



Supplementary Materials for

Geometry of sequence working memory in macaque prefrontal cortex

Yang Xie *et al.*

Corresponding authors: Shiming Tang, tangshm@pku.edu.cn; Bin Min, bin.min@bsbii.cn; Liping Wang, liping.wang@ion.ac.cn

Science **375**, 632 (2022)
DOI: 10.1126/science.abm0204

The PDF file includes:

Materials and Methods
Figs. S1 to S12
Tables S1 to S5
References

Other Supplementary Material for this manuscript includes the following:

MDAR Reproducibility Checklist

Material and Methods

Experimental model preparation

Subjects

Two rhesus monkeys M1 and M2 (*Macaca mulatta*, male, 5.0–6.1 kg, 4 years old) from Beijing Institute of Xieerxin Biology Resource were used in the study. To express the Ca²⁺ indicator GCaMP6s (39), we injected AAV1.Syn.GCaMP6s.WPRE.SV40 (titer 2.3e13 GC/ml, Penn Vector Core) into the prefrontal cortex in each monkey. All experimental protocols followed the Guide of the Institutional Animal Care and Use Committee (IACUC) of Peking University Laboratory Animal Center and were approved by the Peking University Animal Care and Use Committee (LSC-TangSM-3).

Surgeries

Three sequential surgeries were performed on each animal under general anesthesia. In the first surgery, head posts were implanted on each animal's skull. Two head posts were implanted on the back of the head and one on the forehead. A Y-shaped steel frame was fixed to these head posts for head stabilization during training and recording. Both monkeys were then trained on the behavioral task over 7 months. After reaching the performance criteria, the second surgery was performed. In the second surgery, a 20 mm craniotomy was made on the skull over the region of lateral prefrontal cortex located along the principal sulcus (PS) until the arcuate sulcus (AS), roughly corresponding to Walker's areas 46 and 8, and the dura was opened. Then, the virus injections were performed. Around 200 nL AAV1.Syn.GCaMP6S.WPRE.SV40 (Penn Vector Core, titer 2.7e13 GC/ml) was injected at a depth of around 400 μm. The injection sites were located alongside the PS, roughly corresponding to Walker's area 46 (see fig. S2 for details). After virus injection, the dura was sutured, and the removed skull bone was replaced. In the first week of recovery, ceftriaxone sodium antibiotic was administered.

The last surgery was performed 8–9 weeks later to implant an imaging window. The skull was re-opened and the subjacent dura was cut off (diameter 16 mm) to expose the cortex. A glass coverslip (15 mm in diameter and 0.17 mm in thickness) was glued with a titanium ring (12 mm outer diameter and 10 mm

inner diameter). A ring-shape GORE membrane (12 mm inner diameter and 20 mm outer diameter) was mounted onto the titanium ring and glued on the outer part of the glass coverslip. The coverslip, titanium ring, and membrane constituted a window unit. The window unit was gently pushed onto the cortical surface. The GORE membrane was inserted under the dura. The titanium ring was glued to the skull with dental acrylic. The entire imaging window was then covered with a steel shell to protect the coverslip. Please refer to Li et al. 2017 (40) for more details.

Behavioral task and analysis

Sequences

Two naïve monkeys M1 and M2 were trained to perform a delayed-sequence reproduction task (Fig. 1A) (22). The spatial sequences were created from stimuli at six locations (set: [1, 2, 3, 4, 5, 6]), all of which were 7° away from the screen center and together formed a symmetrical hexagon. Each location was sampled only once in a single sequence. As a result, there are 30 different length-2 sequences and 120 different length-3 sequences in total. Sequences were presented on the screen and monkeys were required to memorize and reproduce each sequence. Due to the trial number limitation during a day, it is impossible to collect enough trials to cover all 120 length-3 sequences each day. Therefore, we had to select a subset of the length-3 sequences for use during a day to make sure that we collected enough trials for each selected length-3 sequence in each imaging session. More specifically, a random set of 36 length-3 sequences was selected each day. These length-3 sequences and all 30 length-2 sequences comprised a sequence pool. In each trial, one sequence was randomly selected from this pool.

Task structure

During experiments, the monkeys were seated in a primate chair facing a 20-inch LCD monitor. Eye positions were monitored with an infrared oculometer system at a sampling rate of 500 Hz (Eyescan). Each trial began with the appearance of a fixation point in the center of the screen. Monkeys were required to maintain fixation until the “go” signal was presented. After monkeys gazed at the fixation point for 500 ms, six circles 1.2° in diameter were presented 7° away from the fixation point. Together, the six circles

composed a symmetrical hexagon. After another 500 ms, two or three red targets were presented sequentially. Each target was located in one of the six circles. Each target was presented for 200 ms and the time interval between targets (the inter-stimulus interval, ISI) was 400 ms. These targets constituted a specific spatial sequence. After a random delay (2500 - 4000 ms, based on an exponential decay time), the fixation point disappeared (the “go” signal), and then the monkeys were required to saccade to the location of each memorized target (the item information of the sequence) in the correct order (the rank information of the sequence). When reproducing the spatial sequence, monkeys were required to maintain fixation at each memorized location for 300 ms. As feedback, the circle at a fixated location disappeared during fixation (regardless of whether it was the correct location or not). Monkeys were rewarded only when they correctly repeated the whole sequence. The reward was 3 drops of juice per trial (0.5 ml); reward size was controlled by a computer-controlled solenoid.

Behavioral Analysis

We used NIMH MonkeyLogic (41) to control behavior and collect behavioral data. In each imaging session, we collected at least 700 trials in which monkeys completed the task, regardless of behavioral performance. After removal of unstable complete trials during calcium-imaging extraction preprocessing, about 500 complete trials remained from each imaging session for further neural data analysis. The behavioral data included 20 sessions (M1) and 13 sessions (M2).

All complete sequences in which monkeys reproduced the entire sequence were pooled for behavioral analysis. In the task, monkeys did not receive feedback about the correctness of their responses until the reproduction period was over and the chance level for each rank was equal to 1/6. By calculating the proportion of correct responses to different ranks, we could separately evaluate monkeys’ performances on different ranks in sequences with different lengths (Fig. 1B). Monkeys could make mistakes on both item and rank. For example, for a target sequence of [2 3 5] and a response sequence of [1 3 5], the monkey made an item error on rank 1. As another example, for a target sequence of [2 3 5] and a response sequence of [3 2 5], the monkey made two rank errors on ranks 1 and 2. We evaluated monkeys’ item errors and rank errors separately for all length-3 sequences and length-2 sequences (Fig. 1C).

Two-photon calcium imaging

Set-up

We performed *in vivo* two-photon imaging using a Thorlabs two-photon microscope and a Ti:Sapphire laser (Mai Tai eHP, Spectra Physics). A 16 \times objective lens (0.8 N.A., Nikon) imaged an area of 512 μm \times 512 μm or 800 μm \times 800 μm at 30 frames per second. The recording depth ranged from 250 μm to 350 μm below the pia. Each imaging session lasted for 2–3 hours. We tried to cover most of the regions in the recording windows located alongside the PS. There were 33 recording sessions in total (20 for M1 and 13 for M2). In each recording session, around 160 neurons on average were found in the corresponding imaged Field-of-view (FOV). These FOVs were distributed alongside the PS, including its anterior, posterior, ventral, and dorsal sides (fig. S2).

Data processing

The image processing pipeline was implemented in Python and JupyterLab. First, two-photon images were temporally down-sampled and spatially smoothed by a Gaussian filter. Then, the images were motion-corrected using a non-rigid motion-correction algorithm (42). However, there were some periods when motion artifacts were too large to be corrected or the motion was in the direction orthogonal to the imaging plane. These data were excluded by examining the spatial shifts and correlations with the reference image. Source extraction was performed with the CAIMAN package (42) based on constrained non-negative matrix factorization. A set of scores were calculated for each extracted component. Regions of interest were selected by thresholding these scores or, in ambiguous cases, human inspection (38). The resulting traces had a frame rate of 7.5 Hz for M1 and 10 Hz for M2.

In Fig. 1, $\Delta F/F$ was obtained by built-in function in CAIMAN. For other analysis, the calcium traces were first detrended and then normalized by the noise variance, which was estimated by calculating the variance of the residual signal after subtracting a Savitzky-Golay filtered version. For decoding, before feeding into the decoders, the calcium traces were further convolved with a half gaussian filter of 0.5 s width and de-means. See supplementary code for more details (38).

State space analysis (Related to Figure 2)

Linear regression

A multi-variable linear regression model was used to determine how various task variables affect the average neural response during the late delay period. A length-3 sequence can be represented as an 18-dimensional three-hot vector. For example, sequence [5 3 1] can be represented by

$(0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T$. We thus defined 18 one-hot vectors $S_{r,l}$ as task variables, where $r = 1, 2, 3$ and $l = 1, \dots, 6$. In our model, the average neural response of the i^{th} neuron in one trial during the late delay period (1 s before the onset of the “go” signal) was assumed to be a linear combination of the task variables, such that

$$y_i = \sum_{r=1}^3 \sum_{l=1}^6 \beta_i(r, l) S_{r,l} + \epsilon_i,$$

where $\beta_i(r, l)$ are unknown regression coefficients and ϵ_i is the trial-by-trial noise. To avoid overfitting, a Lasso regularization term was added to the linear regression model and the regularization amplitude with maximum likelihood was selected. For each neuron, trials were randomly split in half for 100 times, and for each time the linear regression model was applied separately on each half. As a result, we obtained 200 estimations for each regression coefficient.

Regression rank subspaces (Fig. 2A)

The regression coefficients $\beta_i(r, l)$ (averaged across 200 estimations) were used to identify the low dimensional subspaces containing most task-related variance. Specifically, with N neurons in total collected from all FOVs, an N -dimensional vector $\boldsymbol{\beta}(r, l) = [\beta_1(r, l), \dots, \beta_N(r, l)]^T$ was used to represent rank-item combination (r, l) at the neural population level. To capture the response variance due to item variation at each rank in this neural state space, we divided 18 $\boldsymbol{\beta}(r, l)$ ($r = 1, 2, 3$, $l = 1, \dots, 6$) into three groups along the rank index. For each group $\{\boldsymbol{\beta}(r, l)\}_{l=1,\dots,6}$, principal component analysis was performed to identify the first two axes that captured the most response variance (see fig. S3, A and J, for explained variance). In this way, we can further make the approximation that

$$\boldsymbol{\beta}(r, l) \approx V_r \boldsymbol{\kappa}(r, l) + \langle \boldsymbol{\beta}(r, l) \rangle_l,$$

where $V_r = [\mathbf{v}_r^1 \ \mathbf{v}_r^2]$ (of size $N \times 2$) with \mathbf{v}_r^1 and \mathbf{v}_r^2 ($\|\mathbf{v}_r^1\|_2 = \|\mathbf{v}_r^2\|_2 = \sqrt{N}$) forming an orthogonal basis for rank- r subspace, $\boldsymbol{\kappa}(r, l)$ (a vector of length 2) is the projected value onto this rank- r subspace under the basis $\{\mathbf{v}_r^1, \mathbf{v}_r^2\}$, and $\langle \boldsymbol{\beta}(r, l) \rangle_l$ (of dimension N) is the mean value independent of items. In other words, we obtained the collective variable $\boldsymbol{\kappa}(r, l)$ by projecting the de-meanned $\boldsymbol{\beta}(r, l)$ onto rank- r subspace. For ease of use in the following sections, we denoted V_r/\sqrt{N} as \hat{V}_r , a normalized version of V_r . To better compare spatial item representations across different rank subspaces, instead of using \mathbf{v}_r^1 and \mathbf{v}_r^2 as rank- r subspace axes, we used \mathbf{q}_r^1 and \mathbf{q}_r^2 (introduced in **Sequence representation** and denoted as rPC1 and rPC2 in Fig. 2A, respectively) as rank- r subspace axes.

Principal angles between rank subspaces (Fig. 2B)

To characterize how different rank subspaces were oriented in neural state space, we computed their principal angles, a measure for quantifying the alignment of two manifolds (43). Specifically, given two rank subspaces a and b and the associated two-dimensional bases \hat{V}_a and \hat{V}_b , a singular value decomposition (SVD) was performed onto the inner product matrix $\hat{V}_a^T \hat{V}_b$ (of size 2×2), such that $\hat{V}_a^T \hat{V}_b = P_a C P_b^T$, where both P_a and P_b are 2×2 orthogonal matrices and C is a 2×2 diagonal matrix whose elements are the ranked cosines of the principal angles θ_1 and θ_2 :

$$C = \text{diag}(\cos(\theta_1), \cos(\theta_2)).$$

The first principal angle was reported in the figures. For M1, the degrees of first principal angles between each rank subspaces were 74.7 ± 0.39 (rank 1-2), 69.7 ± 0.58 (rank 2-3), and 84.8 ± 0.39 (rank 1-3), while the degrees of first principal angles within each rank were 17.99 ± 0.26 (rank 1), 30.5 ± 0.53 (rank 2) and 34.4 ± 0.49 (rank 3). For M2 length-2 trials, the degree of first principal angle between rank subspaces was 70.2 ± 0.51 , while the degrees of first principal angles within each rank were 17.4 ± 0.39 (rank 1) and 27.3 ± 0.66 (rank 2). All values here follow the format of *mean \pm s.t.d.*

Variance accounted for (VAF) ratio (Fig. 2C)

To further quantify the alignment of different rank subspaces, we defined a measure called variance-accounted-for (VAF) ratio (43) as follows. For given spatial location l and rank r , denote $V_r \boldsymbol{\kappa}(r, l)$ as $\mathbf{g}_r(l)$. The VAF ratio for rank subspace pair (a, b) was defined as

$$VAF_{ab} = \frac{\text{Var}(\hat{V}_b \hat{V}_b^T G_a)}{\text{Var}(G_a)},$$

where $G_a = [\mathbf{g}_a(1), \dots, \mathbf{g}_a(6)]$ (of size $N \times 6$). Note that the VAF ratio depends on the order of a and b .

It equals 0 if the two subspaces are orthogonal and equals 1 if they completely overlap with each other.

