

Published in final edited form as:

Nature. 2019 June 27; 571(7765): 361–365. doi:10.1038/s41586-019-1346-5.

High-dimensional geometry of population responses in visual cortex

Carsen Stringer^{*,1,2}, Marius Pachitariu^{*,1,3}, Nicholas Steinmetz^{3,†}, Matteo Carandini^{‡,4}, and Kenneth D. Harris^{‡,3}

¹HHMI Janelia Research Campus, Ashburn, Virginia, 20147, USA

²Gatsby Computational Neuroscience Unit, UCL, London W1T 4JG, UK

³UCL Institute of Neurology, London WC1E 6DE, UK

⁴UCL Institute of Ophthalmology, London EC1V 9EL, UK

Abstract

A neuronal population encodes information most efficiently when its stimulus responses are high-dimensional and uncorrelated, and most robustly when they are correlated and lower-dimensional. Here, we analyzed the dimensionality of the encoding of natural images by large visual cortical populations recorded from awake mice. Evoked population activity was high dimensional, with correlations obeying an unexpected power-law: the n^{th} principal component variance scaled as $1/n$. This scaling was not inherited from the $1/f$ spectrum of natural images, because it persisted after stimulus whitening. We proved mathematically that if the variance spectrum decayed any slower, the population code could not be smooth, allowing small changes in input to dominate population activity. The theory also predicts larger power-law exponents for lower-dimensional stimulus ensembles, which we validated experimentally. These results suggest that coding smoothness may represent a fundamental constraint governing correlations in neural population codes.

Introduction

The visual cortex contains millions of neurons, and the patterns of activity that images evoke in these neurons form a “population code”. The structure of this code is largely unknown,

Correspondence and requests for materials should be addressed to stringerc@hhmi.org, kenneth.harris@ucl.ac.uk.

[†]Present address: Department of Biological Structure, University of Washington, Seattle, WA 98195, USA.

*equal first authors

‡equal senior authors.

Author Contributions

Conceptualization, C.S., M.P., N.S., M.C. and K.D.H.; Methodology, C.S., M.P., N.S. and K.D.H.; Software, C.S. and M.P.; Investigation, C.S., M.P., N.S. and K.D.H.; Writing, C.S., M.P., N.S., M.C. and K.D.H.; Resources, M.C. and K.D.H. Funding acquisition, M.C. and K.D.H.

Author Information

Reprints and permissions information is available at www.nature.com/reprints.

The authors declare no competing interests.

Materials

All of the processed deconvolved calcium traces are available on figshare[36] (figshare.com/articles/Recordings_of_ten_thousand_neurons_in_visual_cortex_in_response_to_2_800_natural_images/6845348), together with the image stimuli. The code is available on github (github.com/MouseLand/stringer-pachitariu-et-al-2018b).

due to the lack of techniques able to record from large populations. Nonetheless, the population code is the subject of long-standing theories.

One such theory is the “efficient coding hypothesis” [1, 2, 3], which maintains that the neural code maximizes information transmission by eliminating correlations in natural image inputs. Such codes are high-dimensional and sparse, which can allow complex features to be read out by simple downstream networks [4, 5, 6].

However, several studies have suggested that neural codes are confined to low-dimensional subspaces (“planes”) [7, 8, 9, 10, 11, 12, 13, 14, 15]. Codes of low planar dimension are correlated and redundant, allowing for robust computations of stimuli in the face of noise [16, 17]. Nevertheless, low planar dimension is inevitable given stimuli or tasks of limited complexity [18]: the responses to a set of n stimuli, for example, have to lie in an n dimensional subspace. The planar dimension of the cortical code thus remains an open question, which can only be answered by recording the responses of large numbers of neurons to large numbers of stimuli.

Here, we recorded the simultaneous activity of ~10,000 neurons in mouse visual cortex, in response to thousands of natural images. We found that stimulus responses were neither uncorrelated (“efficient coding”) nor low-dimensional. Instead, responses occupied a multidimensional space with the variance in the n^{th} dimension scaling as a power law $n^{-\alpha}$, where $\alpha \approx 1$. We showed mathematically that if variances decay slower than a power law with exponent $\alpha = 1 + 2/d$, where d is the dimension of the input ensemble, then the space of neural activity must be non-differentiable – i.e. not smooth. We varied the dimensionality of the stimuli d and found that the neural responses respected this lower bound. These findings suggest that the population responses are constrained by efficiency, to make best use of limited numbers of neurons, and smoothness, which allows similar images to evoke similar responses.

Simultaneous recordings of ~10,000 neurons

To obtain simultaneous recordings of ~10,000 cells from mouse V1, we employed resonance-scanning two-photon calcium microscopy, using 11 imaging planes spaced at $35\mu\text{m}$ (Fig. 1a). The slow timecourse of the GCaMP6s sensor allowed activity to be detected at a 2.5 Hz scan rate, and an efficient data processing pipeline [19] allowed large numbers of cells to be detected accurately (Fig. 1b). Natural image scenes (Imagenet database [20]) were presented on an array of 3 monitors surrounding the mouse (Fig. 1c), at an average of 1 image/s. Cells were tuned to these natural image stimuli: in experiments in which responses to 32 images were averaged over 96 repeats (Fig. 1d), stimulus responses accounted for $55.4 \pm 3.3\%$ (SE, $n=4$ recordings) of the trial-averaged variance. Consistent with prior reports [21, 22, 23], neuronal responses were sparse: only a small fraction of cells ($13.4 \pm 1.0\%$ SE, $n=4$ recordings) were driven more than two standard deviations above their baseline firing rate by any particular stimulus.

For our main experiments, we assembled a sequence of 2,800 image stimuli, and presented these stimuli twice in the same order, to maximize the number of images presented while

still allowing analyses based on cross-validation (Fig. 1e). Most neurons ($81.4 \pm 5.1\%$ SE, $n=7$ recordings) showed correlation between repeats at $p < 0.05$ (Extended Data Fig. 1a,b). Nevertheless, consistent with previous reports [24], responses showed substantial trial-to-trial variability. Cross-validation showed that stimulus responses accounted for on average $13.2 \pm 1.5\%$ of the single-trial variance (Extended Data Fig. 1c), and the average signal-to-noise ratio was $17.3 \pm 2.4\%$ (Fig. 1f). This level of trial-to-trial variability was not due to our particular recording method: measuring responses to the same stimuli electrophysiologically yielded a similar signal-to-noise ratio (Extended Data Fig. 2). Despite trial-to-trial variability, however, population activity recorded on a single trial contained substantial information about the sensory stimuli. A simple nearest-neighbor decoder, trained on one repeat and tested on the other, was able to identify the presented stimulus with up to 75.5% accuracy (Fig. 1g; range 25.4%-75.5%; median 41.7% compared to chance level of 0.036%, $n=7$ recordings). Decoding accuracy did not fully saturate at a population size of 10,000, suggesting that performance would increase further with more neurons.

Neurons had similar visual properties to previous reports [23, 25], and their responses were only partially captured by classical linear-nonlinear models, consistent with previous visual cortex studies [26, 27, 28, 29, 30]. We calculated a receptive field (RF) for each cell from its responses to natural images in two ways: by fitting linear RFs regularized with a reduced rank method; or by searching for an optimal Gabor filter that was rectified/quadrature filtered to simulate classical simple/complex cell responses. As expected from retinotopy, the RF locations of simultaneously recorded neurons overlapped but there was a high diversity of receptive field sizes and shapes (Fig. 1h; Extended Data Fig. 3). Both RF models, however, explained less than 20% of the stimulus-related variance (the linear model explained $11.4 \pm 0.7\%$ SE, and the Gabor model explained $18.5 \pm 1.0\%$ SE, $n=7$ recordings each).

Power-law scaling of dimensionality

To characterize the geometry of the population code for visual stimuli, we developed a method of cross-validated principal component analysis (cvPCA). cvPCA measures the reliable variance of stimulus-related dimensions, excluding trial-to-trial variability from unrelated cognitive/behavioral variables or noise. It accomplishes this by computing the covariance of responses between training and test presentations of an identical stimulus ensemble (Fig. 2a). Because only stimulus-related activity will be correlated across repeats, cvPCA provides an unbiased estimate of the stimulus-related variance. In simulations using the same noise statistics as our recordings, we confirmed that this technique can recover the true variances (Extended Data Fig. 4, Extended Data Fig. 5, Supplementary Discussion 1).

This method revealed that the visual population responses did not lie on any low-dimensional plane within the space of possible firing patterns. The amount of variance explained continued to increase as further dimensions were included, without saturating at any dimensionality below the maximum possible (Fig. 2b). As a control analysis, we applied cvPCA to neural responses to only 32 images shown many times – whose responses must by definition lie in a 32-dimensional subspace – and indeed observed a saturation of the variance after 32 dimensions.

This analysis revealed an unexpected finding: the fraction of neural variance in planes of successively larger dimensions followed a power law. The eigenspectrum, i.e. the function summarizing the variance of the n^{th} principal component, had a magnitude approximately proportional to $1/n$ (Fig. 2c), reflecting successively less variance in dimensions encoding finer stimulus features (Extended Data Fig. 6). Its power-law structure did not result from averaging over experiments: analysis of data from each mouse individually revealed power-law behavior in every case (Fig. 2d). The scaling exponent of the power law had a peak just above 1 (1.04 ± 0.02 SE, $n = 7$ recordings, Fig. 2e). This eigenspectrum reflected correlations between neurons, and was not the consequence of a lognormal distribution of firing rates or signal variance (Extended Data Fig. 7). In addition, this result could not be explained by classical models of visual cortical receptive fields: a model of simple/complex Gabor receptive fields with parameters fit to single cell responses (Fig. 1h) had lower dimensionality than the neural responses (Fig. 2f).

The range of dimensions over which the power law held grew larger the more neurons and stimuli were considered. To show this, we repeated the analyses on randomly-chosen subsets of neurons or stimuli (Fig. 2g-h). To quantify the accuracy of the power law, we fit the log variance of the n^{th} dimension to $\log(n)$ using linear regression over the range 11-500 dimensions. Both the correlation coefficient and slope (which reflects the power law exponent) approached 1 for increasing subset sizes (Fig. 2i,j, Extended Data Fig. 8). Electrophysiological recordings – with fewer recorded neurons and fewer presented stimuli – showed the same eigenspectrum as a similarly-sized subset of the two-photon data (Extended Data Fig. 9). We conclude that the power law held accurately over ~ 2 orders of magnitude in these recordings, and infer it would likely extend ever further if more yet neurons and stimuli could be analyzed.

Power-law and stimulus statistics

Natural images have a power-law structure[31, 32] (Fig. 3a), but this did not cause the neural code's power-law. To show this, we removed the image power law by spatially whitening the images, and presented the whitened stimuli to 3 of the mice. Although the power law in the image pixels was abolished, the power law in the neural responses remained (Fig. 3b). Furthermore, the eigen-spectrum of neural responses could not be explained by straightforward receptive field properties: the responses of a simple/complex Gabor model produced eigenspectra that decayed more quickly than the actual responses, and were worse fit by a power-law ($p < 10^{-3}$, Wilcoxon ranksum test on Pearson correlations, Fig. 3a,b).

The power law eigenspectra also did not arise from other characteristics of natural images. To investigate the role of long-range image correlations, we constructed spatially localized image stimuli, in which the region outside the classical RF was replaced by gray. Again, the power law persisted with exponent close to 1 (Fig. 3c). Finally, we removed any natural image structure and showed sparse noise stimuli to the mice (Fig. 3d). Again, we observed a power-law spectrum, with exponent close to 1, although slightly larger than for the natural image stimuli (1.13 ± 0.044 SE, $n=3$ recordings; $p = 0.067$, Wilcoxon two-sided rank sum test). We conclude that power-law spectra do not reflect neural processing of a property specific to natural images.

Power-law and stimulus dimensionality

Power law eigenspectra are observed in many scientific domains, and are related to the smoothness of the underlying functions. For example, the Fourier spectrum of a differentiable function of one variable must decay at least as fast as a power law of exponent 1 (see e.g. Ref. [33]). We therefore theorized that the variance power law might be related to smoothness of the neural responses. We showed mathematically that if the sensory stimuli presented can be characterized by d parameters, and if the mapping from these parameters to (noise-free) neural population responses is differentiable, then the population eigenspectrum must decay asymptotically faster than a power law of exponent $\alpha = 1 + 2/d$ (Supplementary Discussion §2). Conversely, if the eigenspectrum decays slower than this, a smooth neural code is impossible as the neural responses must lie on a fractal rather than a differentiable manifold.

This mathematical analysis gave rise to an experimental prediction. For a high-dimensional stimulus ensemble such as natural images, d will be large so $1 + 2/d \approx 1$, close to the exponent we observed. However for smaller values of d , the power-law must have larger exponents if fractality is to be avoided. We therefore hypothesized that lower-dimensional stimulus sets would evoke population responses with larger power-law exponents.

We obtained stimulus ensembles of dimensionality $d = 8$ and $d = 4$ by projecting the natural images onto a set of d basis functions (Fig. 3e,f). For a stimulus ensemble of dimensionality $d = 1$ we used drifting gratings, parameterized by a single direction variable. Consistent with the hypothesis, stimulus sets with $d = 8, 4$, and 1 yielded power-law scaling of eigenvalues with exponents of 1.49, 1.65, and 3.43, near but above the lower bounds of 1.25, 1.50, and 3.00 predicted by the $1 + 2/d$ exponent (Fig. 3h). These results suggest that the neural responses lie on a manifold that is almost as high-dimensional as possible without becoming fractal. The eigenspectra of responses simulated from the simple/complex cell model decayed even faster, suggesting a differentiable, but lower dimensional representation (Fig. 3i).

