Synaptic diversity naturally arises from neural decoding of heterogeneous populations

Jacob L. Yates[1], Benjamin Scholl[2]

[1]University of Maryland, College Park, Maryland, USA
[2]University of Pennsylvania, Department of Neuroscience, Perelman School of Medicine, Philadelphia, PA, USA

Running title: Synaptic diversity results from heterogeneity

Corresponding author: Benjamin Scholl, benjamin.scholl@pennmedicine.upenn.edu

Number of pages: 25
Number of figures: 5
Number of tables: 0
Number of words Abstract: 198
Number of words Introduction: 675
Number of words Results: 3017
Number of words Discussion: 677

Conflict of Interest: The authors declare no competing financial interests.

Author contributions statement: J.Y and B.S. conceived of the experiment. J.Y. derived analytic solutions, J.Y and B.S. preformed analysis, J.Y and B.S. wrote the manuscript.

Abstract

The synaptic inputs to single cortical neurons exhibit substantial diversity in their sensory-driven activity. What this diversity reflects is unclear, and appears counter-productive in generating selective somatic responses to specific stimuli. We propose that synaptic diversity arises because neurons decode information from upstream populations. Focusing on a single sensory variable, orientation, we construct a probabilistic decoder that estimates the stimulus orientation from the responses of a realistic, hypothetical input population of neurons. We provide a straightforward mapping from the decoder weights to real excitatory synapses, and find that optimal decoding requires diverse input weights. Analytically derived weights exhibit diversity whenever upstream input populations consist of noisy, correlated, and heterogeneous neurons, as is typically found *in vivo*. In fact, in silico weight diversity was necessary to accurately decode orientation and matched the functional heterogeneity of dendritic spines imaged in vivo. Our results indicate that synaptic diversity is a necessary component of information transmission and reframes studies of connectivity through the lens of probabilistic population codes. These results suggest that the mapping from synaptic inputs to somatic selectivity may not be directly interpretable without considering input covariance and highlights the importance of population codes in pursuit of the cortical connectome.

Introduction

Cortical neurons are driven by large populations of excitatory synaptic inputs. Synaptic populations ultimately shape how sensory signals are encoded, decoded, or transformed. The sensory representation or functional properties of an excitatory input population will define and constrain the operations a neuron can perform and reflects the rules neurons use to form connections. Electrophysiological and anatomical studies suggest that connections between excitatory neurons exhibits functional specificity, where inputs are tuned for similar features as the soma (Cossell et al., 2015; Ko et al., 2011; Lee et al., 2016; Reid and Alonso, 1995). In contrast, synaptic imaging techniques have revealed that synaptic populations exhibit functional diversity, deviating from canonical connectivity rules, such as 'like-connects-to-like' (Scholl and Fitzpatrick, 2020). This functional diversity within input populations has been observed in a variety of mammalian species, from rodents to primates, and for a variety of sensory cortical areas (Chen et al., 2013, 2011; Iacaruso et al., 2017; Jia et al., 2011, 2010; Ju et al., 2020; Kerlin et al., 2019;

2

34 Scholl et al., 2017; Wertz et al., 2015; Wilson et al., 2018, 2016). This apparent discrepancy

35 challenges our understanding about how synaptic inputs drive the selective outputs of cortical

36 neurons and leads to a simple fundamental question: If the goal is to produce selective somatic

37 responses, why would a neuron have excitatory synaptic inputs tuned far away from the somatic

38 preference?

39 To answer this question, we turn to population coding theory; starting with the idea that to

40 accurately represent sensory signals, cortical neurons must decode the activity of upstream

41 populations. This decoding is likely accomplished by combing signals across neural populations

42 (Graf et al., 2011; Jazayeri and Movshon, 2006). Many studies have examined how sensory

43 variables might be decoded from cortical populations (Butts and Goldman, 2006; Graf et al., 2011;

44 Shamir and Sompolinsky, 2006), an endeavor increasingly applied to larger population sizes with

45 innovative recording techniques (Rumyantsev et al., 2020; Stringer et al., 2019). These decoding

46 approaches are often used as a tool to quantify the information about a stimulus available in a

47 neural population, carrying the assumption that downstream areas could perform such a process

48 (Berens et al., 2011; DiCarlo et al., 2012). In real brain circuits, decoders must be composed of

49 individual neurons, driven by sets of synaptic inputs, akin to a decoder's weights over a given

50 input population. To date, few studies have explicitly examined the weight structure of population

51 decoders (Jazayeri and Movshon, 2006; Rust et al., 2006; Zavitz and Price, 2019).

52 In this paper, we investigate the weights of a simple population decoder and how they

53 compare to real synaptic inputs measured *in vivo*. Focusing on a single sensory variable,

54 orientation, we derive the maximum-likelihood readout for a simulated input population that

55 encodes stimuli with noisy tuning curves (e.g., Ecker et al., 2011). Under reasonable assumptions,

56 the decoder weights can be interpreted as the synaptic connectivity between the input population

57 and the downstream decoder neurons. This allows us to examine how synaptic connectivity

58 depends on properties of the input population and to directly compare population decoders to

59 synaptic input measured *in vivo*. We then test a hypothesis that an optimal decoder will show

60 substantial heterogeneity in its synaptic weights given a biologically realistic input population. We

61 find that when input populations are shifted copies of the same tuning curve, the synaptic

62 excitatory inputs closely resemble the somatic output. However, with a biologically realistic input

63 population, the expected inputs onto readout neurons exhibit functional diversity. We then

64 compare the orientation tuning of simulated inputs with large populations of dendritic spines

65 (excitatory synaptic inputs) onto individual neurons of ferret primary visual cortex (V1), recorded

66 with two-photon calcium imaging *in vivo*. This comparison revealed similar diversity in the

67 orientation tuning of dendritic spines on ferret V1 neurons and simulated decoder weights. The

3

68 similarity between the synaptic populations of actual V1 neurons and the optimal neural decoder

69 suggest that diversity and heterogeneity observed in dendritic spines across sensory cortices are,

70 in fact, expected when considering how information is propagated through neural circuits in the

71 presence of noise.

72

73 **Results**

74

75 Following several decades of work on population coding theory, we derive a Bayesian

76 decoder to report the probability of a visual stimulus given inputs from a neural population (Fig.

77 1). With this framework, and given the specifics of the *encoding* population, we can analytically

78 derive the optimal *decoding* weights of a population of readout neurons. Here, we use "optimal"

79 to refer to the maximum-likelihood solution. Previous work has shown that a population of neurons
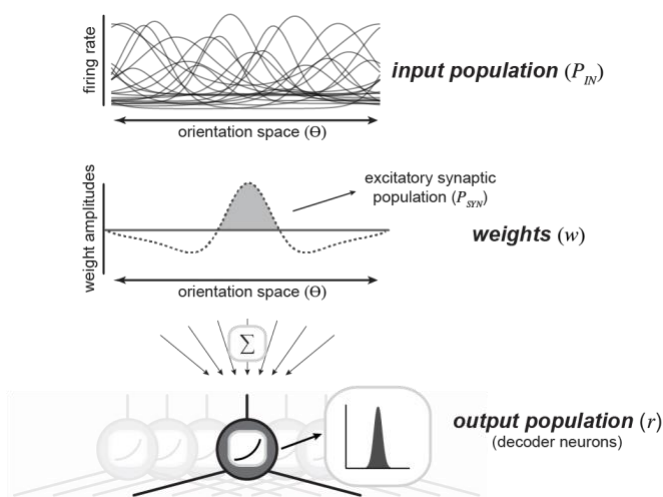

Figure 1: A population decoding framework to study synaptic diversity. An upstream population of neurons are tuned for a single stimulus variable (orientation) (*top*). This input population is readout by downstream decoder neurons (*bottom*). Downstream neurons decode stimulus identify by reading out spikes from the upstream input population. Each decoder neuron is defined by a set weights (*middle*) over the upstream population, which are summed and rectified to produce an output.

80 could perform such probabilistic decoding with weighted summation and divisive normalization,

81 as long as their inputs exhibit Poisson-like noise (Jazayeri and Movshon, 2006; Ma et al., 2006).

