Encoding manifolds constructed from grating responses organize responses to natural scenes in cortical visual areas

Luciano Dyballa[1,4], Greg D. Field[2,5], Michael P. Stryker[3,6], and Steven W. Zucker[1,7]

[1]Department of Computer Science, Yale University, New Haven, USA
[2]Department of Neurobiology, Duke University, Durham, NC, USA
[3]Department of Physiology, University of California, San Francisco, CA, USA
[4]School of Science and Technology, IE University, Madrid, Spain
[5]Stein Eye Institute, Department of Ophthalmology, David Geffen School of Medicine, University of California, Los Angeles, CA, USA
[6]Kavli Institute for Fundamental Neuroscience, University of California, San Francisco, CA, USA
[7]Department of Biomedical Engineering, Yale University, New Haven, CT, USA

Abstract

We have created "encoding manifolds" to reveal the overall responses of a brain area to a variety of stimuli. Encoding manifolds organize response properties globally: each point on an encoding manifold is a neuron, and nearby neurons respond similarly to the stimulus ensemble in time. We previously found, using a large stimulus ensemble including optic flows, that encoding manifolds for the retina were highly clustered, with each cluster corresponding to a different ganglion cell type. In contrast, the topology of the V1 manifold was continuous. Now, using responses of individual neurons from the Allen Institute Visual Coding–Neuropixels dataset in the mouse, we infer encoding manifolds for V1 and for five higher cortical visual areas (VISam, VISal, VISpm, VISlm, and VISrl). We show here that the encoding manifold topology computed only from responses to various grating stimuli is also continuous, not only for V1 but also for the higher visual areas, with smooth coordinates spanning it that include, among others, orientation selectivity and firing-rate magnitude. Surprisingly, the encoding manifold for gratings also provides information about natural scene responses. To investigate whether neurons respond more strongly to gratings or natural scenes, we plot the log ratio of natural scene responses to grating responses (mean firing rates) on the encoding manifold. This reveals a global coordinate axis organizing neurons' preferences between these two stimuli. This coordinate is orthogonal (i.e., uncorrelated) to that organizing firing rate magnitudes in VISp. Analyzing layer responses, a preference for gratings is concentrated in layer 6, whereas preference for natural scenes tends to be higher in layers 2/3 and 4. We also find that preference for natural scenes dominates the responses of neurons that prefer low (0.02 cpd) and high (0.32 cpd) spatial frequencies, rather than intermediate ones (0.04 to 0.16 cpd). Conclusion: while gratings seem limited and natural scenes unconstrained, machine learning algorithms can reveal subtle relationships between them beyond linear techniques.

Research supported by NIH Grant EY031059, NSF CRCNS Grant 1822598, the Swartz Foundation (LD), the RPB Disney Award for Amblyopia Research (MPS), and an unrestricted fund to the The Jules Stein Eye Institute from Research to Prevent Blindness and P30 - EY000331 (GDF). We thank the Allen Institute for the use of their data and Cris Neill for discussions.

1 Introduction

Understanding how populations of neurons are organized to represent external stimuli is a core goal of sensory neuroscience. This is made challenging by the fact that both artificial and sensory stimuli are typically high dimensional as are the patterns of spiking activity that these stimuli produce across populations of neurons. This presents a challenge to visualizing, let alone understanding, how complex stimuli are represented by neural population activity. To overcome this challenge, we have recently developed *encoding manifolds* to organize neurons according to their responses to an ensemble of visual stimuli [25]; these manifolds allow visualizing how neural populations encode ensembles of stimuli. In the encoding manifold, each point is a neuron, and nearby neurons respond similarly in time to members of the stimulus ensemble. The encoding manifold differs from more standard applications of manifolds in neuroscience, in which each point on the manifold is a stimulus (or a point along a trajectory) embedded in neural response coordinates (see, e.g., [16] for a review). This distinction is relevant as the latter approach emphasizes reading out the stimulus (or movement trajectory) from the population response (i.e., decoding), while the former organizes neurons according to their functional properties (i.e., encoding), which can drive hypotheses about cell types and their anatomical and functional connectivity in a circuit.

One feature of encoding manifolds exploited in prior work [25] is they allow for comparing the topology of sensory encoding at different stages of sensory processing. When comparing the encoding manifolds between retina and V1 in mouse to an ensemble of artificial stimuli (drifting gratings) and more naturalistic stimuli (flows), they were strikingly different: the encoding manifold inferred from retinal responses consisted of largely separate clusters of neurons that corresponded to known retinal ganglion cell types. This result served to lend credence to the encoding manifold as a useful approach for revealing underlying biological structure because the stimuli used to generate the manifold were distinct from those used to classify the retinal ganglion cell types. The V1 manifold derived from responses to the same stimuli was much more continuous than the retinal manifold, with neurons effectively carpeting stimulus/response space. Like the retinal manifold, the structure of the V1 manifold was quite informative about the physiological organization of V1. While no information about laminar position or putative excitatory versus putative inhibitory neurons (inferred from spike waveforms) was used in its construction, these properties were not randomly located across the manifold; rather, functional cell types were distributed in a complementary fashion across different layers and in different regions.

This prior work raised several important questions: to what extent does the structure of the V1 encoding manifold depend on the stimulus set? To what extent are the encoding manifolds of higher visual areas similar to that observed in V1? Finally, can the encoding manifold reveal aspects of how natural scenes are encoded by visual cortex? Fortunately, these three questions could be answered by utilizing the Allen Institute's large data set of visual cortical responses of V1 and higher cortical visual areas to a variety of stimuli [62, 20, 21]. The stimulus set includes stationary gratings of a wide range of orientations, spatial frequencies, and spatial phases, a set of a similar number of stationary natural images, and an ensemble of moving gratings across a range of spatial frequencies and orientations (Fig. 1a,b). Using these data, we have constructed encoding manifolds for each of the higher visual areas using only the responses to static and drifting gratings. The resulting V1 encoding manifold was similar to that in our earlier work despite substantive differences in the stimulus sets. Namely, both the present and previous manifolds were continuous, with foci for different cortical layers and excitatory versus inhibitory cell types. The higher visual areas also exhibited similar manifold structures. One major similarity among the manifolds generated from each visual area was a preference for low spatial frequency gratings in layer 6, a layer that projects back to sub-cortical structures and elsewhere [54]; cells in all other layers, across all visual areas, tended to prefer natural scenes.

Further analysis of the natural scene responses using the encoding manifold revealed a particularly interesting geometry. The resulting manifolds had a coordinate that ran from cells with strong to weak

orientation and stimulus selectivity. However, those cells that responded better to natural images relative to gratings formed a band in the manifold orthogonal to the orientation axis in VISp, which suggests that there exist natural-scene preferring cells over the full range of selectivity for stimulus orientation. Furthermore, these cells preferred either high or low rather than the intermediate spatial frequencies that were preferred by the cells driven best by gratings. This structure was also present in each of the higher visual areas, although the axes differed from VISp. These results further validate the use of encoding manifolds for assaying the structure and topology of sensory coding across distinct stimulus ensembles and stages of sensory processing.

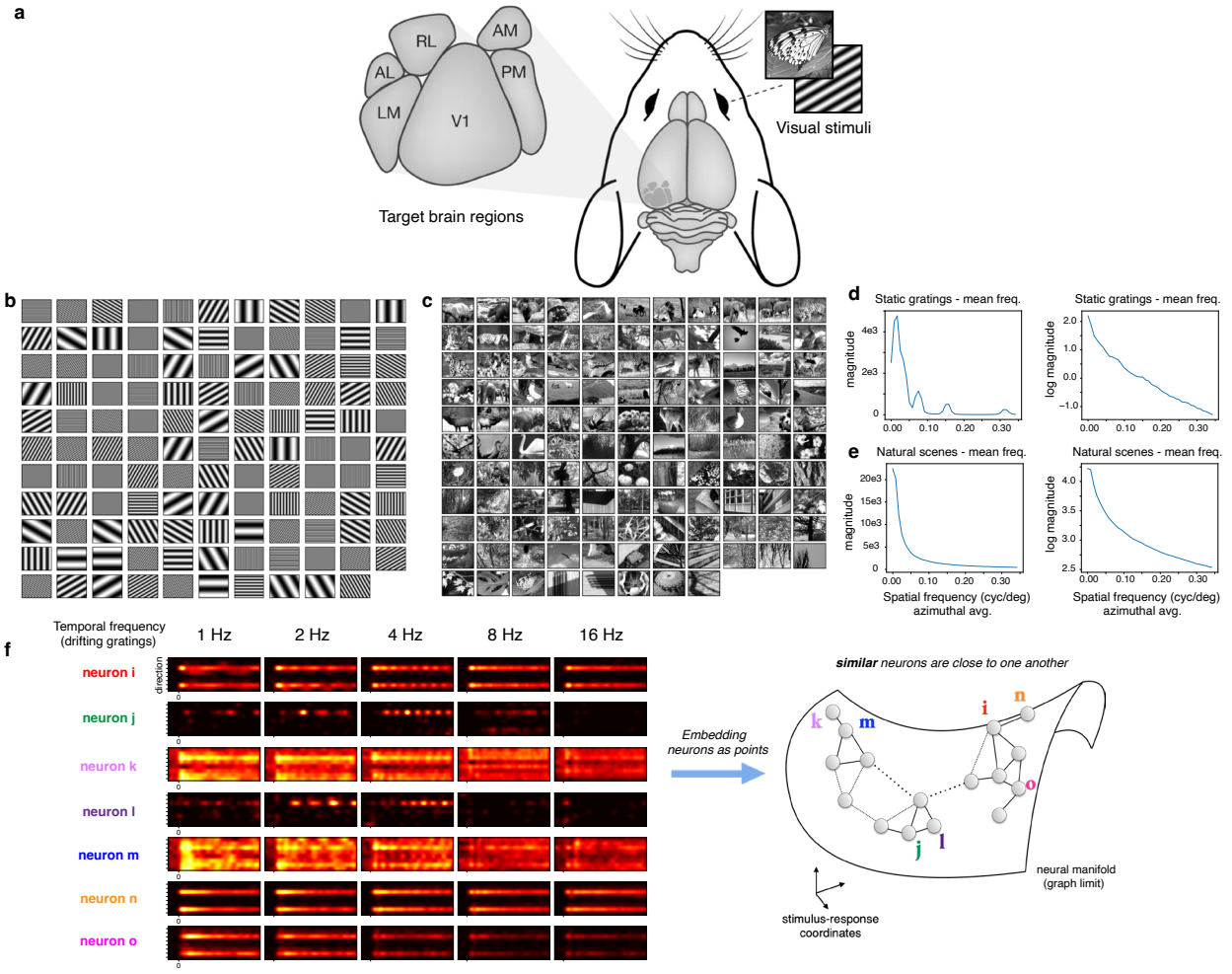

Fig. 1: **Experimental setup, stimulus ensemble, and method.** **a**, The high-order visual areas recorded at the Allen Institute surround V1 in the mouse. The stimuli used include drifting gratings, flashed static gratings, and flashed natural scenes. **b**, The ensemble of static spatial frequency gratings and **c** natural scenes from the Allen dataset. Note that the spatial frequency contents of the natural scenes and flashed gratings are similar (**d**, **e**). **f** The encoding manifold organizes neurons so that those nearby (on the manifold) respond similarly in time. It is inferred using a kernel that evaluates similarity across the PSTH responses to an ensemble of spatial frequency gratings. Neurons i and k exhibit similar activity patterns, and are close, while neurons i and j are different. This is illustrated by the underlying data graph (see Methods) superimposed on the manifold. Part **a** modified from [21]; part **f** after [25].

2 Results

We began by determining the extent to which the fundamental structure of the encoding manifold of mouse primary visual cortex (VISp) generalizes from our prior visual stimulus set to that utilized by the Allen Institute [5]. In previous work, we analyzed responses to drifting gratings and naturalistic 'flow' stimuli, composed of points and line segments moving semi-coherently [24, 25]. However, the previous data were limited to VISp and did not thoroughly examine how the structure of the encoding manifold depended on the temporal frequency composition of the stimulus. The Allen Institute data set has several features that allow a useful comparison (Fig. 1). First, it consists of measurements not only from VISp but also from five higher cortical visual areas in mouse: VISam, VISal, VISpm, VISlm, VISrl (Fig. 1). Second, the stimulus set contains drifting gratings presented at five temporal frequencies. Third, it contains static gratings presented at multiple spatial frequencies (five), orientations (six), and phases (four) (Fig. 1b). Finally, the Allen Institute dataset contains static images of natural scenes (Fig. 1c). Thus, we used the Allen Institute dataset to generate and analyze encoding manifolds along the mouse cortical visual hierarchy.

2.1 The encoding manifold reveals novel relationships between stimulus selectivity and neuronal subtypes

Encoding manifolds were calculated from the Allen Institute data using the responses to both drifting and static gratings [25]. In brief, the encoding manifold arises from a machine learning algorithm that learns to place individual neurons in a relatively low-dimensional space according to their similarity in stimulus selectivity and response dynamics: neurons nearby on the manifold exhibit similar responses to similar stimuli. Responses to the natural scene stimuli were not included in the calculations of the encoding manifolds.

Consistent with prior results [25], the encoding manifold of VISp was relatively smooth and continuous, with neurons uniformly carpeting stimulus space (Fig 2a). Note that this topology is quite distinct from the manifold produced by this analysis when applied to retinal responses [25]. Representative peri-stimulus histograms (PSTHs) for preferred stimuli of 5 temporal frequencies are shown for a number of neurons; differences in selectivity for stimulus orientation and in the temporal responses to drifting gratings are evident across the manifold.

