Stimulus-driven population activity patterns in macaque primary visual cortex

Benjamin R. Cowley[a,*], Matthew A. Smith[b], Adam Kohn[c], Byron M. Yu[d,**]

[a]*Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA*
[b]*University of Pittsburgh, Pittsburgh, PA*
[c]*Albert Einstein College of Medicine, Bronx, NY*
[d]*Carnegie Mellon University, Pittsburgh, PA*

Abstract

Background:

Recent advances in neural recording technology have allowed experimenters to record from tens of neurons simultaneously, and many analysis techniques have been developed to leverage the co-variations between neurons. One promising techinique, dimensionality reduction, has been applied in various brain areas to study the activity of populations of neurons. However, no systematic study of the dimensionality (i.e., the number of dimensions of the reduced space) of population activity has been carried out. Thus, before assessing dimensionality in brain areas where less is known about the properties of single neurons, an important first step is verifying that the outputs of dimensionality reduction are interpretable in a brain area where much is known about single-neuron properties.

Objective:

Our objective is to develop a framework which provides a solid foundation for interpreting the outputs of dimensionality reduction applied to population activity. This framework should be applicable to population activity recorded from many brain areas across different experimental conditions.

Method:

To assess dimensionality, we applied principal component analysis (PCA) to trial-averaged neural activity in macaque primary visual cortex (V1) in response to different visual stimuli, including sinusoidal gratings, a natural movie, and white noise. Dimensionality reduction was applied for each stimulus, separately. Since we also sought to compare population activity basis patterns between stimuli, we developed a novel statistical method to compute the similarity of basis patterns across different stimuli and time. We also used this same framework on outputs of well-established V1 receptive field models.

Data:

We recorded spiking activity from 60-80 neurons of five monkeys (recorded in primary visual cortex, V1) in response to individual drifting sinusoidal gratings and movies of sinusoidal gratings differing in orientation, a natural scene, and white noise. We repeated

presentations of each stimulus for hundreds of trials, and trial-averaged the spike counts, taken in 20 ms bins. We also considered the visual stimuli themselves, converting images into vectors of pixel intensities. Lastly, our data include outputs from common receptive field models (with the same number of neurons and time points as that of the neural data) in response to the same visual stimuli presented to each monkey.

Results:

We found that the neural dimensionality increased as stimulus complexity increased, and that the population activity occupies similar dimensions to encode different stimuli. These results were consistent with known properties of single neurons and with the results of the receptive field models, suggesting that the outputs of dimensionality reduction reflect many different single-neuron properties in a unified manner. However, there were notable discrepancies between real data and model, and our dimensionality results indicated improvements that could be made to the models.

Conclusion:

We showed that our dimensionality reduction framework was interpretable and consistent with many single-neuron properties in a well-known area. Future work is to apply this framework to population activity recorded in brain areas where less is known about how populations of neurons encode sensory and task information.

Introduction

Dimensionality reduction has been applied to neural population activity to study decision making [1, 2], motor control [3, 4], olfaction [5], working memory [6, 7], visual attention [8], audition [9], rule learning [10], speech [11], and more [for review, see 12]. In many cases, dimensionality reduction is applied in brain areas for which the relationship between neural activity and external variables, such as the sensory stimulus or behavior, is not well characterized. This is indeed the setting in which dimensionality reduction may be most beneficial because it allows one to relate the activity of a neuron to the activity of other recorded neurons, without needing to assume a moment-by-moment relationship with external variables.

To aid in interpreting the outputs of dimensionality reduction in such settings, it is important to systematically vary the inputs to a brain area and ask whether the outputs of dimensionality reduction change in a systematic way. This is most readily done for a brain area close to the sensory periphery, such as the primary visual cortex (V1). Here, we apply dimensionality reduction to V1 and ask two fundamental, population-level questions. First, how is neural complexity related to task complexity? Previous studies using dimensionality reduction analyze population activity in

---

*In fulfillment of Machine Learning DAP requirement, submitted Dec. 16, 2015

**Advisor of Benjamin Cowley

Email address: bcowley@cs.cmu.edu (Benjamin R. Cowley)

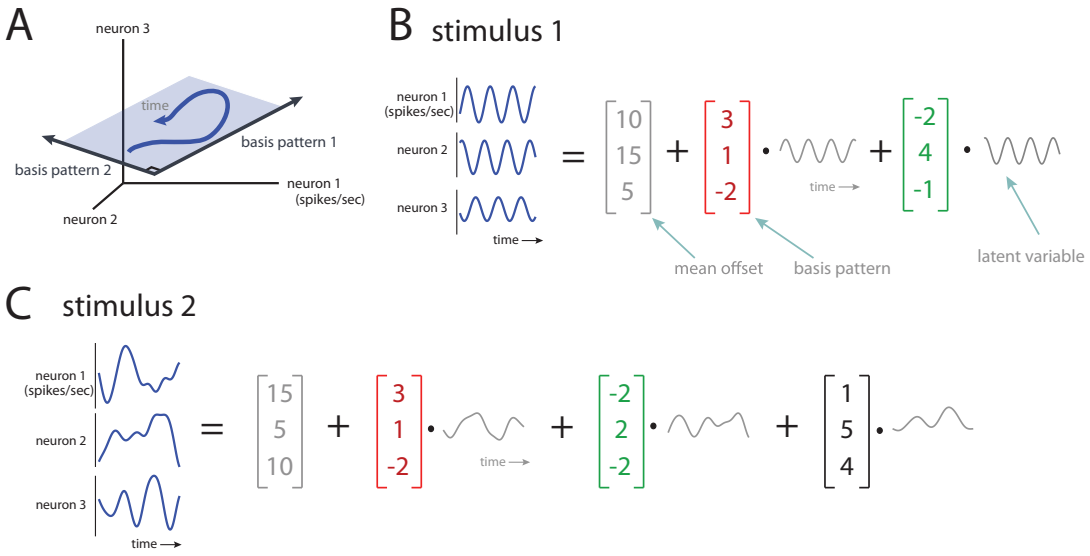

Figure 1: Conceptual illustration of dimensionality and basis patterns. *A*: The activity of three neurons can be plotted in a 3-d population firing rate space, where each axis represents the firing rate of one neuron. The population activity evolves over time (blue trace), and occupies a 2-d plane (blue shade). The basis patterns are axes of this 2-d plane. *B*: The activity of three neurons can also be represented as time-varying firing rates or peri-stimulus time histograms, PSTHs. The activity can be decomposed into a weighted sum of basis patterns (red and green) and a mean offset (gray). Each basis pattern is weighted by a time-varying latent variable. Note that basis patterns are mutually orthogonal, by definition. *C*: The activity of the same three neurons as in *A*, but for a different stimulus. Same conventions as in *A*.

a reduced space [e.g., 5, 3, 4]. Implicit in these studies is the appropriate dimensionality of the reduced space, which is a measure of neural complexity. It is currently unknown how neural complexity scales with task complexity for a given population of neurons [13]. Second, how does a neural circuit flexibly encode (or "multiplex") the representation of the vast number of stimuli encountered in the natural world? Recent studies suggest that it may be possible to take advantage of the multi-dimensional properties of the population activity space [14, 6, 2, 15, 16]. In particular, the population activity representing different stimuli might occupy similar dimensions of the population activity space [9, 4]. It is currently unknown how the similarity of the dimensions being occupied by the population activity changes with the similarity of the stimuli.

The concept of dimensionality is illustrated in Fig. 1. Consider a high-dimensional space (termed the *population firing rate space*) in which each axis represents the firing rate of a recorded neuron (Fig. 1*A*). The goal of dimensionality reduction is to identify i) how many dimensions are occupied by the neural population activity (Fig. 1*A*, blue), i.e., the *dimensionality* of the population activity, and ii) how these dimensions are oriented within the population firing rate space (Fig. 1*A*, basis patterns). In this three-neuron example, the population activity is two-dimensional, where the dimensions are defined by the orthogonal basis patterns. Equivalently, we can think of dimensionality reduction in terms of decomposing the population activity as a weighted sum of basis patterns and a mean offset (Fig. 1*B*). A basis pattern describes a characteristic way in which activity of the neurons covaries. Each basis pattern is fixed and is weighted by a time-varying latent variable, which represents the contribution of the basis pattern at each point in time.

We can compare the outputs of dimensionality reduction for two different stimuli (Fig. 1*B* and 1*C*). The population activity for stimulus 1 is two-dimensional because it can be described by two basis patterns (Fig. 1*B*), whereas that for stimulus 2 is three-dimensional (Fig. 1*C*). Thus, the population response to stimulus 2 would be deemed more complex than the population response to stimulus 1. In addition, we can ask whether the population responses to different stimuli occupy similar dimensions within the population firing rate space. This is assessed by comparing the basis patterns across stimuli. In this example, there is one basis pattern that is shared by both stimuli (red), one basis pattern that is similar between stimuli (green), and one basis pattern (black) that is employed only by stimulus 2.

In this work, we systematically characterize how neural complexity varies with stimulus complexity in macaque V1 by applying principal components analysis (PCA) to the trial-averaged neural responses to different classes of visual stimuli, including sinusoidal gratings, a natural movie, and white noise. In addition, we develop a new method (termed the pattern aggregation method) to measure how the basis patterns extracted from the population responses to each stimulus relate to each other. This provides a systematic characterization of how the similarity of the dimensions being occupied by the population activity changes with the similarity of the stimuli. A key advantage of studying these questions in V1 is that there are well-established receptive field (RF) models. By applying the same dimensionality reduction methods to the outputs of RF models, we can more deeply understand the relationship between the outputs of dimensionality reduction and known properties of V1 neurons. The results described in this work provide a solid foundation upon which to interpret the outputs of dimensionality reduction applied to other brain areas.

Materials and Methods

Neural recordings

Details of the neural recordings have been described previously [17, 18]. We recorded from the primary visual cortex (V1) of anesthetized, paralyzed macaque male monkeys. V1 has been studied for decades, and single-neuron properties are well-known ([19]). Thus, it is an appriopriate area from which to record for our study, as we desire to interpret the dimensionality results. Anesthesia was administered throughout the experiment with a continuous intravenous infusion of sufentanil citrate (6-18 μg/kg/hr). This anelgesic drug targets higher cortical areas, but does not directly affect the synaptic receptors of neurons in V1. Previous studies have confirmed that trial-averaged neural activity under anesthesia strongly resembles that of an awake monkey [e.g., 20]. Thus, we believe the dimensionality results found here also characterize data recorded from an awake monkey. Eye movements were minimized with a continuous intravenous infusion of vecuronium bromide (100-150 μg/kg/hr). Experiments typically lasted 5-7 days. All experimental procedures followed guidelines approved by the Institutional Animal Care and Use Committee of the Albert Einstein College of Medicine at Yeshiva University, and were in full compliance with the guidelines set forth in the US Public Health Service Guide for the Care and Use of Laboratory Animals.

Neural activity was recorded using 96-channel multi-electrode arrays (Blackrock Microsystems, Salt Lake City, Utah), which covered 12.96 mm^2 (about an eighth the size of a penny) and

had an electrode length of 1 mm. After the monkey was anesthetized, we removed a small portion of skull, and implanted the array through the dura (the thick layer of membrane that protects the brain). The electrodes were inserted to a nominal depth of 0.6 mm to confine recordings mostly to layers 2-3. Recordings were performed in parafoveal V1, with RFs within 5 degrees of the fovea. While implantation is invasive and may cause damage to brain tissue, this damage is typically minimal for short-term recordings (less than a month). There is more concern for long-term recordings (greater than a month) in which the build-up of scar tissue and jostling of the array can deteriorate neural recordings. However, our experiment was short-term (7 days), anesthesia prevented any jostling of the array, and our recorded activity strongly resembles that of previous studies. Thus, we are confident that our results reflect those of an awake, un-implanted monkey.

Acquisition of spike trains

Our goal was to extract the times of spiking events of the neurons from the recorded voltage traces. From cellular studies, the "spike" of the neuron (in which the voltage of the neuron rapidly depolarizes, then has a refactory period of hyperpolarization before returning to a baseline level) is the most salient property of a neuron, and allows the neuron to communicate with other neurons through the release of neurotransmitters. Thus, by analyzing the spikes of neurons, we discard the small fluctuations of the voltage traces as a layer of abstraction. This abstraction is rooted from the findings of many previous studies, which have correlated the timing of spikes with behavior [e.g., **?**], encoding of stimulus information [e.g., 21], and even learning [4].

