker $D > S$ signals in the active task compared to the passive task (Figure 2.9D; Kolmogorov-Smirnov test: $p = 0.0011$ on average CE in window). In contrast to the clear task effects on CEs for the DDI and SDI neurons, the IDI group of neurons showed no significant difference in CE magnitudes between the two tasks (Figure 2.9A-B). IDI neurons did show a trend for more prevalent $D > S$ neurons, but this effect did not stand up to correction for multiple comparisons (Figure ??D, $\chi^2 = 6.21$, $p = 0.013$). When viewing the whole distribution of comparison effects vs direction information, it becomes increasingly clear that neurons with strong comparison effects convey less direction information (Figure ??E), despite being tuned for direction (Figure 2.8B-C). The higher proportion of $S > D$ neurons in the DDI and SDI groups, along with their increased CE signaling, indicates that while the heightened task demands did not improve their stimulus encoding, these neurons may have been recruited to aid the comparison process.

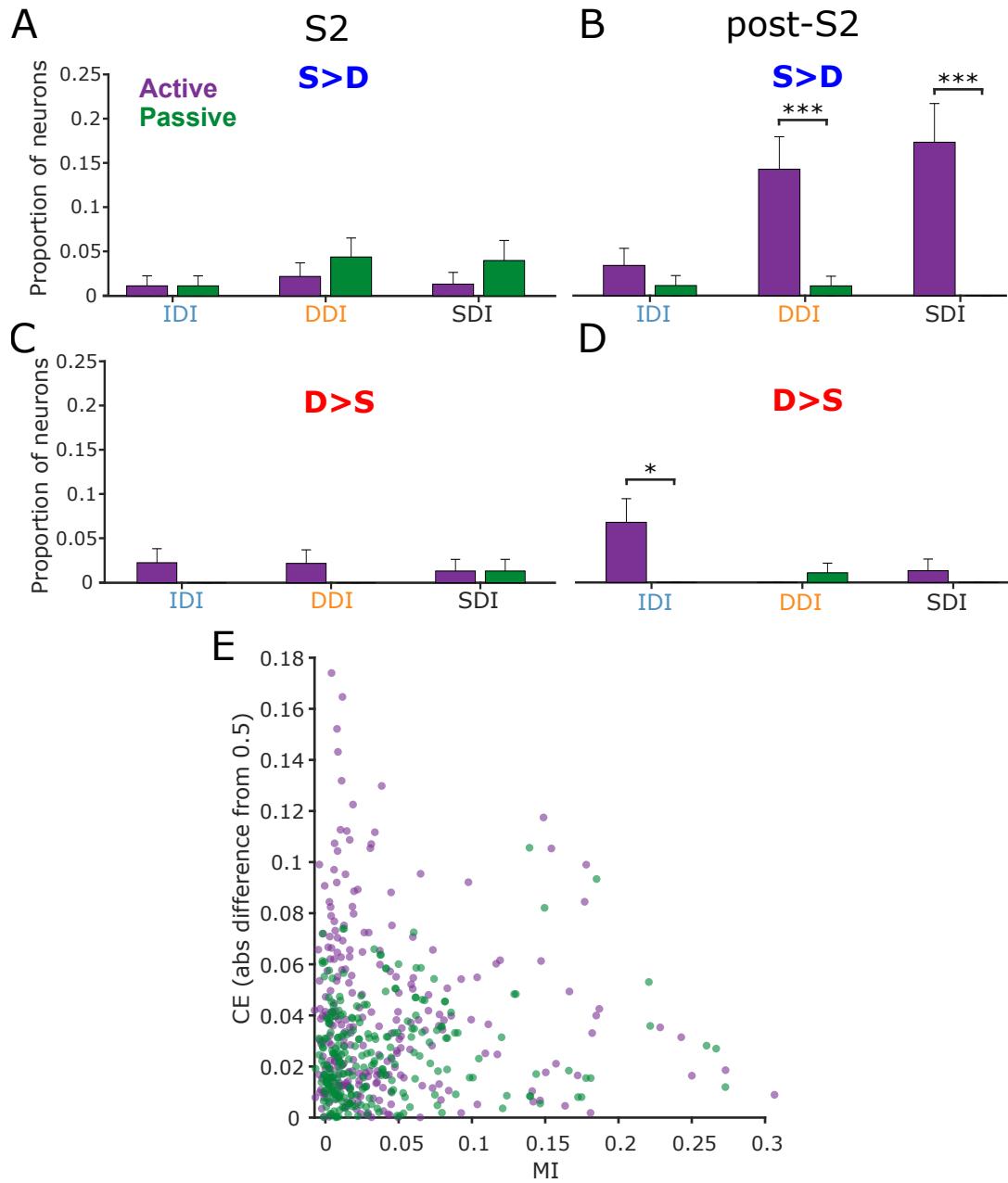


FIGURE 2.10: *Proportion of neurons with significant comparison signals during active and passive tasks* A.) Proportion of neurons within each information group with significant $S > D$ signal during the S2. B.) Same as panel A for the post-S2 interval. C) Same as in panel A for the $D > S$ neurons. D) Same as in panel B for the $D > S$ neurons. (***: $p < 0.001$, χ^2 test). E.) Comparison Effects (absolute difference from 0.5 AUC) vs mutual information for each task. Neurons with the largest CE in the active task also conveyed the least direction information. Error bars = 95% binomial confidence intervals.

2.4 Discussion

Task Demands and Sensory Processing

We hypothesized that neurons in the primate Middle Temporal Cortex (MT) area behave differently when task demands increase in order to improve information processing for goal-oriented behavior. We tested this hypothesis by directly comparing responses to motion during two identically structured tasks: the active task that required monkeys to correctly report a decision for reward, and the passive task in which monkeys were rewarded on each trial without making the report. As the neurons were matched across tasks, we could directly compare how each neurons' tuning and variability were impacted by task demands. Mutual information (MI) analysis used to quantify motion signals suggested heterogeneity of task effects distributed across roughly three groups of neurons: Increased Direction Information (IDI), Decreased Direction Information (DDI), and Similar Direction Information (SDI), each with different contributions to sensory and cognitive aspects of the motion comparison task. During the active direction comparison task, the IDI neurons showed a relative amplitude gain in direction tuning while the direction tuning of the DDI neurons became relatively weakened (Figure 2.8). This difference in tuning is likely to be one of the factors contributing to the change in the altered motion information during the two tasks. These effects on responses to motion are consistent with the classification of neurons based on changes to motion direction signaling in MI, but may not account for the full picture, as MI is also affected by trial-to-trial response variability.

The extent to which responses to motion and variability are each influenced by cognitive factors has been a subject of intense research (Renart and Machens, 2014; Gilbert and Li, 2013), and each has different implications for computation and neural mechanisms. We found a strong relationship between direction preference and trial-to-trial response variability, which did not depend on group (Figure 2.8D-F). Neurons had the largest reduction in trial-to-trial variability during presentations of their preferred stimulus, with weaker effects as direction diverged from preferred. This cannot be due to

the increased firing rate for those directions, as entropy factor (HF) accounts for rate effects (see section [Methods](#)). Although the DDI neurons showed reduced tuning for direction they did not alter their trial-to-trial variability. Previous work has suggested that a substantial portion of trial-to-trial variability can be understood in terms of neurons' susceptibility to changes in cognitive state (Ecker et al., 2016; Denfield et al., 2018); thus, one interpretation of this pattern of results is that these neurons decrease their contribution to sensory processing while maintaining their participation in more cognitive processes. The observed reduction in trial-to-trial stimulus response variability in IDI neurons with higher task demands is reminiscent of the finding of Mitchell, Sundberg, and Reynolds (2007), who found general reductions in Fano Factor in V4 responses when animals selectively attended to a stimulus compared to when that stimulus was not attended, as well as Hussar and Pasternak (2010) in the LPFC using a similar behavioral paradigm to the current experiment. The novel and more surprising result we found is that the strength of variability reduction depended on whether the stimulus was more or less preferred. The improved tuning for direction during the active task suggests that elevating task demands may increase reliability in the information transmission.

Task Demands Affected Anticipatory Signals During the Delay

We found a significant difference in ramping of the delay activity recorded during the active (relative to passive) task (Figure 2.8G-I). Ramping of neural firing rates has traditionally been explored in the context of motor preparation (Ding, 2015; Narayanan, 2016), and decision making (Shadlen and Newsome, 2001), where it has been typically linked to anticipation. Finding anticipatory increases in firing rate for sensory neurons in MT is consistent with previous reports (Bisley et al., 2004; Zaksas and Pasternak, 2006) and is not likely explained during the pre-S2 interval of our tasks through either decision making or motor planning, as neither a correct decision nor appropriate motor plan can be made before first seeing S2. More likely, the ramping we observed is related to preparing cognitive and attentional resources for the perceptual phases of the

task, which had highly predictable timing, as has also been found for LPFC neurons (Hussar and Pasternak, 2010; Hussar and Pasternak, 2013). This interpretation of ramping activity as preparing perceptual resources also aligns with our finding that pre-S1 ramping was restricted to IDI neurons, which enhance perceptual signaling in the active task. This line of thought has also been explored in the context of another nearby visual region, V4 (Snyder, Yu, and Smith, 2018; Luck et al., 1997), where anticipatory ramping has been linked to interactions between V4 and prefrontal cortex (Snyder, Yu, and Smith, 2021). Taken together, these results suggest anticipatory signals may originate in prefrontal cortex and influence visual cortex via top-down feedback.

Task Demands and Comparison Effects

In addition to the more strictly sensory aspects of the task, the active task contains a comparison component which requires the subject not only to process the current stimulus direction (S2) but also retrieve the direction of the previously seen S1, perform the comparison and report it 1 second later. During the passive task, all the components of the task are present, but the animal is rewarded without having to report its decision. Previous work revealed that during the direction comparison task, similar to that used in the present study, responses of many MT neurons during S2 reflected the direction of the preceding S1, with some neurons showing stronger responses to S2 when its direction matched S1 ($S > D$) and other neurons showing stronger responses to S2 when its direction was different from S1 ($D > S$) (Lui and Pasternak, 2011). Here, we examined if such signals, termed comparison effects, are also present when the task does not require the animals to make direction comparisons. We would expect that if CEs in MT are explained primarily by “passive” sensory effects such as adaptation/facilitation (Kohn and Movshon, 2003), we should find similar effects in the active and passive tasks. In contrast, a difference in the magnitude or incidence of CEs between the active and passive tasks would indicate a greater role for memory-guided comparison processes underlying CEs. We found no difference in magnitude of comparison effects for IDI neurons at a population level, but a slight increase in incidence for significant

$D > S$ signals in this group. Taken together with the gain on tuning, it may be the case that this result reflects heightened adaptation in the active task, which would be consistent with a primarily sensory explanation of CEs for these neurons. However, we found significantly stronger (Figure 2.9) and more prevalent (Figure ??) CEs for the DDI and SDI groups, suggesting a greater cognitive role for CEs for these cells. These cognitive effects being isolated to neurons that became worse at signaling motion direction when that information became task-relevant further supports a division of labor in area MT for completing the task. That is, these neurons shed direction signaling in favor of increased CEs when active comparison was required for the task.

We found that comparison effects emerged most clearly during the period after the offset of S2 and before the execution of the response saccade (the ‘post-S2’ interval; Figure 2.9). This underscores the non-sensory, cognitive nature of these effects (since the sensory stimulus was absent when effects were strongest) and reduces the potential that response adaptation or repetition suppression could explain D>S comparison effects (Lui and Pasternak, 2011; Kohn and Movshon, 2003). This relatively late time-course of comparison effects in MT differs from earlier reports, which found comparison effects emerging shortly after S2-onset (Lui and Pasternak, 2011; Zaksas and Pasternak, 2006). One reasonable explanation for this difference in the timing of comparison effects across studies is the difference in timing between the onset of S2 and the animals’ reports. Earlier studies required the animals to respond immediately following the offset of the S2, whereas in the current study animals must withhold responding until after the post-S2 delay. Thus, if the timing of comparison effects was more tightly related to the animals’ reports than to the time of S2 onset, then that would be consistent with the results across all studies of comparison effects in MT and further underscores the cognitive role of comparison signals. Finally, in earlier studies decisions were reported with a manual button press (Lui and Pasternak, 2011; Zaksas and Pasternak, 2006), whereas the current study required animals to report with a saccade, therefore CEs are independent of the mode of decision reporting, reinforcing the notion that comparison effects are related to cognitive decisions rather than the formation of particular motor plans.

Class	Tuning			HF			Delay Ramping
	Task	Dir.	Stim.	Task	Dir.	Stim.	
IDI	$9.54 \cdot 10^{-5}$	$5.47 \cdot 10^{-126}$	0.238	$1.18 \cdot 10^{-14}$	$8.14 \cdot 10^{-8}$	0.018	$5.66 \cdot 10^{-5}$
DDI	$1.97 \cdot 10^{-5}$	$1.49 \cdot 10^{-140}$	0.187	0.879	$2.98 \cdot 10^{-8}$	0.011	0.557
SDI	0.005	$2.70 \cdot 10^{-66}$	0.531	0.001	0.044	0.021	0.010

TABLE 2.1: Task effects summary. Significance of task effects (p value) separated by information clusters. Statistics for Tuning and entropy (HF) were a three-way anova between task, motion direction, and stimulus (S1 vs. S2). Delay ramping was compared using t-tests on the slopes (Passive minus active: (2.8 G-I).

2.5 Conclusion

We investigated the effects of task demands on information processing in area MT of the macaque. We found a diversity of neurons that were affected in different ways in the active compared to the passive task. Neurons that increased their signaling of task-relevant sensory information (IDI neurons) did not show changes in comparison effects, which reflect more cognitive memory and comparison processes, whereas neurons that either decreased (DDI) or did not change (SDI) direction information in the active task showed changes in their comparison signals. These results suggest a distributed allocation of labor of cognitive and perceptual processes across a population of neurons in MT and highlight the role of MT populations in all phases of memory-guided comparisons of visual motion direction.

Chapter 3

Relating Neural Data to Stimulus Parameters in 3+ Dimensions

In this section I will introduce various analysis approaches for quantifying neural population tuning to stimulus spaces of arbitrary dimensionality. These approaches will first be applied to a simulated population of neurons in order to build intuition for the neural data in chapter 4.

3.1 Simulated Neural Populations

I begin by referring back to the 1D tuning curve from chapter 1, figure 1.2. I use the simple gaussian tuning curve as a template, and apply it to a small population of 3 neurons along 3 arbitrary stimulus dimensions. The choice of 3D neural state space and 3D stimulus space comes from the curse of dimensionality; in order to completely sample the spaces, 3 dimensions is the maximum due to combinatorics. For the sake of simplicity I will begin with a set of axioms with which to constrain my artificial neurons:

1. Neurons have a minimum (0) and maximum (R_{max}) firing rate
2. Neurons are gaussian tuned to each stimulus dimension.
3. There is one global maximum stimulus per neuron, which lies at a point in the stimulus space ($\vec{\rho}_n$).

4. Neurons' responses to stimuli is some function of distance between its' preference and the stimuli
5. Tuning to each stimulus dimension is independent of other dimensions (e.g. color and orientation tuning are assumed independent).
6. Neurons' responses to stimuli are independent from other neurons

Later on, I will reduce the number of assumptions by adding in neural correlations and tuning dependence, but for now start with the simplest model.

3.1.1 Defining the Population Response Function

Due to axiom 6, I can calculate the responses of each neuron independently and define a single response function $f_n(\vec{\theta})$. This function is bounded by $f_n(\vec{\theta}) \in [0, R_{max}]$ according to axiom 1. If we combine axioms 2 and 4 with these natural boundaries, we get equation 3.1, which defines the neural response in terms of the stimulus values ($\vec{\theta}$), peak firing rate (R_{max}), and decay term ($D_n(\vec{\theta})$; i.e. tuning width) for a stimulus space of arbitrary dimensionality.

$$f_n(\vec{\theta}) = R_{max} \cdot e^{D_n(\vec{\theta})} \quad (3.1)$$

The decay term, $D_n(\vec{\theta}) \in [-\infty, 0]$, depends on both the euclidean distance between $\vec{\theta}$ and $\vec{\rho}_n$, and the neurons' tuning matrix (τ_n). τ_n determines the rate at which the firing rate decays as $\vec{\theta}$ diverges from $\vec{\rho}_n$.

$$D_n(\vec{\theta}) = (\vec{\rho}_n - \vec{\theta})^\top \cdot \tau_n \cdot (\vec{\theta} - \vec{\rho}_n) \quad (3.2)$$

Neurons therefore respond maximally to their preferred stimulus $\vec{\rho}_n$ at a rate of R_{max} . As the stimulus parameters differ along any dimension, this maximum response decays exponentially at a rate defined by τ_n .

