



Published in final edited form as:

Nature. 2021 July ; 595(7865): 80–84. doi:10.1038/s41586-021-03652-7.

## Geometry of abstract learned knowledge in the hippocampus

Edward H. Nieh<sup>1,\*</sup>, Manuel Schottdorf<sup>1,\*</sup>, Nicolas W. Freeman<sup>1</sup>, Ryan J. Low<sup>1</sup>, Sam Lewallen<sup>1</sup>, Sue Ann Koay<sup>1</sup>, Lucas Pinto<sup>1,5</sup>, Jeffrey L. Gauthier<sup>1</sup>, Carlos D. Brody<sup>1,2,3,#</sup>, David W. Tank<sup>1,2,4,#</sup>

<sup>1</sup>Princeton Neuroscience Institute, Princeton University, Princeton, NJ, 08544, USA

<sup>2</sup>Department of Molecular Biology, Princeton University, Princeton, NJ, 08544, USA

<sup>3</sup>Howard Hughes Medical Institute, Princeton University, Princeton, NJ, 08544, USA

<sup>4</sup>Bezos Center for Neural Dynamics, Princeton University, Princeton, NJ, 08544, USA.

<sup>5</sup>Present address: Department of Physiology, Feinberg School of Medicine, Northwestern University, Chicago, IL, 60611, USA.

### Abstract

Hippocampal neurons encode physical variables<sup>1–7</sup> such as space<sup>1</sup> or auditory frequency<sup>6</sup> in cognitive maps<sup>8</sup>. In addition, human fMRI studies have shown that the hippocampus can also encode more abstract, learned variables<sup>9–11</sup>. However, their integration into existing neural representations of physical variables<sup>12,13</sup> is unknown. Using 2-photon calcium imaging, we show that individual dorsal CA1 neurons jointly encode accumulated evidence with spatial position in mice performing a decision-making task in virtual reality<sup>14–16</sup>. Nonlinear dimensionality reduction<sup>13</sup> showed that population activity was well described by ~4–6 latent variables, suggesting that neural activity is constrained to a low-dimensional manifold. Within this low-dimensional space, both physical and abstract variables were jointly mapped in an orderly fashion, creating a geometric representation that we demonstrate to be similar across animals. The existence of conjoined cognitive maps suggests that the hippocampus performs a general computation – to create geometric representations of learned knowledge instantiated by task-specific low-dimensional manifolds.

---

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence:** David W. Tank, [dwtank@princeton.edu](mailto:dwtank@princeton.edu), Carlos D. Brody, [brody@princeton.edu](mailto:brody@princeton.edu).

\*these authors contributed equally to this work

#these senior authors contributed equally to this work

**Author Contributions.** E.H.N., C.D.B., and D.W.T. designed the experiments. E.H.N. performed the experiments. E.H.N., M.S., and N.W.F. analyzed the data. S.A.K. wrote code for processing the calcium imaging data. L.P. provided code for behavioral analysis. J.L.G. provided code for single cell analysis. M.S., R.J.L., and S.L. developed the MIND algorithm. E.H.N., M.S., C.D.B., and D.W.T. wrote the paper. The authors declare no competing interests.

**Supplementary Information** is available for this paper.

Code availability

The code used for all analyses in this study is available on Github ([https://github.com/BrainCOGS/HPC\\_manifolds](https://github.com/BrainCOGS/HPC_manifolds)). All other code is available upon reasonable request.

## Introduction

Since the discovery of place cells in the CA1 of dorsal hippocampus that increased their firing rates when rats moved through specific locations in a given environment<sup>1</sup>, hippocampal neurons have also been shown to encode time<sup>17,18</sup>, auditory frequency<sup>6</sup>, odors<sup>4,7</sup>, and taste<sup>5</sup>. Together, these studies support the view that the hippocampus constructs task-dependent cognitive maps<sup>8,19</sup>, where hippocampal neurons not only encode spatial position, but whichever environmental variable is relevant to the task at hand. Furthermore, fMRI studies in humans have shown that the hippocampus can encode more cognitive variables, such as the sequential nature of a non-spatial task<sup>9</sup> or social structures<sup>10,11</sup>. Cognitive variables can be characterized by geometric properties such as adjacency and distance<sup>20–22</sup>, suggesting the neural encoding of these variables at the cellular level may also be manifest with geometric structure.