To obtain the spiking events, we set a voltage threshold separately for each electrode channel, based on the magnitudes of the depolarizing spikes. When the voltage trace surpassed the threshold, we saved a 1 ms segment of the voltage waveform which ensured that we captured the voltage dynamics of the spike (which occur within 1 ms). Because one electrode can record from multiple neurons (where each neuron has a unique spike waveform segment), we used semi-automated spike-sorting software [18] that applies PCA to the spike waveforms and then performs clustering. The sorts of the waveforms are then manually verified and re-sorted to discard classified waveforms that appear to be noise artifacts rather than neural in nature. This process yields single "unit" activity (i.e., spikes that we believe are from a single neuron) and multi-unit activity (i.e., spikes that we believe are from mulitple neurons that could not be differentiated). For all analyses, we used both single units and small clusters of units (see [22] for comparison of these signals). We included as a unit any waveform with a signal-to-noise ratio (SNR) greater than 1.5, where SNR is defined as the ratio of the average waveform amplitude to the standard deviation of the waveform noise [23]. This SNR level was chosen a priori of viewing the data to obtain as many units as possible to analyze. We also removed neurons with mean firing rates less than 1 spike per second to prevent ill-conditioned matrices when performing dimensionality reduction, where we chose this threshold a priori.

Visual stimuli

We used two sets of visual stimuli. The first set (termed the individual gratings set) consisted of individual presentations of drifting sinusoidal gratings with different orientations. The second

set (termed the movies set) consisted of different classes of visual stimuli, including a sequence of drifting sinusoidal gratings with different orientations, a contiguous sequence of natural scenes, and white noise. All stimuli were presented on a CRT monitor with a frame rate of 100 or 120 Hz, and had mean luminance of approximately 40 cd/m^2. We used a look-up table for all stimuli to correct for the nonlinearity between input voltage and output luminance in the monitor.

Individual gratings. Full-contrast (100%) drifting sinusoidal gratings with 12 equally-spaced orientations (30° between adjacent orientations, covering 360°) were presented. We use "orientation" to refer to the angle of drift and "direction" to refer to two orientations that are 180° apart (i.e., opposing drift) [24]. Spatial frequency (1.3 cpd) and temporal frequency (6.25 Hz) were chosen to evoke robust responses from the population as a whole, and the position and size (8-10 degrees) were sufficient to cover the RFs of all recorded neurons. Gratings were block-randomized across the 12 orientations and presented for 1.28 seconds each, followed by a 1.5 second intertrial interval consisting of an isoluminant gray screen. We conducted 200 trials for each of the 12 orientations.

Movies. We presented three different 30-second grayscale movies: gratings, natural, and noise, as described previously [18]. Each movie comprised 750 unique images (each presented for four consecutive video refreshes of the CRT, for an effective framerate of 25 Hz). The movie frames were surrounded by a gray field of average luminance.

The gratings movie was a pseudo-randomly chosen sequence of full-contrast drifting sinusoidal gratings with 98 equally-spaced orientations (3.67° between adjacent orientations, covering 360°). The presentation of each drifting grating of a particular orientation lasted 300 ms. Spatial frequency (1.3 cpd), temporal frequency (6.25 Hz), position and size (8 degrees or 640 pixel diameter circular aperature) were chosen to evoke high firing rates and to sufficiently cover RFs of all recorded neurons. Two 300 ms periods of blank screen frames were included at a randomly-chosen position in the sequence, bringing the total to 100 stimuli in a block lasting 30 seconds. A 30-second movie with the same random sequence of gratings and blank screens was repeated 120 times.

The natural movie was a 30 second consumer film of a contiguous sequence of natural scenes converted to grayscale (a monkey wading through water). The movie was displayed in a square of 5 degrees (400×400 pixels) to cover all RFs, and repeated 120 times.

The noise movie was 30 seconds of white noise, where each 4 degree (320×320 pixel) frame comprised a 40×40 grid of 8-pixel squares. Each 8×8-pixel square displayed a random intensity drawn from a uniform distribution independently (in space and time) of other squares. The entire 40×40 grid was randomly shifted or "jittered" k pixels ($1 \leq k \leq 8$) horizontally and vertically between consecutive frames to avoid high-contrast grid effects. One noise movie was randomly generated, and the same movie was repeated 120 times.

Preprocessing of population activity and visual stimuli

Population activity for individual gratings. We presented the individual gratings stimulus set to three monkeys (101r, 102l, 103r). After removing neurons that did not satisfy the SNR and firing rate criteria, the lowest number of neurons across monkeys was 61. Because measurements of

dimensionality can be influenced by the size of the recorded population, we selected a random subset of 61 neurons from the data sets with larger populations, in order to combine results across monkeys (dimensionality trends were similar across monkeys for all analyses). We considered neural activity in a 1 second window starting at stimulus onset (discarding the remaining 0.28 seconds of response). Results were similar for a 1 second window starting 100 ms after stimulus onset to avoid onset transient effects. For each neuron and orientation, we took spike counts in 20 ms bins and averaged them across trials to create a peri-stimulus time histogram (PSTH), yielding 50 time points. Because a step in PCA subtracts from each PSTH the mean across the 50 time points, we can compare (across different orientations) fluctuations of trial-averaged activity around its mean.

Population activity for movies. We presented the movies stimulus set to two monkeys (monkey 1: 103l, 61 neurons, monkey 2: 102l, 81 neurons). For each neuron, we took spike counts in 20 ms bins during the 30-second movie presentation and averaged them across trials to create a PSTH. For each movie, this yielded a PSTH with 1,500 time points for each neuron. PCA then centers each PSTH by subtracting its mean, enabling the study of fluctuations of the firing rate around the mean. We ensured the neural activity was centered just once for all analyses (i.e., for non-overlapping one-second time windows, the "local" mean was not subtracted again). In this way, all dimensionality measurements are made using a common reference frame, thereby allowing comparison of dimensionality across different time windows (analysis of Fig. 6*B*).

Visual stimuli for movies. To relate neural complexity to stimulus complexity, we performed PCA on the movie stimuli. This required the following pre-processing steps in order to match the sizes of images across stimuli and remove incidental spatial correlations in the noise movie. The processed movies were only used for PCA analysis; the original, unprocessed movies were shown to subjects. We first cropped the grating and natural image frames to 320 x 320 pixels to match the size of the noise images. The noise movie had two incidental spatial correlations from the experimental design that prevented it from being "true" white noise: spatial correlations caused by noise images being comprised squares of 8×8 pixels with identical intensities and spatial correlations caused by jittering the squares to avoid grid effects. To weaken these correlations, we converted the images of all movies to 8×8 pixel squares. The average pixel intensity was computed for each pixel square, and each image was compressed from a matrix of 320×320 values to a matrix of 40×40 values. Because we averaged over fixed pixel squares, we did not eliminate all incidental spatial correlations in the noise movie, as the borders between the 8×8 pixels with identical intensities were jittered randomly from frame-to-frame. Thus, the red eigenspectrum curve in Fig. 4*B* is not perfectly diagonal, which is expected from true white noise. Each 40×40 pixel image was then reshaped into a vector (with size $1,600 \times 1$). Each movie stimulus was therefore represented as 750 vectors (one for each image) in a 1,600-dimensional space, where each axis corresponds to the average intensity of one pixel block.

Assessing dimensionality and similarity of patterns

Assessing dimensionality. We describe here how we assessed the dimensionality of the population activity and that of the visual stimuli. Conceptually, the dimensionality is the number of basis

patterns needed to describe the population activity (Fig. 1) or the pixel intensities of the visual stimuli. There are many ways to assess dimensionality, including linear [25] and nonlinear [26, 27] methods. Here, we consider the most basic linear dimensionality reduction method, principal component analysis (PCA). PCA is appropriate for use with trial-averaged neural activity because trial averaging removes much of the Poisson-like spiking variability, consistent with PCA's implicit assumption of no observation noise [12]. PCA has also been applied to pixel intensities in previous studies [28, 16].

Dimensionality was assessed throughout this study as the number of basis patterns needed to explain 90% of the variance in the population activity or visual stimuli. The threshold of 90% is arbitrary, and we verified that the results were qualitatively similar for other high-percentage thresholds (e.g., 70% and 80%). It is possible to use a data-driven approach to determine the variance threshold [16]. Under this approach, data which require a small or large number of dimensions to explain the majority of the variance can be deemed low-dimensional. To aid in interpreting dimensionality comparisons, we chose to use a pre-determined (90%) variance threshold in this work. It is important to note that the dimensionality depends on many factors, including the dimensionality reduction method (here, PCA), metric for assessing dimensionality (here, 90% variance explained), the number of neurons, and the number of data points used. Thus, we do not attempt to interpret the dimensionality in an absolute sense. Rather, we make relative comparisons of dimensionality across experimental conditions, while keeping these other factors fixed. We also considered fixing the mean firing rate by removing neurons that showed large differences in mean firing rates across stimuli. Presumably, this would allow us to measure how much dimensionality changes for a fixed number of spikes. However, it is unclear how the remaining neurons are representative of the entire recorded population. For example, when assessing dimensionality across sinusoidal gratings with different orientations, this procedure would remove neurons whose firing rates are modulated by orientation. We would most likely be left with neurons who have low firing rates for all orientations.

Assessing similarity of patterns. In addition to computing the number of basis patterns needed for each condition, we sought to compare the patterns across conditions. The following is the intuition for performing this comparison. Consider two sets of patterns: condition A requires k_A patterns and condition B requires k_B patterns, where $k_A > k_B$. At one extreme, the space spanned by the condition A patterns includes the space spanned by the condition B patterns (i.e., the spaces overlap). In this case, the joint space is described by $k_{AB} = \max(k_A, k_B) = k_A$ patterns. At the other extreme, the space spanned by the condition A patterns is orthogonal to the space spanned by the condition B patterns. In this case, the joint space is described by $k_{AB} = k_A + k_B$ patterns. In general, k_{AB} will lie between $\max(k_A, k_B)$ and $k_A + k_B$. The closer k_{AB} is to $\max(k_A, k_B)$, the more similar the patterns are. Conversely, the closer k_{AB} is to $k_A + k_B$, the more dissimilar the patterns are.

To compute k_{AB}, we tried several different approaches. In the first approach, we aggregated the population activity (or visual stimuli) for different conditions, then applied the same PCA method to find the number of dimensions that explained 90% of the variance. A problem with this approach is that it does not account for possibly different variances of the population activity (or visual stimuli) across conditions. The consequence is that k_{AB} obtained by this method can be less

than $\max(k_A, k_B)$, thereby violating the intuition laid out above. To see this, consider a scenario where $k_A > k_B$ and condition B has much larger variance than condition A. The aggregated population activity (or visual stimuli) would be dominated by condition B, essentially ignoring the patterns of condition A. As a result, the aggregated dimensionality k_{AB} would be close to k_B, which is less than $\max(k_A, k_B) = k_A$. This motivated us to consider a second approach in which we normalized the population activity (or visual stimuli) such that the direction of greatest variance was 1 for each condition. However, there are still scenarios for which k_{AB} obtained by this method is less than $\max(k_A, k_B)$, due to how the variance is distributed across patterns.

Pattern aggregation method. To overcome the issues described above, we developed an alternative approach which guaranteed that k_{AB} would be between $\max(k_A, k_B)$ and $k_A + k_B$. This method, termed the pattern aggregation method, is based on first identifying the k_A patterns for condition A (represented by $U_A : N \times k_A$, whose columns are orthonormal and where N is the number of pixels or the number of neurons) and k_B patterns for condition B (represented by $U_B : N \times k_B$, whose columns are orthonormal) separately using PCA. Then, we aggregate the patterns into a single matrix $V = [U_A U_B]$ (where $V : N \times (k_A + k_B)$) and designate k_{AB} to be the effective column rank (defined below) of V. In the case where $k_A > k_B$ and the space spanned by the condition A patterns includes the space spanned by the condition B patterns, then k_{AB} will be $\max(k_A, k_B) = k_A$. On the other hand, if the space spanned by the condition A patterns is orthogonal to the space spanned by the condition B patterns, then k_{AB} will be $k_A + k_B$. For more than two conditions, this method is easily extended by aggregating the patterns for all conditions into $V = [U_A U_B U_C \ldots]$.