Symbol	Description
$\vec{\theta}$	Vector of stimulus values for presented stimulus
$\vec{\rho}_n$	Vector of preferred stimulus values for neuron n
$f_n(\vec{\theta})$	Response in $\frac{spk}{s}$ for neuron n to stimulus $\vec{\theta}$
$D_n(\vec{\theta})$	Decay term for neuron n to stimulus $\vec{\theta}$
R_{max}	Maximum firing rate
τ_n	Tuning matrix for neuron n

3.1.2 Deterministic Simulations

If we ignore for the moment neural response variability, we can analyze the expected values of the population response in stimulus space. I begin with a relatively simple approach to quantifying linear relationships between the two 3D spaces, Canonical Correlations Analysis (CCA). CCA rotates the two spaces and orders the new bases by descending correlation. In this way, I can pull out which stimulus dimension is most linearly related to the neural dimensions (CCA pair 1). This answers the question: "which stimulus dimension is best at predicting neural responses." A variety of neural parameters will ultimately affect these relationships, which I will expand on in this section.

we can see clearly from figure 3.1, as tuning narrows for a neural population the correlations drop significantly. One way to interpret this is that a linear-readout of stimulus encoding is only possible in the wide-tuning regime. Once the neural data becomes too "peaky" (neurons respond profoundly to a single stimulus but weakly to most stimuli), a linear approach is no longer appropriate. In real-world situations this is not an issue. As discussed previously, natural visual scenes contain innumerable variability and dimensionality, and as the dimensionality of a stimulus increases, wider tuning is known to be more effective at conveying accurately and robustly (Brown and Bäcker, 2006).

Another important feature of neurons that can affect these analyses is the dimensionality of the neural tuning space relative to the dimensionality of the stimulus space. In a naturalistic setting, there will undoubtedly be stimulus features that neurons are not tuned to. For example, if color is present for a purely orientation-tuned neuron, or visual features in general for an auditory neuron. When working in 10+ feature dimensions, having stimulus variability that is irrelevant for neurons allows for CCA to capitalize on

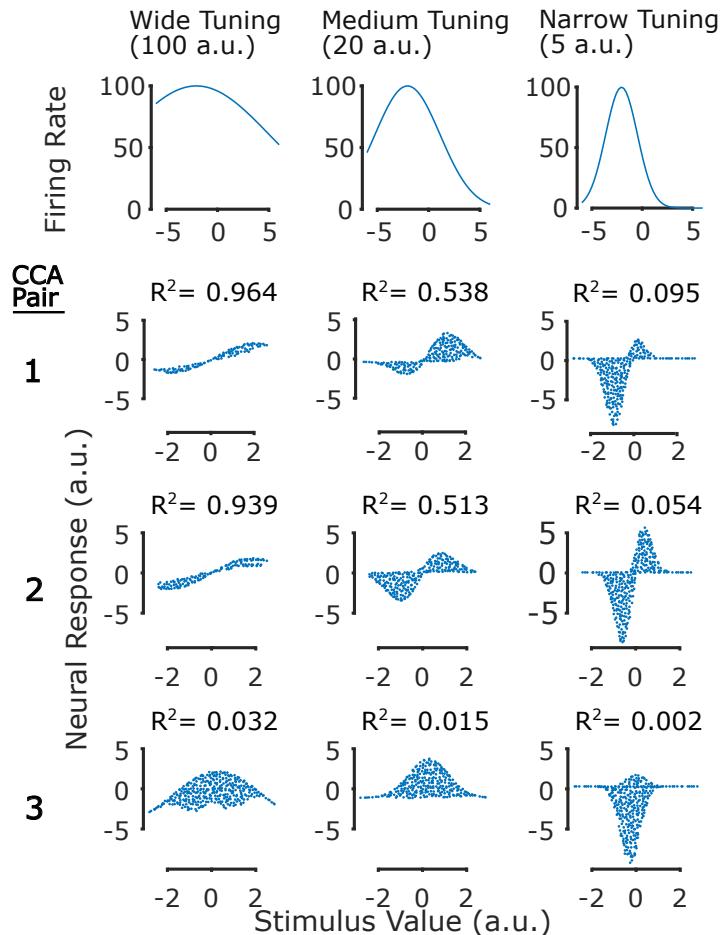


FIGURE 3.1: *The effect of tuning width on CCA.* Top row.) Example tuning curves for three simulations with descending tuning widths. First column.) Canonical correlation variable pairs for a population of 3 neurons with wide tuning. Neurons were assumed to have the same tuning width but different preferred stimuli along the three dimensions. Second&Third column.) Same as first column but for medium and narrow tuning respectively.

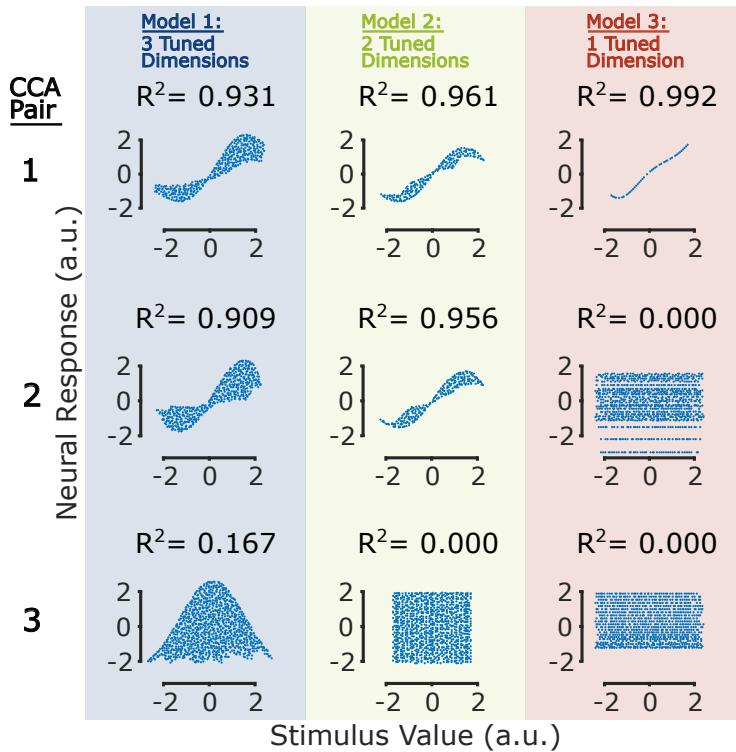


FIGURE 3.2: *The effect of irrelevant stimulus dimensions on CCA.* Canonical correlation variable pairs for three models relating neural population space to stimulus space. Model 1.) Similar to models shown in figure 3.1, with all three stimulus dimensions being relevant to the neurons. Tuning for neurons was set at 50 a.u. across models. Model 2.) Same as model 1, except one stimulus parameter (consistent across neurons) was set to have no impact on the firing rate of the neurons. Model 3.) Same as model 2 but with multiple irrelevant stimulus dimensions.

the extra dimensionality. This becomes clear in figure 3.2, where the embedding dimensionality remains the same across models but the dimensionality of tuning decreases. As the number of relevant stimulus dimensions for neurons decreases, the R^2 for the tuned variables increases.

3.1.3 Poisson Variability Simulations

As mentioned previously, real neurons display a property called response variability. When shown the same stimulus multiple times, they may respond more or less across trials. This is known to occur most similarly to a Poisson distribution, and therefore I added in Poisson response variability to the tuning simulations from 3.1 to create figure

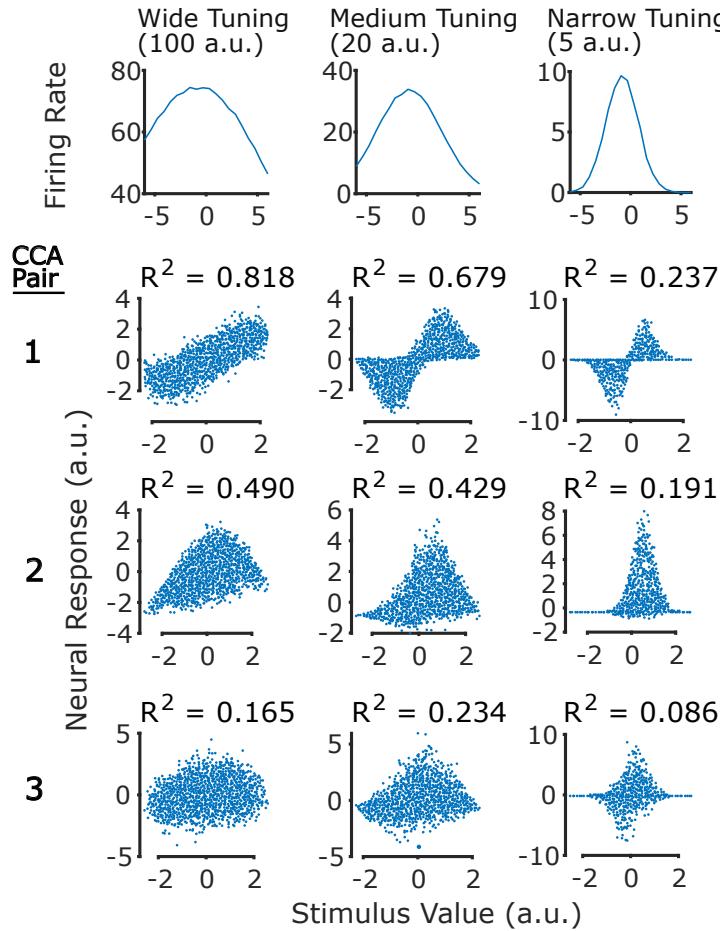


FIGURE 3.3: *The effect of Poisson variability on CCA.* Same as figure 3.1 with the exception of added Poisson variability to the neural responses.

3.3. While the relationships from the deterministic setup hold, there are some interesting differences that arise from this added variability. Poisson variability appears to function as an equalizer across tuning profiles. In the narrow tuning regime, all three canonical variable pairs become more linear (higher R^2) after adding variability (right columns of figures 3.1 and 3.3). In the wide tuning regime where neurons already encompass large swaths of the stimulus space, correlations decrease likely due to the added uncertainty.

In this section I have explored the relationship between tuning properties of neural populations and linear readouts of stimulus parameters through a series of simulations and canonical correlations analysis. Most importantly, CCA was able to correctly identify tuning properties of the neural population despite both a wide range of tuning differences and a lack of true linearity in the data. I will further apply the principles

explored in this section later in chapter 4 on real neural data. By only simulating 3D neural populations with 3D stimulus spaces, I have avoided a central problem that arises in even as low as 4 dimensions; what stimuli to show. In the simulations from this section I equally sampled all three stimulus parameters independently, but this becomes nearly impossible in animal experiments where time is limited. I address how to make this choice of which parts of stimulus space to sample in the next section, using techniques from machine learning called optimization. I first give a brief overview of optimization before bringing it back to the specific problem of optimizing stimuli for neural populations.

3.2 Optimization

In its simplest form, optimization begins with an unknown function of inputs \vec{x} :

$$f(\vec{x}) : \mathbb{R}^n \rightarrow \mathbb{R} \quad (3.3)$$

where the output of the function is deterministic of its inputs \vec{x} . The goal of optimization is to then find the set of inputs which minimizes (or maximizes) the output of the function $f(\vec{x})$.

We can illustrate this using the simple example of a parabola.

In the realm of low-noise and few variables, optimization is a rather simple problem of calculating the gradient and following it. In Figure 3.4, we test the function by using two random inputs (points 1&2), calculate the gradient (3), and follow the gradient for the next input (4). This is an iterative process whereby each time we sample the function, we obtain more knowledge about it. Continuing this process will eventually result in $x = 2$ as the global minimum. While this may appear obvious, in real world problems the blue line is unknown, necessitating the use of optimization techniques. In this chapter, I will first describe the two main classes of optimization. Second, I will link these concepts to functional properties of neurons, focusing heavily on the visual system.

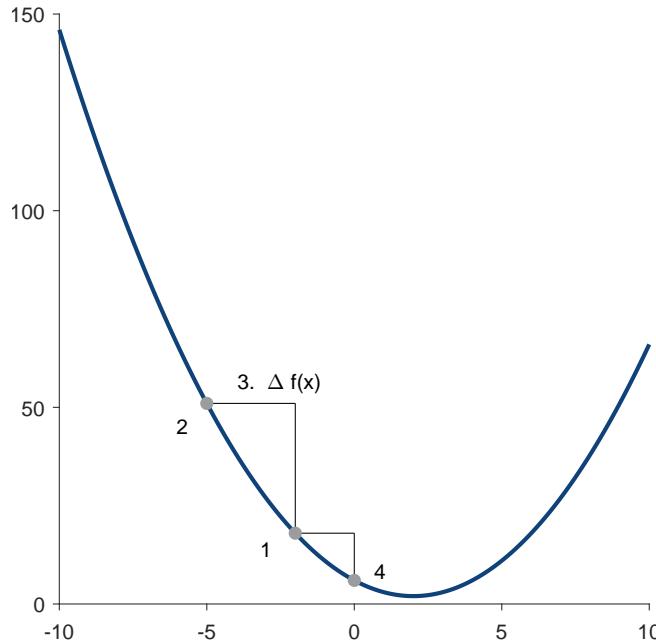


FIGURE 3.4: *Optimization Example*. This example demonstrates the simplest case of optimization. Some linear, deterministic function (blue line) is sampled repeatedly and the gradient is followed to find the minimum.

3.2.1 Derivative-Based Optimization

Consider again the example in figure 3.4. This is a demonstration of derivative-based optimization problem in which the algorithm successfully converged on the correct answer. Several parameters had to be appropriately selected for even this simple problem to converge. If, for example, the initial two points chosen were $x = [-5, 7]$ the derivative would've evaluated to 0 resulting in no solution. Luckily for a problem of this magnitude we simply need to sample a third data point. While this may seem like an absurd reason to discount the approach (the probability of randomly choosing $x = -5$ or $x = 7$ is 0), it becomes a real problem with higher-dimensional data, realistic variability, and limited experimental time (see: Section 3.1). Unfortunately, because data is often high-noise and high-dimensional (in the parabola example there is only one input x , but most problems have many), calculating gradients quickly becomes computationally intractable in finite time. In addition to these constraints, real-world problems are often non-convex, meaning they have many local minima where gradient descent fails (figure 3.5).

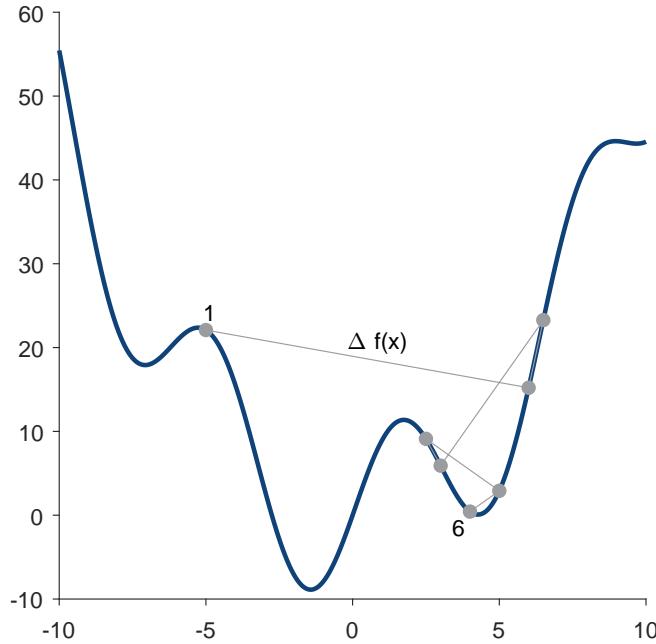


FIGURE 3.5: *Non-Convex Optimization Example.* Here I show how gradient-based optimization can fail in even simple 1D problems. When a function is non-convex (multiple minima/maxima), the local gradient can often be misleading about the global behavior of the function.

In figure 3.5, we can again run gradient descent beginning from $x = -5$ and follow it to the final solution. Starting from point 1 and following through point 6 in the figure we can see how strictly following the gradient leads to the optimization getting trapped in a local minimum, instead of finding the global solution. In fact running this exact problem using a quasi-Newton method ("fminunc" function from the Matlab 2022b Optimization toolbox) results in this failure of derivative-based optimization. In more natural contexts (more variables, high noise, large datasets, etc.) derivative based methods are all but intractable, which lead to the development of derivative-free optimization approaches that I expand on in the next section.