Discussion

We found that the variance of n^{th} dimension of visual cortical population activity decays as a power of n , with exponent $\alpha \approx 1 + 2/d$ where d is the dimensionality of the space of sensory inputs. The population eigenspectrum reflects the fraction of neural variance devoted to representing coarse and fine stimulus features (Supplementary Discussion §2, §3; Extended Data Fig. 6). If the eigenspectrum decayed slower than $n^{-1-2/d}$ then the neural code would emphasize fine stimulus features so strongly that it could not be differentiable. Our results therefore suggest that the eigenspectrum of the visual cortical code decays almost as slowly as possible consistent with smooth neural coding.

To illustrate the consequences of eigenspectrum decay for neural codes, we simulated various one-dimensional coding schemes in populations of 1,000 neurons, and visualized them by random projection into 3d space (Fig. 4). The stimulus was parameterized by a single circular variable, like the direction of a grating. A low-dimensional code with two

nonzero eigenvalues produced a circular neural manifold (Fig. 4a). An uncorrelated, high-dimensional code where each neuron responded to a different stimulus produced 1,000 equal variances, consistent with the efficient coding hypothesis (Fig. 4b). However this code did not respect distances: responses to stimuli separated by just a few degrees differed as much as responses to diametrically opposite stimuli, and the neural manifold appeared as a spiky, discontinuous ball. Powerlaw codes (Supplementary Discussion §2.7, Example 2) show a scale-free geometry, whose smoothness depends on the exponent α (Fig. 4c-e). A power-law code with $\alpha = 2$ (Fig. 4c) was a non-differentiable fractal because the many dimensions encoding fine stimulus details (blue) together outweighed the few dimensions encoding large-scale stimulus differences (red). At the critical exponent of $\alpha = 3$ (which equals $1 + 2/d$ since $d = 1$), the neural manifold was on the border of differentiability; the code represents fine differences between stimuli while still preserving large-scale stimulus features (Fig. 4d). A higher exponent led to a smoother neural manifold (Fig. 4e).

What are the computational consequences of these different coding geometries, and what advantage might the brain gain from a power-law code with a close-to-critical exponent? The efficient coding hypothesis suggests information is optimally encoded when responses to different stimuli are as different as possible. However, such codes carry a cost in terms of generalization and robustness: if the neural responses to any pair of stimuli were orthogonal, then stimuli that differ only in tiny details would have completely different representations (Supplementary Discussion §2.1). Similar behavior can be seen in deep neural network architectures that misclassify “adversarial images” differing only very slightly from training examples [34, 35]. We suggest that a power-law code just above the critical exponent represents a balance between efficient, high-dimensional codes, and smooth codes enabling generalization.

Methods

All experimental procedures were conducted according to the UK Animals Scientific Procedures Act (1986). Experiments were performed at University College London under personal and project licenses released by the Home Office following appropriate ethics review.

Animals and surgery

We used mice bred to express GCaMP6s in excitatory neurons in our recordings: 13 recordings from TetO-GCaMP6s x Emx1-IRES-Cre mice (available as JAX 024742 and JAX 005628); 3 recordings from a Camk2a-tTA, Ai94 GCaMP6s 2tg x Emx1-IRES-Cre mouse (available as JAX 024115 and JAX 005628); and 2 recordings from a Camk2a-tTA, Ai94 GCaMP6s 2tg x Rasgrf-Cre mouse (available as JAX 024115 and JAX 022864). We also used mice bred to express tdTomato in inhibitory neurons (GAD-IRES-Cre x CAG-tdTomato, available as JAX 010802 and JAX 007909) in 14 recordings. In this case, GCaMP6s was expressed virally, and excitatory neurons were identified by lack of tdTomato expression. These mice were male and female, and ranged from age 2 to 8 months. We recorded from sufficiently many mice to draw scientific conclusions (8 mice in total). There was no randomization or blinding done in the study.

Surgical methods were similar to those described elsewhere [19, 37]. Briefly, surgeries were performed in adult mice (P35–P125) under isoflurane anesthesia (5% for induction, 0.5–1% during the surgery) in a stereotaxic frame. Before surgery, Rimadyl was administered as a systemic analgesic and lidocaine was administered locally at the surgery site. During the surgery we implanted a head-plate for later head-fixation, and made a craniotomy of 3–4 mm in diameter with a cranial window implant for optical access. In Gad-Cre x tdTomato transgenics, we targeted virus injections (AAV2/1-hSyn-GCaMP6s, University of Pennsylvania Vector Core, 50–200 nl, 1–3 x 10¹² GC/ml) to monocular V1 (2.1–3.3 mm laterally and 3.5–4.0 mm posteriorly from Bregma), using a beveled micropipette and a Nanoject II injector (Drummond Scientific Company, Broomall, PA 1) attached to a stereotaxic micromanipulator. To obtain large fields of view for imaging, we typically performed 4–8 injections at nearby locations, at multiple depths (~500 μ m and ~200 μ m). Rimadyl was then used as a post-operative analgesic for three days, delivered to the mice via their drinking water.

Data acquisition

We used a 2-photon microscope (Bergamo II multiphoton imaging microscope, Thorlabs, Germany) to record neural activity, and ScanImage [38] for data acquisition, obtaining 10622 ± 1690 (standard deviation) neurons in the recordings. The recordings were performed using multi-plane acquisition controlled by a resonance scanner, with planes spaced 30–35 μ m apart in depth. Ten or twelve planes were acquired sequentially, scanning the entire stack repeatedly at 3 or 2.5 Hz. Because plane scanning was not synchronized to stimulus presentation, we aligned the stimulus onsets to each of the planes separately, and computed stimulus responses from the first two frames acquired after stimulus onset for each plane.

The mice were free to run on an air-floating ball and were surrounded by three computer monitors arranged at 90° angles to the left, front and right of the animal, so that the animal's head was approximately in the geometric center of the setup. Data from running and non-running periods was analyzed together.

For each mouse, recordings were made over multiple days, always returning to the same field of view (in one mouse, two fields of view were used). For each mouse, a field of view was selected on the first recording day such that 10,000 neurons could be observed, with clear calcium transients and a retinotopic location (identified by neuropil fluorescence) localized on the monitors. If more than one potential field of view satisfied these criteria, we chose either a horizontally and vertically central retinotopic location, or a lateral retinotopic location, at 90° from the center, but still centered vertically. The retinotopic location of the field of view (central or lateral) was unrelated to variance spectra. We also did not observe significant differences between recordings obtained from different modes of GCaMP expression (transgenic vs viral injection). Thus, we pooled data over all conditions.

Visual stimuli

During two-photon recordings, all stimuli other than sparse noise stimuli were presented for 0.5 sec, alternating with a gray-screen inter-stimulus interval lasting a random time between

0.3 and 1.1 s. During electrophysiological recordings, all stimuli were presented for 400ms, with a uniformly-distributed inter-stimulus interval of 300-700ms.

Image stimuli were selected from the ImageNet database [20], from ethologically-relevant categories: “birds”, “cat”, “flowers”, “hamster”, “holes”, “insects”, “mice”, “mushrooms”, “nests”, “pellets”, “snakes”, “wildcat”. Images were chosen manually to ensure that less than 50% of the image was a uniform background, and to contain a mixture of low and high spatial frequencies. The images were uniformly contrast normalized. This was achieved by subtracting the local mean brightness and dividing by the local mean contrast (standard deviation of pixel values); the local mean and standard deviation were both computed using a Gaussian filter of standard deviation 30° . Each presented stimulus consisted of a different normalized image from ImageNet (2800 different images) replicated across all 3 screens, but at a different rotation on each screen (Fig. 1c).

For the main two-photon recordings, these 2800 stimuli were presented twice, in the same order each time. In the electrophysiological recordings, 700 of these same stimuli were presented twice in the same order each time. Additionally, in a subset of imaged mice (4 out of 6), we presented a smaller set of 32 images, presented in a randomized order 90-114 times, to enable more accurate estimation of trial-averaged responses.

We also presented partially spatially whitened versions of the 2800 natural images. To compute spatially whitened images, we first computed the two-dimensional Fourier spectrum for each image, and averaged the spectra across images. We then whitened each image in the frequency domain by dividing its Fourier transform by the averaged Fourier spectrum across all images. The rescaled Fourier transform of the image was transformed back into the pixel domain by computing its inverse two-dimensional Fourier transform and retaining the real part. Each image was then intensity-scaled to have the same mean and standard deviation pixel values as the original.

The eight- and four-dimensional stimuli were constructed using a reduced-rank regression model. We first used reduced rank regression to predict the neuronal population responses R from the natural images I (N_{pixels} by N_{stimuli}) via a d -dimensional bottleneck :

$$R = A^T B I$$

where A is a matrix of size d by N_{neurons} and B is a matrix of size d by N_{pixels} . The dimensionality d was either eight or four depending on the set of stimuli being constructed. The columns of B represent the image dimensions which linearly explain the most variance in the neural population responses. The stimuli were the original 2800 natural images projected onto the reduced-rank subspace B : $I_{\text{low-D}} = B^T B I$.

In addition to natural image stimuli, we also presented drifting gratings and sparse noise. Drifting gratings of 32 directions, spaced evenly at directions evenly spaced at 11° , were presented 70-128 times each, lasting 0.5 sec each, and with a gray-screen inter-stimulus interval between 0.3 and 1.1 s. They were full-field stimuli (all three monitors) and their spatial frequency was 0.05 cycles per degree and their temporal frequency was 2Hz.

Sparse noise stimuli consisted of white or black squares on a gray background. Squares were of size 5°, and changed intensity every 200 ms. On each frame, the intensity of each square was chosen independently, as white with 2.5% probability, black with 2.5% probability, and gray with 95% probability. The sparse noise movie contained 6000 frames, lasting 20 minutes, and the same movie was played twice to allow cross-validated analysis.

Spontaneous activity was recorded for 30 minutes with all monitors showing a gray or black background, in all but six of 32 image set recordings. In three recordings of 32-natural image responses and three recordings of drifting grating responses, we interspersed the spontaneous activity, recording 30 seconds of spontaneous gray screen activity in between each set of 32 stimuli. In all recordings but these six, there were also occasional blank stimuli (1 out of every 20 stimuli in the 2800 natural image stimuli). The activity during these non-stimulus periods was used to project out spontaneous dimensions from the neuronal population responses (see below).

Calcium imaging processing

Calcium movie data was processed using the Suite2p toolbox [19, 37], available at www.github.com/cortex-lab/Suite2P.

Briefly, the Suite2p pipeline consists of registration, cell detection, ROI classification, neuropil correction, and spike deconvolution. Movie frames are registered using 2D translation estimated by regularized phase correlation, subpixel interpolation and kriging. To detect regions of interest (ROIs; corresponding to cells), Suite2p clusters correlated pixels, using a low-dimensional decomposition of the data to accelerate processing. The number of ROIs is determined automatically via a threshold on pixel correlations. Finally, ROIs were classified as somatic or non-somatic using a classifier trained on a set of human-curated ROIs. The classifier reached 95% agreement with training data, thus allowing us to skip manual curation for most recordings. For neuropil correction, we used the approach of Ref. [39], subtracting from each ROI signal the surrounding neuropil signal scaled by a factor of 0.7; all pixels attributed to an ROI (somatic or not) were excluded from the neuropil trace. After neuropil subtraction, we further subtracted a running baseline of the calcium traces with a sliding window of 60 seconds to remove long timescale additive shifts in the signals. Fluorescence transients were estimated using non-negative spike deconvolution [40] with a fixed timescale of calcium indicator decay of 2 seconds, a method which we found to outperform others on ground truth data [41]. Finally, each cell's deconvolved trace was z-scored with respect to the mean and standard deviation of that cell's trace during a 30-minute period of gray-screen spontaneous activity before or after the image presentations.

All of the processed deconvolved calcium traces are available on figshare[36] (https://figshare.com/articles/Recordings_of_ten_thousand_neurons_in_visual_cortex_in_response_to_2_800_natural_images/6845348), together with the image stimuli.

Data acquisition and processing (electrophysiology)

Neuropixels electrode arrays [42] were used to record extracellularly from neurons in six mice. The mice were between 8 and 24 weeks old at the time of recording, and were of

either gender. The genotypes of the mice were Slc17a7-Cre;Ai95, Snap25-GCaMP6s, TetO-GCaMP6s;CaMKIIa-tTA, Ai32;Pvalb-Cre (two mice), or Emx1-Cre;CaMKIIa-tTA;Ai94. In some cases, other electrophysiological recordings had been made from other locations in the days preceding the recordings reported here. In all cases, a brief (<1 hour) surgery to implant a steel headplate and 3D-printed plastic recording chamber (12mm diameter) was first performed. Following recovery, mice were acclimated to head-fixation in the recording setup. During head-fixation, mice were seated on a plastic apparatus with forepaws on a rotating rubber wheel (five mice) or were on a styrofoam treadmill and able to run (one mouse). Three 20x16cm TFT-LCD screens (LG LP097QX1) were positioned around the mouse at right angles at a distance of 10cm, covering a total of 270x78 degrees visual angle. On the day of recording, mice were again briefly anesthetized with isoflurane while eight small craniotomies were made with a dental drill. After several hours of recovery, mice were head-fixed in the setup. Probes had a silver wire soldered onto the reference pad and shorted to ground; these reference wires were connected to a Ag/AgCl wire positioned on the skull. The craniotomies as well as the wire were covered with saline-based agar. The agar was covered with silicone oil to prevent drying. Probes were each mounted on a rod held by an electronically position-able micromanipulator (uMP-4, Sensapex Inc.) and were then advanced through the agar and through the dura. Once electrodes punctured dura, they were advanced slowly ($10\mu\text{m}/\text{sec}$) to their final depth (4 or 5 mm deep). Electrodes were allowed to settle for approximately 15 minutes before starting recording. Recordings were made in external reference mode with LFP gain=250 and AP gain=500, using SpikeGLX software. Data were preprocessed by re-referencing to the common median across all channels. 6 recordings were performed in 6 different mice, with a total of 14 probes in visual cortex across all experiments.

