

RESEARCH ARTICLE

Computational Neuroscience

The geometry of correlated variability leads to highly suboptimal discriminative sensory coding

Jesse A. Livezey,^{1,2} Pratik S. Sachdeva,³ Maximilian E. Dougherty,⁴  Mathew T. Summers,⁵ and
 Kristofer E. Bouchard^{1,2,6,7}

¹Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States; ²Redwood Center for Theoretical Neuroscience, University of California, Berkeley, California, United States; ³Department of Physics, University of California, Berkeley, California, United States; ⁴Department of Neurology, University of California, San Francisco, California, United States; ⁵Department of Molecular and Cell Biology, University of California, Berkeley, California, United States; ⁶Scientific Data Division, Lawrence Berkeley National Laboratory, Berkeley, California, United States; and ⁷Helen Wills Neuroscience Institute, University of California, Berkeley, California, United States

Abstract

The brain represents the world through the activity of neural populations; however, whether the computational goal of sensory coding is to support discrimination of sensory stimuli or to generate an internal model of the sensory world is unclear. Correlated variability across a neural population (noise correlations) is commonly observed experimentally, and many studies demonstrate that correlated variability improves discriminative sensory coding compared to a null model with no correlations. However, such results do not address whether correlated variability is optimal for discriminative sensory coding. If the computational goal of sensory coding is discriminative, then correlated variability should be optimized to support that goal. We assessed optimality of noise correlations for discriminative sensory coding in diverse datasets by developing two novel null models, each with a biological interpretation. Across datasets, we found that correlated variability in neural populations leads to highly suboptimal discriminative sensory coding according to both null models. Furthermore, biological constraints prevent many subsets of the neural populations from achieving optimality, and subselecting based on biological criteria leaves red discriminative coding performance suboptimal. Finally, we show that optimal subpopulations are exponentially small as the population size grows. Together, these results demonstrate that the geometry of correlated variability leads to highly suboptimal discriminative sensory coding.

NEW & NOTEWORTHY The brain represents the world through the activity of neural populations that exhibit correlated variability. We assessed optimality of correlated variability for discriminative sensory coding in diverse datasets by developing two novel null models. Across datasets, correlated variability in neural populations leads to highly suboptimal discriminative sensory coding according to both null models. Biological constraints prevent the neural populations from achieving optimality. Together, these results demonstrate that the geometry of correlated variability leads to highly suboptimal discriminative sensory coding.

correlated variability; neurophysiology; null models; sensory coding

INTRODUCTION

The brain represents the sensory world through the coordinated firing of neural populations. But what is the computational goal of the brain's sensory representations? Traditionally, neural populations in early sensory areas have been thought to transform features of stimuli and transmit them to downstream

cortical areas where they may be read out (i.e., decoded, discriminated). Indeed, many studies of sensory areas seek to analyze what sensory features are transmitted in the brain and with what fidelity. Alternatively, a generative sensory code would enable neural population activity to generate an internal model of the sensory world that may support diverse downstream tasks. Understanding population neural activity necessitates



Correspondence: K. E. Bouchard (kebouchard@lbl.gov).
Submitted 19 July 2024 / Revised 30 October 2024 / Accepted 30 October 2024



analyzing the joint activity of many neural units, beyond single-neuron analysis. Normative theories, which formalize optimality criteria, are powerful tools in these analyses, as they can establish principles for explaining features of experimentally observed neural activity at the population level. Therefore, it is important to develop methods for quantitatively assessing normative theories based on the features observed in neural data. One prominent feature of neural activity is variability: neural recordings exhibit trial-to-trial fluctuations in response to the same stimulus. From a normative perspective, the geometry of variability in neural activity impacts how optimally a population of neurons can discriminate stimuli (1, 2). However, the optimality of correlated variability for discriminative sensory coding in experimental neural population data has not been assessed.

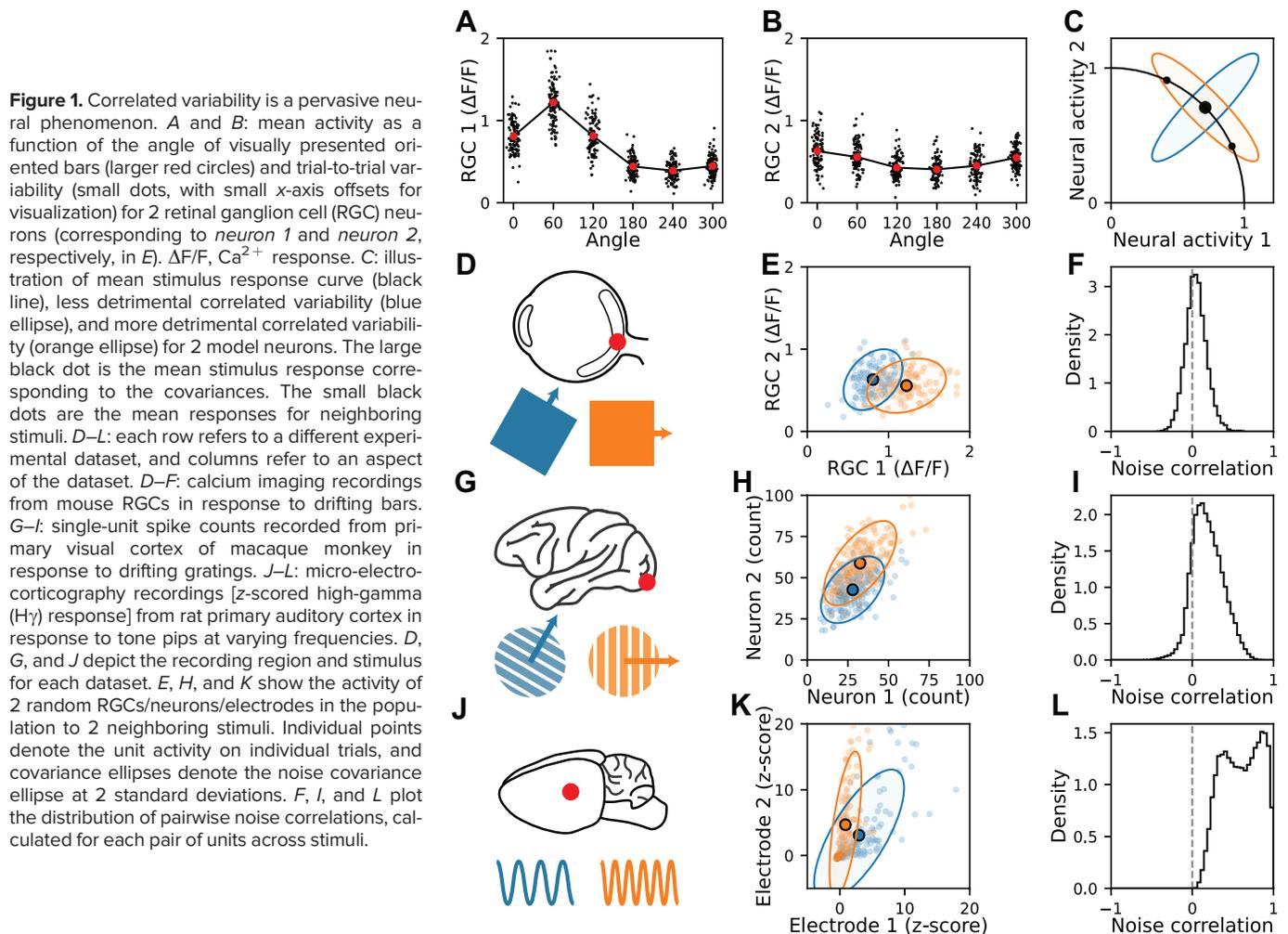
Many studies have characterized pairwise correlations in the trial-to-trial variability of the firing rates of simultaneously recorded neurons, often called correlated variability or noise correlations (3–9). The correlated variability observed in experimental studies typically depends on how neuronal responses are modulated by features of sensory stimuli (i.e., sensory tuning) (10–12). Although correlated variability is typically considered in simultaneous single-neuron electrophysiology measurements, it has been observed in calcium imaging recordings (13) and larger-scale measurements such as electrocorticography recordings (9). For example, Fig. 1, A and B, show the single-trial variability in Ca^{2+} responses ($\Delta F/F$) for two simultaneously recorded mouse retinal ganglion cells (RGCs; each dot is the neuronal response on a single trial) in response to drifting bar stimuli (Fig. 1D). The RGCs' correlated variability for a single stimulus is shown in Fig. 1E, which plots the single-trial responses (dots) for these neurons against each other for two stimuli (colors; large dots indicate means, ovals correspond to 2 SD covariances). Correlated variability has many possible biological sources in neural populations (6, 7, 14–18), some of which the nervous system may be able to modify. Understanding the impact of correlated variability on population coding is important for revealing the principles governing neural computation (1, 2, 4). In particular, if the computational goal of sensory representations is to enable discriminating stimuli, then the sensory code should be optimal for that goal.

Correlated variability can have diverse impacts on the fidelity of a neural code for discriminating stimuli. Theoretical and computational studies have determined how the interplay between correlated variability and tuning properties affects population coding (2, 8, 15, 19–23). Figure 1C illustrates the mean response curve (black line, defined by the mean firing rate of the neurons in response to various stimuli) from two hypothetical simultaneously recorded neurons across a range of stimulus values (3 neighboring stimuli are demarcated with black dots). From a discriminative perspective, if the correlated variability has low variance (Fig. 1C, blue ellipse) perpendicular to the mean stimulus response curve (Fig. 1C, black line), the impact on coding will be less detrimental than having high variance (Fig. 1C, orange ellipse) parallel to the mean stimulus response curve. This is because the trial-by-trial fluctuation (blue ellipse) in response to the central stimulus (large black dot) will minimally overlap with the response to the nearby stimuli (small black dots). In early sensory areas, such

as retina and primary visual cortex (V1), studies have found that correlated variability enhances discriminative coding (6, 16, 24–28). Outside of early sensory areas, both the structure of correlated variability and its impact on discriminative coding are heterogeneous (4, 12, 29). Brain states can change correlated variability and therefore its effect on discriminative coding (30–32). Rummyantsev et al. (33) showed a case where correlated variability was substantially worse than the shuffle null model, suggesting that noise correlations can be detrimental to discriminative coding. These studies leave open the possibility that correlated variability is optimal for discriminative coding in sensory areas; however, this has not been evaluated.

The impact of correlated variability on discriminative coding is typically assessed by comparing the linear Fisher information (LFI) of the experimentally observed correlations to the distribution of LFI under the *shuffle null model*, a null distribution with the same per-neuron variability but no correlations across neurons (1, 2, 29). LFI quantifies how accurately neural population activity can be used to discriminate (i.e., decode) two stimuli. Many previous studies have shown, by using the shuffle null model, that the structure of correlated variability can benefit discriminative coding (6, 16, 24–28). However, comparing the experimentally observed correlated variability with the zero-correlation version is only one comparison for determining optimality; there are potentially other covariance structures that are not captured by the shuffle null model. In principle, the brain's correlated variability could have produced better (or worse) discriminative coding properties. Furthermore, it is unclear whether zero-correlation population activity is the only reasonable null distribution, highlighting the importance of developing other null models (34) that span the space of covariance structures. Testing normative theories of stimulus coding in neural datasets requires understanding whether the geometry of experimental correlated variability is optimal; however, a methodological framework for testing the optimality of correlated variability is currently lacking.

Here, we ask the specific question: does the experimentally observed geometry of correlated variability in neural population data result in optimal discriminative coding? To test the optimality of correlated variability in experimentally observed neural responses, we developed two null models. The *uniform correlation (UC) null model* and the *factor analysis (FA) null model* each define a null distribution of correlated variability and have particular biological interpretations. Using these null models, we test the optimality of neural coding in newly acquired data recorded from retinal ganglion cells (RGCs, Retina), previously recorded neurons in primary visual cortex (V1), and newly acquired electrocorticography (ECoG) data on primary auditory cortex (PAC) (Fig. 1, D–L). These datasets span species, brain areas, and recording modalities used in many previous studies. Our main finding is that the experimentally observed geometry of correlated variability leads to highly suboptimal discriminative coding across all datasets and both null models. Furthermore, the degree of suboptimality worsens as a function of the number of neural units considered in the neural population. We find that for a large fraction of subpopulations of the recorded units, achieving optimality would push the neural responses into regimes that



violate biological constraints. However, even when neural units are subsampled to optimize for biological criteria, they remain highly suboptimal. Finally, direct selection of optimal discriminative subpopulations shows that the optimal population is exponentially small as a function of neural dimensionality. Our results demonstrate that the traditional null model of correlated variability cannot be used to assess the optimality of discriminative coding in data and that biological constraints limit the ability of neural activity to achieve optimal correlated variability. Together, our results show that the geometry of correlated neural response variability leads to highly suboptimal discriminative sensory coding.

METHODS

Data Availability

The preprocessed and trialized Retina and trialized PAC datasets are available at <https://zenodo.org/records/14342290>. The raw PAC data are available at <https://crcns.org/data-sets/ac/ac-7> (R32_B7.nwb). The V1 data are available at <https://crcns.org/data-sets/vc/pvc-11>.

Code Availability

Code to reproduce the analysis and figures is available at https://github.com/BouchardLab/noise_correlations.

