bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Variations in neuronal selectivity create efficient representational geometries for perception

Sonica Saraf[1], J. Anthony Movshon[1], and SueYeon Chung[1,2]

[1]Center for Neural Science, New York University
[2]Center for Computational Neuroscience, Flatiron Institute

Keywords: vision, psychophysical discrimination, psychophysical identification, neural geometry, population coding

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Significance

The visual system supports many perceptual tasks, such as discriminating and identifying specific visual stimuli. This flexibility is supported by the coordinated responses of neurons in the visual cortex. What neuronal response properties affect our perceptual abilities? Here, we show how two types of response diversity shape the geometry of neural responses to improve the population code for two types of perceptual tasks. Each diversity type creates a different geometric transformation. One transformation improves our ability to discriminate similar targets, while the other improves our ability to find one target among many distractors.

Abstract

Our visual capabilities depend on neural response properties in visual areas of our brains. Neurons exhibit a wide variety of selective response properties, but the reasons for this diversity are unknown. Here, we related the distribution of neuronal tuning properties to the information capacity of the population. Our results from theory, simulations, and analysis of recordings from macaque primary visual cortex (V1) reveal that diversity of amplitude and bandwidth drive complementary changes to the representational geometry of a population. Amplitude diversity pushes the centers of the representations further apart, whereas bandwidth heterogeneity decorrelates the center locations. These geometric changes separate out representations for distinct stimuli, creating more efficient encoding. We study how both types of diversity affect the population code for two different perceptual tasks: discrimination and identification. While both types of diversity improve encoding for both tasks, their distinct impacts on geometry make each more beneficial for one of the two tasks. Amplitude diversity impacts coding efficiency more for discrimination than it does for identification, while bandwidth diversity has a stronger impact on identification. These complementary effects indicate the importance of both types of diversity for perception. Finally, because tuning diversity exists across species and brain areas, our results suggest a fundamental neural coding strategy that may be applicable to a wide range of behavior.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Introduction

Across species and brain areas, neurons exhibit varying levels of stimulus selectivity. Some are broadly responsive, while others are more specific about the stimuli to which they respond (Eliav et al., 2021; Fitzgerald et al., 2006; Goris et al., 2015; Ringach et al., 2002). What purpose does this diversity serve? Does it improve neural populations' encoding of sensory stimuli (Goris et al., 2015; Kang et al., 2004)? Previous work by Shamir and Sompolinsky (2006) and Ecker et al. (2011) has used analytical methods to study how amplitude diversity affects coding efficiency for fine discrimination tasks. Here, we use a new analytical framework to study coding efficiency for different tasks and types of tuning diversity.

We suppose that a neural population's coding efficiency is related to the geometric structure of its responses to stimuli, termed a population's representational geometry. Population responses are represented in a high-dimensional space where the axes each correspond to a neuron's firing rate. A point in this space represents the population's response to a presentation of a stimulus. Points from various trials or slight variations in the stimulus (depending on the scientific question at hand) are grouped together and form a representation encoding the stimulus (DiCarlo and Cox, 2007; Roweis and Saul, 2000; Tenenbaum et al., 2000; Yerxa et al., 2023). Analyzing the geometry of these representations has emerged as a powerful technique for understanding population coding in many brain areas (Bernardi et al., 2020; Chung and Abbott, 2021; DiCarlo and Cox, 2007; Genkin et al., 2023; Mehrpour et al., 2021; Tian et al., 2024; Wakhloo et al., 2024). More recently, an approach has emerged that focuses on the connection between the high-dimensional geometry of neural representations and the performance of a downstream decoder, termed the *classification capacity* (also referred to as *capacity*) (Chung et al., 2018). The decoding assumption for capacity is discrimination – the downstream populations measure the separation between the neural representations for various stimuli.

Inspired by this framework, we revisit the question of why tuning diversity is beneficial to population coding, using data from macaque V1 and simulated neural populations. We use previously established connections between geometry and capacity (Chou et al., 2025; Chung et al., 2016, 2018; Wakhloo et al., 2023) to relate the geometric transformations due to diversity to improvements in coding efficiency.

Specifically, we consider diversity in two response properties, amplitude and bandwidth, and its impact on coding efficiency for two different perceptual tasks, discrimination and identification. *Fine* discrimination involves detecting differences between stimuli, such as distinguishing a vertically- and off-vertically oriented stimulus. Coding efficiency for fine discrimination has been studied both theoretically and experimentally (Green et al., 1966; Jazayeri and Movshon, 2006; Seung and Sompolinsky, 1993) because it helps us understand the limits of our perception. We also consider identification – when one stimulus must be distinguished from all other potential stimuli. Identification has a special role in everyday vision: for example, reaching for the correct utensil, searching for a familiar face in a crowd, reading. Our results establish that both amplitude and bandwidth diversity increase the efficiency of population codes for both discrimination and identification. However, they create distinct geometric changes, and each diversity type improves efficiency more for one of the two tasks than the other. Amplitude diversity has more impact on discrimination, while bandwidth diversity has more impact on identification.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Results

Geometric framework

We examined how well the responses of populations of real and simulated V1 neurons can be used to discriminate the orientation of drifting sinusoidal gratings. We term the effectiveness of the population in discriminating stimuli its *efficiency*. Our goal was to use population geometry-based analyses to study how tuning diversity impacts efficiency, so we defined population-level representations in the same experimental paradigm used for analyzing tuning curves, illustrated in Figure 1. The left section of Figure 1A shows the experimental design. Briefly, monkeys were shown 50 trials each of 72 drifting gratings that differed in their orientation (0 deg, 5 deg, ..., 175 deg). Recordings were made with Utah arrays placed in V1 of 5 anesthetized macaque monkeys (Graf et al., 2011). Our simulated population responses followed a similar paradigm – see *Methods* for further detail. The right side of Figure 1A shows schematics of neural tuning curves of various amplitudes and bandwidths, illustrative of the diversity observed in V1 (Ringach et al., 2002).

We define the neural representations using the population's trial by trial responses to the stimuli (Figure 1B). The neural representation (or *neural manifold*) encodes an orientation and is created by collecting the population activity for every trial where one orientation was presented. If P different orientations are presented for S trials each, there are P different manifolds, each containing S points. The shapes of the manifolds are determined by trial-to-trial variability and covariability in the population, because each manifold encodes a single orientation. The neural population's coding efficiency, referred to as its *capacity*, is determined by how well each of the P manifolds can be separated from any nearby manifold (discrimination) or all of the other manifolds (identification) with a hyperplane. See panels A and B of Figure 2 for a depiction of discrimination and identification. Capacity is the ratio of the number of manifolds to the minimum number of neurons needed to achieve a threshold level of separability (Chou et al., 2025; Chung et al., 2018). See *Methods* for details on measuring capacity.

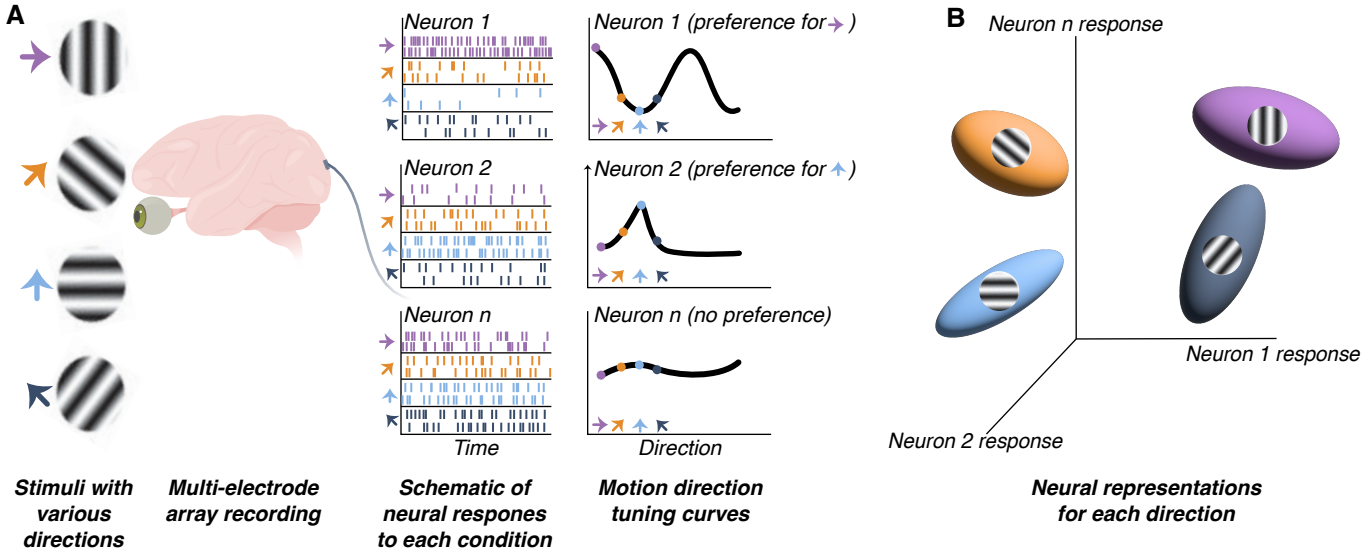


Figure 1. Schematic of tuning curves and neural manifolds. **A** Experimental procedure, simulated rasters and tuning curves. Neural data were previously reported in Graf et al. (2011), and we made analogous measurements in our simulations. A neuron's average response to each condition gives a point on the neuron's tuning curve. **B** The population's responses to each stimulus condition form the neural manifold encoding that stimulus. A point on the manifold represents the population's response to one trial, and the centers of the manifolds are the average responses to each stimuli (i.e. the values on the tuning curves).