The effective column rank of V is the number of "large" singular values of V. The definition of "large" requires setting a threshold (termed the rank threshold, between 0 and 1) for the singular values. The rank threshold determines how different two basis patterns (cf. Fig. 1, green) need to be before they define separate dimensions. To gain intuition for the rank threshold, we ran a simulation where we rotated a 2-d unit vector (v_B) from a position of overlap with another unit vector (v_A, 0°) to a position of orthogonality (90°) (Fig. 2*A*). We varied the rank threshold for different angles between v_A and v_B, and assessed the effective column rank of V. A rank threshold $t = 0$ means that slight deviations of v_B away from v_A lead to $V = [v_A v_B]$ having an effective column rank of two (Fig. 2*B*). In other words, if two patterns are slightly different, the dimensionality of the data will be two. On the other hand, a rank threshold $t = 1$ means that v_B needs to be orthogonal to v_A for V to have an effective column rank of two (Fig. 2*B*). In other words, two patterns need to be orthogonal for the dimensionality of the data to be two; else, the dimensionality will be one. As a compromise between the two extremes, we used a rank threshold $t = 0.5$ throughout this study. This threshold means that the transition from one to two dimensions occurs when the angle between v_A and v_B is near 45 degrees (Fig. 2*B*). We also verified that this intuition holds for higher-dimensional spaces by rotating a subspace towards an orthogonal subspace by taking a convex combination of the two subspaces and measuring the rank at intermediate rotations.

Computing dimensionality expected by chance. To assess whether two sets of patterns are similar or dissimilar, it is necessary to compare k_{AB} to a chance level, rather than simply asking whether

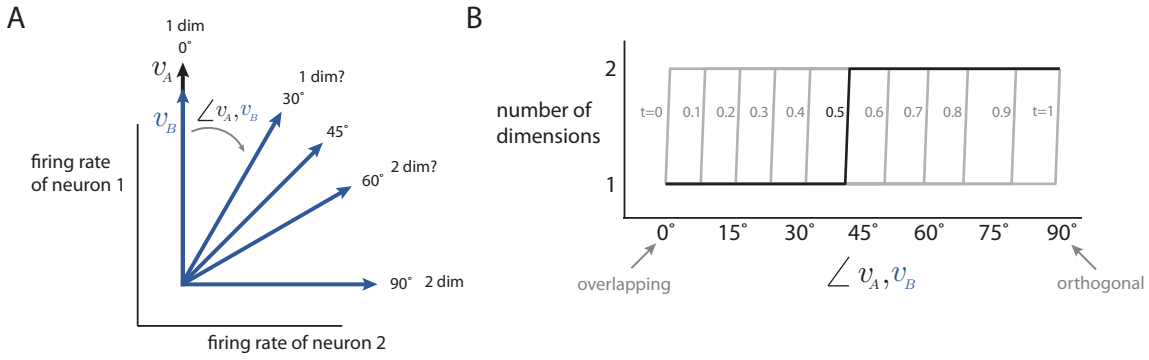

Figure 2: Assessing similarity of basis patterns. *A*: Conceptual illustration for two neurons, where v_A denotes the basis pattern for one condition and v_B denotes the basis pattern for another condition. As v_B is rotated away from v_A, the question is at which point do we consider v_A and v_B to span two dimensions. *B*: The transition from one to two dimensions in *A* depends on the rank threshold t. For $t = 0.5$, the transition occurs when the angle between v_A and v_B is near 45 degrees.

k_{AB} is closer to $\max(k_A, k_B)$ or $k_A + k_B$. We compute the chance level by drawing patterns randomly from the N-dimensional space, and compute a distribution of aggregated dimensionalities. Specifically, we repeatedly draw k_A orthonormal patterns and k_B orthonormal patterns in an N-dimensional space, and measure their aggregated dimensionality using the same method as applied to the data.

The above method makes the implicit assumption that all patterns within the N-dimensional space can be shown in the data. In some cases, there is reason to believe that not all patterns in the N-dimensional space can be achieved. For example, the underlying neural circuitry may constrain the space of activity patterns that a population of neurons is capable of producing [9, 4]. In such settings, the chance level should be computed by drawing patterns randomly from an M-dimensional space, where $M < N$. Although the exact value of M that should be used for a particular neural population is unknown, M should lie between the largest dimensionality exhibited empirically by the neural population for the stimuli shown and N. Thus, we can assess whether the results hold for all values of M within this range.

Similarity index. The similarity of two sets of patterns depends on whether k_{AB} is larger or smaller than the chance level, relative to the two extremes ($\max(k_A, k_B)$ and $k_A + k_B$). We summarized this dependence with a single number, the similarity index. The similarity index is defined as $s = \frac{\hat{k}_{AB} - k_{AB}}{k_A + k_B - \max(k_A, k_B)}$. Intuitively, we take the difference between the mean chance dimensionality ($\hat{k}_{AB}$) and the actual dimensionality for the aggregated patterns of A and B (k_{AB}). Then, we normalize this difference by the difference between the two extreme dimensionalities, yielding an index between -1 and 1. A similarity index of $s > 0$ means that the patterns are more similar (i.e., more overlapping) than expected by chance, whereas $s < 0$ means that the patterns are less similar (i.e., closer to orthogonal) than expected by chance.

Statistical assessment of dimensionality

Error bars for the dimensionality of the population activity were computed by subsampling from all time points. We chose subsampling over bootstrapping, since bootstrapping led to biased estimates due to small sample size relative to the number of neurons. For population responses to the stimulus movies, we randomly subsampled 750 of the 1,500 time points, computed the dimensionality of the subsampled points, and repeated this 100 times (Fig. 5*C*). Similarly, for the dimensionality analysis in 1 second windows, we randomly sampled 25 of the 50 time points 100 times (Fig. 6*A*). We did not compute error bars for the visual stimuli because subsampling was not possible (only 750 available time points), and bootstrapping was not possible due to the small number of time points relative to the dimensionality (1,600) of the pixel space. All p-values were computed from 10^5 runs of a random permutation test.

Receptive field models

We considered three mathematical operations — linear filtering, pointwise nonlinearity, and divisive normalization — that are common to many RF models of V1 neurons [19]. These particular models were chosen for their simplicity in interpretation and the build-up of non-linearities. We applied these operations to the movies stimulus set and examined how the dimensionality of the outputs of each operation compares to that of the stimuli for non-overlapping one-second time windows. Note that our goal was not to find the RF model that best reproduces the recorded neural responses. Rather, our goal was to determine how mathematical operations common to many RF models alters dimensionality, in a way that informs the relationship between the dimensionality of the visual stimuli and that of the recorded population activity. The parameters of the operations were chosen to be representative of the properties of the recorded neurons. For each operation, we computed model responses based on RF properties measured for multi-unit activity for each monkey. The results were similar for these different RF ensembles, so we present only the results using RF properties of monkey 1.

Linear filtering. To explore the consequence of applying V1-like linear filters, we implemented a bank of Gabor filters, where each Gabor filter is defined by a 2-d Gaussian function multiplied by a 2-d sine wave [29, 30]. For pixel location (x, y), the Gabor filter value is:

$$F(x_\theta, y_\theta) = \exp\big(-\frac{x_\theta^2}{2\sigma} - \frac{y_\theta^2}{2\sigma/\gamma}\big)\sin(\frac{2\pi}{\lambda}x_\theta + \phi) \tag{1}$$

where $x_\theta = x\cos(\theta) + y\sin(\theta) + x_{\text{loc}}$ and $y_\theta = x\sin(\theta) + y\cos(\theta) + y_{\text{loc}}$. The parameters of the Gabor filters were chosen to be representative of the recorded neurons. We did this by using the measured parameters of the RFs for multi-unit activity of each electrode by presenting sinusoidal gratings (for 250 ms to 1 sec) that varied in spatial location $(x_{\text{loc}}, y_{\text{loc}})$ (15×15 locations spaced 0.5 degrees apart for monkey 1; 21×21 locations for monkey 2), spatial frequency $1/\lambda$ (11 different values from 0.098 cpd to 3.125 cpd), and orientation θ ($\{0°, 45°, 90°, 135°\}$). For each electrode, the RF parameters $(x_{\text{loc}}, y_{\text{loc}}, 1/\lambda, \theta)$ were chosen based on maximal observed counts of threshold crossings. We used RF spatial sizes (σ) of 0.5-1.0 degrees, consistent with previously reported RF sizes [31]. The spatial aspect ratio ($\gamma = 0.9$) and phase (ϕ =0°) were fixed.

We created a Gabor filter whose parameters corresponded with those of each electrode. Then, to match the number of neurons in the actual recordings (61 neurons for monkey 1), we randomly selected 61 Gabor filters without replacement for analysis. We then computed the output of each Gabor filter (F) for a given input image (I): $L = \sum_x \sum_y F(x,y)I(x,y)$.

Pointwise nonlinearity. The output of the Gabor filters may be negative, unlike neural responses, which can only take on non-negative values. Thus, RF models typically pass the output of each Gabor filter through a pointwise nonlinear function, such as a squaring function, squared hinge function, or a linear hinge function [19]. For the purposes of assessing dimensionality, we chose to focus on the squaring function ($R = L^2$) because it appears in the divisive normalization model described below. We also tried the other nonlinearities listed above, which gave similar results.

Divisive normalizaton. The third operation that we implemented was divisive normalization, where the activity of each neuron was normalized by the activity of the neighboring neurons. Divisive normalization is often included in RF models due to its ability to explain basic nonlinear response properties of V1 neurons [32, 33]. Here, we used the divisive normalization model proposed by [34]. The normalized response R of a filter is

$$R = \frac{L^2}{\sigma_L^2 + \sum_{i=1}^{M} w_i L_i^2} \tag{2}$$

where L is the output of a Gabor-like filter, L_i is the output of the ith neighboring Gabor-like filter ($i = 1, \ldots, M$), w_i is the weight of the ith neighboring filter, and σ_L^2 is an offset term.

Details about the model and fitting procedures can be found in [34]. Briefly, we used the same filter bank (i.e., the steerable pyramid) as in [34], whose filters are qualitatively similar to the Gabor-like RFs of V1 neurons [34]. The filter bank comprised 104 neighboring filters (25 spatial locations with ~1/8 cycles/pixel spatial frequency, spaced 3 pixels apart vertically, horizontally, and diagonally, each with 4 different orientations; and 4 different orientations at the center spatial location with ~1/16 cycles/pixel spatial frequency). We trained the weights on 1,000 images of a natural image dataset not used in the neurophysiological recordings [35]. We verified that our fitting procedure was correct by accurately recovering weights from data generated by a model where we know the ground truth weights. Further, the values of the fitted weights were reasonable: for example, a selected filter that was vertically-oriented had weights with large magnitudes for other vertically-oriented filters that were directly above or below the selected filter's location. We then computed the output of 61 filters, where each filter's RF parameters (spatial location and orientation) were matched with that of a randomly-selected electrode from the experiment.

Results

Dimensionality of population responses to gratings

We first asked how changing the complexity of the visual stimulus changes the dimensionality of trial-averaged population responses using the individual gratings. To change the stimulus complexity, we included different numbers of consecutive grating orientations in the analysis (ranging

from 1 to 12 orientations). For example, the least complex stimulus included a single orientation, and more complex stimuli included two or five consecutive orientations (Fig. 3*A*).

We asked how quickly the dimensionality of the population activity grows as the number of orientations increases. At one extreme, it may be that the population response to each orientation uses an entirely different set of basis patterns (i.e., dimensions). In this case, the dimensionality for two orientations would be two times the dimensionality for one orientation. At the other extreme, it may be that the population response to each orientation resides in the same set of dimensions. In other words, the population response is formed using the same basis patterns, but linearly combined using different weights for different orientations. In this case, the dimensionality for two orientations would be the same as the dimensionality for one orientation.

We first computed the basis patterns for the trial-averaged population response (taken in 20 ms bins) of each orientation individually. Then, to assess dimensionality of the population response to multiple orientations, we aggregated the basis patterns for each constituent orientation and computed the effective rank (termed the pattern aggregation method, see Methods). Using a 90% variance threshold, we found that the dimensionality for two orientations was 1.62 times the dimensionality for one orientation (Fig. 3*B*, '90%' curve), and significantly smaller than what would be expected had the basis patterns been randomly chosen (Fig. 3*B*, 'chance 90%' curve, $p < 10^{-5}$). In other words, for consecutive orientations separated by 30°, the population responses share about half of their basis patterns. As more orientations were included, the dimensionality of population responses remained lower than expected by chance (Fig. 3*B*), indicating that population responses to oriented gratings separated by angles larger than 30° also tend to use similar basis patterns. Similar trends were found using different variance thresholds (Fig. 3*B*, '70%', '80%' curves), so we use a 90% threshold in the rest of this work.