Neural Recordings

We examined correlated variability in a diverse set of datasets, spanning distinct brain regions, animal models, and recording modalities. We used calcium imaging recordings from mouse retinal ganglion cells, single-unit recordings from macaque primary visual cortex, and micro-electrocorticography (μ ECoG) recordings from rat auditory cortex. We briefly describe the experimental and preprocessing steps for each dataset. See Fig. 1 and Table 1 for summaries of the datasets, including the recording modalities, stimulus types, number of simultaneously recorded units, number of distinct stimuli, and number of trials per stimulus. For the retinal and auditory cortex datasets, all procedures were performed in accordance with established animal care protocols approved by the University of California-Berkeley/Lawrence Berkeley National Lab Institutional Animal Care and Use Committee.

Recordings from mouse retina.

Mouse retina data were collected via ex vivo two-photon calcium imaging in an isolated retina preparation (35). The retina was bulk loaded with Cal-520 AM dye with a previously described multicell bolus loading technique (36) and then imaged with ScanImage software (37) at 2.96 Hz in the ganglion cell layer of a $425 \times 425\text{-}\mu\text{m}$ area of ventral retina.

Table 1. Experimental dataset summary

Dataset	Animal	Recording	Stimulus	Units	Stimuli	Trials/Stim
Retina	Mouse (isolated)	Calcium imaging	Drifting bars	54	6	114
V1	Macaque	Utah array	Drifting gratings	106	12	200
PAC	Rat	μ ECoG	Tone pips	65	30	60

PAC, primary auditory cortex; V1, primary visual cortex; μ ECoG, micro-electrocorticography.

Visual stimuli were delivered via an ultraviolet LED (375 nm) coupled to a digital micromirror device and were presented on the flyback of the fast-axis scanning mirror during a scan to interleave the stimuli with imaging (35, 38). Visual responses were elicited via 600×600 - μ m bars drifting for 2.93 s at 750 μ m/s in one of six directions (spanning 0° to 300°), with a 5-s intertrial interval. Each direction was presented 114 times, for a total of 684 trials per cell. Fluorescence signals from 832 manually selected regions of interest were baseline subtracted and normalized to calculate a $\Delta F/F_0$ time series. Of these regions of interest, 54 were used for further analysis after determination of directional tuning via permutation testing and manual screening. Per-trial RGC activity used in the analysis here is the maximum $\Delta F/F_0$ value. Retina data were collected by M. T. Summers. Further details on surgical, experimental, and preprocessing steps can be found in Refs. 38, 39.

Recordings from macaque primary visual cortex.

Primary visual cortex (V1) data comprised spike-sorted units simultaneously recorded in anesthetized macaque monkey. The data were obtained from the Collaborative Research in Computational Neuroscience (CRCNS) data sharing website (40) and were recorded by Kohn and Smith (41). This dataset contains recordings from three monkeys, of which the main text presents results from the first (results on the 2 other monkey datasets are qualitatively similar). Recordings were obtained with a 10×10 grid of silicon microelectrodes spaced 400 μ m apart and covering an area of 12.96 mm². The monkey was presented with grayscale sinusoidal drifting gratings, each for 1.28 s. Twelve unique drifting angles (spanning 0° to 330°) were each presented 200 times, for a total of 2,400 trials per monkey. Spike counts were obtained in a 400-ms bin after stimulus onset. A total of 106 units were isolated in the monkey presented in the main text. These units were chosen by the original authors such that 1) their signal-to-noise ratio (the ratio of the average waveform amplitude to the standard deviation of the waveform noise) was at least 2.75, 2) the best grating stimulus evoked at least 2 spikes/s, and 3) the variance-to-mean response ratio did not exceed 10. Further details on the surgical, experimental, and preprocessing steps can be found in Refs. 11, 42.

Recordings from rat primary auditory cortex.

Primary auditory cortex (PAC) data comprised cortical surface electrical potentials (CSEPs) recorded from rats with a custom fabricated micro-electrocorticography (μ ECoG) array. The μ ECoG array consisted of an 8×16 grid of 40- μ m-diameter electrodes. Anesthetized rats were presented with 50-ms tone pips of varying amplitude (8 different levels of attenuation, from 0 dB to -70 dB) and frequency (30 frequencies equally spaced on a log scale from 500 Hz to 32 kHz). We only used samples for the lowest three levels of

attenuation, since these evoked the largest responses. Each frequency-amplitude combination was presented 20 times, for a total of $3 \times 30 \times 20 = 1,800$ samples. The response for each trial was calculated as the z -scored to baseline, high- γ band amplitude of the CSEP, calculated using a constant-Q wavelet transform. The maximum of the per-trial high- γ activity was used in the analysis here. Of the 128 electrodes, we used 65, selecting those that recorded from primary auditory cortex. Data was recorded by M. E. Dougherty and K. E. Bouchard. Further details on the surgical, experimental, and preprocessing steps can be found in Refs. 43, 44.

Linear Fisher Information Measures Coding Fidelity

A commonly used measure of coding fidelity in the context of decoding is the Fisher information, which provides a limit on how accurately a readout of a neural representation can be used to determine the value of the stimulus (45). Formally, the Fisher information is a lower bound on the variance of an unbiased estimator for the stimulus. In practice, the Fisher information is analytically intractable. An alternative measure is the linear Fisher information (LFI). The LFI acts as a suitable lower bound to the Fisher information and is the most commonly used measure of coding fidelity in correlated variability analyses (1, 6, 15, 16, 19, 46–48).

Experimental neuroscience datasets only consider discrete sets of stimuli. In particular, the derivative of the average neural activity must be estimated by considering the neighboring pairs of stimuli. Thus, in practice, we calculate the coarsened linear Fisher information (49), which is defined for two stimuli s_1 and s_2 as

$$\mathcal{I}_{\text{coarse}}(\mathbf{f}_1, \mathbf{f}_2, \Sigma_1, \Sigma_2) = \left(\frac{\mathbf{f}_1 - \mathbf{f}_2}{\Delta s} \right)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} \left(\frac{\mathbf{f}_1 - \mathbf{f}_2}{\Delta s} \right) \quad (1)$$

where $\mathbf{f}_1 = \mathbf{f}(s_1)$, $\mathbf{f}_2 = \mathbf{f}(s_2)$, $\Sigma_1 = \Sigma(s_1)$, $\Sigma_2 = \Sigma(s_2)$, and Δs is the stimulus difference between s_1 and s_2 , whose form may depend on the stimulus structure. In addition, we use the unbiased LFI estimator (48) for the observed LFI values as well as for those sampled from null models. Note that since the corrections to the naive estimator only depend on the dimensionality of the neural population and number of samples, the corrections only impact the raw LFI values and not percentiles. In this work, we use the terms “coarsened LFI” and “LFI” interchangeably.

Assessing the Optimality of Neural Data with Null Models

Information theoretic analyses of neural data often ask whether the observed neural data are “optimal.” In the case of correlated variability, the question can be posed as: are the observed covariances optimal from a decoding perspective? We quantified the coding fidelity with the linear Fisher information (LFI, Eq. 1). In this case, LFI can be infinitely large if

$\Sigma \rightarrow 0$ [or at least if the subspace of Σ^{-1} defined by $\frac{df(s)}{ds}$ diverges]. This answer is likely unsatisfying because neural systems have many sources of variability, and so expecting a neural system to become noiseless or exactly remove noise from a subspace seems implausible. Therefore, when assessing the optimality of correlated variability, one must decide which aspects of the correlated variability the neural system could modify and which aspects will remain fixed.

In this section, we develop the formalism that will allow us to assess the optimality of observed correlated neural variability. The formalism consists of first defining a covariance parameterization for Σ , which is composed of constraints (fixed parameters) and degrees of freedom (free parameters). These constraints and degrees of freedom define the space of allowed correlated variability. Ideally, these constraints and degrees of freedom have some biological interpretation, e.g., fixed private variability or input from other regions of the brain (7, 14). Then, a null model is defined by combining a covariance parameterization with a null distribution over the degrees of freedom. The distribution of some measure, such as the LFI, under the null model can be used to assess the optimality of the observed neural data.

We first review the commonly used fixed-marginal constraint for correlated variability using our formalism and then define the commonly used shuffle and novel uniform correlation (UC) null models. Finally, we propose the factor analysis covariance parameterizations and associated null model for assessing optimality, which has more biological interpretability. In the following sections we use the terminology that we define here:

- Covariance parameterization: a parameterization of Σ that can combine various constraints (fixed parameters) and degrees of freedom (free parameters)
- Constraints: elements of the covariance parameterization that are estimated from data and fixed
- Degrees of freedom: elements of the covariance parameterization that can potentially be modified or optimized to analyze a null model or optimality
- Optimality: values for the degrees of freedom in a covariance parameterization that maximize a specified objective. Here we assess optimality using the linear Fisher information (LFI), although this formalism can be applied to other objectives
- Null distribution: distribution of a covariance parameterization's degrees of freedom
- Null model: combines a covariance parameterization with a baseline or uniform correlation null distribution over the degrees of freedom.

The standard constraint considered for understanding correlated neural variability is to keep the per-neuron marginal distributions fixed. Since the LFI only depends on the covariance of the correlated variability, the fixed-marginal parameterization is equivalent to constraining the per-neuron variances to be constant [equivalently, the diagonal of Σ is kept constant, $\text{diag}(\Sigma) = \sigma^2$]. The corresponding degrees of freedom in this parameterization are the positive-definite pairwise correlation matrix, ρ , specifically the symmetric, off-diagonal entries, ρ_{ij} for $i \neq j$, which can vary. Under this

parameterization, the observed covariance structure can be compared to other proposed distributions of correlations.

When considering the structure that generates Σ , it is desirable that the constraints and degrees of freedom be biologically interpretable. This can be achieved by considering the equations that define the mean-centered, single-trial response in terms of the degrees of freedom being considered. For the fixed-marginal parameterization, the distribution of the single-trial responses, $\mathbf{f}_i(s)$, can be written in terms of a multivariate normal distribution with the mean response, $\mathbf{f}(s)$, where the covariance is the element-wise product of the constrained marginal standard deviations, $\sigma\sigma^T$, and the free correlations, ρ ,

$$\begin{aligned} \mathbf{f}_i(s) &= \mathbf{f}(s) + \boldsymbol{\epsilon} \\ \boldsymbol{\epsilon} &\sim \mathcal{N}(0, \sigma\sigma^T \odot \rho) \end{aligned} \quad (2)$$

This equation is difficult to directly interpret as a network model, but the correlations could be seen as coming from recurrent activity within the observed neurons.

Given a parameterization (fixed-marginal) and a measure of coding fidelity (LFI), it is possible to find optimal covariance structures as a function of the free parameters. In general, the value (or distribution of values) for the degrees of freedom that lead to optimality can be derived analytically or optimized numerically. For the fixed-marginal parameterization, this corresponds to finding the points $\hat{\rho}$, such that

$$\hat{\rho} = \arg \max_{\rho} \text{LFI} \left(\frac{df(s)}{ds}, \text{diag}(\Sigma), \rho \right) \quad (3)$$

Hu et al. (22) characterize the optima of the fixed-marginal parameterization, although they do not provide a constructive way of finding the global optima. We optimize ρ numerically to find optima. We find that the optimization process finds many local maxima for $\hat{\rho}$ in practice.

Novel Null Models Allow the Assessment of Optimality in Neural Data

So far, we have laid out a formalism to define the optimal degrees of freedom for a specified covariance parameterization. However, it is unlikely that observed neural data will precisely match the predicted optimal degrees of freedom, even if the biological system is behaving optimally, so the predictions from Eq. 3 cannot be used directly to assess optimality in data. To assess the optimality of a observed population of neurons, a null model must be constructed for a corresponding parameterization. In this formalism, constructing a null model corresponds to assuming a null distribution for the degrees of freedom of the covariance parameterization. The null distribution should correspond to some notion of “uniform” or “baseline” for the degrees of freedom.

For example, the shuffle null model, based on the fixed-marginal parameterization, posits that the baseline distribution of correlations is zero. The shuffle null model compares the LFI of the observed response to the distribution of LFIs where the individual neural responses are independently trial shuffled, that is, with fixed-marginal variability, no underlying pairwise correlations, and empirical pairwise correlations only arising from finite sampling effects. Under this choice of null model, the observed LFI can be beneficial if it has a high percentile under the null distribution that has no correlations. The shuffle null model provides a limited

baseline comparison for the observed LFI. To assess optimality, the distribution of parameters should be uniform over the space of allowed covariance matrices, which is the motivation for the uniform correlation null model.

Across a population, the median observed percentile across subpopulations and stimuli can be used to categorize a dataset as optimal (median percentile $\geq 2/3$), near-chance (median percentile between $1/3$ and $2/3$), or suboptimal (median percentile $< 1/3$). This categorization is motivated by simplicity in having few categories. However, it is also desirable to not have the optimal and suboptimal categories share a boundary. If they do, small changes in percentiles can switch between optimal and suboptimal. In our case, since the null model defines “near-chance,” having three categories is natural. The near-chance boundaries could be set in a number of ways besides the choice for an even division into thirds. A Kolmogorov–Smirnov test could compare the distribution of percentiles to a uniform distribution. However, given the large number of subpopulations and stimuli we use, empirically no distributions of percentiles in these datasets would be near-chance for P value thresholds in sensible ranges. Said another way, almost no empirical distributions of percentiles are statistically similar to a uniform distribution. A looser test could be to test whether a binomial distribution with $P = 0.5$ would lead to the observed distribution of percentiles categorically above and below 0.5. We find that with P values in sensible ranges this gives boundaries comparable to the division into thirds, but the boundaries differ across datasets because of the variation in the number of subpopulations and stimuli.