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

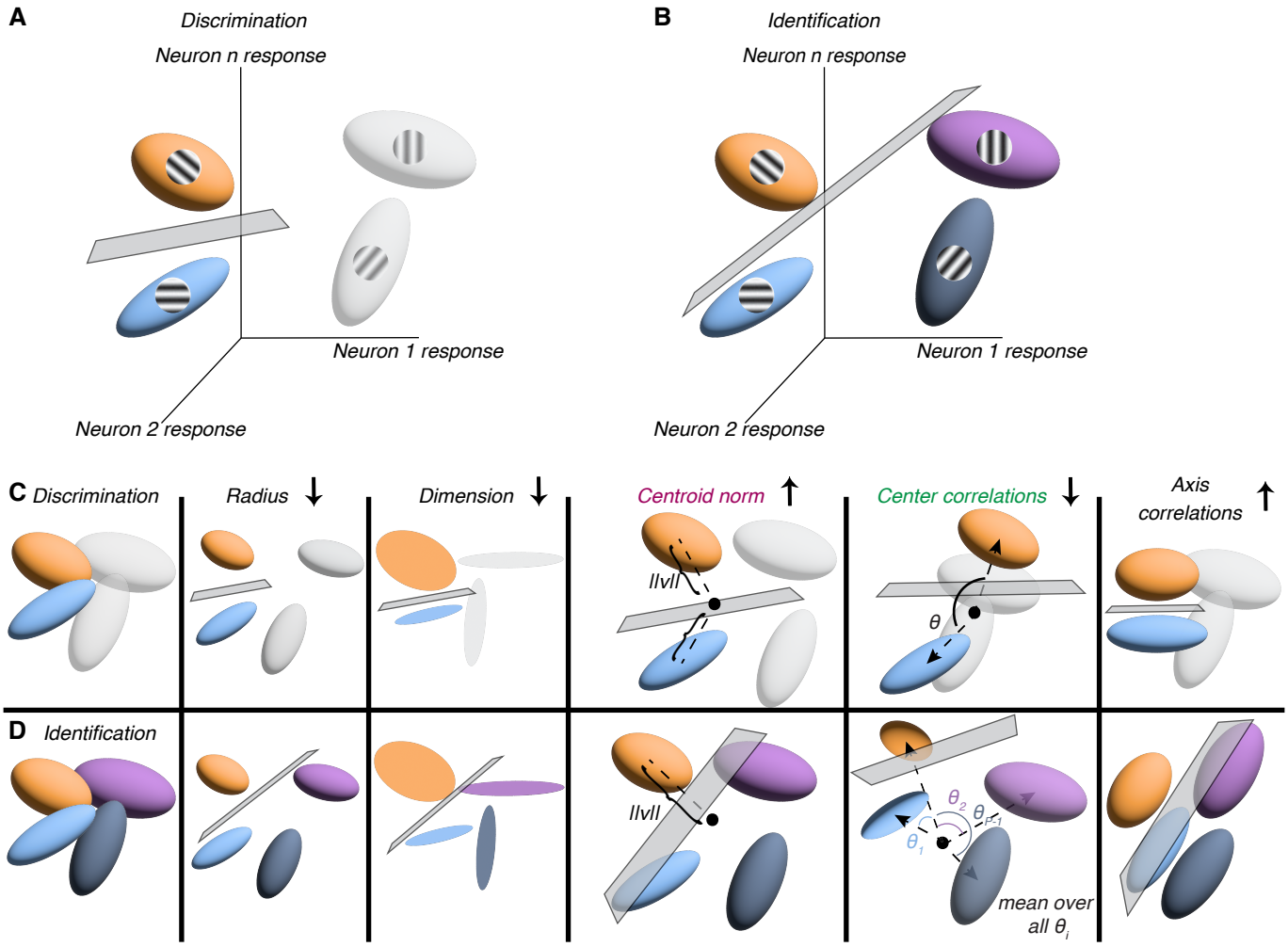


Figure 2. Tasks and geometric changes that increase capacity. **A** Discrimination task. Capacity, a measure of neural coding efficiency, is measured by linearly separating the manifolds for neighboring stimuli. **B** Identification task. Here, capacity is measured by linearly separating each manifold from the other $P-1$ manifolds. **C** Geometric properties and their relations to capacity for both tasks. The radius and dimension capture the overall size of the two manifolds being separated. The radius measures the average distance between a manifold's boundary points and its center. The dimension measures the average projection of random vectors onto the manifold's boundary points. The centroid norm is the average Euclidean distance ($||v||$), between the center of each of the two manifolds and the global mean of all of the manifolds, which is indicated by the black dot. Center correlations measure the cosine similarity between the centers of the two manifolds. Finally, the axis correlations capture the similarity of the principal axes of the two manifolds. **D** Analogous to panel C, but for the identification task. Here, the radius, dimension, and centroid norms are defined as for the discrimination task, except instead of averaging over the measures for the two manifolds to discriminate, we report the values for only the target manifold. The black dot indicates the global mean of all manifolds. Center correlations are the average of the cosine similarities between the target manifold's center and every other manifold's center. The axis correlations are an average of the similarity of the target manifold's axes and all other manifold's axes. Throughout the paper we report the averages of the geometric measures over all possible fine discrimination and identification tasks.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Geometric properties that relate to capacity

First consider the discrimination task. Radius and dimension are properties characterizing the average size of each individual manifold, defined here by trial to trial response variation. Smaller manifolds are less likely to overlap, so they are easier to separate from each other. Thus, lower radii and dimensions increase capacity. Three geometric properties capture the spatial relationship among manifolds. The centroid norm is a measure of the distance between a manifold's mean point (or "center") and the global mean of the neural responses to all P stimuli. A higher centroid norm increases separability (Chou et al., 2025; Chung et al., 2018). The center correlations measure the similarity between the manifolds' center vector directions and lower center correlations increase separability. Finally, axis correlations quantify the similarity in the manifolds' principal components and higher axis correlations increase separability (Chou et al., 2025; Chung et al., 2018; Wakhloo et al., 2023). Figure 2C shows visualizations of how the geometric properties affect capacity for discrimination.

An analogous set of geometric measures exists for the identification task, which is implemented as a discrimination of one target from many. We define the radius, dimension, and centroid norm in the same way as for discrimination, except that instead of averaging the measures over the two manifolds being separated, we only consider the target manifold's measures. Center and axis correlations are the average pairwise center and axis correlations, respectively, between the target and all other manifolds.

There are several ways to define manifold geometry, each with distinct advantages and limitations. Some measures focus on the "worst-case" points – the points on the manifolds that determine separability (Chou et al., 2025; Chung et al., 2018). Here we use geometric definitions that characterize the full structure of point-cloud representations, to facilitate intuitive connections between tuning properties and the geometry of the neural population responses in high-dimensional activity space.

Table 1. Multiple regression results for simulations. We show the standardized coefficients for each of the geometric properties as predictors of capacity.

	Amplitude diversity		**Bandwidth diversity**	
Geometric property	Discrimination	Identification	Discrimination	Identification
Centroid norm	1.31	1.56	0.22	0.21
Center correlations	-0.34	-0.45	-1.19	-1.18
Radius	-0.048	-0.31	-0.057	-0.14
Dimension	-0.037	-0.15	-0.071	-0.045
Axis correlations	0.028	-0.080	0.010	0.076

Amplitude and bandwidth diversity for discrimination and identification

We simulated populations of neurons and observed how diversity affects manifold geometry and coding efficiency. We followed the approach of (Shamir and Sompolinsky, 2006): we modeled neural tuning curves as von Mises functions. The preferred stimuli of the neurons were evenly spaced across orientations (180 deg). In the homogeneous populations, the tuning curves for all cells had the same amplitude and bandwidth. To create populations with diverse amplitudes, we randomly chose each neuron's tuning curve amplitude from a lognormal distribution whose mean matched the homogeneous population. To implement higher levels of diversity, we increased the variance of the lognormal distribution. We chose lognormally distributed amplitudes to match what is experimentally observed in macaque V1 (see the distribution of amplitudes in Figure 5A and SI Figure 2). Amplitude variance ranged from 0 to 97 $(\text{spikes/s})^2$. We selected this range of amplitude variance to match the variance exhibited in our analyzed neural data populations as well as possible (the amplitude variances of the four datasets were 110, 154, 528, and 74 $(\text{spikes/s})^2$). Because finite samples from distributions with long tails are vulnerable to the mean scaling with variance, we also ensured that the variability range allowed our samples to have constant mean even at the highest