When individual VISp neurons (points) are colored on the manifold according to their preferred stimuli and/or response features, clear gradients or geometries emerged (Fig. 2b,c). For example, a gradient in the orientation selectivity index (OSI) of each cell was revealed by the manifold (Fig. 2b). This can be seen directly in the example PSTHs (Fig. 2a) and is not surprising, given earlier reports [24][25]. However, not all stimulus features were so well organized. Entropy, a measure of the range of stimuli for which the neurons give good responses, exhibited generally lower values on one side of the manifold, but was distributed more widely than OSI (Fig. 2c). Meanwhile, neurons were not particularly well organized with respect to their spatial frequency preference (Fig. 2d), a point to which we shall return later.

Broadly, the organization of stimulus features across the manifold had a similar topology to that from our previous studies that used distinct stimulus ensembles [24][25]. This suggests that the topology of the VISp encoding manifold is not strongly dependent on the chosen stimulus set, provided that the set is sufficiently diverse. Of course, we cannot rule out that other topological features (i.e., distinct clusters) might emerge with yet-to-be tested stimuli.

The organization of features that were not used in the calculation of the manifold is more interesting. Neurons were classified on the basis of their electrical waveforms as putative excitatory (broad-spiking) or inhibitory (narrow-spiking). The putative excitatory neurons were concentrated on the right on this view of the manifold, with a secondary concentration salient at the lower left, and were present at lower concentration throughout the middle of the manifold (Fig. 2e). In contrast, the putative inhibitory neurons were concentrated in the upper left corner and were also present at a lower concentration throughout the middle of

the manifold; they were less well represented in the regions of the manifold in which the excitatory neurons were concentrated (Fig. 2f). The encoding manifold recapitulates the well known observation that inhibitory neurons tend to be less selective for stimulus orientation and to fire at higher rates than excitatory neurons in mouse V1.

The encoding manifold for V1 also organized neurons according to cortical layers (Fig. 2g,h). Putative excitatory cells in layer 2/3 were concentrated in the lower right, as are those in layer 4 though less tightly, while excitatory cells in layer 6 were much more tightly concentrated in the upper right. It is known that layer 6 excitatory cells tend to have low firing rates [73], which is confirmed by our manifolds. Such cells in layer 5 have two foci of density: one together with those of layer 6, but more tightly focused, and another in the salient concentration at the lower left. This bimodal distribution reveals that layer 5 contains (at least) two populations of excitatory neurons with quite different response properties (cf. [39]). The distributions of putative inhibitory cells in layers 5 and 6 along the top of the manifold toward the left are similar to each other, whereas those in layers 2 and 4 are concentrated lower down except for a cluster at the very top left corner. While the layers differ from one another in the distributions of both types of neurons, within each layer these two major types of neurons are largely complementary. Again, we emphasize that these features of the recording were not utilized in generating the encoding manifold, and thus they demonstrate the ability of the manifold approach to reveal novel relationships between stimulus selectivity, response dynamics and neuronal subtypes.

2.2 Encoding manifolds organize natural scene responses relative to gratings responses

A novel feature of the Allen Institute dataset relative to previous applications of the encoding manifold is the inclusion of responses to natural scenes. This feature provided an opportunity to examine the topology of neural responses to natural scenes relative to gratings. While we did not utilize the responses to natural scenes in generating the encoding manifolds, we could display how neurons were organized on the grating-derived manifold according to their responses to natural scenes [37]. Given the (perhaps) unbounded variability of natural scenes [69, 46], it is difficult to imagine how to arrange them, so we began by simply plotting the mean firing rate to natural scenes across the manifold (Fig. 3a). Surprisingly, the topology of the natural scene responses, like that of the grating responses, varied smoothly across the manifold from right to left (Fig 3b). Comparing the kinetics of the natural scene to grating responses in individual neurons reveals similar dynamics (Fig. 3c,d), but some neurons exhibited larger responses to natural scenes (Fig. 3c), while others responded more strongly to gratings (Fig. 3d).

It should be noted that the static grating stimuli can be said to tile the (static) stimulus space for gratings, encompassing the relevant range of orientations and spatial frequencies so as to stimulate all grating-responsive neurons well, regardless of the location of their receptive fields on the stimulus screen. Therefore, the best response measured with these stimuli for each neuron is likely to be the near-optimal grating response. However, the natural scenes are a small selection from all possible natural scenes and, since they are flashed (not drifted), the responses to them presumably depend strongly on the location of each neuron's receptive field in the visual field.

Nevertheless, the comparable co-variation in the topologies of natural scenes to gratings in the manifolds (Fig. 3a,b) inspired us to consider the ratio of each neuron's responses to the two types of stimuli. This comparison produced a striking result: the relative responses to natural scenes were organized along an axis *orthogonal* to the the gradient of orientation selectivity on the manifold (Fig. 2b). Specifically, the average response of each neuron to the 118 natural scenes (Fig. 1b) was compared to its average response to all the static grating stimuli. Note that the two types of stimuli were presented identically, in separate blocks that were interleaved between the repetitions of each block. The neurons preferring natural scenes (green) were distributed along one side of the manifold while those preferring static gratings were more prevalent on the other side (magenta). An alternative analysis compared the maximum response to any individual stimulus

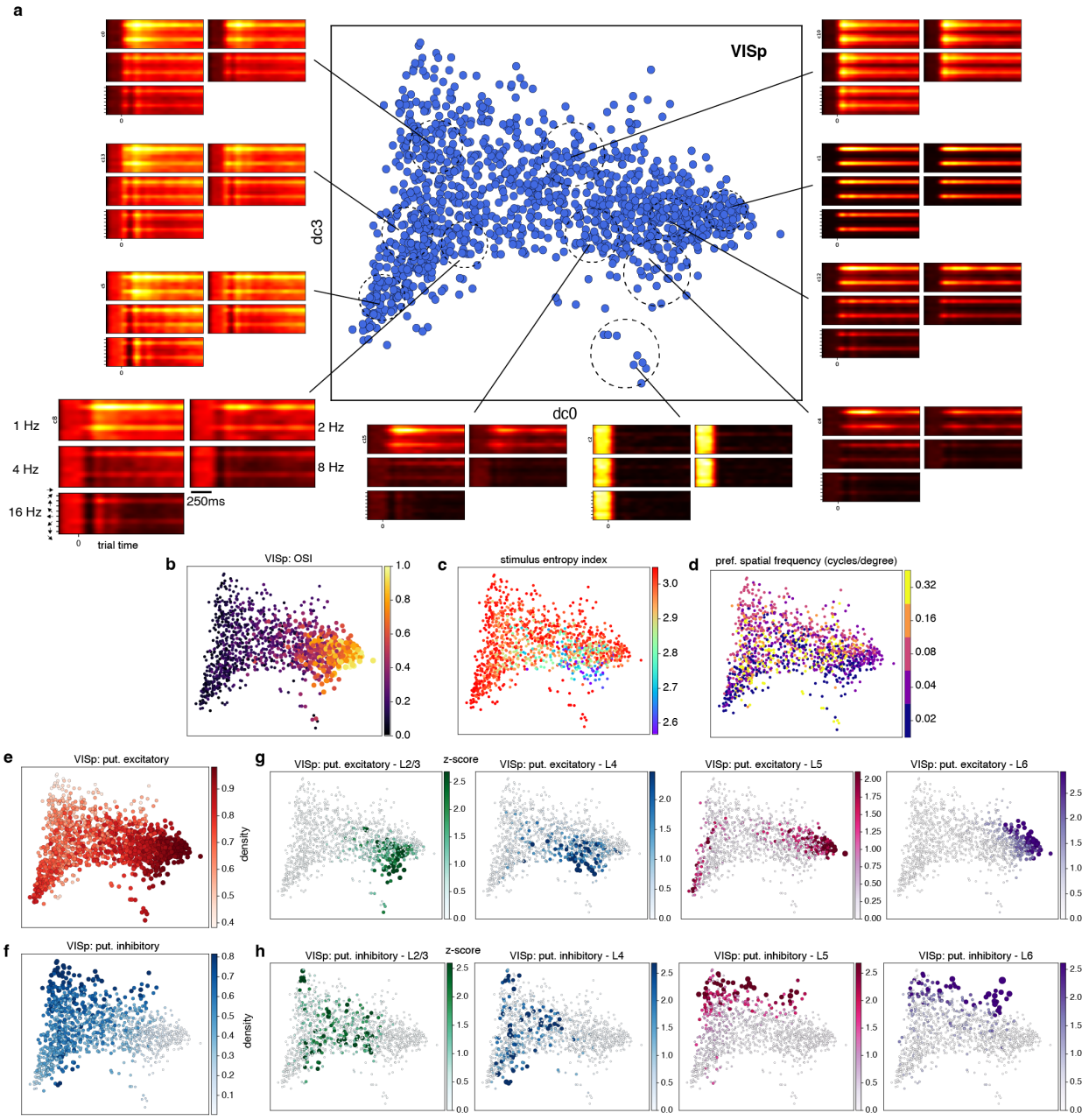

Fig. 2: **The encoding manifold for V1 (VISp) computed from responses to drifting gratings.** **a**, A projection onto the first and fourth diffusion coordinates shows a smooth transition in ensemble PSTHs, from neurons sharply tuned for orientation (right side) to those broadly tuned (left side), indicating a continuous topology. The mean PSTHs for 10 different groups of neurons indicated on the manifold are arranged by temporal frequency from left to right then top to bottom at 1, 2, 4, 8, and 15Hz. **b**, orientation selectivity index (OSI) varies smoothly and continuously across the manifold, while **c** stimulus entropy and **d** preferred spatial frequency are not well organized. **e**, Putative excitatory neurons are concentrated on the right, while **f** putative inhibitory neurons are concentrated on upper left. The distributions of putative excitatory and inhibitory neurons across layer roughly follow this plan, although there are significant differences in detail.

from among the sets of 120 static gratings and 118 natural scenes; this is potentially a noisier measure of selectivity than using the average responses to the entire stimulus set, but it reveals the same result: when the manifold is colored by the ratio of each cell's response to the best natural scene versus its response to the best static grating, the picture is very similar (Fig. S1). This surprising topology across the manifold for natural scenes (relative to gratings) is not a trivial consequence of differences in firing rates; the distributions of average firing rates to static gratings and to natural scenes are similar (Fig. 3a,b).

Despite the marked difference in the two stimulus sets, a relationship between grating preferences and natural scene preferences was revealed by the encoding manifold. We note that it has been notoriously difficult to find any consistent characterization of natural scene responses in visual cortex [29, 28, 58, 62, 11, 80], especially when considering average responses [52]. Yet the encoding manifold has found an organization whereby at any given range of grating preferences, there appears to be an ensemble of neurons that display a range of preferences for gratings versus natural scenes.

2.3 Higher cortical areas exhibit a relationship between natural scenes and gratings similar to VISp

We next computed encoding manifolds from responses to the same gratings for the 5 higher cortical visual areas with the most data in the Allen Neuropixel database: VISam, VISal, VISpm, VISlm, and VISrl; see Fig. 4. The manifolds from all areas were continuous in a manner similar to VISp and distinct from the retina [25]. When the orientation selection index (OSI) and firing rate are plotted as functions on the manifolds, they vary smoothly and similarly, as they did for VISp. Once again, high firing rates are organized with low OSI, and conversely. However, spatial frequency preference does not vary smoothly for any of the higher visual areas, suggesting it is not a well-organized feature in these diffusion coordinates. However, the ratio of the firing rates elicited by flashed natural scenes versus flashed gratings does vary smoothly for all areas (Fig. S2), a point to which we shall return in the Discussion.

We now focus our discussion on two of the areas VISpm and VISal, thought to be representative of the two major (ventral and dorsal, respectively) higher cortical visual streams [30, 49], but see [45].

The responses of VISpm neurons to high spatial and low temporal frequencies suggest that it is for the mouse a part of what is referred to in the primate as the "ventral stream" of higher cortical visual areas, which lead to the inferior temporal lobe and are thought to participate in object recognition. The bias toward lower temporal frequencies is indicated by the distribution of PSTHs shown (Fig. 5(a), especially toward to lower portion). As mentioned above, OSI and firing rate are well-organized coordinates across this projection of the manifold, as is the complementary distribution of excitatory vs inhibitory densities. Note how cells with high OSI are on the left with a gradient to lower OSIs on the right. And again, as in VISp, there is a complementary distribution of excitatory vs. inhibitory neuronal densities for each layer, with the inhibitory neurons tending to align with higher firing rates. All of this confirms at a population level much of what was known for VISpm at the individual cell level.

There are interesting differences in the layer concentrations. In VISpm, L2/3 has significantly different concentrations of excitatory cells from L4, and L5 shows two populations of excitatory cells. L6, on the other hand, is closer to L2 in distribution. We note that previous analyses of layer responses did not distinguish cell types [75].

We next compared the relative responses of natural scenes (NS) to static gratings (SG) for VISpm (Figure 4a). As with VISp, a distributed organization was observed for the ratio of natural scene responses to grating responses, but it is no longer orthogonal to orientation selectivity: unlike VISp, the main concentration of SG-preferring cells overlaps with the most orientation-selective cells on the left side. Examining the layer distribution, these grating-preferring cells tend toward L2/3 and L6. The inhibitory neurons exhibited little preference for SG or NS, responding about equally to the two. These relationships held both when comparing the mean response to all natural scenes compared to the mean response for gratings, or when

comparing the maximum responses for both stimuli. The most curious feature of the layer distributions was the bi-lobed distributions of L5 putative excitatory cells, suggesting that this layer contains two populations of neurons with distinct responses.

We repeated this analysis for VISal (Fig. 6), an area that prefers higher temporal frequencies and lower spatial frequencies, suggesting that it may function in the mouse like the "dorsal stream" areas in the primate (cf. the PSTHs with those in Fig. 5). As expected, orientation selectivity was generally less precise than in VISpm, however, there remains a clear gradient of OSIs across the manifold (Fig. 6(h)). Furthermore, similar to VISpm, a gradient was present for neurons preferring natural scenes vs. gratings. About half the cells, particularly those responding more strongly to natural scenes than to gratings, show little evidence of orientation selectivity. The cells most selective for orientation form the left most border of the manifold, similar to the low-entropy cells. In general, then, there is substantial similarity in functional organization between VISpm and VISal, as has been reported previously, even though the projections from VISp are largely segregated [40].