In some cases, it may also be possible to define a distribution over optimal covariances and categorize whether the observed LFI is likely under the optimal covariance distribution. For instance, if there is a unique optimal covariance, the Wishart distribution could be used to create a sampling distribution of optimal LFIs that the observed LFIs could be compared against. This is not possible in our case since there is not generally a unique optimal covariance. The Wishart distribution could be used to understand the sampling distribution of optimal covariance matrices, which is required for assessing real data.

Uniform correlation and factor analysis null models.

We derive null distributions for linear Fisher information (LFI) by defining null distributions over covariance matrices. For the factor analysis (FA) null model, the covariance constraints are on the diagonal private variability and the magnitude of the shared variability. This leaves the rotation of the shared variability as the free parameter to form a distribution over. The maximum entropy distribution over rotations can then be defined uniquely by the uniform distribution according to the Haar measure (50). In this sense, the FA null model was derived according to the maximum entropy principle. For the uniform correlation (UC) model, our understanding is that there is no equivalent Haar measure to uniquely define a uniform distribution over correlation matrices (51), so we cannot claim that the UC null model provides a maximum entropy distribution. The particular uniform distribution used is specifically designed to have uniform distributions of joint pairwise correlations (52). However, the uniform distribution defined this way is intuitively higher entropy than the shuffle null model.

Our first contribution is the uniform correlation null model based on the fixed-marginal parameterization, where the correlations are chosen randomly from a uniform distribution over correlation matrices (52). This tests whether the observed correlations are optimal with respect to all possible correlations, rather than only comparing against zero correlations. To our knowledge, this null model has not been considered before. Evaluating data under this null model provides a stronger assessment of the optimality of the observed correlated variability than the shuffle null model.

At another extreme, we could attribute all trial-to-trial variability to external sources that the network can shape or filter. To prevent trivial solutions, we can restrict the network to only changing the loading of the variability onto the neurons (through a rotation, \mathbf{R}). This model was previously discussed (22), but not analyzed because of its incompatibility with the fixed-marginal constraint.

As a parsimonious combination of the fixed-marginal constraint and pure rotation degrees of freedom, we propose using a factor analysis (FA) model to parameterize the correlated variability. Factor analysis decomposes the observed correlated variability into two components: the first is per-neuron private variability, represented as a diagonal matrix $\text{diag}(\sigma_{\text{FA}}^2)$, and the second is a low-rank shared variability component, $\mathbf{L}_{\text{FA}}^T \mathbf{L}_{\text{FA}}$, where $\mathbf{L}_{\text{FA}} \in \mathbb{R}^{k \times d}$, $k < d$. The FA null model has private variability and the spectrum of the shared component as constraints and the rotation of the shared components as the degrees of freedom, combining aspects of the fixed-marginal and rotation null models. The single-trial response can be written as a function of the mean response $\mathbf{f}(s)$, private variances σ_{FA}^2 , low-rank external sources \mathbf{z}_{FA} , loading matrix \mathbf{L}_{FA} , and rotation matrix \mathbf{R}

$$\begin{aligned} \mathbf{f}_i(s) &= \mathbf{f}(s) + \mathbf{R}^T \mathbf{L}_{\text{FA}}^T \mathbf{z}_{\text{FA}} + \boldsymbol{\epsilon}_{\text{FA}} \\ \mathbf{z}_{\text{FA}} &\sim \mathcal{N}(\mathbf{0}, \mathbf{I}) \\ \boldsymbol{\epsilon}_{\text{FA}} &\sim \mathcal{N}(\mathbf{0}, \text{diag}(\sigma_{\text{FA}}^2)) \end{aligned} \quad (4)$$

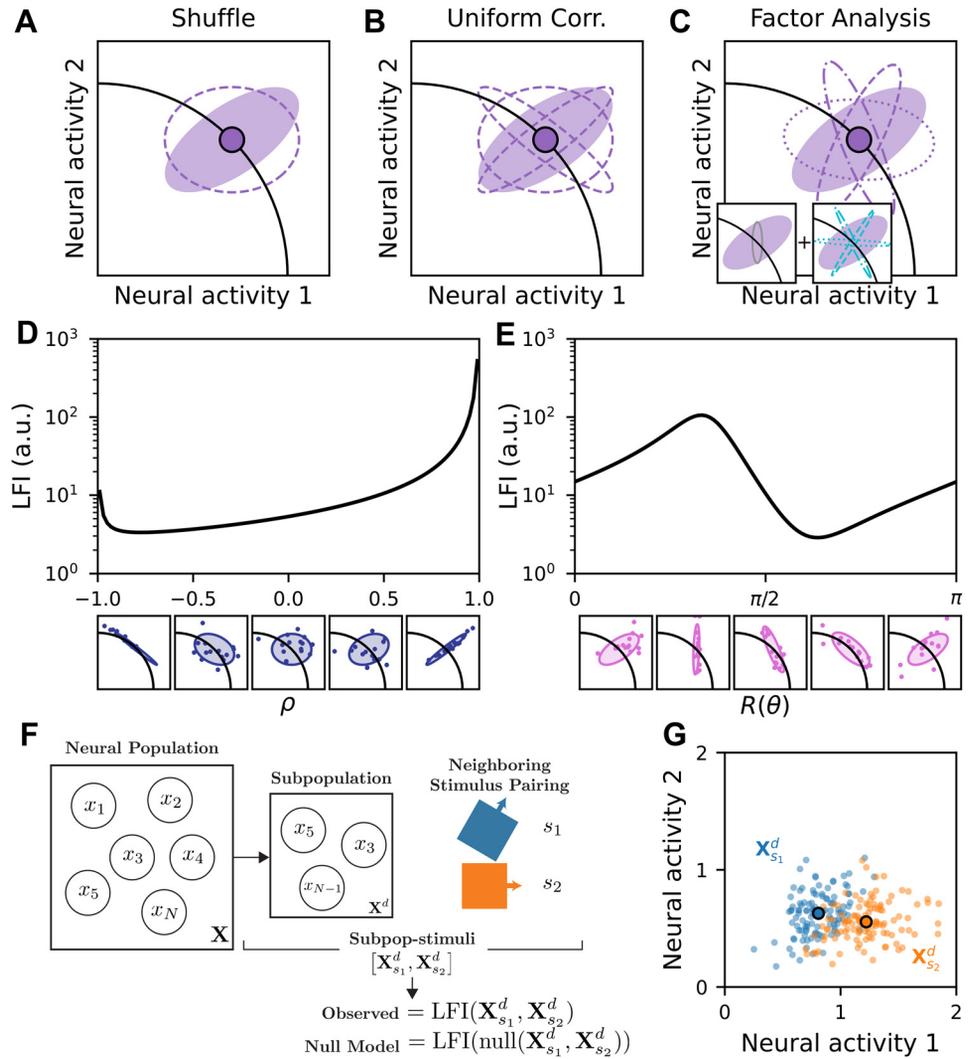
To our knowledge, there is no closed-form solution for $\hat{\mathbf{R}}$ in the FA model to maximize LFI. Instead, to optimize the FA model, the rotation can be numerically optimized by gradient ascent. To construct the FA null model, a uniform distribution (Haar distribution) over special orthogonal rotations (53) is applied to the rotations.

To estimate the initial σ_{FA}^2 and \mathbf{L}_{FA} , we fit a factor analysis model to the samples (54). In fitting the model we had two requirements. The first is that we wanted the dimensionality of the shared component, k , to be as large as possible so that the observed covariance can be modeled as accurately as possible. In opposition to this, we wanted the factor analysis model parameters to be identifiable, meaning that the private variance estimate is unique, which places a limit, which depends on d , on how large k can be (55). In practice, we find the largest k that is lower than the identifiability bound where different initializations return the same parameters. Note that factor analysis is never identifiable in two dimensions, so we do not consider $d = 2$.

Population Statistics across Subpopulations and Stimuli Measure Optimality under a Null Model

Each dataset can be described by a $D \times N$ design matrix \mathbf{X} , where D is the total number of samples and N is the number of

Figure 2. Novel methods for assessing the optimality of neural codes. **A–C:** null models of correlated variability. Solid purple ellipses denote idealized trial-to-trial variability observed about the mean stimulus activity (solid point). Samples from the null models are depicted by dashed ellipses. **A:** the shuffle null model maintains per-neuron variance and samples correlations near 0. **B:** the uniform correlation (UC) null model maintains per-neuron variance and samples uniform correlations. **C:** the factor analysis (FA) null model combines a fixed private variability (estimated from the experimental data, *left inset*) with shared variability (*right inset*) that can be rotated to form null samples (dash styles are consistent between the teal shared variabilities in the *right inset* and the purple null samples in the main panel). **D:** for a synthetic 2-dimensional (2-D) dataset, the linear Fisher information [LFI, arbitrary units (a.u.)] for the UC parameterization as a function of the pairwise correlation, ρ , is shown at *top*; *bottom* plots are the covariance and samples as a function of ρ . **E:** for a synthetic 2-D dataset, the LFI for the FA parameterization as a function of the rotation angle, $R(\theta)$, is shown at *top*; *bottom* plots are the covariance and samples as a function of θ . **F:** to calculate an observed LFI or percentile under a null model, d units were randomly subsampled from the population. Then, 2 neighboring stimuli, s_1 and s_2 , were chosen. The subpopulation and stimulus pairing together constitute a pair of design matrices $[\mathbf{X}_{s_1}^d, \mathbf{X}_{s_2}^d]$. These matrices are the inputs into a LFI calculation or null model analysis and form the basis for the distributions of calculated quantities. **G:** responses in the retinal data for the depicted stimulus pairing (colors) from **F**.



units in the population (Fig. 2F). We considered distributions of LFI across subpopulations and stimuli, or subcomponents of the design matrix. To create subpopulations and stimuli, we first selected a subpopulation of size d by subsampling d units from the population at random, resulting in the $D \times d$ design matrix \mathbf{X}^d (Fig. 2F). Next, we created the subpopulation and stimulus by further subsampling the design matrix according to a specific stimulus pairing. Specifically, we chose two neighboring stimuli, s_1 and s_2 (Fig. 2F), and isolated the samples of \mathbf{X}^d corresponding to those stimuli, thereby creating a pair of design matrices $[\mathbf{X}_{s_1}^d, \mathbf{X}_{s_2}^d]$. The subpopulation and stimulus maps to the task of discriminating between two neighboring stimuli using a subpopulation's responses across trials to those stimuli, which can be visualized in the neural space (Fig. 2G).

For each dataset, we considered subpopulation dimensions $d = 3-20$. As we only allowed neighboring stimulus pairings, the number of available stimulus pairings for a subpopulation was 6 (retinal), 12 (V1), and 29 (PAC). Note that the retinal and V1 stimulus sets are circular, providing an additional stimulus pairing. In the retinal and V1 datasets, we drew 1,000 subpopulations for each dimension d and

considered all stimulus pairings per subpopulation, resulting in $1,000 \times 6 = 6,000$ subpopulations and stimuli for the retinal dataset and $1,000 \times 12 = 12,000$ subpopulations and stimuli for the V1 dataset. To manage computation time, we considered 3,000 unique subpopulations and stimuli for the PAC dataset, selecting both the subpopulation and stimulus pairing at random for each subpopulation and stimulus.

For each subpopulation and stimulus, we calculate its observed LFI, defined as $\mathcal{I}_{\text{coarse}}(\mathbf{f}_1, \mathbf{f}_2, \Sigma_1, \Sigma_2)$. Specifically, we computed

$$\mathcal{I}_{\text{obs}}(\mathbf{X}_{s_1}^d, \mathbf{X}_{s_2}^d) = \mathcal{I}_{\text{coarse}}\left(\text{mean}(\mathbf{X}_{s_1}^d), \text{mean}(\mathbf{X}_{s_2}^d), \text{cov}(\mathbf{X}_{s_1}^d), \text{cov}(\mathbf{X}_{s_2}^d)\right) \quad (5)$$

$$= \left(\frac{\mathbf{f}_{s_1}^d - \mathbf{f}_{s_2}^d}{\Delta s}\right)^T \left(\frac{\Sigma_{s_1}^d + \Sigma_{s_2}^d}{2}\right)^{-1} \left(\frac{\mathbf{f}_{s_1}^d - \mathbf{f}_{s_2}^d}{\Delta s}\right) \quad (6)$$

where $[\mathbf{f}_{s_1}^d, \mathbf{f}_{s_2}^d]$ are the subpopulation and stimulus average responses, $[\Sigma_{s_1}^d, \Sigma_{s_2}^d]$ are the subpopulation and stimulus covariances, and Δs is the stimulus difference, or $\Delta s = |s_1 - s_2|$. When necessary, the stimulus difference was taken as a

circular difference (retinal and V1 datasets). Since the LFI is scaled by the units of the stimulus difference, it is only meaningful to compare observed LFIs within a particular stimulus type. In this work, since all datasets use a different stimulus, the LFIs may not have a meaningful relationship across datasets.