82 Starting from that basic framework, we derived a decoder that represents the probability that each

83 possible stimulus orientation was present given the responses of a large population of upstream,

84 input neurons (P_{IN}). This is effectively a categorical decoder, where each possible orientation is a

85 different category. Similar decoders have been used throughout the literature to estimate how

86 much information is in a neural recording and suggest how downstream neurons might read it out

87 (Graf et al., 2011; Stringer et al., 2019). Our decoder has weight vectors for each possible stimulus

4

88 orientation, which integrate across P_{IN} and are passed through a static nonlinearity (the
89 exponential function) and normalized. As we will show below, given specific assumptions about
90 the variability in P_{IN}, the weights over P_{IN} depend systematically on the tuning functions and
91 covariance of P_{IN}. Following a characterization of this decoding framework, we will make direct
92 comparisons with real data: defining an effective "synaptic input population" (P_{SYN}) as nonzero,
93 positive weights over P_{IN}. Although our strategy applies to any one-dimensional stimulus variable,
94 we describe this model in the context of orientation of drifting gratings presented to V1 neurons
95 for a direct comparison with *in vivo* measurements.

96
97 *A neural population as a probabilistic decoder*
98
99   A categorical probabilistic decoder reports the probability that a particular stimulus
100 orientation, θ_k, was present given the spiking responses of an input population, $\boldsymbol{R}$. This can be
101 expressed as a normalized exponential function of the log-likelihood plus the log prior for each θ_k,

$$p(\theta_k|R) = \frac{p(R|\theta_k)p(\theta_k)}{\sum_i p(R|\theta_i)\,p(\theta_i)} = \frac{e^{L(\theta_k)}}{\sum_i e^{L(\theta_i)}}$$

103 where

$$L(\theta_k) = ln\big(p(R|\theta_k)\big) + ln\big(p(\theta_k)\big)$$

105 The likelihood, $p(\boldsymbol{R}\,|\,\theta_k)$, is the probability of the observed responses in an input population given
106 the stimulus k and $p(\theta_k)$ is the prior probability of that stimulus class. If $p(\boldsymbol{R}\,|\,\theta_k)$ is in the exponential
107 family, then $L(\theta_k)$ can be written as a weighted sum of the input population response vector plus
108 an offset, which can be estimated numerically via multinomial logistic regression (Ma et al., 2006).
109 For simplicity, we assume the input population has a response that is a function of the stimulus
110 plus Gaussian noise, and equal covariance across all stimulus conditions. Although this
111 assumption about the covariance structure deviates from real neural activity, this assumption
112 means the weights and offset can be solved analytically (see Methods), and as will be shown
113 below, such a simple model makes substantial headway in explaining biological phenomena. Our
114 goal here is to provide a plausible alternative to "somatic selectivity" for the connectivity rules in
115 cortex. Under the Gaussian assumption, the decoder amounts to:

$$p(\theta_k|R) = \frac{e^{(\boldsymbol{R}^T w_k + \beta_k)}}{\sum_{i=1}^{K} e^{(\boldsymbol{R}^T w_i + \beta_i)}} \qquad (1)$$

117 where

$$w_k = Q^{-1}f(\theta_k)$$

5

$$\beta_k = -\left(\frac{1}{2}\right) f(\theta_k)^T Q^{-1} f(\theta_k) + \ln p(\theta_k)$$

Here, $f(\theta_k)$ is the mean input population response to stimulus orientation, k, K is the total number of orientations, and Q is the covariance matrix. The covariance term captures the influence of each neuron's response variance (diagonal elements) and the variability shared with other neurons (off-diagonal elements). Intuitively, in the absence of covariability (i.e., off-diagonal elements are zero), the weights are proportional to the signal-to-noise ratio of the neuron (the mean divided by the variance). The term $\boldsymbol{R}^T w_k$ is the dot product between the population response and weights. The second term, β_k, is an offset for each stimulus. $\ln p(\theta_k)$ is a constant reflecting the log prior probability of stimulus k. It is worth noting that if the covariance depends on the stimulus, the optimal readout is no longer a linear function of $\boldsymbol{R}$ and is quadratic, which can be interpreted as a complex-cell (Jaini and Burge, 2017; Pagan et al., 2016) and is a potentially fruitful future direction.

In this study, we focus on the weights of this simple Gaussian, equal covariance decoder in order to examine how synaptic tuning from such a simple decoder would arise. Because the optimal weights have an analytic solution (eq. 1), we can see how they depend on the parameters of P_{IN}. The simplifying assumptions we use to derive the maximum-likelihood weights help build intuitions about what can be expected in biological circuits, and linear weights such as these could be learned by real neural systems (Dayan and Abbott, 2001). A key difference here from prior work is that rather than focus on discrimination (Haefner et al., 2013), we treat orientation estimation as a multiclass identification problem, discretizing θ such that for each possible θ_k, there is a separate weight vector. Thus, in this derivation, the optimal weights depend on the tuning curves themselves, not the derivative.

Characteristics of a neural population decoder

To understand how synaptic weights depend on input statistics, we derived maximum-likelihood weights for input populations, P_{IN}, with different tuning and covariance. To generate P_{IN} we simulated N neurons responding to K oriented stimuli (θ = [-90°:K/180:+90°]). We briefly describe the construction of P_{IN} here (full details are described in the Methods). Each neuron is defined by a tuning function and noise term, describing trial-by-trial variability, which are summed to generate stimulus-driven responses. We compared two fundamentally different types of input populations that have been used in the literature, homogeneous and heterogenous, as well as the role of correlated variability in shaping readout weights. A homogeneous P_{IN} consists of shifted

6

152   copies of a single tuning curve (Fig. 2a). Heterogenous P_{IN} have diverse tuning functions and

153   were generated to match measurements from macaque V1 (Ringach et al., 2002). The

154   heterogenous P_{IN} consisted of tuning curves closely resembling V1 physiology in terms of the

155   variation in peak firing rate, bandwidth, and baseline firing rate (Fig. 2d). Varying amounts of

156   limited-range correlations were included such that the noise correlation between two neurons
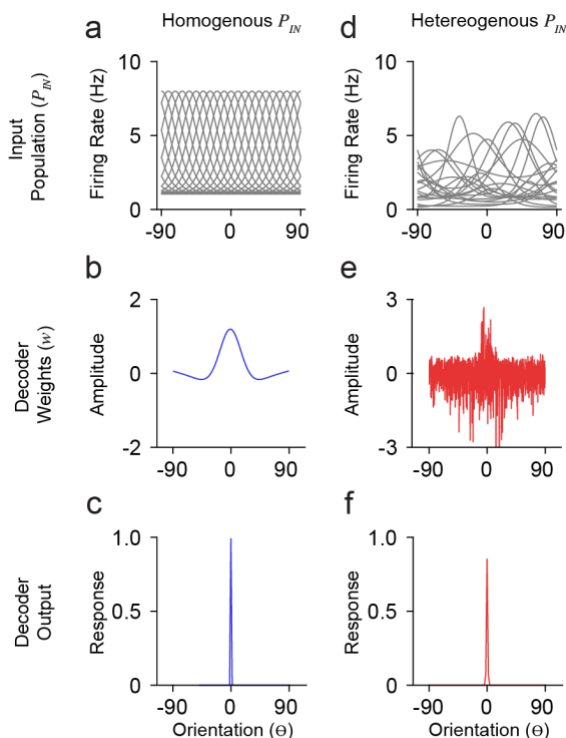

Figure 2: Model simulations with homogenous and heterogeneous input populations.

(a) Orientation tuning of a homogenous input population. Shown is a subset of the total population (n = 20/1000). Ordinate is orientation preference, restricted between -90º and 90º. (b) Derived weights for a single decoder neuron (preferring 0º) reading out the homogenous (*blue*) input population in (a). Weights for homogenous populations smoothly vary over orientation space. (c) Response output of the decoder neuron whose weights are shown in (b). (d-f) Same as in (a-c) for a heterogeneous input population with moderate correlation (c_o = 0.25). Note that decoder weights for heterogeneous input populations are not smooth.