The distributions of putative excitatory and inhibitory cell types in the different cortical layers of VISal are strikingly diverse (Fig. 6(k,l)). In general, within any one layer, the concentrations of excitatory and inhibitory cells are largely but not fully complementary. Excitatory cells in layer 6 are distributed like the cells that prefer gratings over natural scenes, whereas the layer 6 inhibitory cells are concentrated in the upper right portion of the manifold, with few over most of its area. Layer 2/3 (L2 in the figure) is similar to layer 6, but both cell type, and in particular the inhibitory cells, are more broadly distributed. Layer 4 excitatory cells extend into the center of the manifold, different from any of the other layers, and its inhibitory cell distribution resembles that of layer 2/3. As in VISpm, layer 5 in VISal has the most unusual distribution. Its excitatory cells are concentrated in two regions, one to the lower right, where most cells prefer natural scenes, and one to the lower left, where cells strongly prefer gratings over natural scenes. Layer 5 inhibitory cells also occupy a distinct region of the manifold, different from that of any other layer.

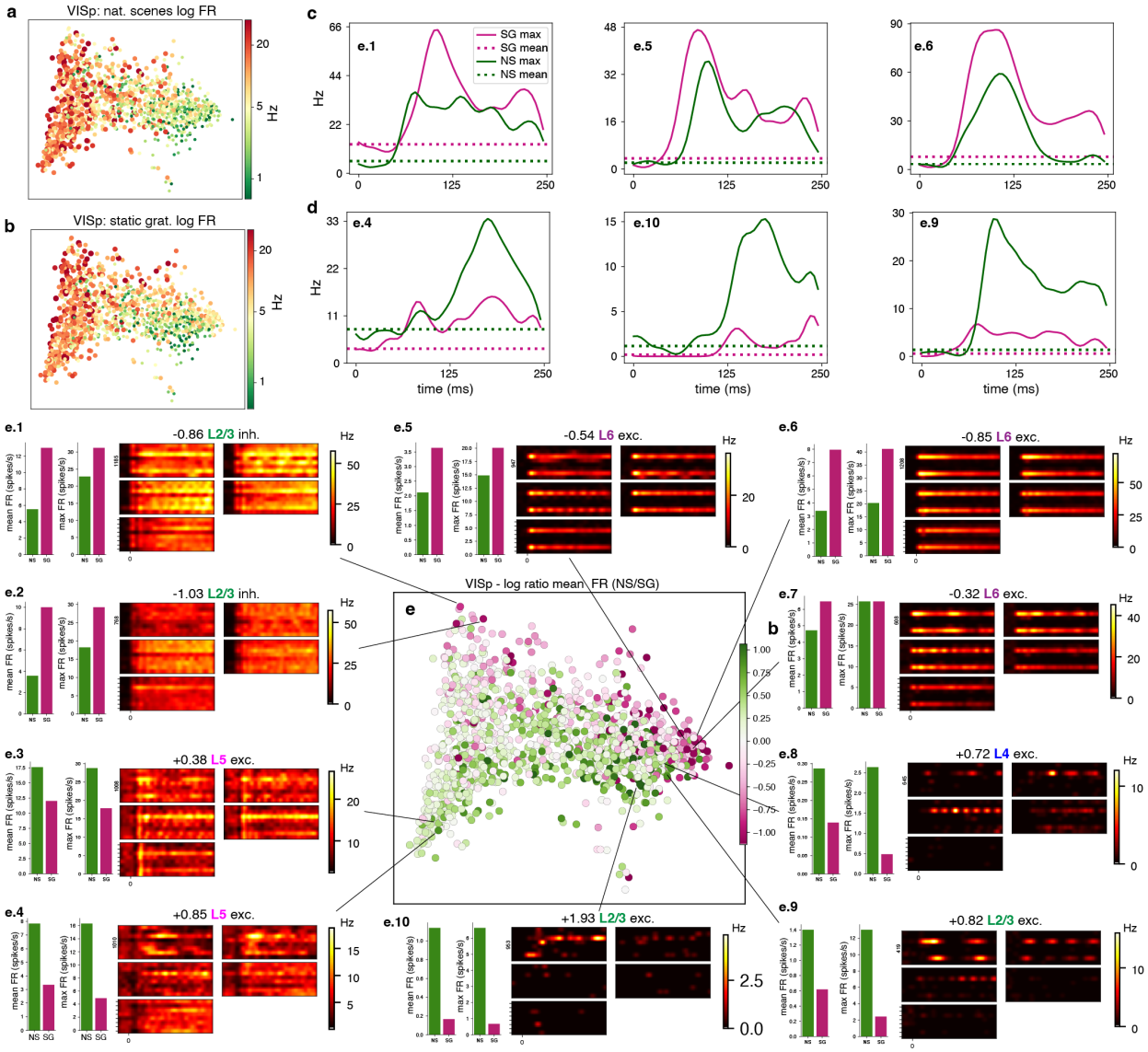

Fig. 3: **The encoding manifold inferred from gratings organizes the ratio of natural scene to static grating responses. a**, Neurons on the encoding manifold colored by log mean firing rate to all natural scenes and **b** log mean firing rate to all static gratings. **c**, Firing rates as a function of time for 3 representative neurons that respond better to static gratings than to natural scenes. **d**, Firing rates of 3 neurons that respond better to natural scenes than to static gratings. **e**, Coloring each neuron on the encoding manifold by the ratio of its mean response to gratings vs. natural scenes reveals a global organization: neurons "preferring" gratings are colored in magenta (toward top of manifold in this projection) while neurons responding more strongly to natural scenes are colored in green. **e.1–e.10** PSTHs as in Fig. 2 and bar plots of firing rates to gratings and natural scenes are shown for representative neurons whose positions are indicated on the manifold. Bar plots show the result is robust across both mean and max firing rates. The top-to-bottom organization of natural scene preference is approximately orthogonal to firing rate and orientation selectivity to drifting gratings (cf. Fig. 2). This organization is thus an emergent property of the encoding manifold; it is not the individual firing rates that matter, but their ratio.

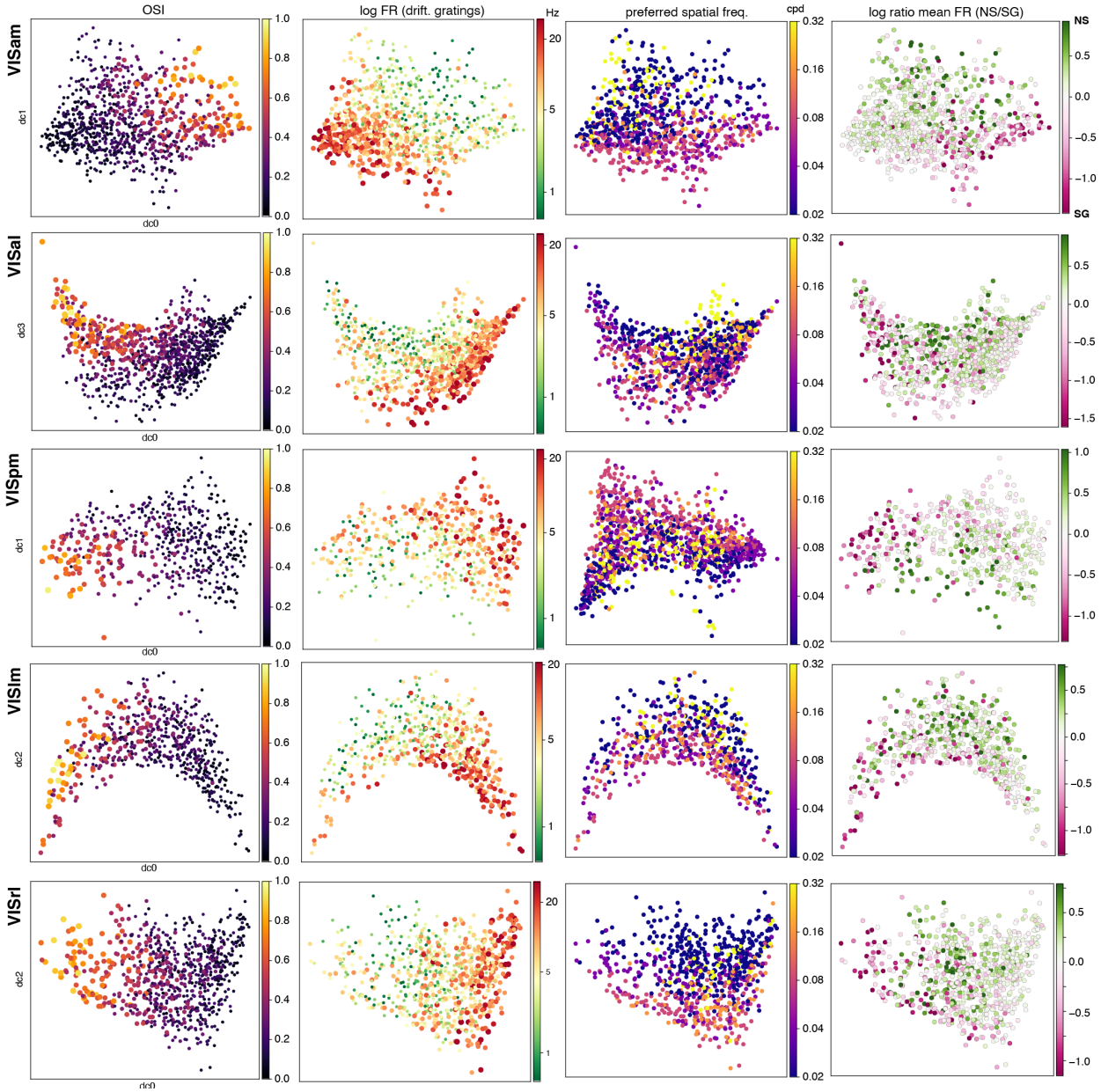

Fig. 4: **Encoding manifolds for 5 higher visual areas (HVAs) computed from grating responses.** Rows indicate the different HVAs and columns indicate features shown in color scales. In the first two columns points on the manifolds are colored by orientation selection index (OSI) and log firing rate to drifitng gratings; note similarity of organization of these properties. As in VISp (V1), preferred spatial frequency (third column) is not well organized in the HVA manifolds, but preference for natural scenes vs. gratings (fourth column, log ratio of mean firing rates to entire stimulus ensembles) is well organized. Unlike in VISp, neurons in the HVAs that prefer gratings tend to have high orientation selectivity (OSI) and lower firing rates (FR).

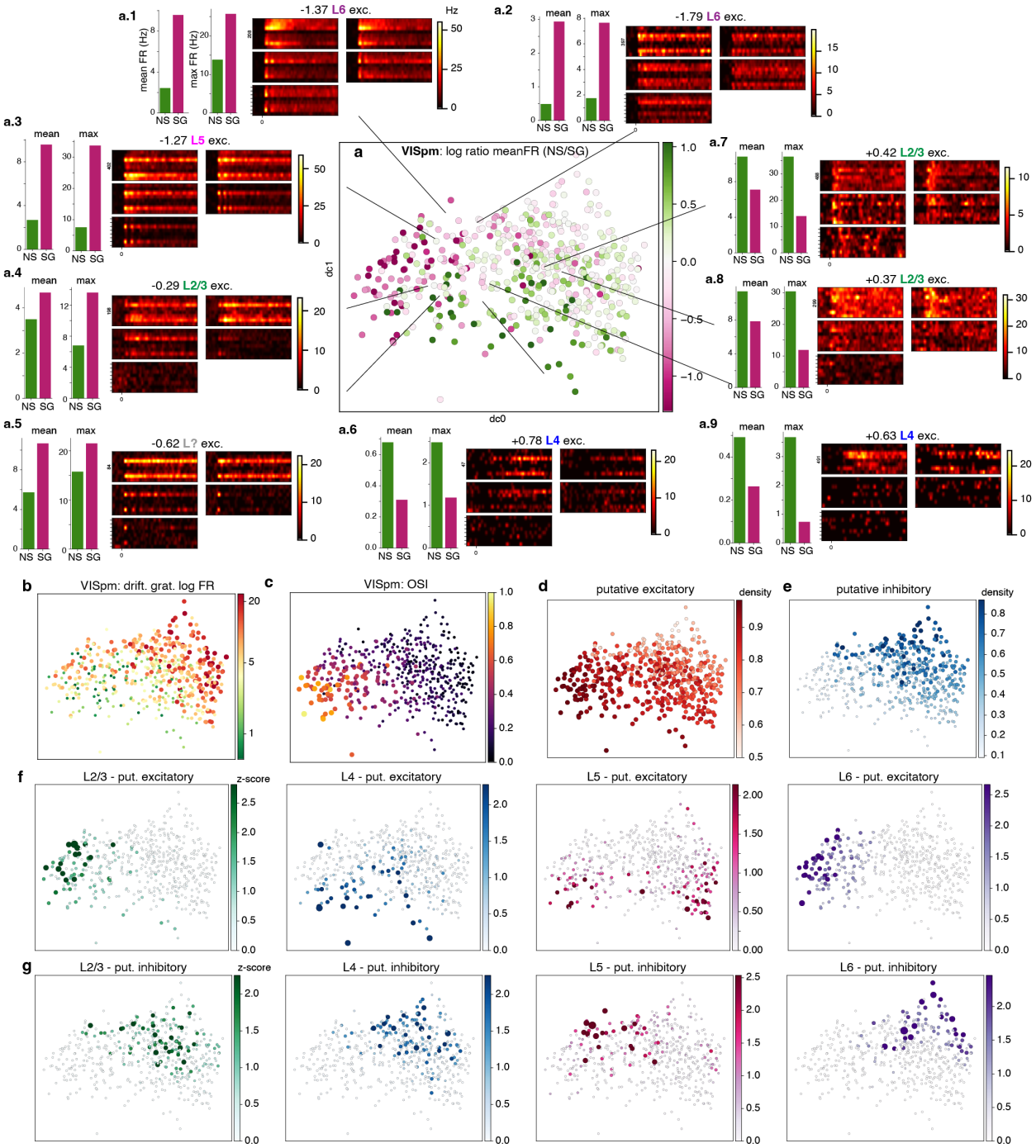

Fig. 5: **The VISpm encoding manifold computed from grating responses.** **a** The encoding manifold for the HVA VISpm colored to indicate ratio of responses to natural scenes versus gratings is again continuous, varying from right (preference for gratings) to left (preference for natural scenes). Details of responses of 9 representative neurons shown as in Fig 3. **b** log mean firing rates to all driftng gratings. **c** orientation selectivity index (OSI) for responses to drifting gratings **d** distribution of putative excitatory (broad-spiking) neurons. **e** distribution of putative inhibitory (narrow-spiking) neurons. **e** distribution of putative excitatory neurons by layer. **g** distribution of putative inhibitory neurons by layer.