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

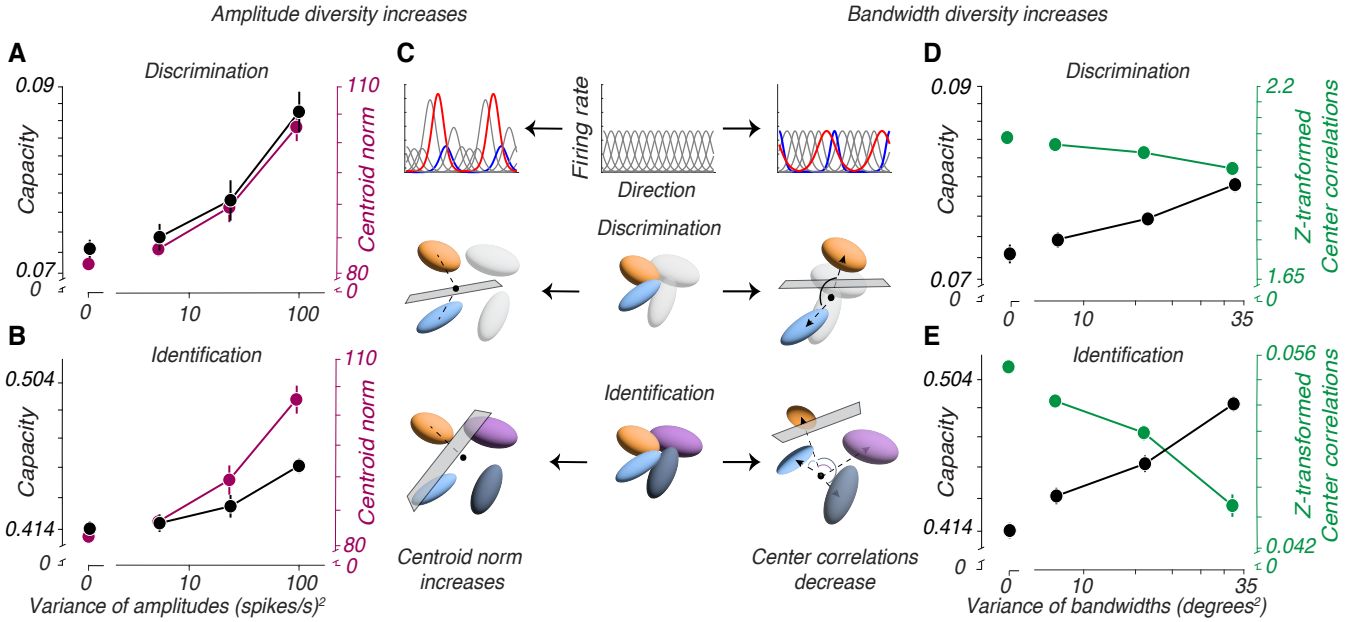


Figure 3. Amplitude and bandwidth diversity increase capacity for discrimination and identification via distinct geometric transformations. **A** Capacity and centroid norm for populations with varying levels of amplitude diversity for the discrimination task. We ran 10 trials each of different realizations of the population and its responses for four levels of amplitude variability in the population (0, 5, 23, and 97 $(\text{spikes/s})^2$. An amplitude variability of 0 corresponds to populations that are homogeneous in their amplitudes. Error bars indicate the standard deviation over the 10 trials. As the amplitude diversity increases, the capacity increases. This increase is due to larger centroid norms in the neural representations. **B** Analogous to panel A, but for identification. The range of the axis for capacity values is 0.405 to 0.5202 to match the ratio of increase in the capacity axis of panel A. This allows for the fractional increases across the tasks to be easily compared. Note that capacity increases more for discrimination than identification, indicating that larger centroid norms do not help identification as much as discrimination. **C** The left side of the upper panel is a schematic of tuning curves with amplitude diversity. As amplitude diversity increases, the centroid norms of the manifolds increase. See *Supporting Information* for a mathematical proof. The right side of the upper panel is a schematic of tuning curves with bandwidth diversity. As bandwidth diversity increases, the center correlations of the manifolds decrease. **D** Capacity and center correlations for populations with varying levels of bandwidth diversity for the discrimination task. We carried out an analogous set of simulations to the ones used in panel A. The four levels of bandwidth diversity were 0, 8, 16, and 32 deg^2. Error bars indicate the standard deviation over the 10 trials. Higher diversity of bandwidths increases the capacity of the population, supported by a reduction of center correlations. We applied Fisher's Z-transform to the correlations to enable comparisons among correlation values. **E** Analogous to panel D, but for identification. The ranges shown on the center correlations axes are in equivalent ratios. Note that capacity increases more for identification than for discrimination. This is supported by a steeper decrease in center correlations for the identification task.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

level of diversity. We introduced bandwidth diversity into neural populations in a similar fashion, but used gamma distributed bandwidths – again, to mimic what is experimentally observed (see Figure 5A and SI Figure 2). Bandwidth variance ranged from 0 to 32 deg^2. As with the range of variances for amplitude diversity, we selected an experimentally reasonable range for bandwidth variance that also allowed us to keep the mean bandwidth constant across the levels of diversity (the total bandwidth variances for the four arrays were 25, 69, 84, and 239 deg^2).

We simulated the responses of populations of 300 neurons to 50 trials each of 36 different oriented drifting gratings (5 deg spacing between stimuli). We assumed a multiplicative noise model for each cell – meaning that the response variance increases linearly as the mean response rate increased, and a population noise correlation that decayed with the difference in the preferred directions for a pair of cells (Shamir and Sompolinsky, 2006).

For each population, we measured the geometry and coding efficiency for discrimination and identification. We analyzed: 1) the relationship between diversity and coding efficiency for each perceptual task, 2) the relationship between diversity and the geometric properties for each task, and 3) the geometric properties that explained the changes in coding efficiency.

For the first, we found that both amplitude and bandwidth diversity improve coding efficiency for discrimination and identification (Figure 3A, B, D, and E). However, amplitude diversity impacts discrimination more, while bandwidth diversity impacts identification more. There was a 20% increase in capacity for discrimination due to amplitude diversity. This increase corresponds to 0.015 additional separable manifolds per neuron, which indicates that the diverse population of 300 neurons can separate 4.4 more manifolds than the homogeneous one (see Figure 3A). Capacity also increased with amplitude diversity for identification, but only by 8.8% (an increase of 0.036 manifolds/neuron, or 10.8 manifolds for the population; see Figure 3B). Note that capacity for identification is higher than for discrimination because there are 36 manifolds being separated rather than 2, and some of these are much easier to separate from the target manifold. We performed a multiple regression among the five geometric properties and capacity to infer the geometric changes due to diversity that improve capacity. We first ensured that there were no optimal non-linear relationships between each property and capacity by using Tukey's ladder of transformations. The coefficients for the multiple regression are shown in Table 1. For both discrimination and identification, the centroid norm predicts the capacity increase due to amplitude diversity. The center correlations also have some effect on capacity. However, center correlations increase with amplitude diversity (see SI Figure 1A) and they have a negative coefficient, which indicates that they slightly decrease capacity.

In Figure 3A and B, we show the change in the centroid norm due to diversity. There was a 26% increase in the centroid norms of the manifolds, which indicates that amplitude diversity moves the manifold center locations further from each other (see *Supporting Information* for a formal proof of why the centroid norm increases with amplitude diversity). The centroid norm changes in the same way for discrimination and identification, but the capacity increase for discrimination is greater, suggesting that larger centroid norms are not as helpful for identification. Consistent with this, the multiple regression revealed that the change in centroid norm predicted only a 14% capacity increase for identification, but 27% for discrimination.

There was a 9.4% increase in capacity for discrimination due to bandwidth diversity (0.0068 additional separable manifolds/neuron, 2 more manifolds for the entire population; see Figure 3D). Capacity increases more for identification – by 18% (0.074 manifolds/neuron, 22.1 more manifolds for the population; see Figure 3E). We again performed a multiple regression with the five geometric properties as explanatory variables for capacity (see Table 1). For both discrimination and identification the center correlations predict the capacity increase, indicating that bandwidth diversity decorrelates the manifold centers to improve coding efficiency. In addition, the centroid norm decreases (see SI Figure 1B) which slightly decreases capacity.

In Figure 3D and E, we show the change in the center correlations due to diversity. There was a 4.5% and 19% decrease in center correlations for discrimination and identification, respectively. This suggests that bandwidth diversity affects identification more because of larger decorrelations among all manifold centers compared to just the nearest pairs.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Comparing amplitude and bandwidth diversity

Altogether, our simulation results reveal that amplitude and bandwidth diversity improve coding efficiency through distinct transformations to representational geometry. For both discrimination and identification, both types of diversity increase the capacity. However, each diversity type impacts one of the two tasks more: Amplitude diversity improves discrimination more, while bandwidth diversity improves identification more (compare Figure 4A and B).

We have only considered amplitude and bandwidth diversity separately. However, biological neural populations exhibit both types of diversity which raises the question of how amplitude and bandwidth diversity interact with one another. We studied this by measuring the capacity for a population with amplitude variance (97 (spikes/s)^2) and bandwidth variance (32 deg^2) – the highest levels of diversity used in the simulations depicted in Figure 3. We compared the capacity of this population with the capacities of the populations with only amplitude or bandwidth diversity. We found an additive relationship between capacity improvements due to amplitude and bandwidth diversity for both discrimination and identification (Figure 4C, which suggests that populations should have both types to maximize their coding efficiencies.

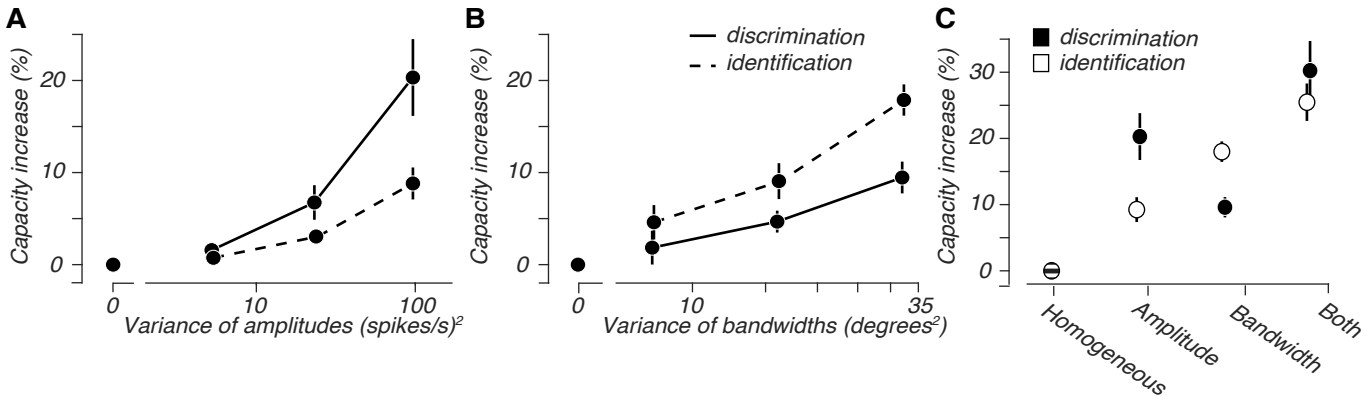


Figure 4. How tuning diversity affects capacity for discrimination and identification. Panels A and B replot the data from Figure 3, but with a different axis (percentage increase of capacity) to show comparisons between different capacity trends. **A** Amplitude diversity increases capacity more for discrimination than for identification. Solid line shows the percentage increase in capacity from the homogeneous case (represented by the dot) for the three diversity levels (5, 23, and 97 (spikes/s)^2) for the discrimination task. Dashed line shows the same but for the identification task. **B** Same as panel A, but for bandwidth diversity with diversity levels of 0, 8,16, and 32 deg^2. We see that bandwidth diversity increases capacity more for identification than for discrimination. **C** Capacity improvement in populations with both amplitude (97 (spikes/s)^2 and bandwidth (32 deg^2) diversity. For populations with both types of diversity, capacity increases by 30% for discrimination and 25% for identification. Populations with just amplitude or bandwidth diversity increase capacity for discrimination by 20% and 10%, respectively. Amplitude or bandwidth diversity alone increase capacity for identification by 9% and 18%, respectively. Hence, the two types of diversity have approximately additive effects on capacity.