We spike sorted the data using a modification of Kilosort [43] that tracks drifting clusters, called Kilosort2 [37, 44], available at www.github.com/MouseLand/Kilosort2. Without the modifications, the original Kilosort and similar algorithms can split clusters according to drift of the electrode. Kilosort2 in comparison tracks neurons across drift levels and for longer periods of time (~ 1 hour in our case).

Removal of ongoing activity dimensions

As shown previously [37], approximately half the variance of visual cortical population activity is unrelated to visual stimuli, but represents behavior-related fluctuations. This ongoing activity continues uninterrupted during stimulus presentations, and overlaps with stimulus responses only along a single dimension. Because the present study is purely focused on sensory responses, we projected out the dimensions corresponding to ongoing activity prior to further analysis. The top 32 dimensions of ongoing activity were found by performing principal component analysis on the z-scored ongoing neural activity recorded during a 30-minute period of gray screen stimuli before or after the image presentations. To remove these dimensions from stimulus responses, the stimulus-driven activity was also first z-scored (using the mean and variance of each neuron computed from ongoing activity), then the projection onto the 32 top spontaneous dimensions was subtracted (Extended Data Fig. 4).

In the electrophysiological recordings, we considered stimulus responses in a window of 50 ms or 500 ms following stimulus onset. Therefore, we computed the ongoing activity using these two different bin sizes (50 ms or 500 ms). Then we z-scored the stimulus responses by this ongoing activity. Next we computed the top 10 principal components of the ongoing activity (in both bin sizes) and then subtracted the projection of the stimulus responses onto these dimensions.

Receptive field estimation

We estimated the receptive fields of the neurons, either using a reduced-rank regression model or using a simple/complex Gabor model. In both cases, the model was fit to the mean response of each neuron to half of the 2800 images (I_{train}) over the two repeats. The performance of the model was tested on the mean response of each neuron to the other half of the 2800 images (I_{test}).

Reduced-rank receptive field estimation—To estimate a linear receptive field for each neuron, we used reduced rank regression [45], a self-regularizing method which allowed us to fit all neurons' responses to a single repeat of all 2800 image stimuli. Reduced rank regression predicts high-dimensional outputs from high-dimensional inputs through a linear low-dimensional hidden “bottleneck” representation. We used a 25-dimensional hidden representation to predict each neuron's activity from the image pixel vectors, taking the resulting regressor matrices as the linear receptive fields. These receptive fields explained $11.4 \pm 0.7\%$ (SE, $n=7$ recordings) of the stimulus-related variance on the test set. These were z-scored prior to display in Fig. 1h and Extended Data Fig. 3a.

Model-based receptive field estimation—To fit classical simple/complex receptive fields to each cell, we simulated the responses of a convolutional grid of Gabor filters to the natural images, and fit each neuron with the the filter response most correlated to its response.

The Gabor cell filters $G(x)$ were parametrized by a spatial frequency f , orientation θ , phase ψ , size α and eccentricity β . Defining \mathbf{u} and \mathbf{v} to be unit vectors pointing parallel and perpendicular to the orientation θ .

$$G(x) = \cos(2\pi f x \cdot \mathbf{u} + \psi) e^{-((x \cdot \mathbf{u})^2 + \beta(x \cdot \mathbf{v})^2)/2\alpha^2}$$

We constructed 12288 Gabor filters, with centers spanning a 9 by 7 grid spaced at 5 pixels, and with parameters f , θ , ϕ , α , and β ranging from (0.01, 0, 0, 3, 1) to (0.13, 157, 315, 12, 2.5) with (7, 8, 8, 4, 4) points sampled of each parameter respectively. The parameters were equally spaced along the grid (e.g. f was sampled at 0.01, 0.03, 0.05, 0.07, 0.09, 0.11, 0.13).

Simple cell responses were simulated by passing the dot product of the image with the filter through a rectifier function $r(x) = \max(0, x)$. Complex cell responses were simulated as the root-mean-square response of each unrectified simple cell filter and the same filter with phase ψ shifted by 90° . A neuron's activity was predicted as a linear combination of a simple cell and its complex cell counterpart, with weights estimated by linear regression.

Each neuron was assigned to the filter which best predicted its responses to the training images (Extended Data Fig. 3b-h). This simple/complex Gabor model explained $18.4 \pm 0.1\%$ of the stimulus-related variance on the test set.

We also evaluated a model of Gabor receptive fields including divisive normalization [46]. To do so, the response of each of modeled simple or complex cell filter was divided by the summed, normalized responses of all the other simple and complex cells at this retinotopic location. The experimentally measured response of each neuron was then predicted as a linear combination of simple and complex responses to the best-fitting Gabor, with weights estimated by linear regression. $45.4\% \pm 1.0\%$ (mean \pm SE) of cells were better fit by the divisive normalization model. However, while divisive normalization changed the optimal parameters fit to many cells (Extended Data Fig. 3i-n), the resulting eigenspectra were indistinguishable from a model with no normalization (Extended Data Fig. 3o-u).

Sparseness estimation

To estimate the sparseness of single-cell responses to the image stimuli, we counted how many neurons were driven more than two standard deviations above their baseline rate by any given stimulus. This was estimated using 4 experiments in which 32 natural images were repeated <90 times. We computed each neuron's tuning curve by averaging over all repeats. The standard deviation of the tuning curve is computed for each neuron across stimuli. Baseline rate was defined as the mean firing rate during all spontaneous activity periods, without visual stimuli. A neuron was judged as responsive to a given stimulus if its response was more than two times this standard deviation plus its baseline firing rate.

Decoding accuracy from 2,800 stimuli

To decode the stimulus identity from the neural responses (Fig. 1g), we built a simple nearest neighbor decoder based on correlation. The first stimulus presentation was used as training set while the second presentation was used as test set. We correlated the population responses for a individual stimulus in the test set with the population responses from all stimuli in the training set. The stimulus with the maximum correlation was then assigned as our prediction. We defined the decoding accuracy as the fraction of correctly labelled stimuli.

Signal-to-noise ratio and explained variance

To compute the tuning-related signal-to-noise ratio (SNR; Fig. 1f), we first estimated the signal variance of each neuron \hat{V}_{sig} as the covariance of its response to all stimuli across two repeats (for neuron c , $\hat{V}_{\text{sig}} = \text{Cov}_s[f_1(c, s), f_2(c, s)]$, see Supplementary Discussion §1). The noise variance $\hat{V}_{\text{noise}} = V_{\text{tot}} - \hat{V}_{\text{sig}}$ was defined as the difference between the within-repeat variance (reflecting both signal and noise) and this signal variance estimate, and the SNR as their ratio. The SNR estimate is positive when a neuron has responses to stimuli above its noise baseline; note that as \hat{V}_{sig} is an unbiased estimate, it can take negative values when he true signal variance is zero.

To compute the percentage of explained variance for each neuron (Extended Data Fig. 1c), we divided the estimated signal variance by the total variance across trials (averaged across the repeats):

$$EV = \frac{\hat{v}_{\text{sig}}}{\frac{1}{N_r} \sum_r \text{Var}_t[f_{t,r}]}$$

Note that this formula is similar to the Pearson correlation of a neuron's responses between two repeats. In Pearson correlation, the numerator is the same, equal to the covariance between repeats, but the denominator is the geometric rather than arithmetic mean of the variances of the two repeats.

cvPCA method

The cvPCA method is fully described in Supplementary Discussion §1, characterized mathematically in §1.1 and §3.6, and analyzed in simulation in Extended Data Fig. 5. In brief, the difference between this approach and standard PCA (e.g. Refs. [47, 48]) is that it compares the activity on training and test repeats to obtain an estimate of the stimulus-related (“signal”) variance, discounting variance from trial-to-trial variability (“noise”).

Denote the response of neuron c to repeat r of stimulus s by $f_r(c, s)$, define the signal as the expected response, which will be equal for both repeats: $\phi(c, s) = \mathbb{E}[f_r(c, s)]$, and the noise on repeat r to be the residual after the expected response is subtracted: $v_r(c, s) = f_r(c, s) - \phi(c, s)$. By definition, the noise has zero expectation: $\mathbb{E}v_r[v_r(c, s)/c, s] = 0$ for all r, c , and s . Let $\hat{\mathbf{u}}_n$ denote the n^{th} principal component eigenvector, computed from repeat 1.

If we estimated the variance of the projection of activity onto $\hat{\mathbf{u}}_n$ using a single repeat, it would contain a contribution from both the signal and the noise. However, because stimulus-independent variability is by definition uncorrelated between repeats of the same stimulus, we can obtain an unbiased estimate of the signal variance, from the covariance across these independent repeats:

$$\begin{aligned} & \mathbb{E}_{v_1, v_2} \left[\frac{1}{N_s} \sum_{s_i=1}^{N_s} (\mathbf{f}_1(s_i) \cdot \hat{\mathbf{u}}_n)(\mathbf{f}_2(s_i) \cdot \hat{\mathbf{u}}_n) \right] \\ &= \frac{1}{N_s} \sum_{s_i=1}^{N_s} (\phi(s_i) \cdot \hat{\mathbf{u}}_n)^2 \end{aligned}$$

Thus, if \mathbf{u}_n is an eigenvector of the population signal variance, the cvPCA method will produce an unbiased estimate of the signal principal component variances. As shown in Supplementary Discussion 1.1, this will occur if response variability comprises a mixture of multiplicative response gain changes, correlated additive variability orthogonal to the stimulus dimensions, and uncorrelated noise. Although additive variability in the signal space could in principle downwardly bias the estimated signal variance, other work confirms

that under conditions similar to those analyzed here there is little additive variability in the signal space [37]; furthermore, simulations confirm that the amount of such variability present in our recordings does not substantially bias the estimation of signal eigenspectra with cvPCA (Extended Data Fig. 5).

We ran cvPCA 10 times on each dataset, on each iteration randomly sampling each stimulus' population response from the two repeats without replacement. Thus, $\mathbf{f}_1(s)$ could be the population response from either the first or second repeat, with $\mathbf{f}_2(s)$ being the response from the other. The displayed eigenspectra are averages over the 10 different runs.

Simulations

To verify that cvPCA method was able to accurately estimate signal eigenspectra in the presence of noise, we analyzed simulated data where the true eigenspectrum was known by construction, and stimulus responses were corrupted by noise. Mathematical analyses (Supplementary Discussion 1.1 and 3.6) showed that noise consisting of multiplicative gain modulation, additive noise orthogonal to signal dimensions, or independent additive noise should not bias the expected eigenspectrum estimate, but that correlated additive noise in the stimulus dimensions could potentially lead the eigenspectrum to be underestimated. We therefore first concentrated on this possibility.

To create the test data, we first simulated noise-free sensory responses whose eigenspectrum followed an exact power law, with three possible exponents: $\alpha = 0.5, 1.0, \text{ or } 1.5$. To simulate the responses of $N_c = 10,000$ neurons to $N_s = 2,800$ stimuli with this exact eigenspectrum, we first constructed a set of random orthogonal eigenvectors by performing singular value decomposition on a $N_c \times N_s$ matrix A of independent standard Gaussian variates: $A = USV^T$. We created a diagonal matrix D_α , whose n^{th} diagonal entry was $n^{-\alpha/2}$, and created the $N_c \times N_s$ matrix of simulated noise-free responses as $\phi = UDV^T$.

Additive noise

To simulate correlated additive noise in the stimulus space (Extended Data Fig. 5c), we constructed noise whose eigenspectrum matched that observed experimentally. To find the empirical noise eigenspectrum, we first estimated the total variance of the n^{th} principal component as

$$\hat{\Lambda}_n = \frac{1}{2} \left[\frac{1}{N} \sum_{s_i=1}^{N_s} (\mathbf{f}_1(s_i) \cdot \mathbf{u}_n)^2 + \frac{1}{N} \sum_{s_i=1}^{N_s} (\mathbf{f}_2(s_i) \cdot \mathbf{u}_n)^2 \right]$$

and estimated the signal variance using cvPCA as

$$\hat{\lambda}_n = \frac{1}{N} \sum_{s_i=1}^{N_s} (\mathbf{f}_1(s_i) \cdot \mathbf{u}_n)(\mathbf{f}_2(s_i) \cdot \mathbf{u}_n).$$

The estimated noise spectrum was the difference between total variance and noise variance: $\hat{\delta}_n = \hat{\Lambda}_n - \hat{\lambda}_n$. This spectrum reflects the summed magnitude of both correlated and uncorrelated noise in the signal dimensions, and is shown in Extended Data Fig. 5b. Responses corrupted by additive noise were simulated as $\phi + b_\alpha U V^T$, where U is a diagonal matrix with entries δ_n , and the scale factor b_α ensured that, as in the data, the simulation showed a total of 14% reliable variance. The scale factors were found by search to be 2.62, 2.52, and 2.41 for the signal eigenspectrum exponents $\alpha = 0.5, 1.0, \text{ and } 1.5$, respectively.

Multiplicative noise

To simulate multiplicative noise (Extended Data Fig. 5d), responses were multiplied by an amplitude factor that was constant across neurons, but was drawn independently for each stimulus and repeat. To simulate an appropriately skewed distribution of gains, the scale factor was distributed as 0.5 plus an exponential random variate with mean parameter c_α . The values of c_α were found by search as those matching the observed 14% reliable variance, yielding 1.55, 1.52, and 1.40 for the signal eigenspectrum exponents $\alpha = 0.5, 1.0, \text{ and } 1.5$, respectively.