3.2.2 Derivative-Free Optimization

In this section I will give brief overviews of two algorithms: particle swarm optimization and genetic algorithms. I expand further on the specifics of the algorithms in chapter 4 where I apply both to optimizing neural data.

Particle Swarm

Particle Swarm Optimization (PSO) is essentially a population-approach to problem solving where a set of possible solutions are all transformed across iterations according to the relative fit both within and across the history of each tested solution. The algorithm is again initiated by testing a set of randomly-selected inputs and recording the outputs. Each input vector is called a particle (think back to the geometric representation of vectors from figure 1.4) due to its conception as a point in the possible solution space. These particles are then moved through the solution space across iterations (step: S_t) according to their personal history of function evaluations (local best solution: L_p), the evaluations from other particles (global function behavior: G_p) and a momentum term to alleviate the influence of local minima (M_p).

$$S_p = c_1 L_p + c_2 G_p + c_3 M_p \quad (3.4)$$

Over the course of many optimization iterations, PSO can balance exploration and exploitation through the use of these three terms and their assigned weights.

Particle swarm has the benefit of scaling well with high-dimensional and complex search spaces, but suffers from the need to optimize it's own parameters. Depending on the dimensionality of the problem and complexity of the input-output relationship, choices of coefficients ($c_{1,2,3}$) can make or break the algorithms ability to optimize.

Symbol	Description
S_t	The step each particle takes for the next generation
c_1	Weight for local gradient component
L_p	Vector pointing towards the best solution within each particles history
c_2	Weight for global component
G_p	Vector pointing towards the best solution across particle histories
c_3	Weight for momentum component
M_p	Particle Velocities from last generation (S_{t-1})

Genetic Algorithm

Another powerful derivative-free optimization approach is the genetic algorithm, which takes its inspiration for how DNA is recombined during reproduction of organisms. Much like PSO, it begins with evaluating an initial set of randomly-selected input vectors. These vectors (particles in PSO) are then paired up as parents, recombined and mutated into the next generation of inputs following their relative fitness as solutions (figure 3.6).

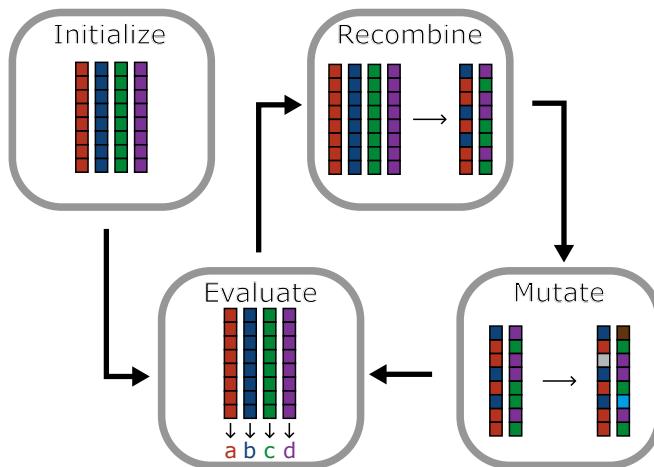


FIGURE 3.6: Genetic Algorithm Flowchart. The general process of how a genetic algorithm optimizes.

Much like particle swarm, genetic algorithms have proven to be a powerful optimization approach that is much less susceptible to complex high-dimensional geometry and peakiness of the solution space (Whitley, 1994; Mirjalili, 2019). In fact these optimization techniques have already begun showing promise in optimizing stimuli for individual neurons (Ponce et al., 2019; Bashivan, Kar, and DiCarlo, 2019). In the next section I describe a set of experiments that compare optimization approaches to intelligently explore neural response functions in multiple brain areas.

Chapter 4

Using Generative Models of Natural Images to Define Neural Tuning Manifolds

4.1 Introduction

Research into sensory coding in the visual system focuses on determining what visual features neuronal activity covaries with, i.e. what information is it encoding. Traditional experiments looked at neural responses to simple stimuli such as small patches of light or oriented bars (Hubel and Wiesel, 1959). These hallmark studies uncovered the simple/complex models of V1 neurons, which describe the way in which neurons integrate inputs from the Lateral Geniculate Nucleus (LGN). Therefore, gabor stimuli are effective at driving neurons in early visual areas due to the precise arrangement of their inputs. Later in the visual hierarchy, such as V4, it becomes much more difficult to infer the exact tuning properties of neural inputs, and how they combine to form emergent tuning of these neurons. Neurons in area V4 only weakly respond to such stimuli as gabors, as they prefer complex combinations of many features such as contrast, color, orientation, curvature, etc (Sani et al., 2013; Tanigawa, Lu, and Roe, 2010; Nandy et al., 2016). It is not feasible to systematically test every combination of visual features to accurately describe feature tuning in such brain regions. Here we describe a method for efficiently exploring rich stimulus spaces to find relationships between sensory features

and neural responses.

In these experiments, we used a generative adversarial network (GAN) as our high-dimensional feature space. A GAN is a type of artificial neural network trained to represent the high-order statistics from a training set of data using a lower-dimensional set of ‘latent’ variables. In a GAN trained on faces for example, one of these variables may be the age of the person (Karras, Laine, and Aila, 2019). The complex combinations of visual features that makes someone look older or younger actually represents a one-dimensional variable, and GANs can capture that. We sought to investigate if neurons in visual areas V1 and V4 are tuned to the image statistics captured by this latent variable model. In order to effectively quantify neural tuning in such large stimulus spaces, an efficient search method is required. To this end, we utilized two methods to efficiently explore parts of the GANs latent space that are most likely to result in large neural responses from the population, PSO and genetic algorithms.

While the idea of optimizing stimuli for neurons is not new, most of these studies focus on maximizing the response of single neurons (Ponce et al., 2019; Bashivan, Kar, and DiCarlo, 2019). It is unlikely that a single unique stimulus exists such that it maximally excites a diverse neural population. This is especially true in areas like V4, where neurons have complex tuning across many modalities (Pasupathy and Connor, 2002). In fact, there are likely many unique stimuli that drive the neural population at similar levels due to contributions from different neurons. To contend with this difficulty, we use two distinct optimization approaches both utilizing the L^2 norm of the population response vector. Cowley et al (2017) demonstrated that optimizing for the L^2 norm results in a reasonable estimation of the response manifold (Cowley et al., 2017).

We found several strong relationships (linear and nonlinear) between the latent space of our GAN and neural activity in both V1 and V4. These relationships arose even when random stimuli were shown, but they were systematically stronger when stimuli were optimized using either PSO or genetic algorithms. By directly comparing the stimulus embeddings within session and brain area, we found that V1 is consistently more linearly-related to stimulus parameters than V4. This holds true across both GAN

latents and the actual pixel values. We show here that stimulus tuning in visual neurons can be elucidated using a single experiment, removing the need for assumptions about stimulus parameters, brain region differences, and optimization approaches. The flexibility afforded by this single-experiment approach for describing neuron preferences is a powerful tool for future studies of neural tuning.

4.2 Methods

Anesthetized Experiment

The subject used in this study was one adult male rhesus macaque (*Macacca Mulatta*), Tw. All surgical and experimental procedures were performed in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and were approved by the University of Rochester Committee for Animal Research.

Surgical Procedures.

TW was given an initial dose of ketamine (10mg/kg, IM) to induce anesthesia, as well as Atropine (0.05mg/kg, IM) and dexmethazone (1-4ml/kg, IM). The animal was then maintained on fentanyl (10–25micrograms/kg/h, i.v.), isoflurane (0.25 – 1.5%), and nitrous oxide (1:2, in oxygen) throughout the experiment. Animals were placed in a stereotaxic frame and body temperature was maintained with a thermostatically controlled heating blanket. Animals were monitored for ECG, temperature, expired CO₂, SPO₂, and blood pressure to maintain the anesthetic plane. Animals were also given lactated Ringer's with dextrose (5 – 20ml/kg/h, IV) to prevent dehydration, cefazolin (11mg/kg/h, IV) antibiotic. The scalp was first incised and retracted to uncover the right hemisphere from V1 to V4. Craniotomies were made over V1 and V4 in the right hemisphere. The dura was reflected and then the craniotomies filled with agar (1% in

sterile saline). Animals were then fit with contact lenses and paralyzed with vencuronium bromide (0.1-0.6mg/kg/h, IV) before recording. After neurophysiological recordings were complete at 100 hours post initial anesthesia, animals were euthanized by an injection of sodium pentobarbital (80mg/kg, IV).

Neural Recordings.

Neurophysiological recordings were done using stereotactic implantation of Phase 3B neuropixels (Jun et al., 2017) and recorded using SpikeGLX. Neuropixels were lowered normal to the surface of the cortex in V4 and allowed to baseline for around 30 minutes prior to recording. Following several sessions in V4, we recorded from V1 following the same procedures. Receptive fields were mapped by hand on a whiteboard and listening for spikes. Stimuli were then sized to best cover the conglomerate receptive fields of simultaneously recorded neurons.

Visual Stimuli.

Naturalistic full-color images (as described in methods section: [Generative Adversarial Network](#)) were displayed on a gamma-calibrated 100Hz CRT monitor (ViewSonic, Brea, CA) placed 50-60cm from the animals' eyes. Stimulus control was done with ViSaGe system (Cambridge Research Systems, Rochester, UK) using crs 24-bit color mode in coordination with custom-written Matlab (Mathworks, Natick, MA; for optimization and display) and Python (GAN image generation) routines. Images were displayed on a grey background for 500ms with 1000ms inter-stimulus-intervals.

Data Preprocessing: Anesthetized.

Spiking data was first run through Kilosort (Steinmetz et al., 2021) to extract individual neurons across channels. Neuron quality was then manually evaluated using Phy before being converted to NeuroData without borders prior to further analyses. Neurons

were included for analysis if they were identified as clean, independent units by Kilosort, totaling 1084 neurons in V1 from three optimization experiments (PSO, genetic, random) and 588 neurons in V4 from three separate optimization experiments.

Alert Experiment

Surgical Procedures

Prior to any electrical recordings, the subject (Ro) was first implanted with a titanium head holder, trained on behavioral tasks, and then implanted with chronic electrode arrays. All surgeries we performed with isoflurane anesthesia, aseptic technique, and perioperative opiate analgesics in accordance with the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and were approved by the University of Rochester Committee for Animal Research.

Head Holder Implantation

The subject was initially sedated with a 10mg/kg intramuscular injection of ketamine and administered with 0.25mg/kg midazolam, 0.011mg/kg glycopyrrolate, 25mg/kg cefazolin, and 0.2mg/kg meloxicam, all intramuscular. Once anesthetized, we intubated the animal with an endotrachial tube, shaved the animals head, inserted a catheter in the small saphenous vein for infusion of lactated Ringer's solution, then positioned the head in a stereotaxic frame. The anesthesia was maintained with 1.5% isoflourane throughout the surgery. The surgical site was then thoroughly scrubbed with povidone iodine solution in the preparation of the sterile field. We made a horseshoe-shaped incision (6cm wide by 10cm anterior-posterior) using a #10 scalpel blade starting from the right brow, cutting caudally parallel to the midline, and ending at the left brow. We then used bone curettes to retract the tissue and clear a large enough surface of skull for the head holder. Sterile gauze soaked in saline was used to keep the tissue moist throughout the surgery. The prefabricated titanium head holder (custom design machined by the university of Pittsburgh cite mat?) attaches to the skull using 16 6mm titanium screws (Veterinary Orthopedic Implants, Saint Augustine, FL) that go through

6 flange straps radially extending from the center of its base. We next bent the straps of the head holder so it fit tightly on the skull. We marked with pencil the final position of the head holder on the skull and made two small incisions in an x through the retracted tissue where the head holder would protrude. We then pushed the exterior portion of the head holder through the dermostomy in the retracted tissue and repositioned it back on the skull. We used a 2mm surgical drillbit and a custom drill stop set to the thickness of the skull to drill through the skull, starting with the lateral-most location of the head holder. Screws were implanted one at a time, alternating across the midline, lateral to medial. Once each screw was implanted we used Geristore (DenMat, Lompoc, CA, USA) to fill any gaps between the skull and head holder. After the geristor fully dried, we sutured the wound using a continuous running stitch to attach the subcutaneous fascia and simple interrupted stitches on top to fully connect the skin around the wound margin. Intramuscular injections of 25mg/kg cefazolin were administered every 12h for seven days post op, and 0.2mg/kg meloxicam every 24h for 3 days post op. The subject was allowed one month to recover before the start of training.

Behavioral Training

Prior to array implantation, the animal first underwent fixation training. The monkey sits 50cm away from a 120Hz ViewPixx/3D monitor (VPixx Technologies, Saint-Bruno, QC, Canada). We used Matlab (The MathWorks, Inc) and Psychtoolbox to control the experiments and present visual stimuli. Eye position was tracked with an Eyelink 1000 IR eye tracking camera (SR Research, Ottawa, Ontario, Canada), and the monkey was given water reward for fixating on a central dot. Once the monkey understood the basic reward contingencies he was implanted with two 128 channel matrix electrode arrays (NeuroNexus, Ann Arbor, MI, USA).

Array Implantation

The monkey was given intramuscular injections of ketamine, medazolam, glycopyrrolate, cefazolin, and meloxicam at the same doses described earlier for the headpost

holder surgery. Similarly, the monkey was intubated, shaved, catheterized, and positioned in the stereotactic frame in the same manner. Again the animal was maintained on 1.5% anesthesia throughout the surgery. We marked the stereotactic coordinates for prefrontal cortex (30mm anterior, 21mm lateral, 25mm dorsal) and visual area V4 (0mm anterior, 0mm lateral, 27mm dorsal) as well as estimated locations for the two array pedestals (NeuroNexus, Ann Arbor, MI, USA) prior to making the incision. Once we planned out the location of the pedestals and craniotomies we made the incision with a size 10 scalpel blade, starting on the posterior surface of the cranium. The incision was made just off the midline (away from the hemisphere we implanted in) and continued until about 1cm away from the margin around the head holder, leaving room to ensure that enough healthy tissue remained between the incision and head holder. We then cut a hemicircle around the head holder, again leaving enough healthy tissue to be able to suture the wound and continued just off the midline up to the brow. The final incision started halfway through the hemicircular incision, perpendicular to the midline, and continued 10cm laterally. We used bone curettes to retract the tissue while minimizing muscle damage, and again kept the retracted tissue hydrated with sterile gauze and saline. Once the skull was sufficiently cleaned, we marked the location of the craniotomies in pencil using stereotactic coordinates. The pedestals were both placed on the midline, one anterior to the head holder (PFC), and the other posterior to the head holder (V4). Similar to the head holder surgery, we then bent the legs of the two pedestals to fit tightly onto the skull and marked the location of each screw with a pencil. The animal was then administered a second intramuscular dose of cefazolin. We used 8 and 10 6mm titanium screws (Veterinary Orthopedic Implants, Saint Augustine, FL) for the PFC and V4 pedestals respectively. After the pedestals were secured we used a 19mm diameter trephine to do the PFC craniotomy followed by kerrison punches to remove any pieces of bone around the perimeter. The animal was then administered a 0.5mg/kg intramuscular dose of dexamethasone. Afterward, we used a drimmel to smooth down a trench going from the pedestal to the craniotomy in order to prevent the wire bundle from snagging or bending. We next cut the dura on three sides of a 1cm

square and retracted it. A 128 channel matrix electrode array (NeuroNexus, Ann Arbor, MI, USA) was inserted dorsal to the principal sulcus, using a microdriver (Zaber, Vancouver, British Columbia, Canada) attached to an all-angle manipulator (NeuroNexus, Ann Arbor, MI, USA). Once the array was in place, we inserted the reference wire under the dura, sutured it closed with nurolon sutures (5-0 dissolving sutures), and covered the craniotomy with duragen (Integra Life Sciences co.). The animal was injected with Buprinex SR IM every two hours, starting 6 hours into the surgery until we finished. We then used two titanium plates to cover and protect the craniotomy by screwing them into the skull in the same manner as the pedestals and headpost. After the craniotomy was secure we covered the trench and wire bundle with kwik-sil and filled any gaps under the PFC pedestal with geristore. We then sutured around the anterior wound margin back to the lateral incision using the same procedure as the head holder sutures. These same steps from craniotomy to suturing were replicated for the implantation of the V4 array. Post operative care included administration of 25mg/kg cefazolin every 12h for seven days, and 0.2mg/kg meloxicam every 24h for 3 days.

Data Preprocessing: Alert.