157   depends on the difference in their tuning preferences (Ecker et al., 2011; Kohn et al., 2016).

158   The statistics of P_{IN} responses, **R,** will dictate the weight structure for neurons in a decoding

159   population. For a homogeneous P_{IN}, the weights are smooth across orientation space and exhibit

160   three primary features: a prominent peak about the preferred orientation of the output tuning,

161   slight negative weights for orientations just outside the preferred, and near-zero weights at

162   orthogonal orientations (Fig. 2c). With more realistic tuning diversity (heterogenous P_{IN}), optimal

163   weights are no longer smooth (Fig. 2d-e). While the optimal weights appear to roughly have the

164   same overall shape as for homogeneous P_{IN}, there is considerable positive and negative

7

165   weighting across orientation space. Despite substantial changes in optimal weight vectors, the

166   decoder output (i.e. somatic response) tuning was narrow (Fig. 2f), similar to the output for the

167   homogeneous case (Fig. 2e).

168   To explore the importance of decoding weight diversity, we imposed a smoothing penalty

169   on weight vectors (Park and Pillow, 2011). We calculated cross-validated decoder accuracy using

170   the mean-squared error between the maximum a posteriori estimation and true stimulus (see

171   Methods). Different degrees of smoothing are shown for an example set of weights in Figure 3a.

172   We simulated a range of population sizes ($N = 2$ - 2048) and correlations ($c_o = 0, 0.25, 0.50$).
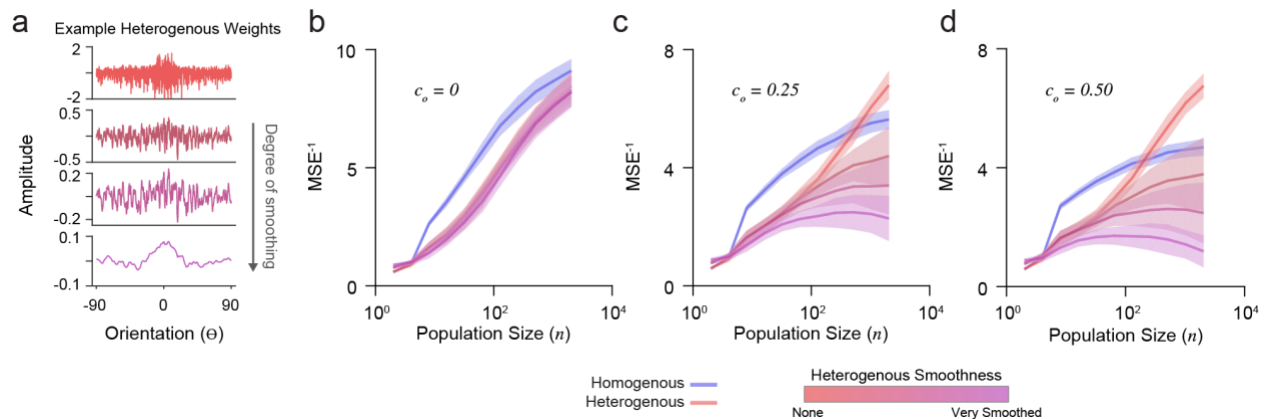

Figure 3: Decoder performance of heterogeneous input populations depends on population size, correlations, and weight diversity. (a) Example weight distribution for a decoder neuron reading out a heterogeneous input population (*top*). Shown are the effects of progressively smoothing weights. Smooth parameters (see Methods) from top to bottom: (0,0), (0.1,1), (0.2, 2), (1,10). Ordinate is orientation preference, restricted between -90º and 90º. (b) Decoder performance (inverse mean-squared-error) plotted for homogenous and heterogeneous input populations of increasing size. Simulations here include no correlations ($c_o = 0$). Shading indicates standard error. (c) Same as in (b) for input populations with moderate correlation ($c_o = 0.25$). (d) Same as in (b) for input populations with stronger correlation ($c_o = 0.50$).

173   Without noise-correlations, the accuracy of all decoders increases with population size,

174   with a homogenous P_{IN} preforming best (Fig. 3b). In the presence of noise-correlations, accuracy

175   saturates for large homogenous P_{IN} (Fig. 3c-d). As previously shown (Ecker et al., 2011), accuracy

176   for heterogeneous populations with limited-range correlations does not saturate (Fig. 3c-d).

177   However, this depends on weight diversity. Smoothing the weights for heterogeneous P_{IN} caused

178   saturation and decreased accuracy (Fig. 3c-d), demonstrating that weight amplitude diversity in

179   analytically derived weights distributions are critical for the decoder performance.

180

181   *Simulating excitatory weight tuning*

182

8

183   In order to compare analytically derived weights with the synaptic inputs onto V1 neurons
184   measured *in vivo*, we generated excitatory synaptic input populations (P_{SYN}). Under the
185   assumption that synaptic integration is linear, two synapses of equal weight are the same as one
186   synapse with double that weight. This creates a degeneracy where synapse count and size trade
187   off. Because current spine imaging techniques typically capture large synapses and there is no
188   relationship between strength and orientation preference (Scholl et al., 2021), we can assume
189   size is fixed and convert the derived weights into a frequency distribution of 'synaptic inputs' (Fig.
190   4a). The tuning curve for such a synapse is the tuning curve of the input and thus, the synaptic
191   input population, $P_{SYN,}$ is the input population resampled with probabilities given by the derived
192   weights. An example P_{SYN} for a single decoder neuron is shown in Figure 4b (drawn from the
193   heterogeneous P_{IN} in Figure 2). P_{SYN} in this example displays some specificity in orientation tuning
194   relative to the somatic output, indicated by a larger proportion of simulated synapses with similar
195   orientation preference as the somatic output (0°) of the decoder neuron.
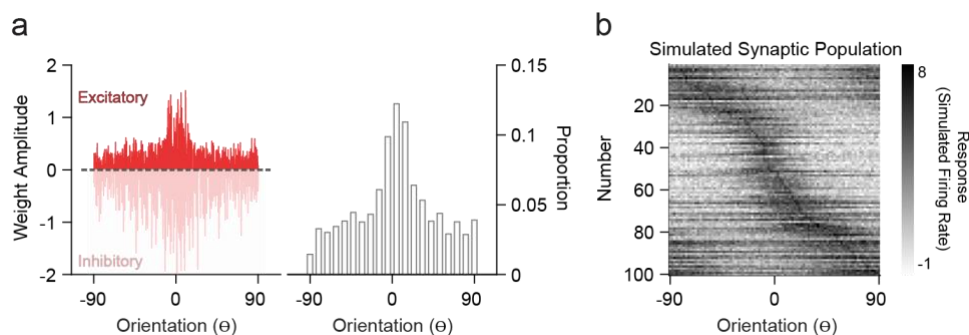

Figure 4: Simulation of synaptic populations from decoder neuron weight distributions. (a) Example weight distribution for a single decoder neuron tuned to 0° (*left*). Ordinates are orientation preference, restricted between -90° and 90°. Dashed line separates excitatory (positive) and inhibitory (negative) weights. Excitatory weight distribution over the input population is transformed into a frequency distribution, whereby greater amplitude equates to greater frequency of occurrence (*right*). (b) Example simulated synaptic population (n = 100 spines) from the weight distribution in (a). Shown are the orientation tuning curves of each simulated synapse (normalized).