Each null model acts on the design matrices of a subpopulation and stimulus and outputs a distribution of covariance matrices. For example, the fixed-marginal null model shuffles the data within the design matrix, producing new design matrices $[\mathbf{X}_{s_1}^d, \mathbf{X}_{s_2}^d]$ and corresponding covariances $[\Sigma_{s_1}^d, \Sigma_{s_2}^d]$. We then calculate the LFI using the new covariance matrices. Each null model can be summarized as such: a sampled transformation is applied to the observed subpopulation and stimulus, producing new sampled covariance matrices and therefore a sample of LFI from the null. The shuffle null model transformed the data directly, so we write its LFI as

$$\mathcal{I}_{FM}(\mathbf{X}_{s_1}^d, \mathbf{X}_{s_2}^d) = \mathcal{I}_{\text{obs}}(\text{shuffle}(\mathbf{X}_{s_1}^d), \text{shuffle}(\mathbf{X}_{s_2}^d)). \quad (7)$$

Meanwhile, the uniform and factor analysis null models transform the covariance parameterization directly, so we write their LFIs as

$$\mathcal{I}_U(\mathbf{X}_{s_1}^d, \mathbf{X}_{s_2}^d) = \mathcal{I}_{\text{coarse}}(\mathbf{f}_{s_1}^d, \mathbf{f}_{s_2}^d, \text{sample}_U(\Sigma_{s_1}^d), \text{sample}_U(\Sigma_{s_2}^d)) \quad (8)$$

$$\mathcal{I}_{FA}(\mathbf{X}_{s_1}^d, \mathbf{X}_{s_2}^d) = \mathcal{I}_{\text{coarse}}(\mathbf{f}_{s_1}^d, \mathbf{f}_{s_2}^d, \text{rotate}_{FA}(\Sigma_{s_1}^d), \text{rotate}_{FA}(\Sigma_{s_2}^d)). \quad (9)$$

Equations 7 and 9 capture a single application of a null model. Specifically, $\text{shuffle}(\cdot)$ shuffles the neural data, $\text{sample}_U(\cdot)$ samples a random off-diagonal correlation structure and applies it to the covariance, and $\text{rotate}_{FA}(\cdot)$ applies a rotation to the shared component of the covariance. However, we were interested in characterizing the entire distribution of the null model. Thus, for each subpopulation and stimulus, we applied 1,000 samples of the null model to obtain a null model distribution of LFIs. We then calculated observed percentiles as the fraction of samples for which the observed LFI exceeded the null model LFI. Thus, each observed subpopulation and stimulus has its own corresponding observed percentile per null model. When summary statistics are reported such as the median LFI, median percentile, or optimal fraction, 95% bootstrap confidence intervals from 1,000 boot strap resamples are reported (56).

Estimating the population means and covariance will have some variability due to the finite number of trials used for estimation. Although we use the unbiased LFI estimator (48) for the observed LFI values, there may still be a dependence of the LFI or percentile medians or variability on the number of trials. Empirically, we find that the UC and FA null models' LFI and percentile estimates are insensitive to trial subsampling in the V1 dataset (200 trials per stimulus) when sampled to 50% (100 trials) and 25% (50 trials) of the original number of trials. The shuffle null model percentiles show a qualitative chance, but this does not change the interpretation of the analysis (results not shown). We do expect the estimates to break down as the number of trials approaches the subpopulation dimensionality ($n = 20$),

which does constrain the types of experimental design required for this analysis.

Optimal Fraction Calculation

The optimal fraction of a population was calculated in the following way. Given a set of subpopulations and stimuli at a particular dimension, the observed percentiles were calculated. Then, the percentiles were sorted from largest to smallest. The optimal fraction of the percentiles is initially set as the largest single percentile. Starting from this, the median percentile of the optimal fraction is calculated. If the median is $\geq 2/3$, the next smallest percentile is included in the optimal fraction and the process continues to iterate. If the optimal fraction is $< 2/3$, the process terminates. This defines the largest possible fraction of the percentiles that can be retained and have their median be $\geq 2/3$. For reference, the top 2/3 of a uniform distribution (i.e., $[1/3, 1]$) of percentiles has median equal to 2/3.

Measures of Biological Plausibility

We calculated the mean Fano factors (FFs) for a subpopulation and stimulus, based on the per-unit variance and response means

$$\text{FF} = \frac{1}{d} \sum_{i=1}^d \frac{\Sigma_{ii}(s)}{f(s)_i} \quad (10)$$

of the observed and optimal covariances matrices directly from the mean response and covariance matrix parameters.

We calculated the negative density (ND) as follows. For each subpopulation and stimulus, we calculated $f_i^{1\%}$, the neural activity at the 1st percentile, for each neuron i . We then computed $\text{CDF}_i(f_i^{1\%})$, the cumulative density at $f_i^{1\%}$, for a Gaussian obtained from either the observed covariance or the optimal covariance under the null model. The ND, then, was defined as the maximum CDF_i among the neurons in the subpopulation.

Distance and Tuning Ranking Subpopulations and Stimuli for Subselection

For the Retina and PAC datasets, we have access to the spatial locations of the RGC/electrode. For distance-based subselection, we compute the average pairwise distance between neural units for each subpopulation and stimulus. The subpopulations and stimuli are ranked by this distance, and the 10% of subpopulations and stimuli with the smallest average distance are selected.

For tuning-based subselection, the stimuli are ranked for each neural unit based on the mean neural activity (tuning). The rank was used because it is less sensitive to absolute firing rates compared to using the activity per stimuli, which would bias the subselection toward subpopulations and stimuli that contain neural units with high firing rates. We then sort the subpopulations and stimuli by their average tuning rank across subpopulations and calculate percentile statistics for the 10% of subpopulations and stimuli that have the highest tuning ranking.

RESULTS

To assess the optimality of correlated variability for discriminative sensory coding in neural population data, we

used three neural datasets that span animal models, sensory recording areas, and recording modalities (Fig. 1). The newly recorded Retina dataset is calcium imaging from mouse retinal ganglion cells (RGCs) (Fig. 1, D–F). The stimuli are drifting bars at 6 angles, with each stimuli being presented 114 times. The previously recorded V1 dataset is spike-sorted, single-unit electrophysiology recordings in macaque V1 (Fig. 1, G–J) (41). The stimuli are drifting gratings at 12 angles, with each stimuli being presented 200 times. The newly recorded primary auditory cortex (PAC) dataset is high-gamma amplitude ($H\gamma$) from μ ECoG over rat primary auditory cortex (Fig. 1, J–L). The stimuli are tone pips at 30 different frequencies, with each stimuli being presented 60 times. We refer to RGCs/neurons/electrodes as neural units. The neural units have various levels of pairwise noise correlations, ρ , across datasets (Fig. 1, F, I, and L), which is a key quantity for analyzing correlated variability. This diversity likely encompasses the diversity in the brain and allows us to interpret potentially heterogeneous results across datasets in terms of the heterogeneity of pairwise noise correlations. See METHODS for more details on dataset recording and preprocessing.

Novel Methods for Assessing the Optimality of Neural Data for Discriminative Sensory Coding

An abundance of work has aimed to assess whether experimentally observed correlated variability is beneficial or detrimental for discriminative coding (4, 6, 9–12, 16, 19, 24–32). These studies often quantify the discriminability or fidelity of a neural code with the linear Fisher information (LFI; see METHODS) (48), which is a measure of how well the neural activity could be used to discriminate (i.e., decode) different stimuli. We used LFI as it provides a tractable limit on the amount of stimulus “information” that can be transmitted to downstream brain areas (e.g., retina to V1, V1 to V2, etc.) for (linear) decoding, which is important for application to experimental data and connects with prior literature (see METHODS). Typically, the impact of correlated variability is assessed by comparing the experimentally observed LFI to a distribution of LFIs generated from the shuffle null model. Trial-shuffling the data will produce a distribution over covariance matrices $[\Sigma(s)]$ where the pairwise correlations are all centered near zero (Fig. 2A, observed covariance ellipse is filled, corresponding shuffle covariance ellipse is dashed). The neurobiological interpretation of the shuffle null model is that the causes of correlated variability have been removed: recurrent connections among the observed population or between the larger partially observed population have been severed or shared sources have been removed. However, the shuffle null model does not compare the observed correlations to a broad range of potential nonzero correlations. In principle, neural circuits that are not optimal for discriminative coding of sensory stimuli could support a range of covariance structures with nonzero pairwise correlations, many of which could produce higher LFI than having zero correlations. In this case, using the shuffle null model would overestimate the level of optimality in neural data, and therefore it cannot be used to assess the optimality of the experimentally observed correlations. To our knowledge, the optimality of correlated variability for discriminative sensory coding has not been directly evaluated on neural data before.

Our goal was to determine the optimality of correlated variability for discriminative coding in experimental data. To assess optimality of data, one must compare observed data with distributions from null models. A null model should be chosen to adequately span covariance structures. Defining the covariance structures can be motivated from the maximum entropy principle: we can choose features of interest from the data to constrain in the null model distribution and then let the unconstrained parts of the distribution vary according a maximum entropy distribution (i.e., a distribution that is least structured). Defining and choosing the features of interest to constrain may depend on the experimental context, including the types of neurons being recorded, their location in the brain, or the recording modality. Thus, it is beneficial if the parameters of the null model have a biological interpretation. We developed two null models that allow us to assess the optimality of experimental neural responses: the uniform correlation (UC) null model and the factor analysis (FA) null model. These null models have different assumptions about the relationships between neurons. As described below, although there are potentially many possible null models, the UC and FA null models were chosen based on maximum entropy principle considerations of the covariance matrix, the ability to sample from these covariance matrices for hypothesis testing, as well as their interpretation in terms of connectivity.

In the UC null model, neural units maintain their private mean and variance for a particular stimulus but have the freedom to change their multivariate pairwise correlations (ρ). The uniform correlation null model maintains the per-neural unit distributions of activity (i.e., means and variances), like the shuffle null model. In contrast to the shuffle null model, which samples the correlations around zero (Fig. 2A), the uniform correlation null model samples the multivariate correlations uniformly (Fig. 2B, dashed lines are samples with different correlations) (52). Depending on the correlations, the sampled covariances could achieve a range of coding fidelity as assessed by the LFI [Fig. 2D, covariance structures shown below the plot lead the LFI as a function of the scalar pairwise correlation (ρ)]. At extreme values of correlation, the LFI can take on the highest values (Fig. 2D) (22). Biologically, changing the pairwise correlations could be achieved through changing recurrent connectivity within the network of observed neural units (or their larger partially observed network) or by restructuring shared sources of variability. In the uniform null model only the correlations are independent of the mean firing rate. Our biological interpretation is that rather than “breaking” or zeroing-out the recurrent connections as in the shuffle null model, the UC null model instead randomly draws those connection strengths from a uniform distribution.

Motivated by experimental findings that the variability in population responses has private and shared components (7, 14), we also developed a factor analysis (FA) null model. The FA null model decomposes the experimentally observed covariance into independent private variances and shared variability (7, 15). The private variance is fixed (Fig. 2C, gray ellipse in *left inset*), and the shared variability’s weighting on different neural units can be changed (i.e., a rotation; Fig. 2C, dashed teal ellipses in *right inset* are sampled rotations of the shared variability). Biologically, this models each

neuron having fixed private variability and incoming shared variability that could be weighted in different ways. As the shared variability is rotated, the covariance structure varies, and the LFI takes on a smaller range of values than in the UC null model (Fig. 2E, covariance structures shown below the plot generate the LFI as a function of rotation angle, $R(\theta)$). The factor analysis null model weakens the relationship between mean firing rates and variability compared to the uniform null model; however, the private variability is held fixed, which will induce a relationship between the means and the covariance structure. Our biological interpretation of this is that low-rank shared variability from the factors is being generated by unobserved neural activity, unrelated to the task. Analogously to the shuffle and UC null models, the interpretation of rotating this activity is that we are randomizing the connections between the unobserved neurons and the observed neurons while keeping the private variability of the observed neurons fixed. Together, these null models define the potential space of covariances based on two different, complementary biological motivations and constraints, and therefore provide two suitable tests of optimality.

To use the null models for hypothesis testing in neural data, for each neural dataset (retinal ganglion cells, V1 neurons, electrodes over primary auditory cortex), we randomly subsampled the recorded populations of neural units of varying population size (d). We combined the subpopulation with a variety of neighboring stimulus pairings to obtain a subset of the neural responses (Fig. 2F; see METHODS). This subsample of neural responses would be the input to the task of constructing a decoder for neighboring stimuli across trials (Fig. 1, E, H, and K, and Fig. 2G). Rather than explicitly constructing a decoder, we quantified an optimal linear decoder's performance with linear Fisher information (LFI). LFI quantifies the amount of stimulus information that is available in the subpopulation that could be linearly read out by downstream populations. Specifically, for each dataset, we subsampled populations of varying numbers of neural unit population sizes ($d = 3-20$; see METHODS) and calculated the LFI for each. We refer to this quantity as the observed LFI. Next, we sampled the null models 1,000 times for each experimental subpopulation and calculated the LFI for each sample (see METHODS). Thus, for each experimental subpopulation, we obtained a single experimentally observed LFI and a corresponding distribution of LFIs for each null model. The 1,000 null LFIs constitute a null distribution to compare the experimentally observed LFI against. In particular, we define the "null percentile" as the fraction of the 1,000 null LFIs that are less than or equal to the observed LFI. Higher null percentiles indicate that the observed LFI is larger than more samples from the null model.

The Geometry of Correlated Variability Leads to Suboptimal Discriminative Sensory Coding

With the uniform correlation (UC) and factor analysis (FA) null models, we assessed the optimality of the neural code for discriminating stimuli. To characterize the optimality of a wide range of subpopulation and stimulus settings, we performed a large-scale analysis evaluating the LFI in both the experimentally observed data and null models (see METHODS).