The pattern aggregation method requires a choice for the rank threshold t to determine how different basis patterns need to be before they define separate dimensions. We repeated the above analysis for different choices of t and a fixed variance threshold of 90% (Fig. 3*C*). We found that the dimensionality trends are similar for rank thresholds t near 0.5, so we use $t = 0.5$ in the rest of this work. Because the absolute dimensionality depends on the variance and rank thresholds, we make no claims about absolute dimensionality. Rather, we focus on relative comparisons of dimensionality for fixed variance and rank thresholds.

We observed a change in the rate of increase of dimensionality after six orientations (Fig. 3*B*, black curve). Because consecutive orientations were separated by 30°, the first and seventh orientations were 180° apart and differed only in their drift direction. Thus, the seventh to twelfth orientations were identical to the first to sixth orientations respectively, but drifted in opposite directions. Only a small proportion of V1 neurons are known to be direction selective [36, 24], so the change in slope of the dimensionality curve might be due to the population activity using similar basis patterns for opposite drift directions.

To test this possibility, we performed two analyses. First, we assessed the direction selectivity of each neuron (direction index = 1 - null response / preferred response), and found that only 16 of 183 neurons had a direction index greater than 0.5, consistent with [24]. If none of the neurons were direction selective, we would expect the dimensionality curve to be flat beyond 6 orientations in Fig. 3*B*. The increase in dimensionality after 6 orientations is consistent with the finding that at least some neurons show direction selectivity. A potential concern is that the increase in

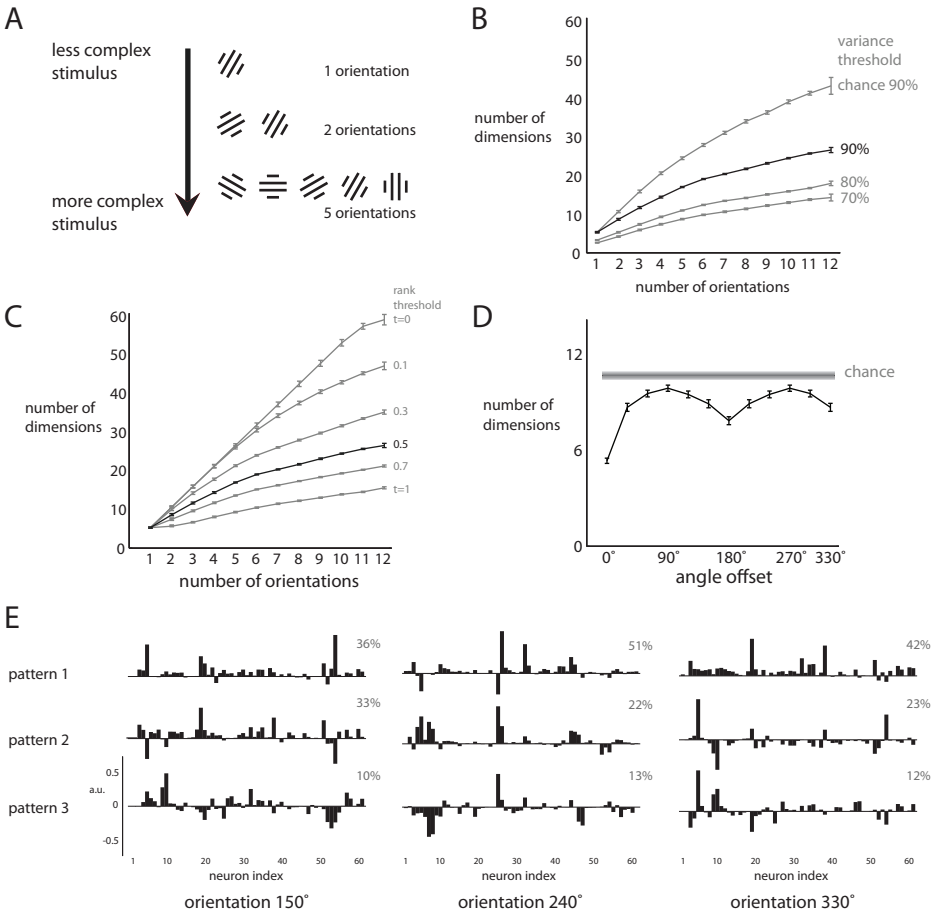

Figure 3: Dimensionality of population responses to individual gratings. *A*: The complexity of the stimulus was varied by combining a different number of consecutive orientations. The least complex stimulus consisted of a single orientation, and more complex stimuli included two or five orientations. *B*: Dimensionality of population activity versus number of orientations. Bottom three curves correspond to the number of dimensions needed to explain 70%, 80%, and 90% of the variance. Top gray curve corresponds to the number of dimensions expected by chance for the 90% variance threshold. Error bars represent the standard error across monkeys and all possible combinations of consecutive orientations. *C*: Varying the rank threshold for a fixed variance threshold (90%). Same data as in *B*. Each curve represents the dimensionality of the population response as the number of consecutive orientations varies, for a particular rank threshold t. Error bars are computed as in *B*. *D*: Dimensionality of population activity versus angle offset between two orientations (bottom black curve). Chance dimensionality (top gray curve) and error bars are computed as in *B*. *E*: Basis patterns describing the largest percentage of variance for the population responses to three example orientations (90° apart) for one monkey. Each pattern is a unit vector with a norm of 1. Percentages denote percent variance explained by each pattern.

dimensionality beyond 6 orientations arises from the fact that a larger number of patterns are being aggregated for a larger number of orientations. To address this, we performed a control analysis in which the patterns were randomly drawn from a large pool, and therefore equalized the number of patterns being aggregated across different numbers of orientations. We found the same trend as shown in Fig. 3*B* (results not shown).

Second, we assessed how the dimensionality of the population activity for two orientations varies with the angular offset between the orientations (Fig. 3*D*). This indicates how the similarity of the dimensions being occupied by the population activity changes with the similarity of the stimuli. We found that the dimensionality increases with angular offset up to 90°, indicating that the population activity differs the most for two orientations with 90°offset. Then, the dimensionality decreases as the angular offset increases, reaching a minimum at a 180° offset, where gratings drift in opposite directions. Thus, the population activity uses more similar basis patterns for gratings drifting in opposite directions (180° offset) than to gratings of different orientations (angular offsets other than 180°). The dimensionality for 180° offset was higher than that for 0° offset ($p < 10^{-5}$), indicating that the population activity does not use identical basis patterns for opposite drift directions. This remained true when we controlled for the number of patterns being aggregated. This result, combined with the change in slope of the dimensionality curve (Fig. 3*B*), indicates that the population activity tends to use similar (but not identical) basis patterns for opposite drift directions. Taken together, the analyses in Fig. 3*B* and 3*D* provide a systematic characterization of how the outputs of dimensionality reduction vary with the sensory input to V1.

We also visualized the basis patterns describing the largest percentage of variance for three different orientations (Fig. 3*E*). These basis patterns extracted from the trial-averaged population activity are akin to the hypothetical basis patterns shown in Figure 1 (red, green, black). For a given basis pattern, the absolute height of each bar indicates the degree to which each neuron contributes to that basis pattern. The following are two salient properties of the identified basis patterns. First, most of the neurons in the recorded population contribute to each basis pattern to a greater or lesser extent. Thus, the basis patterns capture changes in firing rates that are shared broadly across the population, rather than reflecting the activity of only a small number of neurons. Second, the basis patterns capture both positive and negative signal correlations between neurons. A basis pattern describes positive signal correlation between a pair of neurons if the neurons have coefficients of the same sign. Conversely, a basis pattern describes negative signal correlation for coefficients of opposite sign. We can relate these basis patterns to the results shown in Fig. 3*B-D* by asking how similar are the linear combinations of each set of basis patterns across different stimulus orientations. This is difficult to assess by eye, and so we rely on the quantifications shown in Fig. 3*B-D* to determine how similar are the basis patterns across stimulus orientations.

Dimensionality of population responses to different classes of visual stimuli

We next sought to determine how the dimensionality of the trial-averaged population activity varies with the complexity of the visual stimulus for a wider range of stimuli. We presented three movie stimuli (Fig. 4*A*): a sequence of sinusoidal gratings (‘gratings movie’), contiguous natural scenes (‘natural movie’), and white noise (‘noise movie’). In contrast to Fig. 3*A* where the order

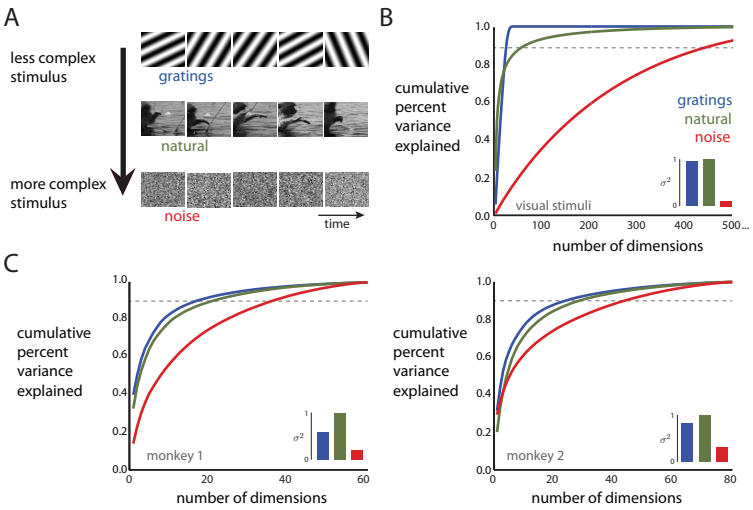

Figure 4: Dimensionality of movie stimuli and population responses to those movies. *A*: Example frames of the movies for the gratings movie, natural movie, and noise movie. *B*: For each stimulus, cumulative percent variance of the stimulus pixel intensities explained by different numbers of dimensions. Dashed line corresponds to 90% of the variance explained. Inset: summed variance of the pixel intensities, normalized by the maximum variance across movies. *C*: For each stimulus, cumulative percent variance of the population activity explained by different numbers of dimensions. Dashed lines correspond to 90% of variance explained. Left panel: monkey 1, 61 neurons. Right panel: monkey 2, 81 neurons. Insets: summed variance of each neuron's activity, normalized by the maximum variance across movies.

of stimulus complexity is clear (i.e., a larger number of orientations is more complex), here we needed first to assess the relative complexity of the three movie stimuli. By applying PCA to the pixel intensities, we found that 40 dimensions could explain nearly 100% of the variance of the pixel intensities for the gratings movie (Fig. 4*B*). For the natural movie, the top few dimensions explained a large percentage of the variance due to overall contrast changes caused by zooming and panning the camera, and a large number of additional dimensions were needed to explain the remaining variance. For the noise movie, each dimension explained only a small percentage of the total variance. We summarized these cumulative percent variance curves by finding the number of dimensions (gratings movie: 24, natural movie: 64, noise movie: 459) needed to explain 90% of the variance (Fig. 4*B*, dashed line). Based on these values, the gratings movie was least complex, followed by the natural movie, then the noise movie. Similar results were obtained when first transforming the pixel intensities using V1 receptive field models, then applying PCA (see "Comparing to V1 receptive field models"). We further asked how much the pixel intensities varied for each movie stimulus, and found that the noise movie had a smaller variance than the other two movie stimuli (Fig. 4*B*, inset). Together, this indicates that the distribution of pixel intensities for the gratings movie and natural movie is akin to an elongated ellipsoid (low dimensionality, high variance), whereas that for the noise movie is akin to a small ball (high dimensionality, low variance).

Having measured the relative complexity of the movie stimuli, we then asked how the dimensionality of the population responses to the movie stimuli varies with stimulus complexity. We found that the dimensionality of the trial-averaged population responses (20 ms bins) was or-

dered in the same way as the stimulus complexity (Fig. 4*C*); namely, the population responses to the gratings movie had the lowest dimensionality, followed by the natural movie, then the noise movie. This ordering did not simply follow from the ordering of the mean population firing rates for the different movies (monkey 1: 4.2, 6.4, 5.4 spikes/sec, monkey 2: 6.6, 8.2, 6.7 spikes/sec for gratings, natural, and noise movies, respectively). We also assessed how much the firing rates varied in response to each movie stimulus. Note that this measures how much the mean firing rate (averaged across experimental trials) varies over time. As with pixel intensities, we found that the population response to the noise movie had the smallest variance, followed by the gratings movie, then the natural movie (Fig. 4*C*, inset). Overall, the dimensionality and variance ordering in the visual stimuli and the population responses were the same, indicating that the population activity in V1 retains much of the statistical structure present in the visual stimuli themselves.