Gain-modulation approximation of collective variables (Fig. 2D)

Here, we asked how well a simple gain-modulation model can approximate the full linear regression model at the collective variable level. Specifically, we approximated $\kappa(r, l)$ with a gain modulation formula, such that

$$\kappa(r, l) = \lambda_r O_r \mathbf{f}_l + \zeta(r, l), \quad (\text{Eq. S1})$$

where λ_r is the rank modulation factor, \mathbf{f}_l is the spatial location vector (of length 2) independent of rank r , O_r (of size 2×2) is an orthogonal matrix and $\zeta(r, l)$ is the approximation error. After approximation, the rank modulation factors were rescaled based on the mean of λ_1 values. And the spatial location vectors were accordingly adjusted. For M1, the rank modulation factors were 1.00 ± 0.045 , 0.64 ± 0.029 , 0.57 ± 0.026 , respectively. For M2, the rank modulation factors were 1.00 ± 0.08 , 0.74 ± 0.06 , respectively. All values here follow the format of *mean \pm s.t.d.*

In practice, we obtained O_r , λ_r and \mathbf{f}_l through minimizing the Frobenius norm of the errors ζ . For rank r , the similarity score between $\kappa(r, \cdot)$ (of size 2×6) and the approximation is defined as

$$1 - \frac{\|\zeta(r, \cdot)\|_2^2}{\|\kappa(r, \cdot)\|_2^2}.$$

The similarity scores for each rank in M1 were 0.96 ± 0.0044 , 0.99 ± 0.0009 and 0.97 ± 0.0048 . The scores in M2 were 0.99 ± 0.0007 and 0.98 ± 0.0013 . All values here follow the format of *mean \pm s.t.d.*

Sequence representation (Fig. 2E)

For ease of interpretation, we denote $V_r O_r$ (of size $N \times 2$) as $Q_r = [\mathbf{q}_r^1 \mathbf{q}_r^2]$, where \mathbf{q}_r^1 and \mathbf{q}_r^2 ($\|\mathbf{q}_r^1\|_2 = \|\mathbf{q}_r^2\|_2 = \sqrt{N}$) form an orthogonal basis of rank-r subspace. By doing so, the gain modulation approximation of the demeaned $\beta(r, l)$ can be written as $\lambda_r Q_r \mathbf{f}_l$, which states that the same filler \mathbf{f}_l is shared across different rank subspaces and ordinal rank acts as a gain modulator at the collective variable

level. Under this approximation, for sequence $[l_1 l_2 l_3]$, the average neural response $\mathbf{y}([l_1 l_2 l_3])$ during the late delay period in the neural state space can be written as

$$\mathbf{y}([l_1 l_2 l_3]) = \sum_{r=1}^3 \lambda_r Q_r \mathbf{f}_{l_r} + \sum_{r=1}^3 \langle \boldsymbol{\beta}(r, l) \rangle_l. \quad (\text{Eq. S2})$$

Decoding analysis (Related to Figure 3)

Cross-temporal decoding

In Fig. 3, each decoder was trained on the data within a half-second time window. Here we use length-3 trials as an example. Within one trial, there were 19 sampled time windows in total with starting points as follows: 1 s before the first target onset, the first target onset and 0.3 s after that, the second target onset and 0.3 s after that, the third target onset and 8 time points at 0.3 s intervals following the third target onset, 0.1 s after fixation off (the go signal), the moments when monkey's eyes moved away from each of the three saccade targets, and one more time point 0.3 s after the last response. Decoders trained in one time window were tested in all 19 time windows to obtain the cross-temporal decoding graphs. In leave-one-trial-out settings, we used the decoders trained with remaining trials to predict target item of the left-out trial in all time windows. The decoded items were then pooled across folds to calculate decoding accuracy. Additionally, the cross-temporal heat maps were smoothed by a two-dimensional Gaussian kernel with 1/3 bin-width when plotting.

To determine the efficacy of decoding, we compared the performance of each trained decoder with a dummy decoder that makes random predictions while respecting the original output distribution. The distribution was calculated for each time window by feeding the trained decoder with all the data (including error trials) at that time. We utilized pixel-based statistics to correct for multiple comparisons (44). Fifty thousand heat maps were generated with dummy decoders and the maximum of each heat map was gathered to construct a null distribution. The yellow contour encloses the area in which decoding accuracy was greater than the $1 - p$ percentile of this distribution, where p is the p-value we reported. For more details, please refer to supplementary code (38).

Decoder architecture (Fig. 3A)

We used linear decoders of the same structure throughout the paper, implemented in PyTorch. For a given trial and time window, to decode rank-r item, the calcium activity \mathbf{x} of N simultaneously recorded neurons was first projected to a low-dimensional decoding space through a linear transformation

$$\mathbf{h}_r = \mathbf{W}_r \mathbf{x} + \mathbf{b}_r,$$

where \mathbf{W}_r is an $N_h \times N$ matrix and \mathbf{b}_r is an N_h -dimensional bias term. Next, we computed the dot product of \mathbf{h}_r and N_s target vectors. In matrix terms, \mathbf{h}_r was multiplied by an $N_s \times N_h$ target matrix M , which in our case was of size 6×2 . The outcome was then softmaxed to obtain the scores

$$\mathbf{p}_r = \text{softmax}(M\mathbf{h}_r)$$

for N_s possible results. The item with the largest score was selected as the decoded item. Note that the target matrix M was assumed to be rank invariant. That is, the activity patterns in decoding subspaces shared similar spatial layout across ranks. This assumption was based on the finding that activity in regression-based rank subspace can be well approximated by the same geometric structure (Fig. 2).

Decoder training

For a given time window, the parameters \mathbf{W}_r , \mathbf{b}_r and M were trained simultaneously by gradient descent with standard Adam algorithm, where the loss function for rank-r consisted of cross-entropy loss and L2 penalty on \mathbf{W}_r and M :

$$\text{Loss} = \text{CELoss} + \frac{\gamma}{m} (\text{L2}(\mathbf{W}_r) + \rho \text{L2}(M)),$$

where γ is the overall factor for balancing the two losses, m is the mini-batch size, and ρ is the scaling factor so that less constraints were imposed on M . We set $\gamma = 20$ and $\rho = 0.1$ in all decoders. As the trial numbers were not balanced across different experimental conditions, loss for each trial was weighted by the corresponding condition when training in batches. The weighting factor for trials of condition s was calculated as $\alpha_s = (\theta \cdot C_s + \bar{C})^{-1}$, where C_s is the trial count of condition s , \bar{C} is the mean trial count of all conditions, and $\theta = 3$ is the coefficient to control the degree of balancing. The resulting α_s values were then normalized across all trials so that they summed to the total trial count.

In practice, an ensemble of 50 decoders was trained for each situation. \mathbf{W}_r was initialized with different

random realizations, while \mathbf{b}_r and M were initialized with zero vectors and a 2-dimensional ring structure, respectively. We fixed the learning rate at 1e-3 and progressively increased the batch size (from 32 to 128) when training loss plateaus, mimicking the effect of reducing learning rate. The decoders with the 10th- to 20th-smallest training loss were used in the decoding analysis and the results displayed in the figures are averages across these ten decoders. Data from FOV2 was set aside for tuning hyperparameters, which were selected based on cross-validation results of correct length-3 trials. See supplementary code (38) for all the parameter used and more training details.

Decoder-based state space trajectories (Fig. 3D)

To plot the trajectories in Fig. 3D, we first trained decoders in a 1.33 s time window during the late delay period to get projection matrices W_r . We then performed SVD on these projection matrices, such that

$$W_r = U_r^D \Sigma_r^D V_r^{DT},$$

where V_r^D is an $N \times 2$ matrix with two orthonormal column vectors. We refer to the subspace spanned by these two orthonormal column vectors as the decoding-based rank-r subspace. Next, we took continuous calcium traces of 4.67 s long and aligned them to 0.4 s before the first target onset. The traces were averaged across trials of the same sequence before being projected to different decoding-based rank subspaces. The resulting sequence-wise trajectories were then rotated to an appropriate viewing angle for illustrative purposes.

Single neuron basis of rank subspace (Related to Figure 4)

Projecting single neuron unit vector onto rank subspaces (Fig. 4A)

To characterize the geometric relationship between single neuron axes and rank subspaces, we retrieved the entries for single neuron from the basis of rank subspaces. In other words, we projected unit vector \mathbf{e}_i of the axis of neuron i onto rank- r subspace. We found that it is useful to introduce the following two quantities, the scalar projection A_{ri} and the angle φ_{ri} , such that

$$(\mathbf{q}_r^1)^T \mathbf{e}_i = A_{ri} \cos(\varphi_{ri}),$$

$$(\mathbf{q}_r^2)^T \mathbf{e}_i = A_{ri} \sin(\varphi_{ri}),$$

in which A_{ri} measured the degree of alignment between neuron i 's axis and rank-r subspace, and φ_{ri} specified the spatial item preference of neuron i in rank-r subspace, as illustrated in Fig. 4A.

Participation ratio (PR) of rank subspaces (Fig. 4B)

To characterize how distributed the rank-r subspace is, we introduced a quantity called PR (45) based on the distribution of square scalar projection across the neural population as follows

$$PR_r = \frac{(\sum_{i=1}^N A_{ri}^2)^2}{\sum_{i=1}^N A_{ri}^4}.$$

Normalized PR value is PR normalized by N . To get an intuitive understanding of PR_r , let us consider the following two scenarios. In the first scenario, rank-r subspace is evenly distributed across the population, such that A_{ri} is constant. Since $\sum_{i=1}^N A_{ri}^2 = Nd_r$ where d_r is the subspace dimension, it follows that $A_{ri} = (d_r)^{1/2}$ and $PR_r = N$. In the second scenario, rank-r subspace relies on the minimum number of d_r neurons. $A_{ri} = 1$ for these neurons and otherwise 0, so that $PR_r = d_r$. These two scenarios correspond to the most distributed and most localized rank subspaces, respectively. In general, the larger the PR is, the more distributed the rank subspace is.

Examining rank-dependence of single neuron item preference (Fig. 4C)

To examine if the above-defined item preference φ_{ri} is rank-dependent, for any given rank pair, we computed single neuron item preference difference and plotted its distribution over the neural population. To control the effect of noise, we only selected neurons with the significantly large square scalar projection in the population for both ranks (estimated according to the corresponding normalized PR value) in the pair. Specifically, for rank-1 (2 and 3), the significantly large square scalar projection means the top 38% (34% and 32%), respectively. Under this criterion, there were 531 (406 and 564) neurons left for 1-2 (1-3 and 2-3) rank pairs, respectively. For M2 (fig. S8E), there were 321 neurons for 1-2 rank pair.

Relationship of two quantities characterizing item preference (Fig. 4E)

Since the item preference φ_{ri} is defined with respect to rank-r subspace, it is in principle different from item preference directly extracted from the raw regression coefficient $\beta_i(r, l)$. To explore the potential

relationship between them, we projected the population activity $Q_r(\lambda_r \mathbf{f}_l)$ under gain-modulation approximation in rank- r subspace onto neuron i 's axis: $\mathbf{e}_i^T Q_r(\lambda_r \mathbf{f}_l)$, where $\mathbf{e}_i^T Q_r = \mathbf{e}_i^T [\mathbf{q}_r^1 \mathbf{q}_r^2]$ can be rewritten as $A_{ri}[\cos(\varphi_{ri}), \sin(\varphi_{ri})]$. Under the assumption that $\mathbf{f}_l = [\cos(\theta_l), \sin(\theta_l)]$, we have

$$\mathbf{e}_i^T Q_r(\lambda_r \mathbf{f}_l) = \lambda_r A_{ri} \cos(\theta_l - \varphi_{ri}).$$

which states that the tuning curve extracted from the rank subspace can be approximated by a cosine function. Therefore, the angle φ_{ri} , a quantity characterizing the geometric relationship between neuron i 's axis and rank- r subspace under gain-modulation approximation, may have a good correspondence with the item preference from the raw tuning $\beta_i(r, l)$, which was confirmed by our neural data (Fig. 4E). The correspondence was confirmed by one-way ANOVA with multiple comparison correction (Tukey's honest significant difference criterion).

Characterizing the functional map (Related to Figure 5)

To examine whether neurons with similar degree of alignment with subspaces are spatially clustered in monkey cortex, a clustering metric was computed. Given rank r and cortical distance d in the recording plane, the metric $C_{align}(r, d)$ takes the expected normalized subspace-to-neuron alignment (SNA) difference at rank r between cell pairs separated by d . The normalized SNA difference at rank r between neurons i and j is defined as

$$SNA_{i,j}(r) = \frac{|A_{ri}^2 - A_{rj}^2|}{|A_{ri}^2 + A_{rj}^2|}.$$

Then, for a given FOV, the clustering index can be defined as

$$C_{align}(r, d) = \frac{\mathbb{E}[SNA_{resampled}(r)]}{\mathbb{E}[SNA_{actual}(r)|d]},$$

where the denominator is the average normalized SNA difference at rank r between cell pairs separated by d and the nominator is the average normalized SNA difference at rank r between randomly selected cell pairs. As a control we shuffled the locations of all neurons 1000 times, then calculated the C_{align} values for the shuffled data. Only when the minimum actual C_{align} value was larger than 99.9% of the shuffled data, these C_{align} values were regarded as significantly larger than shuffle level. Only when the maximum actual C_{align} value was smaller than 99.9% of the related shuffled data, these C_{align} values were regarded as

significantly smaller than shuffle level.