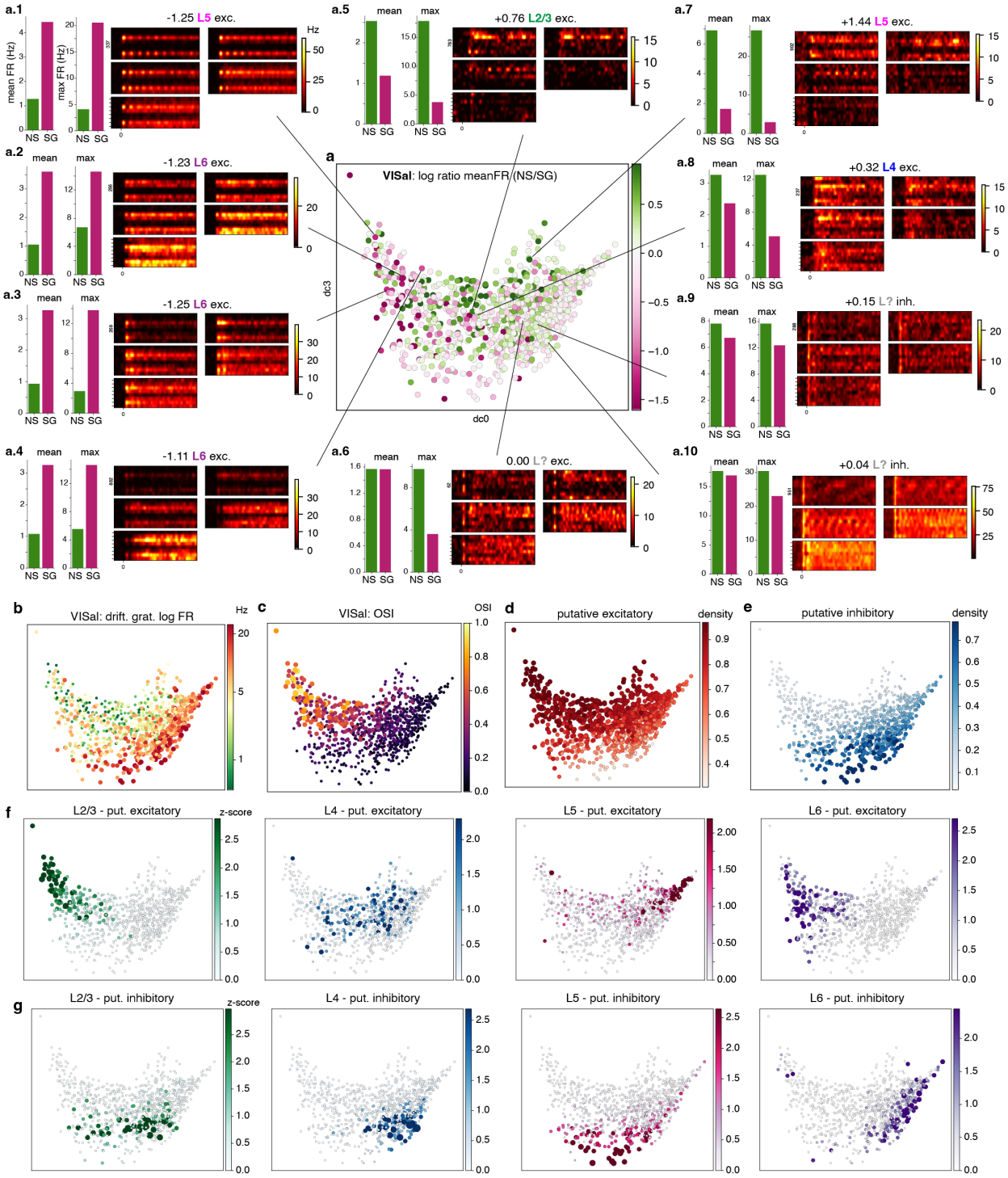

Fig. 6: **The VISal manifold computed from responses to drifting gratings.** All details as in Fig. 5.

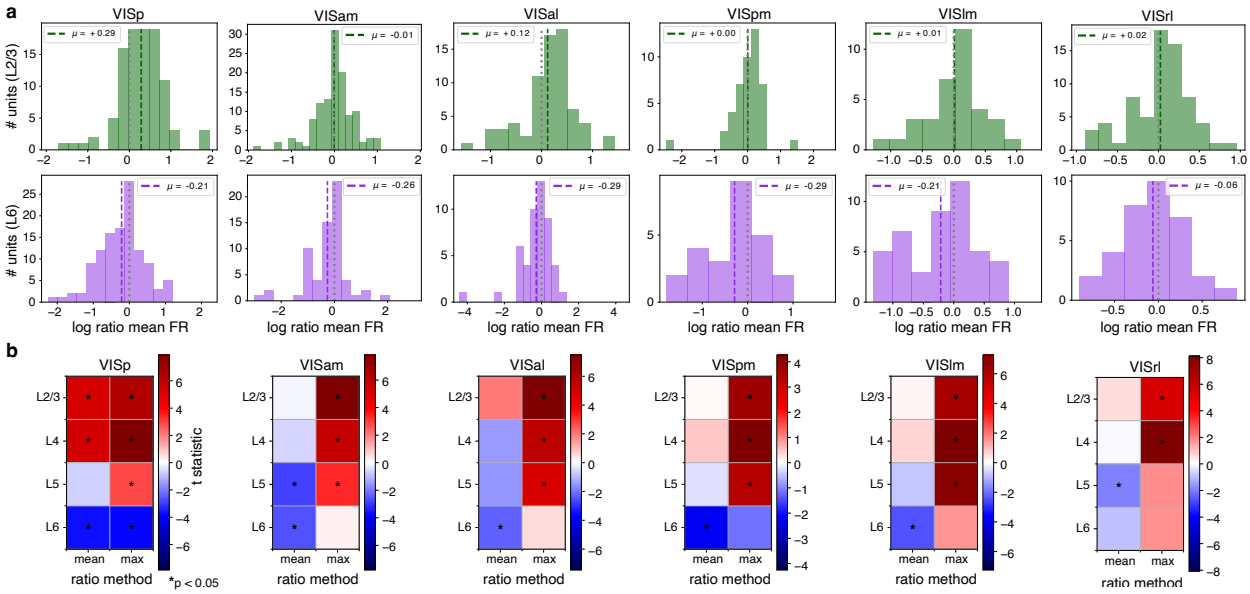

Fig. 7: **Preference for natural scenes vs gratings differs across layers. a**, Histograms of the distribution of ratios (log mean firing rate) for layer 2/3 (**top row**) and layer 6 (**bottom row**) neurons in different visual cortical areas; μ indicates the mean. In layer 2/3 (green histograms), the VISp histogram is shifted toward positive values, indicating a significant preference for natural scenes; all other areas have histograms centered around approximately 0, indicating no preference for natural scenes compared with gratings. By contrast, neurons in layer 6 (purple histograms) show a significant preference for gratings in all areas (negative shift) except for VISrl. Two-tailed t-tests were performed to determine if each distribution differed significantly from a normal distribution (same variance) centered at 0. Significance ($p < 0.05$) is indicated by $\star$. **b**, Summary of t-test statistics comparing the log ratio between natural scenes (positive, red) and static gratings (negative, blue), per cortical layer across areas. Both the mean (left column) and maximum firing rates (right) are shown on the grids. Curiously, a preference for gratings dominates in L6 across all visual cortical areas, especially considering the average responses, while most neurons in other layers respond more strongly to at least one natural scene, especially when considering the maximum response to any one stimulus of each class.

3 Discussion

Our algorithm for inferring the encoding manifold organizes large populations of neurons based on their responses to an ensemble of stimuli, thereby incorporating both circuit and stimulus properties. Responses to gratings collected by the Allen Institute confirmed the continuous topology of the encoding manifold for VISp that we had found earlier using a different stimulus ensemble [25] and extended this topological observation to higher visual areas. Operationally these findings confirm the robustness of our algorithms to neurophysiological data recorded in a different laboratory. We discuss below two curious aspects of the Allen data revealed by the encoding manifold.

3.1 Layer preferences

Others have observed differences in grating selectivity between visual areas [30, 49], and interpreted this as a hint toward dorsal/ventral stream models [76, 68, 45]. Our encoding manifolds allowed us to refine the population analysis to individual layers, where further, novel differences emerged. In VISp, for example, the greatest concentrations of high orientation selectivity of putative excitatory cells were in layers 5 and 6. In the hierarchy of visual cortical areas, the upper layers of lower areas project to layer 4 of higher higher areas, the outputs of which go to the upper layers [62]. In VISal, however, the excitatory neurons of layer 2/3 lie in the regions of the manifold with high orientation selectivity, while those in L4 do not (Fig. 6(h,l)). The same is true to a lesser extent in VISpm (Fig. 5(b,f)), whereas in the VISp manifold the L2/3 and L4 excitatory cells occupy a region with similar orientation selectivity (Fig. 2(g,h)). This is surprising if the dominant feedforward projection from superficial layers of VISp is to layer 4 of the higher cortical visual areas; one would expect the recipient L4 populations not to lose selectivity.

The deeper layers are also interesting, since they involve both cortical and sub-cortical projections [73]. Recall that projections to the dLGN may sharpen receptive fields in the mouse [12] and cat [53, 63, 8]), while the cortico-cortical projection could be involved in generating endstopping and other surround effects [22]. L5 also projects to the pulvinar [42], a projection that could be relevant to coordinating natural scene responses [82, 10].

We observed, on our manifolds, that L5 and L6 are very similarly organized for inhibitory cell types, but not for excitatory ones. In VISp, L5 exhibits two distinct population concentrations, one of which overlaps almost completely with L6 in the region of highest orientation selectivity. In the higher visual areas, L5 in the VISpm manifold is similarly organized in two clusters, although the overlap of the one cluster with L6 is less complete. On the other hand, the L5 cluster of high orientation selectivity is completely missing in the VISal manifold, while the highly tuned L6 cluster remains. This difference between VISpm and VISal is striking, and consistent with the anatomical observations above (cf. [40]).

The distributions of other properties by layer for VISam is in Fig. S3; for VISlm in Fig. S4; and for VISrl in Fig. S5. Layer 5 in all three cases indicates a multi-lobed concentration distribution, one of which remains consistent with the density of (putative) excitatory neurons in L6. Perhaps this reflects processing involving the horizontal pathways within an area [2, 3, 79].

The preference for natural scenes vs. gratings also differs between cortical layers. Fig. 7 compares the selectivity for natural scenes vs. gratings for the entire populations of the different cortical layers of the 6 areas. Histograms are shown of the log mean firing rate for neurons in layers 2 and 6. In these histograms a shift above 0 indicates a preference for natural scenes, while a shift below 0 indicates a preference for gratings. A two-tailed t-test was used to determine if the distribution differed significantly $(p < 0.05)$ from a normal distribution (same variance) centered at 0. In layer 2, only units in VISp show a preference for natural scenes, while the others show no significant preference. In layer 6, however, almost all visual areas show a preference for gratings.

These results are summarized graphically in Fig. 7b, with blue indicating a shift toward gratings and red

toward natural scenes. The left column of each plot was made from the mean responses above; the preference for gratings in layer 6 is clear from the blue squares. The right column was made from the maximum response for each neuron to the best grating and the best natural scene, to illustrate a complementary norm. In VISp, but not in the higher cortical visual areas, layers 2/3 and 4 have significantly stronger responses to natural scenes in the mean sense. VISp granular and supragranular layers seem particularly tuned to some generic feature of the natural scene images. When the responses to the best grating and the best natural scene for each cell are considered, a highly significant majority of cells in layers 2/3, 4, and 5 of all areas find a natural scene that drives them most strongly (except for area VISrl, where layer 5 is not significant). In VISp, the maximum response of layer 6 cells still favors gratings, but in all other areas, neither stimulus class drives the population significantly better. It is to be expected that there might exist one particularly "exciting" piece of natural scene among the 118 used that will drive any neuron more strongly than the gratings do. Surprisingly, however, that is typically not the case for layer 6 in any of the areas. Surround suppression is thought to be lacking or nearly so in layer 6, at least in experiments done with anaesthesia [61]; other experiments show surround feedback effects in layers 5 and 6 [38].

3.2 Natural scenes, spatial frequency, and "U"-shaped preferences

Natural scenes can be used to estimate receptive fields, since a large ensemble of them can approximate a random stimulus set, but such an analysis does not inform about the processing of any individual natural scene [26]. Nevertheless, there is information beyond the RF in the neural population representation, because it can be decoded [69]. An advantage of the encoding approach using diffusion geometry is that certain coordinates can identify subtle organizing trends for natural scenes. For the Allen dataset the diffusion manifolds uncovered a relationship between natural scene responses relative to grating responses, not only from VISp (Fig. 3) but also for the higher visual areas (Fig. 4). We now seek to understand the relationship identified by the difference in the location on the encoding manifolds of the neurons preferring one versus the other stimulus class.