A initial automatic step of maximum likelihood estimation of gaussian distribution parameters. After this automatic step, results were manually refined using custom routines for MatLab (Kelly et al., 2007)

Data Analysis

All analyses were done with custom routines for Matlab 2022b (The MathWorks, Inc). Due to limitations in real-time spike sorting, a simplified definition of neural response was used during optimization. During the experiment, a neuron's response was defined as the number of threshold crossings (-4 standard deviations below the median noise) that occurred in a 500ms window from stimulus onset.

Generative Adversarial Network

Generative adversarial networks (GANs) are a type of artificial neural network that learn low-dimensional representations of higher-dimensional data, such as images Karras, Laine, and Aila, 2019. GANs trained to generate images learned to map high-order image statistics onto a set of lower-dimensional latent variables. The GAN used in these experiments had a 128-dimensional input (latent) space, and was trained on the Cifar-10 image data set as described previously Fruend and Stalker, 2018.

Particle Swarm

Particle swarm utilizes a hive-mind approach to solve optimization problems. Each ‘particle’ corresponds to a point in the high dimensional stimulus space (an image), and travels through the space in order to maximize the L^2 norm of the population response. By moving through the stimulus space, the images change along latent feature dimensions, resulting in smooth image manipulation. Particles move through the stimulus space by integrating information about local gradients and neural responses to other particles. Particles therefore both compete and collaborate in order to explore the stimulus and response spaces.

Manifold Approximation with Particle Swarm (MAPS) is initialized with three generations of random points for each of 64 particles. This gives the kernel regression a history to estimate gradients. For subsequent generations, each particle takes a step S_p according to three terms: the estimated local gradient ∇e_p , the weighted sum of vectors towards particles that resulted in better neural responses G_p , and a momentum term M_p (Equation 1).

$$S_p = c_1 r_{1p} \nabla e_p + c_2 r_{2p} G_p + b M_p \quad (4.1)$$

The constants c_1 and c_2 are learning rates for the gradient and global information components respectively, while $b \in [0, 1]$ is a decay term for the momentum. The stochastic scaling factors r_{1p} and $r_{2p} \sim U(0, 1)$ help circumvent the problem of choosing

a correct step size. Too large and particles will jump over maxima, but too small and it will take too long to converge, so step sizes are pulled from a uniform distribution for each particle each generation.

We used kernel regression to estimate the gradient according to each particles' personal history. This way, each particle travels along its' local gradient, independent of how each other particle moves.

$$\nabla e_p(x^*) = \frac{1}{t-1} \sum_{k=1}^{t-1} \nabla w_k(x^*) A_k \quad (4.2)$$

$$\nabla w_k(x^*) = \frac{2K(x^*, x_k) \sum_{l=1}^{t-1} (x_k - x_l) K(x^*, x_l)}{h^2 \left(\sum_{l=1}^{t-1} K(x^*, x_l) \right)^2} \quad (4.3)$$

where

$$A_k = \sum_{j=1}^{t-1} ||r_k|| + ||r_k - r_j|| \quad (4.4)$$

and

$$K(x, y) = e^{-\frac{||x-y||}{h^2}} \quad (4.5)$$

The next component (Equation 4.6) is a weighted sum that uses information about the global response manifold to predict where good parts of the space is. Stimuli that resulted in a large neural response from many neurons pull the particle density toward them. Again, x_p and r_p represent the stimulus embedding vector and the neural response vector respectively for particle p . G_p is the sum of unit vectors towards all the points which resulted in better neural responses $x_k, k = 1, \dots, u$, weighted by how much better that response $r_k, k = 1, \dots, u$ was than the current point r_p .

$$G_p = \sum_{k=1}^u \frac{(||r_k|| - ||r_p||)}{\sum_{j=1}^u ||r_j|| - ||r_p||} \frac{x_k - x_p}{||x_k - x_p||} \quad (4.6)$$

Genetic Algorithm

Another common optimization algorithm for non-convex problems is the genetic algorithm (Katoch, Chauhan, and Kumar, 2021). This approach draws inspiration from how organisms recombine their DNA to optimize offspring for their environment. The process is quite straightforward and consists of four distinct steps each generation:

1. Sampling
2. Selection
3. Crossover
4. Mutation

The genetic algorithm in this experiment was adapted from the XDREAM algorithm presented by Ponce et al., 2019. Briefly, the algorithm is initiated with a random set of 64 stimuli and neural responses to each were recorded. Individual stimuli were then evaluated for optimality during the selection step by z-scoring neural responses within generation and then passing through a softmax function to turn them into probabilities. These are then used as the probability of being a parent during the next step, crossover. The top ten best stimuli are kept unaltered and shown again in the next generation along with 54 generated children. Once parents are selected from the probability distribution defined during selection, they contribute to the child in a 75%/25% ratio. During the final step, 25% of the childrens' latents (genes) are randomly mutated. These mutations are drawn from a 0-centered gaussian with $\sigma = 0.75$. Finally these new stimuli are shown to the neural population and the process repeats for the next generation.

Individual Neuron Models

Neurons were modeled using generalized linear regression with various stimulus parameters on spike count responses, assuming a Poisson distribution of responses. Two sets of predictors were considered for each neuron: GAN Latent variables and Pixel value predictors. Due to the difference in variable counts (128 latents and 3072 Pixel

values) we first used PCA to reduce the dimensionality to 128 for the pixel predictors. We then used the difference in Akaike-Information Criterion (AIC) to determine which set of predictors better explain the neural data within neuron. Negative values indicate that the GAN latent model was better, where positive values indicate that the pixel model was better.

Relative Linearity

In order to directly compare the linearity of embedding across brain regions and sessions, one must control for the number of neurons recorded and the specific stimuli shown.

We controlled for these with a metric we termed Relative Linearity (RL). We first calculated distance covariance on the raw data ($dCov_{raw}$) according to Cowley et al., 2017. We then used CCA to rotate the raw data according to strength of linear relationship between stimulus parameters and neural responses. Then, we again calculated distance covariance on the residuals of the linear CCA analysis ($dCov_{residual}$). This provided us with a set of predictors in which the linear relationships were removed. Finally, relative linearity was calculated as:

$$RL = \frac{dCov_{raw} - dCov_{residual}}{dCov_{raw} + dCov_{residual}} \quad (4.7)$$

This provides us with a bounded measure ($RL \in [-1, 1]$) where -1 indicates that 100% of the discovered relationships were nonlinear, and 1 indicates that 100% linear for each pair of stimulus/neural dimensions.

4.3 Results

Two adult male Rhesus Macaques (*Macaca mulatta*) were shown a series of stimuli generated with a generative adversarial network (GAN) while recording in either V1 or V4. Subject Ro was alert, and had to simply fixate on a central dot until a series of



FIGURE 4.1: *Example GAN Stimuli.* Four example images generated by the 128 GAN latents (top row) and their pixel reconstruction after dimensionality reduction with PCA (bottom row).

stimuli were shown and he was rewarded. Subject Tw was put through a similar experimental protocol with the exception that he was anesthetized throughout. In these studies we were interested in understanding neural tuning to high-dimensional stimuli, as real-world vision vastly differs from laboratory conditions. Most studies of vision investigate either natural images or overly-simplistic shapes and colors. While natural images are rich in information, even a small image will have thousands of pixel values (variables), making them intractable for the majority of scientific questions about stimulus coding. Simpler stimuli on the other hand, can provide clear answers to narrow questions, but rely on a vast set of assumptions, and often fail in the context of population codes. We hoped to bridge this gap with the use of a GAN trained on natural images. These GANs provide rich, colorful, and complex stimuli (as shown in Figure 4.1) while maintaining natural world image statistics (Fruend and Stalker, 2018), and providing a lower-dimensional parameter space (in our case, 128 latent dimensions).

We first looked at the relationship between stimulus parameters and individual neuron responses. We fit generalized linear models for our 128 GAN latent variables to

spike count responses. For each neuron we then also used the actual pixel values shown as predictors for separate models and then compared the AICs in figure 4.2A. As lower AIC indicates a better model, we found that GAN latents are better predictors of neural activity for the majority of neurons (Figure 4.2B).

One of the benefits of our approach is the use of high-density electrode recordings afford us population-level analyses of high-dimensional stimuli. The first approach we used to understanding these neural manifolds is through Canonical Correlations Analysis (CCA). CCA is a powerful tool in situations where we have high-dimensional predictor variables (latents; Figure 4.3A) and high-dimensional response variables (neurons; Figure 4.3B). Normally, with a standard basis set, we may find only weak relationships (Figure 4.3C). CCA allows us to take advantage of the dimensionality by rotating and organizing the predictor and response spaces to front-load all linear relationships (Figure 4.3D). This demonstrates the principal that while the predictor and response spaces in figure 4.3 don't appear that related at first glance, a simple change of perspective can clarify much.

We solved the problem of which stimuli to show using one of three optimization algorithms: Particle Swarm Optimization (PSO), genetic algorithm, or randomly selected stimuli. We then used CCA on 6 sessions, one for each algorithm in each brain area, and calculated the correlation between neural activity and GAN latent variables. Directly comparing these values poses a problem though, as CCA is sensitive to the original dimensionality. So if one session has more or fewer neurons, their correlations won't be comparable. To correct for this we bootstrapped CCA on permuted data within session 100 times and subtracted out what correlations would be expected at chance (Figure 4.4). In both V1 and V4, the genetic algorithm outperformed both PSO and random in finding significantly tuned latent dimensions. Linear stimulus tuning was uncovered even when no optimization was performed, lending credence to the idea that this tuning to GAN latent dimensions is real and not an artifact of optimization.

We next wanted to take a closer look at the specific neuron-focused stimulus dimensions that were uncovered in these experiments. In figure 4.5, we show the strongest

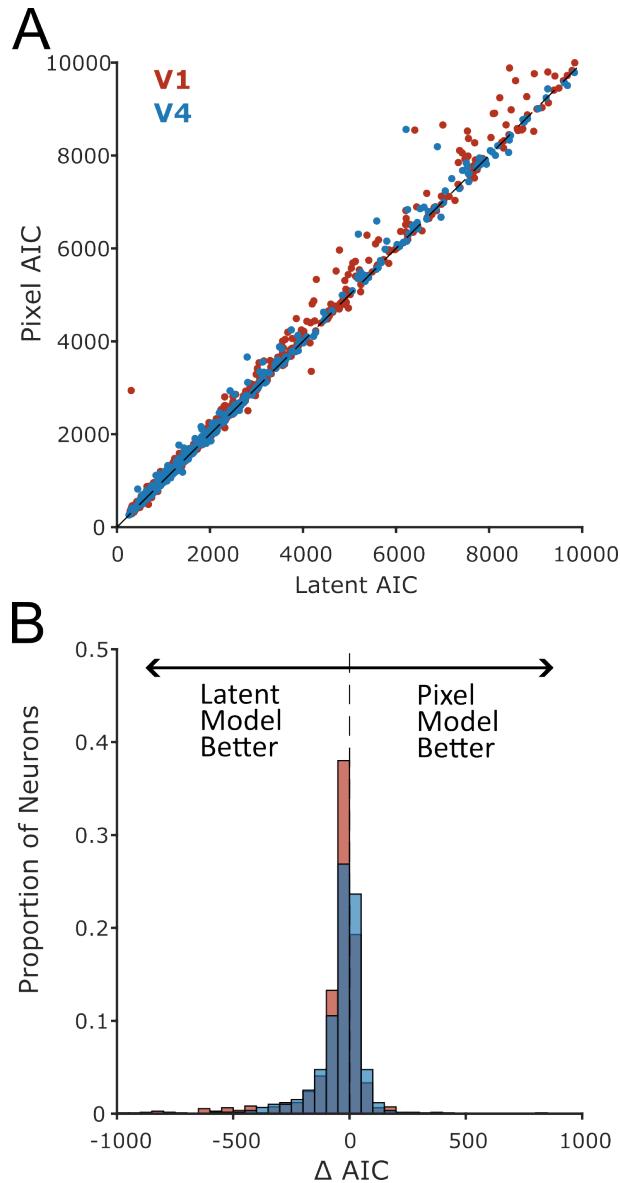


FIGURE 4.2: *Individual neuron model fits* A.) AIC for GLMs using the GAN latents as predictors (x-axis) and the pixel predictors (y-axis) on neuron spike counts. B.) Relative change in AIC across models for individual neurons. Negative values indicate that the GAN latents are better predictors of neural data while positive values indicate that the pixel model was better. The latent model was better on average in both brain regions (V1: $t = -10.26, p = 1.54 \cdot 10^{-23}$, V4: $t = -6.02, p = 3.46 \cdot 10^{-9}$; t-test).

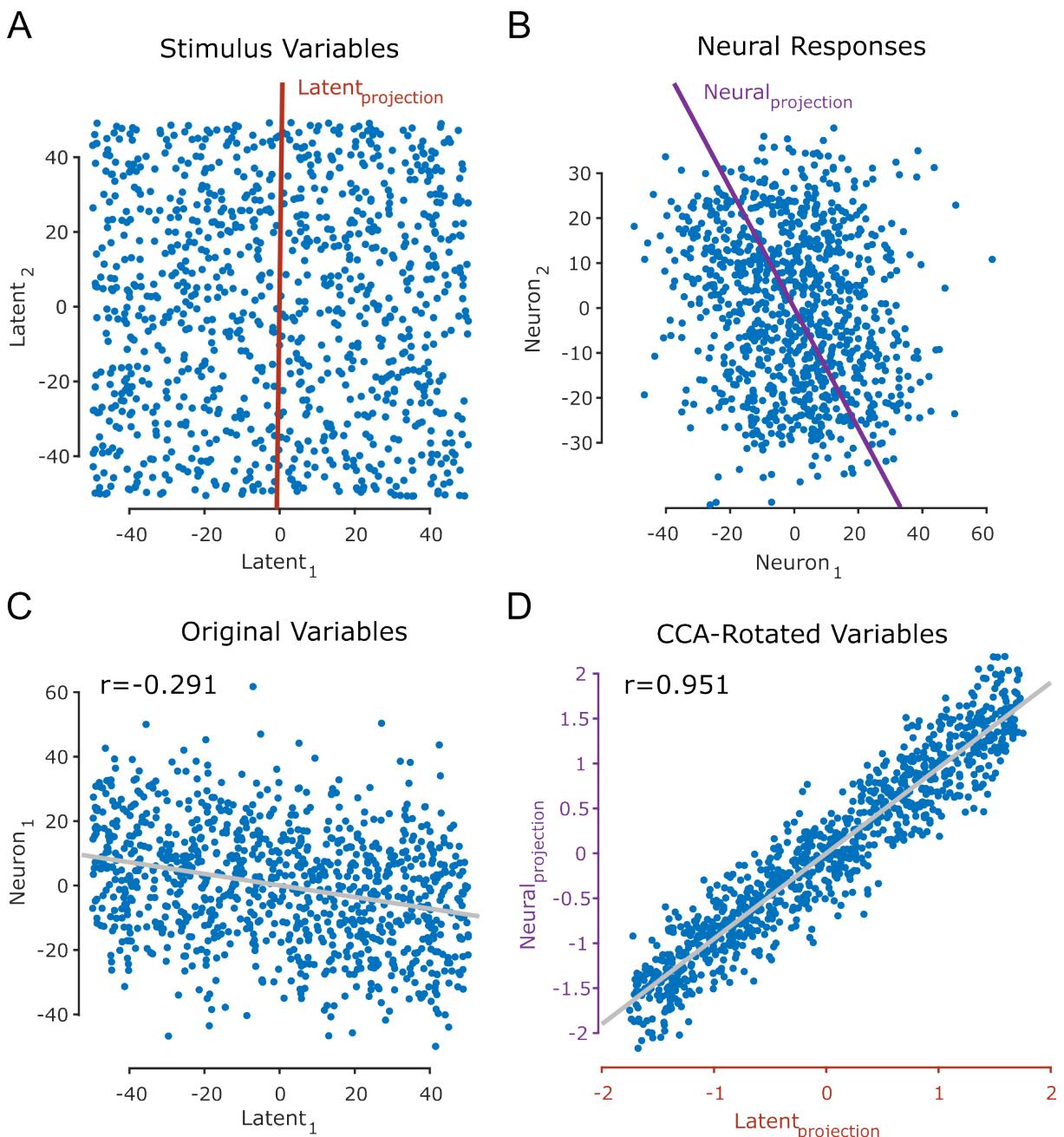


FIGURE 4.3: Canonical Correlations in neural data A.) Simulated stimulus data where points were randomly selected to span an arbitrary space. The red line indicates the axis chosen by CCA for panel D. B.) Neural state space of stimuli from panel A, where each axis is the firing rate of a neuron. Data was centered for visualization purposes as CCA also centers data. The purple line indicates the axis in neural space chosen by CCA for panel D. C.) Original correlation between neuron 1 (panel B) and latent 1 (panel A). D.) First canonical variable pair from the data in panels A and B. The X and Y axes are now linear combinations of the latent and neural dimensions respectively.