Similarly, to examine whether neurons with similar tuning property are spatially, another clustering metric was computed. Given rank r and cortical distance d in the recording plane, the metric $C_{tuning}(r, d)$ takes the normalized expected spatial location tuning dissimilarity at rank r between cell pairs separated by d (46). The spatial location tuning similarity at rank r between neurons i and j is defined as

$$\rho_{i,j}(r) = \frac{\text{Cov}_l(\hat{\beta}_i(r, l), \hat{\beta}_j(r, l))}{\sqrt{\text{Var}_l(\hat{\beta}_i(r, l))} \sqrt{\text{Var}_l(\hat{\beta}_j(r, l))}},$$

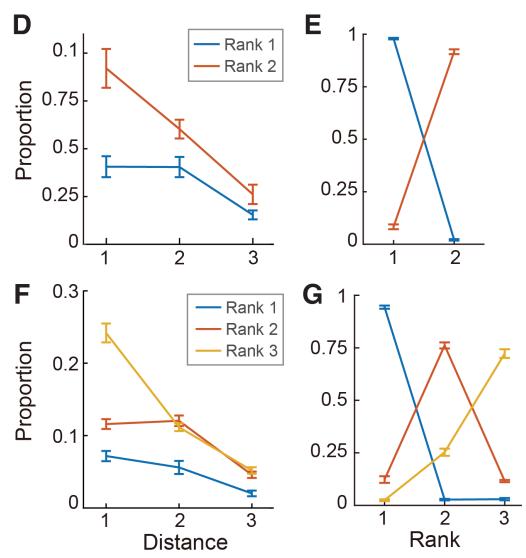
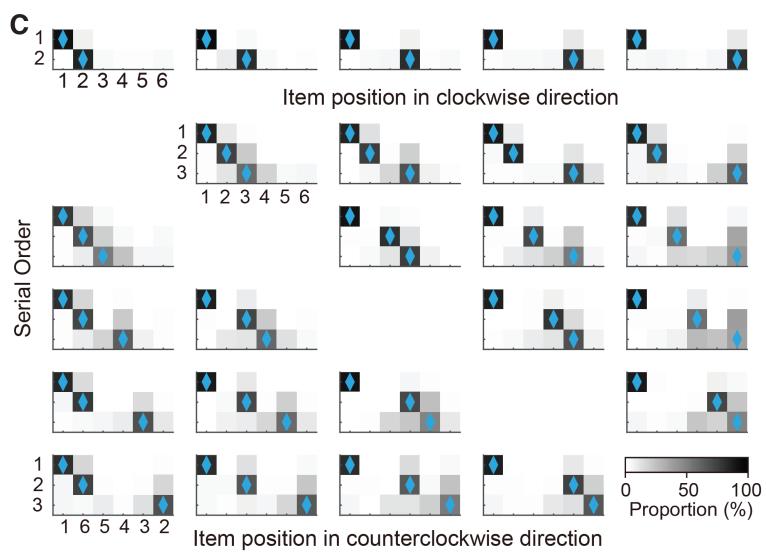
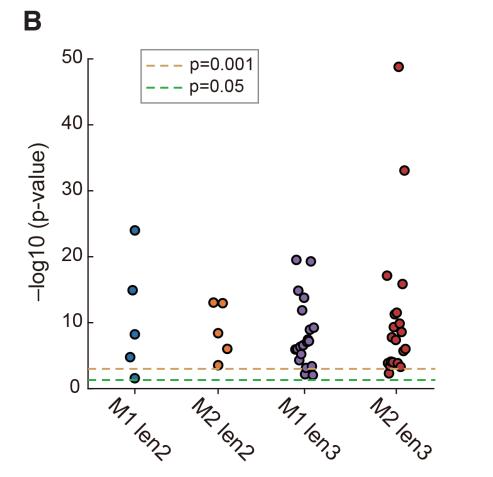
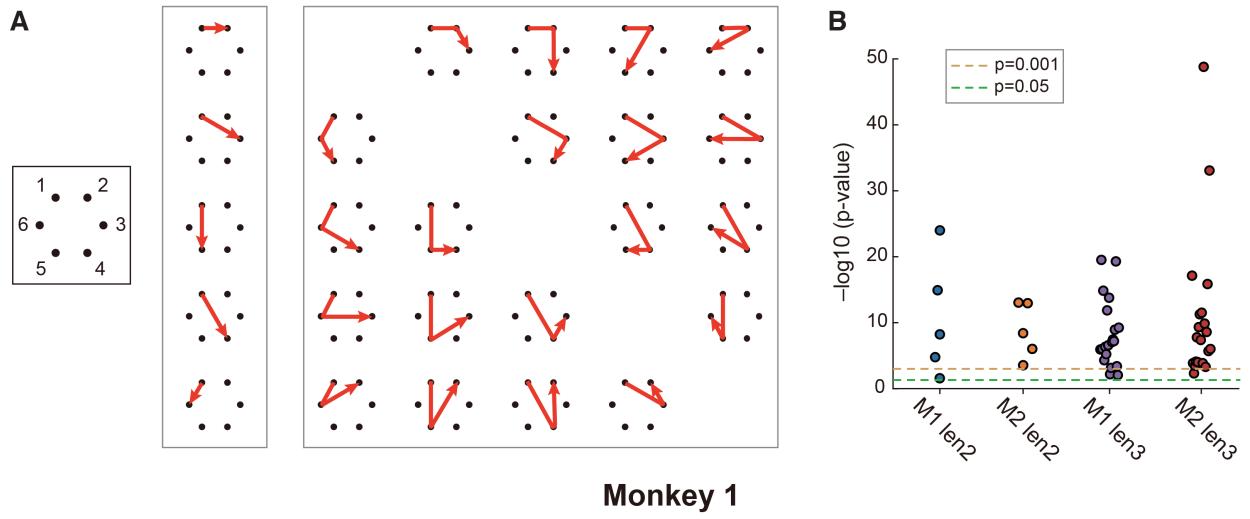
where $\hat{\beta}_i(r, l) = \mathbf{e}_i^T Q_r(\lambda_r \mathbf{f}_l)$. Then, the clustering index can be defined as

$$C_{tuning}(r, d) = \frac{\mathbb{E}[1 - \rho_{resampled}(r)]}{\mathbb{E}[1 - \rho_{actual}(r)|d]},$$

where the denominator is the average tuning dissimilarity at rank r between cell pairs separated by d and the nominator is the average tuning dissimilarity at rank r between randomly selected cell pairs. The measurement of signficancy was similar with C_{align} index.

References

39. T. W. Chen *et al.*, Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
40. M. Li, F. Liu, H. Jiang, T. S. Lee, S. Tang, Long-Term Two-Photon Imaging in Awake Macaque Monkey. *Neuron* **93**, 1049–1057.e3 (2017).
41. J. Hwang, A. R. Mitz, E. A. Murray, NIMH MonkeyLogic: Behavioral control and data acquisition in MATLAB. *J. Neurosci. Methods* **323**, 13–21 (2019).
42. A. Giovannucci *et al.*, CaImAn an open source tool for scalable calcium imaging data analysis. *eLife* **8**, e38173 (2019).
43. J. A. Gallego *et al.*, Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nat. Commun.* **9**, 4233 (2018).
44. M. X. Cohen, ‘Nonparametric Permutation Testing’ in *Analyzing Neural Time Series Data: Theory and Practice* (MIT Press, 2014), chap. 33.
45. P. Gao *et al.*, A theory of multineuronal dimensionality, dynamics and measurement. *bioRxiv* 214262 [Preprint] (2017).
46. I. Nauhaus, K. J. Nielsen, A. A. Disney, E. M. Callaway, Orthogonal micro-organization of orientation and spatial frequency in primate primary visual cortex. *Nat. Neurosci.* **15**, 1683–1690 (2012).



Monkey 2

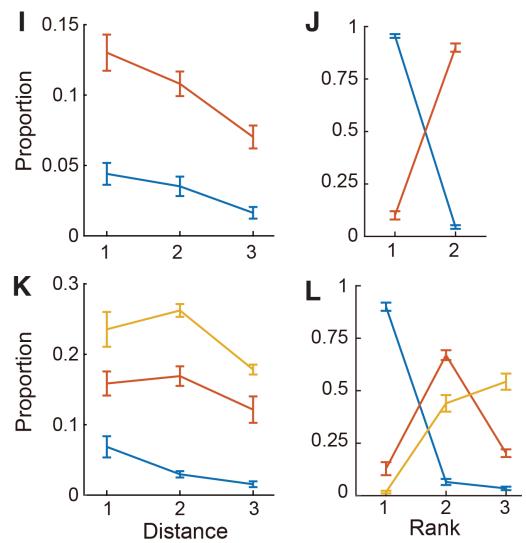
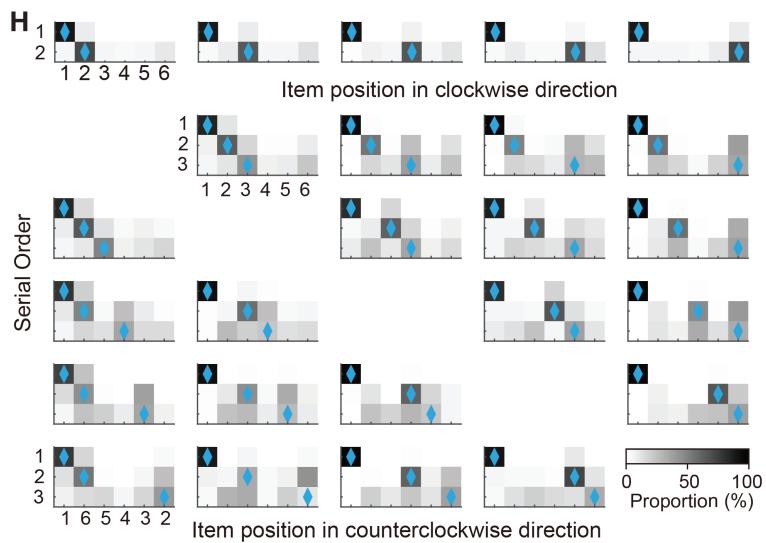


fig. S1. Illustration of sequences and behavioral performance for both monkeys.

- (A) Each polygonal line with an arrow corresponds to a set of sequences that share the same geometric pattern but have different starting points. The arrowhead indicates the end point.
- (B) For any given geometric pattern, the performance deviation of specific sequence from the pattern average was significant (one-way ANOVA within each geometric pattern group) for length-2 and -3 sequences in both monkeys, suggesting that the monkeys did not learn geometric patterns.
- (C) Response distribution of different sequence patterns illustrated in (A). The gray level indicates the percentage of responses for locations in each rank, where diamonds mark the correct location. For simplicity, performance was averaged after normalizing the starting point to 1.
- (D and E) Spatial (left) and ordinal (right) transposition of Monkey 1 (M1) for length-2 sequences. Spatial transposition error is shown as a function of spatial error distance, averaged across different ranks. Ordinal response pattern is shown as a function of ordinal rank, averaged across different spatial locations. Error bars, SEM.
- (F and G) Spatial (left) and ordinal (right) transposition error patterns of M1 for length-3 sequences.
- (H to L) Results from Monkey 2 (M2), the same format as M1.

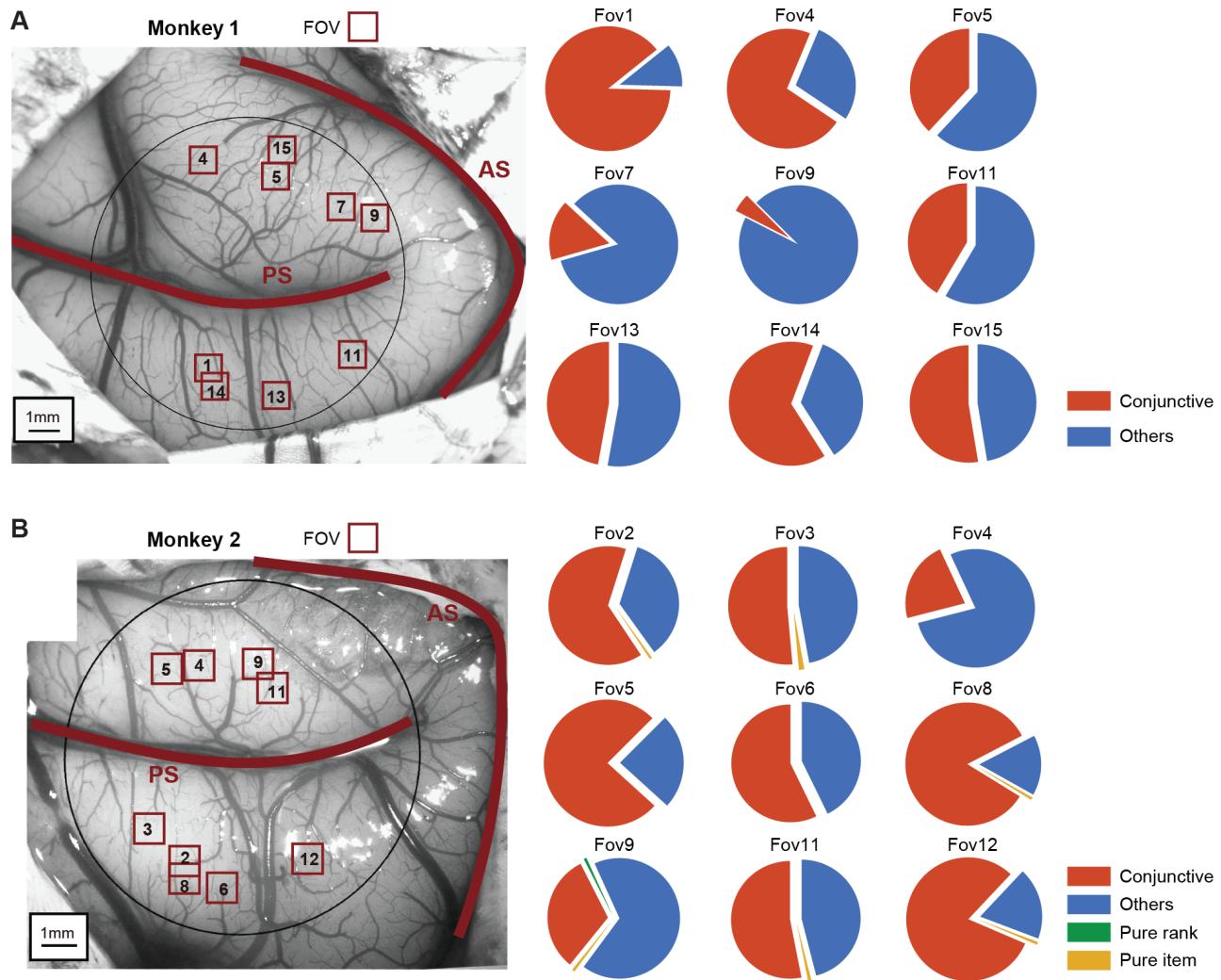


fig. S2. Recording sites and proportion of conjunctive neurons in individual FOVs.

(A) Example recording sites in M1. The left panel showed the distribution of FOVs labeled by red squares (not to scale with the reference bar). Principal sulcus (PS) and arcuate sulcus (AS) are indicated by red curves. The right panels display the distribution of different functional types of neurons within each FOV. Anatomically, no FOV shared recording plane. Recording sites not shown were in close proximity to the marked FOVs (FOV 16 and 15, FOV 6 and 5, FOV 8 and 7, FOV 10 and 9, FOV 12 and 13, FOVs 19, 20 and 14, FOVs 2, 3, 17, 18 and 4).