A plausible place to start is with the selectivity to spatial frequencies (Fig. 8a). As expected from natural scene statistics, there are more cells preferring lower spatial frequencies than higher ones, a predominance consistent with optimal coding theory [59, 71, 64]. However the distribution of spatial frequency preference is not well organized on the manifold (Fig. 2d). Moreover, the different spatial frequency bands need not be taken as independent. Given the selectivity revealed by plotting the log ratio of mean firing rate to natural scenes relative to static gratings on the manifold, as described previously, we next plotted this statistic as a function of the spatial frequency preference for different neuronal groups, to check whether the relative ratio would form a pattern. The result was a new invariant: a "U"-shaped distribution, showing that cells in the group preferring either low or high spatial frequencies had the largest responses to natural scenes, while those preferring intermediate frequencies responded more vigorously to gratings (Fig. 8(b)). A t-test was used to determine if the distribution differed significantly $(p < 0.05)$ from a normal distribution (same variance) centered at 0. Importantly this distribution held not only in VISp but also in all five of the higher visual areas studied. To check whether it held at the population level and considering only the grating responses, we combined the low- and high-frequency preferring cells together, versus those preferring intermediate frequencies, and plotted this new statistic on the encoding manifold. The density of high + low preferring cells complements beautifully the density for intermediate-preferring cells, and their distributions over the manifold vary almost identically to the distributions of cells preferring natural scenes versus gratings (Fig. 8g–i, and Fig. S6). To repeat: the population of natural scene preferring cells distribute identically, on the manifold, to cells preferring high or low spatial frequencies.

How might this curious statistic be explained (cf. [80, 11])? One possibility is that it somehow has to do with object content, or at least a proxy for it. Examining the statistics of image content is instructive (Fig. 8(c–f). Low frequency filtering signals "where" an object might be, while high frequency signals

"what" it might be made of. Such an identification of spatial frequency content with scene layout has long been proposed [70, 56], and blobs do indeed matter at low frequencies [60]. But it is unlikely that simply extracting filtered features is the complete story (e.g., [77]).

Object perception, figure/ground discrimination and hunting behavior [36, 67] are complex tasks, implicating network effects across multiple visual areas [38], including feedback within and between them [31, 33] and recurrence [41, 81]. A recent study [66] found direct low- and high-frequency dynamics in natural scene viewing, extending classical results on dynamical frequency tuning [4, 14, 51, 74]. Since our algorithm (see Methods in [25]) uses the full PSTH, it captures these dynamics—a possible signature of coarse-to-fine processing [48, 35]—in the manifold inference process.

4 Summary and conclusions

We here applied the encoding manifold approach to an extensive dataset collected by the Allen Institute, and were able to verify that the continuous topology of VISp agreed with an earlier assessment [25] obtained using different stimuli, different electrodes, and done in a different laboratory. We further showed that the higher visual areas in mouse shared this continuous topology, and elaborated population distributions across layer and cell type. A distinct preference among layer 6 (putative) excitatory neurons for gratings held across all areas, although only weakly so in VISrl.

Natural scenes have been among the most difficult stimuli to analyze, and the encoding manifold revealed a unique invariant at the population level. Although the manifolds computed in this study were based solely on static scenes, neglecting potential contributions of motion and navigation [47, 65, 27], a surprising and consistent relationship emerged between the natural scene preferences of high- and low-spatial frequency preferring cells compared to cells preferring intermediate frequencies. Classically, scene analysis has been thought to proceed by a coarse-to-fine visual processing [34]. Perhaps while building biologically-motivated perceptual systems, following this study one should consider coarse-*and*-fine processing.

Encoding manifolds capture similarity in the responses over time of neurons exposed to a wide range of stimuli. They display neurons in relation to as many different stimulus features as one is able to illustrate. No other procedure of which we are aware lets one simultaneously visualize how the neurons that respond particularly well to one stimulus respond to many others. They disclose visually the relationships between, for example in the present data, firing rate, selectivity for stimulus orientation, spatial and temporal frequencies, cortical laminae, electrical waveform (related to excitatory versus inhibitory class), overall selectivity (stimulus entropy), and preference for gratings versus natural scenes. This property should make encoding manifolds widely useful in across biology when many relationships may be present but go unnoticed.

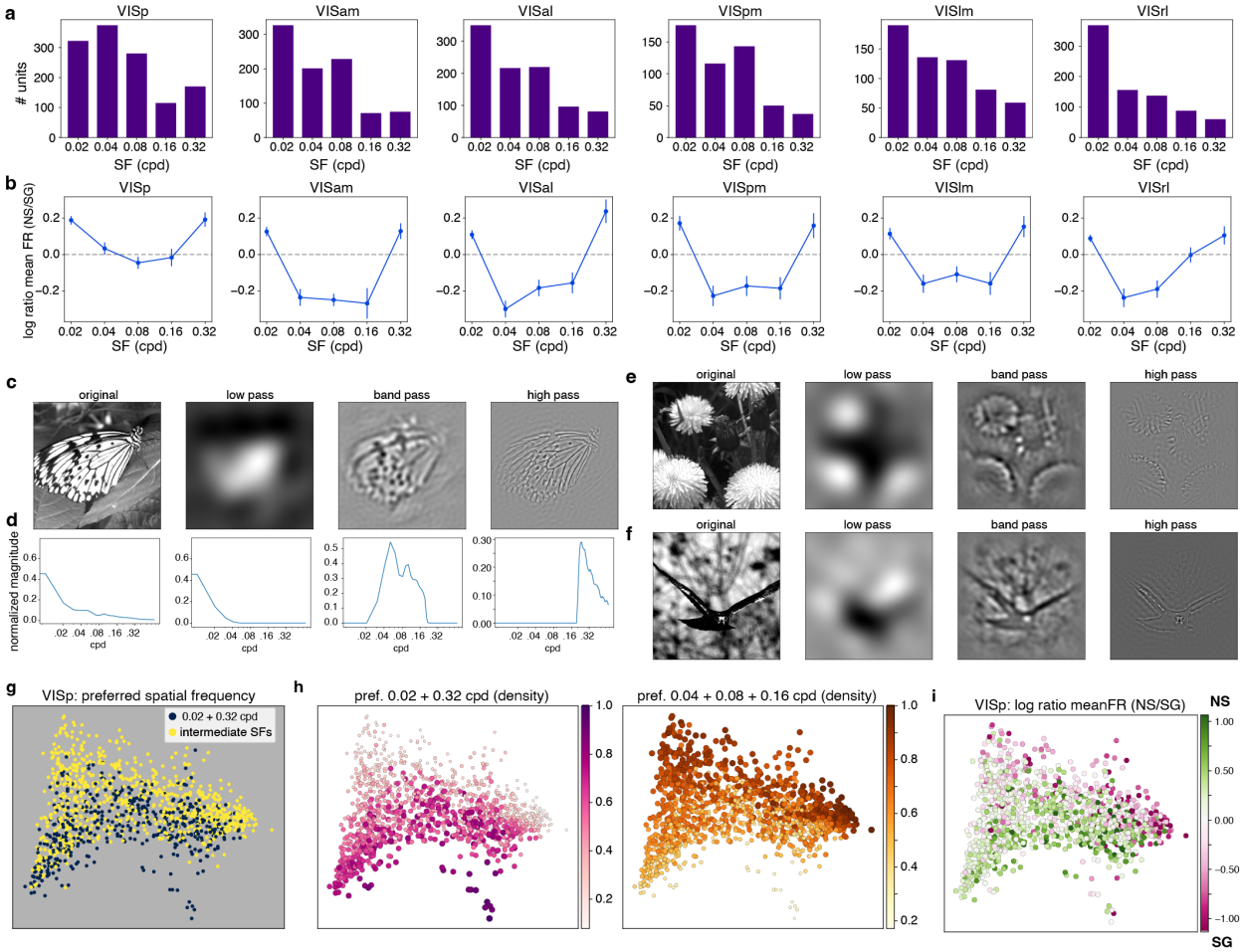

Fig. 8: **Natural scene statistics and neuronal preference on encoding manifold. a** Distribution of numbers of units preferring a given spatial frequency in all visual areas. As expected, it drops off with increasing spatial frequency, although many cells prefer higher spatial frequencies; cf. Fig. 1d. **b**, Mean $\pm$ s.e.m. for log ratios of mean firing rates of the neuronal populations in each area as a function of their preferred spatial frequency (see Methods). *Neurons that prefer low or high SF tend to respond more strongly to natural scenes, while those that prefer intermediate frequencies respond more strongly to gratings.* This U-shaped curve holds across all areas, and is perhaps surprising compared with the plots above in **a**. **c** A natural image and its low-pass, intermediate-, and high-pass filtered versions, and **d** associated spectra. **e,f** Two more natural images plus their filtered versions. **g** The different spatial frequency bands perhaps provide some insight into the U-shape curves, suggesting a grouping of the low and high frequency content signals separately from the intermediate ones. Here we show how the frequency representations associate with natural scene preferences on the encoding manifold for V1 (VISp). Cells are now labeled according to their spatial frequency preference for static grating stimuli, grouping 0.02+0.32 (extreme frequencies, in yellow) vs. 0.04+0.08+0.16 (intermediate frequencies, in blue), agreeing with the natural scene ratios plotted on the encoding manifold in Fig. 3e. In more detail, this composite plot can be decomposed into **h** the density over the data graph (see Methods) for preference for extreme frequencies (left) and for intermediate frequencies (right). **i** The log ratios from Fig. 3e, here for comparison with the distributions in **h**. For comparable plots in the higher visual areas, see Fig. S6.

5 Methods

5.1 Dataset

The "Neuropixel" single neuron recording dataset shared by the Allen Institute includes recordings from many cortical visual areas in alert wild type (C57BL6/J) mice Harris et al. [33]. Recording sites were localized to specific cortical areas by reference to intrinsic signal imaging maps and to cortical layer by transformation into the Allen Common Coordinate Framework. Details of all procedures are described at links from [5].

To summarize, visual stimuli were were displayed on an LCD monitor at a resolution of 1920×1200 pixels at 60 Hz refresh rate. Stimuli were presented monocularly, and the monitor spanned $120° \times 95°$ of visual space prior to stimulus warping. Each monitor was gamma corrected and had a mean luminance of 50 cd/m^2. To account for the close viewing angle of the mouse, a spherical warping was applied to all stimuli to ensure that the apparent size, speed, and spatial frequency were constant across the monitor as seen from the mouse's perspective.

In our experiments, we analyzed responses to a subset of their "Brain Observatory 1.1" stimulus set, including 40 different drifting gratings, 120 stationary gratings, and 118 natural scenes [7]. The drifting gratings were shown at a spatial frequency of 0.04 cycle/deg, 80% contrast, 8 directions of motion (0, 45, 90, 135, 180, 225, 270, and 315 degrees) and 5 temporal frequencies (1, 2, 4, 8, and 15 Hz), with 15 repeats per condition. The stationary grating set, also presented at 80% contrast, consisted of 120 gratings of 6 different orientations (0, 30, 60, 90, 120, and 150 degrees, clockwise from 0 = vertical), 5 spatial frequencies (0.02, 0.04, 0.08, 0.16, 0.32 cycle/degree), and 4 phases (0, 0.25, 0.5, 0.75), each presented for 0.25 seconds with no intervening gray period. The natural scene stimuli consisted of 118 natural images taken from the Berkeley Segmentation Dataset [50], the van Hateren Natural Image Dataset [72], and the McGill Calibrated Colour Image Database [57]. The images were presented in grayscale and were contrast normalized and resized to 1174 x 918 pixels. The images were presented in a random order for 0.25 seconds each, with no intervening gray period.

5.1.1 Data preprocessing

As in Dyballa et al. [25], responses of a neuron to each stimulus were kernel-smoothed using the Improved Sheather-Jones algorithm [13], using a bandwidth of 25 ms (typical value obtained for all visual areas when using automatic bandwidth selection. The `KDEpy` Python implementation [55] was used.

Responses to the drifting gratings were tested for significance by performing a two-tailed Mann-Whitney test comparing the mean activity during the second half of the ISI immediately preceding the stimulus versus mean activity over any interval of the same length within the stimulus presentation. This strategy allowed transient responses to be able to reach the same level of significance as sustained responses. Responses were deemed significant if they had a p-value of at most 0.001.

5.2 Data analysis

5.2.1 Neural encoding manifolds

Neural encoding manifolds were constructed from the spikes elicited by drifting gratings in VISp and each of 5 higher cortical visual areas using the procedure described in Dyballa et al. [25]. Briefly, it involves a dimensionality reduction pipeline consisting of two main steps: applying a permuted non-negative tensor factorization (pNTF) to a tensor built from the temporal responses of each neuron to each variation of grating stimuli, followed by manifold inference using iterated adaptive neighborhoods (IAN) [23] in combination with diffusion maps [19] to produce an embedding in which each point is a neuron. We briefly overview this

material here for completeness; please refer to Dyballa et al. [25] for details and further discussion. Code is available at `https://github.com/dyballa/NeuralEncodingManifolds`.

Tensor decomposition We organize the data as a *3-way tensor*, NEURONS$^{(1)}\times$ STIMULI$^{(2)}\times$ RESPONSE IN TIME$^{(3)}$. (For an introduction to tensors, see [78].) In standard notation [43]:

$$\tilde{\boldsymbol{T}} = [\![X^{(1)}, X^{(2)}, X^{(3)}]\!] \equiv \sum_{r=1}^{R} v_r^{(1)} \circ v_r^{(2)} \circ v_r^{(3)}, \tag{1}$$

where $\circ$ denotes vector outer product and the *factors* are collected into the factor matrices $X^{(k)}$, with individual factors $v_r^{(k)}$ as columns. A *component* is an associated set of factors, one from each tensor mode.