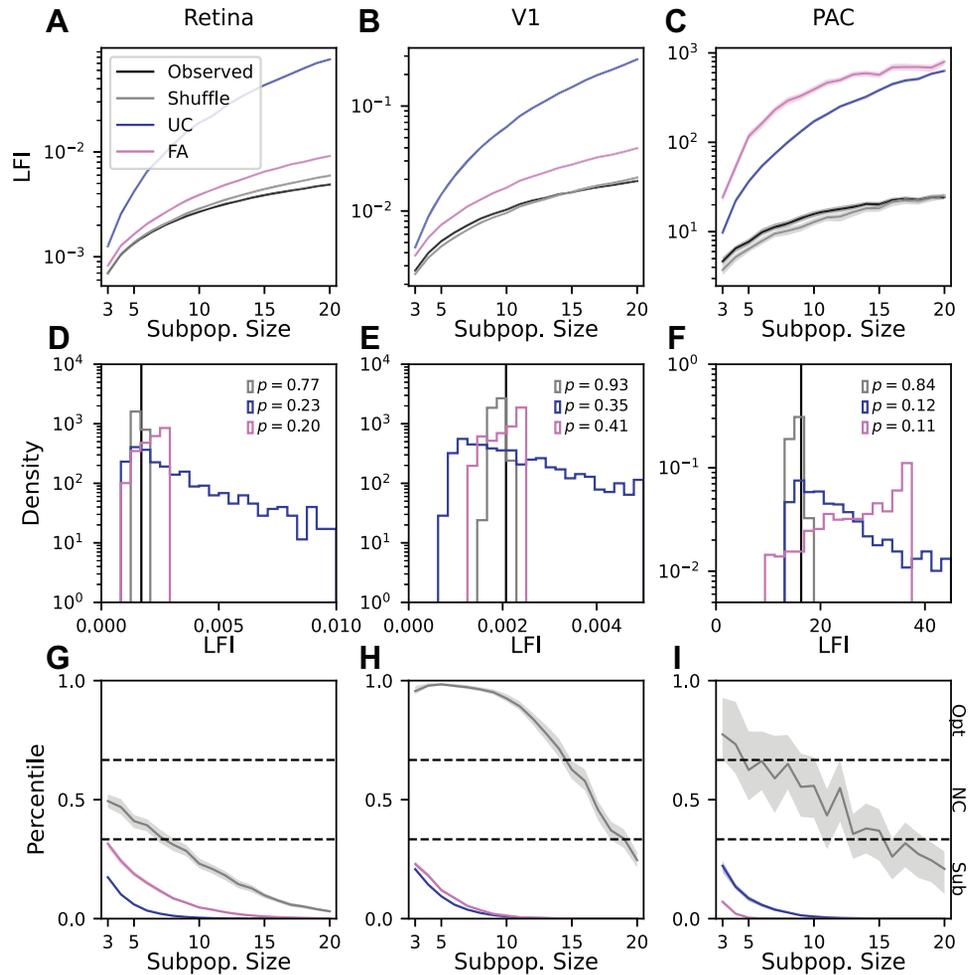
We initially compared the experimentally observed LFI to the distribution of LFI from the null models on a population level. Specifically, for the experimental data, we compute the median LFI across subpopulations and stimuli for varying sizes of subpopulations (Fig. 3, A-C, black lines). For the shuffle, uniform correlation (UC), and factor analysis (FA) null models, we first calculate the median LFI from the null distribution for each subpopulation and stimulus and then report the median across subpopulations and stimuli (Fig. 3, A-C, gray, blue, and orchid lines, respectively).

As expected, the experimentally observed LFIs across subpopulations and stimuli grew with the subpopulation size, indicating that increasing the size of the neural population improved stimulus decoding (Fig. 3, A-C, black lines). Raw LFI values were highest for PAC, lower for V1, and lowest for retina (Fig. 3, A-C, black lines), although LFI carries inverse stimulus parameter-squared units and so care must be taken when comparing LFI values across stimuli. Similar to the experimental data, the median null model LFIs grew with subpopulation size. The shuffle null model exhibited comparable discriminability relative to the experimental LFI at lower subpopulation sizes (Fig. 3, A-C, gray lines). In contrast, both the uniform correlation and factor analysis null models exhibited considerably larger median LFIs than the observed data, with the disparity increasing with subpopulation size. Therefore, on average, the stimuli were more easily discriminable using the covariances sampled from the UC and FA null models than the experimentally determined covariances. We further observed differences across datasets. For example, the factor analysis null model (Fig. 3, A-C, orchid lines) exhibited LFIs similar to the uniform correlation null model for the PAC dataset. However, in the Retina and V1 data, the factor analysis LFIs were more comparable to the observed and shuffle LFIs. Overall, Fig. 3, A-C, demonstrates that the uniform correlation and factor analysis null models produce LFIs that generally exceed the LFIs of the observed data, suggesting that the neural code is suboptimal for sensory discrimination.

Although the differences between the null model LFIs and observed LFIs were large, the preceding analysis was done at a population level rather than comparing each subpopulation and stimulus LFI with its own null distribution. Therefore, we quantified the optimality per subpopulation and stimulus, relative to a null model, with its observed null percentile. To calculate the population optimality measure, the median null percentile across subpopulations and stimuli is taken. A higher null percentile indicates that the observed LFIs are greater than a larger fraction of the null LFIs. To operationalize the notion of population coding optimality, we define three categories for optimality based on the median of the experimental distribution of null percentiles. If the median is $>2/3$ the population is optimal (Opt), if the median is between $1/3$ and $2/3$ the population is near-chance (NC), and if the median is $<1/3$ the population is suboptimal (Sub). Alternative categorizations could be used, but we chose the even splitting into thirds for simplicity (see METHODS for details).

We found that each null model exhibits distinct LFI distributions, with further variation depending on the dataset and subpopulation and stimulus. Example null model distributions for individual three-neuron subpopulations and stimuli are depicted in Fig. 3, D-F (vertical black line indicates

Figure 3. The geometry of correlated variability leads to suboptimal discriminative sensory coding. Each column corresponds to one of the datasets. Color key is preserved across all panels. **A–C:** the median linear Fisher information (LFI) is plotted (solid lines, log scale y-axis) as a function of the subpopulation size (x -axis) for the observed correlated variability and null model samples (colors in legend). Shaded regions indicate the 95% confidence interval (CI) of the median LFI (note that CIs are often comparable to the median line width). FA, factor analysis; PAC, primary auditory cortex; UC, uniform correlation; V1, primary visual cortex. **D–F:** histograms of null LFIs are shown for the shuffle, UC, and FA null models for 1 subpopulation and stimulus for each dataset. The observed LFI is denoted by the black vertical line in each plot. Null percentiles for each null model are reported. **G–I:** median observed subpopulation and stimulus null percentiles are shown (solid lines) as a function of subpopulation size, for each dataset and null model. Shaded regions indicate the 95% CI of the median observed null percentile (note that CIs are often comparable to the median line width). Black dashed lines divide optimal (Opt), near-chance (NC), and suboptimal (Sub) regions.



the experimental LFI; gray, blue, and orchid lines are the shuffle, UC, and FA null model LFI distributions, respectively). The uniform correlation null distributions often have long tails at high LFI and are truncated for visualization. The broader range of LFIs from the UC null model is likely due to the wide range of possible null correlation structures. These examples highlight that the null percentiles can vary across null models for a dataset (Fig. 3, *D–F*, insets). The heterogeneity in observed null percentiles motivated examining their distribution across all subpopulations and stimuli. Thus, for each dataset we computed the distribution of observed null percentiles across the subpopulations and stimuli per subpopulation size ($d = 3$ to $d = 20$). The median observed null percentile (calculated across subpopulations and stimuli) as a function of subpopulation size is shown in Fig. 3, *G–I*. Consistent with other studies (6, 13, 16), we found that the shuffle null model (gray lines) often had large observed null percentiles, indicating that the shuffle null model often suggests benefits of experimentally observed correlations versus having no correlations. However, it would be misleading to interpret these results as a test of optimality. Indeed, compared to the uniform correlation (blue lines) and factor analysis (orchid lines) null models, the experimental data exhibited suboptimal null percentiles (Fig. 3, *G–I*, blue and orchid lines). All null percentiles decreased with subpopulation size, implying that the neural representations become

less optimal for sensory discrimination as the number of neurons increases, becoming essentially 0 by dimension 15 (Fig. 3, *G–I*, blue and orchid lines), indicating that considering larger populations was not necessary. In theory, this decrease is expected, as eventually differential correlations induce information saturation in the populations. However recent work indicates that we should not expect to see the impact of differential correlations until much higher dimensions (33, 49, 57). Indeed, saturation of the LFI was not evident in Fig. 3, *A–C*. This indicates that the suboptimality observed in Fig. 3, *G–I*, is not due to differential correlation but from some other biological cause that is occurring at small population sizes.

Figure 3, *G–I*, also highlight differences across datasets. The shuffle null model had the lowest observed null percentiles among the three datasets for the Retina data, starting near-chance for small subpopulation sizes and dropping below 1/3 around $d = 7$ (Fig. 3*G*, gray lines). For the V1 data, the shuffle null model clearly exhibited the highest observed null percentiles, indicating the coding benefits of correlations compared to zero correlations for small subpopulation sizes up to $d = 15$ (Fig. 3*H*, gray lines). In the primary auditory cortex data, the shuffle null model exhibited intermediate observed null percentiles, with a larger spread in confidence intervals, indicating a higher heterogeneity in the null percentiles (Fig. 3*I*, gray shaded region). Meanwhile,

the observed null percentiles for the uniform correlation and factor analysis null models were more similar across the three datasets, with slightly different magnitudes. In particular, the retinal data exhibited the largest observed null percentiles for the factor analysis null model (orchid), whereas the PAC data exhibited the smallest, going to zero around $d = 5$. The uniform correlation null model (blue) had the lowest null percentiles for the Retina dataset and similar null percentiles for the V1 and PAC datasets. This behavior roughly tracked the distribution of pairwise correlations among the three datasets (Fig. 1, E, H, and K): the retinal data had the lowest average noise correlation, and the PAC data had the highest average noise correlation. Critically, across all datasets and subpopulation sizes, the null percentiles for both the uniform correlation and factor analysis null models were $<1/3$. This indicates that the geometry of correlated variability leads to suboptimal discriminative sensory coding and that the suboptimality becomes more pronounced with increasing neural population size.

Optimal Correlated Variability for Sensory Discrimination Is Typically Biologically Inaccessible

The results of the preceding section indicate that the geometry of correlated variability is highly suboptimal, as opposed to near-chance or optimal. We next sought to understand why this was the case. For the uniform correlation model, we summarize findings about optimal correlations from Hu et al. (22). For the factor analysis model, we compared the structure of the observed covariances to those of the optimal covariances.

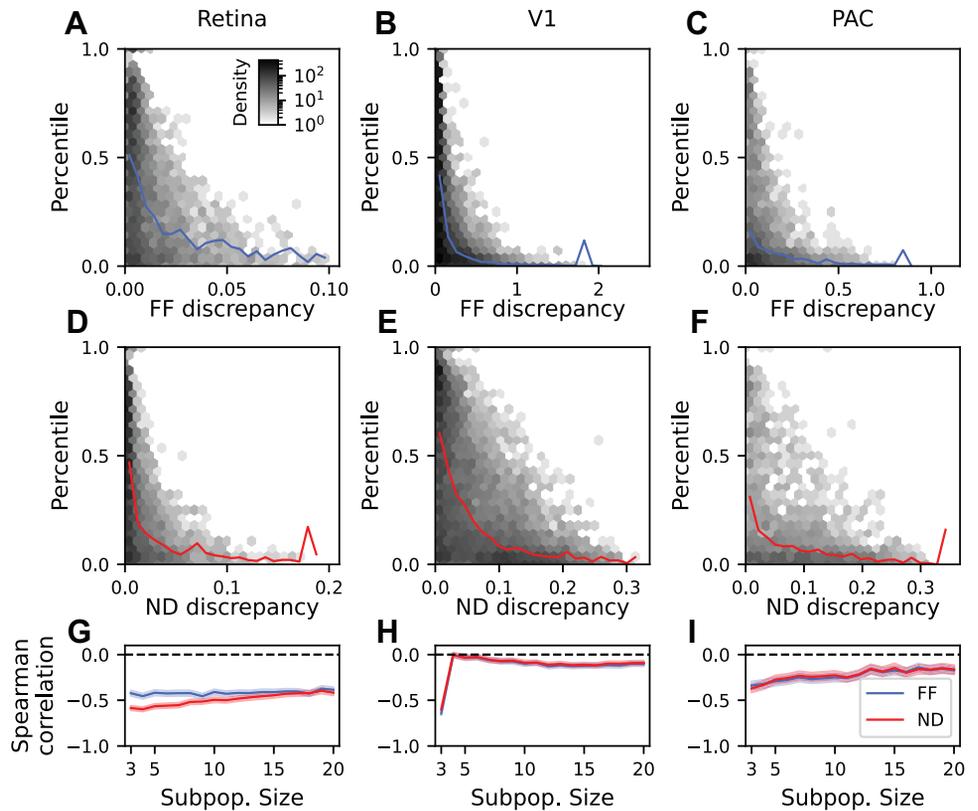
When the per-neural unit variability is fixed, as in the shuffle and uniform correlation null models, Hu et al. (22) showed that the optimal covariance structure will lie on the boundaries of the allowed values of ρ for several measures of coding fidelity, including the LFI (Fig. 2D). The authors discussed that points on the boundary may fall outside of biologically allowed regions. Consistent with this, we found that optimal correlation matrices for the uniform correlation null model often had absolute pairwise correlations close to 1, which was never observed in the single-unit (Retina and V1) experimental datasets (data not shown) although the PAC datasets had more pairwise correlations near 1. Thus, the optimal correlated variability structure suggested by the uniform correlation null model may be biologically inaccessible. Meanwhile, the factor analysis model allows the distribution of highest pairwise correlations to be modified (and generally increased) but does not extend near 1, suggesting that the distribution of noise correlations achieved by the factor analysis null model is more biologically realistic.