(B) The same result for M2. No FOV shared recording plane, though a couple of FOVs were largely overlapping (FOV 10 and 9, FOV 1 and 2, FOV 7 and 6). In order to characterize how neural responses were affected by rank and item information, we used two-way ANOVA analysis (rank \times item) on our regression coefficients. If a neuron's regression coefficients only varied across items (rank), it is a pure item (rank) neuron. A neuron is conjunctive if its regression coefficients varied across both rank and item. The remaining ones were grouped as 'others'.

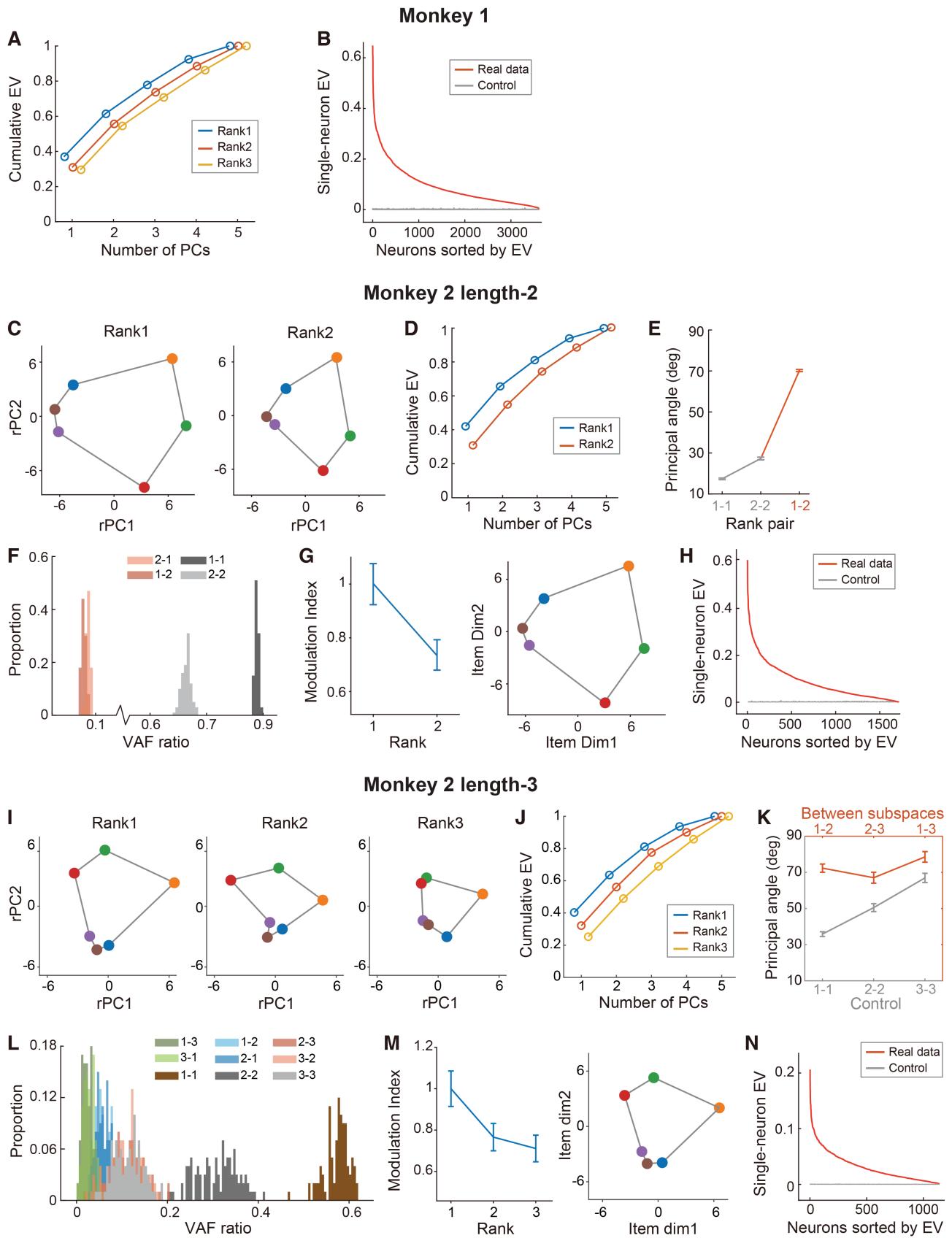


fig. S3. Explained variance analysis of M1 and neural state space analysis for M2.

- (A) The cumulative explained variance along difference ranks in neural state space.
- (B) Histogram of single neuron explained variance in our linear regression model.
- (C) The averaged population response for a given rank-location combination projected to the corresponding rank subspace. Analysis based on length-2 sequences of M2.
- (D) The cumulative explained variance along difference ranks in neural state space.
- (E) Principal angle of different rank subspace pairs.
- (F) The VAF ratio for different rank subspace pairs.
- (G) Gain modulation approximation of collective variables in rank subspaces.
- (H) Histogram of single neuron explained variance.
- (I to N) Results for length-3 sequences of M2. Note that, due to the relatively low precision of rank subspace estimation for the length-3 trials of M2, some of the subsequent analyses (e.g., single neuron basis of rank subspace and characterizing the functional map) did not include the length-3 trials.

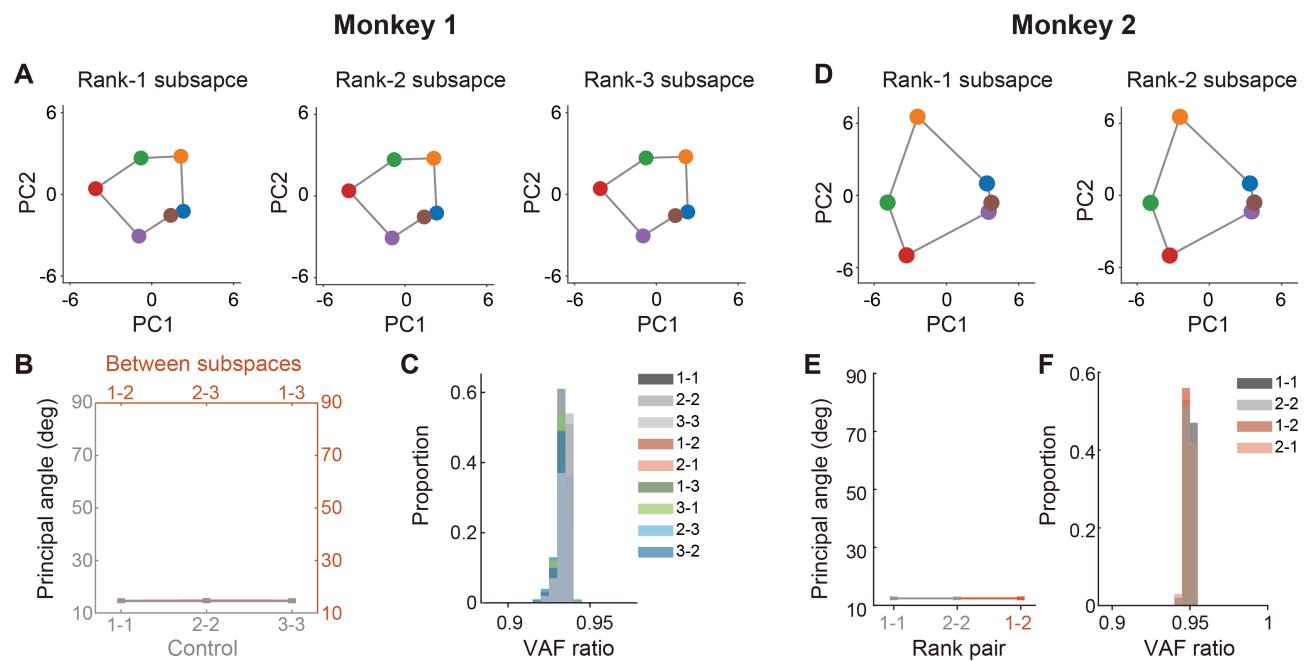
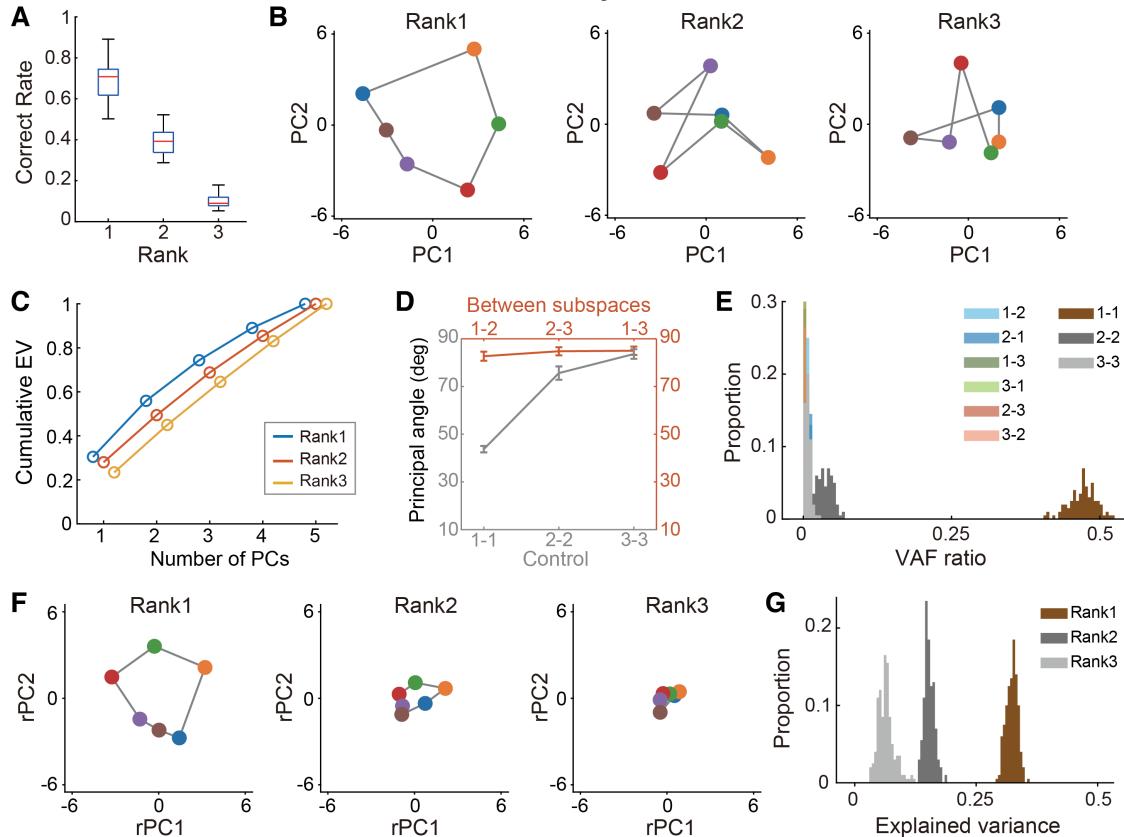


fig. S4. State space analysis of rank-shuffled data.

(A to C) Population analysis based on regression with 18 conjunctive variables when ranks are shuffled while items are held to be constant for M1. Spatial items in different ranks were shown in corresponding rank subspaces (A). The principal angles and VAF ratios between different rank subspaces were shown in (B) and (C), respectively. For each trial, the related items' ranks were randomly shuffled for 50 times, then the state space analysis was carried out.

(D to F) Results for M2.

Monkey 1



Monkey 2

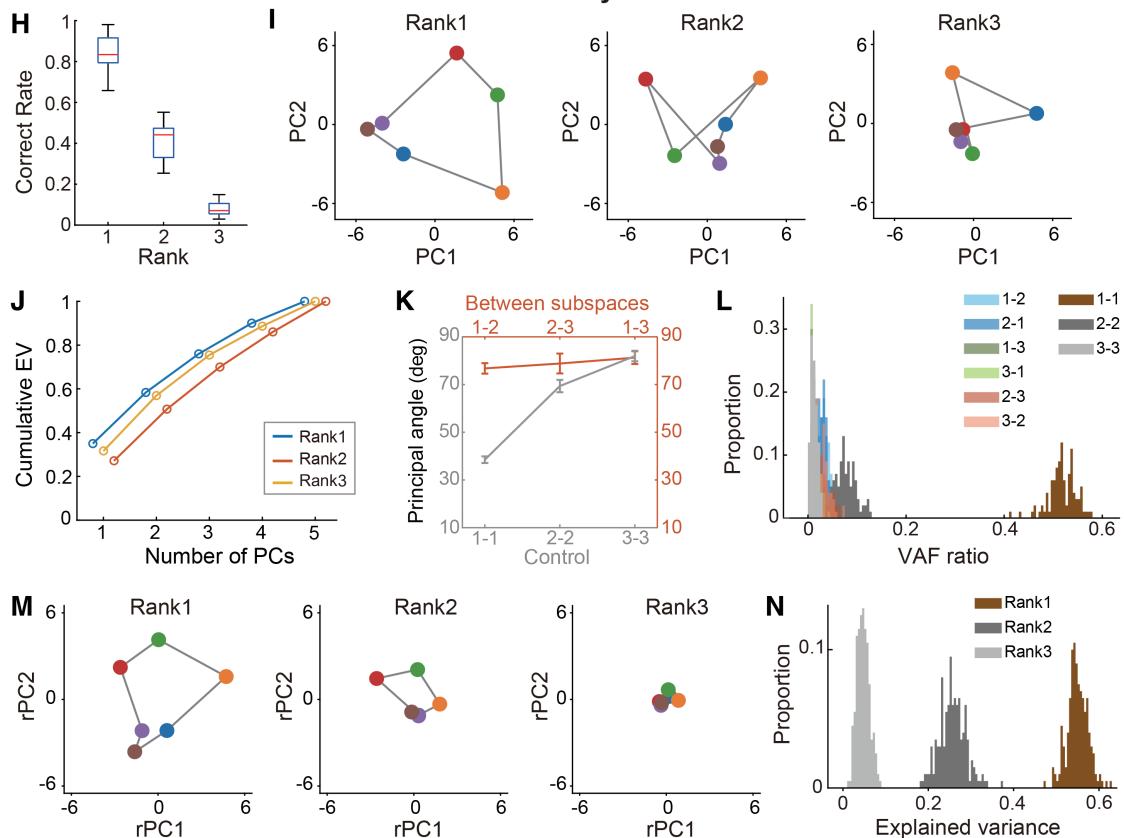


fig. S5. Neural state space analysis of error trials.