Factors are normalized by computing $\lambda_r = \prod_{k=1}^{3} \|v_r^{(k)}\|$, where r indexes the component and k the factor mode. Collecting these scalars into a vector $\lambda \in \mathbb{R}^R$, we have:

$$\tilde{\boldsymbol{T}} = [\![\lambda\,;\, X^{(1)}, X^{(2)}, X^{(3)}]\!] \equiv \sum_{r=1}^{R} \lambda_r v_r^{(1)} \circ v_r^{(2)} \circ v_r^{(3)}. \tag{2}$$

Since neuronal firing rates are non-negative, we adopt a non-negative tensor factorization (NTF) algorithm [15] based on CP decomposition. This minimizes the squared reconstruction error [17]: $\min_{X^{(1)},X^{(2)},X^{(3)}} \frac{1}{2}\|\boldsymbol{T} - \tilde{\boldsymbol{T}}\|^2$ subject to the non-negativity constraint $X^{(k)} \geq 0, k \in [1,2,3]$ [44, 17]. Non-negative tensor factorization is further advantageous in the sense that it imposes a part-like decomposition [44, 17].

For implementation, we use a gradient-based direct optimization approach (OPT) [1] in the Tensor Toolbox [9], modified to allow permutation of the response factors. This allowed us to exploit data limitations, since two narrowly-tuned simple cells (for example) are similar except in their preferred orientations, effectively collapsing the orientation-preference coordinate on the manifolds.

Because drifting and static gratings have different trial lengths, two tensors were constructed for each cortical area: one containing the drifting grating responses (in which the stimulus mode consisted of the 5 temporal frequencies) and another for the static grating responses (in which the stimulus mode consisted of the 5 spatial frequencies). The temporal mode was a concatenation of the various directions/orientations (8 directions of motion for the drifting gratings and 5 orientations at the optimal phase for the static gratings). The responses of a neuron to each stimulus were kernel-smoothed using the Improved Sheather-Jones algorithm [13], which does not assume normally distributed data; we used the `KDEpy` Python implementation[55] and a bandwidth of 25 ms.

Neural encoding space. The first mode of each decomposition is a neural factor; together they can be composed into a neural matrix (factors as columns). We interpret this as a linear space spanned by the outer product of the other modes (stimulus and response) after vectorization and normalization. Thus, each neuron can now be viewed as a function of its response to (weighted combinations of) the stimuli. The number of tensor components — technically the dimension of the neural encoding space — is determined by the "explained variance" condition discussed in Dyballa et al. [25], *Methods*.

Similarity kernel and diffusion maps. The data graph for manifold inference is built in the neural encoding space. Since neighborhoods may vary, we use the IAN kernel [23], a multiscale Gaussian kernel that adapts to local geometry in a (relaxed) optimal fashion. The IAN weighted graph yields the similarity matrix on which the diffusion maps algorithm [19, 18] is based. Diffusion maps are a nonlinear spectral embedding method based on the eigendecomposition of the normalized graph Laplacian that tends to better preserve the topology of the data [23]. The standard parameters α = 1 (Laplace-Beltrami approximation)

and diffusion time $t = 1$ were used. The resulting diffusion coordinates are then used to embed the neurons in (reduced) stimulus-response coordinates (see above).

Our figures show only a projection of the diffusion maps into two dimensions, while the intrinsic dimensionality was between 5 and 6 for all cortical areas; see [23] for computation details. Just as in principal component analysis, where different coordinates — principal components — emphasize different aspects of the data, so too do lower-dimensional projections of diffusion maps illustrate different features of organization. We chose the two-dimensional projections of the manifolds shown in the figures to emphasize how they organize the properties such as orientation selectivity noted in the text. Other coordinates organize different properties, such as temporal frequency preference.

Laminar and putative type densities around each neuron in a manifold were computed as the fraction of adjacent nodes in the non-weighted IAN graph belonging to the same putative type and layer.

5.2.2 Natural scene vs. static grating selectivity ratio

Selectivity ratios for individual neurons were computed as the natural logarithm of the ratio between their firing rate to natural scenes and their firing rate to static gratings. Thus a selectivity ratio of 0 means the neuron produced exactly the same magnitude of response to both stimulus classes; and a positive (respectively, negative) value indicates a stronger response to natural scenes (resp., static gratings). The magnitudes of responses to these two stimulus classes are readily comparable because both consist of static images presented during the same time interval of 250 ms.

Following the methodology from the Allen Brain Observatory SDK [6], mean firing rates (FR) for each stimulus class (namely, drifting gratings, static gratings, and natural scenes) were computed as the mean number of spikes per second of stimulus presentation across all trials of the same stimulus, averaged across all stimulus variations within the same class. For static gratings, variations mean different orientation (6), spatial frequency (5), and phase (4); for natural scenes, variations mean each of 118 scenes used. We also computed the maximum FR for a given stimulus class as the maximum trial-averaged FR across all stimulus variations. Thus, two ratios were computed: one using the mean responses ("mean FR ratios"), and another using the maximum responses ("max FR ratios").

5.2.3 Additional metrics

We used precomputed stimulus metrics available from the Allen Brain Observatory SDK [6] for firing rates, global orientation selectivity index (OSI), preferred spatial frequency (drifting gratings), and preferred phase (static gratings) for each unit. In addition, a stimulus entropy index (Fig. 2c) was defined as 2^H, where H is the base-2 entropy of the vector containing the relative response magnitudes (divided by their sum) of a neuron to the 5 temporal frequencies used in the drifting gratings experiments. It therefore ranges between 0 (case in which the neuron responds to a single stimulus) and 5 (when it responds with uniform magnitude to all stimuli).

5.2.4 Natural scene filtering

Filtered versions of the natural scenes were created by first computing fast Fourier transforms of each image using the FFT module in the `Numpy` Python library [32]. For the low-pass filtered version, the Fourier domain was cropped with a disk of radius 0.04 cpd; for the band-pass version, an annulus with inner radius 0.04 cpd and outer radius 0.22 cpd; and for the high-pass version a disk with radius 0.22 cpd (see example in Fig. S7).

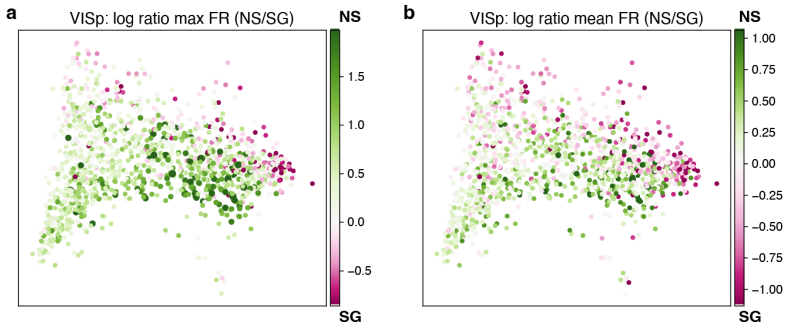

Fig. S1: **Ratio distribution is robust to maximum firing rate.** **a**, Log ratios of the maximum firing rate for all natural scenes divided by the maximum firing rate over all static gratings are plotted on the VISp encoding manifold. **b**, For comparison, an analogous plot using mean firing rates when computing log ratios (cf. Fig. 3 in main text).

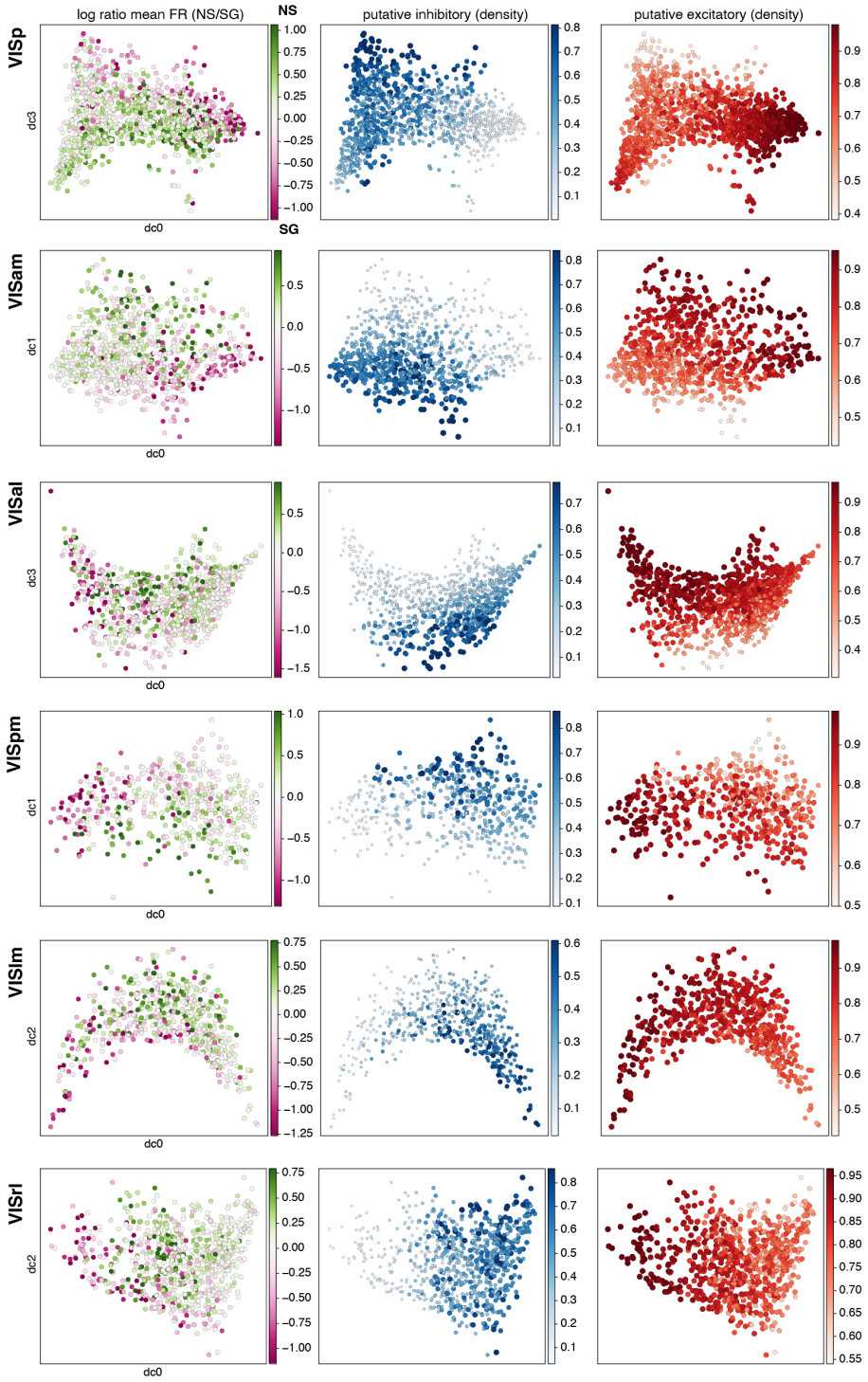

Fig. S2: **Smooth property distributions across all visual areas. left** Encoding manifolds colored by the ratio of natural scene responses to static grating responses. **middle** Encoding manifolds colored by the density of putative inhibitory and **right** putative excitatory neurons.

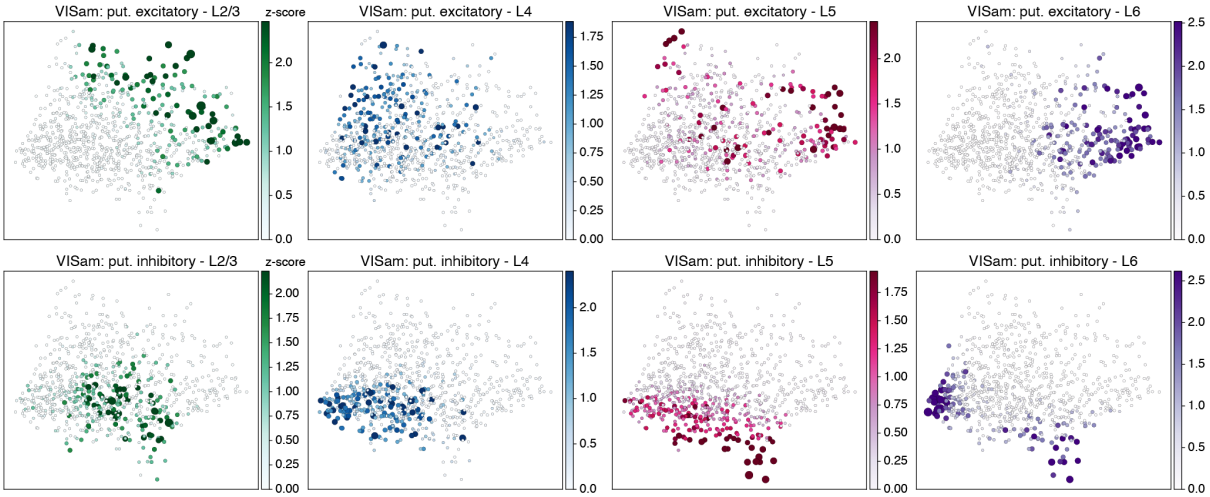

Fig. S3: **Laminar density across the VISam manifold.**. Density of putative excitatory and inhibitory neurons by layer. Note the distinctness of layer 2 vs layers 5 and 6 for excitatory neurons, and the multi-lobed density of inhibitory neurons in layer 6.