To simulate a combination of additive and multiplicative noise (Extended Data Fig. 5e), responses were modulated by the additive mechanism described above and then modulated multiplicatively. The gain factors were $b_\alpha = 0.55, 0.53, 0.51$ and $c_\alpha = 0.65, 0.64, 0.59$ for $\alpha = 0.5, 1.0, 1.5$ respectively.

Two-photon noise

To investigate whether our two-photon deconvolution method could be biasing the estimated eigenspectrum, we simulated the effect of passing noise through this algorithm (Extended Data Fig. 5f).

To do so, we extended the simulations above to apply in the time domain. When simulating the additive noise, we allowed it to vary across all simulated 2-photon imaging frames (replacing the matrix A used to compute the eigenvectors U and V by a 10000×8400 matrix providing 3 simulated frames per stimulus presentation). The gain modulation factor was assumed equal for all three frames corresponding to a single stimulus. The magnitudes of the additive noise and the gain factor giving 14% signal variance were found by search to be $b_\alpha = 0.50, 0.50, 0.49$ and $c_\alpha = 0.68, 0.67, 0.66$, for $\alpha = 0.5, 1.0, 1.5$ respectively.

To simulate the GCaMP6s response, we convolved these responses with an exponentially decaying kernel with a timescale of 2 seconds (since each timepoint in the data is 0.4 seconds, this corresponds to a decay timescale of 5 timepoints). To simulate shot noise, we added Gaussian white noise with a standard deviation of 0.5. Next we deconvolved these noisy traces using OASIS [40], with a timescale of 5 timepoints and no sparsity constraints. The reduction in signal variance from this procedure was roughly 1%.

For all noise simulations, we estimated the signal eigenspectrum from two repeats using cvPCA. We found that cvPCA, but not ordinary PCA, correctly estimated the ground-true eigenspectrum, for all simulated power-law exponents α (Extended Data Fig. 5g).

Estimation of power-law exponent

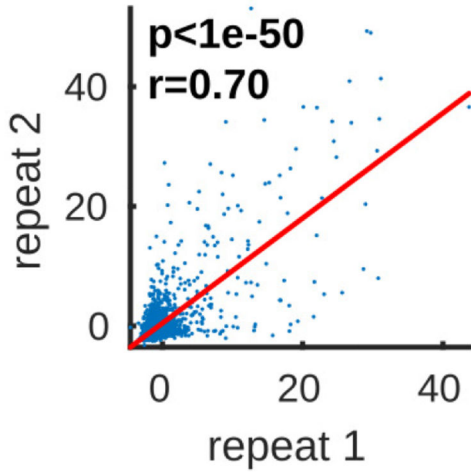
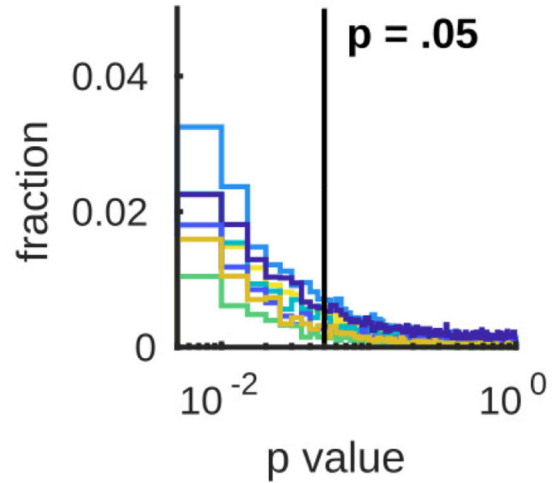
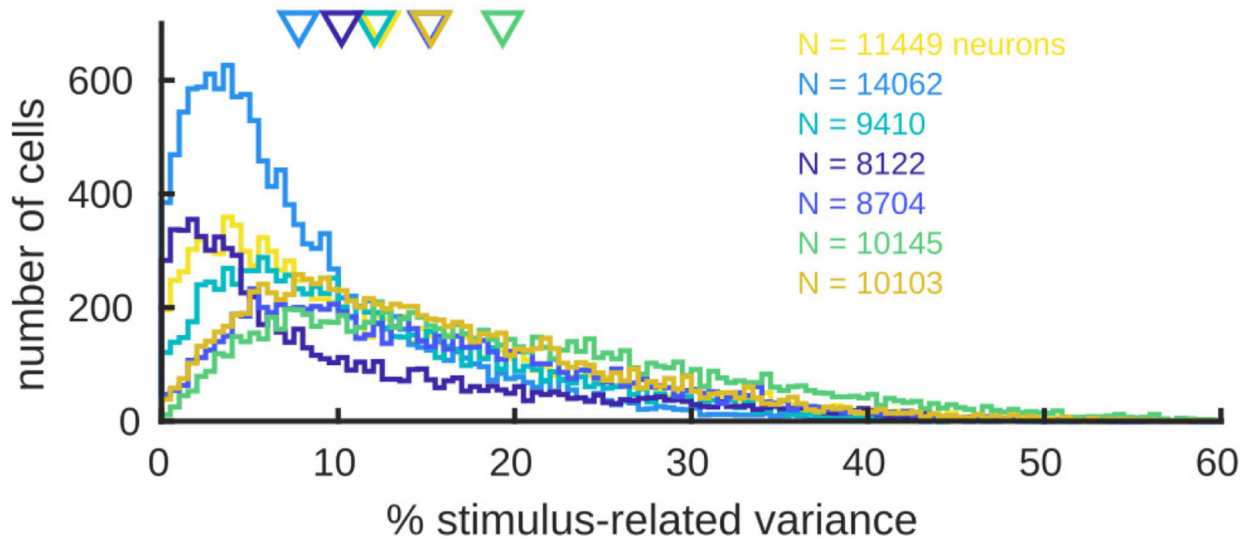
We computed the linear fit of the eigenspectrum over the range of 11 to 500 dimensions for all recordings (and model fits) other than the 32 drifting grating recordings. For the 32-grating recordings, due to noise and the length of the spectrum, we computed the power-law exponent from 5 to 30. The linear fit was performed in log-log space: the range of $\log(11)$ to $\log(500)$ was regressed onto the log of the eigenspectrum.

Sorting neurons and stimuli by correlations

In Extended Data Fig. 6, neurons and stimuli were sorted so they were close to other neurons and stimuli which they were correlated.

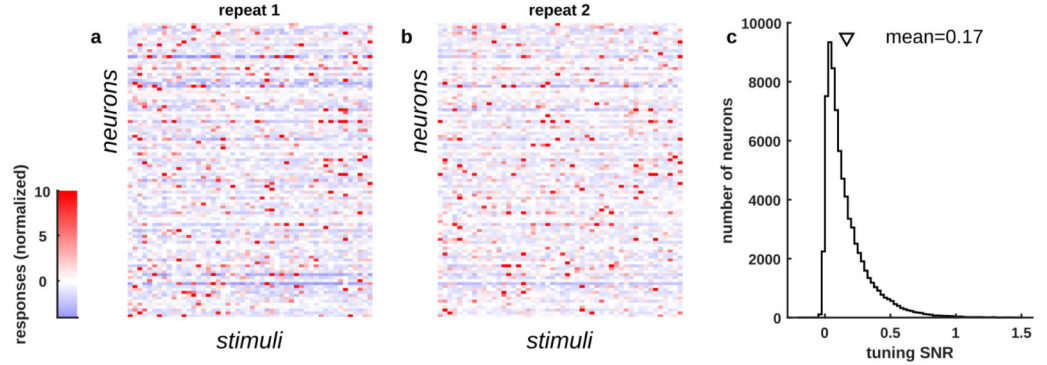
To do this, we first z-scored the binned activity of each neuron and computed principal components of its averaged activity across repeats. Each panel shows this for different PC projections of the data: 1,2, 3-10, 11-40, 41-200, and 201-1000. Stimuli were re-ordered so that each stimulus's pattern of evoked population activity was most similar to the average of its neighbors. The stimulus order was initialized by sorting stimuli according to their weights on the top principal component of activity, then dividing them into 30 clusters of equal size along this ordering. For 50 iterations, we computed the mean activity of each cluster and smoothed this activity across clusters with a Gaussian whose width annealed from 6 clusters to 1 over the first 25 iterations. Each stimulus was then reassigned to the cluster it was most correlated with. On the final pass, we upsampled the stimuli's correlations with each cluster by a factor of 100 via kriging interpolation (smoothing constant of 1 cluster), resulting in a continuous assignment of stimuli along the 1D axis of clustering algorithm. After sorting across stimuli, we smoothed across them to reduce noise, recomputed the principal components on the activity smoothed across stimuli, and repeated the procedure to sort neurons. The algorithm is available in Python and MATLAB at www.github.com/MouseLand/RasterMap. These plots were made using the MATLAB version of the code.

Extended Data

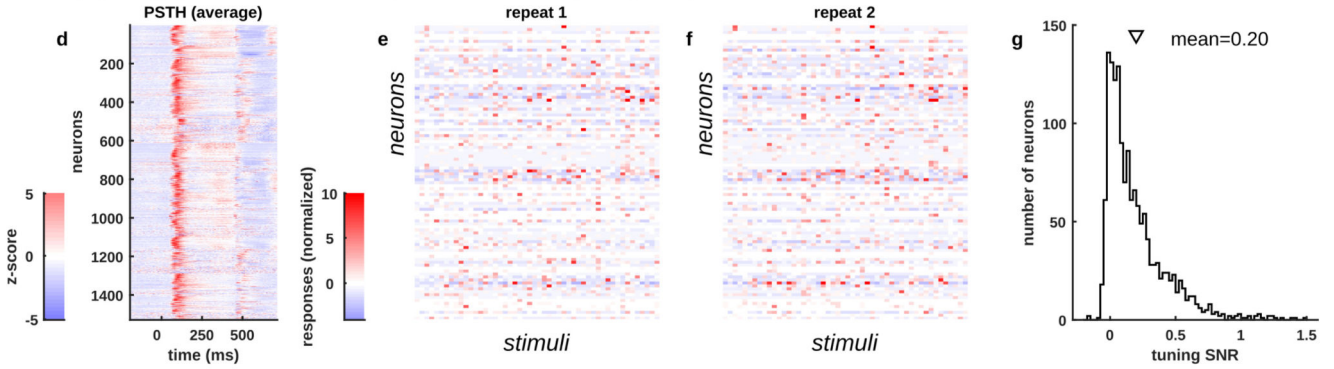
a Responses of one cell**b** Tuning significance**c****Extended Data Fig. 1. Reliability of single neuron responses.**

a, A single neuron's response to the first repeat of 2800 stimuli plotted against its responses to the second repeat of the same stimuli. **b**, Histograms of p-values for Pearson correlation of responses on the two repeats. Each colored histogram represents a different recording. $81.4 \pm 5.1\%$ (SE, $n=7$ recordings) of cells were significant at $p < 0.05$. **c**, Histogram of the single neuron percentage of stimulus-related variance across the population. Each colored histogram represents a different recording; arrowheads (top) represent the mean for each experiment.

Two-photon calcium imaging (74,353 neurons from 7 experiments)

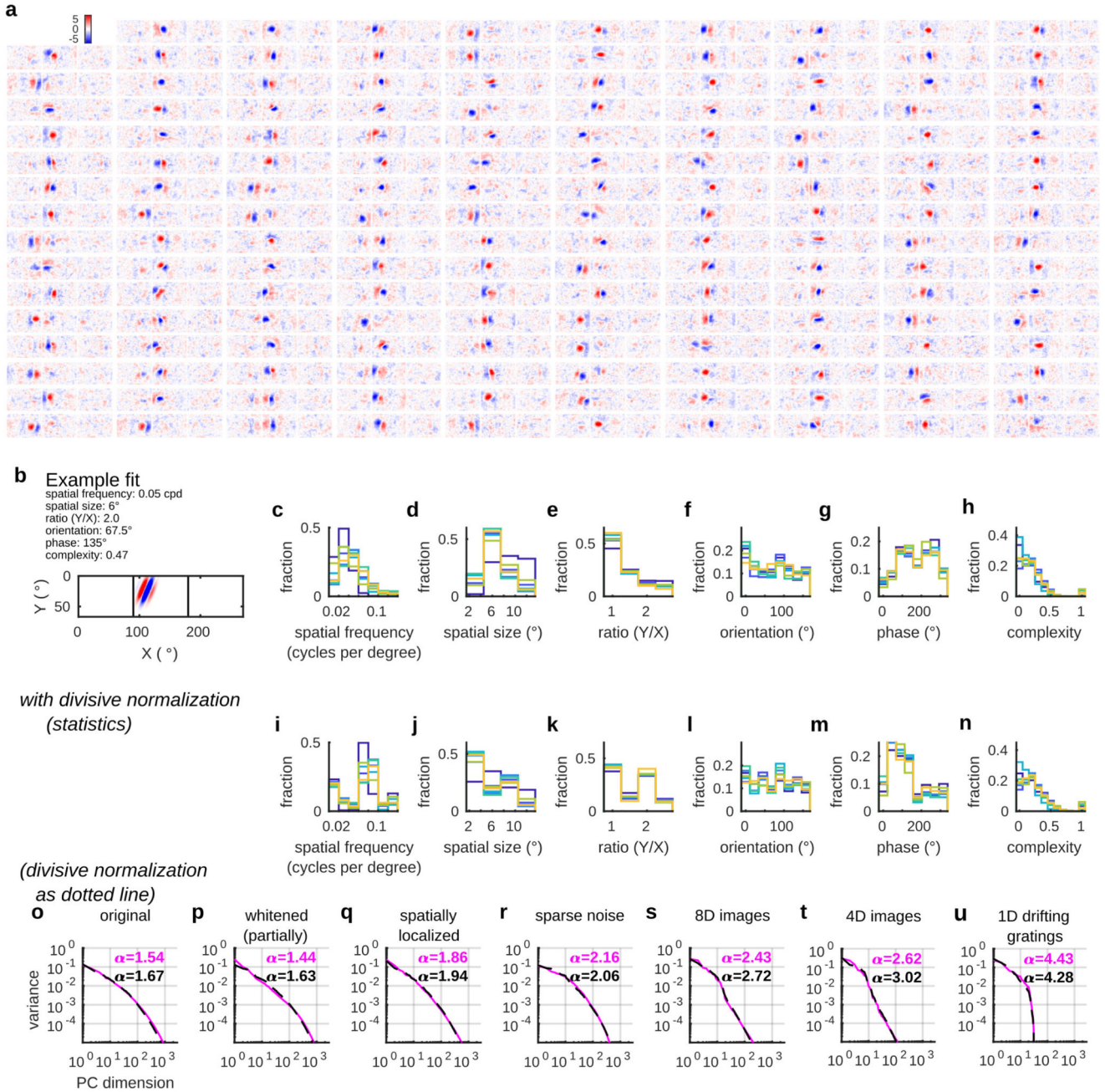


Electrophysiology (1,529 neurons from 6 experiments / 14 Neuropixels probes)