196

Empirical distribution of dendritic spine tuning is consistent with decoding of a heterogeneous
input population

199   We analyzed two-photon calcium recordings from soma and corresponding dendritic
200   spines on individual neurons in ferret V1 during the presentation of oriented drifting gratings (see
201   Methods). While our model draws from a P_{IN} matched to measurements from macaque V1, the
202   orientation tuning of layer 2/3 neurons in ferret V1, as measured by two-photon cellular imaging,
203   exhibit a similar range in selectivity (Wilson et al., 2017). Visually responsive and isolated dendritic
204   spines (see Methods) typically exhibit diverse orientation tuning relative to the somatic output,

9

205   although some individual cells show greater overall diversity (Fig. 5b) than others (Fig. 5a). To

206   characterize P_{SYN} diversity, both for real dendritic spines and simulated inputs, we computed the

207   Pearson correlation coefficient between the tuning curves of individual inputs and the

208   corresponding somatic output (Scholl et al., 2021). For these comparisons, we sampled

209   orientation space in the model spines to match our empirical measurements (22.5 deg

210   increments) and the number of total excitatory inputs recovered for each simulated downstream

211   neuron was set to 100, similar to the average number of visually-responsive spines recorded for
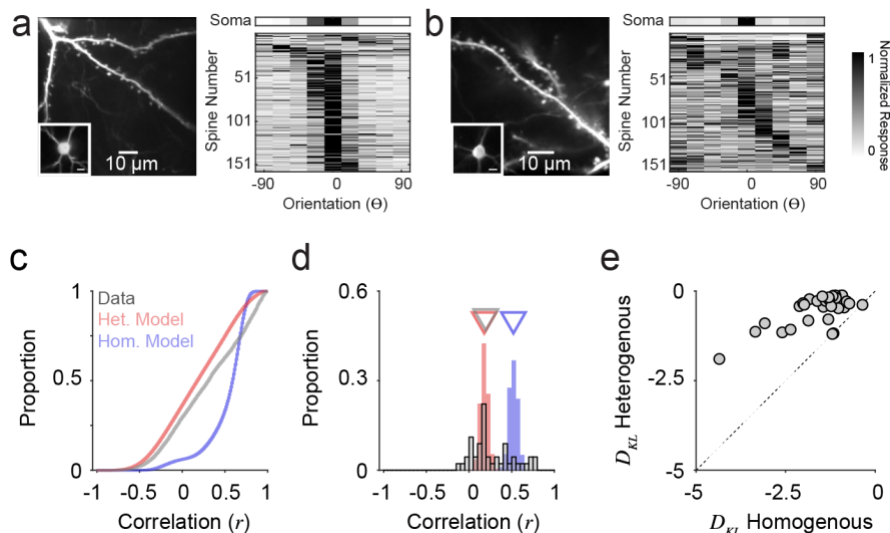

Figure 5: Orientation tuning diversity of dendritic spine populations in ferret V1 match simulations with correlated, heterogeneous input populations. (a) Two-photon standard-deviation projection of example dendrite and spines recorded from a single cell (*left*). Inset: Two-photon standard-deviation projection of corresponding soma. Scale bar is 10 microns. Orientation tuning of soma (*top*) and all visually-responsive dendritic spines from this single cell (n = 159) are shown (*right*). Spine responses are normalized peak ΔF/F. Orientation preferences are shown relative to the somatic preference (aligned to 0°). (b) Same as in (a) for another example cell (n = 162 visually-responsive spines). (c) Cumulative distributions of tuning correlation between individual dendritic spines or simulated synaptic inputs with corresponding somatic tuning or decoder output. Shown are correlations of simulations of homogenous (*blue*) or heterogeneous (*red*) input populations, compared to empirical data (*gray*). (d) Distributions of average tuning correlation between synaptic input and somatic output across measured cells (n = 45). Also shown are distributions of average tuning correlation for simulated cells. Triangles denote median values for each distribution. (e) Comparison of Kullback-Leibler divergence (D_{KL}) between data and each model type. Each data point represents an individual cell's population of dendritic spines.

212   each ferret V1 neuron (n = 45, n = 158.9 ± 73.2 spines/cell). Simulations were run 10,000 times,

213   with $N = 1,000$ for P_{IN} and $c_o = 0.20$.

214   Across all simulated inputs, input-output tuning correlation was higher for homogeneous

215   P_{SYN} compared to heterogeneous P_{SYN} (median $r_{hom} = 0.60$, median $r_{hom} = 0.18$, n = 900,000; Fig.

216   5c). Tuning correlation between all imaged dendritic spines and soma was low (median $r_{cell} =$

217   0.31, n = 7,151 spines from 45 cells), more closely resembling our model with a heterogenous

218   P_{IN}. As somatic orientation selectivity (i.e., tuning bandwidth) varies for single cells in ferret V1

219   (Goris et al., 2015; Wilson et al., 2016), we next examined the average input-output tuning

220   correlation across individual cells (Fig. 5d). Here, the homogeneous model exhibited greater

221   specificity then the heterogeneous model (median r_{hom} = 0.52; median r_{het} = 0.18, n = 90,000). For

222   ferret V1 cells, we observed similar spine-soma correlation as the heterogeneous simulation

223   (median r_{cell} = 0.20, n = 45). Ferret V1 cells were not statistically different from neural decoders

224   with a heterogeneous P_{SYN} (p = 0.19, Mann-Whitney test), while neural decoders with a

225   homogeneous P_{SYN} were significantly more correlated with the inputs (p < 0.0001, Mann-Whitney

226   test). A small percentage of imaged cells (17.9%, n = 5/28) had synaptic populations whose mean

227   tuning correlation were within the 95% confidence interval of the homogeneous model distribution.

228   Additionally, some cells had negative average correlations with their spines, which never occurred

229   in the models—potentially indicating nonlinearities between the spines and soma. It is also

230   important to emphasize that both synaptic populations and the heterogeneous model exhibit a

231   positive bias in tuning correlations, illustrating that while inputs are functionally diverse, they are,

232   on average, more similarly tuned to the cell/decoder output.

233   Given the differences between ferret V1 neurons, we quantified the degree to which

234   synaptic populations on *each* neuron matched tuning correlation distributions from models of

235   homogenous and heterogenous P_{IN}, by calculating the Kullback-Leibler divergence (D_{KL}, bin size

236   = 0.05, see Methods). Across our population, imaged neurons more closely resembled

237   simulations with heterogenous, compared to homogenous, P_{IN} (93.3%, n = 42/45; Fig. 5e) and D_{KL}

238   from a heterogenous model was consistently larger (p < 0.0001, sign rank Wilcoxon test). This

239   trend held for a range of histogram bin sizes (0.001 – 0.20). Importantly, the models are not fit to

240   data. They are derived entirely from the statistics of the input population, so this correspondence

241   between the heterogenous model and the data results from no free parameters.

242   In addition to the similarity in input-output tuning correlation, we observed several trends

243   predicted by the heterogeneous model that were evident in synaptic populations imaged *in vivo*.

244   Simulated excitatory inputs correlated with the decoder output were not more selective for

245   orientation (see Methods) (bootstrapped PCA slope = 0.001 ± 0.004 s.e., n = 10,000 simulation

246   runs). For two-photon data, a minuscule, but significant, trend was evident (bootstrapped PCA

247   slope = 0.03 ± 0.1 s.e., n = 7151). So while selective inputs are proposed to provide more

248   information about encoded stimulus variables (Seriès et al., 2004; Shamir and Sompolinsky,

249   2006; Zavitz and Price, 2019) and unselective (or poorly selective) inputs could convey

250   information through their covariance with selective neurons (Zylberberg, 2017), our model and

experimental data suggest co-tuned and orthogonally-tuned inputs exhibit a wide range of tuning selectivity. Response variability (i.e. standard deviation) across trials for simulated excitatory inputs was significantly smaller for 'null' orientations (± 90 deg) than at the 'preferred' (median = 0.30 and IQR = 0.14, median = 0.38 and IQR = 0.22, respectively; p < 0.001, Wilcoxon ranksum test). This trend was also observed in our two-photon data (null: median = 0.13 and IQR = 0.15; preferred: median = 0.23 and IQR = 0.31, respectively; p < 0.001, Wilcoxon ranksum test). As both modeled and imaged neurons had "null"-tuned excitatory inputs that exhibited less response variability, these inputs may carry useful information about when the preferred stimulus is *not* present.

Taken together, our decoding framework with a realistic (i.e. heterogenous orientation tuning), noisy input populations suggest the collection of orthogonally-tuned excitatory inputs in cortical neurons *in vivo* are not unexpected. Instead, the synaptic architecture of layer 2/3 neurons in ferret visual cortex are likely optimized for the readout of upstream populations tuned to orientation.