Basis patterns of population responses

Having compared the dimensionality of the population activity across stimuli, we next asked how the basis patterns of the population activity (corresponding to the dimensions being occupied by the population activity) compare across stimuli. Previous studies have found that the ability of a RF model to predict a neuron's response can depend on the stimulus class on which the model was trained [37, 38, 39], suggesting that the population activity might use somewhat distinct basis patterns for different stimulus classes. On the other hand, if basis patterns are influenced by the shared underlying network structure [9, 4], then we would expect them to be shared across responses to different stimuli.

We first asked whether there are qualitative differences in the coefficients of the basis patterns for population responses to the stimulus movies (Fig. 5*A*). As in Fig. 3*E*, we found that most of the basis patterns represented activity across a large number of neurons and described both positive and negative signal correlations. However, there were two notable exceptions. First, the basis pattern describing the most variance for the population response to the gratings and noise movies involved primarily two neurons (neuron indices 26 and 27). For these movies, the two neurons had the highest firing rate modulation (maximum - minimum firing rate) across the recorded population. Second, the basis pattern describing the most variance for the population response to the natural movie had coefficients of the same sign. This is due to the entire population increasing or decreasing its firing rates together in response to overall contrast changes prevalent in natural movies. Other than the similarity of the top basis pattern for the gratings and noise movies, it was difficult to determine by eye whether the basis patterns were being shared across stimuli. Thus, we used the pattern aggregation method to quantitatively assess the similarity of the identified basis patterns.

As a baseline, we assessed the extent to which the visual stimuli themselves reside in the same dimensions in pixel space. We compared the dimensionalities of the individual movies (Fig. 5*B*, teal dots, consistent with Fig. 4*B*) to those of combinations of movies (Fig. 5*B*, orange and purple dots). If the stimuli reside in overlapping dimensions, then the resulting dimensionality would be the maximum of the dimensionalities for the individual movies. However, if the stimuli reside in completely non-overlapping (i.e., orthogonal) dimensions, then the resulting dimensionality would be the sum of the dimensionalities for the individual movies. There are two main observations.

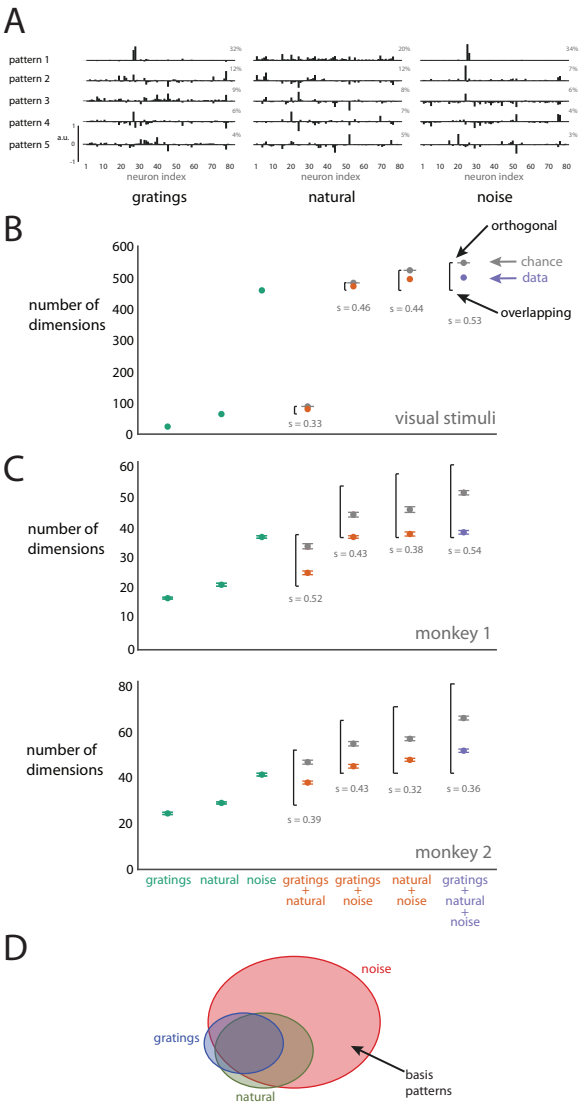

Figure 5: Similarity of basis patterns across stimuli. *A*: Basis patterns describing the largest percentage of variance for the population responses to the gratings, natural, and noise movies for monkey 2. Each pattern is a unit vector with a norm of 1. Percentages denote percent variance explained by each pattern. *B*: Dimensionality of visual stimuli for individual movies (teal dots), and combinations of two (orange dots) or three (purple dots) movies. The teal dots correspond to where the curves intersect the dashed lines in Fig. 4*B*. Black brackets denote the range of possible dimensionalities (bottom of the black bracket corresponds to overlapping patterns; top of the black bracket corresponds to orthogonal patterns). Gray dots indicate dimensionalities expected by chance, and error bars represent the standard deviation of 100 random samples. The similarity index s indicates if patterns overlap more than expected by chance ($s > 0$) or are closer to orthogonal than expected by chance ($s < 0$). *C*: Dimensionality of population activity, for individual and combinations of movies. Same conventions as in *B*. Error bars represent standard deviations of the subsampled estimates. *D*: Illustrative Venn diagram of basis patterns. The size of each ellipse indicates the number of basis patterns, and the overlap indicates the extent to which basis patterns are shared by different stimuli.

First, the dimensions occupied by the gratings movie overlap with those for the natural movie. To see this, note that the dimensionality for the gratings and natural movies together (80 dimensions) was less than the sum of the individual dimensionalities for the two movies (88 dimensions). To ensure that the overlap in patterns was meaningful, we confirmed that the aggregated dimensionality of 80 was less than the dimensionality (88 dimensions) that would be expected by combining randomly-chosen dimensions (Fig. 4*B*, gray, $p < 10^{-5}$). Second, the dimensions occupied by the noise movie include many of the dimensions for the other two stimuli. This is indicated by the fact that the dimensionality corresponding to any combination of stimuli that included the noise movie (gratings + noise: 472, natural + noise: 495, and gratings + natural + noise: 500 dimensions) was less than the dimensionality that would be expected by chance ($p < 10^{-5}$ for all cases). Note that, in all cases, the chance dimensionality was near the maximum dimensionality (indicating orthogonality between the two sets of randomly-chosen patterns) because the dimensionality of the pixel space (1,600 dimensions) was much larger than the dimensionalities of the individual movies.

We used the same approach to analyze the population responses as we did the visual stimuli. We compared the dimensionality of the population responses to individual movies (Fig. 5*C*, teal dots, consistent with Fig. 4*C*) to that of population responses to combinations of movies (Fig. 5*C*, orange and purple dots). We found that the relationship of the basis patterns employed by the population activity across stimuli (Fig. 4*C*) was similar to the relationship between the stimuli themselves (Fig. 4*B*). First, the dimensions occupied by the population responses to the gratings movie are overlapping with those for the natural movie. This is indicated by the fact that the aggregated dimensionality for the population responses to the gratings and natural movies (monkey 1: 25, monkey 2: 38) was less than the dimensionality if the patterns were orthogonal (top of the black brackets) and the dimensionality expected by chance ($p < 10^{-5}$). Second, the dimensions occupied by the population response to the noise movie include most of the dimensions for the other two stimuli. This is indicated by the fact that the dimensionality corresponding to any combination of stimuli that included the noise movie was less than the dimensionality expected by chance ($p < 10^{-5}$ in all cases), and close to entirely overlapping.

The chance dimensionality in Fig. 4*C* is computed by assuming that any population activity pattern within the N-dimensional population firing rate space can be achieved (where N=61 for monkey 1 and N=81 for monkey 2). Previous studies indicate that the population activity may only be able to occupy a subset of dimensions due to underlying network constraints [9, 4]. Although it is unknown exactly how many dimensions can be occupied by the population activity, it can occupy at least M dimensions (where $M = 39$ for monkey 1 and $M = 52$ for monkey 2), which is the largest dimensionality observed in response to any combination of stimuli (Fig. 4*C*). If the chance level is computed instead by drawing random patterns from this M-dimensional space, we still find that the population activity tends to occupy more similar dimensions than expected by chance for all pairs of movies ($p < 0.05$). This is a conservative assessment because larger values of M would only make the comparison more highly significant.

These results are summarized by the schematic in Fig. 5*D*, where the size of each ellipse represents the dimensionality (i.e., number of basis patterns) of the population response to the corresponding stimulus and the overlap between ellipses represents the extent to which the population responses share basis patterns. We found that the basis patterns for the gratings movie were overlapping with those of the natural movie, and that the basis patterns for the noise movie largely

contain the basis patterns for the other two stimuli. However, there were a small number of basis patterns that were unique to each stimulus, shown by the small areas of non-overlap among the ellipses. Overall, this suggests that a neural circuit is capable of expressing a limited repertoire of basis patterns, and that a subset of those basis patterns is expressed for any given stimulus. Furthermore, the statistical properties of the visual stimuli are largely maintained in their neural population representation in V1.

Time-resolved dimensionality of population responses to movie stimuli

In the preceding sections, the analyses of the movie stimuli and the corresponding neural activity used the entire 30-second movie (i.e., 750 time points) together. Here, we consider time-resolved measurements of dimensionality using one second windows, each comprising 25 time points. This allows us to assess how dimensionality changes over time, and compare the basis patterns employed by the trial-averaged population activity during different parts of the 30-second movies.

For the visual stimuli (Fig. 6*A*, left panel) and the population responses (Fig. 6*A*, center and right panels), we found that the dimensionality corresponding to the noise movie was higher than the dimensionality corresponding to the gratings and natural movies in each one-second window, consistent with the results of analyzing each 30-second movie in its entirety (Fig. 4). However, the dimensionality of the natural movie was not greater than that of the gratings movie (cf. colored triangles in Fig. 6*A*, which indicate average dimensionality over time), in contrast to Fig. 4. We hypothesized that this was due to temporal correlations in the natural movie, in which frames within a one-second window tend to be self-similar. In contrast, for the gratings and noise movies, there are at least three different grating orientations and 25 different frames of white noise within each one-second window.

To reconcile the results for short and long time windows, we performed three analyses. First, we tested the hypothesis that temporal correlations in the natural movie result in lower dimensionalities (relative to the other movies) for short time windows. To break the temporal correlations, we shuffled the time points (20 ms resolution) across each 30 second period and performed the same analysis as in Fig. 6*A*. We found that, in a one-second window, the mean dimensionality corresponding to the natural movie was higher than that corresponding to the gratings movie for the shuffled data. This was true for the visual stimuli ($p < 0.01$) and for the population responses (monkey 1: $p < 10^{-3}$, monkey 2: $p < 0.05$). This indicates that the range of basis patterns expressed by the population responses to the natural movie within a short time window is limited by the temporal correlations in the natural movie itself.