Comparison with neuronal data

Amplitude diversity

We tested if our results from simulations held for neural data from macaque V1. For amplitude diversity, we specifically focus on the discrimination task because our simulations showed that amplitude diversity affects it more than identification. We analyzed previously collected microelectrode array recordings from V1 of four hemispheres from three macaques (Graf et al., 2011). We sampled from each recorded population to create subpopulations with varying levels of amplitude diversity – if we did not subsample, we would only have one population per dataset for measuring the relationship between amplitude diversity and

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

capacity. We enforced the subsets to have the same mean amplitude, within a small error, to isolate the effect of amplitude diversity. The dataset shown in Figure 5B,C, had 59 cells, so we formed 100 subsets of 30 neurons each. This particular dataset had an outlier cell with a very high amplitude (see the second subpanel of Figure 5A), so we enforced that the cell was one of the 30 neurons in every subset. Otherwise, the subsets would separate into clusters of low and very high amplitude variance.

The recordings yielded an uneven tiling of the orientation space – for example, see the first subpanel of Figure 5A. Because of this, manifolds for different subsets of the population display geometric differences caused by inconsistent coverage of the stimulus space. This prevents isolating the effects of amplitude diversity. To mitigate this, we measured the geometry and capacity for only one pair of stimuli, instead of all 36 nearby pairs. We chose the pair such that the mean firing rate of the population across the two orientations did not differ too much and the variance of the mean firing rates across the subsets was low. We used discrimination between 290 and 295 deg for the case of the dataset shown in Figure 5B and C. Both the centroid norm and capacity show an increasing trend (Spearman rank correlation of $r = 0.57$ and $r = 0.57$, respectively) which supports our results from the simulated populations. Here, the centroid norm is measured with respect to the global mean of just the two manifolds to reflect that we only consider one pair to discriminate. These trends also hold in three other neural datasets which are shown in SI Figure 2.

Note that there are far fewer cells in the neural subsets than are in the simulated datasets. Because of this, it is more difficult to linearly separate manifolds with the same number of points as were in the simulations. We downsampled to 12 of the 50 points per manifold to combat this issue. We then averaged the capacity over the four sets of 12 points per manifold (first 12 points, second 12, and so on). Further note that the range of capacity values in Figure 5C for the experimental data is 0.18 to 0.3 which is higher than the 0.07 to 0.09 range for the simulated data. We confirmed that this is due to the manifolds only having 12 points, because the capacity for simulations with 12 points per manifold is 0.2 to 0.23. Lastly, while we attempted to control for other variables in the subpopulations to isolate the effects of amplitude diversity, the cells were not as uniform as our simulated neural populations were. Hence, the trends in the real data are weaker than those in the simulations, likely due to variations among the subpopulations in properties other than amplitude diversity.

Bandwidth diversity

We tested that our results about bandwidth diversity from the simulations held in neural data from macaque V1 – the same datasets that we used for the neural data analysis for amplitude diversity above. We focused on identification because our simulations show that bandwidth diversity impacts it more. The same limitations of the data apply here as did for amplitude diversity. For the dataset depicted in Figure 5, we used 100 subsets of 30 neurons where the mean bandwidth across the subsets was held within a small range to isolate the effects of bandwidth diversity. As above, we aimed to create an identification task for the portion of the stimuli that had relatively constant mean response among the various orientations and low variance in mean response across the subsets. Following these criteria, we separated 285 deg from the manifolds for the four orientations surrounding it – 275, 280, 290, and 295 deg. The global mean for calculating center correlations was the average of the five manifolds. In Figure 5D and E, we see that the center correlations decrease and capacity increases with bandwidth diversity (Spearman rank correlation of $r = -0.68$ and $r = 0.39$, respectively). These trends hold in the other three neural datasets as well (see SI Figure 2) and when identification was performed by discriminating one target from the remaining 35 (see SI Figure 3). This supports our simulated results – bandwidth diversity decorrelates the centers of the representations to enhance coding efficiency.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

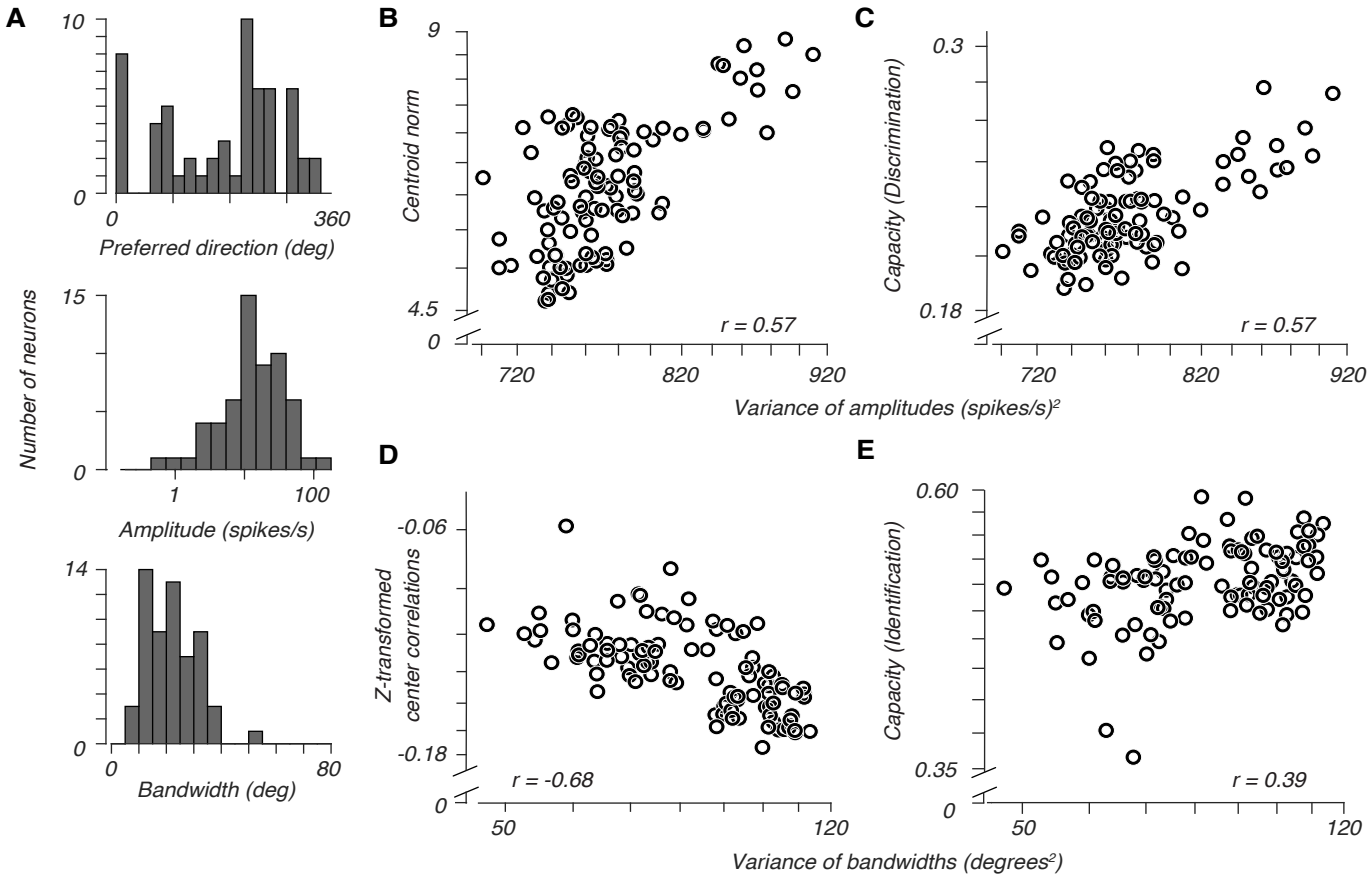


Figure 5. Amplitude and bandwidth diversity improve geometry and coding efficiency in neural populations from macaque V1. **A** Distribution of preferred stimuli, peak to trough amplitudes, and half-bandwidths of the cells in the dataset (dataset 3 from Graf et al. (2011)). See the Methods section for definitions of peak to trough amplitude and half-bandwidth. Observe from the top panel that the neural preferences do not evenly tile the stimulus space. Due to the uneven tiling, we chose one fine discrimination and one identification task to measure the geometry and capacity for, instead of averaging over all possible tasks. For panels B, C, D, and E, we select 100 random subsets of 30 neurons from the total population of 59 neurons, ensuring that one outlier cell with high amplitude was always included. **B** The centroid norm versus amplitude variance of 100 subsets. The centroid norm was calculated for the discrimination task of separating 290 and 295 deg. **C** Capacity for discrimination versus the amplitude variance of the same 100 subsets used in panel B. **D** The center correlations versus bandwidth variance of 100 subsets. The center correlations are measured for the identification task of separating 285 deg from the other manifolds for 275, 280, 290, and 295 deg. We selected 5 manifolds instead of all 36 due to the uneven tiling. See the main text for more detail on these choices. **E** Capacity for identification versus bandwidth variance of the same subsets used in panel C. The range for capacity is of the same ratio as the range for capacity in panel C.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Discussion

Diversity of selectivity is commonplace across mammals and brain areas (e.g. place field sizes in bat (Eliav et al., 2021) and rat (Rich et al., 2014) hippocampus, finger pad selectivity in macaque somatosensory cortex (Fitzgerald et al., 2006), head direction tuning bandwidths in mouse postsubiculum (Clark et al., 2025; Duszkiewicz et al., 2024), among numerous others). Our results reveal that tuning diversity alters population representations to enhance coding efficiency. Moreover, amplitude and bandwidth diversity improve coding via distinct mechanisms : amplitude diversity pushes the representations for different stimuli further apart in the neural space, while bandwidth diversity decorrelates the neural subspaces where the centers of different representations lie. Both amplitude and bandwidth diversity increase the separability among signals, but in complementary ways that enhance their impact on different perceptual tasks. Amplitude diversity affects coding efficiency for discrimination more, while bandwidth diversity affects identification more.