Discussion

We used a population decoding framework (Jazayeri and Movshon, 2006; Kohn et al., 2016; Pouget et al., 2000; Shamir, 2014) to elucidate a possible source of synaptic diversity in functional response properties. We find that even simple decoders exhibit substantial heterogeneity in their weights when the inputs are noisy, correlated neural populations with heterogeneous orientation tuning. We argue that this could naturally explain the heterogeneity in synaptic inputs measured *in vivo* if these cortical neurons are decoding information from upstream input populations. We compared two neural decoders: one with homogenous input (Jazayeri and Movshon, 2006) and one with heterogenous input (Ecker et al., 2011). We show that empirical measurements from dendritic spines recorded within individual cortical neurons in ferret V1 exhibit a similar amount of diversity in orientation tuning as simulated inputs (i.e. excitatory weights) from heterogeneous input populations. It may appear trivial that heterogeneous input populations would produce heterogeneous weights, but it was neither immediately obvious that the weights would not be smooth nor that excitatory weights would be evident for orthogonal orientations. Orthogonally-tuned or non-preferred inputs are often considered to be aberrant; to be pruned away during experience-dependent plasticity or development (Holtmaat and Svoboda, 2009). Our decoding approach suggests these inputs are purposeful and emerge through development as cortical circuits learn the statistics of their inputs (Avitan and Goodhill, 2018). Taken together, our

results shed light on synaptic diversity that has been puzzling, suggesting that it is, in fact, expected given known properties of the input population.

We believe our study is a significant step forward in combing population coding theory (Averbeck et al., 2006; Pouget et al., 2000) and functional connectomics (Wilson et al., 2016). The ability to measure receptive field properties and statistics of sensory-driven responses of synapses *in vivo* provides a new testing bed for population codes. The individual neurons which synapses converge on are the real components of what has long been a hypothetical downstream population decoder. While we did not set out to build a computational or biophysical model of a neuron, we believe simplistic approaches such as ours are fruitful for understanding basic principles.

To limit complexity, our decoder did not account for many aspects of cortical networks such as stimulus-dependent correlations or recurrent connections. In the case of stimulus dependent covariance, the optimal decoder is no longer linear, however, that decoder closely resembles a complex cell (Jaini and Burge, 2017; Pagan et al., 2016). Extending a decoding framework to include realistic noise has been used to capture many nonlinear features of neural responses including divisive normalization, gain control, and contrast-dependent temporal dynamics— all features which fall naturally out from a normative framework (Chalk et al., 2017). These more sophisticated approaches may be able to make predictions about the synaptic organization itself, whereby local clusters of synapses act as nonlinear subunits (Ujfalussy et al., 2018).

Our model does not describe a cortical transformation. Instead, to limit complexity, we focused on the propagation of orientation selectivity from one neural population to another, akin to the propagation of basic receptive field properties from V1 to higher-visual areas. Our approach was chosen to provide a starting point for predicting the tuning diversity of synaptic input populations as compared to the tuning output or downstream cells. However, this model could be modified to study the convergence and transformation of cortical inputs. An obvious case study would be complex cells in layer 2/3 V1 (Hubel and Wiesel, 1962; Movshon et al., 1978; Spitzer and Hochstein, 1988), which are thought to integrate across presynaptic cells with similar oriented receptive field with offset spatial subunits to produce polarity invariance. This extension would be better suited for a nonlinear quadratic decoder (Jaini and Burge, 2017; Pagan et al., 2016), rather than the linear one used here. We hope that future studies build upon this modeling framework, exploring quadratic decoders and work towards using richer visual stimuli and neural models (Chalk et al., 2017). We believe this will be critical for gaining insight into how information

13

318   propagates between cortical areas and from largescale measurements of cortical functional

319   connectivity.

14

Materials and Methods

All procedures were performed according to NIH guidelines and approved by the Institutional Animal Care and Use Committee at Max Planck Florida Institute for Neuroscience.

Derivation for a Bayesian probabilistic decoder

We construct a probabilistic decoder, represented by a population of neurons, that reports or estimates the identity of a stimulus from the spiking response of an input population of neurons. We assume an input population with responses that are a function of the stimulus, $f(\theta_k)$, plus Gaussian noise, $f(\theta_k)$, and the covariance (Q) is equal for all stimulus conditions (Q) such that $Q = Q_k = Q_i$. Then, the posterior distribution can be written as

$$p(\theta_k|R) = \frac{p(\theta_k) \, N(f(\theta_k), Q)}{\sum_i^K p(\theta_i) \, N(f(\theta_i), Q)}$$

where a multivariate Gaussian is

$$N(f(\theta_k), Q) = 2\pi^{-\frac{n}{2}} |Q|^{-\frac{1}{2}} \, e^{\left(-\frac{1}{2}(R - f(\theta_k))^T Q^{-1}(R - f(\theta_k))\right)}$$

This can be expanded and simplified such that

$$p(\theta_k|R) = \frac{p(\theta_k) 2\pi^{-\frac{n}{2}} |Q|^{-\frac{1}{2}} \, e^{\left(-\frac{1}{2}(R - f(\theta_k))^T Q^{-1}(R - f(\theta_k))\right)}}{\sum_i^K p(\theta_i) 2\pi^{-\frac{n}{2}} |Q|^{-\frac{1}{2}} \, e^{\left(\left(-\frac{1}{2}\right)(R - f(\theta_i))^T Q^{-1}(R - f(\theta_i))\right)}}$$

$$p(\theta_k|R) = \frac{p(\theta_k) \, e^{\left(-\frac{1}{2}\left(R^T Q^{-1} R - f(\theta_k)^T Q^{-1} R - R^T Q^{-1} f(\theta_k) + f(\theta_k)^T Q^{-1} f(\theta_k)\right)\right)}}{\sum_i^K p(\theta_i) \, e^{\left(-\frac{1}{2}\left(R^T Q^{-1} R - f(\theta_i)^T Q^{-1} R - R^T Q^{-1} f(\theta_i) + f(\theta_i)^T Q^{-1} f(\theta_i)\right)\right)}}$$

$$p(\theta_k|R) = \frac{p(\theta_k) \, e^{-\frac{1}{2} R^T Q^{-1} R} \, e^{\left(f(\theta_k)^T Q^{-1} R \, - \frac{1}{2} f(\theta_k)^T Q^{-1} f(\theta_k)\right)}}{\sum_i^K p(\theta_i) \, e^{-\frac{1}{2} R^T Q^{-1} R} \, e^{\left(f(\theta_i)^T Q^{-1} R - \frac{1}{2} f(\theta_i)^T Q^{-1} f(\theta_i)\right)}}$$

$$p(\theta_k|R) = \frac{e^{\left(R^T Q^{-1} f(\theta_k) \, - \frac{1}{2} f(\theta_k)^T Q^{-1} f(\theta_k) + \ln p(\theta_k)\right)}}{\sum_i^K e^{\left(R^T Q^{-1} f(\theta_i) - \frac{1}{2} f(\theta_i)^T Q^{-1} f(\theta_i) + \ln p(\theta_i)\right)}}$$

$$p(\theta_k|R) = \frac{e^{(R^T w_k + \beta_k)}}{\sum_i^K e^{(R^T w_i + \beta_i)}}$$

where

$$w_k = Q^{-1} f(\theta_k)$$

343
$$\beta_k = -\left(\frac{1}{2}\right) f(\theta_k)^T Q^{-1} f(\theta_k) + ln\, p(\theta_k)$$

344 Here, w are the weights over k for each neuron in the decoder population and β is a constant term

345 for each k. Importantly, because we assume Gaussian input, with this formulation, w and β

346 are derived closed form. More generally, w and β, can be estimated numerically using multinomial

347 logistic regression and this form remains optimal for any input population statistics within the

348 exponential family (e.g., Poisson noise).