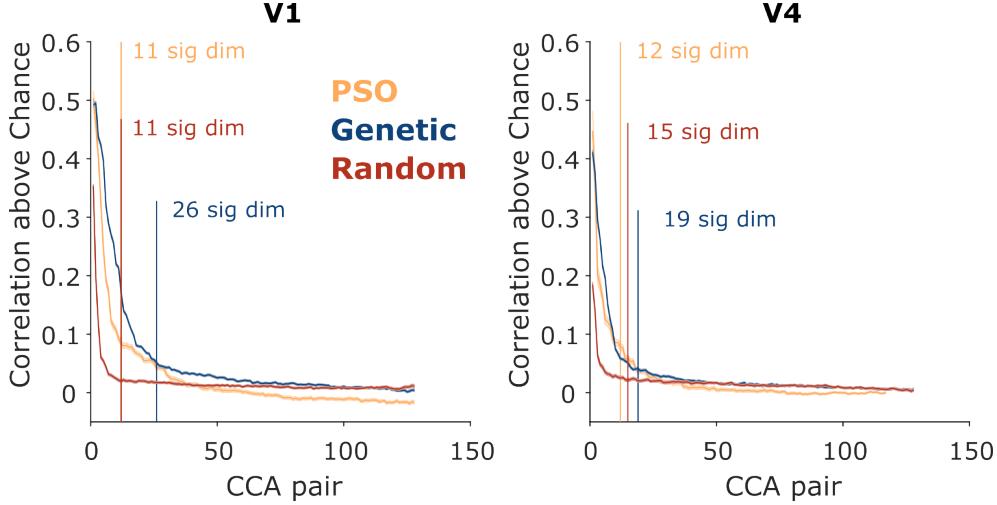


FIGURE 4.4: *Baseline-Corrected R values* The baseline-corrected r values for each canonical variable pair across brain region and algorithm. Baseline correction was calculated by subtracting out the distribution of r values calculated from randomly permuting predictors and calculating CCA 100 times. The number of significant CCA pairs was calculated using Rao's approximate F statistic.

linear relationships between neurons and GAN latents across the algorithms. Interestingly, both optimization algorithms found similar ideal dimensions characterized by high contrast dark patches on yellow-green backgrounds (Figure 4.5G&H). We also see that randomly presenting stimuli results in a much more coarse feature dimension that is largely just light vs dark patches (Figure 4.5I). This is in line with previous studies in V1, and our understanding of simple cells responding best to monochromatic, high-contrast gabor. As for how the algorithm explored the neural manifolds, PSO transitioned much more smoothly (Figure 4.5D) while the genetic algorithm jumped around more sporadically (Figure 4.5E). Despite the different approaches, it is reassuring to see similar features arise across algorithms that tie in neatly to prior studies in V1.

Similar to V1, our experiment found strong linear relationships between the GAN latent space and neural responses. Again, PSO and genetic algorithms found starkly similar feature dimensions (Figure 4.5G&H), while random stimulus presentations found low-contrast global color patches, this time from yellow to blue (Figure 4.6I).

In addition to linear stimulus-response functions, it is well known that mid-level visual areas like V4 demonstrate nonlinear tuning. We decided to use an approach

V1

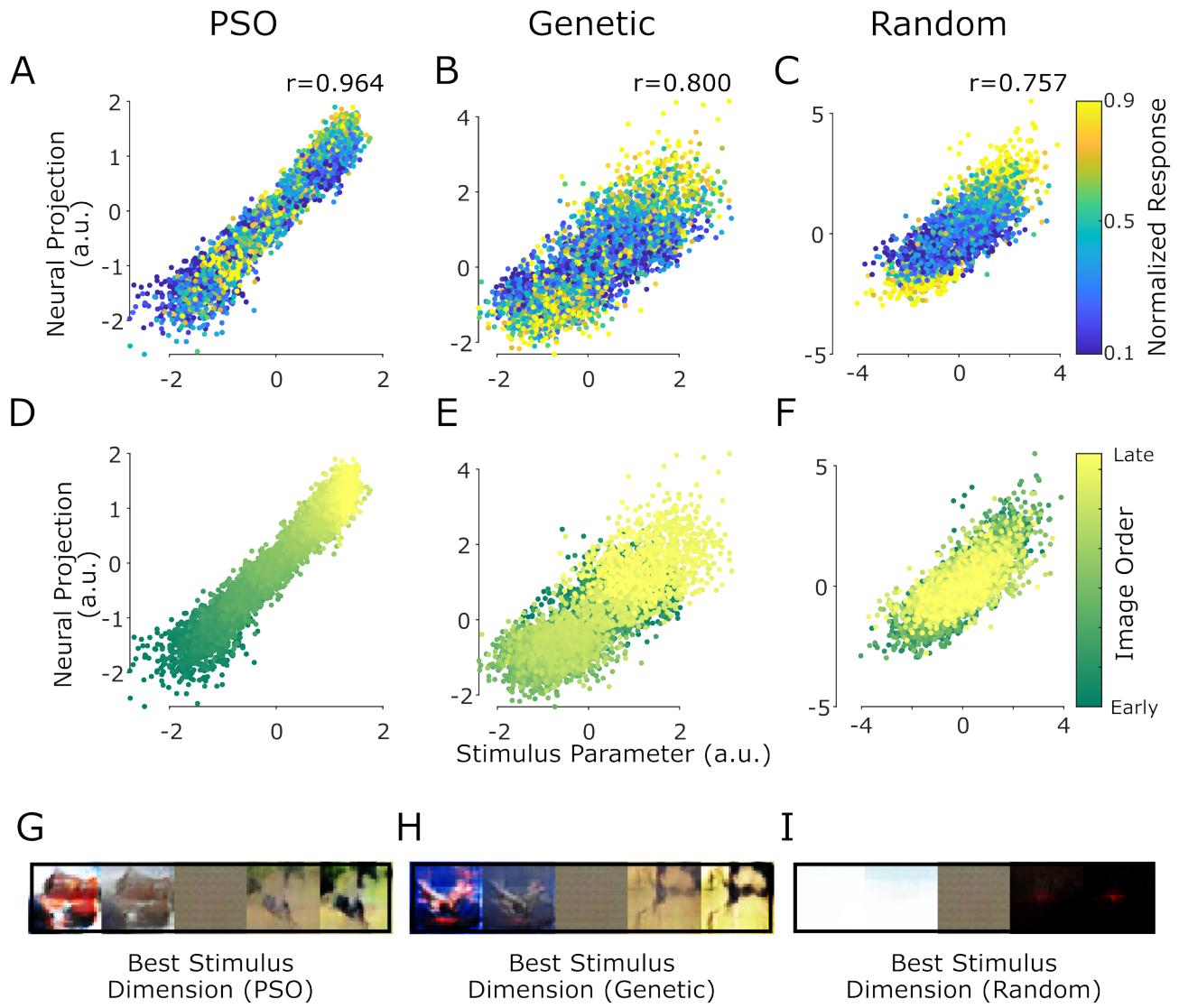


FIGURE 4.5: *Linear Relationships in V1* A.) Top CCA pair for a PSO session in V1. Each point is one stimulus and color indicates the normalized L^2 norm of the population response vector to that image. B.) Same as panel A for a session optimized with the genetic algorithm. C.) Same as panels A&B for a session in which no optimization was performed, and instead random images were shown. D-F.) Same as panels A-C except the color now indicates the order of stimuli in the session. G-I.) The GAN latent dimension that CCA found was most linearly related to neural activity for each algorithm. These dimensions correspond to the X-axes in panels A-F.

V4

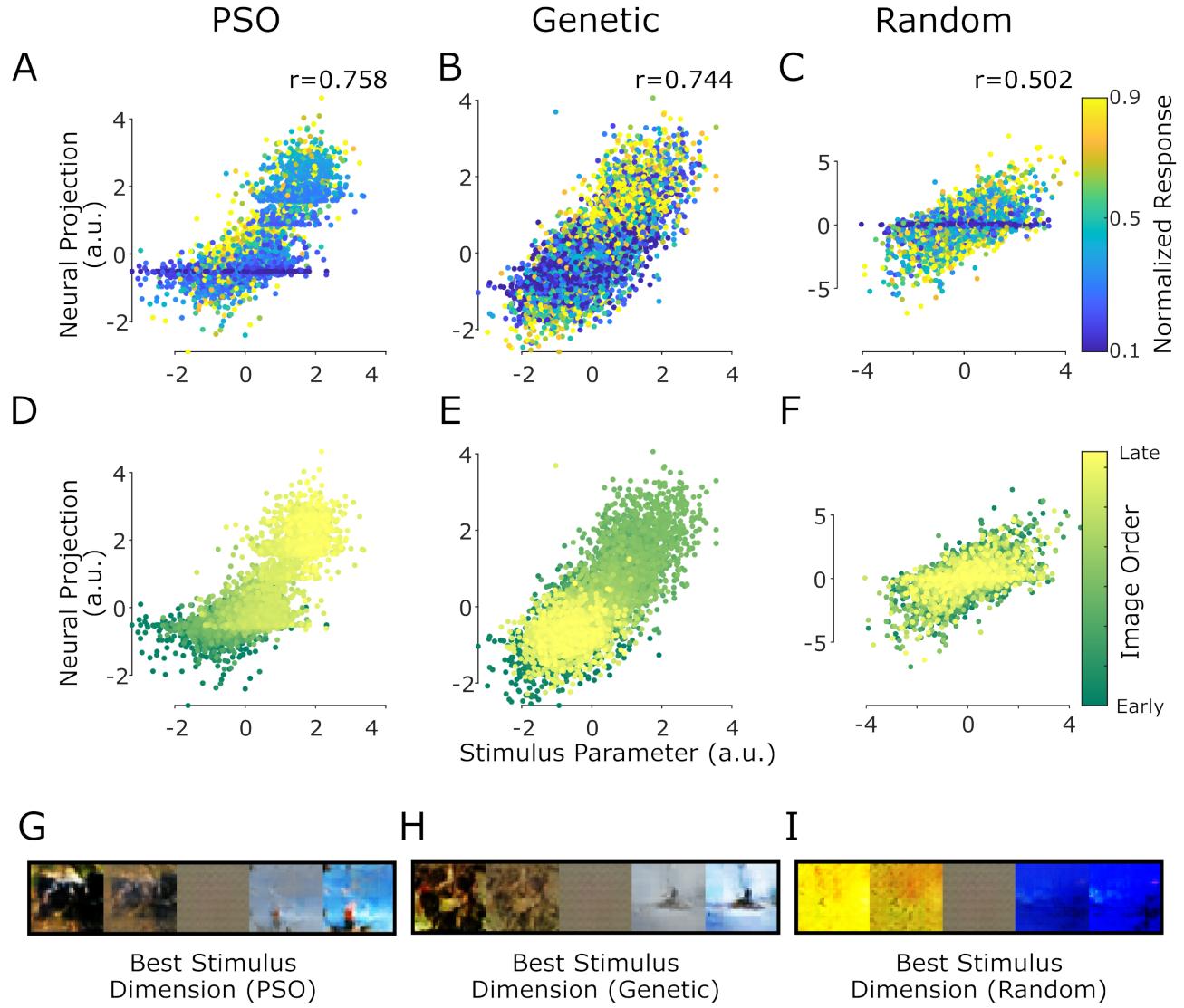


FIGURE 4.6: *Linear Relationships in V4* A.) Top CCA pair for a PSO session in V4. Each point is one stimulus and color indicates the normalized L^2 norm of the population response vector to that image. B.) Same as panel A for a session optimized with the genetic algorithm. C.) Same as panels A&B for a session in which no optimization was performed, and instead random images were shown. D-F.) Same as panels A-C except the color now indicates the order of stimuli in the session. G-I.) The GAN latent dimension that CCA found was most linearly related to neural activity for each algorithm. These dimensions correspond to the X-axes in panels A-F.

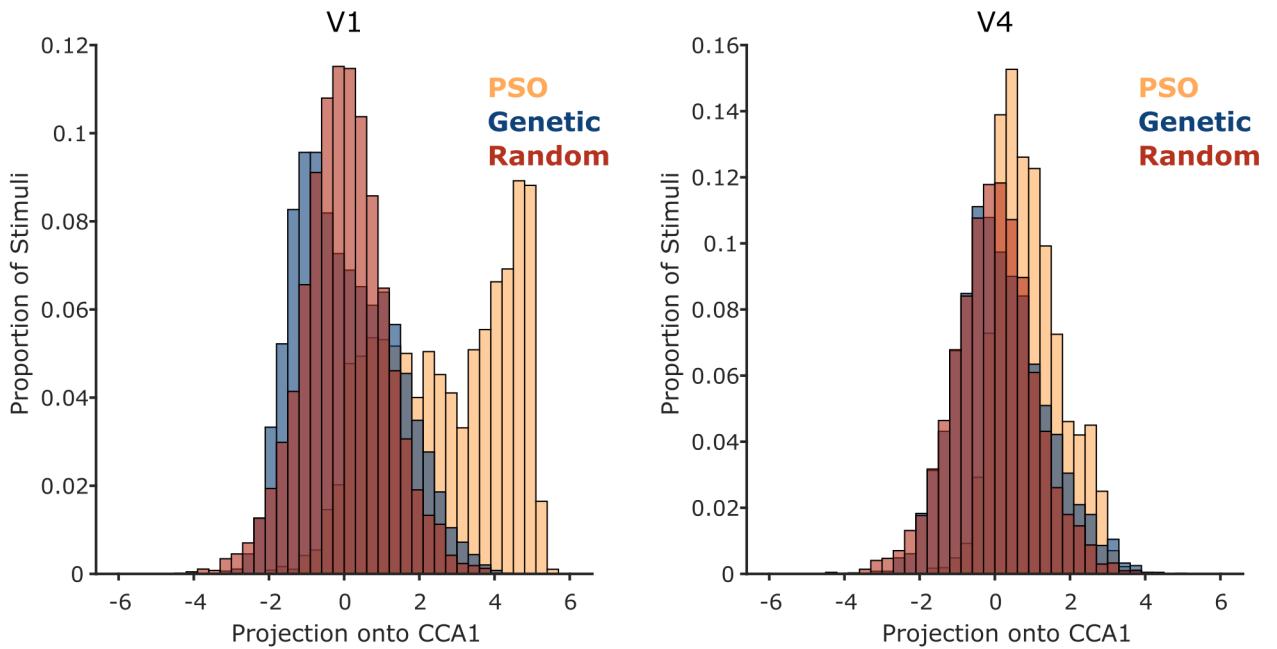


FIGURE 4.7: Projections onto CCA1 The value of projections onto Canonical variable pair 1 for each stimulus shown. PSO tends to pick one side of the latent space that it knows has strong optimization evaluations, whereas the genetic algorithm explores more.

similar to CCA, Distance Covariance Analysis (DCA). Instead of maximizing the correlation between pairs of stimulus/response dimensions, DCA maximizes the distance covariance. One of the drawbacks of pearson correlation is that a value of 0 does not necessarily mean the variables are independent, a drawback that distance correlation does not share. In this way, DCA can be conceptualized as a version of CCA where both linear and nonlinear relationships are taken into account. We ran DCA on the same sessions as CCA and calculated the distance correlation and significance following Shen, Panda, and Vogelstein, 2022. All three session in both brain areas resulted in significant relationships between the GAN and neural responses, though to different degrees.