- (A) Behavioral performance on error trials of length-3 sequences averaged across sessions for M1. Note that the error trial number used in this analysis was comparable with the correct trial number for M1 in Fig. 2.
- (B) The population responses for different rank-location combinations in error trials projected to the corresponding error-trial-based rank subspaces for M1. Error-trial-based rank subspaces were defined as those two-dimensional linear manifolds in the neural state space that captured the most neural response variance due to spatial location at the corresponding ranks.
- (C to E) Cumulative explained variance as a function of PC number for different rank subspaces (C), principal angle of different rank subspace pairs (D) and VAF ratio of different rank subspace pairs (E) for M1.
- (F) The population responses for different rank-location combinations in error trials projected to the corresponding correct-trial-based rank subspaces (subspaces used in Fig. 2A) for M1.
- (G) The explained variance histogram when projecting the population responses for different rank-location combinations in error trials to the corresponding correct-trial-based rank subspaces.
- (H to N) Results for M2. Note that the error trial number used in this analysis was comparable with the correct length-3 trial number for M2 in Fig. S3.

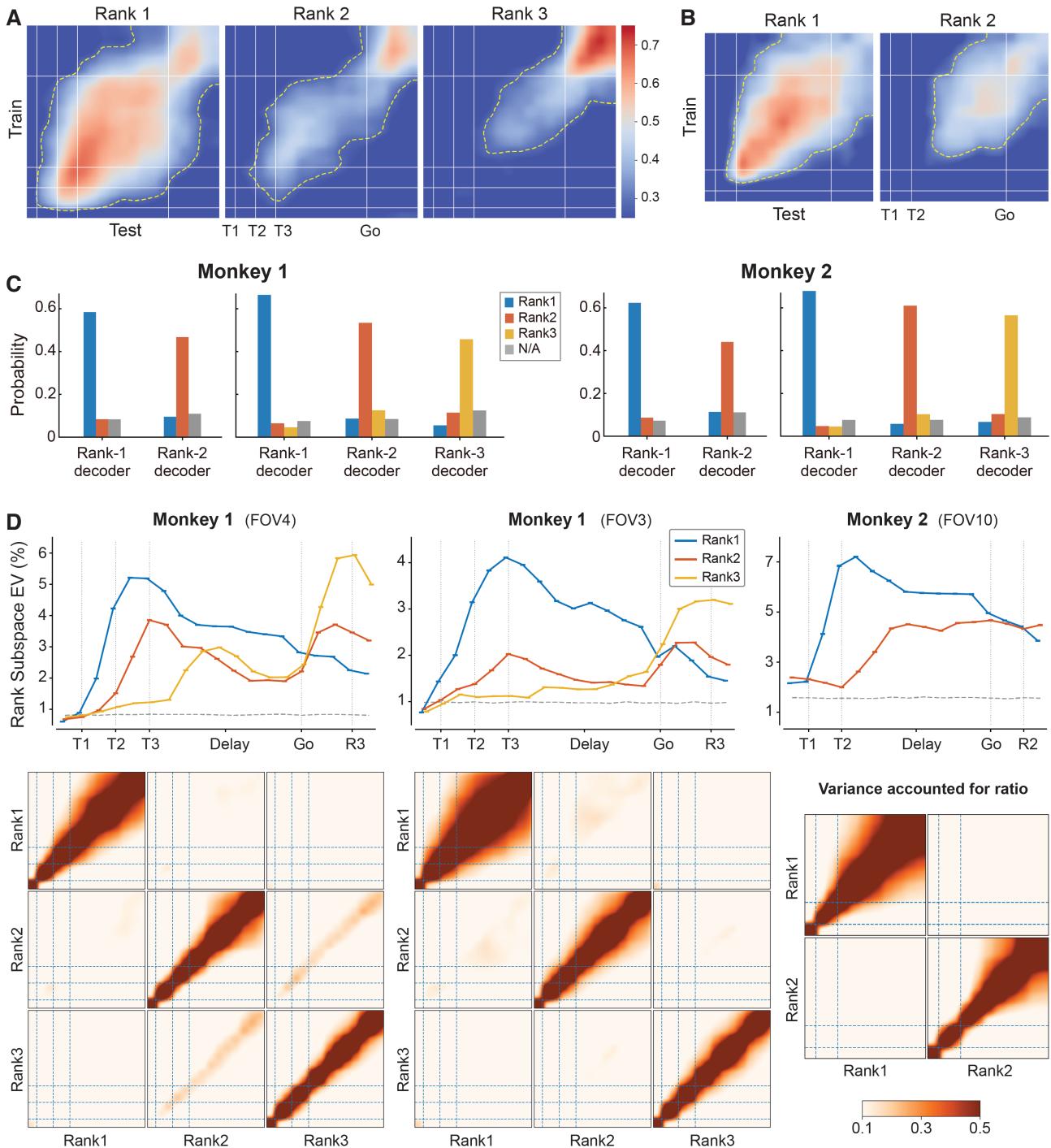


fig. S6. Cross-temporal decoding and decoder analysis.

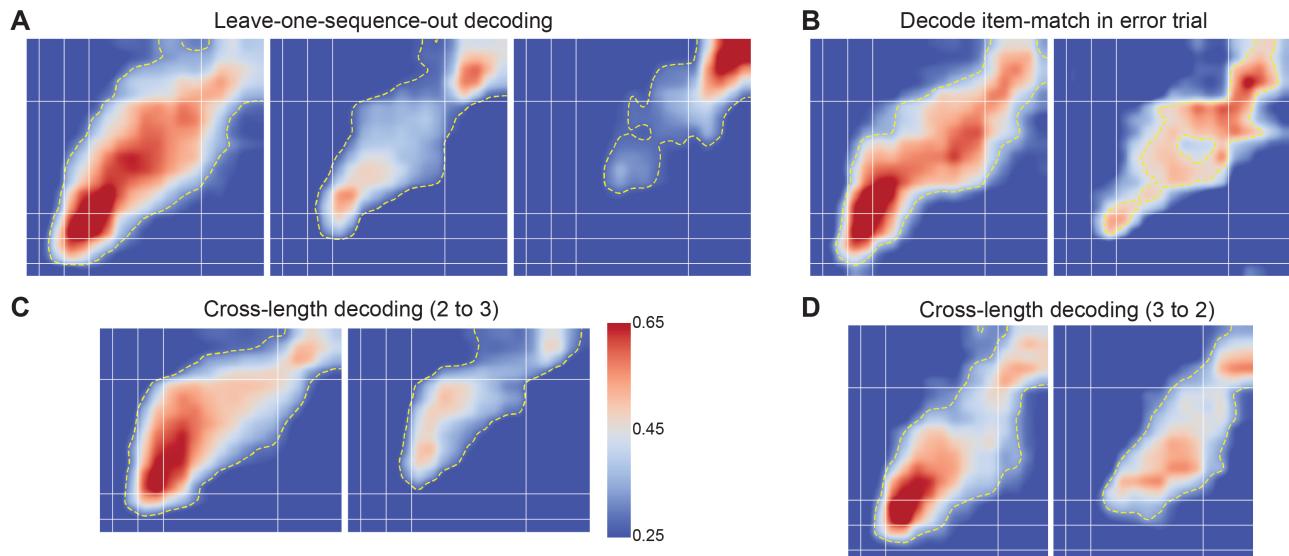
(A) Cross-temporal decoding in another session of M1 (FOV3). Colormap, 0.25-0.75. Contour, $p < 0.001$.

(B) Same as in (A) for M2 (FOV10), but for length-2 sequences.

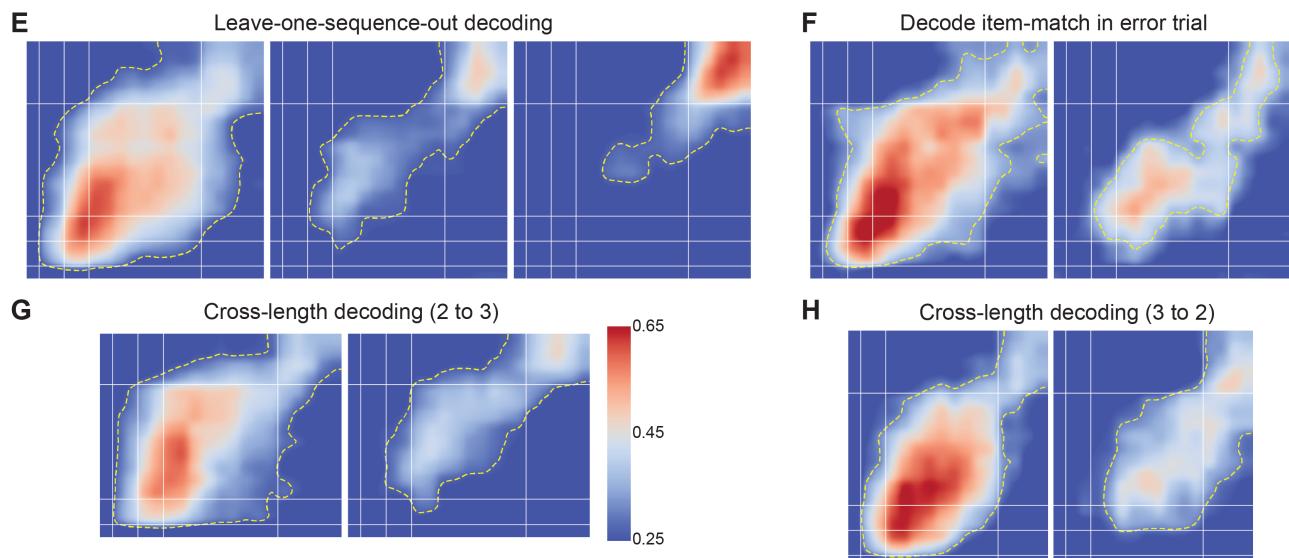
(C) The probability of decoded item matching the target of each rank (rank-1, -2 and -3) or none of the items in the sequence (N/A). Decoders were trained in late delay with length-2 or length-3 trials for both monkeys. The N/A probability is normalized by the number of items that were not present in the sequence. The results were averaged across FOVs with significant decoding performance ($p < 0.01$) for all the ranks involved (see table S1 and S2).

(D) Analysis of decoding-based subspace for three example sessions. Similar to Fig. 3D, we performed SVD to W_r for each decoder and defined the decoding-based rank- r subspace as the space spanned by the first two right singular vectors. Top, variance explained by rank-specific subspaces of decoders trained at different time windows for each rank. Each data point averaged from all 50 trained decoders with standard deviation error bars. Bottom, cross-rank and cross-temporal variance accounted for ratio (see Methods). The signal in decoding-based subspaces in the y-axis was projected to subspaces in the x-axis to calculate the ratio.

Monkey 1 (FOV4)



Monkey 1 (FOV3)



Monkey 2 (FOV10)

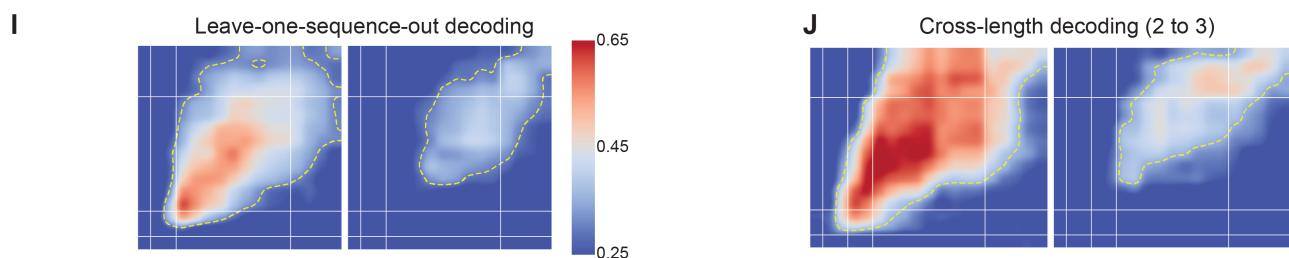


fig. S7. Extended results for the decoding of spatial location.

(A to D) Cross-temporal decoding results spanning the entire trial for M1 (FOV4), related to Fig. 3, E to H. Contour, $p < 0.005$.

(E to H) Results for another recording session (FOV3) of M1.

(I and J) Results of M2 (FOV10).

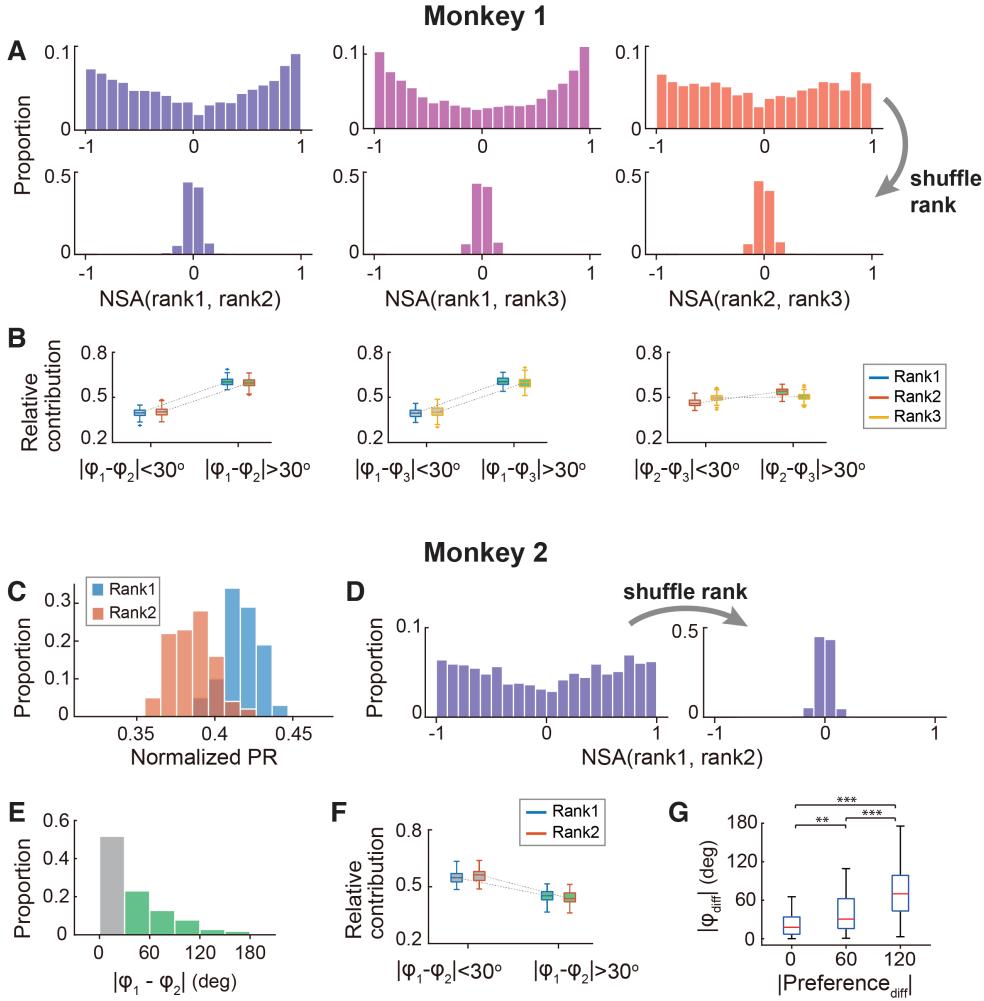


fig. S8. Single neuron basis of rank subspaces.