349

350 *Input population model*

351

352 To generate input populations (P_{IN}), we simulated N neurons responding to a stimulus

353 characterized by orientation ($\theta_k \in [-\pi/2:\pi/2K:\pi/2]$). The response of each neuron, r_i, depends on

354 a tuning function, $f_i(\theta)$, and an additive noise term, ε_i, describing trial-to-trial variability. Noise is

355 correlated across the population, generated from a multivariate Gaussian distribution with zero

356 mean and covariance C. Orientation tuning functions were defined as:

357
$$f_i(\theta) = \alpha_i + \beta_i e^{\kappa_i [cos(\theta - \phi_i)^2 - 1]}$$

358 Here, α is the baseline firing rate, β scales the tuned response, κ scales the tuning bandwidth,

359 and ϕ is the orientation preference of each neuron. For homogeneous P_{IN} all parameters except ϕ

360 were fixed: (α, β, κ) = (0, 5, 4). For heterogenous P_{IN}, we sampled parameters to match

361 measurements from macaque V1 (Ringach et al., 2002) and our ferret V1 data. Tuning bandwidth

362 was generated by converting half-width at $1/\sqrt{2}$ height (γ) values from a lognormal distribution (μ

363 = -1, σ = 0.6):

364
$$\kappa = -log(\sqrt{2})/(cos(\gamma^2) - 1)$$

365 Limited-range correlations were included so neural noise correlation depends on tuning

366 preference difference (Ecker et al., 2011). A correlation matrix, C, was specified by the difference

367 between preferred orientations of neurons and the maximum pairwise correlation, c_o:

368
$$A_{ij} = c_o e^{-|\delta(\phi_i - \phi_j)|}$$

369 where δ is the circular difference and

370
$$C_{ij} = A_{ij} + (1 - c_o)I$$

371 where I is the identity matrix of size N. We scaled the correlation matrix by the mean firing rate of

372 each neuron to produce Poisson-like noise (Ecker et al., 2011).

373 Derived weights for a given P_{IN} were artificially smoothed using the following equation

374 from (Park and Pillow, 2011):

$$S_{ij} \mathrel{+}= e^{-\left(\rho_1 + \left(\frac{\delta\left(\phi_i - \phi_j\right)}{\rho_2}\right)\right)}$$

375

376 Here, S^+ is the pseudoinverse of S, δ $(\phi_i - \phi_j)$ is the circular difference between preferred

377 orientations of neurons, ρ_1 scales the amplitude of smoothing, and ρ_2 scales functional range of

378 smoothing.

379

380 *Population decoder estimation accuracy*

381

382 Decoding accuracy was calculated with the mean-squared-error of the maximum a

383 posterior probability (MAP) estimate across t simulated trials of each stimulus (k):

384
$$error(k) = \left(\frac{1}{t}\right) \sum_1^t angle\left(e^{i(MAP(w_k) - \theta_k)}\right)^2$$

385 Here, w_k are the weights for a given decoder neuron and θ_k is the true stimulus.

386

387 *Viral Injections*

388

389 Briefly, female ferrets aged P18-23 (Marshall Farms) were anesthetized with isoflurane

390 (delivered in O2). Atropine was administered and a 1:1 mixture of lidocaine and bupivacaine was

391 administered SQ. Animals were maintained at an internal temperature of 37 degrees Celsius.

392 Under sterile surgical conditions, a small craniotomy (0.8 mm diameter) was made over the visual

393 cortex (7-8mm lateral and 2-3mm anterior to lambda). A mixture of diluted AAV1.hSyn.Cre

394 (1:25000 to 1:50000) and AAV1.Syn.FLEX.GCaMP6s (UPenn) was injected (125 - 202.5 nL)

395 through beveled glass micropipettes (10-15 micron outer diameter) at 600, 400, and 200 microns

396 below the pia. Finally, the craniotomy was filled with sterile agarose (Type IIIa, Sigma-Aldrich)

397 and the incision site was sutured.

398

399 *Cranial Window*

400

401 After 3-5 weeks of expression, ferrets were anesthetized with 50mg/kg ketamine and

402 isoflurane. Atropine and bupivacaine were administered, animals were placed on a feedback-

403 controlled heating pad to maintain an internal temperature of 37 degrees Celsius, and intubated

404 to be artificially respirated. Isoflurane was delivered throughout the surgical procedure to maintain

405 a surgical plane of anesthesia. An intravenous cannula was placed to deliver fluids. Tidal CO2,

17

external temperature, and internal temperature were continuously monitored. The scalp was retracted and a custom titanium headplate adhered to the skull (Metabond, Parkell). A craniotomy was performed and the dura retracted to reveal the cortex. One piece of custom cover-glass (3mm diameter, 0.7mm thickness, Warner Instruments) adhered using optical adhesive (71, Norland Products) to custom insert was placed onto the brain to dampen biological motion during imaging. A 1:1 mixture of tropicamide ophthalmic solution (Akorn) and phenylephrine hydrochloride ophthalmic solution (Akorn) was applied to both eyes to dilate the pupils and retract the nictating membranes. Contact lenses were inserted to protect the eyes. Upon completion of the surgical procedure, isoflurane was gradually reduced and pancuronium (2 mg/kg/hr) was delivered IV.

Visual Stimuli

Visual stimuli were generated using Psychopy (Peirce, 2007). The monitor was placed 25 cm from the animal. Receptive field locations for each cell were hand mapped and the spatial frequency optimized (range: 0.04 to 0.25 cpd). For each soma and dendritic segment, square-wave drifting gratings were presented at 22.5 degree increments (2 second duration, 1 second ISI, 8-10 trials for each field of view).

Two photon imaging

Two photon imaging was performed on a Bergamo II microscope (Thorlabs) running Scanimage (Pologruto et al., 2003) (Vidrio Technologies) with 940nm dispersion-compensated excitation provided by an Insight DS+ (Spectraphysics). For spine and axon imaging, power after the objective was limited to < 50 mW. Cells were selected for imaging on the basis of their position relative to large blood vessels, responsiveness to visual stimulation, and lack of prolonged calcium transients resulting from over-expression of GCaMP6s. Images were collected at 30 Hz using bidirectional scanning with 512x512 pixel resolution or with custom ROIs (frame rate range: 22 - 50 Hz). Somatic imaging was performed with a resolution of 2.05 - 10.24 pixels/micron. Dendritic spine imaging was performed with a resolution of 6.10 -15.36 pixels/micron.

Two Photon Imaging Analysis

Imaging data were excluded from analysis if motion along the z-axis was detected. Dendrite images were corrected for in-plane motion via a 2D cross-correlation based approach in

440   MATLAB or using a piecewise non-rigid motion correction algorithm (Pnevmatikakis and

441   Giovannucci, 2017). ROIs (region of interest) were drawn in ImageJ; dendritic ROIs spanned

442   contiguous dendritic segments and spine ROIs were fit with custom software. Mean pixel values

443   for ROIs were computed over the imaging time series and imported into MATLAB (Hiner et al.,

444   2017; Sage et al., 2012). $\Delta F/F_o$ was computed by computing F_o with time-averaged median or

445   percentile filter (10th percentile). For spine signals, we subtracted a scaled version of the dendritic

446   signal to remove back-propagating action potentials as performed previously (Wilson et al., 2016).

447    $\Delta F/F_o$ traces were synchronized to stimulus triggers sent from Psychopy and collected by Spike2.

448   Spines were included for analysis if the SNR of the preferred response exceeded 2 median

449   absolute deviations above the baseline noise (measured during the blank) and were weakly

450   correlated with the dendritic signal (Spearman's correlation, r < 0.4). Some spine traces contained

451   negative events after subtraction, so correlations were computed ignoring negative values. We

452   then normalized each spine's responses so that each spine had equal weight. Preferred

453   orientation for each spine was calculated by fitting responses with a Gaussian tuning curve using

454   lsqcurvefit (Matlab). Tuning selectivity was measured as the vector strength index (v) for each

455   neuron's response:

456
$$v_i = \frac{\sqrt{\sum(r_i cos\theta_k)^2 + \sum(r_i sin\theta_k)^2}}{\sum r_i}$$

457   Here r is the mean responses over the orientations (θ_k) presented for each spine (i). Note, this

458   same index is used to characterize simulated input selectivity.