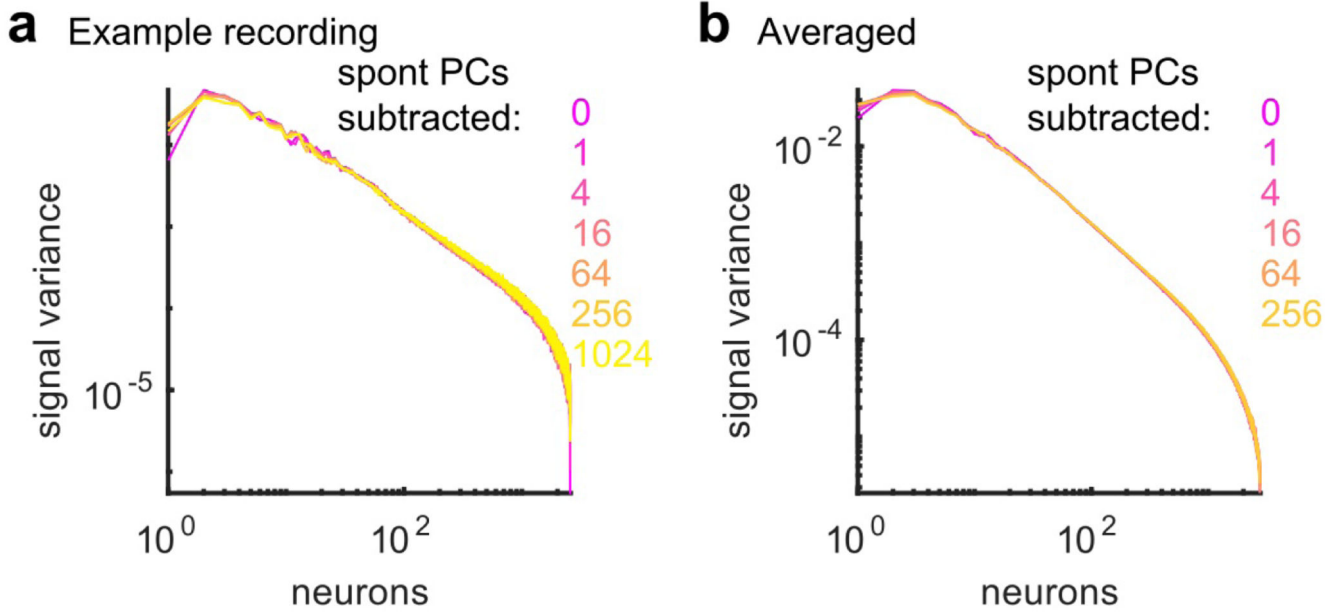
Extended Data Fig. 2. Comparison with electrophysiology.

a,b, Single trial responses of 100 neurons to two repeats of 50 stimuli, recorded by two-photon calcium imaging. **c**, Distribution of tuning SNR for 74,353 neurons recorded by two-photon calcium imaging. **d**, Average peri-stimulus time histogram of spikes recorded electrophysiologically in a separate set of experiments. The images shown were a random subset of 700 images out of the total 2,800. The PSTH reflects the average over all stimuli. The responses are z-scored across time for each neuron. **efg** Same as (abc) for the electrophysiologically recorded neurons.



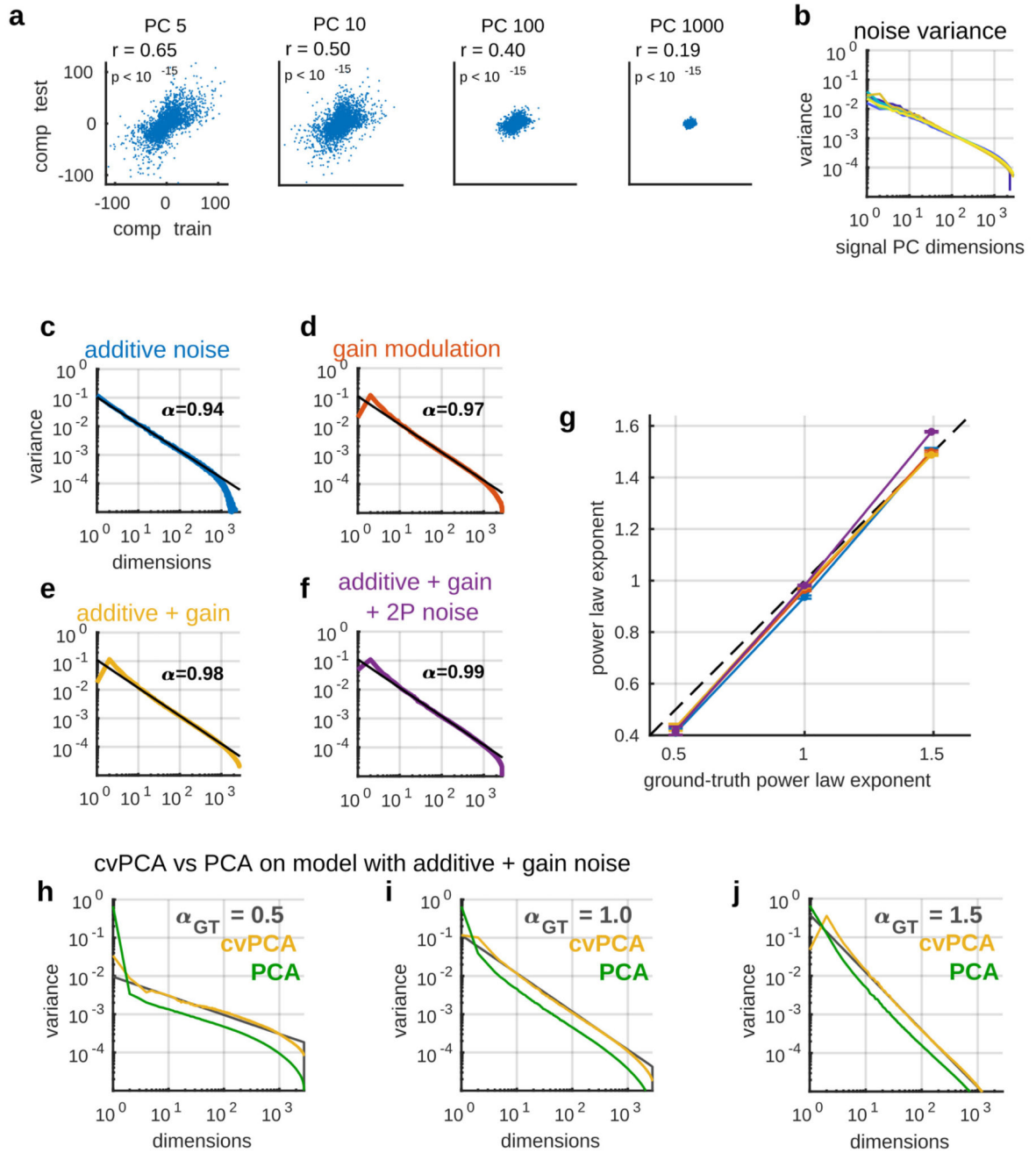
Extended Data Fig. 3. Single neuron receptive fields estimated using reduced-rank regression and Gabor models.

a, 159 randomly chosen neurons' receptive fields estimated using reduced-rank regression. The receptive field map is z-scored for each neuron. **b**, An example Gabor fit to a single cell. **c-h**, Histograms showing the distribution of model parameters across cells. Each color represents cells from one recording. **i-n**, Histograms showing the distribution of model parameters across cells when model also has divisive normalization. **o-u**, Eigenspectra of Gabor population model responses to the different stimulus sets. The unnormalized Gabors are shown in magenta, and the model with divisive normalization in black.



Extended Data Fig. 4. Stimulus-independent activity does not affect the measured eigenspectrum.

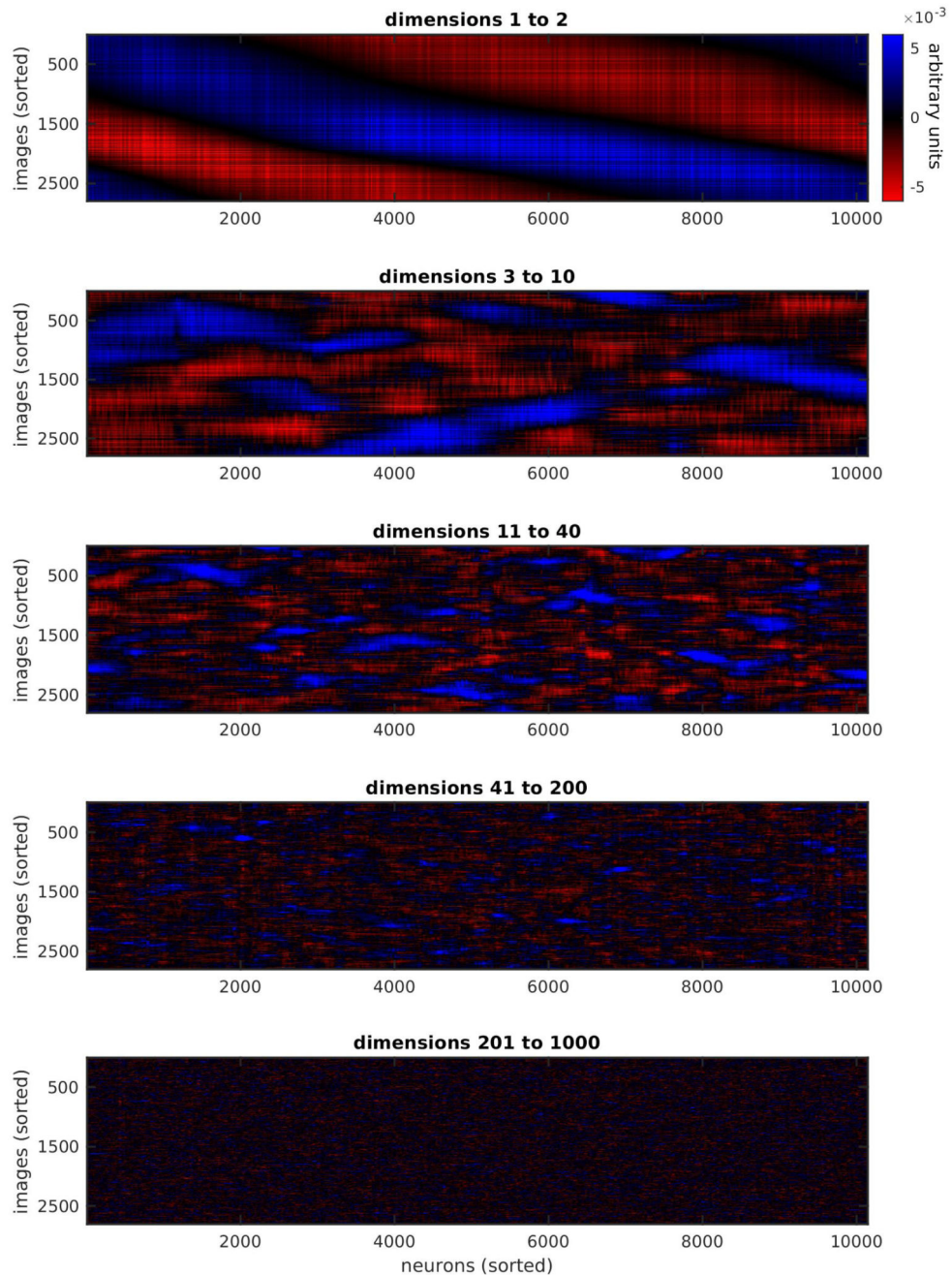
a, To measure the effects of correlated noise variability on eigenspectra estimated by cvPCA, we examined the effect of projecting out different numbers of noise dimensions (estimated during periods of spontaneous gray-screen) from the responses in an example experiment. **b**, Same analysis, averaged over all recordings. The presence of these noise dimensions made little difference to the estimated signal eigenspectrum other than to slightly reduce estimated eigenvalues in the highest and lowest dimensions. For the main analyses, 32 spontaneous dimensions were subtracted.