Second, we asked how the dimensionality grows when increasing the window size progressively from one second to 30 seconds, where each window starts at the beginning of the movie (Fig. 6*B*). If the dimensionality increases with window size, this would indicate that new patterns (of pixels or of population activity) are being used throughout the length of the movie. However, if the dimensionality plateaus, then the patterns are being reused and no new patterns are being expressed. The leftmost points on these curves (one-second window) correspond to the leftmost points on the corresponding curves in Fig. 6*A*. The rightmost points on these curves (30-second window) correspond to the dimensionalities shown in Fig. 4. Although the dimensionality cor-

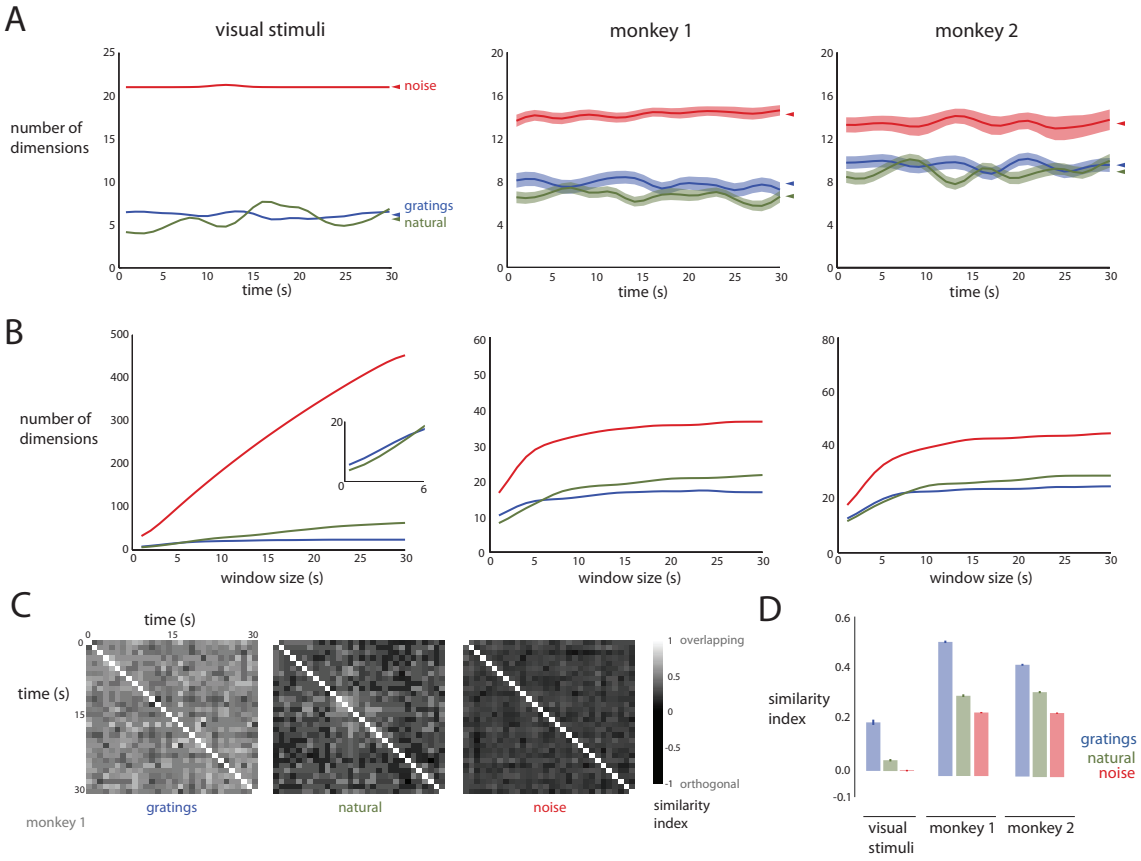

Figure 6: Time-resolved dimensionality of movie stimuli and population responses. *A*: Dimensionality versus time. Left panel: visual stimuli. Center panel: monkey 1 (61 neurons). Right panel: monkey 2 (81 neurons). Error bars are standard deviations of subsampled estimates. Triangles denote mean dimensionality across time for each movie. Curves were smoothed with a Gaussian kernel with a width of 1.5 s. *B*: Dimensionality with growing time windows starting at the beginning of each movie. Left panel: visual stimuli. Inset: zoomed portion of the bottom-left plot. Center panel: monkey 1. Right panel: monkey 2. Curves were smoothed as in *A*. Due to the smoothing, the dimensionalities for the 1 second window do not exactly match the leftmost point in *A*, and those for the 30 second window do not exactly match the dimensionalities in Fig. 4. *C*: Similarity of the basis patterns employed by the population responses across time (monkey 1). Each element in the similarity matrix corresponds to the similarity index between the sets of patterns for the pair of time points. *D*: Mean similarity index for visual stimuli and population responses. Error bars denote standard error over similarity indices.

responding to the natural movie (green) is lower than that corresponding to the gratings movie (blue) for short time windows, the dimensionality corresponding to the natural movie grows more quickly and surpasses that corresponding to the gratings movie for longer time windows. We found this to be true for both the visual stimuli (Fig. 6*B*, left panel) and the population responses (Fig. 6*B*, center and right panels). This indicates that the patterns for the natural movie tend to be self-similar within a short time window, and new patterns are expressed over longer time windows. In contrast, the dimensionality corresponding to the gratings movie does not grow as quickly with window size. This is because once a few different grating orientations are included, patterns corresponding to additional grating orientations can be represented approximately as linear combinations of patterns corresponding to existing grating orientations (both for the visual stimuli and the population responses), and therefore do not increase the dimensionality appreciably.

Third, we directly measured the similarity of the patterns between different one-second windows using the pattern aggregation method. We computed a similarity index s, for which $s > 0$ indicates that the patterns are more similar (i.e., more overlapping) than expected by chance and $s < 0$ indicates that the patterns are less similar (i.e., closer to orthogonal) than expected by chance (see Methods). Fig. 6*C* shows the matrix of similarity values for the population responses to each of the three movies (monkey 1). By averaging the values of the off-diagonal similarity indices across the matrix, we found that the patterns corresponding to the gratings movie tended to be more similar across time than those for the natural and noise movies (Fig. 6*D*). This was true for the visual stimuli (left panel), as well as for the population responses (middle and right panels). This result indicates that new patterns tend to be expressed (for both the visual stimuli and the population responses) as the movies play out over time more for the natural and noise movies than for the gratings movie. This result also supports the finding in Fig. 6*B* that the dimensionality corresponding to the natural movie grows more quickly than that corresponding to the gratings movie.

Taken together, these results indicate that the noise movie drives a large number of basis patterns in the population activity in a short time window, relative to the gratings and natural movies. As the movies play out over time, the noise and natural movies tend to drive new basis patterns, whereas the grating movie tends to recruit the same patterns.

We also asked whether the second-by-second fluctuations of the dimensionality of the visual stimulus during the 30-second movie (Fig. 6*A*, left panel) were related with that of the population responses (Fig. 6*A*, middle and right panels) and found weak correlations (mean across movies $\rho = 0.20$). Although the dimensionality fluctuations were larger for the natural movie than the other two movies (Fig. 6*A*, left panel), the same effect was less salient in the population activity (Fig. 6*A*, middle and right panels). These two observations underscore that, although there are similarities in the statistical properties of the visual stimuli and population responses, there are also differences that remain to be understood.

Comparing to V1 receptive field models

One of the key advantages of performing this study in V1 is that much is known about the stimulus-response relationship of individual neurons, as described by the many RF models that have been proposed in the literature [19]. Although the RF models do not capture every aspect of

V1 neuronal activity [40, 37, 38, 39], we can apply the same dimensionality reduction methods to the activity generated by RF models to help interpret the outputs of dimensionality reduction. We consider it a strength that, in many cases (described below), the outputs of dimensionality reduction applied to real data show the same trends as when applied to activity from RF models. It shows how single-neuron properties are reflected in population metrics, such as dimensionality and similarity of basis patterns, which was previously not known. In cases where there are discrepancies between the outputs of dimensionality reduction for real data and RF models, our results provide insight into the nature of the discrepancies and can help to improve the RF models.

Although a complete study of the many V1 RF models available is beyond the scope of this work, here we focus on three mathematical operations that are commonly included in RF models – linear filtering, pointwise nonlinearity, and divisive normalization. Specifically, we consider the following three RF models that are increasingly complex (Fig. 7*A*): i) a bank of Gabor filters only, ii) a bank of Gabor filters, each with a squaring nonlinearity, and iii) a bank of Gabor-like filters, each with a squaring nonlinearity, and then divisive normalization [34]. Below are results for which the RF model parameters were matched to the properties of the recorded neurons for monkey 1. Similar results were obtained for RF model parameters matched to monkey 2 (not shown).

For the individual gratings, we observed similar trends between the real data (Fig. 3*B* and *D*) and RF models (Fig. 7*B* and 7*C*), with two notable discrepancies. Whereas the dimensionality continues to grow after 6 orientations in the real data (Fig. 3*B*), the dimensionality plateaus after 6 orientations for the linear and nonlinear Gabor models, and only slightly increases for divisive normalization (Fig. 7*B*). Furthermore, the dimensionality for two orientations with 180° offset is higher than that for 0° offset in real data (Fig. 3*D*), but they are not different for the models (Fig. 7*C*). Both of these discrepancies can be explained by the fact that none of the models we implemented capture direction selectivity, which may require extending the models to include suppressive filters [41].

For the movies, the dimensionalities for the models (Fig. 8*A*) were lower than the dimensionalities for the real data (Fig. 4*C*, left panel). This can be seen by noting the number of dimensions at which each curve crosses the 90% variance threshold (dotted line). The linear and nonlinear Gabor models (Fig. 8*A*, left and middle panels) showed the same dimensionality ordering as the real data (Fig. 4*C*, left panel), but the divisive normalization model (Fig. 8*A*, right panel) did not. The discrepancy for the divisive normalization model is likely related to the objective of this particular model [34], which was to make the responses independent across neurons, and not necessarily representative of all divisive normalization models [33]. None of the models (Fig. 8*A*, insets) reproduced the variance trends seen in the real data (Fig. 4*C*, left panel, inset). However, all models had high similarity indices (Fig. 8*B*), consistent with the population responses occupying overlapping dimensions for the real data (Fig. 5*C*, top panel). The similarity of basis patterns for the model repsonses is to be expected, as neurons are modelled with rigid filters that can capture only a finite number of dimensions in pixel space. The rigidity in filters overwhelms the ability of the nonlinearities to express different basis patterns, although we did see that the basis patterns were less similar for the divisive normalization model (Fig. 8*B*, right panel) than those for the linear and nonlinear Gabor models (Fig. 8*B*, left and middle panels).

For the time-resolved analysis of movies, only the nonlinear Gabor model (Fig. 8*C* and *D*,

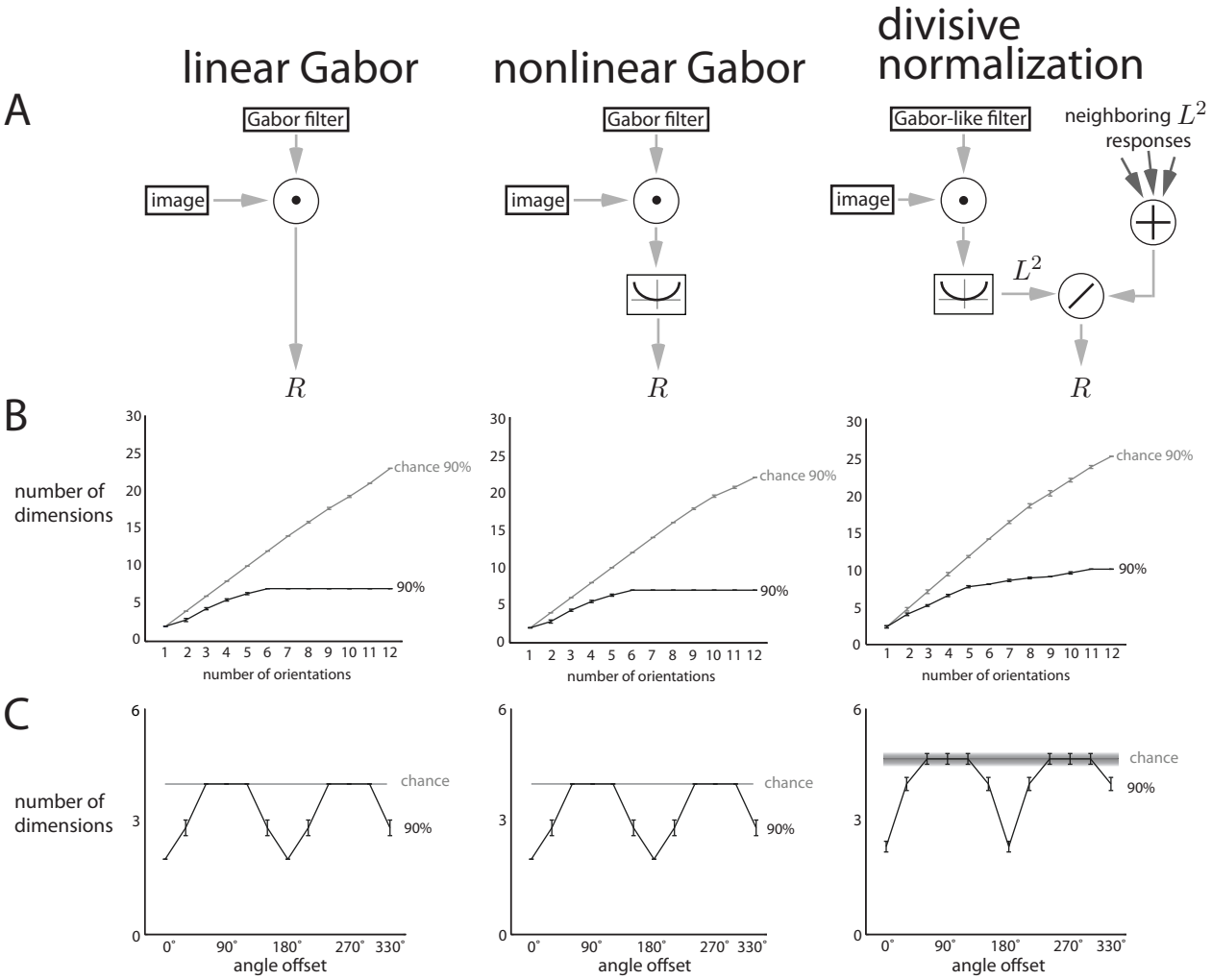

Figure 7: Dimensionality of model responses to individual gratings. *A*: Flow diagram of the three different models, which transform an input image to model responses, R. We then assess the dimensionality of the model responses. *B*-*C*: Dimesionality results for individual gratings, computed in the same manner as in Fig. 3*B* and *D*, respectively. Results are based on 61 filters whose RF parameters are matched to those of the 61 neurons for monkey 1.