Both the shuffle and uniform correlation null models will necessarily reproduce the observed single-unit statistics, because they only change the correlations. Therefore, both of these null models will reproduce the Fano factors (FFs, $\frac{\text{variance}}{\text{mean}}$) and negative densities (ND, fraction of activity below the smallest responses of the experimental activity) of the observed data. The factor analysis null model, however, can produce covariance ellipses that have different single-unit distributions. Thus, some FA-optimal covariances may orient variance in the negative or low-activity regions of the

neural response. For the factor analysis null model, we quantified the degree to which the biological inaccessibility of optimal covariances related to the percentiles of the experimental data for each subpopulation and stimulus. The Fano factor quantifies the variability of neural units relative to their average activity. Typically, Fano factors for single-unit firing rates have been observed to be near 1 (58–61), in line with the approximately Poisson nature of firing rates. Thus, a large deviation from the Fano factors observed in the experimental data indicates that the single-unit properties of the optimal covariances are biologically implausible. First, we examined whether the observed Fano factors diverged from the Fano factors achieved by the FA-optimal covariance on each subpopulation and stimulus via their absolute log ratio (see METHODS). Large values of this quantity indicate greater difference between optimal and experimental single-unit distributions, suggesting less biological plausibility. Relatedly, a sample covariance that has negative neural activity can be interpreted as less biologically plausible, because negative activity is either unachievable (for single-unit count variables) or highly unlikely (calcium imaging $\Delta F/F$ or baseline z -scored μECOG). Therefore, the second quantity we examined was the absolute difference in negative density (ND), which captures the degree to which the FA-optimal covariance had negative neural activity (see METHODS). Larger values of the negative density imply less biological plausibility. We used these two measures of biological plausibility to assess when the observed neural responses can be optimal according to the FA null model.

We determined whether the Fano factor (FF) and negative density (ND) distributions of the optimal covariances from the FA null model are related to the suboptimality of the experimentally observed neural code. To do this, we directly compared the optimal FA null model Fano factors to the experimental Fano factors in Fig. 4, A–C. Across subpopulations and stimuli, for $d = 3$, Fig. 4, A–C shows two-dimensional (2-D) histograms of the absolute log ratio of Fano factors against the FA percentile, with darker colors corresponding to higher log density of samples. For each histogram, we additionally plot the median percentile as a function of the log ratio in blue. We found that when the Fano factors closely matched (i.e., the log ratio was close to 0), the percentiles spanned a broad range between 0 and 1 (median percentiles: 0.51, 0.41, 0.15 for the lowest bin across datasets). However, FA-optimal covariances commonly deviated from the observed Fano factors, and when they did, the observed percentiles dropped below 0.5 and were often near 0. Likewise, for negative density (ND), we directly compared the optimal FA null model NDs to the experimentally observed NDs in Fig. 4, D–F. Across subpopulations and stimuli, for $d = 3$, Fig. 4, D–F, show 2-D histograms of the absolute difference in NDs against the FA percentile, with darker colors corresponding to higher log density. For each histogram, we additionally plot the median percentile as a function of ND difference in red. We found that when the difference was close to zero the percentiles spanned a broad range between 0 and 1 (median percentiles: 0.47, 0.60, 0.31 for the lowest bin across datasets). However, the NDs of FA-optimal covariances commonly deviated from the observed ND, and when they did, the experimentally observed percentiles were typically closer to 0. Thus, as the biological

Figure 4. Optimal correlated variability for sensory discrimination is typically biologically inaccessible. Each column corresponds to a separate dataset. Two-dimensional (2-D) histograms are plotted with a log-density color scale with shared color bar. Color key in *I* is shared across panels. *A–C*: 2-D histogram across subpopulations and stimuli of the observed null percentile under the factor analysis (FA) null model vs. the absolute log ratio of the observed and FA-optimal covariance Fano factors (FFs) for $d = 3$. Blue line is the median binned null percentile as a function of the absolute log ratio of observed and FA-optimal covariance FFs. PAC, primary auditory cortex; V1, primary visual cortex. *D–F*: 2-D histogram across subpopulations and stimuli of the null percentile under the FA null model vs. the absolute difference of negative densities (NDs) of the observed and FA-optimal covariance FFs for $d = 3$. Red line is the median binned null percentile as a function of the absolute difference in NDs. *G–I*: the Spearman correlation coefficient between the null percentile and absolute log FF ratio or absolute difference of NDs, respectively, is shown as a function of subpopulation size. Dashed black line indicates zero correlation.



accessibility of the optimal covariance decreased, so did the optimality of the observed neural code.

We summarized the relationship between biological plausibility and percentile for both FF and ND. At each subpopulation size d , we calculated the Spearman rank correlation between the observed percentile and each measure of biological plausibility (Fig. 4, *G–I*). For each dataset, we observed negative correlations that were significantly lower than zero across subpopulation sizes ($P < 10$, 1-sample t test). These negative correlations imply that observed percentiles are smaller (i.e., the neural code is more suboptimal) when optimal correlated variability is biologically inaccessible. Together, these results indicate that the optimal covariances under the FA null model for neural populations of 3 or more neurons ($d \geq 3$) are not biologically accessible.

Optimal Subpopulations for Sensory Discrimination Are Exponentially Small

The results in the preceding section show that a majority of experimental subpopulations and stimuli could not attain optimal covariances according to the UC and FA null models because of biological constraints. However, it is possible that although a majority of experimental subpopulations and stimuli are suboptimal there is a subset that are optimal, and these specific subpopulations are somehow utilized by the nervous system. If this was the case, the uniform sampling strategy over neural units might underestimate optimality as utilized by the nervous system. For example, in the retina, if we are imagining that a downstream region like V1 is discriminating the stimuli, then a more retinotopic sampling strategy, where retinal ganglion cells are more likely to be considered in a subpopulation if they are located spatially near each other in the

retina, would be preferable. Alternatively, synaptic learning rules (e.g., Hebbian plasticity, “neurons that fire together, wire together”) in downstream areas may select for neural populations that are tuned for similar stimuli. The responses to the preferred stimuli would be high, and therefore we expect less Fano factor and negative density violation. Thus, it is possible that subpopulations and stimuli subselected by these criteria will be more optimal than subpopulations and stimuli sampled uniformly.

To test whether biologically motivated sampling of subpopulations and stimuli improved the null percentiles, we performed distance- and tuning-based subselection of the neural populations. For the Retina and PAC datasets, we had access to the spatial locations of the RGCs/electrodes. We subselected 10% of subpopulations and stimuli with the smallest average physical distance. Similarly, we subselected the 10% of subpopulations and stimuli that had the most similar preferred stimuli (see METHODS for details on subselection). We found that distance-based subselection did not reveal an optimal or near-chance subset of subpopulations and stimuli (Fig. 5, *A* and *C*, dotted lines and hatched shaded regions). Similarly, for the Retina and V1 datasets, the tuning-based subselection did not reveal an optimal subset of subpopulations and stimuli and the percentiles only improved to near-chance for the PAC dataset at $d = 3$ (Fig. 5, *A–C*, solid lines and shaded regions). Furthermore, subselection directly based on the FF and ND criteria also did not find optimal or near-chance percentiles (results not shown).

Although these subselection criteria are biologically motivated, the previous results do not address whether any subpopulation of the neural units across stimuli has optimal null percentiles and, if so, how small the subpopulation is.

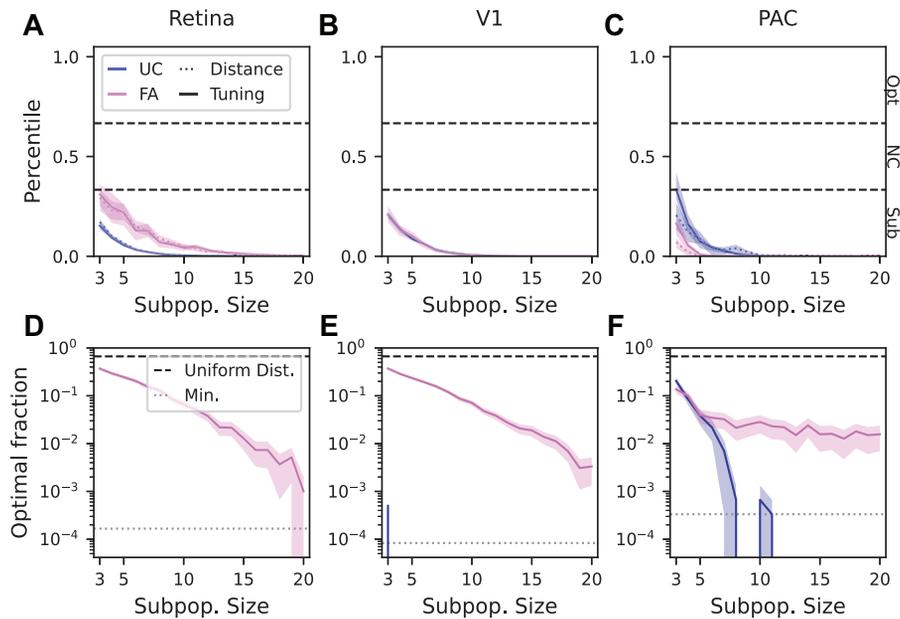


Figure 5. Optimal subpopulations are exponentially small. Color key in A is shared across panels. A–C: for the uniform correlation (UC) and factor analysis (FA) null models, subpopulations and stimuli were subselected to maximize the units' tuning (solid lines, highest 10% subselected). Additionally, for the Retina and primary auditory cortex (PAC) datasets, subpopulations and stimuli were subselected to minimize the average pairwise distance between the retinal ganglion cell (RGC) regions of interest (ROIs) in a subpopulation and stimulus (dashed lines, lowest 10% subselected). The median percentiles are shown as a function of dimension. Black dashed lines indicate the 1/3 and 2/3 null percentile range. Shaded regions indicate the 95% confidence interval (CI) of the median percentiles. NC, near-chance; Opt, optimal; Sub, suboptimal; V1, primary visual cortex. D–F: for each subpopulation size, the largest possible fraction of subpopulation and stimulus percentiles such that their median is $\geq 2/3$ is plotted. Shaded regions indicate 95% CI. For the uniform correlation null model, subpopulation sizes where no samples exceeded the 2/3 threshold are not plotted. Black dashed line indicates the optimal fraction if null percentiles were drawn from a uniform distribution. Gray dotted line indicates the minimum nonzero optimal fraction that can be estimated because of finite sampling.

Intuitively, given the combination of a large enough neural population, variety of stimuli, and enough subpopulations and stimuli, one would expect at least a small fraction of the subpopulations and stimuli to have optimal null percentile statistics by chance. To estimate the size of the optimal subpopulation, we calculated the optimal fraction of the neural population, that is, the largest fraction of subpopulations and stimuli that could be retained and still achieve optimal null percentile statistics (median $\geq 2/3$) (Fig. 5, D–F). If the optimal fraction is smaller, optimal subpopulations are more rare. As a reference, if the distribution of null percentiles was uniform, the largest two-thirds of the percentiles could be retained and their median would be 2/3, which is optimal (Fig. 5, D–F, black dashed line). At $d = 3$ for the FA null model (Fig. 5, D–F, orchid line), across datasets between 14% and 37% of the entire population was optimal if subselected. The optimal fraction according to the FA null model dropped below 10% by $d = 4$ –9 and below 2% by $d = 13$ –15 across datasets. At larger subpopulation sizes, the optimal subpopulation continued to become exponentially small, although the PAC dataset had a slower decrease. According to the uniform correlation null model, for the Retina and V1 datasets, less than $\sim 0.1\%$ of the population was optimal for sensory discrimination since almost no subpopulation was found from the finite samples. At $d = 3$ for the PAC dataset, 20% of subpopulations and stimuli would be considered optimal, but that dropped below 1% by $d = 7$ and continued to decrease to the smallest possible estimated value by $d = 12$, since no subpopulations were found for larger

subpopulation sizes. Finally, an alternative analysis of peaks in the null percentiles near 1 in excess of what would be expected from a uniform distribution confirmed that there were exponentially small optimal subpopulations. Together these results show that correlated variability is suboptimal for sensory discrimination in the neural recordings considered here. Furthermore, biologically motivated selection criteria are not able to find the exponentially small optimal subpopulations.

DISCUSSION

Determining the principles of the neural code is critical for a complete understanding of brain function (62). Correlated variability is prevalent in neural recordings and has been the subject of numerous studies seeking to understand its mechanistic sources and implication for neural coding. Many previous studies have found that the experimentally observed correlations can be a benefit to discriminative sensory coding compared to having zero correlations (6, 13, 16, 24, 25, 27). This suggests that the correlated variability could in fact be optimal. However, the shuffle null model used in these studies is not able to assess optimality. Rummyantsev et al. (33) showed a case where correlated variability was substantially worse than the shuffle null model, suggesting that noise correlations can be detrimental to sensory coding. Since the shuffle null model covers a more restrictive distribution of covariances than the FA or UC null models used here, if

observed noise correlations have LFIs lower than those from the shuffle null model, it is likely that they will also have LFIs lower than those from the FA or UC null model. However, this is only an indirect test of optimality according to the FA or UC null model. Thus, to the best of our knowledge, the optimality of correlated variability for discriminative sensory coding in neural data has not previously been directly assessed.