(A) To determine whether a single neuron axis primarily aligns with one particular subspace or equally aligns with multiple subspaces, we defined a quantity called neuron-to-subspace alignment (NSA) index. Given neuron i and two rank subspaces a and b , the NSA index is defined as: $\text{NSA}_i(a, b) = (A_{ai}^2 - A_{bi}^2)/(A_{ai}^2 + A_{bi}^2)$. The value of NSA index ranges between -1 and 1 . Having an $\text{NSA}_i(a, b)$ value close to 1 (or -1) indicates that neuron i primarily aligned with rank- a (or b) subspace. By contrast, when $\text{NSA}_i(a, b)$ is close to 0 , this indicates that neuron i equally aligned with rank- a and b subspaces.

Histograms of NSA index for different rank pairs show that many neurons were mostly aligned to only one subspace. Such disentanglement disappeared when the rank information was shuffled (bottom). Specifically, we performed linear regression on rank shuffled data and calculated the NSA indices, which were found to be distributed round 0 .

(B) To quantify the contribution of those neurons exhibiting location-preference-shift property (Figure 4C, green bars) to different rank subspaces, for given rank pair, we separated neurons into two groups, one (group 1) consisting of neurons with preference shift less than 30 degree and the other (group 2) consisting of neurons with preference shift larger than 30 degree. Then, for each group, we defined a contribution index $\sum_{i \in G \cap S_{(a,b)}} A_{ai}^2 / \sum_{i \in S_{(a,b)}} A_{ai}^2$, where G is the group index and $S_{(a,b)}$ is the set of neurons satisfying the criterion used in Figure 4C for rank pair (a, b) . The larger this value is, the larger the contribution of the corresponding neural group to rank- a subspace is. Here, this index was shown as a function of neural groups

for different rank pairs (1-2, left; 1-3, middle; 2-3, right). For example, the left panel showed that, for both rank 1 (blue) and 2 (red), the contribution index of group 2 is significantly larger than that of group 1, manifesting the significant contribution of those neurons exhibiting location-preference-shift property to both rank-1 and -2 subspaces. The similar trend holds for other panels. The relative contributions neurons with preference-shift larger than 30-degree were significantly larger than 0 (one-tailed t-test, $p<<0.001$).

(C) Normalized PR for different rank subspaces for M2.

(D) Histogram of NSA index for original data (left) and rank-shuffled data (right).

(E) The φ_r difference between 1-2 rank pair.

(F) Quantification of the contribution of those neurons exhibiting location-preference-shift property to different rank subspaces.

(G) The correspondence of φ_r difference and preference difference.

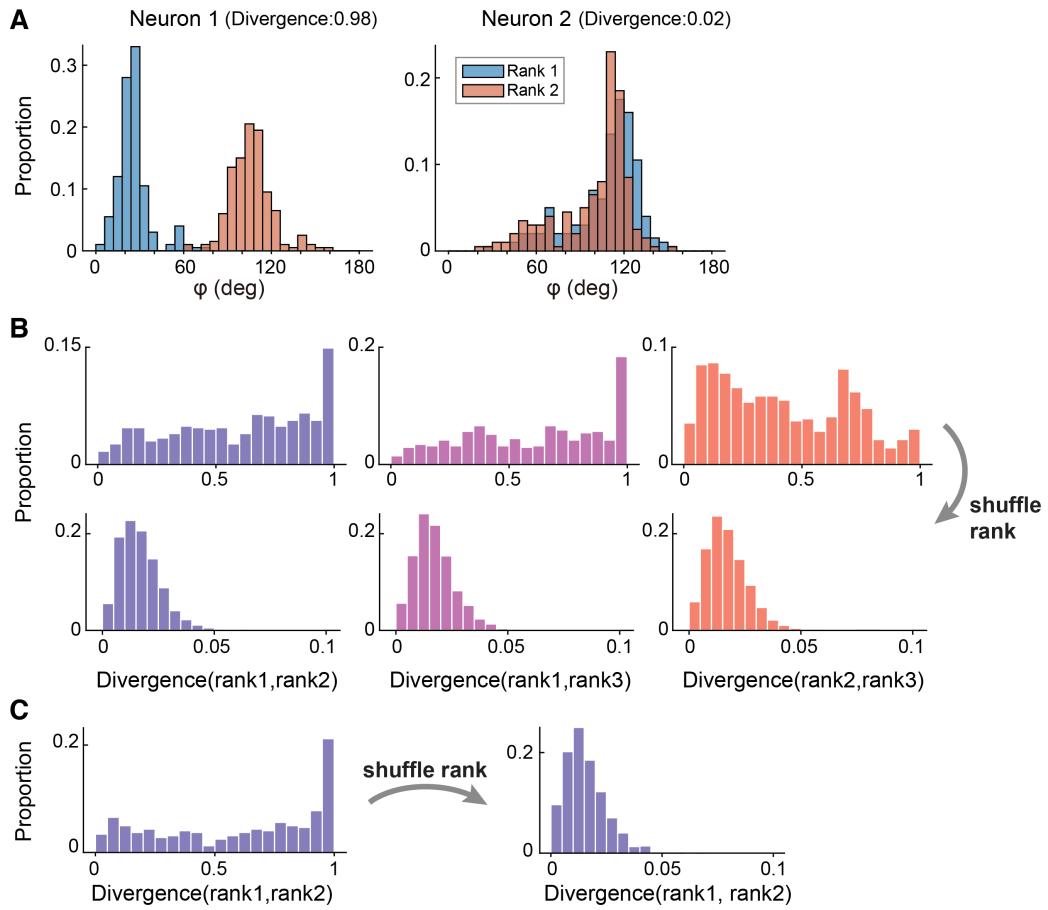


fig. S9. Angle estimation distribution comparison for different rank pairs.

(A) The distribution preferred angles for two example neurons in rank-1 and -2. We randomly split trials in half, performed regression analysis, found the rank subspaces, and computed the angle for each rank. After repeating this operation 100 times, we obtained the distribution of angle estimation (200 values) for each rank. For a given rank pair, Jensen-Shannon (JS) divergence was used to measure the distance between the two angle estimation distributions. A close-to-one divergence indicates the two distributions show little overlap (for example, neuron 1) while a close-to-zero divergence suggests the two distributions become almost identical (for example, neuron 2).

(B) Histograms of JS divergence between among neurons for each rank pairs in M1, estimated by original data (top) or rank-shuffled data (bottom). There were many neurons having a close-to-one JS divergence for the original data, indicating that they had different angles for different ranks even considering the dispersion within each rank. In contrast, the distribution of the JS divergence became concentrated around 0 when the rank information was shuffled, consistent with the fact that rank subspaces became almost identical in this control scenario.

(C) JS divergence results for M2. The same conclusion as M1 holds.

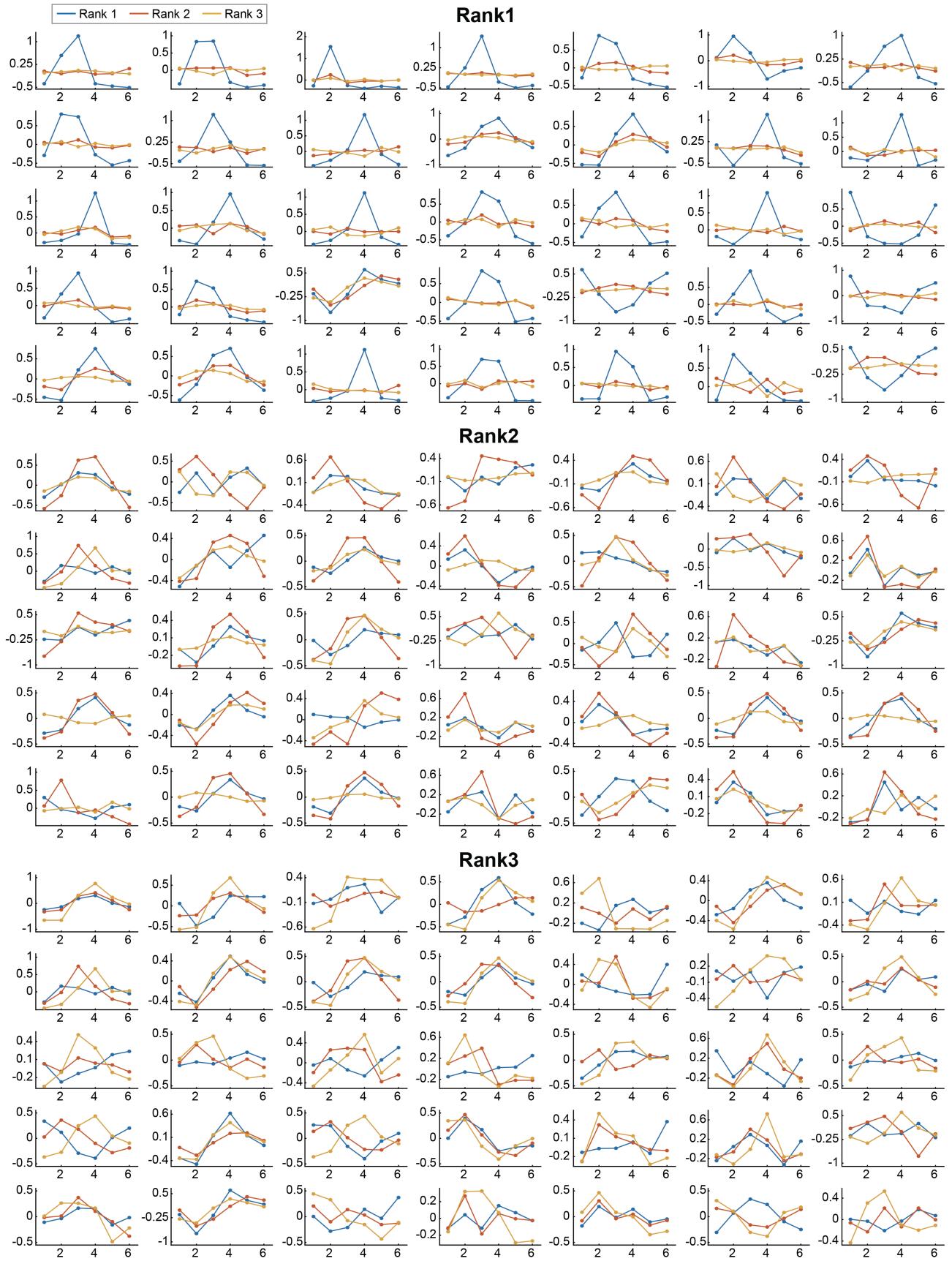


fig. S10. Single neuron tuning curves for M1.

Tuning curves of 35 neurons with the highest degree of alignment to rank-1 (**Top**), rank-2 (**Middle**) or rank-3 (**Bottom**) subspace.