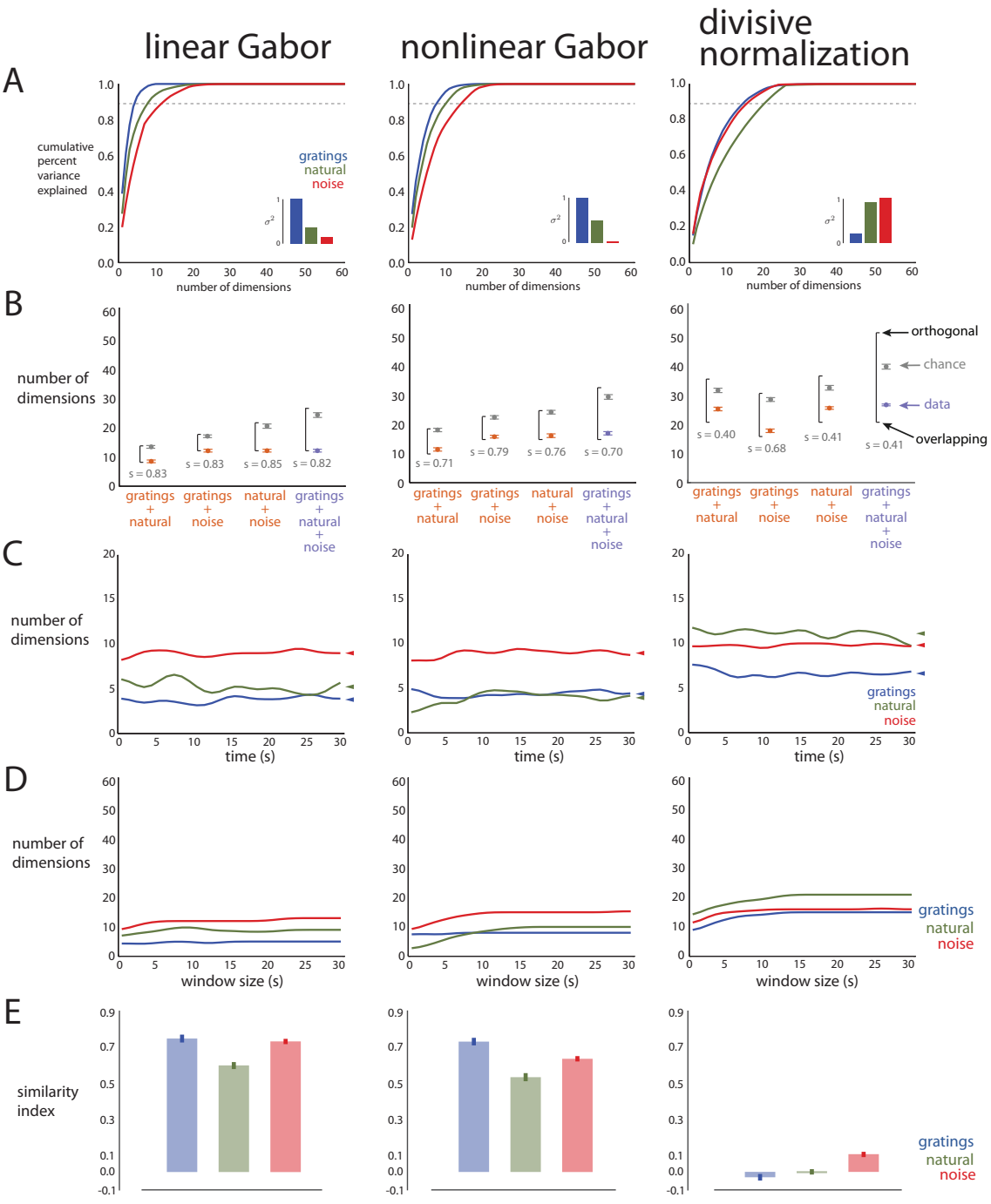

Figure 8: Dimensionality of model responses to movie stimuli. *A-E*: Dimensionality results for movies, computed in the same manner as in Fig. 4*C*, Fig. 5*C*, and Fig. 6*A-B* and *C*, respectively. Results are based on 61 filters whose RF parameters are matched to those of the 61 neurons for monkey 1.

middle panels) showed the same dimensionality trends as the real data (Fig. 6*A* and *B*, middle panels). However, the nonlinear Gabor model showed these trends for a different reason than the real data. Unlike for the real data, shuffling the nonlinear Gabor responses across time kept the same dimensionality ordering (natural ¡ gratings, $p < 0.01$) as without shuffling. This indicates that the dimensionality trends seen in Fig. 8*C* and *D* (middle panels) are not due to temporal correlations in the natural movie, but rather due to properties of the nonlinearity used. Compared to the linear and nonlinear Gabor models, the divisive normalization model showed higher overall dimensionalities (Fig. 8*C* and *D*, right panels), which better matched that of the real data (Fig. 6*A* and *B*, middle panels). Finally, the linear and nonlinear Gabor models used more similar basis patterns across time than the real data, as the mean similarity indices were larger for the models (Fig. 8*E*, left and middle panels) than for the real data (Fig. 6*D*, middle panel). However, the divisive normalization model typically used new basis patterns across time, as its mean similarity indices were close to zero (Fig. 8*E*, right panel). None of the models matched the ordering of the mean similarity indices for the visual stimuli and the population activity.

These results show how altering the properties of single neurons via RF models changes the outputs of dimensionality reduction. This informs the interpretation of the outputs of dimensionality reduction in other brain areas, many of which do not have well-established models of single-neuron function. In addition, these results show how RF models compare to V1 recordings from a novel population-level perspective. We found that the RF models reproduce some, but not all, of the population-level properties of the V1 recordings. Although we chose reasonable model parameters, we did not finely measure the RF parameters of the recorded neurons, which could be a source of the discrepancies between the models and real data. Furthermore, there are many model variants that we did not implement that could resolve these discrepancies [e.g., 38, 41, 42, 43, 39]. Overall, the results presented here should only be interpreted in the context of the particular models we implemented, and not necessarily representative of all models within each class.

Discussion

In this study, we systematically varied the inputs to V1 and asked whether the outputs of dimensionality change in a systematic way. By applying PCA to trial-averaged population responses to different classes of visual stimuli, including sinusoidal gratings, a natural movie, and white noise, we found that the dimensionality of the population responses grows with the stimulus complexity. In addition, we assessed whether the population responses to different stimuli occupy similar dimensions of the population firing rate space using a novel statistical method (the pattern aggregation method). We found that the population responses to stimuli as different as gratings and natural movies tended to occupy similar dimensions. For comparison, we applied the same analyses to the outputs of well-established V1 receptive field models to understand how single-neuron properties are reflected in population metrics.

A recent study showed that the dimensionality of the population activity in prefrontal cortex collapses during error trials [6], suggesting that a smaller number of distinct variables are encoded by the neural population during error trials than on correct trials. While this previous work indicates a link between dimensionality and behavior (correct vs. incorrect responses), our work shows that complexity in V1 – an area tightly linked to the sensory periphery – can be traced to the com-

plexity of stimulus drive. Thus, our findings demonstrate that dimensionality is a well-behaved attribute of neural population activity, and lay the foundation for measuring dimensionality in higher-level brain areas that are not closely coupled to sensory input or motor output.

Different basis patterns may be used by the population activity during different task epochs, suggesting that certain basis patterns drive downstream areas more effectively than others [15, 44]. Thus, the identification of which basis patterns are used may be critical for understanding how different brain areas interact on a population level [45]. Furthermore, the activity patterns across a neural population have been used to characterize the effects of normalization [46]. In the present work, we have developed a statistical framework (the pattern aggregation method) to measure the similarity of basis patterns across any number of stimulus conditions or time points, and validated this framework using recordings in V1 and well-established V1 receptive field models. This framework can be applied broadly to other brain areas.

Many previous studies have compared visual cortical responses to natural and artificial stimuli on a single-neuron level [e.g., 37, 38, 42, 43, 39]. Their predominant approach was to define a parameterized RF model to relate a neuron's activity to the visual stimulus. These studies found that, although RF models derived from natural stimuli share properties with those derived from artificial stimuli [37, 42], there can be important differences [37, 38, 39]. Here, rather than relating each neuron's activity to the stimulus, we relate the activity of the recorded neurons to each other. We can then ask how this relationship (i.e., the covariation of trial-averaged activity among neurons) changes for different classes of visual stimuli. This approach has been adopted for pairs of neurons (i.e., signal correlation) [47], and we extend this work to characterize the signal correlation among all pairs of neurons at once. We found that the gratings, natural, and noise stimuli elicit many common basis patterns, consistent with previous studies showing similarities between RFs measured with gratings and natural stimuli [37] and those measured with natural and noise stimuli [42]. Finally, our finding of some unique basis patterns for each stimulus class suggests that estimates of RFs will best capture the responses to the same type of stimulus used in estimating the RFs, as reported in previous studies [38, 39].

At a population level, several studies have compared visual cortical activity evoked by natural and artificial stimuli to spontaneous activity [48, 49, 50]. These studies focused on the raw neural activity, which includes both the trial-averaged component (i.e., the PSTHs) and trial-to-trial variability. Here, we focused on the trial-averaged component. Because trial-to-trial variability can be substantial relative to the trial-averaged component [51, 48], it is difficult to directly compare results of these previous studies to those reported here. Future work investigating the complexity of the raw neural activity, using methods related to those proposed here, can provide a more direct comparison.

Although we believe that the results shown here are representative of a wide range of gratings, natural, or noise stimuli, they should be interpreted in the context of the particular visual stimuli used. For example, the dimensionality of the gratings movie and its population response could be increased by including more than one spatial and temporal frequency. Similarly, the dimensionality of the natural movie and its population response likely depends on the particular movie clip shown. If the scenes in the movie change more quickly (or slowly) over time, then we would expect the dimensionality over a 30-second time window to be larger (or smaller). For the noise movie, our results in Fig. 6*B* indicate that showing more instances of white noise is not likely to

further increase the dimensionality of the population response. However, changing the statistics of the noise in the pixels could change the dimensionality of the population response.

The measurement of dimensionality depends on many factors, including the choice of dimensionality reduction method, the metric for assessing dimensionality, the number of neurons, and the number of data points, as previously mentioned. Two other factors could also influence dimensionality. First, dimensionality may be influenced by the locations of the RFs of the recorded neurons. If two neurons have nearby RFs relative to the spatial frequency content of the stimulus, then the neurons would be driven by similar stimulus patches. This would likely lead to a lower dimensionality of the population response than if the two neurons had RFs that were far apart, where the neurons would be driven by different stimulus patches. Second, dimensionality may be influenced by the size of the visual stimulus. We presented large visual stimuli that extended outside of the classical RF of most neurons. Previous studies have shown that stimulation outside of the classical RF tends to increase the sparseness of V1 responses [52, 53, 54], which may affect the dimensionality of the population response. Sparseness leads to independence in the responses between neurons [55, 52, 53], as well as increased discriminability of the population activity [56]. Independence implies that each basis pattern only captures modulations of a single neuron (i.e., only one element of each pattern is non-zero), and the dimensionality (as assessed by PCA) depends on the relative variances captured by the basis patterns (in this case, the relative variances of the neurons). To our knowledge, there is no general relationship between sparseness and dimensionality. For all these reasons, it is not possible to make absolute statements about the dimensionality of V1. Instead, we made relative comparisons where all of the factors affecting dimensionality are fixed, except for the stimulus content.