Here, we developed two null models that allow the discriminative optimality of experimentally observed correlated variability to be directly assessed: the uniform correlation (UC) and factor analysis (FA) null models. Using these null models, we found that the experimentally observed neural activity across three datasets was consistently suboptimal. As more neural units were included in the neural population, the neural populations became more suboptimal, often reaching the floor of our sampling by neural unit subpopulations of size 15. This implies that considering larger populations would not change our conclusions. To more fully understand the suboptimality, we evaluated the characteristics of the optimal covariance and found that a consistent picture emerged: for a majority of neural subpopulations, the optimal covariance is biologically inaccessible. We then used biologically motivated subselection criteria to assess whether there were subpopulations with optimal coding statistics. We found that subsampling using criteria based on the tuning of units or the spatial location of the units does not result in increased discriminative coding optimality. Finally, we showed that optimal subpopulations based on post hoc selection became exponentially small as the size of the neural population increases. Thus, we conclude that in the early sensory areas studied here, the geometry of correlated variability leads to highly suboptimal discriminative coding.

We observed suboptimal discriminative coding performance as assessed by both the uniform correlation and factor analysis null models. However, the magnitude of the suboptimality, as measured by the observed null percentiles, differed across null models and datasets. The observed null percentiles for the uniform correlation null model had a small trend from low to high for the retina data, the V1 data, and the PAC data, respectively. This trend tracks with the distribution of noise correlations in each dataset (Fig. 1, *F*, *I*, and *L*), with the Retina dataset exhibiting, on average, the smallest magnitude noise correlations and the PAC datasets exhibiting the largest. The smaller range of noise correlations exhibited by the retina suggests that there may be stronger biological restrictions on its correlated variability compared to V1 and PAC. The observed null percentiles for the factor analysis null model trend from just below near-chance to highly suboptimal from retina to PAC. Thus, the larger correlations and more suboptimal discriminative coding performance indicate that shared variability in V1 and PAC is more likely to interfere with discriminative sensory coding. The retina and V1 recording modalities (calcium imaging and single-unit electrophysiology, respectively) measure putative single-unit activity where correlated variability in the recordings corresponds to correlated single-neuron activity. A discriminative model describes the ability to predict an external signal or stimulus parameter from a population of neural activity. The impact of correlated

variability across a population of neurons in response to a stimulus is typically conceptualized as a question about how much information is contained about the stimulus in that population that may be available to a downstream brain area (e.g., information in V1 that is available in V2) (1, 2, 4, 5, 9, 13, 19, 25, 29). As such, understanding the (sub)optimality of the neural code at the level of multiple single units directly addresses discrimination of sensory signals as a normative theory in early sensory areas. On the other hand, the correlated variability in the μ ECoG recordings in PAC is likely due to a combination of the correlations between the neural populations under each electrode and local tissue conduction (43, 44). Because of this, the optimality of the high-gamma correlated variability recorded with μ ECoG is a coarse-grained signal that may not be read out by any downstream cortical area. However, these results are important for understanding when correlated variability in brain signals may limit the accuracy of clinical ECoG-based brain-computer interfaces in humans. Indeed, our results suggest that ECoG is a viable recording methodology for brain-computer interfaces for extracting information about the external world from the brain. This is amplified by the fact that ECoG recordings are regularly performed in human neurological patients (63, 64).

Our results indicate that correlated variability across diverse sensory neural populations is highly suboptimal for transmitting information to subsequent brain areas for discriminating the stimulus. This raises a question about what computational role variability plays in sensory processing and what insights have been gained by comparison to null models. Many studies of correlated variability, including ours, consider the impact of correlated variability from a discriminative sensory coding perspective. However, other normative perspectives exist. In generative (e.g., Bayesian) models of sensory processing (65), correlated variability could correspond to sampling from a relevant (posterior) distribution. In this case, correlated variability would be informative for understanding the structure of uncertainty in sensory processing, rather than nuisance variability as in the decoding perspective. Likewise, neural systems likely have other important constraints or ethological goals. Making decisions or generating behavior based on sensory information may be optimized by different correlation structures versus a purely discriminative framework (66). For example, Valente et al. (66) found that single-trial responses in posterior parietal cortex that have higher noise correlations also have more correct choices, contrary to expectation. They model this finding with a readout network that computes an additional nonlinear “consistency” value across the population in addition to the linear sensory information for use in decision making. Huang and Lisberger (67) showed that correlated variability in middle temporal visual area could plausibly be the cause of variability in smooth-pursuit eye movements. Even within the normative discriminative framework, correlated variability that facilitates discrimination as assessed by the LFI may not be the same as the correlated variability that facilitates information propagation or learning in more realistic nonlinear, noisy networks (5, 66, 68). LFI is a local measure, and if we consider the so-called larger threshold errors (69) then results from the different null models could change. In these contexts,

our formalism for creating null models could be used to test the optimality of neural codes, although as the assumptions on linear decoding are relaxed, it may become difficult to make theoretical predictions that hold generally. The null models we developed could be utilized to test optimality of other proposed computations of sensory populations and may provide insights into the constraints (embodied in the degrees of freedom in the null models) for realizing that optimality. Regardless of the role, if any, of correlated variability in neural computations, an important contribution of our work is to identify what role it is not playing: correlated variability is not optimally enhancing information transmission for discriminating sensory stimuli.

The null models we proposed both have parameterizations that are interpreted in a fully Gaussian model. There are likely many other possible combinations of covariance parameterizations and constraints that could be used to generate null distributions, including a FA-like model where the marginal statistics are preserved. Although we have not explored this particular direction in detail, it is unclear that this has an interpretation as a maximum entropy distribution or a simple interpretation in terms of data-estimated parameters that are kept fixed while others are varied. Generalized linear models (70, 71) or correlated multivariate distributions with binary-spike or spike-count distributions (72–74) could potentially better model nonlinearities between the parameters of the model and the non-Gaussian neural responses, which can impact estimates of neural coding optimality. To assess optimality in these models when fit to data, a similar formalism for generating null models is needed, where certain parts of the parameterization are fixed and others are given a null distribution. However, the independent parameterization of the mean responses (tuning) and correlated variability is a unique feature of the multivariate Gaussian distribution. Therefore, new mathematical results would be needed to directly study the impact of non-Gaussian correlated variability on neural coding. A broader set of null distributions could similarly be used in phenomenological models of correlated variability that combine tuning and various types of (correlated) noise (6, 21, 28, 75) or in mechanistic models, which attempt to simulate some aspects of the neural circuit that lead to correlated variability (5, 15, 16, 76).

Future directions that build off of our results could investigate the issue of optimality of correlated variability with null models that target other neurobiological properties or scales of organization. Some examples include learning rules with naturalistic input statistics, network models with neurobiologically motivated recurrent connectivity, or sensory inputs of dimensionality different from the target area. Also of interest is understanding how variability of neural responses can be used as a way to explore during learning (77). We note that the results of such investigations will certainly depend on the details of, e.g., recurrent connectivity, the neurobiology of which is poorly understood. For example, one approach to understand population dynamics is to set up a recurrent neural network with random connectivity and train it with backpropagation to perform a task. However, it is known that connectivity is highly structured in the brain (78) and depends on the cell type. Additionally, the biological learning rules that shape neural circuits to perform specific functions are poorly understood. A better

understanding of the neurobiological constraints in these models would likely be required to generate null models that tightly estimate optimality of various functions in experimental neural datasets.

Correlated variability has been shown to be impacted by behavior and brain states. For example, it has been observed that behavior such as running, whisking, and pupil diameter are encoded in V1 and other brain areas (7). In these contexts, the behavioral subspaces could be estimated directly (as in Ref. 7) and their optimality could be assessed with the FA null model. In experiments with visual attention, it has been shown that attention can modulate both the within-area and between-area correlated variability (30, 79, 80), which can lead to better sensory discrimination or better communication of information as assessed by the shuffle null model. The null models developed here could be used to assess whether the modulation due to attention or learning changes the optimality of the correlated variability. Emerging neural recording technologies will allow neuroscientists to simultaneously record from a larger fraction of neurons in a region and more regions, all while the animals are performing naturalistic behaviors. Given these possibilities, the biological origins of correlated variability and how they are modulated by neural circuitry can be further traced and evaluated.

In summary, we found that the geometry of correlated variability in sensory areas leads to highly suboptimal coding for transmission of information to discriminate the stimulus. Given the consistency of the findings across datasets, we expect that our results would hold true in other organisms, sensory areas, and experimental paradigms. Investigated more broadly, understanding the optimality of correlated variability could lead to a better understanding of the sources of variability in neural circuits and biological constraints that lead to suboptimality. Furthermore, quantitatively evaluating normative theories allows us to adjudicate between competing proposed functions of sensory systems, for example, efficient coding versus predictive information coding.

DATA AVAILABILITY

The preprocessed and trialized Retina and trialized PAC datasets are available at <https://zenodo.org/records/14342290>. The raw PAC data are currently available at <https://crcns.org/data-sets/ac/ac-7> (R32_B7.nwb). The V1 data are currently available at <https://crcns.org/data-sets/vc/pvc-11>. Code to reproduce the analysis and figures is available at a publicly available GitHub repository at https://github.com/BouchardLab/noise_correlations.

ACKNOWLEDGMENTS

We thank the Neural Systems and Data Science Lab and Frederic Theunissen for feedback.

GRANTS

J.A.L. was supported by the Lawrence Berkeley National Laboratory (LBNL) Lab Directed Research and Development Grant (LDRD) “Deep Learning for Science” and the Weill Institute for Neuroscience at University of California, San Francisco (UCSF) (Bouchard). P.S.S. was supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate (NDSEG) Fellowship Program. M.E.D. was supported by an LBNL LDRD. M.T.S. was supported by a National Science

Foundation Graduate Research Fellowship (DGE 1752814). K.E.B. was supported by Department of Energy (DOE) Advanced Scientific Computing Research (ASCR) (FP00009697), NIH (RNS118648A), the Kavli Institute, the Weill Institute for Neuroscience, and the LBNL LDRD “Coordination of Distributed Cortical Circuits.” This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

J.A.L., P.S.S., and K.E.B. conceived and designed research; M.E.D., M.T.S., and K.E.B. performed experiments; J.A.L. and P.S.S. analyzed data; J.A.L., P.S.S., and K.E.B. interpreted results of experiments; J.A.L., P.S.S., and K.E.B. prepared figures; J.A.L., P.S.S., and K.E.B. drafted manuscript; J.A.L., P.S.S., and K.E.B. edited and revised manuscript; J.A.L., P.S.S., M.E.D., M.T.S., and K.E.B. approved final version of manuscript