459

460   *Analysis*

461

462   To compare input tuning (derived synaptic population or measured dendritic spine

463   population) with output tuning (downstream readout or measured somatic tuning) we computed

464   the Pearson Correlation coefficient (Matlab). This correlation was computed on trial-averaged

465   responses across different orientations. For dendritic spines and soma, measured responses

466   across stimulus presentation trials were averaged. For simulated synaptic populations and

467   corresponding downstream readout neuron, we simulated trials by adding noise to each synaptic

468   tuning curve.

469

470   **Code Availability**

471

19

472   Matlab code to generate input and readout populations used are provided:

473   https://github.com/schollben/SpineProbablisticModel2020

474   **Figure Legends**

475

476   **Figure 1: A population decoding framework to study synaptic diversity**.

477   An upstream population of neurons are tuned for a single stimulus variable (orientation) (*top*).

478   This input population is readout by downstream decoder neurons (*bottom*). Downstream neurons

479   decode stimulus identify by reading out spikes from the upstream input population. Each decoder

480   neuron is defined by a set weights (*middle*) over the upstream population, which are summed and

481   rectified to produce an output.

482

483   **Figure 2: Model simulations with homogenous and heterogeneous input populations.**

484   (a) Orientation tuning of a homogenous input population. Shown is a subset of the total population

485   (n = 20/1000). Ordinate is orientation preference, restricted between -90° and 90°. (b) Derived

486   weights for a single decoder neuron (preferring 0°) reading out the homogenous (*blue*) input

487   population in (a). Weights for homogenous populations smoothly vary over orientation space. (c)

488   Response output of the decoder neuron whose weights are shown in (b). (d-f) Same as in (a-c)

489   for a heterogeneous input population with moderate correlation (c_o = 0.25). Note that decoder

490   weights for heterogeneous input populations are not smooth.

491

492   **Figure 3: Decoder performance of heterogeneous input populations depends on**

493   **population size, correlations, and weight diversity**.

494   (a) Example weight distribution for a decoder neuron reading out a heterogeneous input

495   population (*top*). Shown are the effects of progressively smoothing weights. Smooth parameters

496   (see Methods) from top to bottom: (0,0), (0.1,1), (0.2, 2), (1,10). Ordinate is orientation preference,

497   restricted between -90° and 90°. (b) Decoder performance (inverse mean-squared-error) plotted

498   for homogenous and heterogeneous input populations of increasing size. Simulations here

499   include no correlations (c_o = 0). Shading indicates standard error. (c) Same as in (b) for input

500   populations with moderate correlation (c_o = 0.25). (d) Same as in (b) for input populations with

501   stronger correlation (c_o = 0.50).

502

503   **Figure 4: Simulation of synaptic populations from decoder neuron weight distributions**. (a)

504   Example weight distribution for a single decoder neuron tuned to 0° (*left*). Ordinates are

505   orientation preference, restricted between -90° and 90°. Dashed line separates excitatory

20

506 (positive) and inhibitory (negative) weights. Excitatory weight distribution over the input population

507 is transformed into a frequency distribution, whereby greater amplitude equates to greater

508 frequency of occurrence (*right*). (b) Example simulated synaptic population (n = 100 spines) from

509 the weight distribution in (a). Shown are the orientation tuning curves of each simulated synapse

510 (normalized).

511

512 **Figure 5: Orientation tuning diversity of dendritic spine populations in ferret V1 match**

513 **simulations with correlated, heterogeneous input populations.**

514 (a) Two-photon standard-deviation projection of example dendrite and spines recorded from a

515 single cell (*left*). Inset: Two-photon standard-deviation projection of corresponding soma. Scale

516 bar is 10 microns. Orientation tuning of soma (*top*) and all visually-responsive dendritic spines

517 from this single cell (n = 159) are shown (*right*). Spine responses are normalized peak ΔF/F.

518 Orientation preferences are shown relative to the somatic preference (aligned to 0º). (b) Same as

519 in (a) for another example cell (n = 162 visually-responsive spines). (c) Cumulative distributions

520 of tuning correlation between individual dendritic spines or simulated synaptic inputs with

521 corresponding somatic tuning or decoder output. Shown are correlations of simulations of

522 homogenous (*blue*) or heterogeneous (*red*) input populations, compared to empirical data (*gray*).

523 (d) Distributions of average tuning correlation between synaptic input and somatic output across

524 measured cells (n = 45). Also shown are distributions of average tuning correlation for simulated

525 cells. Triangles denote median values for each distribution. (e) Comparison of Kullback-Leibler

526 divergence (D_{KL}) between data and each model type. Each data point represents an individual

527 cell's population of dendritic spines.

21

Bibliography

Averbeck, B.B., Latham, P.E., Pouget, A., 2006. Neural correlations, population coding and computation. Nat. Rev. Neurosci. 7, 358–366. doi:10.1038/nrn1888

Avitan, L., Goodhill, G.J., 2018. Code under construction: neural coding over development. Trends Neurosci. 41, 599–609. doi:10.1016/j.tins.2018.05.011

Berens, P., Ecker, A.S., Gerwinn, S., Tolias, A.S., Bethge, M., 2011. Reassessing optimal neural population codes with neurometric functions. Proc. Natl. Acad. Sci. USA 108, 4423–4428. doi:10.1073/pnas.1015904108

Butts, D.A., Goldman, M.S., 2006. Tuning curves, neuronal variability, and sensory coding. PLoS Biol. 4, e92. doi:10.1371/journal.pbio.0040092

Chalk, M., Masset, P., Deneve, S., Gutkin, B., 2017. Sensory noise predicts divisive reshaping of receptive fields. PLoS Comput. Biol. 13, e1005582. doi:10.1371/journal.pcbi.1005582

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

Chen, X., Leischner, U., Rochefort, N.L., Nelken, I., Konnerth, A., 2011. Functional mapping of single spines in cortical neurons in vivo. Nature 475, 501–505. doi:10.1038/nature10193

Cossell, L., Iacaruso, M.F., Muir, D.R., Houlton, R., Sader, E.N., Ko, H., Hofer, S.B., Mrsic-Flogel, T.D., 2015. Functional organization of excitatory synaptic strength in primary visual cortex. Nature 518, 399–403. doi:10.1038/nature14182

Dayan, P., Abbott, L.F., 2001. Theoretical neuroscience: computational and mathematical modeling of neural systems.

DiCarlo, J.J., Zoccolan, D., Rust, N.C., 2012. How does the brain solve visual object recognition? Neuron 73, 415–434. doi:10.1016/j.neuron.2012.01.010

Ecker, A.S., Berens, P., Tolias, A.S., Bethge, M., 2011. The effect of noise correlations in populations of diversely tuned neurons. J. Neurosci. 31, 14272–14283. doi:10.1523/JNEUROSCI.2539-11.2011

Goris, R.L.T., Simoncelli, E.P., Movshon, J.A., 2015. Origin and function of tuning diversity in macaque visual cortex. Neuron 88, 819–831. doi:10.1016/j.neuron.2015.10.009

Graf, A.B.A., Kohn, A., Jazayeri, M., Movshon, J.A., 2011. Decoding the activity of neuronal populations in macaque primary visual cortex. Nat. Neurosci. 14, 239–245. doi:10.1038/nn.2733

Haefner, R.M., Gerwinn, S., Macke, J.H., Bethge, M., 2013. Inferring decoding strategies from choice probabilities in the presence of correlated variability. Nat. Neurosci. 16, 235–242. doi:10.1038/nn.3309

Hiner, M.C., Rueden, C.T., Eliceiri, K.W., 2017. ImageJ-MATLAB: a bidirectional framework for scientific image analysis interoperability. Bioinformatics 33, 629–630. doi:10.1093/bioinformatics/btw681

Holtmaat, A., Svoboda, K., 2009. Experience-dependent structural synaptic plasticity in the mammalian brain. Nat. Rev. Neurosci. 10, 647–658. doi:10.1038/nrn2699

Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. (Lond.) 160, 106–154. doi:10.1113/jphysiol.1962.sp006837

Iacaruso, M.F., Gasler, I.T., Hofer, S.B., 2017. Synaptic organization of visual space in primary visual cortex. Nature 547, 449–452. doi:10.1038/nature23019