The quality of most RF models has been evaluated based on their ability to predict the activity of individual neurons [e.g., 37, 38, 42, 19, 39]. Given that there can be, in some cases, a substantial difference between the predictions of RF models and the recorded neural activity [38, 42, 39], especially for natural scenes, we need to quantify how they are different in an effort to improve the RF models. This is often quantified by computing the percent variance of the recorded activity explained by the model for each neuron individually [38, 19, 39]. Our work provides a complementary way to compare RF models and recorded activity by examining the entire population together. We can compare the many V1 models that have been proposed by assessing which ones best reproduce the relative dimensionalities across stimuli and similarity of basis patterns observed in recorded activity.

Our work lays a solid foundation of to assess the dimensionality and similarily of basis patterns of neural population activity. Because V1 is a well-studied brain area and is close to the sensory input, our results can be compared with expectations based on our intuition and well-established RF models. Moving forward, these methods can be applied broadly to other brain areas and behavioral tasks to examine how the complexity of the population response changes due to conditions such as attentional state, learning, and contextual modulation.

Acknowledgements: We are grateful to Douglas Ruff and Ruben Coen-Cagli for helpful advice and discussion.

Funding: This work was supported by a National Science Foundation Graduate Research Fellowship and a National Defense Science and Engineering Graduate (32 CFR 168a) Fellowship (B.R.C.), NIH EY022928 and P30EY008098, Research to Prevent Blindness, and Eye and Ear Foundation of Pittsburgh (M.A.S.), NIH EY016774, Research to Prevent Blindness, and Irma T. Hirschl Career Scientist Award (A.K.), a Carnegie Mellon University ProSEED / Brainhub seed grant (M.A.S. and B.M.Y.), and the Simons Collaboration on the Global Brain (B.M.Y. and A.K.).

Disclosures: The authors declare no competing financial interests.

References

[1] C. D. Harvey, P. Coen, D. W. Tank, Choice-specific sequences in parietal cortex during a virtual-navigation decision task, Nature 484 (7392) (2012) 62–68.

[2] V. Mante, D. Sussillo, K. V. Shenoy, W. T. Newsome, Context-dependent computation by recurrent dynamics in prefrontal cortex, Nature 503 (7474) (2013) 78–84.

[3] M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, K. V. Shenoy, Neural population dynamics during reaching, Nature.

[4] P. T. Sadtler, K. M. Quick, M. D. Golub, S. M. Chase, S. I. Ryu, E. C. Tyler-Kabara, B. M. Yu, A. P. Batista, Neural constraints on learning, Nature 512 (7515) (2014) 423–426.

[5] O. Mazor, G. Laurent, Transient dynamics versus fixed points in odor representations by locust antennal lobe projection neurons, Neuron 48 (4) (2005) 661–673.

[6] M. Rigotti, O. Barak, M. R. Warden, X.-J. Wang, N. D. Daw, E. K. Miller, S. Fusi, The importance of mixed selectivity in complex cognitive tasks, Nature 497 (7451) (2013) 585–590.

[7] K. Daie, M. S. Goldman, E. R. Aksay, Spatial patterns of persistent neural activity vary with the behavioral context of short-term memory, Neuron 85 (4) (2015) 847–860.

[8] M. R. Cohen, J. H. Maunsell, A neuronal population measure of attention predicts behavioral performance on individual trials, J Neurosci 30 (45) (2010) 15241–15253.

[9] A. Luczak, P. Barthó, K. D. Harris, Spontaneous events outline the realm of possible sensory responses in neocortical populations, Neuron 62 (3) (2009) 413–425.

[10] D. Durstewitz, N. M. Vittoz, S. B. Floresco, J. K. Seamans, Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning, Neuron 66 (3) (2010) 438–448.

[11] K. E. Bouchard, N. Mesgarani, K. Johnson, E. F. Chang, Functional organization of human sensorimotor cortex for speech articulation, Nature 495 (7441) (2013) 327–332.

[12] J. P. Cunningham, B. M. Yu, Dimensionality reduction for large-scale neural recordings, Nat Neurosci.

[13] P. Gao, S. Ganguli, On simplicity and complexity in the brave new world of large-scale neuroscience, Current opinion in neurobiology 32 (2015) 148–155.

[14] J. J. DiCarlo, D. Zoccolan, N. C. Rust, How does the brain solve visual object recognition?, Neuron 73 (3) (2012) 415–434.

[15] M. T. Kaufman, M. M. Churchland, S. I. Ryu, K. V. Shenoy, Cortical activity in the null space: permitting preparation without movement, Nat Neurosci.

[16] S. R. Lehky, R. Kiani, H. Esteky, K. Tanaka, Dimensionality of object representations in monkey inferotemporal cortex, Neural Comput. 26 (10) (2014) 2135–62.

[17] M. A. Smith, A. Kohn, Spatial and temporal scales of neuronal correlation in primary visual cortex, J Neurosci 28 (48) (2008) 12591–12603.

[18] R. C. Kelly, M. A. Smith, R. E. Kass, T. S. Lee, Local field potentials indicate network state and account for neuronal response variability, Journal of computational neuroscience 29 (3) (2010) 567–579.
[19] M. Carandini, J. B. Demb, V. Mante, D. J. Tolhurst, Y. Dan, B. A. Olshausen, J. L. Gallant, N. C. Rust, Do we know what the early visual system does?, J Neurosci 25 (46) (2005) 10577–10597.
[20] A. S. Ecker, P. Berens, R. J. Cotton, M. Subramaniyan, G. H. Denfield, C. R. Cadwell, S. M. Smirnakis, M. Bethge, A. S. Tolias, State dependence of noise correlations in macaque primary visual cortex, Neuron 82 (1) (2014) 235–248.
[21] B. B. Averbeck, P. E. Latham, A. Pouget, Neural correlations, population coding and computation, Nat Rev Neurosci 7 (5) (2006) 358–366.
[22] S. C. Wissig, A. Kohn, The influence of surround suppression on adaptation effects in primary visual cortex, J Neurophysiol 107 (12) (2012) 3370–3384.
[23] R. C. Kelly, M. A. Smith, J. M. Samonds, A. Kohn, A. Bonds, J. A. Movshon, T. S. Lee, Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex, J Neurosci 27 (2) (2007) 261–264.
[24] J. A. Movshon, W. T. Newsome, Visual response properties of striate cortical neurons projecting to area mt in macaque monkeys, J Neurosci 16 (23) (1996) 7733–7741.
[25] J. P. Cunningham, Z. Ghahramani, Linear dimensionality reduction: Survey, insights, and generalizations, Journal of Machine Learning Research in press.
[26] J. B. Tenenbaum, V. De Silva, J. C. Langford, A global geometric framework for nonlinear dimensionality reduction, Science 290 (5500) (2000) 2319–2323.
[27] S. T. Roweis, L. K. Saul, Nonlinear dimensionality reduction by locally linear embedding, Science 290 (5500) (2000) 2323–2326.
[28] E. P. Simoncelli, B. A. Olshausen, Natural image statistics and neural representation, Annual review of neuroscience 24 (1) (2001) 1193–1216.
[29] J. P. Jones, L. A. Palmer, The two-dimensional spatial structure of simple receptive fields in cat striate cortex, J Neurophysiol 58 (6) (1987) 1187–1211.
[30] D. Field, D. Tolhurst, The structure and symmetry of simple-cell receptive-field profiles in the cat's visual cortex, Proceedings of the Royal society of London. Series B. Biological sciences 228 (1253) (1986) 379–400.
[31] J. R. Cavanaugh, W. Bair, J. A. Movshon, Nature and interaction of signals from the receptive field center and surround in macaque v1 neurons, J Neurophysiol 88 (5) (2002) 2530–2546.
[32] D. J. Heeger, Normalization of cell responses in cat striate cortex, Visual Neurosci 9 (02) (1992) 181–197.
[33] M. Carandini, D. J. Heeger, Normalization as a canonical neural computation, Nature Reviews Neuroscience 13 (1) (2012) 51–62.
[34] O. Schwartz, E. P. Simoncelli, Natural signal statistics and sensory gain control, Nat Neurosci 4 (8) (2001) 819–825.
[35] J. H. v. Hateren, A. v. d. Schaaf, Independent component filters of natural images compared with simple cells in primary visual cortex, Proceedings: Biological Sciences 265 (1394) (1998) 359–366.
[36] M. Hawken, A. Parker, J. Lund, Laminar organization and contrast sensitivity of direction-selective cells in the striate cortex of the old world monkey, J Neurosci 8 (10) (1988) 3541–3548.
[37] D. Smyth, B. Willmore, G. E. Baker, I. D. Thompson, D. J. Tolhurst, The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation, J Neurosci 23 (11) (2003) 4746–4759.
[38] S. V. David, W. E. Vinje, J. L. Gallant, Natural stimulus statistics alter the receptive field structure of v1 neurons, J Neurosci 24 (31) (2004) 6991–7006.
[39] V. Talebi, C. L. Baker, Natural versus synthetic stimuli for estimating receptive field models: a comparison of predictive robustness, J Neurosci 32 (5) (2012) 1560–1576.
[40] B. A. Olshausen, D. J. Field, How close are we to understanding v1?, Neural computation 17 (8) (2005) 1665–1699.
[41] N. C. Rust, O. Schwartz, J. A. Movshon, E. P. Simoncelli, Spatiotemporal elements of macaque v1 receptive fields, Neuron 46 (6) (2005) 945–956.
[42] J. Touryan, G. Felsen, Y. Dan, Spatial structure of complex cell receptive fields measured with natural images, Neuron 45 (5) (2005) 781–791.

[43] G. Felsen, J. Touryan, F. Han, Y. Dan, Cortical sensitivity to visual features in natural scenes, PLoS Bio 3 (10) (2005) e342.
[44] D. Raposo, M. T. Kaufman, A. K. Churchland, A category-free neural population supports evolving demands during decision-making, Nat Neurosci 17 (2014) 1784–1792.
[45] J. Semedo, A. Zandvakili, A. Kohn, C. K. Machens, B. M. Yu, Extracting latent structure from multiple interacting neural populations, in: Advances in Neural Information Processing Systems, 2014, pp. 2942–2950.
[46] L. Busse, A. R. Wade, M. Carandini, Representation of concurrent stimuli by population activity in visual cortex, Neuron 64 (6) (2009) 931–942.
[47] K. A. Martin, S. Schröder, Functional heterogeneity in neighboring neurons of cat primary visual cortex in response to both artificial and natural stimuli, The Journal of Neuroscience 33 (17) (2013) 7325–7344.
[48] J. Fiser, C. Chiu, M. Weliky, Small modulation of ongoing cortical dynamics by sensory input during natural vision, Nature 431 (7008) (2004) 573–578.
[49] P. Berkes, G. Orbán, M. Lengyel, J. Fiser, Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment, Science 331 (6013) (2011) 83–87.
[50] M. Okun, P. Yger, S. L. Marguet, F. Gerard-Mercier, A. Benucci, S. Katzner, L. Busse, M. Carandini, K. D. Harris, Population rate dynamics and multineuron firing patterns in sensory cortex, J Neurosci 32 (48) (2012) 17108–17119.
[51] A. Arieli, A. Sterkin, A. Grinvald, A. Aertsen, Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses, Science 273 (5283) (1996) 1868–1871.
[52] W. E. Vinje, J. L. Gallant, Sparse coding and decorrelation in primary visual cortex during natural vision, Science 287 (5456) (2000) 1273–1276.
[53] W. E. Vinje, J. L. Gallant, Natural stimulation of the nonclassical receptive field increases information transmission efficiency in v1, J Neurosci 22 (7) (2002) 2904–2915.
[54] M. Pecka, Y. Han, E. Sader, T. D. Mrsic-Flogel, Experience-dependent specialization of receptive field surround for selective coding of natural scenes, Neuron 84 (2) (2014) 457–469.
[55] B. A. Olshausen, et al., Emergence of simple-cell receptive field properties by learning a sparse code for natural images, Nature 381 (6583) (1996) 607–609.
[56] E. Froudarakis, P. Berens, A. S. Ecker, R. J. Cotton, F. H. Sinz, D. Yatsenko, P. Saggau, M. Bethge, A. S. Tolias, Population code in mouse v1 facilitates readout of natural scenes through increased sparseness, Nat Neurosci.