To illustrate the complementary nature of the geometric transformations, consider the following: the representations exist in a space contained within the entire neuronal population's firing rate space. The size of the space holding the representations is related to their separability – the space is larger when the representations are further apart from one another. One can visualize this space as a box containing the representations. The box's size is defined by its end-to-end distance and dimension, and an increase in either of these quantities increases the overall size of the box. By expanding the centroid norms of the manifolds, amplitude diversity increases the end-to-end distance of the box, e.g. going from a box with smaller sides to one with larger sides. This helps the population code effectively use the range of the firing rates available to the neurons. Bandwidth diversity, on the other hand, increases the dimension of the set of manifold centers, which increases the dimension of the box, e.g. from an envelope to a 3D box. Bandwidth diversity allows the population code to capitalize on the high dimension of the complete neural firing rate space. Amplitude and bandwidth diversity therefore expand the space containing the representations in different ways. With both types of diversity, a larger number of representations are separable than would be if only one type of diversity existed.

Of course, one could achieve these same goals, for example, by increasing the number of neurons or their firing rates. But metabolic constraints prevent neurons from having arbitrarily high firing rates, amplitudes, and bandwidths. Our finding suggests that neural systems can enhance coding efficiency by unevenly distributing spikes across neurons while still respecting the constraints. Future work could involve probing how diversity changes over development or learning, and understanding how the level of diversity differs across species depending on the tasks they perform. Additionally, this work invites the question of *how* neuronal circuits generate and maintain tuning diversity.

Previous work, such as that of Shamir and Sompolinsky (2006), Ecker et al. (2011), and Tian et al. (2024), has explored why tuning diversity is beneficial for neural coding. This research showed that tuning heterogeneity can overcome information saturation in the presence of noise correlations in the population, and that there is a dependence of the information gain due to heterogeneity on the level of noise correlations. These previous reports focused on amplitude diversity's effect on fine discrimination. We, on the other hand, use a representational geometry approach and expand this range of ideas to study how amplitude and bandwidth diversity *differentially* drive efficiency in the population code for discrimination and identification.

In our framework, the manifold size is determined by trial-to-trial variability and noise correlations because there is no variation in the stimulus across presentations. Natural stimuli, however, vary in many ways that require an observer to ignore nuisance variables while performing specific perceptual tasks (e.g. determining the orientation while ignoring variations in size or shape). If we introduce another variable to the stimuli, such as contrast, the new manifolds encoding orientation would be larger than the manifolds created without stimulus variability. This is because of an additional source of response variability to presentations of the same orientation. We expect that tuning diversity would allow these larger manifolds to have higher coding efficiency than they would if the population had homogeneous tuning. We tested

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

this idea by introducing contrast dependent responses and randomized stimulus contrasts into our simulations, and measuring how diversity affected the capacity for discrimination and identification. For both discrimination and identification, homogeneous populations had lower capacity when we added random contrasts. When we introduced amplitude and bandwidth diversity, the capacity increased to match and exceed the capacity for the simulations without nuisance variables (see Figure 6). As before, amplitude diversity improved coding efficiency more for discrimination, and bandwidth diversity for identification. Perhaps tuning diversity is a mechanism by which the neural system flexibly encodes a specific stimulus variable in the presence of other variables that are not important for a given task.

Finally, this work belongs to a broader effort to connect neuronal tuning properties to representational geometry. Representational geometry has often been used to study population coding in areas where single-neuron properties are not well understood. Our work bridges the gap between analytical approaches grounded in tuning and those utilizing population-level geometry (Kriegeskorte and Wei, 2021). While we focused on how tuning diversity shapes the structure of the overall representation, many questions remain open. For example, how do tuning properties influence task-specific geometries (Chou et al., 2025; Chung et al., 2018)? In addition, we did not study how noise and correlational structures affect geometry. By demonstrating a direct link between tuning and representational geometry, we lay the groundwork for a new approach to study coding efficiency—one that unifies tuning, geometry, and task demands under a common framework.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

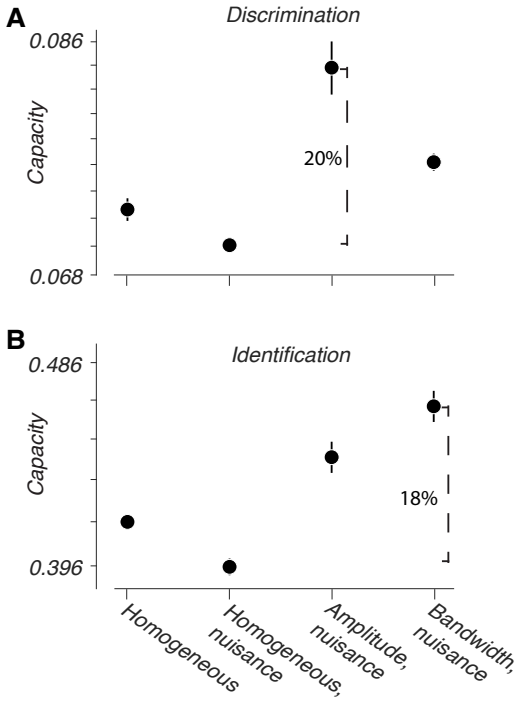


Figure 6. Diversity maintains capacity in the presence of nuisance variables. **A** Capacity for discriminating orientation in the presence of contrast variations. Introducing contrast variations lowers capacity, but capacity is restored with tuning diversity in the population. The amplitude diverse population and bandwidth diverse population had respective variances of 97 $(\text{spikes/s})^2$ and 32 deg^2. **B** Same as panel A, but for identification. The range of capacity values is of the same ratio as that in panel A.

Acknowledgements

We would like to thank Albert Wakhloo, Nga Yu Lo, Artem Kirsanov, Will Slatton, Jenelle Feather, and Dario Ringach for helpful discussion about the work. We also thank Chi-Ning Chou for discussion and a well-managed codebase for measuring capacity (Chou et al., 2025). The Flatiron Institute is a division of the Simons Foundation. The computations reported in this paper were (in part) performed using resources made available by the Flatiron Institute. This work was supported in part by grants from the National Institute of Health– 5T32EY007136, 5T90DA059110, and R01DA059220– and the Simons Foundation, Collaboration on the Global Brain (543019). S.C. is partially supported by a Sloan Research Fellowship and a Klingenstein-Simons Award.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Methods

Definitions of geometric measures

We define geometry that captures the point-cloud structure of the manifolds and their relation to one another in the neural firing rate space. The geometry is designed to capture various aspects of the manifolds that relate to coding efficiency (Chou et al., 2025; Chung et al., 2018; Wakhloo et al., 2023), while maintaining intuitive relationships with tuning distributions. Below, we describe the geometric definitions for both the discrimination and identification task.

Discrimination

Suppose that there are N neurons in the population, let $p = 1, 2$ represent the two manifolds to separate, and suppose that all manifold points have been centered with respect to the overall mean of the P manifolds. C_p is the center point of the manifold p. For manifold p, $t_{i,p}$ is the ith Gaussian random vector originating at C_p, and $s_{i,p}$ is the largest point in the direction of $t_{i,p}$. The vectors $a_{i,p}$ are the unit vector principal components of manifold p and the coefficient $E_{i,p}$ is the proportion of variance explained by component $a_{i,p}$. Geometric definitions: **dimension**, $D_p = \left\langle \left(t_{i,p} \cdot \frac{s_{i,p}}{||s_{i,p}||} \right)^2 \right\rangle_i$; **radius**, $R_p = \langle ||s_{i,p}||^2 \rangle_i^{\frac{1}{2}}$; **centroid norm**, $||C_p||$; **center correlations**, $g\left(\frac{C_1}{||C_1||} \cdot \frac{C_2}{||C_2||} \right)$; **axis correlations**, $\lambda = \sum_i E_{i,1} * g(|a_{i,1} \cdot a_{i,2}|)$

where $g(r) = \frac{1}{2} ln\left(\frac{1+r}{1-r}\right)$ is Fisher's Z-transformation for correlation coefficients. This transformation allows correlation coefficients to be compared, as it linearizes the differences between correlation values. For example, the difference in correlations of 0.1 and 0.2 is smaller than that of 0.85 and 0.95, in terms of strength of association. The Z-transformation accounts for this (Fisher, 1915; Smith et al., 2005).

We averaged over R_p, D_p, and $||C_p||$ to obtain the average radius, dimension, and centroid norm for the two manifolds in the discrimination task. Note that the trends we found with these radius and dimension definitions are similar to the trends shown with more conventional measures such as the L^2 norm of the eigenvalues for radius and the participation ratio of the eigenvalues for dimension. We chose to use these definitions for radius and dimension to capture the sizes of non-ellipsoidal shapes.

For the values reported in Figure 3, and SI Figure 1, we averaged the geometric measures over all possible fine discrimination tasks (between stimuli with 5 degree difference).

Identification

As before, suppose that there are N neurons in the population and that all manifold points have been centered with respect to the overall mean of the P manifolds. The definitions of C_p, $t_{i,p}$, $s_{i,p}$, $a_{i,p}$, and $E_{i,p}$ are the same as for discrimination. Suppose further that we are specifically concerned with separating manifold M_p from the other $P - 1$ manifolds.

Geometric definitions: **dimension**, D_p from above; **radius**, R_p from above; **centroid norm**, $||C_p||$ as above; **center correlations** are defined as : $\psi_p = \left\langle g\left(\frac{C_p}{||C_p||} \cdot \frac{C_k}{||C_k||} \right) \right\rangle_{k \neq p}$; **axis correlations**, $\lambda_p = \langle \sum_i E_{i,p} * g\left(|a_{i,p} \cdot a_{i,k}|\right) \rangle_{k \neq p}$

For discrimination, we measure only one pair of manifold's correlations, while for identification we average over all possible pairs with the pth manifold. The same distinction holds for axis correlations.

For the values reported in Figure 3 and SI Figure 1, we averaged geometric measures over all possible manifolds as p, the target manifold.