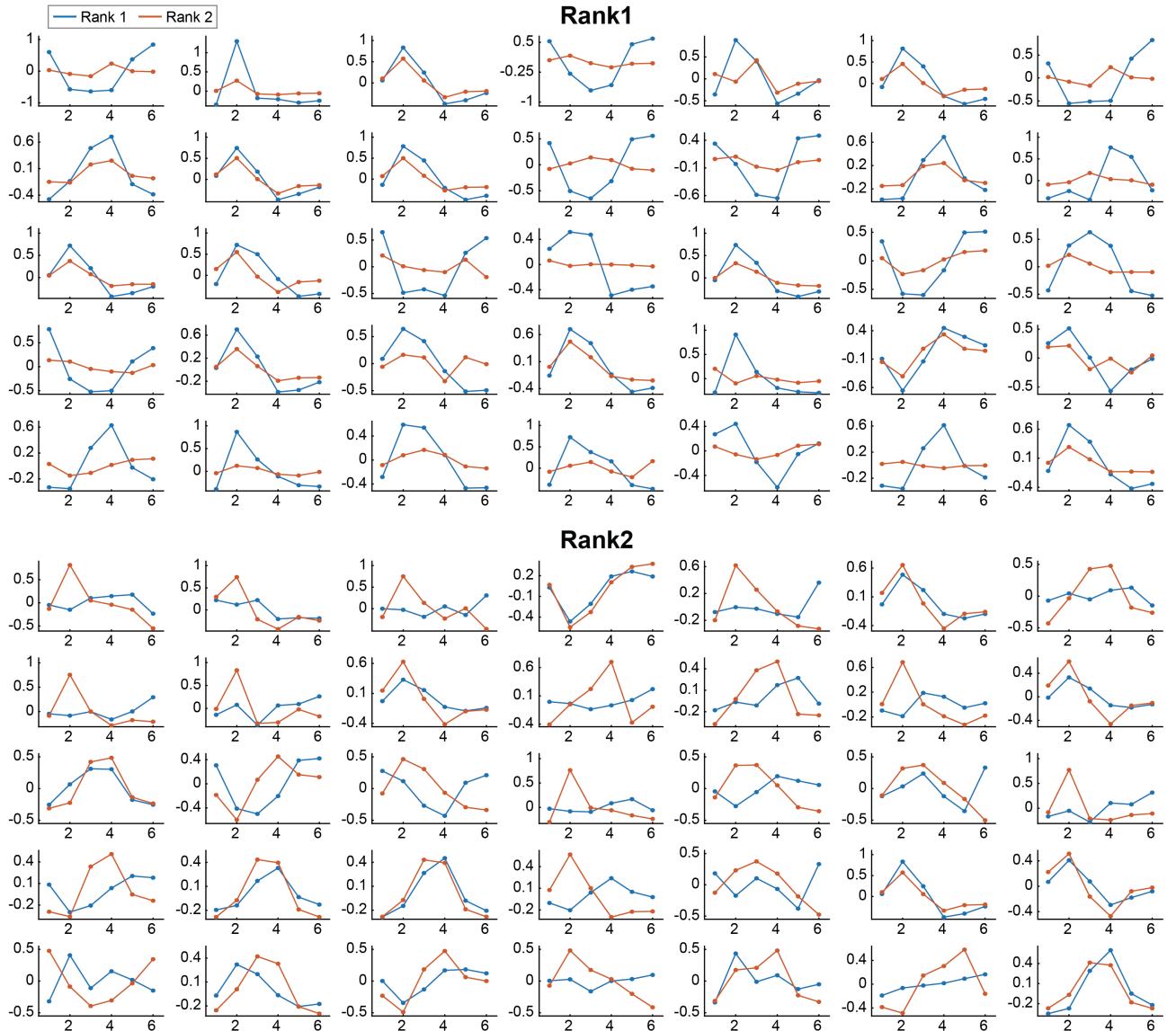
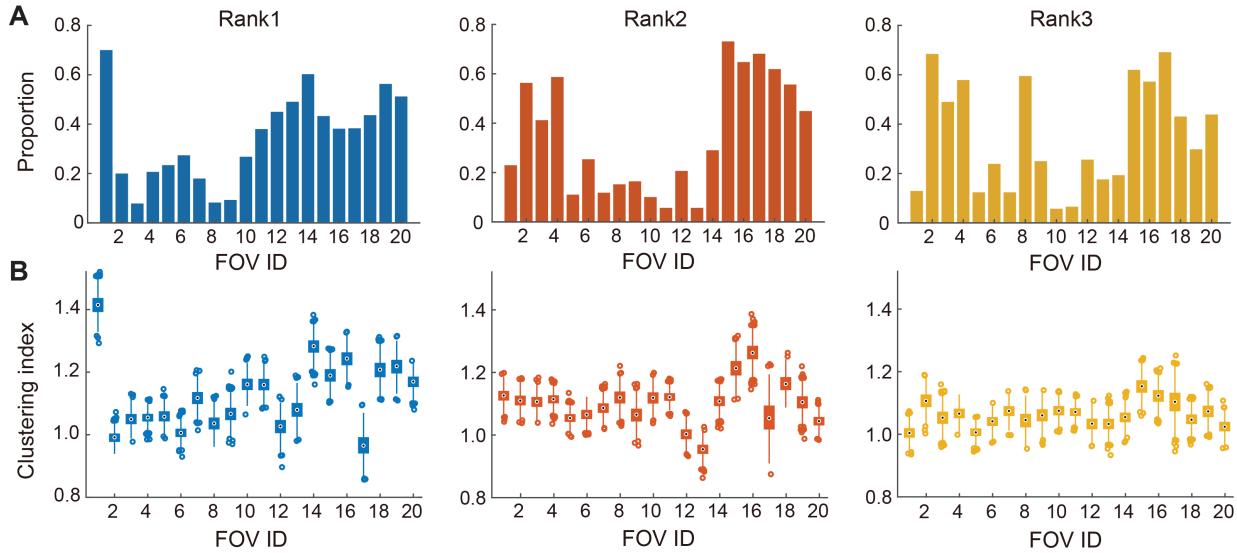


fig. S11. Single neuron tuning curves for M2.

Same as fig. S10, for M2 length-2 sequences.

Monkey 1



Monkey 2

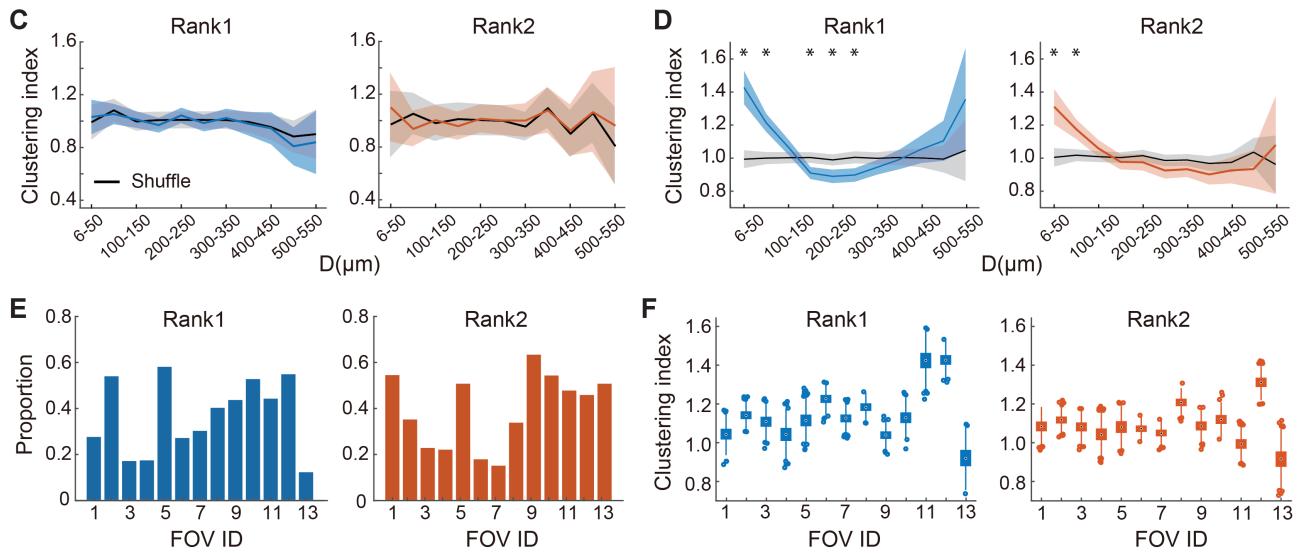


fig. S12. Anatomical organization of SWM code in different FOVs.

(A) The proportions of neurons significantly contributing to different rank subspaces in each FOV in M1. For rank-1 subspace, neural contribution ranking in the top 38% (determined by rank-1 normalized PR value) across the neural population is regarded as significant. For rank-2 and rank-3 subspaces, the thresholds (34% and 32%, respectively) are determined in a similar way.

(B) Functional clustering of spatial preferences along with different ranks for all FOVs in M1. The clustering indices for small cortical distance (6-50 μm) are compared across all FOVs. The results show that the clustering indices of rank-1 spatial tuning curve are the most diversified.

(C) No rank contribution clustering for an example FOV in M2. Same analysis as Fig. 5B.

(D) Spatial preference clustering for another example FOV in M2. Same analysis as Fig. 5D.

(E) The proportions of neurons significantly contributing to different rank subspaces in each FOV in M2. Same analysis as (A).

(F) Functional clustering of spatial preferences for all FOVs in M2. Same analysis as (B).

FOV	#neurons	Length-2 trials			Length-3 trials			
		# trials	Rank 1	Rank 2	# trials	Rank 1	Rank 2	Rank 3
1	170	94	0.784**	0.222	236	0.819**	0.279*	0.216
2	190	90	0.459**	0.528**	130	0.492**	0.360**	0.290
3	204	137	0.511**	0.473**	321	0.538**	0.335**	0.330**
4	242	101	0.431**	0.483**	207	0.528**	0.477**	0.451**
5	218	183	0.582**	0.235	339	0.604**	0.307**	0.259*
6	201	125	0.549**	0.279	170	0.555**	0.331**	0.261
7	178	120	0.420**	0.229	210	0.515**	0.255	0.251
8	158	100	0.251	0.155	151	0.274	0.231	0.230
9	140	87	0.254	0.175	125	0.256	0.167	0.170
10	209	176	0.474**	0.190	198	0.479**	0.232	0.196
11	229	185	0.520**	0.211	270	0.492**	0.259*	0.197
12	160	70	0.555**	0.200	158	0.575**	0.217	0.210
13	159	125	0.521**	0.224	211	0.616**	0.281	0.191
14	176	127	0.871**	0.287*	160	0.883**	0.360**	0.237
15	171	61	0.583**	0.443**	126	0.562**	0.425**	0.392**
16	173	93	0.538**	0.442**	188	0.621**	<u>0.519**</u>	0.455**
17	94	35	0.322	0.407	95	<u>0.344**</u>	—	0.339*
18	181	90	0.538**	0.534**	142	0.543**	<u>0.375**</u>	0.348**
19	151	53	0.745**	0.325	119	0.688**	0.237	0.313
20	205	56	0.831**	0.341	87	0.717**	—	<u>0.302</u>

Table S1. Macro F1-score of decoding from multiple recording sessions for M1.

Correct trials were used to train decoders with leave-one-trial-out protocol. Given the precision P_i and recall R_i for location i , the F1-score for location i is the harmonic mean of precision and recall: $f_i = 2P_iR_i/(P_i + R_i)$, and the macro F1-score is defined as $F = \sum_i f_i/6$. To determine the p value, we construct the shuffle distribution with the dummy decoders as before (see Methods) but restricted to the last four delay-period time windows. Scores listed in the table are averaged across the four decoders. For the p-value we reported, the macro F1-score of all four decoders were required to be at least the value of $1 - p$ percentile of the shuffle distribution. ***, $p < 0.01/0.001$.

Instead of decoding accuracy, we calculated the macro F1-score which is a better metric to quantify the decoding performance in imbalanced classification. This happens in some recording sessions where trial counts of different locations are not evenly distributed. In extreme case where one or more items were missing from a specific rank, the score cannot be calculated, which is indicated by ‘—’. Even if there was at least one trial for each item, the f_i for the item with only one trial will be unreliable, since it is based on the decoder trained with that trial left-out. In this situation, we mark the score with underline.

The result here would serve as a basis for evaluating whether a specific FOV encoded item information at population level. Only the FOVs showing significant decoding in at least one rank are considered for further generalization test in tables S3 to S5.

FOV	#neurons	Length-2 trials			Length-3 trials			
		# trials	Rank 1	Rank 2	# trials	Rank 1	Rank 2	Rank 3
1	123	94	0.501**	0.328	0	—	—	—
2	176	123	0.630**	0.448**	6	—	—	—
3	140	246	0.345**	0.329**	76	0.337	0.290	0.199
4	86	140	0.371**	0.255	48	—	0.253	0.223
5	136	189	0.637**	0.383**	45	0.453*	0.254	—
6	173	184	0.631**	0.420**	56	0.446**	<u>0.207</u>	0.326
7	152	313	0.680**	0.425**	102	0.609**	0.243	<u>0.325</u>
8	139	104	0.739**	0.365**	41	<u>0.556</u> **	0.301	—
9	142	71	0.571**	0.454**	49	—	<u>0.355</u>	<u>0.309</u>
10	127	119	0.571**	0.496**	90	—	0.416**	0.280
11	113	120	0.674**	0.454**	70	0.568**	0.283	0.294
12	144	157	0.630**	0.418**	154	0.634**	0.436**	0.340*
13	65	84	0.193	0.279	64	0.224	<u>0.274</u>	0.207

Table S2. Macro F1-score of decoding from multiple recording sessions for M2.

Legend, same as table S1.

M1, Length-3 trials					M2, Length-2 trials			
FOV	# trials	Rank 1	Rank 2	Rank 3	FOV	# trials	Rank1	Rank2
1	236	0.812**	0.242	0.185	1	94	0.444**	0.270
2	130	0.399**	0.259	0.207	2	123	0.565**	0.289*
3	321	0.447**	0.264**	0.255*	3	246	0.266**	0.248**
4	207	0.446**	0.365**	0.298**	4	140	0.323**	0.202
5	339	0.511**	0.250**	0.169	5	189	0.553**	0.252
6	170	0.505**	0.264*	0.179	6	184	0.566**	0.306**
7	210	0.418**	0.204	0.176	7	313	0.584**	0.310**
10	198	0.432**	0.181	0.149	8	104	0.700**	0.262
11	270	0.394**	0.215	0.128	9	71	0.469**	0.347
12	158	0.508**	0.168	0.147	10	119	0.476**	0.379**
13	211	0.584**	0.208	0.139	11	120	0.551**	0.333**
14	160	0.789**	0.252	0.124	12	157	0.459**	0.253
15	126	0.437**	0.328*	0.205				
16	188	0.556**	<u>0.362</u> **	0.257				
18	142	0.407**	<u>0.295</u> **	0.238				

Table S3. Macro F1-score of leave-one-sequence-out decoding for M1 and M2.

Length-3 correct trials of M1 and length-2 correct trials of M2 were used to train decoders with leave-one-sequence-out cross-validation. For M1, FOVs with more than 120 trials were used. In the situation where all trials for certain item came from one sequence, the score is marked with underline. Otherwise, the notations are the same as table S1.

M1 FOV	2 to 3		3 to 2		M2 FOV	2 to 3	
	Rank1	Rank2	Rank1	Rank2		Rank1	Rank2
1	0.777**	0.309**	0.860**	0.310	3	0.326*	0.321
2	0.474**	0.349**	0.419**	0.330*	4	0.477**	0.185
3	0.430**	0.359**	0.473**	0.395**	5	0.537**	0.370
4	0.475**	0.331**	0.427**	0.435**	6	0.664**	0.272
5	0.594**	0.243	0.590**	0.267	7	0.652**	0.277
6	0.546**	0.231	0.549**	0.302	8	0.755**	0.308
7	0.475**	0.184	0.495**	0.198	9	0.607**	0.304
10	0.494**	0.212	0.490**	0.217	10	0.565**	0.399**
11	0.531**	0.241	0.513**	0.264	11	0.642**	0.413**
12	0.572**	0.237	0.581**	0.260	12	0.455**	0.437**
13	0.649**	0.188	0.546**	0.214		3 to 2	
14	0.827**	0.308	0.806**	0.264	7	0.593**	0.248
15	0.485**	0.397**	0.488**	0.447**	10	0.538**	0.340**
16	0.572**	0.398**	0.595**	0.387**	11	0.538**	0.308
18	0.511**	0.451**	0.442**	0.377**	12	0.541**	0.275

Table S4. Cross-length decoding accuracy for M1 and M2.