Neural activity can be described as a point in a high-dimensional coordinate system, where each coordinate axis represents a single neuron's activity. Underlying properties of the network and its inputs can confine neural trajectories to a subregion of this space, i.e. the neural manifold, which has been proposed to underlie motor movements<sup>23,24</sup>, head direction cells<sup>25</sup>, and hippocampal maps of physical variables<sup>13</sup>. The conceptual ideas in these studies suggest a general principle of hippocampal computation: the construction of organized maps of learned knowledge<sup>26,27</sup> instantiated by neural manifolds. Here, we examine how neurons in the dorsal CA1 integrate neural representations of cognitive and physical variables and whether low-dimensional manifolds underlie these representations.

### Evidence accumulation in virtual reality

We used transgenic GCaMP6f-expressing mice (n=7) performing an evidence accumulation task in virtual reality (VR; Supplementary Video 1)<sup>14,28–30</sup> and 2-photon calcium imaging to record cellular resolution activity of neurons in dorsal CA1 (n=3144 total neurons, 449±64 SEM simultaneously per session; Fig. 1a). The “accumulating towers task”<sup>14</sup> combines navigation with decision-making, such that position, a physical variable, has to be integrated with accumulated evidence<sup>14–16,31,32</sup> - a cognitive variable that is not innate and can only be inferred and calculated after learning the task rules. Mice learned to traverse the stem of an immersive VR T-maze, while visual cues were presented randomly on the left and right walls. Turning to the side with more cues at the end of the maze resulted in the delivery of a liquid reward, while turning to the opposite side resulted in a time-out. Consistent with previously published results<sup>14</sup>, the behavior showed characteristic psychometric curves (Fig. 1b), and mice used evidence (integrated # right towers - # left towers) from throughout the cue period (Fig. 1c).

Figure 1d and 1e illustrate two possibilities for how CA1 neurons may behave in the task. If they behave like previously described place cells that respond differently depending on context<sup>2,3,33,34</sup>, e.g. “splitter cells” that encode turn direction, we would expect reliable place cell sequences specific to right or left turn trials (Fig. 1d). However, if individual CA1 neurons can encode a cognitive variable, such as the amount of accumulated evidence, in addition to position in the maze, the cognitive map might comprise at least two independent axes - a position axis and an accumulated evidence axis<sup>35</sup>. If so, we would expect each right

choice trial to evoke different neural sequences, depending on the time courses of evidence that the animals encountered throughout the maze (Fig. 1e). Importantly, in this second scenario, firing fields evaluated in a single dimension, such as position, would exist, but would appear unreliable across trials with different amounts of accumulated evidence (Fig. 1e bottom). Note that unreliability could appear as either missing activity in the cell's place field or variability in the position at which the cell is active.

### Joint encoding of position and evidence

To distinguish these two possibilities, we examined how neural activity depended upon known behavioral variables such as position, choice, and evidence. We first calculated F/F for each identified hippocampal CA1 neuron following established methods<sup>15,36,37</sup>. We then measured mutual information between each cell's neural activity and the animal's position along the stem of the T-maze (0 to 300 cm) and compared it to a shuffled dataset where each cell's activity within a trial was circularly shifted. CA1 neurons exhibited choice-specific place cell sequences when activity was sorted by the position of peak activity (Fig. 2a). However, the response of individual cells in these sequences was more variable and unreliable on a trial-by-trial basis in comparison to a simpler alternation task (Extended Data Fig. 1a, b). This is against the prediction of choice-specific cell maps (Fig. 1d), but consistent with maps where evidence and position are jointly encoded (Fig. 1e). We next measured the mutual information between accumulated evidence and each cell's neural activity and found that CA1 neurons formed firing fields in evidence space that spanned small segments of evidence values (Fig. 2b, Extended Data Fig. 1c), consistent with Figure 1e.