Measuring capacity

We measured capacity directly by following the methods described in Chung et al. (2018); Cohen et al. (2020) and improved upon in Chou et al. (2025). We briefly describe the procedure here: Suppose we have

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

a population of N neurons and their responses to P different stimuli. For simplicity, we will describe the algorithm with reference to a specific discrimination task, separating the manifold for θ_1 from the manifold for θ_2. Capacity is empirically measured by finding the smallest N_c such that the manifolds for θ_1 and θ_2 are separable for 50% of the random projections of the data onto neural subspaces of dimension N_c. We perform binary search to find N_c, and capacity is measured to be $\frac{2}{N_c}$, where is 2 is the number of manifolds to separate. The units for capacity are manifolds/neuron. We measure capacity analogously for identification for θ_1, but instead measure the separability between the manifold for θ_1 and $\mathcal{M} := \bigcup_{i \neq 1} M_i$. Capacity is $\frac{P}{N_c}$ since there are P manifolds to separate.

Above we described measuring capacity for a specific discrimination or identification task. Since there are $P = 36$ possible discriminations (or identifications) in our simulations, we report the mean of the capacity values for the P discriminations of θ_i and θ_{i+1}. For the edge case of $i = P$, we measure the separability between the manifolds for stimulus P and the first stimulus because orientation is a circular variable. We report an analogous averaged capacity for identification on our simulated data. Capacity is higher for identification than for discrimination because all P manifolds are being separated, and the further away manifolds are easier to separate from the target manifold.

Setup for simulated neural populations

Tuning curves

We followed the tuning curve model from Shamir and Sompolinsky (2006). The ith neuron's tuning curve follows the function

$$f_i(\theta) = A_i e^{2(cos(B_i(\theta - \theta_i)) - 1)} + s$$

where s is the baseline firing rate for all neurons and θ_i, A_i, and B_i are the preferred stimulus, amplitude, and bandwidth parameters for neuron i, respectively. For a population of N neurons, the preferred stimuli were evenly distributed following the formula $\theta_i = -\pi \frac{N+1}{N} + \frac{2\pi}{N} i$. For all simulations shown, $s = 3.3$. The f_i are modified to create direction selective (f_i has one peak within 360 deg), directionally biased (f_i has one main peak, and one smaller peak 180 deg away at 60% of the main peak height), and orientation selective cells (f_i has two equally sized peaks 180 deg apart). We roughly matched the proportions of each type of selectivity with the neural data such that 10% were direction selective, 20% were directionally biased, and the rest were orientation selective.

In the homogeneous populations, $A_i = 15$ and B_i is chosen such that the half-bandwidth of the cells is 16 deg. The half-bandwidth is defined as half of the width of the tuning curve at $\frac{1}{\sqrt{2}}$ of the peak of the curve. For example, if the peak of the tuning curve is $15 + s$, the half-bandwidth would be measured between the points where the tuning curve has the value $\frac{1}{\sqrt{2}}(15 + s)$.

We implemented amplitude diversity by assuming the amplitudes are lognormally distributed with a mean of 15 spikes/s. We create more diversity by increasing the variance of the lognormal distribution from which the A_i are selected. Please refer to the figures for the specific diversity levels used in the simulations.

Bandwidth diversity was introduced in a similar fashion, however we assume a gamma distribution for bandwidths. The mean of the half-bandwidths is set to 16 deg to match our biological neural populations, and bandwidths are drawn as follows:

Let κ_b be a scale parameter in $\{0.5, 1, 2\}$, and select $b_i \sim \Gamma(\frac{16}{\kappa_b}, \kappa_b)$. The expected value of b_i is 16, and the variance is $16\kappa_b$. Note that as κ_b increases, the population's bandwidth heterogeneity increases. We then set

$$B_i = \frac{cos^{-1}\left(\frac{1}{2} log\left(\frac{1}{\sqrt{2}}\right) + 1\right)}{b_i},$$

which gives f_i the desired half-bandwidth of b_i. Throughout the text, we use the term "bandwidth" to mean the "half-bandwidth" as defined above.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Noise and correlations

Trial-to-trial variability and noise correlations are introduced in a similar manner to the method used in Shamir and Sompolinsky (2006). Correlations between neurons depended on the distance between their preferred stimuli, and exponentially decayed as the distance increased. However, we used multiplicative noise while Shamir and Sompolinsky (2006) assumed an additive noise model for their cells. The correlation between neurons i and j where $i \neq j$, was given by $C_{ij} = 0.4e^{-|\theta_i - \theta_j|}$.

We assumed Poisson-like statistics, and enforced that the trial-to-trial variability of each cell was equal to it's mean response to a given stimulus. We then scaled the correlation matrix by the neural variabilities to create the covariance matrix, $\tilde{C_\theta}$, where θ indicates the stimulus direction.

Finally, we simulated the population's responses on a trial of stimulus θ as $(r_1, ..., r_N) \sim \mathcal{N}(f_\theta, \tilde{C_\theta})$, where $f_\theta = (f_1(\theta), ..., f_N(\theta))$.

Contrast response functions

To model contrast dependent responses, we redefined $f_i(\theta)$ as

$$f_i(\theta) = R(c)(f_i(\theta) - s) + s$$

where $0 \leq c \leq 1$ and $R(c) = \frac{(3c)^3}{(3c)^3+0.5}$. The parameter c scales the stimulus's contrast, and we chose the function R to match the shapes of contrast response curves in V1 (Albrecht and Hamilton, 1982). See Supporting Figure 10 for a plot of $R(c)$.

Contrast as a nuisance variable

For each trial response, we uniformly selected c from the range of 0.183 to 0.35 which ensured that $R(c)$ fell within the range of $0.25to0.7$, with a mean of 0.5 (see SI Figure 4). We also doubled the mean amplitude of our populations so that the mean responses were consistent with the simulations that did not have contrast variations.

Neural data analyses

The data used here were previously reported in Graf et al. (2011). Neural data were recorded with Utah arrays implanted in V1 of 5 hemispheres of 3 macaques. Animals viewed drifting sinusoidal gratings of 72 different directions with 50 trials each. The spatial and temporal frequencies of the stimuli were chosen to maximize the population activity. The gratings were presented for 1,280 ms with a 1,280 ms blank screen in between stimuli. Data were spike sorted, and neural responses to the stimuli were taken from all 1,280 ms of each presentation. Further details of the data collection can be found in Graf et al. (2011). One array, dataset 1, was reported by the authors of Graf et al. (2011) to have been under an altered experimental regime. Through our analyses, we also noticed a lower number of cells and lower overall firing rates in this dataset than the other datasets showed. For this reason, we excluded dataset 1 from our analyses.

Criteria for including neurons in analyses

For our analyses, we wanted to use cells that were visually responsive and well-tuned. For each dataset, we applied the following to criteria to decide whether cells were included for analysis or not: Cells were visually responsive if they displayed an evoked response that was at least one standard deviation above or below their spontaneous rate. The average and standard deviation of the spontaneous rate were computed from the final 500 ms of the blank screen presentation in between trials. Cells were well-tuned if they were approximated by the sum of two Von mises functions with an $r^2 > 0.75$ (Graf et al., 2011).

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Measuring amplitude and bandwidth

For each cell, we first found the maximum value of the tuning curve. We then considered the half of the tuning curve where the maximum value was (0-175 deg or 180-355 deg), and convolved this with $[0.05, 0.25, 0.4, 0.25, 0.05]$ for smoothing. We measured amplitude as the difference between the maximum and minimum values of the smoothed tuning curve. To measure bandwidth, we first found the two points on either side of the peak that were closest to $\frac{1}{\sqrt{2}}$ of the peak value. Then, we defined the bandwidth as half of the difference in deg of two points.

Applying the geometric measures

We apply the same geometric measures defined above, with a minor change to take into account the uneven tiling of stimulus space. Recall from the main text that we specify the subset of manifolds to discriminate or identify from for the neural data analyses. Thus, for centroid norm and center correlation calculations, we center the data with respect to the mean of the subset of manifolds used in the task, rather than all P manifolds as we do for the simulated data.

Multiple regression analyses

We performed multiple regression analyses using the five geometric properties as explanatory variables and capacity as the dependent one. We first performed Tukey's ladder of transformations to find the appropriate relationship between each geometric variable and capacity. For each variable x, we computed $\log(x)$, $x^{\frac{1}{2}}$, x, x^2, ..., x^7 and found the Pearson correlation coefficient between the transformed values and capacity. If there was no optimal transformation (i.e. all of the correlation coefficients were about the same) we assumed no transformation. This was the case for every geometric property, so we performed the multiple regression without any transformation on the variables. In Table 1 we report the standardized coefficients for all multiple regressions (Kerlinger and Pedhazur, 1973).

References

Albrecht, D. G. and Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function. *Journal of neurophysiology*, 48(1):217–237.

Bernardi, S., Benna, M. K., Rigotti, M., Munuera, J., Fusi, S., and Salzman, C. D. (2020). The geometry of abstraction in the hippocampus and prefrontal cortex. *Cell*, 183(4):954–967.

Chou, C.-N., Kim, R., Arend, L. A., Yang, Y.-Y., Mensh, B. D., Shim, W. M., Perich, M. G., and Chung, S. (2025). Geometry linked to untangling efficiency reveals structure and computation in neural populations. *bioRxiv*.

Chung, S. and Abbott, L. F. (2021). Neural population geometry: An approach for understanding biological and artificial neural networks. *Current Opinion in Neurobiology*, 70:137–144.

Chung, S., Lee, D. D., and Sompolinsky, H. (2016). Linear readout of object manifolds. *Physical Review E*, 93(6):060301.

Chung, S., Lee, D. D., and Sompolinsky, H. (2018). Classification and geometry of general perceptual manifolds. *Physical Review X*, 8(3):031003.

Clark, D. G., Abbott, L., and Sompolinsky, H. (2025). Symmetries and continuous attractors in disordered neural circuits. *bioRxiv*, pages 2025–01.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Cohen, U., Chung, S., Lee, D. D., and Sompolinsky, H. (2020). Separability and geometry of object manifolds in deep neural networks. *Nature communications*, 11(1):1–13.

DiCarlo, J. and Cox, D. (2007). Untangling invariant object recognition. *TRENDS in Cognitive Sciences*, 11(8).