Decoders trained with trials of the length-3 sequence were tested in trials of length-2 sequence (3 to 2), and vice versa (2 to 3). All the data used for training and testing were correct trials (please refer to tables S1 and S2 for trial counts). For M1, FOVs with more than 120 (60) trials in length-3 (-2) sequences were used. In 3 to 2 decoding for M2, FOVs with more than 60 trials were used. We report the decoding accuracy here since the item distribution in training and test set were independent. The statistical testing procedure was the same as in table S1.

FOV	# trials	Rank 1	# trials	Rank 2	# trials	Rank 3
M1-2	77	0.483**	39	0.232	7	0.175
M1-3	133	0.547**	78	0.414**	14	0.451
M1-4	89	0.562**	37	0.516	11	0.265
M1-14	112	0.869**	72	0.341	18	0.273
M1-15	119	0.564**	61	0.597**	10	0.458
M1-16	87	0.610**	48	0.517*	9	0.398
M1-18	129	0.514**	65	0.523**	21	0.310
M2-10	194	0.514**	105	0.361**	25	0.459
M2-12	283	0.645**	109	0.339**	64	0.553**

Table S5. Decoding location match in error trials for M1 and M2.

Decoders were trained with all the length-3 correct trials. Then, for a given rank, we identified error trials in which response was correct at that rank and applied the decoders on these trials. The decoding accuracies were reported for selected FOVs with enough error trial for testing while showing significant decoding for at least two ranks.

References and Notes

1. S. Dehaene, F. Meyniel, C. Wacongne, L. Wang, C. Pallier, The Neural Representation of Sequences: From Transition Probabilities to Algebraic Patterns and Linguistic Trees. *Neuron* **88**, 2–19 (2015). [doi:10.1016/j.neuron.2015.09.019](https://doi.org/10.1016/j.neuron.2015.09.019) [Medline](#)
2. M. Botvinick, T. Watanabe, From numerosity to ordinal rank: A gain-field model of serial order representation in cortical working memory. *J. Neurosci.* **27**, 8636–8642 (2007). [doi:10.1523/JNEUROSCI.2110-07.2007](https://doi.org/10.1523/JNEUROSCI.2110-07.2007) [Medline](#)
3. Y. Liu, R. J. Dolan, Z. Kurth-Nelson, T. E. J. Behrens, Human Replay Spontaneously Reorganizes Experience. *Cell* **178**, 640–652.e14 (2019). [doi:10.1016/j.cell.2019.06.012](https://doi.org/10.1016/j.cell.2019.06.012) [Medline](#)
4. Y. Bengio, A. Courville, P. Vincent, Representation Learning: A Review and New Perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 1789–1828 (2013). [doi:10.1109/TPAMI.2013.50](https://doi.org/10.1109/TPAMI.2013.50)
5. G. Buzsáki, A. Draguhn, Neuronal oscillations in cortical networks. *Science* **304**, 1926–1929 (2004). [doi:10.1126/science.1099745](https://doi.org/10.1126/science.1099745) [Medline](#)
6. C. Baldassano, J. Chen, A. Zadbood, J. W. Pillow, U. Hasson, K. A. Norman, Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron* **95**, 709–721.e5 (2017). [doi:10.1016/j.neuron.2017.06.041](https://doi.org/10.1016/j.neuron.2017.06.041) [Medline](#)
7. R. A. Andersen, G. K. Essick, R. M. Siegel, Encoding of spatial location by posterior parietal neurons. *Science* **230**, 456–458 (1985). [doi:10.1126/science.4048942](https://doi.org/10.1126/science.4048942) [Medline](#)
8. A. Pouget, T. J. Sejnowski, Spatial transformations in the parietal cortex using basis functions. *J. Cogn. Neurosci.* **9**, 222–237 (1997). [doi:10.1162/jocn.1997.9.2.222](https://doi.org/10.1162/jocn.1997.9.2.222) [Medline](#)
9. P. Smolensky, Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artif. Intell.* **46**, 159–216 (1990). [doi:10.1016/0004-3702\(90\)90007-M](https://doi.org/10.1016/0004-3702(90)90007-M)
10. V. Mante, D. Sussillo, K. V. Shenoy, W. T. Newsome, Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013). [doi:10.1038/nature12742](https://doi.org/10.1038/nature12742) [Medline](#)
11. M. Rigotti, O. Barak, M. R. Warden, X.-J. Wang, N. D. Daw, E. K. Miller, S. Fusi, The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013). [doi:10.1038/nature12160](https://doi.org/10.1038/nature12160) [Medline](#)
12. M. M. Churchland, J. P. Cunningham, M. T. Kaufman, J. D. Foster, P. Nuyujukian, S. I. Ryu, K. V. Shenoy, Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012). [doi:10.1038/nature11129](https://doi.org/10.1038/nature11129) [Medline](#)

13. S. Bernardi, M. K. Benna, M. Rigotti, J. Munuera, S. Fusi, C. D. Salzman, The Geometry of Abstraction in the Hippocampus and Prefrontal Cortex. *Cell* **183**, 954–967.e21 (2020). [doi:10.1016/j.cell.2020.09.031](https://doi.org/10.1016/j.cell.2020.09.031) [Medline](#)
14. E. H. Nieh, M. Schottdorf, N. W. Freeman, R. J. Low, S. Lewallen, S. A. Koay, L. Pinto, J. L. Gauthier, C. D. Brody, D. W. Tank, Geometry of abstract learned knowledge in the hippocampus. *Nature* **595**, 80–84 (2021). [doi:10.1038/s41586-021-03652-7](https://doi.org/10.1038/s41586-021-03652-7) [Medline](#)
15. M. F. Panichello, T. J. Buschman, Shared mechanisms underlie the control of working memory and attention. *Nature* **592**, 601–605 (2021). [doi:10.1038/s41586-021-03390-w](https://doi.org/10.1038/s41586-021-03390-w) [Medline](#)
16. J. Wang, D. Narain, E. A. Hosseini, M. Jazayeri, Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* **21**, 102–110 (2018). [doi:10.1038/s41593-017-0028-6](https://doi.org/10.1038/s41593-017-0028-6) [Medline](#)
17. Materials and methods are available as supplementary materials online.
18. A. Nieder, I. Diester, O. Tudusciuc, Temporal and spatial enumeration processes in the primate parietal cortex. *Science* **313**, 1431–1435 (2006). [doi:10.1126/science.1130308](https://doi.org/10.1126/science.1130308) [Medline](#)
19. J. R. Duhamel, F. Bremmer, S. Ben Hamed, W. Graf, Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* **389**, 845–848 (1997). [doi:10.1038/39865](https://doi.org/10.1038/39865) [Medline](#)
20. R. Q. Quiroga, L. Reddy, G. Kreiman, C. Koch, I. Fried, Invariant visual representation by single neurons in the human brain. *Nature* **435**, 1102–1107 (2005). [doi:10.1038/nature03687](https://doi.org/10.1038/nature03687) [Medline](#)
21. J. H. Reynolds, D. J. Heeger, The normalization model of attention. *Neuron* **61**, 168–185 (2009). [doi:10.1016/j.neuron.2009.01.002](https://doi.org/10.1016/j.neuron.2009.01.002) [Medline](#)
22. X. Jiang, T. Long, W. Cao, J. Li, S. Dehaene, L. Wang, Production of Supra-regular Spatial Sequences by Macaque Monkeys. *Curr. Biol.* **28**, 1851–1859.e4 (2018). [doi:10.1016/j.cub.2018.04.047](https://doi.org/10.1016/j.cub.2018.04.047) [Medline](#)
23. K. Oberauer, S. Lewandowsky, E. Awh, G. D. A. Brown, A. Conway, N. Cowan, C. Donkin, S. Farrell, G. J. Hitch, M. J. Hurlstone, W. J. Ma, C. C. Morey, D. E. Nee, J. Schweppe, E. Vergauwe, G. Ward, Benchmarks for models of short-term and working memory. *Psychol. Bull.* **144**, 885–958 (2018). [doi:10.1037/bul0000153](https://doi.org/10.1037/bul0000153) [Medline](#)
24. M. J. Hurlstone, G. J. Hitch, A. D. Baddeley, Memory for serial order across domains: An overview of the literature and directions for future research. *Psychol. Bull.* **140**, 339–373 (2014). [doi:10.1037/a0034221](https://doi.org/10.1037/a0034221) [Medline](#)
25. J. E. Lisman, O. Jensen, The theta-gamma neural code. *Neuron* **77**, 1002–1016 (2013). [doi:10.1016/j.neuron.2013.03.007](https://doi.org/10.1016/j.neuron.2013.03.007) [Medline](#)

26. X. J. Wang, Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol. Rev.* **90**, 1195–1268 (2010). [doi:10.1152/physrev.00035.2008](https://doi.org/10.1152/physrev.00035.2008) [Medline](#)
27. M. Abeles, G. Hayon, D. Lehmann, Modeling compositionality by dynamic binding of synfire chains. *J. Comput. Neurosci.* **17**, 179–201 (2004). [doi:10.1023/B:JCNS.0000037682.18051.5f](https://doi.org/10.1023/B:JCNS.0000037682.18051.5f) [Medline](#)
28. M. Siegel, M. R. Warden, E. K. Miller, Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21341–21346 (2009). [doi:10.1073/pnas.0908193106](https://doi.org/10.1073/pnas.0908193106) [Medline](#)
29. M. A. Wilson, B. L. McNaughton, Reactivation of hippocampal ensemble memories during sleep. *Science* **265**, 676–679 (1994). [doi:10.1126/science.8036517](https://doi.org/10.1126/science.8036517) [Medline](#)
30. K. Diba, G. Buzsáki, Forward and reverse hippocampal place-cell sequences during ripples. *Nat. Neurosci.* **10**, 1241–1242 (2007). [doi:10.1038/nn1961](https://doi.org/10.1038/nn1961) [Medline](#)
31. S. Funahashi, C. J. Bruce, P. S. Goldman-Rakic, Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J. Neurophysiol.* **61**, 331–349 (1989). [doi:10.1152/jn.1989.61.2.331](https://doi.org/10.1152/jn.1989.61.2.331) [Medline](#)
32. A. Compte, N. Brunel, P. S. Goldman-Rakic, X. J. Wang, Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* **10**, 910–923 (2000). [doi:10.1093/cercor/10.9.910](https://doi.org/10.1093/cercor/10.9.910) [Medline](#)
33. P. Barone, J. P. Joseph, Prefrontal cortex and spatial sequencing in macaque monkey. *Exp. Brain Res.* **78**, 447–464 (1989). [doi:10.1007/BF00230234](https://doi.org/10.1007/BF00230234) [Medline](#)
34. S. Funahashi, M. Inoue, K. Kubota, Delay-period activity in the primate prefrontal cortex encoding multiple spatial positions and their order of presentation. *Behav. Brain Res.* **84**, 203–223 (1997). [doi:10.1016/S0166-4328\(96\)00151-9](https://doi.org/10.1016/S0166-4328(96)00151-9) [Medline](#)
35. Y. Ninokura, H. Mushiake, J. Tanji, Representation of the temporal order of visual objects in the primate lateral prefrontal cortex. *J. Neurophysiol.* **89**, 2868–2873 (2003). [doi:10.1152/jn.00647.2002](https://doi.org/10.1152/jn.00647.2002) [Medline](#)
36. K. S. Lashley, *The Problem of Serial Order in Behavior* (Bobbs-Merrill, 1951).
37. Y. Xie, P. Hu, J. Li, J. Chen, W. Song, X.-J. Wang, T. Yang, S. Dehaene, S. Tang, B. Min, L. Wang, Data repository for ‘Geometry of Sequence Working Memory in Macaque Prefrontal Cortex’, Zenodo (2021); <https://doi.org/10.5281/zenodo.5739376>.
38. P. Hu, Y. Xie, Code for ‘Geometry of Sequence Working Memory in Macaque Prefrontal Cortex,’ Zenodo (2021); <https://doi.org/10.5281/zenodo.5746184>.
39. T. W. Chen, T. J. Wardill, Y. Sun, S. R. Pulver, S. L. Renninger, A. Baohan, E. R. Schreiter, R. A. Kerr, M. B. Orger, V. Jayaraman, L. L. Looger, K. Svoboda, D. S. Kim,

Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013). [doi:10.1038/nature12354](https://doi.org/10.1038/nature12354) [Medline](#)

40. M. Li, F. Liu, H. Jiang, T. S. Lee, S. Tang, Long-Term Two-Photon Imaging in Awake Macaque Monkey. *Neuron* **93**, 1049–1057.e3 (2017). [doi:10.1016/j.neuron.2017.01.027](https://doi.org/10.1016/j.neuron.2017.01.027) [Medline](#)
41. J. Hwang, A. R. Mitz, E. A. Murray, NIMH MonkeyLogic: Behavioral control and data acquisition in MATLAB. *J. Neurosci. Methods* **323**, 13–21 (2019). [doi:10.1016/j.jneumeth.2019.05.002](https://doi.org/10.1016/j.jneumeth.2019.05.002) [Medline](#)
42. A. Giovannucci, J. Friedrich, P. Gunn, J. Kalfon, B. L. Brown, S. A. Koay, J. Taxidis, F. Najafi, J. L. Gauthier, P. Zhou, B. S. Khakh, D. W. Tank, D. B. Chklovskii, E. A. Pnevmatikakis, CaImAn an open source tool for scalable calcium imaging data analysis. *eLife* **8**, e38173 (2019). [doi:10.7554/eLife.38173](https://doi.org/10.7554/eLife.38173) [Medline](#)
43. J. A. Gallego, M. G. Perich, S. N. Naufel, C. Ethier, S. A. Solla, L. E. Miller, Cortical population activity within a preserved neural manifold underlies multiple motor behaviors. *Nat. Commun.* **9**, 4233 (2018). [doi:10.1038/s41467-018-06560-z](https://doi.org/10.1038/s41467-018-06560-z) [Medline](#)
44. M. X. Cohen, *Analyzing Neural Time Series Data: Theory and Practice* (MIT Press, 2014).
45. P. Gao, E. Trautmann, B. Yu, G. Santhanam, S. Ryu, K. Shenoy, S. Ganguli, A theory of multineuronal dimensionality, dynamics and measurement. bioRxiv 214262 [Preprint] (2017). <https://doi.org/10.1101/214262>.
46. I. Nauhaus, K. J. Nielsen, A. A. Disney, E. M. Callaway, Orthogonal micro-organization of orientation and spatial frequency in primate primary visual cortex. *Nat. Neurosci.* **15**, 1683–1690 (2012). [doi:10.1038/nn.3255](https://doi.org/10.1038/nn.3255) [Medline](#)