Extended Data Fig. 5. Validating the eigenspectrum estimation method using simulations with the true noise distribution.

a. Scatterplots illustrating the noise levels of each estimated PC. Each plot shows population activity projected onto the specified principal component, for the 1st repeat (x-axis) and 2nd repeat (y-axis). Each point represents responses a a single stimulus. **b.** Estimated level of noise variance in successive signal dimensions. Noise variance was estimated by subtracting the cvPCA estimate of signal variance from the total variance (see Methods). **c.** Recovery of ground-truth eigenspectrum in simulated data. We simulated responses of 10,000 neurons to

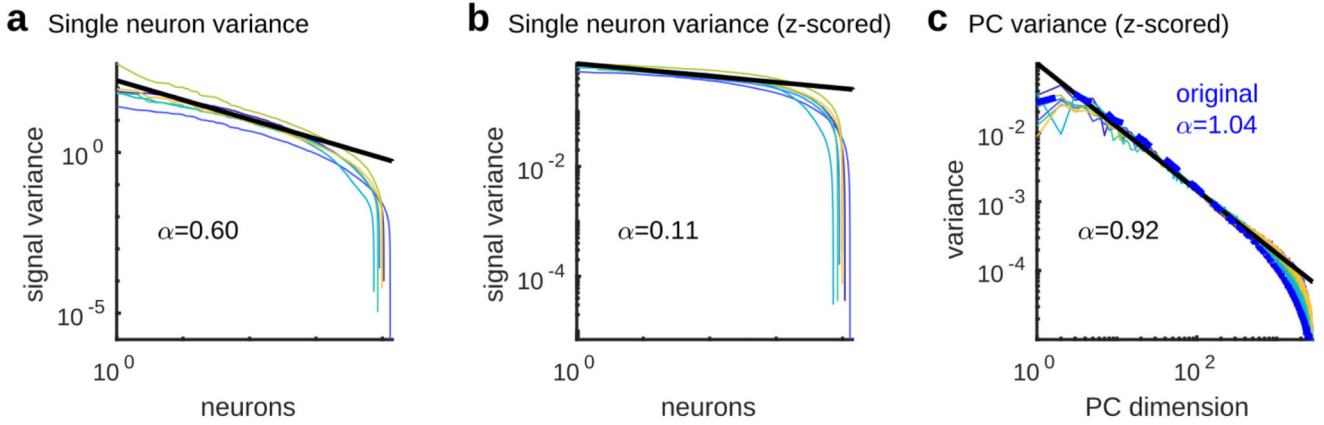
2,800 stimuli with a power spectrum decay of exactly $\alpha = 1$, and added noise in the stimulus space, generated with the spectrum in (b) scaled to produce the same signal-to-noise ratio as in the original neural data. The ground-truth eigenspectrum (black) is estimated accurately by the cvPCA method (blue). **d**, Same analysis with multiplicative noise, in which the responses of all neurons on each trial scaled by a common random factor. The distribution of this factor was again scaled to recover the original signal-to-noise ratio. **e**, Same analysis with a combination of multiplicative and additive noise. **f**, Same analysis, also including simulation of neural and 2-photon shot noise prior to running GCaMP deconvolution algorithm. **g**, 10 instantiations of the simulation were performed with ground-truth exponents of 0.5, 1.0, and 1.5. Error bars represent standard deviations of the power law exponents estimated for each of the 10 simulations. Dashed black line: ground-truth value. **h-j**, Comparison of cvPCA (yellow) and traditional PCA (green) algorithms in the presence of the additive+multiplicative noise combination. While cvPCA recovered the ground truth eigenspectrum (black) exactly, traditional PCA did not, resulting in overestimation of the top eigenvalues and failure to detect the ground-truth power law.



Extended Data Fig. 6. Successive PC dimensions encode finer stimulus features.

Each plot shows the responses of 10,145 neurons to 2,800 natural images, projected onto the specified PCs and then sorted along both axes so correlated neurons and stimuli are close together. We then smoothed the matrix across neurons and stimuli with Gaussian kernels of widths 8 neurons and 2 stimuli respectively. Dimensions 1-2 reveal a coarse, 1-dimensional organization of the neurons and stimuli. Dimensions 3-10 reveal multidimensional structure which involves different neural subpopulations responding to different stimuli. Dimensions 11-40 reveal finer-structured patterns of correlated selectivity among neurons. Dimensions

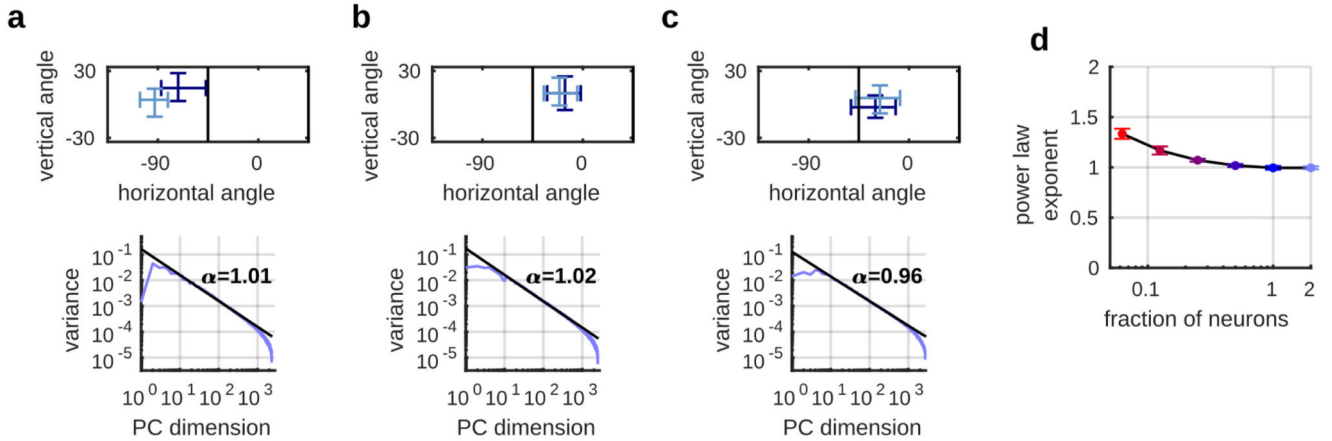
41-200 and 201-1000 reveal even finer-structured selectivity, which contained less neural variance.



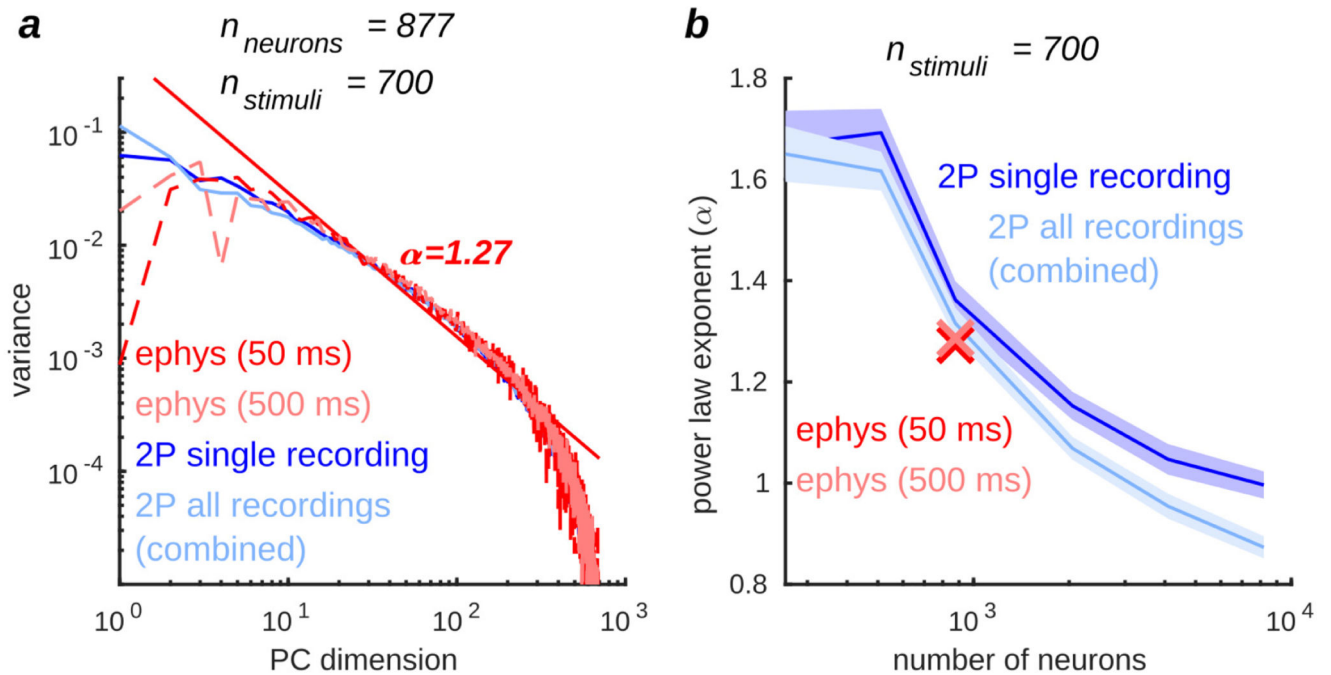
Extended Data Fig. 7. Power law scaling reflects correlation structure, not single-neuron statistics.

a. The signal variance of each neuron's responses are sorted in descending order; they approximately follow a power law with a decay exponent of $\alpha = 0.59$. **b.** Same plot after z-scoring the recorded traces to equalize stimulus response sizes between cells; the distribution of single-neuron variance has become nearly flat. **c.** PC eigenspectra for z-scored data. Each colored line represents a different recording. Dashed blue shows the average eigenspectrum from the original, non-z-scored responses. The fact that the eigenspectrum power-law is barely affected by equalizing firing rates, while the distribution of single cell signal variance is altered, indicates that the power law arises from correlations between cells rather than from the distribution of firing rates or signal variance across cells.

Two recordings with similar RFs concatenated

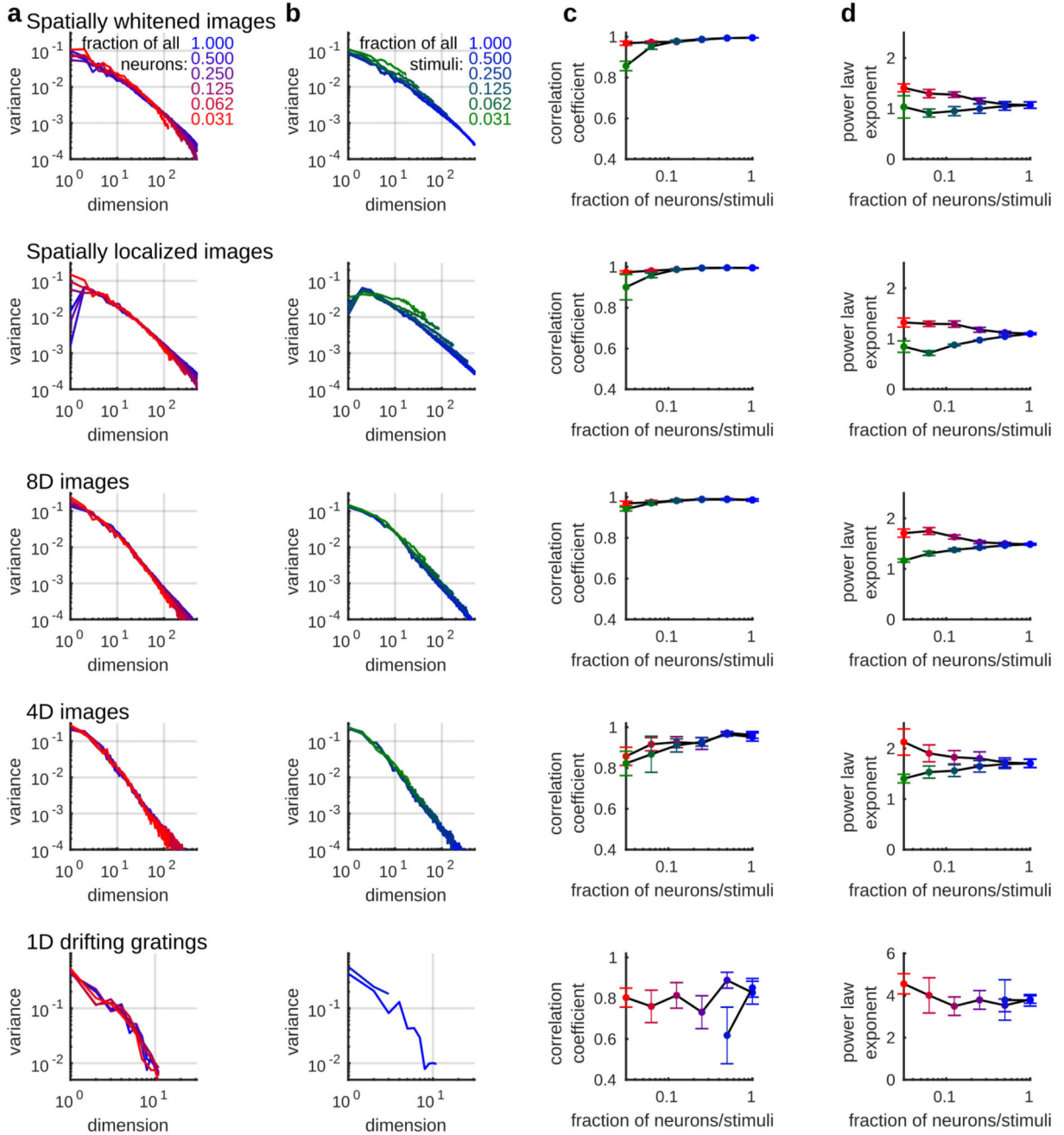
**Extended Data Fig. 8. Power law eigenspectra in concatenated recordings.**

a-c, To ask whether powerlaw eigenspectra apply to even larger populations, we were able to artificially double the number of recorded neurons by combining three pairs of recordings whose imaging fields of view had similar retinotopic locations. Top: retinotopic locations of receptive fields (95% confidence intervals on that recording's mean RF position), with each recording shown in a different shade of blue. Bottom: eigenspectrum of concatenated recordings in response to the 2800 natural image stimuli; total population sizes 19571, 23472, and 18807 cells respectively. Each column represents one pair of recordings. **d**, Eigenspectrum exponents for random subsets of the combined populations (cf. Fig. 2j). X-axis shows population size relative to single recordings, so merged population is 2. Mean power law exponent at "2x" was $\alpha = 0.99 \pm 0.02$ (mean \pm SE).