Duszkiewicz, A. J., Orhan, P., Skromne Carrasco, S., Brown, E. H., Owczarek, E., Vite, G. R., Wood, E. R., and Peyrache, A. (2024). Local origin of excitatory–inhibitory tuning equivalence in a cortical network. *Nature Neuroscience*, 27(4):782–792.

Ecker, A. S., Berens, P., Tolias, A. S., and Bethge, M. (2011). The effect of noise correlations in populations of diversely tuned neurons. *The Journal of Neuroscience*, 31(40):14272–14283.

Eliav, T., Maimon, S. R., Aljadeff, J., Tsodyks, M., Ginosar, G., Las, L., and Ulanovsky, N. (2021). Multiscale representation of very large environments in the hippocampus of flying bats. *Science*, 372(6545):eabg4020.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, 10(4):507–521.

Fitzgerald, P. J., Lane, J. W., Thakur, P. H., and Hsiao, S. S. (2006). Receptive field (rf) properties of the macaque second somatosensory cortex: Rf size, shape, and somatotopic organization. *Journal of Neuroscience*, 26(24):6485–6495.

Genkin, M., Shenoy, K. V., Chandrasekaran, C., and Engel, T. A. (2023). The dynamics and geometry of choice in premotor cortex. *BioRxiv*.

Goris, R. L., Simoncelli, E. P., and Movshon, J. A. (2015). Origin and function of tuning diversity in macaque visual cortex. *Neuron*, 88(4):819–831.

Graf, A., Kohn, A., Jazayeri, M., and Movshon, J. A. (2011). Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature neuroscience*, 14(2):239–245.

Green, D. M., Swets, J. A., et al. (1966). *Signal detection theory and psychophysics*, volume 1. Wiley New York.

Jazayeri, M. and Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience*, 9:690–696.

Kang, K., Shapley, R. M., and Sompolinsky, H. (2004). Information tuning of populations of neurons in primary visual cortex. *Journal of Neuroscience*, 24:3726–3735.

Kerlinger, F. and Pedhazur, E. (1973). *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart Winston.

Kriegeskorte, N. and Wei, X.-X. (2021). Neural tuning and representational geometry. *Nature reviews neuroscience*, 22:703–718.

Mehrpour, V., Meyer, T., Simoncelli, E. P., and Rust, N. C. (2021). Pinpointing the neural signatures of single-exposure visual recognition memory. *Proceedings of the National Academy of Sciences*, 118(18):e2021660118.

Rich, P. D., Liaw, H.-P., and Lee, A. K. (2014). Large environments reveal the statistical structure governing hippocampal representations. *Science*, 345(6198):814–817.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Ringach, D. L., Shapley, R. M., and Hawken, M. J. (2002). Orientation selectivity in macaque v1: diversity and laminar dependence. *Journal of neuroscience*, 22(13):5639–5651.

Roweis, S. T. and Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326.

Seung, H. S. and Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences*, 90:10749–10753.

Shamir, M. and Sompolinsky, H. (2006). Implications of neuronal diversity on population coding. *Neural Computation*, 18:1951–1986.

Smith, M. A., Majaj, N. J., and Movshon, J. A. (2005). Dynamics of motion signaling by neurons in macaque area mt. *Nature neuroscience*, 8(2):220–228.

Tenenbaum, J. B., Silva, V. d., and Langford, J. C. (2000). A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323.

Tian, G. J., Zhu, O., Shirhatti, V., Greenspon, C., Downey, J. E., Freedman, D. J., and Doiron, B. (2024). Neuronal firing rate diversity lowers the dimension of population covariability. *bioRxiv*, pages 2024–08.

Wakhloo, A. J., Slatton, W., and Chung, S. (2024). Neural population geometry and optimal coding of tasks with shared latent structure. *arXiv preprint arXiv:2402.16770*.

Wakhloo, A. J., Sussman, T. J., and Chung, S. (2023). Linear classification of neural manifolds with correlated variability. *Physical Review Letters*.

Yerxa, T., Kuang, Y., Simoncelli, E., and Chung, S. (2023). Learning efficient coding of natural images with maximum manifold capacity representations. *Advances in Neural Information Processing Systems*, 36:24103–24128.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Supporting Information

Proof for the centroid norm increasing with amplitude heterogeneity

First consider the centroid norm of the manifold created by the responses to stimulus θ for a homogeneous population. The neurons have identical tuning curve shapes given by $f(\theta)$. The neurons have different peak preferences of stimuli, so the ith neuron's tuning curve is given by $f_i(\theta) = f(\theta - \theta_i)$ where the ith neuron's preferred stimulus is θ_i. The firing rate for the ith neuron for stimulus θ is $m_i(\theta) = f_i(\theta) + s$ where s is a fixed spontaneous firing rate held constant for all neurons.

The global mean response of each neuron is given by the average of its responses to all stimuli, $g_i = \frac{1}{2\pi}\int_0^{2\pi} f(\theta)d\theta + s$. The centroid vector at stimulus θ_1 is $C = (m_1(\theta_1), ..., m_N(\theta_1))$. The the norm of the global mean subtracted centroid is given by :

$$\|C - G\| = \sqrt{(m_1(\theta_1) - g_1)^2 + ... + (m_N(\theta_1) - g_N)^2}$$

Now consider a population with heterogeneous amplitudes. Suppose for simplicity that heterogeneity is implemented by drawing N independent and identically distributed $\epsilon_i \sim \mathcal{N}(0, \kappa)$ (where κ indicates the level of heterogeneity), and scaling the i-th tuning curve by $1 + \epsilon_i$. The heterogeneous tuning curves are given by $\tilde{m}_i(\theta) := (1 + \epsilon_i)f_i(\theta) + s = (1 + \epsilon_i)m_i(\theta) - \epsilon_i s$.

The new global means are given by $\tilde{g}_i = (1 + \epsilon_i)g_i - \epsilon_i s$ and the new centroid vector for the manifold for stimulus θ_1 is given by $\tilde{C} = (\tilde{m}_1(\theta_1), ..., \tilde{m}_N(\theta_1)) = ((1 + \epsilon_1)m_1(\theta_1) - \epsilon_1 s, ..., (1 + \epsilon_N)m_N(\theta_1) - \epsilon_N s)$ We can write the norm of the global mean subtracted centroid in this case as :

$$\begin{aligned}
\|\tilde{C} - \tilde{G}\|^2 &= (1 + \epsilon_1)^2(m_1(\theta_1) - g_1)^2 + ... + (1 + \epsilon_N)^2(m_N(\theta_1) - g_N)^2 \\
&= \sum_{i=1}^{N}(m_i(\theta_1) - g_i)^2 + 2\sum_{i=1}^{N}\epsilon_i(m_i(\theta_1) - g_i)^2 + \sum_{i=1}^{N}\epsilon_i^2(m_i(\theta_1) - g_i)^2 \\
&= \|C - G\|^2 + 2\sum_{i=1}^{N}\epsilon_i(m_i(\theta_1) - g_i)^2 + \sum_{i=1}^{N}\epsilon_i^2(m_i(\theta_1) - g_i)^2
\end{aligned}$$

Because $\epsilon_i \sim \mathcal{N}(0, \kappa)$, $\mathbb{E}[\epsilon_i^2] = \kappa$. Hence,

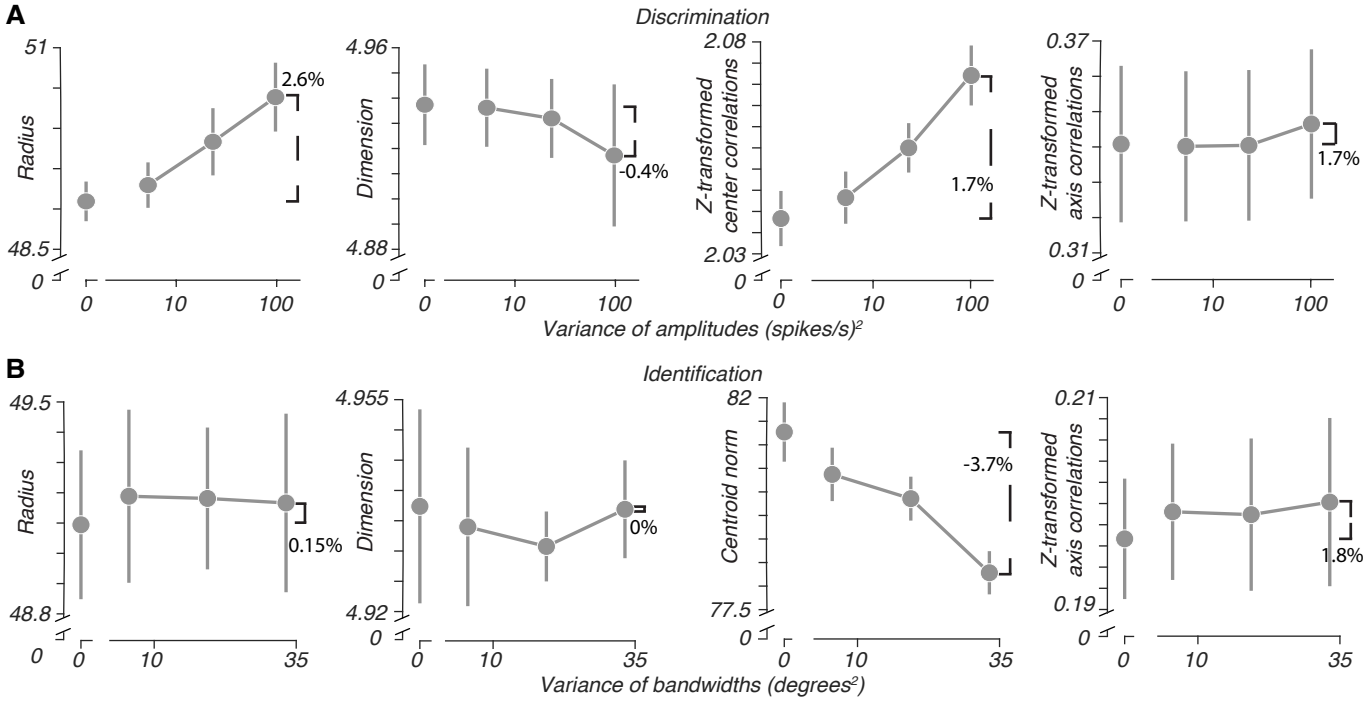
$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{N}\epsilon_i^2(m_i(\theta_1) - g_i)^2\right] &= \kappa\sum_{i=1}^{N}(m_i(\theta_1) - g_i)^2 \\
&= \kappa\|C - G\|^2.
\end{aligned}$$