REFERENCES

- Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A. Correlations and neuronal population information. *Annu Rev Neurosci* 39: 237–256, 2016. doi:10.1146/annurev-neuro-070815-013851.
- Azeredo da Silveira R, Rieke F. The geometry of information coding in correlated neural populations. *Annu Rev Neurosci* 44: 403–424, 2021. doi:10.1146/annurev-neuro-120320-082744.
- Averbeck BB, Latham PE, Pouget A. Neural correlations, population coding and computation. *Nat Rev Neurosci* 7: 358–366, 2006. doi:10.1038/nrn1888.
- Cohen MR, Kohn A. Measuring and interpreting neuronal correlations. *Nat Neurosci* 14: 811–819, 2011. doi:10.1038/nn.2842.
- Zylberberg J, Pouget A, Latham PE, Shea-Brown E. Robust information propagation through noisy neural circuits. *PLoS Comput Biol* 13: e1005497, 2017. doi:10.1371/journal.pcbi.1005497.
- Franke F, Fiscella M, Sevelev M, Roska B, Hierlemann A, da Silveira RA. Structures of neural correlation and how they favor coding. *Neuron* 89: 409–422, 2016. doi:10.1016/j.neuron.2015.12.037.
- Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD. Spontaneous behaviors drive multidimensional, brain-wide activity. *Science* 364: 255, 2019. doi:10.1126/science.aav7893.
- Zohary E, Shadlen MN, Newsome WT. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370: 140–143, 1994 [Erratum in *Nature* 371: 358, 1994]. doi:10.1038/370140a0.
- Dichter BK, Bouchard KE, Chang EF. Dynamic structure of neural variability in the cortical representation of speech sounds. *J Neurosci* 36: 7453–7463, 2016. doi:10.1523/JNEUROSCI.0156-16.2016.
- Kohn A, Smith MA. Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J Neurosci* 25: 3661–3673, 2005. doi:10.1523/JNEUROSCI.5106-04.2005.
- Smith MA, Kohn A. Spatial and temporal scales of neuronal correlation in primary visual cortex. *J Neurosci* 28: 12591–12603, 2008. doi:10.1523/JNEUROSCI.2929-08.2008.
- Ruff DA, Cohen MR. Stimulus dependence of correlated variability across cortical areas. *J Neurosci* 36: 7546–7556, 2016. doi:10.1523/JNEUROSCI.0504-16.2016.
- Montijn JS, Meijer GT, Lansink CS, Pennartz CM. Population-level neural codes are robust to single-neuron variability from a multidimensional coding perspective. *Cell Rep* 16: 2486–2498, 2016. doi:10.1016/j.celrep.2016.07.065.
- Deweese MR, Zador AM. Shared and private variability in the auditory cortex. *J Neurophysiol* 92: 1840–1855, 2004. doi:10.1152/jn.00197.2004.
- Sachdeva PS, Livezey JA, DeWeese MR. Heterogeneous synaptic weighting improves neural coding in the presence of common noise. *Neural Comput* 32: 1239–1276, 2020. doi:10.1162/neco_a_01287.
- Zylberberg J, Cafaro J, Turner MH, Shea-Brown E, Rieke F. Direction-selective circuits shape noise to ensure a precise population code. *Neuron* 89: 369–383, 2016. doi:10.1016/j.neuron.2015.11.019.
- Huang C, Ruff DA, Pyle R, Rosenbaum R, Cohen MR, Doiron B. Circuit models of low-dimensional shared variability in cortical networks. *Neuron* 101: 337–348.e4, 2019. doi:10.1016/j.neuron.2018.11.034.
- Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron* 74: 30–39, 2012. doi:10.1016/j.neuron.2012.03.016.
- Abbott LF, Dayan P. The effect of correlated variability on the accuracy of a population code. *Neural Comput* 11: 91–101, 1999. doi:10.1162/089976699300016827.
- Yoon H, Sompolinsky H. The effect of correlations on the Fisher information of population codes. *Adv Neural Inf Process Syst* 167–173, 1999.
- Ecker AS, Berens P, Tolias AS, Bethge M. The effect of noise correlations in populations of diversely tuned neurons. *J Neurosci* 31: 14272–14283, 2011. doi:10.1523/JNEUROSCI.2539-11.2011.
- Hu Y, Zylberberg J, Shea-Brown E. The sign rule and beyond: boundary effects, flexibility, and noise correlations in neural population codes. *PLoS Comput Biol* 10: e1003469, 2014. doi:10.1371/journal.pcbi.1003469.
- Bujan AF, Aertsen A, Kumar A. Role of input correlations in shaping the variability and noise correlations of evoked activity in the neocortex. *J Neurosci* 35: 8611–8625, 2015. doi:10.1523/JNEUROSCI.4536-14.2015.
- Cafaro J, Rieke F. Noise correlations improve response fidelity and stimulus encoding. *Nature* 468: 964–967, 2010. doi:10.1038/nature09570.
- Graf AB, Kohn A, Jazayeri M, Movshon JA. Decoding the activity of neuronal populations in macaque primary visual cortex. *Nat Neurosci* 14: 239–245, 2011. doi:10.1038/nn.2733.
- Montani F, Kohn A, Smith MA, Schultz SR. The role of correlations in direction and contrast coding in the primary visual cortex. *J Neurosci* 27: 2338–2348, 2007 [Erratum in *J Neurosci* 27: 5835, 2007]. doi:10.1523/JNEUROSCI.3417-06.2007.
- Ruda K, Zylberberg J, Field GD. Ignoring correlated activity causes a failure of retinal population codes. *Nat Commun* 11: 4605, 2020. doi:10.1038/s41467-020-18436-2.
- Lin I-C, Okun M, Carandini M, Harris KD. The nature of shared cortical variability. *Neuron* 87: 644–656, 2015. doi:10.1016/j.neuron.2015.06.035.
- Averbeck BB, Lee D. Effects of noise correlations on information encoding and decoding. *J Neurophysiol* 95: 3633–3644, 2006. doi:10.1152/jn.00919.2005.
- Ruff DA, Cohen MR. Attention increases spike count correlations between visual cortical areas. *J Neurosci* 36: 7523–7534, 2016. doi:10.1523/JNEUROSCI.0610-16.2016.
- Beaman CB, Eagleman SL, Dragoi V. Sensory coding accuracy and perceptual performance are improved during the desynchronized cortical state. *Nat Commun* 8: 1308, 2017. doi:10.1038/s41467-017-01030-4.
- Downer JD, Rapone B, Verhein J, O'Connor KN, Sutter ML. Feature-selective attention adaptively shifts noise correlations in primary auditory cortex. *J Neurosci* 37: 5378–5392, 2017. doi:10.1523/JNEUROSCI.3169-16.2017.
- Rumyantsev OI, Lecoq JA, Hernandez O, Zhang Y, Savall J, Chrapkiewicz R, Li J, Zeng H, Ganguli S, Schnitzer MJ. Fundamental bounds on the fidelity of sensory cortical coding. *Nature* 580: 100–105, 2020. doi:10.1038/s41586-020-2130-2.
- Elsayed GF, Cunningham JP. Structure in neural population recordings: an expected byproduct of simpler phenomena? *Nat Neurosci* 20: 1310–1318, 2017. doi:10.1038/nn.4617.
- Tiriac A, Smith BE, Feller MB. Light prior to eye opening promotes retinal waves and eye-specific segregation. *Neuron* 100: 1059–1065.e4, 2018. doi:10.1016/j.neuron.2018.10.011.
- Stosiek C, Garaschuk O, Holthoff K, Konnerth A. In vivo two-photon calcium imaging of neuronal networks. *Proc Natl Acad Sci USA* 100: 7319–7324, 2003. doi:10.1073/pnas.123223100.

37. **Pologruto TA, Sabatini BL, Svoboda K.** ScanImage: flexible software for operating laser scanning microscopes. *Biomed Eng Online* 2: 13–19, 2003. doi:10.1186/1475-925X-2-13.
38. **Caval-Holme F, Zhang Y, Feller MB.** Gap junction coupling shapes the encoding of light in the developing retina. *Curr Biol* 29: 4024–4035.e5, 2019 [Erratum in *Curr Biol* 30: 185, 2020]. doi:10.1016/j.cub.2019.10.025.
39. **Tiriac A, Bistrong K, Feller M.** Retinal waves but not visual experience are required for development of retinal direction selectivity maps (Preprint). *bioRxiv* 2021.03.25.437067, 2021. doi:10.1101/2021.03.25.437067.
40. **Teeters JL, Harris KD, Millman KJ, Olshausen BA, Sommer FT.** Data sharing for computational neuroscience. *Neuroinformatics* 6: 47–55, 2008. doi:10.1007/s12021-008-9009-y.
41. **Kohn A, Smith MA.** Utah array extracellular recordings of spontaneous and visually evoked activity from anesthetized macaque primary visual cortex (V1). 2016. doi:10.6080/KONC5Z4X.
42. **Kelly RC, Smith MA, Kass RE, Lee TS.** Local field potentials indicate network state and account for neuronal response variability. *J Comput Neurosci* 29: 567–579, 2010. doi:10.1007/s10827-009-0208-9.
43. **Dougherty ME, Nguyen AP, Baratham VL, Bouchard KE.** Laminar origin of evoked ECoG high-gamma activity. *Annu Int Conf IEEE Eng Med Biol Soc* 2019: 4391–4394, 2019. doi:10.1109/EMBC.2019.8856786.
44. **Baratham VL, Dougherty ME, Ledochowitsch P, Maharbiz MM, Bouchard K.** Columnar localization and laminar origin of cortical surface electrical potentials (Preprint). *bioRxiv* 2021.08.16.456540, 2021. doi:10.1101/2021.08.16.456540.
45. **Cover TM, Thomas JA.** *Elements of Information Theory*. Hoboken, NJ: John Wiley & Sons, 2012.
46. **Sompolinsky H, Yoon H, Kang K, Shamir M.** Population coding in neuronal systems with correlated noise. *Phys Rev E Stat Nonlin Soft Matter Phys* 64: 051904, 2001 [Erratum in *Phys Rev E Stat Nonlin Soft Matter Phys* 65: 4, 2002]. doi:10.1103/PhysRevE.64.051904.
47. **Yarrow S, Challis E, Seriès P.** Fisher and Shannon information in finite neural populations. *Neural Comput* 24: 1740–1780, 2012. doi:10.1162/NECO_a_00292.
48. **Kanitscheider I, Coen-Cagli R, Pouget A.** Origin of information-limiting noise correlations. *Proc Natl Acad Sci USA* 112: E6973–E6982, 2015. doi:10.1073/pnas.1508738112.
49. **Kafashan M, Jaffe A, Chetthi SN, Nogueira R, Arandia-Romero I, Harvey CD, Moreno-Bote MR, Drugowitsch J.** Scaling of information in large neural populations reveals signatures of information-limiting correlations (Preprint). *bioRxiv* 2020.01.10.902171, 2020. doi:10.1101/2020.01.10.902171.
50. **Anderson TW, Olkin I, Underhill LG.** Generation of random orthogonal matrices. *SIAM J Sci Stat Comput* 8: 625–629, 1987. doi:10.1137/0908055.
51. **Holmes RB.** On random correlation matrices. *SIAM J Matrix Anal Appl* 12: 239–272, 1991. doi:10.1137/0612019.
52. **Joe H.** Generating random correlation matrices based on partial correlations. *J Multivar Anal* 97: 2177–2189, 2006. doi:10.1016/j.jmva.2005.05.010.
53. **Stewart GW.** The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM J Numer Anal* 17: 403–409, 1980. doi:10.1137/0717034.
54. **Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al.** Scikit-learn: machine learning in Python. *J Mach Learn Res* 12: 2825–2830, 2011.
55. **Bekker PA, ten Berge JM.** Generic global identification in factor analysis. *Linear Algebra Appl* 264: 255–263, 1997. doi:10.1016/S0024-3795(96)00363-1.
56. **Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al.** SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat Methods* 17: 261–272, 2020 [Erratum in *Nat Methods* 17: 352, 2020]. doi:10.1038/s41592-019-0686-2.
57. **Montijn JS, Liu RG, Aschner A, Kohn A, Latham PE, Pouget A.** Strong information-limiting correlations in early visual areas (Preprint). *bioRxiv* 842724, 2019. doi:10.1101/842724.
58. **Eden UT, Kramer MA.** Drawing inferences from Fano factor calculations. *J Neurosci Methods* 190: 149–152, 2010. doi:10.1016/j.jneumeth.2010.04.012.
59. **Softky WR, Koch C.** The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs. *J Neurosci* 13: 334–350, 1993. doi:10.1523/JNEUROSCI.13-01-00334.1993.
60. **Tolhurst DJ, Movshon JA, Dean AF.** The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res* 23: 775–785, 1983. doi:10.1016/0042-6989(83)90200-6.
61. **van Steveninck RR, Lewen GD, Strong SP, Koberle R, Bialek W.** Reproducibility and variability in neural spike trains. *Science* 275: 1805–1808, 1997. doi:10.1126/science.275.5307.1805.
62. **Rieke F, Warland D, Van Steveninck RR, Bialek W.** *Spikes: Exploring the Neural Code*. Cambridge, MA: MIT Press, 1999.
63. **Bouchard KE, Mesgarani N, Johnson K, Chang EF.** Functional organization of human sensorimotor cortex for speech articulation. *Nature* 495: 327–332, 2013 [Erratum in *Nature* 498: 526, 2013]. doi:10.1038/nature11911.
64. **Anumanchipalli GK, Chartier J, Chang EF.** Speech synthesis from neural decoding of spoken sentences. *Nature* 568: 493–498, 2019. doi:10.1038/s41586-019-1119-1.
65. **Doya K, Ishii S, Pouget A, Rao RP.** *Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge, MA: MIT Press, 2007.
66. **Valente M, Pica G, Bondanelli G, Moroni M, Runyan CA, Morcos AS, Harvey CD, Panzeri S.** Correlations enhance the behavioral readout of neural population activity in association cortex. *Nat Neurosci* 24: 975–986, 2021. doi:10.1038/s41593-021-00845-1.
67. **Huang X, Lisberger SG.** Noise correlations in cortical area mt and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements. *J Neurophysiol* 101: 3012–3030, 2009. doi:10.1152/jn.00010.2009.
68. **Nassar MR, Scott D, Bhandari A.** Noise correlations for faster and more robust learning. *J Neurosci* 41: 6740–6752, 2021. doi:10.1523/JNEUROSCI.3045-20.2021.
69. **Lenninger M, Skoglund M, Herman PA, Kumar A.** Are single-peaked tuning curves tuned for speed rather than accuracy? *eLife* 12: e84531, 2023. doi:10.7554/eLife.84531.
70. **Brown EN, Barbieri R, Eden UT, Frank LM.** Likelihood methods for neural spike train data analysis. In: *Computational Neuroscience: a Comprehensive Approach*. Boca Raton, FL: CRC, 2003, p. 253–286.
71. **Kass RE, Ventura V, Brown EN.** Statistical issues in the analysis of neuronal data. *J Neurophysiol* 94: 8–25, 2005. doi:10.1152/jn.00648.2004.
72. **Inouye D, Yang E, Allen G, Ravikumar P.** A review of multivariate distributions for count data derived from the Poisson distribution. *Wiley Interdiscip Rev Comput Stat* 9: e1398, 2017. doi:10.1002/wics.1398.
73. **Schneidman E, Berry MJ, Segev R, Bialek W.** Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature* 440: 1007–1012, 2006. doi:10.1038/nature04701.
74. **Sokoloski S, Aschner A, Coen-Cagli R.** Modelling the neural code in large populations of correlated neurons. *eLife* 10: e64615, 2021. doi:10.7554/eLife.64615.
75. **Goris RL, Movshon JA, Simoncelli EP.** Partitioning neuronal variability. *Nat Neurosci* 17: 858–865, 2014. doi:10.1038/nn.3711.
76. **Brinkman BA, Weber AI, Rieke F, Shea-Brown E.** How do efficient coding strategies depend on origins of noise in neural circuits? *PLoS Comput Biol* 12: e1005150, 2016. doi:10.1371/journal.pcbi.1005150.
77. **Moore E, Chaudhuri R.** Using noise to probe recurrent neural network structure and prune synapses. *Adv Neural Inf Process Syst* 33: 14046–14057, 2020.
78. **Adesnik H, Naka A.** Cracking the function of layers in the sensory cortex. *Neuron* 100: 1028–1043, 2018. doi:10.1016/j.neuron.2018.10.032.
79. **Cohen MR, Maunsell JH.** Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12: 1594–1600, 2009. doi:10.1038/nn.2439.
80. **Ruff DA, Cohen MR.** Attention can either increase or decrease spike count correlations in visual cortex. *Nat Neurosci* 17: 1591–1597, 2014. doi:10.1038/nn.3835.