Extended Data Fig. 9. Eigenspectrum of electrophysiologically recorded data.

We recorded neural activity electrophysiologically in response to 700 out of the 2800 stimuli, and concatenated the recordings, resulting in a total of 877 neurons recorded across 6 experiments. **a**, With this smaller number of stimuli and neurons, convergence to a power law is not complete, and the exponent cannot be estimated accurately (cf. Fig. 2g-j). We therefore compared the electrophysiology data to the responses generated by these stimuli in 877 neurons sampled randomly from either a single two-photon imaging experiment (dark blue) or all experiments combined (light blue). Red and pink show electrophysiology eigenspectra with time bins of 50 or 500ms; red line shows best linear fit to estimate exponent. **b**, Blue curves: power law exponents estimated from the responses of different-sized neuronal subpopulations to this set of 700 stimuli (shading: SE over different random subsets of neurons). Red and pink crosses: estimated exponents from electrophysiology data for 50 and 500 ms bin sizes.



Extended Data Fig. 10. Power law scaling grows more accurate for increasing numbers of neurons and stimuli, for all stimulus ensembles.

a, Eigenspectra estimated from a random subset of the recorded neurons (color-coded by fraction of neurons retained). **b**, Eigenspectra estimated from a random subset of stimuli, color-coded by fraction of stimuli retained. **c**, Correlation coefficient of the spectra plotted in **a,b**. **d**, Power law exponent of the spectra plotted in **a,b**. Each row corresponds to a different ensemble of visual stimuli.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Michael Krumin for assistance with the two-photon microscopes, Charu Bai Reddy for surgeries, and Kenneth Falconer, Arthur Gretton, and Marcus Benna for discussions of mathematics.

This research was funded by Wellcome Trust Investigator grants (108726, 205093, 204915) and by a grant from the Simons Foundation (SCGB 325512). CS was funded by a four-year Gatsby Foundation PhD studentship. MC holds the GlaxoSmithKline / Fight for Sight Chair in Visual Neuroscience. KDH was funded by the European Research Council (694401). NS was supported by post-doctoral fellowships from the Human Frontier Sciences Program and the Marie Curie Action of the EU (656528). CS and MP are now funded by HHMI Janelia.

References

- [1]. Barlow, HB. Possible principles underlying the transformations of sensory messages. *Sensory Communication*. Rosenblith, W, editor. MIT Press; Cambridge, MA: 1961. 217–234.
- [2]. Atick JJ, Redlich AN. Towards a theory of early visual processing. *Neural Computation*. 1990; 2:308–320.
- [3]. Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. *Annual Review of Neuroscience*. 2001; 24:1193–1216.
- [4]. DiCarlo JJ, Zoccolan D, Rust NC. How does the brain solve visual object recognition? *Neuron*. 2012; 73:415–434. [PubMed: 22325196]
- [5]. Rigotti M, et al. The importance of mixed selectivity in complex cognitive tasks. *Nature*. 2013; 497:585. [PubMed: 23685452]
- [6]. Chung S, Lee DD, Sompolinsky H. Classification and geometry of general perceptual manifolds. *Phys Rev X*. 2018; 8
- [7]. Cunningham JP, Yu BM. Dimensionality reduction for large-scale neural recordings. *Nature Neuroscience*. 2014; 17:1500. [PubMed: 25151264]
- [8]. Machens CK, Romo R, Brody CD. Functional, but not anatomical, separation of “what” and “when” in prefrontal cortex. *Journal of Neuroscience*. 2010; 30:350–360. [PubMed: 20053916]
- [9]. Kobak D, et al. Demixed principal component analysis of neural population data. *Elife*. 2016; 5:e10989. [PubMed: 27067378]
- [10]. Archer, EW, Koster, U, Pillow, JW, Macke, JH. Low-dimensional models of neural population activity in sensory cortical circuits. *Advances in Neural Information Processing Systems*. Ghahramani, Z, Welling, M, Cortes, C, Lawrence, ND, Weinberger, KQ, editors. Vol. 27. Curran Associates, Inc; 2014. 343–351.
- [11]. Sadtler PT, et al. Neural constraints on learning. *Nature*. 2014; 512:423. [PubMed: 25164754]
- [12]. Chapin JK, Nicolelis MA. Principal component analysis of neuronal ensemble activity reveals multidimensional somatosensory representations. *Journal of Neuroscience Methods*. 1999; 94:121–140. [PubMed: 10638820]
- [13]. Bathellier B, Buhl DL, Accolla R, Carleton A. Dynamic ensemble odor coding in the mammalian olfactory bulb: Sensory information at different timescales. *Neuron*. 2008; 57:586–598. [PubMed: 18304487]
- [14]. Churchland MM, et al. Neural population dynamics during reaching. *Nature*. 2012; 487:51. [PubMed: 22722855]
- [15]. Mante V, Sussillo D, Shenoy KV, Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*. 2013; 503:78–84. [PubMed: 24201281]
- [16]. Shadlen MN, Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*. 1998; 18:3870–3896. [PubMed: 9570816]

- [17]. Reich DS, Mechler F, Victor JD. Independent and redundant information in nearby cortical neurons. *Science*. 2001; 294:2566–2568. [PubMed: 11752580]
- [18]. Gao, P; , et al. A theory of multineuronal dimensionality, dynamics and measurement. *bioRxiv*. 2017. URL <https://www.biorxiv.org/content/early/2017/11/12/214262>
- [19]. Pachitariu, M; , et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv*. 2016. URL <https://www.biorxiv.org/content/early/2017/07/20/061507>
- [20]. Deng, J; , et al. Imagenet: A large-scale hierarchical image database. *IEEE Conference on Computer Vision and Pattern Recognition; IEEE*; 2009. 248–255.
- [21]. Vinje WE, Gallant JL. Natural stimulation of the nonclassical receptive field increases information transmission efficiency in v1. *Journal of Neuroscience*. 2002; 22:2904–2915. [PubMed: 11923455]
- [22]. Ringach DL. Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. *Journal of Neurophysiology*. 2002; 88:455–463. [PubMed: 12091567]
- [23]. Niell CM, Stryker MP. Highly selective receptive fields in mouse visual cortex. *Journal of Neuroscience*. 2008; 28:7520–7536. [PubMed: 18650330]
- [24]. Softky WR, Koch C. The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps. *Journal of Neuroscience*. 1993; 13:334–350. [PubMed: 8423479]
- [25]. Cossell L, et al. Functional organization of excitatory synaptic strength in primary visual cortex. *Nature*. 2015
- [26]. Smyth D, Willmore B, Baker GE, Thompson ID, Tolhurst DJ. The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *Journal of Neuroscience*. 2003; 23:4746–4759. [PubMed: 12805314]
- [27]. Carandini M, et al. Do we know what the early visual system does? *Journal of Neuroscience*. 2005; 25:10577–10597. [PubMed: 16291931]
- [28]. David SV, Gallant JL. Predicting neuronal responses during natural vision. *Network: Computation in Neural Systems*. 2005; 16:239–260.
- [29]. Touryan J, Felsen G, Dan Y. Spatial structure of complex cell receptive fields measured with natural images. *Neuron*. 2005; 45:781–791. [PubMed: 15748852]
- [30]. de Vries, SEJ; , et al. A large-scale, standardized physiological survey reveals higher order coding throughout the mouse visual cortex. *bioRxiv*. 2018. URL <https://www.biorxiv.org/content/early/2018/06/29/359513>
- [31]. Field DJ. Relations between the statistics of natural images and the response properties of cortical cells. *JOSA A*. 1987; 4:2379–2394.
- [32]. Ruderman DL, Bialek W. Statistics of natural images: Scaling in the woods. *Advances in Neural Information Processing Systems*. 1994:551–558.
- [33]. Tao, T. *An Epsilon of Room, I: Real Analysis*. American Mathematical Society; Providence, RI: 2010.
- [34]. Szegedy, C; , et al. Intriguing properties of neural networks. *arXiv abs/1312.6199*. 2013. URL <http://arxiv.org/abs/1312.6199>
- [35]. Goodfellow, IJ; Shlens, J; Szegedy, C. Explaining and harnessing adversarial examples. *arXiv abs/1412.6572*. 2014. URL <http://arxiv.org/abs/1412.6572>
- [36]. Stringer C, Pachitariu M, Carandini M, Harris K. Responses of ten thousand neurons to 2,800 natural images. *Figshare*. 2018; doi: 10.25378/janelia.6845348.v3
- [37]. Stringer C, et al. Spontaneous behaviors drive multidimensional, brainwide activity. *Science*. 2019; 364:255–255. [PubMed: 31000656]
- [38]. Pologruto TA, Sabatini BL, Svoboda K. Scanimage: flexible software for operating laser scanning microscopes. *Biomedical engineering online*. 2003; 2:13. [PubMed: 12801419]
- [39]. Chen T-W, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature*. 2013; 499:295–300. [PubMed: 23868258]
- [40]. Friedrich J, Zhou P, Paninski L. Fast online deconvolution of calcium imaging data. *PLOS Computational Biology*. 2017; 13:e1005423. [PubMed: 28291787]
- [41]. Pachitariu M, Stringer C, Harris KD. Robustness of spike deconvolution for neuronal calcium imaging. *Journal of Neuroscience*. 2018; 38:7976–7985. [PubMed: 30082416]

- [42]. Jun JJ, et al. Fully integrated silicon probes for high-density recording of neural activity. *Nature*. 2017; 551:232. [PubMed: 29120427]
- [43]. Pachitariu M, Steinmetz NA, Kadir SN, Carandini M, Harris KD. Fast and accurate spike sorting of high-channel count probes with kilosort. *Advances in Neural Information Processing Systems*. 2016:4448–4456.
- [44]. Allen WE, et al. Thirst regulates motivated behavior through modulation of brainwide neural population dynamics. *Science*. 2019; 364
- [45]. Izenman AJ. Reduced-rank regression for the multivariate linear model. *Journal of Multivariate Analysis*. 1975; 5:248–264.
- [46]. Schwartz O, Simoncelli EP. Natural signal statistics and sensory gain control. *Nature neuroscience*. 2001; 4:819. [PubMed: 11477428]
- [47]. Cowley BR, Smith MA, Kohn A, Byron MY. Stimulus-driven population activity patterns in macaque primary visual cortex. *PLoS computational biology*. 2016; 12:e1005185. [PubMed: 27935935]
- [48]. Gallego JA, et al. Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nature communications*. 2018; 9:4233.

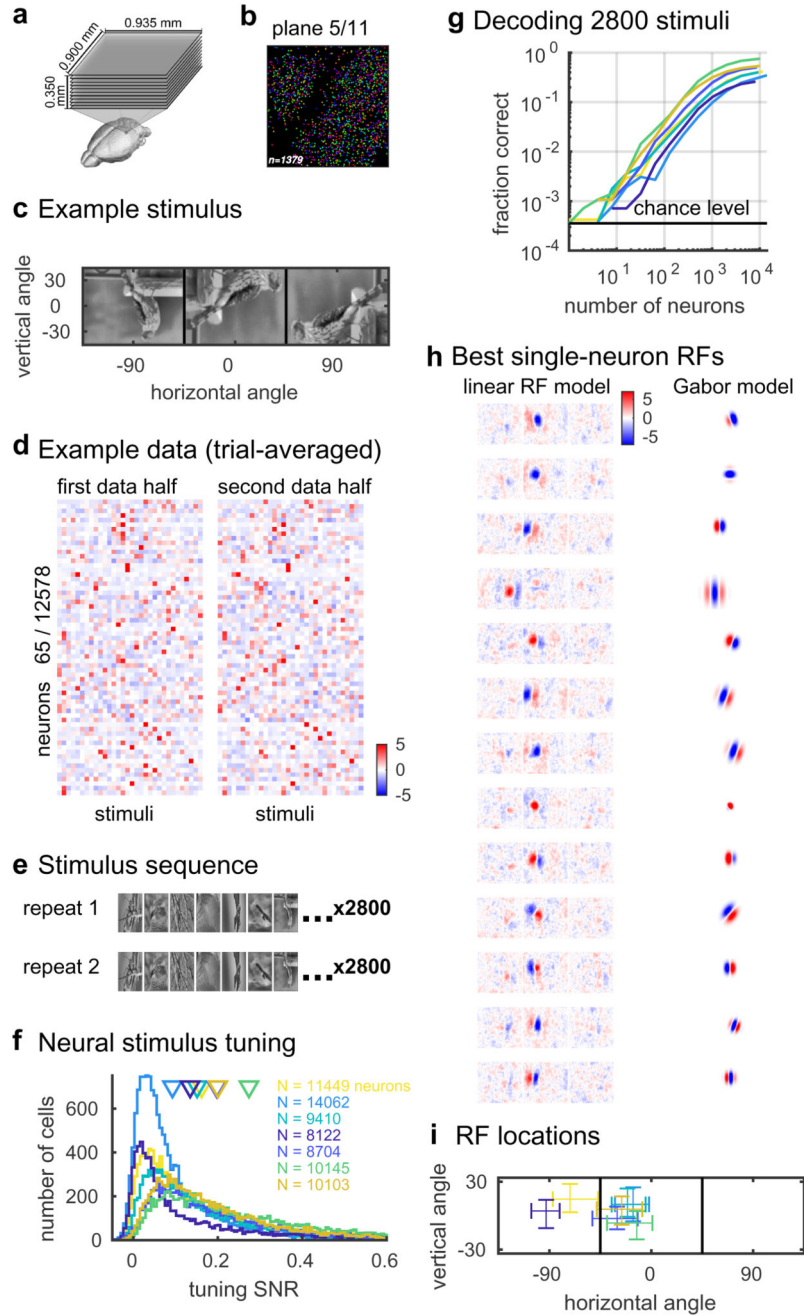


Figure 1. Population coding of visual stimuli.
a, Simultaneous recording of ~10,000 neurons using 11-plane two-photon calcium imaging. **b**, Randomly-pseudocolored cells in an example imaging plane. **c**, Example stimulus spans three screens surrounding the mouse’s head. **d**, Mean responses of 65 randomly-chosen neurons to 32 image stimuli (96 repeats, z-scored, scale bar represents standard deviations, one recording out of four shown). **e**, A sequence of 2800 stimuli was repeated twice during the recording. **f**, Distribution of single-cell signal-to-noise ratios (SNR) (2800 stimuli, two repeats). Colors denote recordings; arrows represent means. **g**, Stimulus decoding accuracy

as a function of neuron count for each recording. **h**, Example receptive fields (RFs) fit using reduced-rank regression or Gabor models (z-scored) (one recording shown, out of 7). **i**, Distribution of the receptive field centers, plotted on the left and center screens (line denotes screen boundary). Each cross represents a different recording, with 95% of neuron's RF centers within error bars.