Another question we had was how linear the stimulus-response relationship was across brain region. In order to test this we calculated relative linearity ([Methods](#)). This metric calculate the distance covariance for the data and the residuals from CCA, and provides a single number $x \in [-1, 1]$ that parses out the contribution of linearity to the distance covariance for each dimension found with DCA. We did this for both GAN and pixel predictors in V1 and V4, and found that regardless of algorithm and predictors,

V1

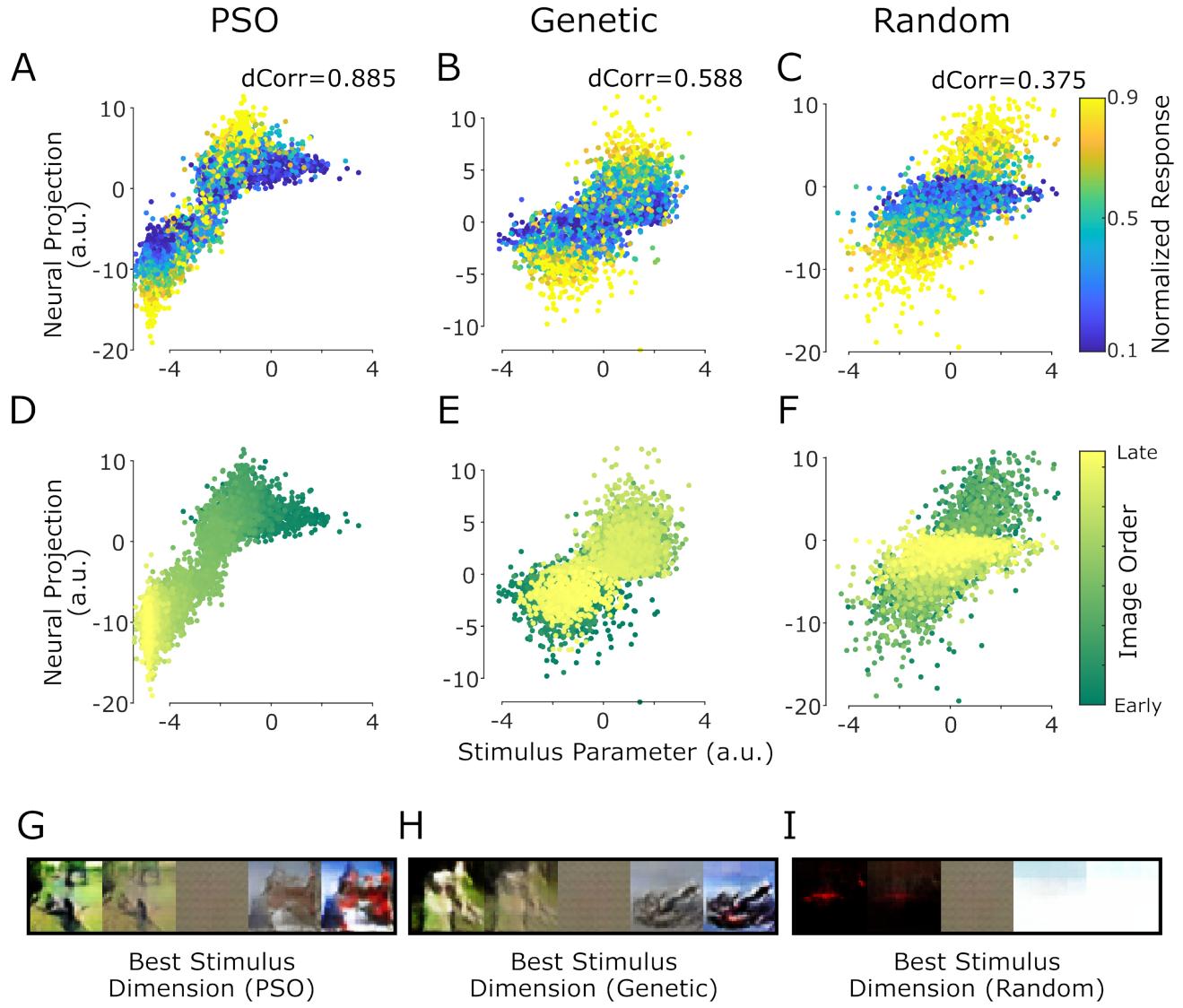


FIGURE 4.8: DCA pair 1 in V1 across Algorithms. A.) Top DCA pair for a PSO session in V4. Each point is one stimulus and color indicates the normalized L^2 norm of the population response vector to that image. B.) Same as panel A for a session optimized with the genetic algorithm. C.) Same as panels A&B for a session in which no optimization was performed, and instead random images were shown. D-F) Same as panels A-C except the color now indicates the order of stimuli in the session. G-I.) The GAN latent dimension that DCA found was most strongly related to neural activity for each algorithm. These dimensions correspond to the X-axes in panels A-F.

V4

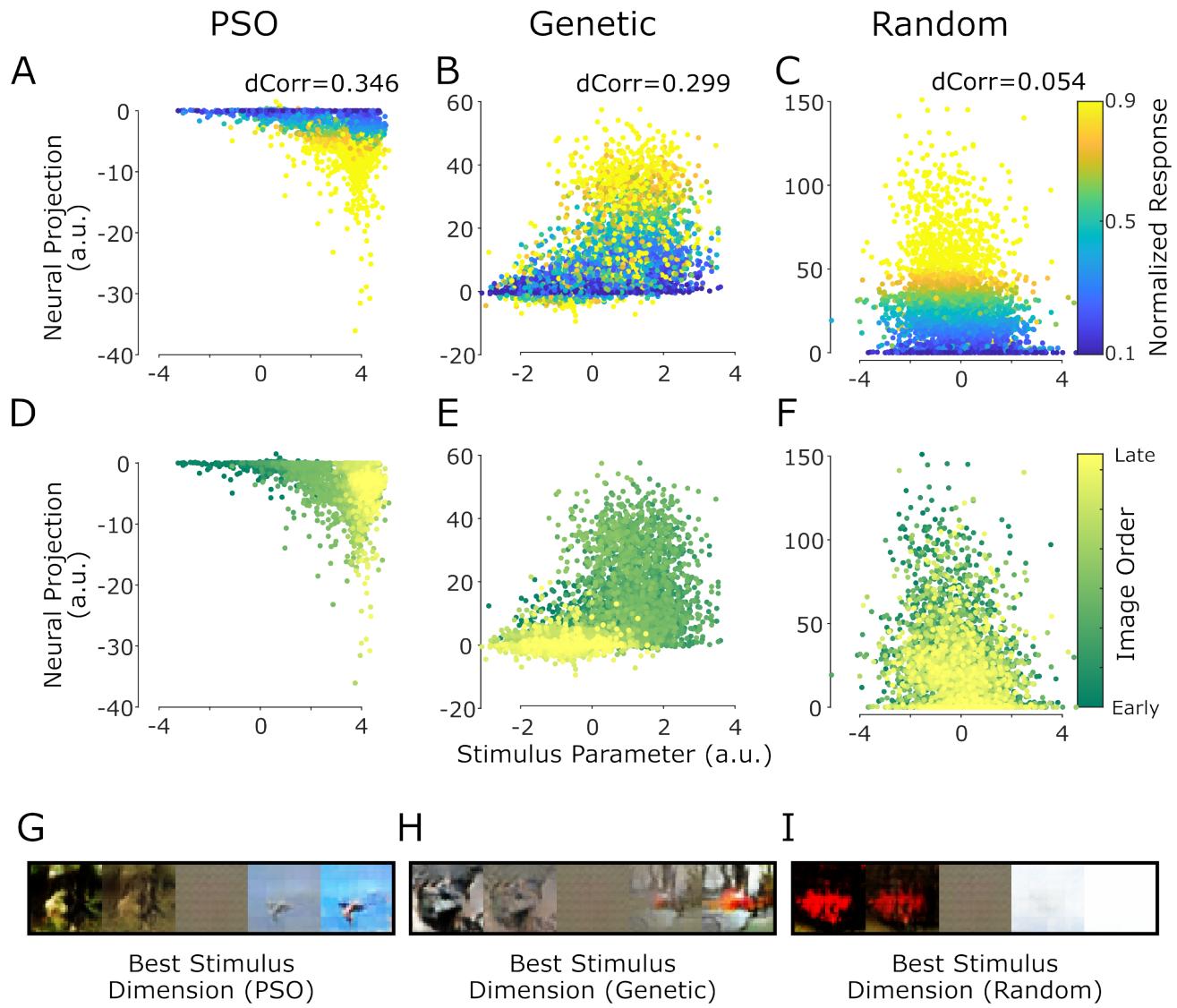


FIGURE 4.9: *DCA pair 1 in V4 across Algorithms* A.) Top DCA pair for a PSO session in V4. Each point is one stimulus and color indicates the normalized L^2 norm of the population response vector to that image. B.) Same as panel A for a session optimized with the genetic algorithm. C.) Same as panels A&B for a session in which no optimization was performed, and instead random images were shown. D-F.) Same as panels A-C except the color now indicates the order of stimuli in the session. G-I.) The GAN latent dimension that DCA found was most strongly related to neural activity for each algorithm. These dimensions correspond to the X-axes in panels A-F.

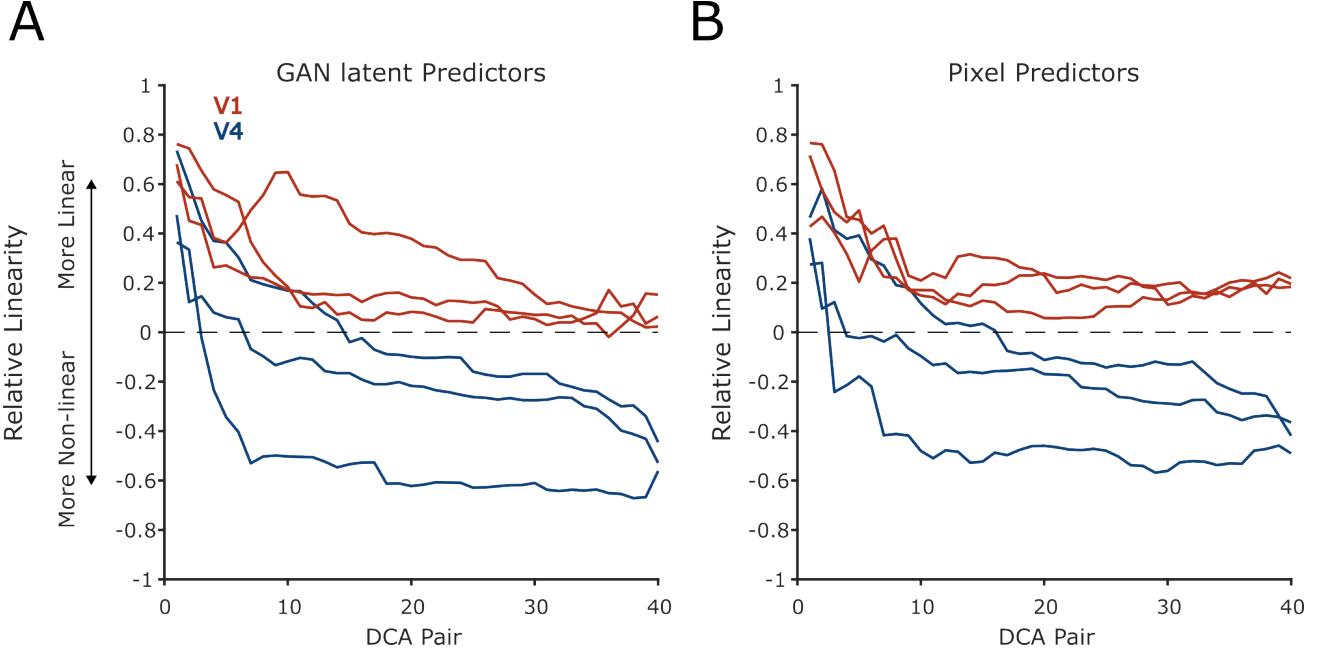


FIGURE 4.10: *Relative Linearity by Brain region* A.) Linearity of neural/stimulus relationships by DCA dimension pairs. Stimulus parameters used were the GAN latent variables. DCA dimension pairs are organized by descending distance covariance and relative linearity was calculated as described in [Methods](#). Each line represents one session, and is colored according to brain area. B.) Same as panel A for pixel predictors.

V1 maintained consistently linear whereas V4 had a much more nonlinear contribution to the stimulus-response function (Figure 4.10).

4.4 Discussion

In this paper we sought to investigate the relationship between neural population activity and the high-dimensional naturalistic stimulus space generated by a generative adversarial network (GAN). Our goal was to develop a pipeline for one-stop-shopping tuning analyses that removed assumptions about brain area, stimulus parameters, and optimization approach. We presented procedurally-generated stimuli to subjects while recording in either V1 or V4, using optimization on the L^2 norm of the population response vector to search the neural response manifold. Stimuli were generated using one of three approaches: Particle Swarm Optimization (PSO), a genetic algorithm, or random stimuli. We found strong linear and nonlinear relationships between the GAN

latent space and neural response space in both V1 and V4, and the GAN was an even better model for neural activity than the actual pixel values shown.

For the majority of individual neurons in both brain areas, the model fits were significantly better in latent space than pixel space. One issue with this approach is the necessary compression in pixel space to 128 dimensions. We can see from figure 4.1 that this compression distorts the image slightly. One may argue that the GAN is simply a better compression algorithm for the stimuli than Principal Component Analysis (PCA) on pixel values, but we believe this wouldn't explain the strong relationships found between the latents and neural data. This is because there is no a-priori reason to believe that the latent dimensions themselves have any meaning to the brain, but we can see from Canonical Correlations Analysis (CCA) that even simple linear relationships are significant (Figure 4.4).

CCA is a relatively simple approach that finds a suitable basis set for both predictor (GAN latents) and response (neurons) variables in which linear relationships are front-loaded. This is analogous to PCA but instead of maximizing variance in the first few dimensions, it maximizes pearson correlation between the spaces (Figure 4.3). We can then look at specifically how many stimulus dimensions are relevant to the neural populations in question. In figure 4.4 we see that all three algorithms (PSO, genetic, random) found significant linear relationships between neural activity and GAN latents, but to various degrees. In both brain areas, the genetic algorithm outperformed PSO and random stimuli in finding neurally-relevant stimulus dimensions. Taken together with figure 4.7, we argue this is due to PSOs tendency to stick to one side of stimulus space that it has determined has large function evaluations whereas the genetic algorithm is free to jump around the space during the recombination and mutation steps. On the other hand, PSO outperformed the other two approaches when it comes to the strength of the relationships found (Figure 4.5A-C; Figure 4.6A-C). This leads to an important choice for experimenters about the goal of using this technique. If, like most optimization problems, the goal is to find the single best stimulus, PSO may be the right choice. If however the goal is to explore the neural response manifold, a genetic algorithm may

be ideal. In any case, it should not go unmentioned that showing a random set of stimuli with no optimization still provides strong evidence of a relationship between the latent space of a GAN and neural response space. This shows that any relationships found are not merely artifacts of optimization, but true linear relationships. Another important point to make here, is the similarity between the stimuli found across algorithms (Figure 4.5G-I & Figure 4.6G-I). It is known that the majority of cells in V1 prefer high contrast light and dark patches, so it is no surprise then that randomly showing naturalistic images pulls out the stimulus dimension in figure 4.5I. The optimization then serves to fill in some of the details around where the contrast edges should be (Figure 4.5G&H). V4 on the other hand, is more tuned to color (Figure 4.6I) and texture (Figure 4.6G&H), similar to prior findings in V4 (Nigam, Pojoga, and Dragoi, 2021; Kim, Bair, and Pasupathy, 2019).

An important next step to consider is that the stimulus-response function may not be a linear one (Prenger et al., 2004; Touryan and Mazer, 2015)). We sought to determine the degree to which V1 and V4 responded linearly and non-linearly to these complex naturalistic images. In order to do so, we utilized a technique similar to CCA called Distance Covariance Analysis (DCA) (Cowley et al., 2017). Instead of maximizing the pearson correlation, DCA finds new basis sets that maximize the distance covariance. We then calculated distance correlation which unlike Pearson correlation, can directly refer to the degree of variable independence (a distance correlation of 0 indicates independence). Again, both V1 and V4 neurons were significantly related to the latent space of our GAN (Figures 4.8 & 4.9). As we were interested in the relative contribution of linear and non-linear neural responses, and distance correlation considers both linear and non-linear relationships, we then needed to parse out the linear component. This would allow us to compare within session and brain region, how linear the neural responses were. We again ran DCA, but this time on the residuals of the CCA analyses, thus providing a distance correlation for the purely non-linear component. By comparing $dCov_{raw}$ and $dCov_{residual}$ (methods 4.2), we can directly compare across sessions and brain regions how linear the neural responses were to these naturalistic images. There

are two possibilities for how this analysis could turn out. The first is that because V1 is lower in the hierarchy, it will respond more linearly to both latent and pixel variables than V4. The other hypothesis, is that V1 will be more linearly related to the pixel predictors and more nonlinear to the latents (and vice versa for V4). This may have been the case due to the captured non-linearities within the architecture of the GAN. Our data supports the former according to figure 4.10A. Here we showed that the strongest relationships between latents and neural activity (low DCA dimensions) in both brain areas were largely linear ($RL > 0$). While all stimulus-response relationships were linear in V1, we found many more dimensions with strong non-linear contributions in V4. These trends held true regardless of whether we used the GAN latent space or pixel space (Figure 4.10B). All together, this study demonstrates an important step toward understanding the relationship between high-dimensional stimulus spaces and neural populations in the visual system.