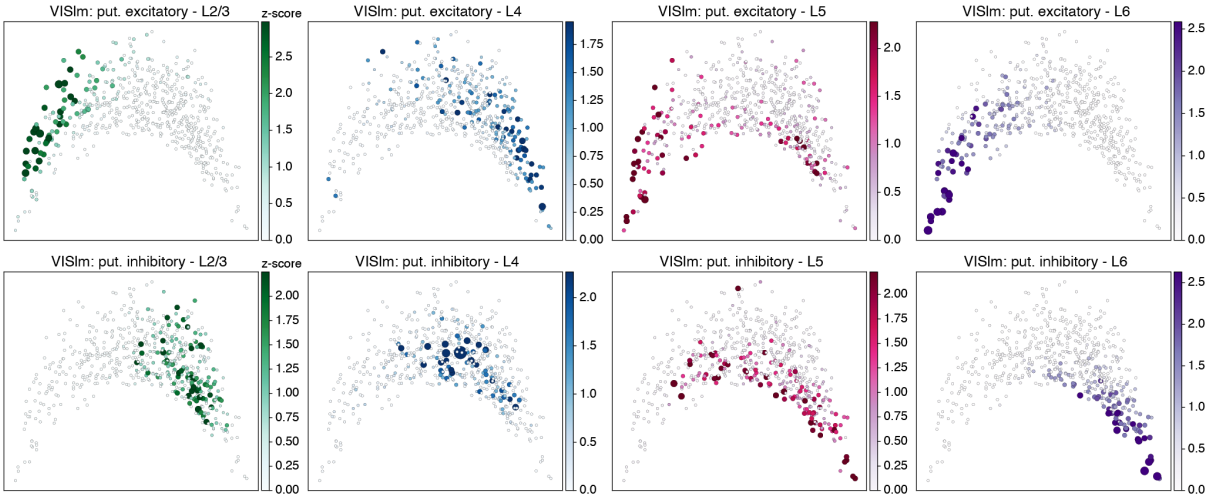

Fig. S4: **Laminar density across the VISlm manifold.** Density of putative excitatory and inhibitory neurons by layer. Note the similarity of layers 2 and 6 for putative excitatory neurons, and the distinctness of layer 4 for putative inhibitory neurons.

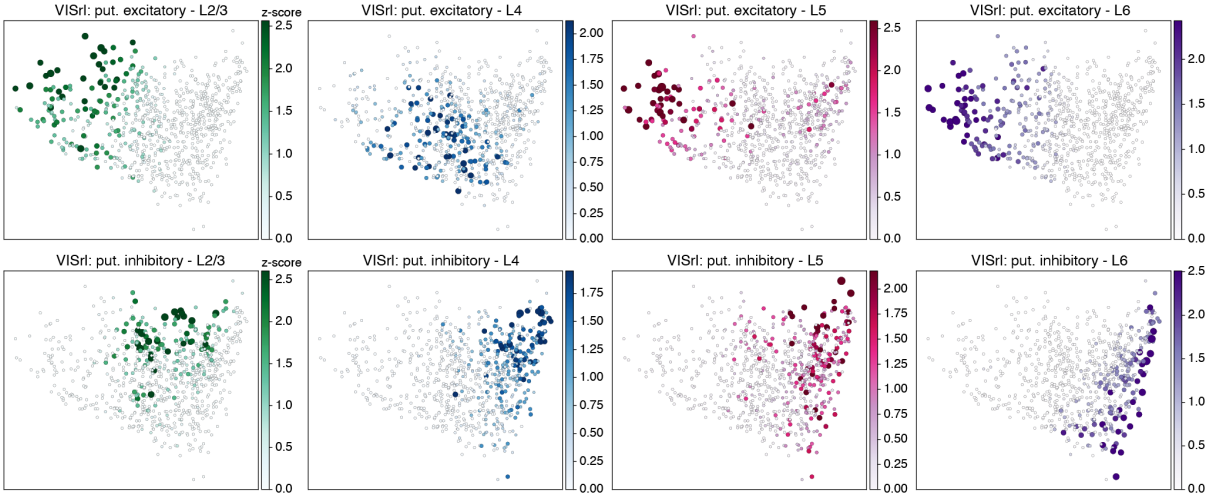

Fig. S5: **Laminar density across the VISrl manifold.** Like VISlm, there is a multi-lobed distribution of putative excitatory neurons in layer 5, and similar agreement between layers 2 and 6.

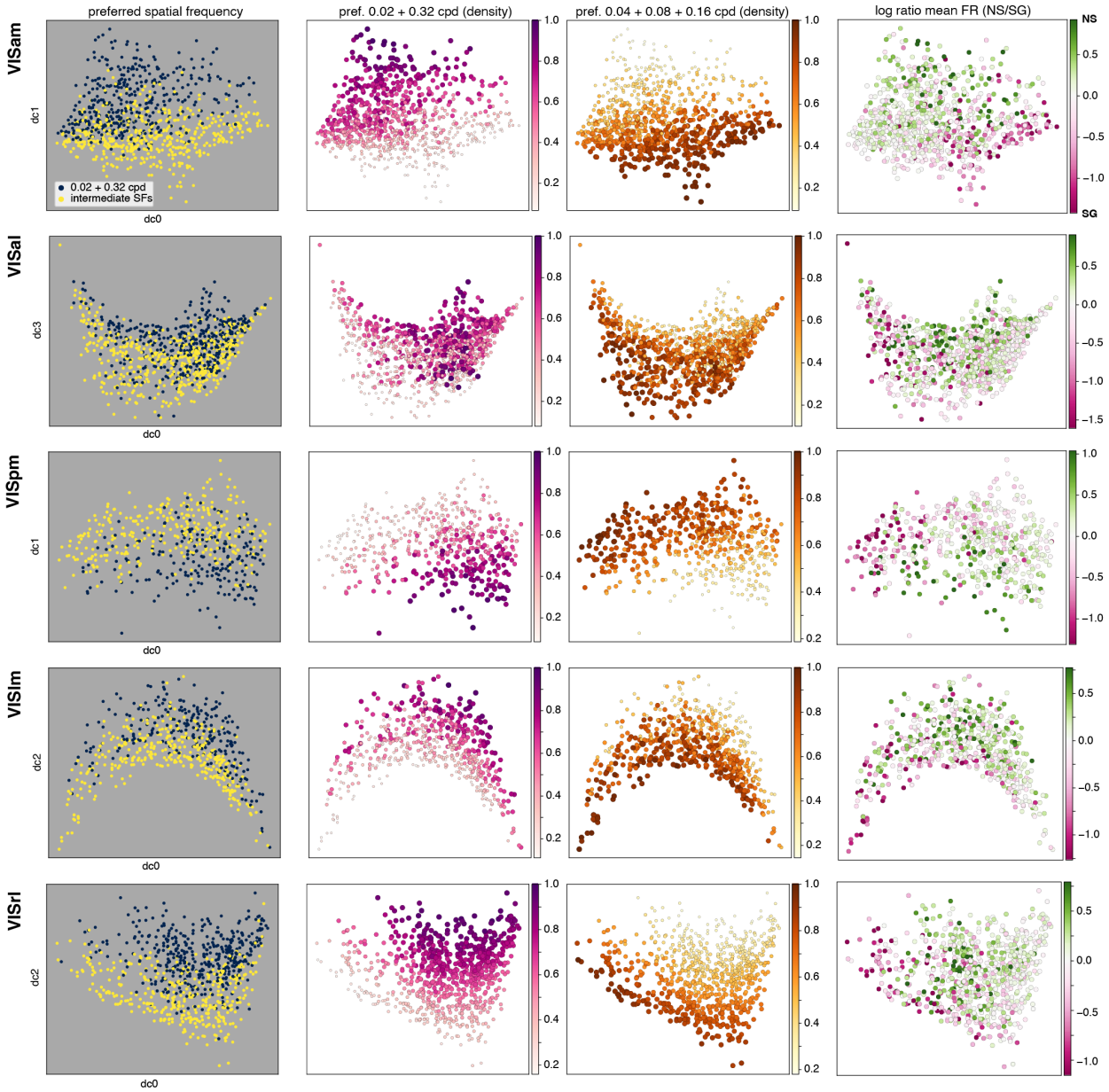

Fig. S6: **Preference for extreme vs. intermediate frequencies.** A grid of colored manifolds, in which each row is a visual area and each column is (from left to right) the ratio of high and low spatial frequencies to intermediate ones; the density of cells preferring high and low; the density of cells preferring intermediate spatial frequencies, and the log ratio plot of mean firing rate to natural scenes (NS) by static gratings (SG). Corresponding plots for VISp are shown in Fig. 8 g,h.

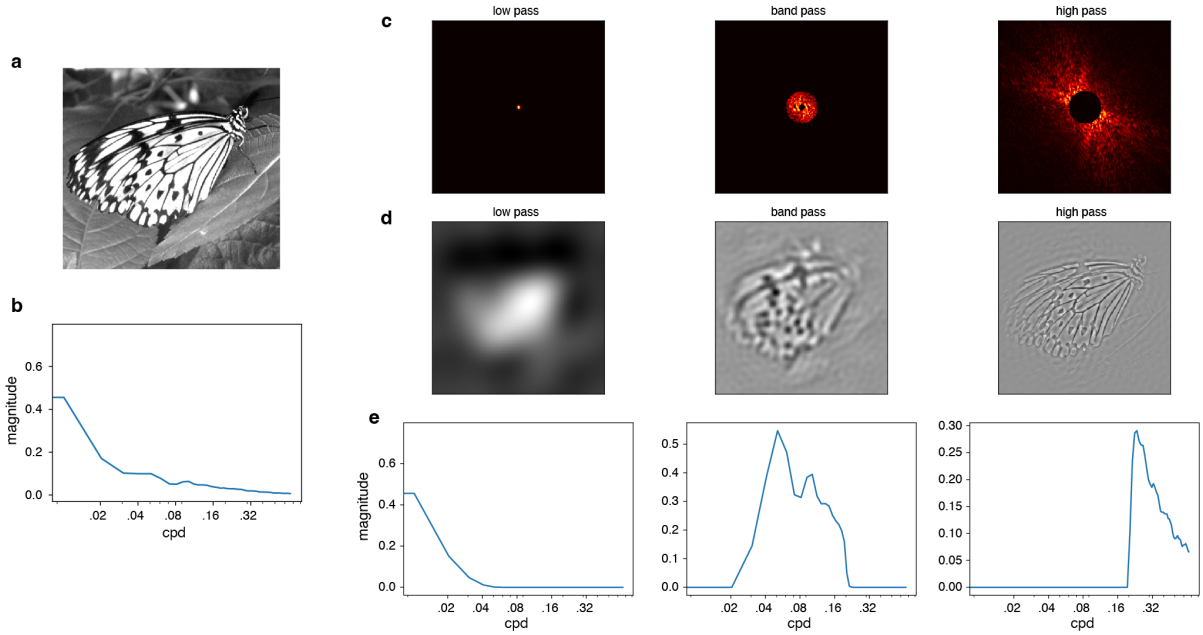

Fig. S7: **Filtering natural scenes.** **a**, An example natural scene and **b** its corresponding spatial frequency spectrum (azimuthal average), in cycles/degree (cpd) of visual angle. **c**, Images were filtered by cropping their spectra in Fourier domain (see Methods). **d** Resulting filtered images and **e** their corresponding spectra (azimuthal averages) after filtering was performed.

References

[1] E. Acar, D. M. Dunlavy, and T. G. Kolda. A scalable optimization approach for fitting canonical tensor decompositions. *J Chemom*, 25(2):67–86, 2011. doi: 10.1002/cem.1335.

[2] H. Adesnik and M. Scanziani. Lateral competition for cortical space by layer-specific horizontal circuits. *Nature*, 464(7292):1155–1160, 2010.

[3] H. Adesnik, W. Bruns, H. Taniguchi, Z. J. Huang, and M. Scanziani. A neural circuit for spatial summation in visual cortex. *Nature*, 490(7419):226–231, 2012.

[4] E. A. Allen and R. D. Freeman. Dynamic spatial processing originates in early visual pathways. *Journal of Neuroscience*, 26(45):11763–11774, 2006.

[5] Allen Institute. Visual coding – neuropixels. `https://portal.brain-map.org/circuits-behavior/visual-coding-neuropixels`, . Accessed: 04–Oct–2024.

[6] Allen Institute. Allen brain atlas software development kit (allensdk). `https://allensdk.readthedocs.io/en/latest/visual_coding_neuropixels.html`, . Accessed: 04–Oct–2024.

[7] Allen Institute. Allen Brain Observatory – Neuropixels Visual Coding. `https://brainmapportal-live-4cc80a57cd6e400d854-f7fdcae.divio-media.net/filer_public/80/75/8075a100-ca64-429a-b39a-569121b612b2/neuropixels_visual_coding_-_white_paper_v10.pdf`, 2019. Version 1.0 (October 1, 2019) [Online; accessed 04–Oct–2024].

[8] J.-M. Alonso, W. M. Usrey, and R. C. Reid. Precisely correlated firing in cells of the lateral geniculate nucleus. *Nature*, 383(6603):815–819, 1996.

[9] B. W. Bader and T. G. Kolda. Efficient MATLAB computations with sparse and factored tensors. *SIAM J Sci Comput*, 30(1):205–231, 2008.

[10] A. Blot, M. M. Roth, I. Gasler, M. Javadzadeh, F. Imhof, and S. B. Hofer. Visual intracortical and transthalamic pathways carry distinct information to cortical areas. *Neuron*, 109(12):1996–2008, 2021.

[11] F. Bolaños, J. G. Orlandi, R. Aoki, A. V. Jagadeesh, J. L. Gardner, and A. Benucci. Efficient coding of natural images in the mouse visual cortex. *Nature Communications*, 15(1):2466, 2024.

[12] G. Born, F. A. Schneider-Soupiadis, S. Erisken, A. Vaiceliunaite, C. L. Lao, M. H. Mobarhan, M. A. Spacek, G. T. Einevoll, and L. Busse. Corticothalamic feedback sculpts visual spatial integration in mouse thalamus. *Nature neuroscience*, 24(12):1711–1720, 2021.

[13] Z. I. Botev, J. F. Grotowski, and D. P. Kroese. Kernel density estimation via diffusion. *Ann Stat*, 38 (5):2916–2957, 2010.