Jaini, P., Burge, J., 2017. Linking normative models of natural tasks to descriptive models of neural response. J. Vis. 17, 16. doi:10.1167/17.12.16

Jazayeri, M., Movshon, J.A., 2006. Optimal representation of sensory information by neural populations. Nat. Neurosci. 9, 690–696. doi:10.1038/nn1691

Jia, H., Rochefort, N.L., Chen, X., Konnerth, A., 2010. Dendritic organization of sensory input to cortical neurons in vivo. Nature 464, 1307–1312. doi:10.1038/nature08947

Jia, H., Rochefort, N.L., Chen, X., Konnerth, A., 2011. In vivo two-photon imaging of sensory-evoked dendritic calcium signals in cortical neurons. Nat. Protoc. 6, 28–35. doi:10.1038/nprot.2010.169

Ju, N., Li, Y., Liu, F., Jiang, H., Macknik, S.L., Martinez-Conde, S., Tang, S., 2020. Spatiotemporal functional organization of excitatory synaptic inputs onto macaque V1 neurons. Nat. Commun. 11, 697. doi:10.1038/s41467-020-14501-y

Kerlin, A., Mohar, B., Flickinger, D., MacLennan, B.J., Dean, M.B., Davis, C., Spruston, N., Svoboda, K., 2019. Functional clustering of dendritic activity during decision-making. Elife 8. doi:10.7554/eLife.46966

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

Kohn, A., Coen-Cagli, R., Kanitscheider, I., Pouget, A., 2016. Correlations and neuronal population information. Annu. Rev. Neurosci. 39, 237–256. doi:10.1146/annurev-neuro-070815-013851

Lee, W.-C.A., Bonin, V., Reed, M., Graham, B.J., Hood, G., Glattfelder, K., Reid, R.C., 2016. Anatomy and function of an excitatory network in the visual cortex. Nature 532, 370–374. doi:10.1038/nature17192

Ma, W.J., Beck, J.M., Latham, P.E., Pouget, A., 2006. Bayesian inference with probabilistic population codes. Nat. Neurosci. 9, 1432–1438. doi:10.1038/nn1790

Movshon, J.A., Thompson, I.D., Tolhurst, D.J., 1978. Receptive field organization of complex cells in the cat's striate cortex. J. Physiol. (Lond.) 283, 79–99. doi:10.1113/jphysiol.1978.sp012489

Pagan, M., Simoncelli, E.P., Rust, N.C., 2016. Neural Quadratic Discriminant Analysis: Nonlinear Decoding with V1-Like Computation. Neural Comput. 28, 2291–2319. doi:10.1162/NECO_a_00890

Park, M., Pillow, J.W., 2011. Receptive field inference with localized priors. PLoS Comput. Biol. 7, e1002219. doi:10.1371/journal.pcbi.1002219

Peirce, J.W., 2007. PsychoPy--Psychophysics software in Python. J. Neurosci. Methods 162, 8–13.

604   doi:10.1016/j.jneumeth.2006.11.017

605 Pnevmatikakis, E.A., Giovannucci, A., 2017. NoRMCorre: An online algorithm for piecewise rigid
606   motion correction of calcium imaging data. J. Neurosci. Methods 291, 83–94.
607   doi:10.1016/j.jneumeth.2017.07.031

608 Pologruto, T.A., Sabatini, B.L., Svoboda, K., 2003. ScanImage: flexible software for operating laser
609   scanning microscopes. Biomed Eng Online 2, 13. doi:10.1186/1475-925X-2-13

610 Pouget, A., Dayan, P., Zemel, R., 2000. Information processing with population codes. Nat. Rev.
611   Neurosci. 1, 125–132. doi:10.1038/35039062

612 Reid, R.C., Alonso, J.M., 1995. Specificity of monosynaptic connections from thalamus to visual cortex.
613   Nature 378, 281–284. doi:10.1038/378281a0

614 Ringach, D.L., Shapley, R.M., Hawken, M.J., 2002. Orientation selectivity in macaque V1: diversity and
615   laminar dependence. J. Neurosci. 22, 5639–5651. doi:20026567

616 Rumyantsev, O.I., Lecoq, J.A., Hernandez, O., Zhang, Y., Savall, J., Chrapkiewicz, R., Li, J., Zeng, H.,
617   Ganguli, S., Schnitzer, M.J., 2020. Fundamental bounds on the fidelity of sensory cortical coding.
618   Nature 580, 100–105. doi:10.1038/s41586-020-2130-2

619 Rust, N.C., Mante, V., Simoncelli, E.P., Movshon, J.A., 2006. How MT cells analyze the motion of visual
620   patterns. Nat. Neurosci. 9, 1421–1431. doi:10.1038/nn1786

621 Sage, D., Prodanov, D., Tinevez, J.Y., 2012. MIJ: making interoperability between ImageJ and Matlab
622   possible. bigwww.epfl.ch.

623 Scholl, B., Fitzpatrick, D., 2020. Cortical synaptic architecture supports flexible sensory computations.
624   Curr. Opin. Neurobiol. 64, 41–45. doi:10.1016/j.conb.2020.01.013

625 Scholl, B., Thomas, C.I., Ryan, M.A., Kamasawa N., Fitzpatrick D., 2021. Cortical selectivity derives
626   from strength in numbers of synapses. Nature 590 (7844), 111-114.

627 Scholl, B., Wilson, D.E., Fitzpatrick, D., 2017. Local Order within Global Disorder: Synaptic
628   Architecture of Visual Space. Neuron 96, 1127–1138.e4. doi:10.1016/j.neuron.2017.10.017

629 Seriès, P., Latham, P.E., Pouget, A., 2004. Tuning curve sharpening for orientation selectivity: coding
630   efficiency and the impact of correlations. Nat. Neurosci. 7, 1129–1135. doi:10.1038/nn1321

631 Shamir, M., 2014. Emerging principles of population coding: in search for the neural code. Curr. Opin.
632   Neurobiol. 25, 140–148. doi:10.1016/j.conb.2014.01.002

633 Shamir, M., Sompolinsky, H., 2006. Implications of neuronal diversity on population coding. Neural
634   Comput. 18, 1951–1986. doi:10.1162/neco.2006.18.8.1951

635 Spitzer, H., Hochstein, S., 1988. Complex-cell receptive field models. Prog. Neurobiol. 31, 285–309.
636   doi:10.1016/0301-0082(88)90016-0

637 Stringer, C., Michaelos, M., Pachitariu, M., 2019. High precision coding in mouse visual cortex. BioRxiv.
638   doi:10.1101/679324

639 Ujfalussy, B.B., Makara, J.K., Lengyel, M., Branco, T., 2018. Global and Multiplexed Dendritic
640   Computations under In Vivo-like Conditions. Neuron 100, 579–592.e5.

641   doi:10.1016/j.neuron.2018.08.032

642   Wertz, A., Trenholm, S., Yonehara, K., Hillier, D., Raics, Z., Leinweber, M., Szalay, G., Ghanem, A.,
643   Keller, G., Rózsa, B., Conzelmann, K.-K., Roska, B., 2015. PRESYNAPTIC NETWORKS.
644   Single-cell-initiated monosynaptic tracing reveals layer-specific cortical network modules.
645   Science 349, 70–74. doi:10.1126/science.aab1687

646   Wilson, D.E., Scholl, B., Fitzpatrick, D., 2018. Differential tuning of excitation and inhibition shapes
647   direction selectivity in ferret visual cortex. Nature 560, 97–101. doi:10.1038/s41586-018-0354-1

648   Wilson, D.E., Whitney, D.E., Scholl, B., Fitzpatrick, D., 2016. Orientation selectivity and the functional
649   clustering of synaptic inputs in primary visual cortex. Nat. Neurosci. 19, 1003–1009.
650   doi:10.1038/nn.4323

651   Zavitz, E., Price, N.S.C., 2019. Weighting neurons by selectivity produces near-optimal population codes.
652   J. Neurophysiol. 121, 1924–1937. doi:10.1152/jn.00504.2018

653   Zylberberg, J., 2017. Untuned but not irrelevant: A role for untuned neurons in sensory information
654   coding. BioRxiv. doi:10.1101/134379

655