Additionally, because $\mathbb{E}[\epsilon_i] = 0$,

$$\begin{aligned}
\mathbb{E}\left[2\sum_{i=1}^{N}\epsilon_i(m_i(\theta_1) - g_i)^2\right] &= 2\sum_{i=1}^{N}\mathbb{E}[\epsilon_i](m_i(\theta_1) - g_i)^2 \\
&= 0.
\end{aligned}$$

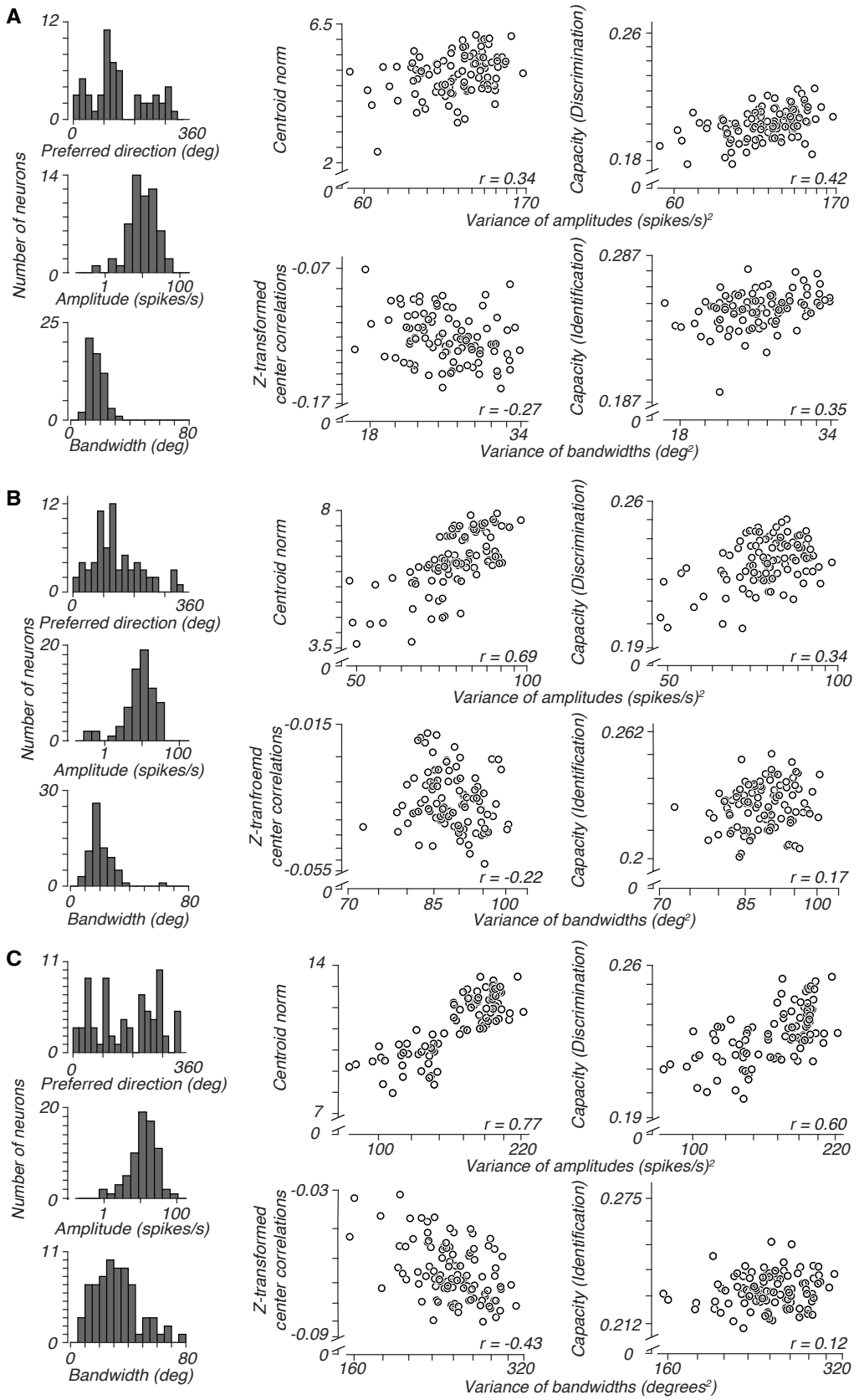
Taken together, the above equations show that on average $\|\tilde{C} - \tilde{G}\|^2 = (1 + \kappa)\|C - G\|^2$, and the centroid norm increases with the heterogeneity level, κ.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



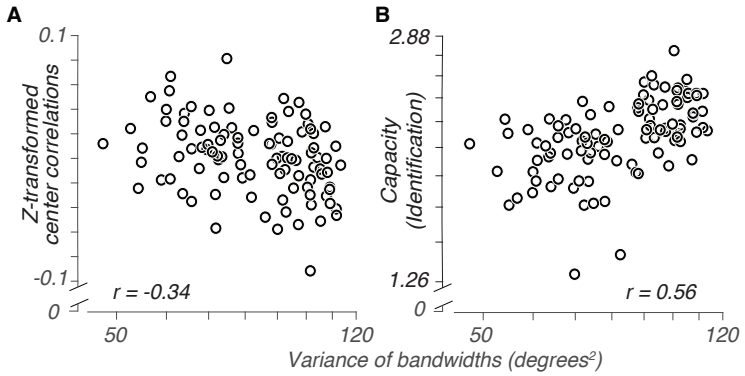
SI Figure 1. The impact of diversity on the remaining geometric properties. Data is from the simulations used for Figures 3A and 3E. We only show the geometry for the task each type of diversity affects more (discrimination for amplitude diversity and identification for bandwidth diversity). **A** The radius, dimension, center correlations, and axis correlations versus the amplitude variance in the simulated populations for discrimination. Error bars indicate the standard deviation over the 10 trials. We performed a multiple regression to show that the change in the centroid norm (shown in Figure 3A) affects capacity the most. The radius and dimension are the same for identification because we average over all possible target manifolds (for identification) and all pairs of nearby manifolds (for discrimination). **B** The radius, dimension, centroid norm, and axis correlations versus the bandwidth variance in the simulated populations for identification. Again, the multiple regression showed that these four properties had a small affect on capacity, while the center correlations drove the capacity increase. The radius, dimension, and centroid norm are the same for discrimination by the logic given in panel A's caption.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

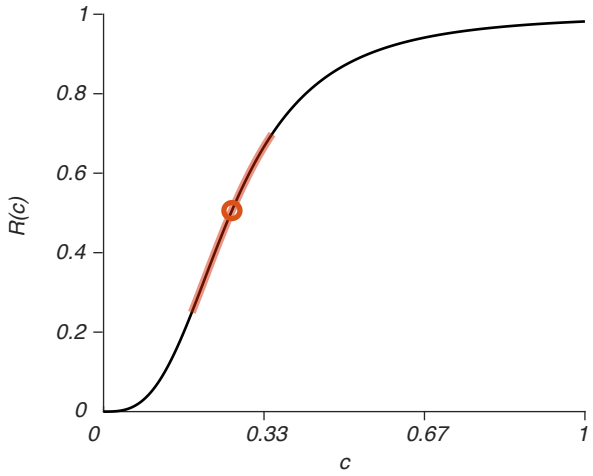


SI Figure 2. Tuning diversity and its relationship to representational geometry and capacity for the three other neural datasets. For each dataset, the range of capacity for discrimination is in the same ratio with the range of capacity for identification. See the SI methods below for details on the analysis of each dataset. **A** Dataset 2 with 56 neurons. **B** Dataset 4 with 68 neurons. **C** Dataset 5 with 71 neurons.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.



SI Figure 3. Identification task with neurophysiological data. Discrimination of one from any (35 other stimuli) instead of one from four other stimuli as was done in the main text and SI Figure 2. This figure uses the same subsets as those used in panels D and E of Figure 5. **A** Center correlations versus bandwidth variance of the 100 subsets. The center correlations are measured for the identifying 315 deg from the rest of 180-355 deg. We selected 315 deg because the population's response to 315 deg was similar to the population's average response across stimuli from 180-355 deg. **B** Capacity for identification versus bandwidth variance of the subsets.



SI Figure 4. Contrast response function for the cells. $R(c)$ multiplies the evoked response of each cell to mimic contrast levels in the stimulus. For the simulations with contrast variations shown in Figure 6, we randomly and uniformly chose c to be within the red highlighted range, with a mean of 0.8 indicated by the red circle. This allowed $R(c)$ to stay within a range of 0.25 to 0.7, with a mean of 0.5. $R(c) = 0.5$ was the multiplier used for the simulations without contrast variations.

bioRxiv preprint doi: https://doi.org/10.1101/2025.06.26.661754; this version posted June 28, 2025. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

SI Methods

Dataset 2: Amplitude variability - 100 subsets of 30 neurons each, we computed discrimination between 260 and 265 deg. Bandwidth variability - 100 subsets of 30 neurons each, we computed identification between 305 deg and all of 295, 300, 310, and 315 deg.

Dataset 4: Amplitude variability - 100 subsets of 40 neurons each, we computed discrimination between 290 and 295 deg. Bandwidth variability - 100 subsets of 40 neurons each, we computed identification between 280 deg and all of 270, 275, 285, and 290 deg. We made sure that each subset included the outlier cell with very high bandwidth.

Dataset 5: Amplitude variability - 100 subsets of 40 neurons each, we computed discrimination between 225 and 230 deg. Bandwidth variability - 100 subsets of 40 neurons each, we computed identification between 265 deg and all of 255, 260, 270, and 275 deg.