4.5 Conclusion

We combined high-density neural recordings in V1 and V4 with techniques pulled from machine learning and artificial intelligence to investigate neural tuning to naturalistic stimuli. We used the latent space of a generative adversarial network (GAN) as a stimulus space and searched it using various optimization techniques to maximize the firing rate of our recorded neural population. This lead to confirmation of both prior studies of stimulus tuning in V1/V4 with fewer assumptions, as well as discovering linear and non-linear contributions to neural manifolds. Future work should continue to explore the geometry of neural manifolds in more and more naturalistic settings.

Chapter 5

General Discussion

In this thesis, I have proposed and executed a series of studies investigating how neurons encode stimulus information across a population.

The first study (Chapter 2), looked at how feature tuning in the Middle Temporal Cortex (MT) can change at an individual and population level, depending on behavioral context. I showed that the amount of stimulus information conveyed by MT neurons changes depending on the animals goals. Counter-intuitively, I found that a substantial population of neurons in area MT actually decrease the amount of conveyed stimulus information when those stimuli are relevant to the animal. While these neurons participated less in the sensory signaling aspect of the task, they also contributed more to the sensory decision making process. This division of labor demonstrates yet again the flexibility and efficiency of the brain in solving sensory problems. This same concept of flexible and efficient sensory coding motivated the second study in this thesis, Chapter 4.

We sought to maximize the information gained about neural sensory-response functions using a single experiment. To this end we leveraged the rich, continuous, and naturalistic feature space of a generative adversarial network (GAN) with optimization techniques such as Particle Swarm Optimization (PSO) on large neural ensembles. By optimizing on the L^2 norm of the population response vector, we found that both PSO and genetic algorithms improved the quality of stimulus-response information gained in the experiment defined by stronger linear and nonlinear relationships present between neural responses and GAN latents. We also found striking similarities across

algorithms in the types of features found to be relevant to the neurons, including color palettes, textures and shapes. Interestingly, GAN latent variables proved to be better at explaining neurons' behavior than even the images themselves at the individual-neuron level. This suggests that the brain may in fact be using some form of latent variable model for representing visual information. So for example, color, may not be relevant to the brain in and of itself, but a nonlinear combination of color, contrast, spatial orientation, etc is a more effective or efficient way to process visual information. As artificial neural networks such as GANs take inspiration from the brain, this follows naturally. Another striking finding from this experiment is the degree to which V1 and V4 respond linearly. We found that despite the complexity of stimuli presented in this study V1 still maintained a linear stimulus embedding, while V4 spanned both linear and nonlinear stimulus-response functions. This result held regardless of predictor, in that pixel values and GAN latent variables had similar trends.

Taken together, these studies form the foundation of an approach to sensory coding that takes into account factors that are not typically accounted for in experimental designs by necessity. By accounting for behavioral context (active vs. passive tasks) we demonstrated that the MT alters sensory processing in many (even detrimental) ways, but that these effects served to aid other parts of the task. Then, by accounting for untested feature dimensions, we demonstrated that neurons are tuned to complex, nonlinear combinations of features, and that these relationships can be parsed out with the use of artificial intelligence. Future work into the relationship between visual stimuli and neural responses should combine the studies to investigate how high dimensional neural manifold geometry changes with task design. Limitations in the first study (few simultaneously recorded neurons) and the second (no behavioral task) could be removed by investigating how neural populations encode complex feature sets, and how that encoding changes according to subject goals.

Bibliography

- Albright, T D and R Desimone (1987). *Local precision of visuotopic organization in the middle temporal area (MT) of the macaque*, pp. 582–592.
- Arandia-Romero, Iñigo et al. (Mar. 2016). “Multiplicative and Additive Modulation of Neuronal Tuning with Population Activity Affects Encoded Information”. In: *Neuron* 89 (6), pp. 1305–1316. ISSN: 10974199. DOI: [10.1016/j.neuron.2016.01.044](https://doi.org/10.1016/j.neuron.2016.01.044).
- Barbas, H (1988). “Anatomic organization of basoventral and mediodorsal visual recipient prefrontal regions in the rhesus monkey”. In: *Journal of Comparative Neurology* 276 (3), pp. 313–342. ISSN: 10969861. DOI: [10.1002/cne.902760302](https://doi.org/10.1002/cne.902760302).
- Barone, Pascal et al. (2000). “Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and intimates the operation of a distance rule”. In: *Journal of Neuroscience* 20 (9), pp. 3263–3281. ISSN: 02706474. DOI: [10.1523/jneurosci.20-09-03263.2000](https://doi.org/10.1523/jneurosci.20-09-03263.2000).
- Bashivan, Pouya, Kohitij Kar, and James J. DiCarlo (May 2019). “Neural population control via deep image synthesis”. In: *Science* 364 (6439). ISSN: 10959203. DOI: [10.1126/science.aav9436](https://doi.org/10.1126/science.aav9436).
- Batardiere, A (2002). “Early Specification of the Hierarchical Organization of Visual Cortical Areas in the Macaque Monkey”. In: *Cerebral Cortex* 12 (5), pp. 453–465. ISSN: 1047-3211. DOI: [10.1093/cercor/12.5.453](https://doi.org/10.1093/cercor/12.5.453).
- Bisley, James W and Tatiana Pasternak (2000). “The multiple roles of visual cortical areas MT/MST in remembering the direction of visual motion”. In: *Cerebral Cortex* 10 (11), pp. 1053–1065. ISSN: 10473211. DOI: [10.1093/cercor/10.11.1053](https://doi.org/10.1093/cercor/10.11.1053).
- Bisley, James W, Daniel Zaksas, and Tatiana Pasternak (2001). “Microstimulation of cortical area MT affects performance on a visual working memory task”. In: *Journal of*

- Neurophysiology* 85 (1), pp. 187–196. ISSN: 00223077. DOI: [10.1152/jn.2001.85.1.187](https://doi.org/10.1152/jn.2001.85.1.187).
- Bisley, James W et al. (2004). “Activity of Neurons in Cortical Area MT during A Memory for Motion Task”. In: *Journal of Neurophysiology* 91 (1), pp. 286–300. ISSN: 00223077. DOI: [10.1152/jn.00870.2003](https://doi.org/10.1152/jn.00870.2003).
- Born, Richard T and David C Bradley (2005). “Structure and function of visual area MT”. In: *Annual Review of Neuroscience* 28, pp. 157–189. ISSN: 0147006X. DOI: [10.1146/annurev.neuro.26.041002.131052](https://doi.org/10.1146/annurev.neuro.26.041002.131052).
- Bredfeldt, C E and D L Ringach (2002). “Dynamics of Spatial Frequency Tuning in Macaque V1”. In: 22 (5), pp. 1976–1984.
- Britten, K H et al. (1992). “The analysis of visual motion: A comparison of neuronal and psychophysical performance”. In: *Journal of Neuroscience* 12 (12), pp. 4745–4765. ISSN: 02706474. DOI: [10.1523/jneurosci.12-12-04745.1992](https://doi.org/10.1523/jneurosci.12-12-04745.1992).
- Britten, K H et al. (1996). *A relationship between behavioral choice and the visual responses of neurons in macaque MT*, pp. 87–100.
- Brown, W Michael and Alex Bäcker (2006). *Optimal Neuronal Tuning for Finite Stimulus Spaces*.
- Butts, Daniel A and Mark S Goldman (2006). “Tuning Curves , Neuronal Variability , and Sensory Coding”. In: 4 (4). DOI: [10.1371/journal.pbio.0040092](https://doi.org/10.1371/journal.pbio.0040092).
- Carandini, Matteo and David J Heeger (2012). “Normalization as a canonical neural computation”. In: *Nature Reviews Neuroscience* 13 (1), pp. 51–62. ISSN: 1471003X. DOI: [10.1038/nrn3136](https://doi.org/10.1038/nrn3136).
- Chen, Yuzhi, Wilson S. Geisler, and Eyal Seidemann (Nov. 2006). “Optimal decoding of correlated neural population responses in the primate visual cortex”. In: *Nature Neuroscience* 9 (11), pp. 1412–1420. ISSN: 10976256. DOI: [10.1038/nn1792](https://doi.org/10.1038/nn1792).
- Chung, Sueyeon, Daniel D Lee, and Haim Sompolinsky (2018). “Classification and Geometry of General Perceptual Manifolds”. In: *Physical Review X* 8 (3), p. 31003. ISSN: 2160-3308. DOI: [10.1103/PhysRevX.8.031003](https://doi.org/10.1103/PhysRevX.8.031003). URL: <https://doi.org/10.1103/PhysRevX.8.031003>.

- Cohen, Marlene R and John H R Maunsell (2009). "Attention improves performance primarily by reducing interneuronal correlations". In: *Nature Neuroscience* 12 (12), pp. 1594–1600. ISSN: 10976256. DOI: [10.1038/nn.2439](https://doi.org/10.1038/nn.2439).
- Cohen, Marlene R and William T Newsome (2008). "Context-Dependent Changes in Functional Circuitry in Visual Area MT". In: *Neuron* 60 (1), pp. 162–173. ISSN: 08966273. DOI: [10.1016/j.neuron.2008.08.007](https://doi.org/10.1016/j.neuron.2008.08.007).
- Cowley, Benjamin et al. (2017). "Adaptive stimulus selection for optimizing neural population responses". In: *Advances in Neural Information Processing Systems 30* (Nips), pp. 1396–1406. ISSN: 10495258. URL: <http://papers.nips.cc/paper/6738-adaptive-stimulus-selection-for-optimizing-neural-population-responses.pdf>.
- Cross, Logan et al. (2021). "Using deep reinforcement learning to reveal how the brain encodes abstract state-space representations in high-dimensional environments". In: *Neuron*, pp. 1–15. ISSN: 10974199. DOI: [10.1016/j.neuron.2020.11.021](https://doi.org/10.1016/j.neuron.2020.11.021). URL: <https://doi.org/10.1016/j.neuron.2020.11.021>.
- Denfield, George H et al. (2018). "Attentional fluctuations induce shared variability in macaque primary visual cortex". In: *Nature Communications* 9 (1). ISSN: 20411723. DOI: [10.1038/s41467-018-05123-6](https://doi.org/10.1038/s41467-018-05123-6).
- Ding, Long (2015). "Distinct dynamics of ramping activity in the frontal cortex and caudate nucleus in monkeys". In: *Journal of Neurophysiology* 114 (3), pp. 1850–1861. ISSN: 15221598. DOI: [10.1152/jn.00395.2015](https://doi.org/10.1152/jn.00395.2015).
- Ecker, Alexander S. et al. (Feb. 2016). "On the structure of neuronal population activity under fluctuations in attentional state". In: *Journal of Neuroscience* 36 (5), pp. 1775–1789. ISSN: 15292401. DOI: [10.1523/JNEUROSCI.2044-15.2016](https://doi.org/10.1523/JNEUROSCI.2044-15.2016).
- Faisal, A. Aldo, Luc P.J. Selen, and Daniel M. Wolpert (Apr. 2008). *Noise in the nervous system*. DOI: [10.1038/nrn2258](https://doi.org/10.1038/nrn2258).
- Felleman, Daniel J and David C Van Essen (1991). "Distributed hierarchical processing in the primate cerebral cortex". In: *Cerebral Cortex* 1 (1), pp. 1–47. ISSN: 14602199. DOI: [10.1093/cercor/1.1.1](https://doi.org/10.1093/cercor/1.1.1).

- Felleman, Daniel J, Youping Xiao, and Evelyn McClendon (1997). "Modular organization of occipito-temporal pathways: Cortical connections between visual area 4 and visual area 2 and posterior inferotemporal ventral area in macaque monkeys". In: *Journal of Neuroscience* 17 (9), pp. 3185–3200. ISSN: 02706474. DOI: [10.1523/jneurosci.17-09-03185.1997](https://doi.org/10.1523/jneurosci.17-09-03185.1997).
- Field, Greg D et al. (2010). "Functional connectivity in the retina at the resolution of photoreceptors". In: *Nature* 467 (7316), pp. 673–677. ISSN: 00280836. DOI: [10.1038/nature09424](https://doi.org/10.1038/nature09424). URL: <http://dx.doi.org/10.1038/nature09424>.
- Fries, Pascal (2005). "A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence". In: *Trends in Cognitive Sciences* 9 (10), pp. 474–480. ISSN: 13646613. DOI: [10.1016/j.tics.2005.08.011](https://doi.org/10.1016/j.tics.2005.08.011).
- Fruend, Ingo and Elee Stalker (2018). "Human sensitivity to perturbations constrained by a model of the natural image manifold". In: *Journal of Vision* 18 (11), p. 20. ISSN: 1534-7362. DOI: [10.1167/18.11.20](https://doi.org/10.1167/18.11.20). URL: <http://jov.arvojournals.org/article.aspx?doi=10.1167/18.11.20>.
- Gegenfurtner, Karl, Daniel Kiper, and Fenstemaker Suzanne (1996). "Processing of color, form and motion in macaque area V2". In.
- Gilbert, Charles D and Wu Li (2013). "Top-down influences on visual processing". In: *Nature Reviews Neuroscience* 14 (5), pp. 350–363. ISSN: 1471003X. DOI: [10.1038/nrn3476](https://doi.org/10.1038/nrn3476).
- Hatsopoulos, N G et al. (1998). "Information about movement direction obtained from synchronous activity of motor cortical neurons". In: *Proceedings of the National Academy of Sciences of the United States of America* 95 (26), pp. 15706–15711. ISSN: 00278424. DOI: [10.1073/pnas.95.26.15706](https://doi.org/10.1073/pnas.95.26.15706).
- Hubel, David and T Wiesel (1959). "Receptive Fields of Single Neurones in the Cat's Striate Cortex". In: *J. Physiol.*, pp. 574–591.

- Hubel, David H and Torsten N Wiesel (1965). "Receptive Fields and Functional Architecture in Two Nonstriate Visual Areas (18 and 19) of the Cat". In: *Journal of Neurophysiology* 28 (2), pp. 229–289. ISSN: 0022-3077. DOI: [10.1152/jn.1965.28.2.229](https://doi.org/10.1152/jn.1965.28.2.229). URL: <http://www.physiology.org/doi/10.1152/jn.1965.28.2.229>.
- Hussar, Cory and Tatiana Pasternak (2010). "Trial-to-trial variability of the prefrontal neurons reveals the nature of their engagement in a motion discrimination task". In: *Proceedings of the National Academy of Sciences of the United States of America* 107 (50), pp. 21842–21847. ISSN: 00278424. DOI: [10.1073/pnas.1009956107](https://doi.org/10.1073/pnas.1009956107).
- Hussar, Cory R and Tatiana Pasternak (2009). "Flexibility of Sensory Representations in Prefrontal Cortex Depends on Cell Type". In: *Neuron* 64 (5), pp. 730–743. ISSN: 08966273. DOI: [10.1016/j.neuron.2009.11.018](https://doi.org/10.1016/j.neuron.2009.11.018). URL: <http://dx.doi.org/10.1016/j.neuron.2009.11.018>.
- (2012). "Memory-guided sensory comparisons in the prefrontal cortex: Contribution of putative pyramidal cells and interneurons". In: *Journal of Neuroscience* 32 (8), pp. 2747–2761. ISSN: 02706474. DOI: [10.1523/JNEUROSCI.5135-11.2012](https://doi.org/10.1523/JNEUROSCI.5135-11.2012).
- Hussar, Cory R. and Tatiana Pasternak (Jan. 2013). "Common rules guide comparisons of speed and direction of motion in the dorsolateral prefrontal cortex". In: *Journal of Neuroscience* 33 (3), pp. 972–986. ISSN: 02706474. DOI: [10.1523/JNEUROSCI.4075-12.2013](https://doi.org/10.1523/JNEUROSCI.4075-12.2013).
- Jun, James J. et al. (Nov. 2017). "Fully integrated silicon probes for high-density recording of neural activity". In: *Nature* 551 (7679), pp. 232–236. ISSN: 14764687. DOI: [10.1038/nature24636](https://doi.org/10.1038/nature24636).
- Karras, Tero, Samuli Laine, and Timo Aila (2019). "A style-based generator architecture for generative adversarial networks". In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 2019-June, pp. 4396–4405. ISSN: 10636919. DOI: [10.1109/CVPR.2019.00453](https://doi.org/10.1109/CVPR.2019.00453).
- Katoh, Sourabh, Sumit Singh Chauhan, and Vijay Kumar (Feb. 2021). "A review on genetic algorithm: past, present, and future". In: *Multimedia Tools and Applications* 80 (5), pp. 8091–8126. ISSN: 15737721. DOI: [10.1007/s11042-020-10139-6](https://doi.org/10.1007/s11042-020-10139-6).