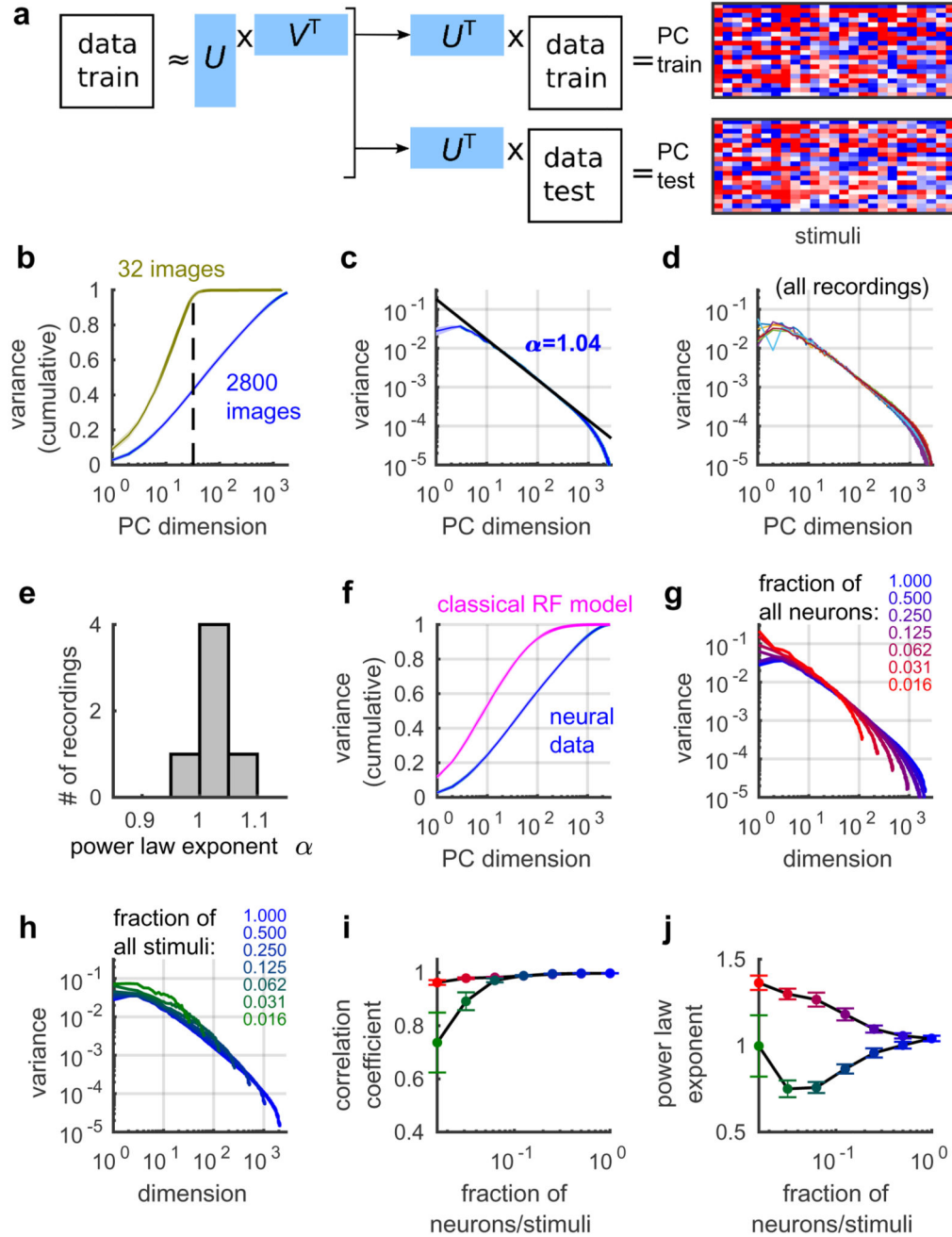


Figure 2. Visual cortical responses are high-dimensional with power-law eigenspectrum.
a. The eigenspectrum of visual stimulus responses was estimated by cross-validated principal component analysis (cvPCA), projecting singular vectors from the first repeat onto responses from the second. **b.** Cumulative fraction of variance in planes of increasing dimension, for an ensemble of 2800 stimuli (blue) and for 96 repeats of 32 stimuli. Dashed line indicates 32 dimensions. **c.** Eigenspectrum plotted in descending order of training set singular value for each dimension, averaged across 7 recordings (shaded error bars represent standard error). Black line denotes linear fit of $1/n^\alpha$. **d.** Eigenspectra of each recording

individually. **e**, Histogram of power law exponents α across all recordings. **f**, Cumulative eigenspectrum for a simple/complex Gabor model fit to the data (pink) superimposed on true data (blue). **g**, Eigenspectra computed from random subsets of recorded neurons, fraction indicated by colors. **h**, Same analysis for random subsets of stimuli. **i**, Pearson correlation of log variance and log dimension over dims 11-500, as a function of fraction analyzed (1 indicates a power law). **j**, Power law exponents of the spectra plotted in **g,h**.

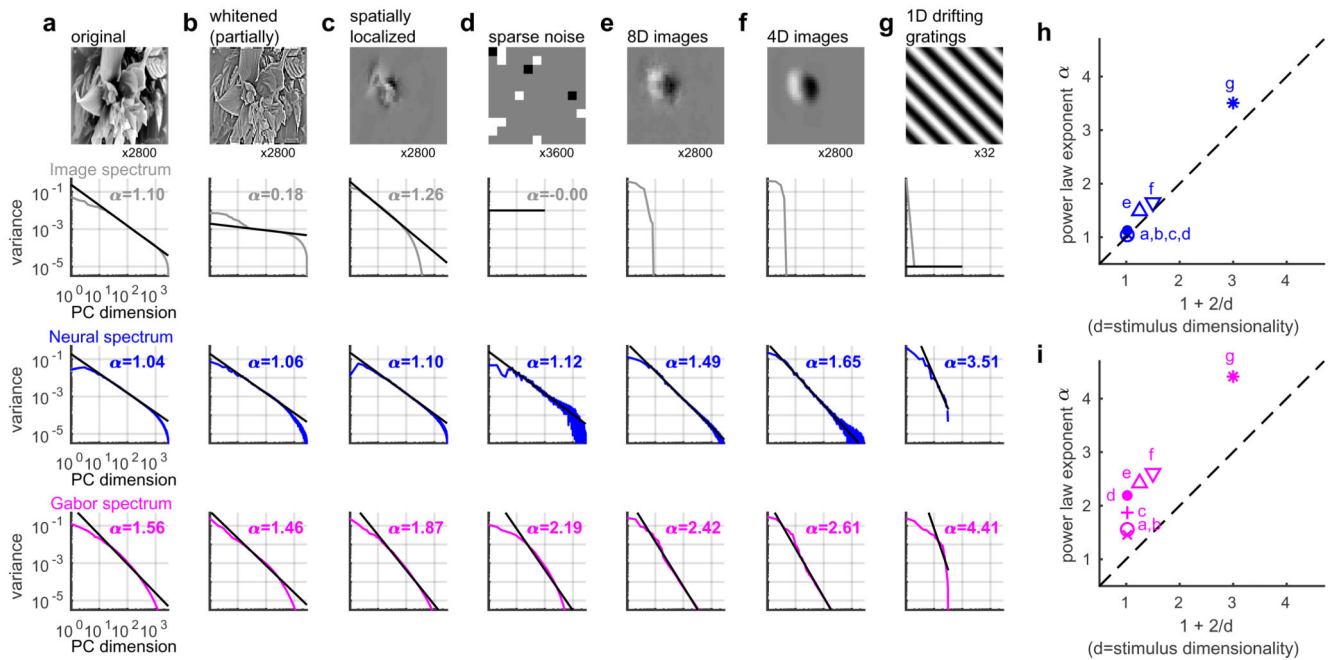


Figure 3. Power law exponent depends on input dimensionality, but not image statistics.
a, Eigenspectra of natural image pixel intensities (gray), of visual cortical responses to these stimuli (blue), and of a simple/complex cell model's response to these stimuli (pink). **b**, Same analysis for responses to spatially whitened stimuli, that lack $1/n$ image spectrum. **c**, Same analysis for images windowed over the RF of the recorded population. **d**, Same analysis for sparse noise stimuli. **e**, Same analysis for images projected into 8 dimensions produces a faster eigenspectrum decay with exponent $\alpha=1.49$. **f**, After projecting images to 4 dimensions, $\alpha=1.65$. **g**, Responses to drifting gratings, a one dimensional stimulus ensemble, show yet faster decay with $\alpha=3.43$. **h,i**, Summary of power law exponents α for neural responses (h) and Gabor model (i), as a function of the dimensionality of the stimulus set d . Dashed line: $\alpha = 1 + 2/d$, corresponding to the border of fractality.

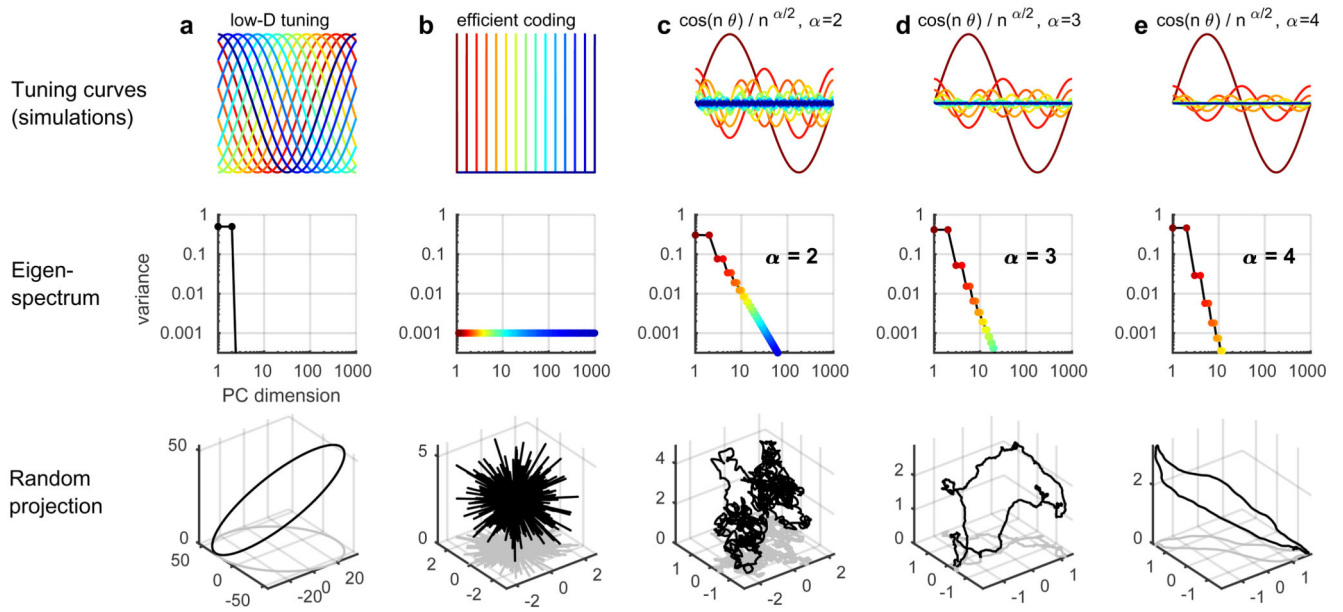


Figure 4. The smoothness of simulated neural activity depends on the eigenspectrum decay. Simulations of neuronal population responses to a 1-dimensional stimulus (x-axis), their eigenspectra, and a random projection of responses in 3D space. **a**, Wide tuning curves, corresponding to a circular neural manifold in a 2-dimensional plane. **b**, Narrow tuning curves corresponding to uncorrelated responses as predicted by the efficient coding hypothesis. **c-e**, Scale-free tuning curves corresponding to power law variance spectra, with exponents of 2, 3 (the critical value for $d=1$), or 4. Tuning curves in c-e represent PC dimensions rather than individual simulated neurons.