[14] C. E. Bredfeldt and D. Ringach. Dynamics of spatial frequency tuning in macaque v1. *Journal of Neuroscience*, 22(5):1976–1984, 2002.

[15] R. Bro and S. De Jong. A fast non-negativity-constrained least squares algorithm. *J Chemom*, 11(5): 393–401, 1997.

[16] S. Chung and L. Abbott. Neural population geometry: An approach for understanding biological and artificial neural networks. *Curr Opin Neurobiol*, 70:137–144, 2021.

[17] A. Cichocki, R. Zdunek, A. H. Phan, and S.-i. Amari. *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.

[18] R. R. Coifman and S. Lafon. Diffusion maps. *Appl Comput Harmon Anal*, 21(1):5–30, 2006.

[19] R. R. Coifman, S. Lafon, A. B. Lee, M. Maggioni, B. Nadler, F. Warner, and S. W. Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proc Natl Acad Sci USA*, 102(21):7426–7431, 2005.

[20] S. E. de Vries, J. A. Lecoq, M. A. Buice, P. A. Groblewski, G. K. Ocker, M. Oliver, D. Feng, N. Cain, P. Ledochowitsch, D. Millman, et al. A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nature neuroscience*, 23(1):138–151, 2020.

[21] S. E. de Vries, J. H. Siegle, and C. Koch. Sharing neurophysiology data from the allen brain observatory. *Elife*, 12:e85550, 2023.

[22] A. Dobbins, S. W. Zucker, and M. S. Cynader. Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature*, 329(6138):438, 1987.

[23] L. Dyballa and S. W. Zucker. IAN: Iterated Adaptive Neighborhoods for manifold learning and dimensionality estimation. *Neural Comput*, 35, 2023.

[24] L. Dyballa, M. S. Hoseini, M. C. Dadarlat, S. W. Zucker, and M. P. Stryker. Flow stimuli reveal ecologically appropriate responses in mouse visual cortex. *Proc Natl Acad Sci USA*, 115(44):11304–11309, 2018.

[25] L. Dyballa, A. M. Rudzite, M. S. Hoseini, M. Thapa, M. P. Stryker, G. D. Field, and S. W. Zucker. Population encoding of stimulus features along the visual hierarchy. *Proceedings of the National Academy of Sciences*, 121(4):e2317773121, 2024.

[26] G. Felsen and Y. Dan. A natural approach to studying vision. *Nature neuroscience*, 8(12):1643–1646, 2005.

[27] T. Flossmann and N. L. Rochefort. Spatial navigation signals in rodent visual cortex. *Current opinion in neurobiology*, 67:163–173, 2021.

[28] E. Froudarakis, P. Berens, A. S. Ecker, R. J. Cotton, F. H. Sinz, D. Yatsenko, P. Saggau, M. Bethge, and A. S. Tolias. Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. *Nat Neurosci*, 17(6):851, 2014.

[29] E. Froudarakis, P. Berens, A. S. Ecker, R. J. Cotton, F. H. Sinz, D. Yatsenko, P. Saggau, M. Bethge, and A. S. Tolias. Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. *Nat Neurosci*, 17(6):851–857, 2014.

[30] L. L. Glickfeld and S. R. Olsen. Higher-order areas of the mouse visual cortex. *Annual review of vision science*, 3:251–273, 2017.

[31] Y. Han, J. M. Kebschull, R. A. Campbell, D. Cowan, F. Imhof, A. M. Zador, and T. D. Mrsic-Flogel. The logic of single-cell projections from visual cortex. *Nature*, 556(7699):51–56, 2018.

[32] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant. Array programming with NumPy. *Nature*, 585(7825): 357–362, Sept. 2020. doi: 10.1038/s41586-020-2649-2. URL `https://doi.org/10.1038/s41586-020-2649-2`.

[33] J. A. Harris, S. Mihalas, K. E. Hirokawa, J. D. Whitesell, H. Choi, A. Bernard, P. Bohn, S. Caldejon, L. Casal, A. Cho, et al. Hierarchical organization of cortical and thalamic connectivity. *Nature*, 575 (7781):195–202, 2019.

[34] J. Hegdé. Time course of visual perception: Coarse-to-fine processing and beyond. *Progress in Neurobiology*, 84:405–439, 2008. URL `https://api.semanticscholar.org/CorpusID:900253`.

[35] J. Hegdé. Time course of visual perception: coarse-to-fine processing and beyond. *Progress in neurobiology*, 84(4):405–439, 2008.

[36] J. L. Hoy, I. Yavorska, M. Wehr, and C. M. Niell. Vision drives accurate approach behavior during prey capture in laboratory mice. *Curr Biol*, 26(22):3046–3052, 2016.

[37] S. Katzner, G. Born, and L. Busse. V1 microcircuits underlying mouse visual behavior. *Current opinion in neurobiology*, 58:191–198, 2019.

[38] A. J. Keller, M. M. Roth, and M. Scanziani. Feedback generates a second receptive field in neurons of the visual cortex. *Nature*, 582(7813):545–549, 2020.

[39] E. J. Kim, A. L. Juavinett, E. M. Kyubwa, M. W. Jacobs, and E. M. Callaway. Three types of cortical layer 5 neurons that differ in brain-wide connectivity and function. *Neuron*, 88(6):1253–1267, 2015.

[40] M.-H. Kim, P. Znamenskiy, M. F. Iacaruso, and T. D. Mrsic-Flogel. Segregated subnetworks of intracortical projection neurons in primary visual cortex. *Neuron*, 100(6):1313–1321, 2018.

[41] L. Kirchberger, S. Mukherjee, U. H. Schnabel, E. H. van Beest, A. Barsegyan, C. N. Levelt, J. A. Heimel, J. A. Lorteije, C. van der Togt, M. W. Self, et al. The essential role of recurrent processing for figure-ground perception in mice. *Science advances*, 7(27):eabe1833, 2021.

[42] M. A. Kirchgessner, A. D. Franklin, and E. M. Callaway. Distinct "driving" versus "modulatory" influences of different visual corticothalamic pathways. *Current Biology*, 31(23):5121–5137, 2021.

[43] T. G. Kolda. Multilinear operators for higher-order decompositions. Technical Report SAND2006-2081, Sandia National Laboratories, April 2006. URL `http://www.osti.gov/scitech/biblio/923081`.

[44] D. D. Lee and H. S. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[45] F. J. Luongo, L. Liu, C. L. A. Ho, J. K. Hesse, J. B. Wekselblatt, F. F. Lanfranchi, D. Huber, and D. Y. Tsao. Mice and primates use distinct strategies for visual segmentation. *Elife*, 12:e74394, 2023.

[46] J. Manley, S. Lu, K. Barber, J. Demas, H. Kim, D. Meyer, F. M. Traub, and A. Vaziri. Simultaneous, cortex-wide dynamics of up to 1 million neurons reveal unbounded scaling of dimensionality with neuron number. *Neuron*, 2024.

[47] T. Marques, M. T. Summers, G. Fioreze, M. Fridman, R. F. Dias, M. B. Feller, and L. Petreanu. A role for mouse primary visual cortex in motion perception. *Current Biology*, 28(11):1703–1713, 2018.

[48] D. Marr and T. Poggio. A computational theory of human stereo vision. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 204(1156):301–328, 1979.

[49] J. H. Marshel, M. E. Garrett, I. Nauhaus, and E. M. Callaway. Functional specialization of seven mouse visual cortical areas. *Neuron*, 72(6):1040–1054, 2011.

[50] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proceedings of the 8th IEEE International Conference on Computer Vision (ICCV 2001)*, volume 2, pages 416–423. IEEE, 2001.

[51] J. A. Mazer, W. E. Vinje, J. McDermott, P. H. Schiller, and J. L. Gallant. Spatial frequency and orientation tuning dynamics in area v1. *Proceedings of the National Academy of Sciences*, 99(3): 1645–1650, 2002.

[52] J. S. Montijn, P. M. Goltstein, and C. M. Pennartz. Mouse v1 population correlates of visual detection rely on heterogeneity within neuronal response patterns. *Elife*, 4:e10163, 2015.

[53] P. Murphy and A. Sillito. Corticofugal feedback influences the generation of length tuning in the visual pathway. *Nature*, 329(6141):727–729, 1987.

[54] C. M. Niell and M. Scanziani. How cortical circuits implement cortical computations: mouse visual cortex as a model. *Annual Review of Neuroscience*, 44:517–546, 2021.

[55] T. Odland. KDEpy: Kernel Density Estimation in Python, Dec. 2018. v0.9.10. Available at https://doi.org/10.5281/zenodo.2392268.

[56] A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, 42:145–175, 2001.

[57] A. Olmos and F. A. Kingdom. A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, 33(12):1463–1473, 2004.

[58] R. V. Rikhye and M. Sur. Spatial correlations in natural scenes modulate response reliability in mouse visual cortex. *J Neurosci*, 35(43):14661–14680, 2015.

[59] D. L. Ruderman and W. Bialek. Statistics of natural images: Scaling in the woods. *Physical review letters*, 73(6):814, 1994.

[60] P. G. Schyns and A. Oliva. From blobs to boundary edges: Evidence for time-and spatial-scale-dependent scene recognition. *Psychological science*, 5(4):195–200, 1994.

[61] M. W. Self, J. A. Lorteije, J. Vangeneugden, E. H. van Beest, M. E. Grigore, C. N. Levelt, J. A. Heimel, and P. R. Roelfsema. Orientation-tuned surround suppression in mouse visual cortex. *J Neurosci*, 34 (28):9290–9304, 2014.

[62] J. H. Siegle, X. Jia, S. Durand, S. Gale, C. Bennett, N. Graddis, G. Heller, T. K. Ramirez, H. Choi, J. A. Luviano, et al. Survey of spiking in the mouse visual system reveals functional hierarchy. *Nature*, pages 1–7, 2021.

[63] A. M. Sillito, H. E. Jones, G. L. Gerstein, and D. C. West. Feature-linked synchronization of thalamic relay cell firing induced by feedback from the visual cortex. *Nature*, 369(6480):479–482, 1994.

[64] E. P. Simoncelli and B. A. Olshausen. Natural image statistics and neural representation. *Annual review of neuroscience*, 24(1):1193–1216, 2001.

[65] K. K. Sit and M. J. Goard. Distributed and retinotopically asymmetric processing of coherent motion in mouse visual cortex. *Nature communications*, 11(1):3565, 2020.

[66] R. Skyberg, S. Tanabe, H. Chen, and J. Cang. Coarse-to-fine processing drives the efficient coding of natural scenes in mouse visual cortex. *Cell reports*, 38(13), 2022.

[67] R. J. Skyberg and C. M. Niell. Natural visual behavior and active sensing in the mouse. *Current Opinion in Neurobiology*, 86:102882, 2024.

[68] I. T. Smith, L. B. Townsend, R. Huh, H. Zhu, and S. L. Smith. Stream-dependent development of higher visual cortical areas. *Nature neuroscience*, 20(2):200–208, 2017.

[69] C. Stringer, M. Michaelos, D. Tsyboulski, S. E. Lindo, and M. Pachitariu. High-precision coding in visual cortex. *Cell*, 184(10):2767–2778, 2021.

[70] E. Switkes, M. J. Mayer, and J. A. Sloan. Spatial frequency analysis of the visual environment: Anisotropy and the carpentered environment hypothesis. *Vision research*, 18(10):1393–1399, 1978.

[71] v. A. Van der Schaaf and J. v. van Hateren. Modelling the power spectra of natural images: statistics and information. *Vision research*, 36(17):2759–2770, 1996.

[72] J. H. Van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1394):359–366, 1998.

[73] M. Vélez-Fort, C. V. Rousseau, C. J. Niedworok, I. R. Wickersham, E. A. Rancz, A. P. Brown, M. Strom, and T. W. Margrie. The stimulus selectivity and connectivity of layer six principal cells reveals cortical microcircuits underlying visual processing. *Neuron*, 83(6):1431–1443, 2014.

[74] S. Vreysen, B. Zhang, Y. M. Chino, L. Arckens, and G. Van den Bergh. Dynamics of spatial frequency tuning in mouse visual cortex. *Journal of neurophysiology*, 107(11):2937–2949, 2012.

[75] H. Wang, O. Dey, W. N. Lagos, and E. M. Callaway. Diversity in spatial frequency, temporal frequency, and speed tuning across mouse visual cortical areas and layers. *Journal of Comparative Neurology*, 530(18):3226–3247, 2022.

[76] Q. Wang, O. Sporns, and A. Burkhalter. Network analysis of corticocortical connections reveals ventral and dorsal processing streams in mouse visual cortex. *Journal of Neuroscience*, 32(13):4386–4399, 2012.

[77] F. A. Wichmann, J. Drewes, P. Rosas, and K. R. Gegenfurtner. Animal detection in natural scenes: Critical features revisited. *Journal of Vision*, 10(4):6–6, 2010.

[78] A. H. Williams, T. H. Kim, F. Wang, S. Vyas, S. I. Ryu, K. V. Shenoy, M. Schnitzer, T. G. Kolda, and S. Ganguli. Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. *Neuron*, 98(6):1099–1115, 2018.

[79] H. Young, B. Belbut, M. Baeta, and L. Petreanu. Laminar-specific cortico-cortical loops in mouse visual cortex. *Elife*, 10:e59551, 2021.

[80] Y. Yu, J. N. Stirman, C. R. Dorsett, and S. L. Smith. Selective representations of texture and motion in mouse higher visual areas. *Current Biology*, 32(13):2810–2820, 2022.

[81] Y. Yu, J. N. Stirman, C. R. Dorsett, and S. L. Smith. Visual information is broadcast among cortical areas in discrete channels. *bioRxiv*, pages 2023–12, 2023.

[82] H. Zhou, R. J. Schafer, and R. Desimone. Pulvinar-cortex interactions in vision and attention. *Neuron*, 89(1):209–220, 2016.