- Katz, Leor N et al. (2016). "Dissociated functional significance of decision-related activity in the primate dorsal stream". In: *Nature* 535 (7611), pp. 285–288. ISSN: 14764687. DOI: [10.1038/nature18617](https://doi.org/10.1038/nature18617).
- Kelly, Ryan C. et al. (Jan. 2007). "Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex". In: *Journal of Neuroscience* 27 (2), pp. 261–264. ISSN: 02706474. DOI: [10.1523/JNEUROSCI.4906-06.2007](https://doi.org/10.1523/JNEUROSCI.4906-06.2007).
- Kim, Taekjun, Wyeth Bair, and Anitha Pasupathy (2019). "Neural coding for shape and texture in macaque area V4". In: *Journal of Neuroscience* 39 (24), pp. 4760–4774. ISSN: 15292401. DOI: [10.1523/JNEUROSCI.3073-18.2019](https://doi.org/10.1523/JNEUROSCI.3073-18.2019).
- Kohn, Adam and J Anthony Movshon (2003). "Neuronal adaptation to visual motion in area MT of the macaque". In: *Neuron* 39 (4), pp. 681–691. ISSN: 08966273. DOI: [10.1016/S0896-6273\(03\)00438-0](https://doi.org/10.1016/S0896-6273(03)00438-0).
- Kole, Maarten H.P. and Greg J. Stuart (Jan. 2012). *Signal Processing in the Axon Initial Segment*. DOI: [10.1016/j.neuron.2012.01.007](https://doi.org/10.1016/j.neuron.2012.01.007).
- Kriegeskorte, Nikolaus and Xue Xin Wei (2021). "Neural tuning and representational geometry". In: *Nature Reviews Neuroscience*. ISSN: 14710048. DOI: [10.1038/s41583-021-00502-3](https://doi.org/10.1038/s41583-021-00502-3).
- Lennie, P, J Krauskopf, and G Sclar (1990). "Chromatic mechanisms in striate cortex of macaque". In: *Journal of Neuroscience* 10 (2), pp. 649–669. ISSN: 02706474. DOI: [10.1523/jneurosci.10-02-00649.1990](https://doi.org/10.1523/jneurosci.10-02-00649.1990).
- Liu, Ye et al. (2020). "Hierarchical Representation for Chromatic Processing across Macaque V1, V2, and V4". In: *Neuron* 108 (3), 538–550.e5. ISSN: 10974199. DOI: [10.1016/j.neuron.2020.07.037](https://doi.org/10.1016/j.neuron.2020.07.037).
- Luck, Steven J et al. (1997). "Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex". In: *Journal of Neurophysiology* 77 (1), pp. 24–42. ISSN: 00223077. DOI: [10.1152/jn.1997.77.1.24](https://doi.org/10.1152/jn.1997.77.1.24). URL: <http://www.physiology.org/doi/10.1152/jn.1997.77.1.24>.

- Lui, Leo L and Tatiana Pasternak (2011). "Representation of comparison signals in cortical area MT during a delayed direction discrimination task". In: *Journal of Neurophysiology* 106 (3), pp. 1260–1273. ISSN: 00223077. DOI: [10.1152/jn.00016.2011](https://doi.org/10.1152/jn.00016.2011).
- Mahon, L E and R L D E Valois (2001). "Cartesian and non-Cartesian responses in LGN , V1 , and V2 cells". In: (2001), pp. 973–981.
- Mirjalili, Seyedali (2019). *Genetic Algorithm*. DOI: [10.1007/978-3-319-93025-1_4](https://doi.org/10.1007/978-3-319-93025-1_4). URL: https://doi.org/10.1007/978-3-319-93025-1_4.
- Mitchell, Jude F, Kristy A Sundberg, and John H Reynolds (2007). "Differential Attention-Dependent Response Modulation across Cell Classes in Macaque Visual Area V4". In: *Neuron* 55 (1), pp. 131–141. ISSN: 08966273. DOI: [10.1016/j.neuron.2007.06.018](https://doi.org/10.1016/j.neuron.2007.06.018).
- Moreno-Bote, Rubén et al. (Oct. 2014). "Information-limiting correlations". In: *Nature Neuroscience* 17 (10), pp. 1410–1417. ISSN: 15461726. DOI: [10.1038/nn.3807](https://doi.org/10.1038/nn.3807).
- Movshon, J Anthony and William T Newsome (1996). *Visual Response Properties of Striate Cortical Neurons Projecting to Area MT in Macaque Monkeys*, pp. 7733–7741.
- Nandy, Anirvan S et al. (2016). "Neurons in Macaque Area V4 Are Tuned for Complex Spatio-Temporal Patterns". In: *Neuron* 91 (4), pp. 920–930. ISSN: 0896-6273. DOI: [10.1016/j.neuron.2016.07.026](https://doi.org/10.1016/j.neuron.2016.07.026). URL: <http://dx.doi.org/10.1016/j.neuron.2016.07.026>.
- Narayanan, Nandakumar S (2016). "Ramping activity is a cortical mechanism of temporal control of action". In: *Current Opinion in Behavioral Sciences* 8, pp. 226–230. ISSN: 23521546. DOI: [10.1016/j.cobeha.2016.02.017](https://doi.org/10.1016/j.cobeha.2016.02.017).
- Nigam, Sunny, Sorin Pojoga, and Valentin Dragoi (2021). "A distinct population of heterogeneously color-tuned neurons in macaque visual cortex". In: *Science Advances* 7 (8). ISSN: 23752548. DOI: [10.1126/sciadv.abc5837](https://doi.org/10.1126/sciadv.abc5837).
- Paninski, Liam et al. (2010). "A new look at state-space models for neural data". In: *Journal of Computational Neuroscience* 29 (1-2), pp. 107–126. ISSN: 09295313. DOI: [10.1007/s10827-009-0179-x](https://doi.org/10.1007/s10827-009-0179-x).

- Pasternak, Tatiana and Duje Tadin (2020). "Linking Neuronal Direction Selectivity to Perceptual Decisions about Visual Motion". In: *Annual Review of Vision Science* 6, pp. 335–362. ISSN: 23744650. DOI: [10.1146/annurev-vision-121219-081816](https://doi.org/10.1146/annurev-vision-121219-081816).
- Pasupathy, Anitha and Charles E Connor (2002). "Population coding of shape in area V4". In: *Nature Neuroscience* 5 (12), pp. 1332–1338. ISSN: 10976256. DOI: [10.1038/nn972](https://doi.org/10.1038/nn972).
- Petrides, M and D N Pandya (2006). "Efferent Association Pathways Originating in the Caudal Prefrontal Cortex in the Macaque Monkey". In: *Journal of Comparative Neurology* 498 (October 2007), pp. 227–251. DOI: [10.1002/cne](https://doi.org/10.1002/cne).
- Ponce, Carlos R et al. (2019). "Evolving Images for Visual Neurons Using a Deep Generative Network Reveals Coding Principles and Neuronal Preferences". In: *Cell* 177 (4), 999–1009.e10. ISSN: 10974172. DOI: [10.1016/j.cell.2019.04.005](https://doi.org/10.1016/j.cell.2019.04.005). URL: <http://dx.doi.org/10.1016/j.cell.2019.04.005>.
- Ponce-Alvarez, Adrián et al. (Aug. 2013). "Stimulus-dependent variability and noise correlations in cortical MT neurons". In: *Proceedings of the National Academy of Sciences of the United States of America* 110 (32), pp. 13162–13167. ISSN: 00278424. DOI: [10.1073/pnas.1300098110](https://doi.org/10.1073/pnas.1300098110).
- Prenger, Ryan et al. (June 2004). "Nonlinear V1 responses to natural scenes revealed by neural network analysis". In: *Neural Networks* 17 (5-6), pp. 663–679. ISSN: 08936080. DOI: [10.1016/j.neunet.2004.03.008](https://doi.org/10.1016/j.neunet.2004.03.008).
- Quiroga, Rodrigo Quian and Stefano Panzeri (Mar. 2009). *Extracting information from neuronal populations: Information theory and decoding approaches*. DOI: [10.1038/nrn2578](https://doi.org/10.1038/nrn2578).
- Rajdl, K, P Lansky, and L Kostal (2017). "Entropy factor for randomness quantification in neuronal data". In: *Neural Networks* 95, pp. 57–65. ISSN: 18792782. DOI: [10.1016/j.neunet.2017.07.016](https://doi.org/10.1016/j.neunet.2017.07.016).
- Renart, Alfonso and Christian K Machens (2014). "Variability in neural activity and behavior". In: *Current Opinion in Neurobiology* 25, pp. 211–220. ISSN: 18736882. DOI: [10.1016/j.conb.2014.02.013](https://doi.org/10.1016/j.conb.2014.02.013).

- Rudolph, Kirsten and Tatiana Pasternak (1999). "Transient and permanent deficits in motion perception after lesions of cortical areas MT and MST in the macaque monkey". In: *Cerebral Cortex* 9 (1), pp. 90–100. ISSN: 10473211. DOI: [10.1093/cercor/9.1.90](https://doi.org/10.1093/cercor/9.1.90).
- Ruff, Douglas A and Marlene R Cohen (2016). "Attention Increases Spike Count Correlations between Visual Cortical Areas". In: 36 (28), pp. 7523–7534. DOI: [10.1523/JNEUROSCI.0610-16.2016](https://doi.org/10.1523/JNEUROSCI.0610-16.2016).
- Rust, Nicole C et al. (2006). "How MT cells analyze the motion of visual patterns". In: *Nature Neuroscience* 9 (11), pp. 1421–1431. ISSN: 10976256. DOI: [10.1038/nn1786](https://doi.org/10.1038/nn1786).
- Salzman, C D et al. (1992). "Microstimulation in visual area MT: Effects on direction discrimination performance". In: *Journal of Neuroscience* 12 (6), pp. 2331–2355. ISSN: 02706474. DOI: [10.1523/jneurosci.12-06-02331.1992](https://doi.org/10.1523/jneurosci.12-06-02331.1992).
- Salzman, C Daniel, Kenneth H Britten, and William T Newsome (1990). "Cortical microstimulation influences perceptual judgements of motion direction". In: *Nature* 346 (6280), pp. 174–177. ISSN: 00280836. DOI: [10.1038/346174a0](https://doi.org/10.1038/346174a0).
- Sani, Ilaria et al. (2013). "Selective Tuning for Contrast in Macaque Area V4". In: 33 (47), pp. 18583–18596. DOI: [10.1523/JNEUROSCI.3465-13.2013](https://doi.org/10.1523/JNEUROSCI.3465-13.2013).
- SCOTT, DAVID W (Dec. 1979). "On optimal and data-based histograms". In: *Biometrika* 66 (3), pp. 605–610. ISSN: 0006-3444. DOI: [10.1093/biomet/66.3.605](https://doi.org/10.1093/biomet/66.3.605). URL: <https://doi.org/10.1093/biomet/66.3.605>.
- Scott, Hayden et al. (Apr. 2023). "Altered Task Demands Lead to a Division of Labor for Sensory and Cognitive Processing in the Middle Temporal Area". In: *European Journal of Neuroscience*. ISSN: 0953-816X. DOI: [10.1111/ejn.15964](https://doi.org/10.1111/ejn.15964).
- Seriès, Peggy, Peter E. Latham, and Alexandre Pouget (Oct. 2004). "Tuning curve sharpening for orientation selectivity: Coding efficiency and the impact of correlations". In: *Nature Neuroscience* 7 (10), pp. 1129–1135. ISSN: 10976256. DOI: [10.1038/nn1321](https://doi.org/10.1038/nn1321).

- Shadlen, Michael N and William T Newsome (2001). "Neural Basis of a Perceptual Decision in the Parietal Cortex (Area LIP) of the Rhesus Monkey". In: *Journal of Neurophysiology* 86 (4), pp. 1916–1936. ISSN: 0022-3077. DOI: [10.1152/jn.2001.86.4.1916](https://doi.org/10.1152/jn.2001.86.4.1916). URL: <http://www.physiology.org/doi/10.1152/jn.2001.86.4.1916>.
- Shen, Cencheng, Sambit Panda, and Joshua T. Vogelstein (2022). "The Chi-Square Test of Distance Correlation". In: *Journal of Computational and Graphical Statistics* 31 (1), pp. 254–262. ISSN: 15372715. DOI: [10.1080/10618600.2021.1938585](https://doi.org/10.1080/10618600.2021.1938585).
- Snyder, A C et al. (2014). "Correlations in V1 are reduced by stimulation outside the receptive field". In: *Journal of Neuroscience* 34 (34). ISSN: 15292401. DOI: [10.1523/JNEUROSCI.0762-14.2014](https://doi.org/10.1523/JNEUROSCI.0762-14.2014).
- Snyder, Adam C, Byron M Yu, and Matthew A Smith (2018). "Distinct population codes for attention in the absence and presence of visual stimulation". In: *Nature Communications* 9 (1). ISSN: 20411723. DOI: [10.1038/s41467-018-06754-5](https://doi.org/10.1038/s41467-018-06754-5).
- Snyder, Adam C., Byron M. Yu, and Matthew A. Smith (Nov. 2021). "A stable population code for attention in prefrontal cortex leads a dynamic attention code in visual cortex". In: *Journal of Neuroscience* 41 (44), pp. 9163–9176. ISSN: 15292401. DOI: [10.1523/JNEUROSCI.0608-21.2021](https://doi.org/10.1523/JNEUROSCI.0608-21.2021).
- Steinmetz, Nicholas A. et al. (Apr. 2021). "Neuropixels 2.0: A miniaturized high-density probe for stable, long-term brain recordings". In: *Science* 372 (6539). ISSN: 10959203. DOI: [10.1126/science.abf4588](https://doi.org/10.1126/science.abf4588).
- Tanigawa, Hisashi, Haidong D Lu, and Anna W Roe (2010). "Functional organization for color and orientation in macaque V4". In: *Nature Neuroscience* 13 (12), pp. 1542–1549. ISSN: 10976256. DOI: [10.1038/nn.2676](https://doi.org/10.1038/nn.2676).
- Touryan, Jon and James A. Mazer (May 2015). "Linear and non-linear properties of feature selectivity in V4 neurons". In: *Frontiers in Systems Neuroscience* 9 (MAY), pp. 1–12. ISSN: 16625137. DOI: [10.3389/fnsys.2015.00082](https://doi.org/10.3389/fnsys.2015.00082).
- Ungerleider, Leslie G and Robert Desimone (1986). "Cortical connections of visual area MT in the macaque". In: *Journal of Comparative Neurology* 248 (2), pp. 190–222. ISSN: 10969861. DOI: [10.1002/cne.902480204](https://doi.org/10.1002/cne.902480204).

- Whitley, Darrell (1994). *A genetic algorithm tutorial*, pp. 65–85.
- Wiesel, Torsten N (1959). “Recording Inhibition and Excitation in the Cat’s Retinal Ganglion Cells with intracellular Electrodes”. In: *Nature* 1, pp. 21–22.
- Wimmer, K et al. (2016). “Transitions between Multiband Oscillatory Patterns Characterize Memory-Guided Perceptual Decisions in Prefrontal Circuits”. In: *Journal of Neuroscience* 36 (2), pp. 489–505. ISSN: 0270-6474. DOI: [10.1523/JNEUROSCI.3678-15.2016](https://doi.org/10.1523/JNEUROSCI.3678-15.2016). URL: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.3678-15.2016>.
- Yoshioka, Takashi, Bruce M Dow, and Robert G Vautin (1996). “Neuronal mechanisms of color categorization in areas V1, V2 and V4 of macaque monkey visual cortex”. In: *Behavioural Brain Research* 76 (1-2), pp. 51–70. ISSN: 01664328. DOI: [10.1016/0166-4328\(95\)00183-2](https://doi.org/10.1016/0166-4328(95)00183-2).
- Yu, Byron M et al. (2009). “Gaussian-process factor analysis for low-dimensional single-trial analysis of neural population activity”. In: *Journal of Neurophysiology* 102 (1), pp. 614–635. ISSN: 15221598. DOI: [10.1152/jn.90941.2008](https://doi.org/10.1152/jn.90941.2008).
- Zaksas, Daniel and Tatiana Pasternak (2006). “Directional signals in the prefrontal cortex and in area MT during a working memory for visual motion task”. In: *Journal of Neuroscience* 26 (45), pp. 11726–11742. ISSN: 02706474. DOI: [10.1523/JNEUROSCI.3420-06.2006](https://doi.org/10.1523/JNEUROSCI.3420-06.2006).