To directly test the hypothesis that CA1 neurons encode evidence and position jointly (Fig. 1e), we measured the amount of mutual information between neural activity and occupancy in a 2D evidence-by-position (E×Y) space and compared this to the amount of mutual information if cells encoded position or evidence independently. The neural activity of an example neuron with significant mutual information between activity and occupancy in E×Y space is shown in Fig. 2c, and 25 of these neurons from a single imaging session are shown in Fig. 2d and Extended Data Fig. 2a. For these neurons jointly encoding position and evidence, mutual information in E×Y space was significantly greater than in 2D spaces in which either evidence or position values were shuffled (Fig. 2e, Extended Data Fig. 2b, c).

### Geometric representation by a neural manifold

While the mutual information metric has been historically used to measure spatial information in single hippocampal neurons<sup>38</sup>, it relies on the manual selection of predetermined behavioral variables. We therefore turned to the unsupervised extraction of neural manifolds using a principled method: manifold inference from neural dynamics (MIND)<sup>13</sup>. While most nonlinear dimensionality reduction techniques focus on the geometric properties of the cloud of population state data, MIND constructs a set of latent variables with a specific emphasis on incorporating temporal dynamics. It is therefore particularly suited to find low-dimensional representations in data with sequential activity.

We first used the distance metric in MIND to estimate the dimensionality of the neural manifold in the hippocampus during the accumulating towers task. We calculated distances from estimated transition probabilities between observed population activity states and counted the cumulative number of population states that fell within spheres of growing radii  $r$ , where  $r$  is an estimate for the inner distance<sup>12,39</sup>. If the manifold has  $d$  dimensions, we expect the number of states to grow as  $r^d$ . We found that the number of states grows approximately as  $d=5.4$  [4.8, 6.0], strongly indicating a low, approximately 4- to 6-dimensional latent geometry (Fig. 3a). Importantly, the dimensionality estimate of a simpler task, where visual cues only appeared on one side of the maze, was significantly lower (Extended Data Fig. 3).

To validate this estimate, we next embedded the manifold into  $d$ -dimensional Euclidean spaces and assessed how well these embedded manifolds described neural data using cross-validation on held-out trials (Extended Data Fig. 4). Figure 3b shows a small portion of the activity from 40 neurons (left) and the reconstruction of that same data from the five latent variables obtained after embedding the manifold into a 5-dimensional Euclidean space (right). We measured the average cross-validated correlation coefficient between the neural data and the reconstruction of the same data from manifolds embedded into 2- to 7-dimensional Euclidean spaces. Consistent with the dimensionality estimate in Fig. 3a, we find that the reconstruction performance saturates at ~5–6 dimensions (Fig. 3c). Using a linear dimensionality reduction technique, principal component analysis (PCA), comparable decoding indices for embedding into 4, 5, and 6 dimensions are reached using 29, 40, and 47 principal components, respectively. This reveals that hippocampal activity is constrained to an intricately shaped low-dimensional manifold that can only be identified with nonlinear dimensionality reduction techniques.

If the neural manifold accurately represents the cognitive map of the task that individual neurons encode, two key predictions should hold true. First, individual neurons should have firing fields that tile the latent space, and second, important variables in the task, such as position and evidence, should be organized in an orderly fashion. The activity of a representative neuron plotted as a heatmap on a 3-dimensional embedding of the manifold is shown in Fig. 3d, demonstrating a localized firing field on the manifold. Plotting the activity of multiple neurons on the same manifold reveals that the manifold is tiled with multiple firing fields (Fig. 3e, Supplementary Video 2). Furthermore, the manifold structure implies the coordinated activity of the entire neural population, such that activity of a single neuron can be well-predicted by activity from the rest of the population (Extended Data Fig. 4).

The second key prediction of our hypothesis is the orderly organization of important task variables on the manifold. Figure 3f reveals that both position (left) and evidence (right) appear organized as gradients in the latent space, in that the neural state trajectory typically progresses along a position direction in the course of a trial, while splitting along an independent, but integrated, evidence direction (Supplementary Video 3) - a structure fundamentally different from the visual inputs that the mouse experiences in the towers task (Extended Data Fig. 5). We then used gaussian process regression to decode position and evidence from the manifold and found that both variables can be decoded with similar accuracy as from neural data (Fig. 3g). In addition, other behavioral variables such as

velocity and view angle could also be decoded from the manifold, as well as binary task variables such as the choice on the previous trial and whether the previous trial was correct (Extended Data Fig. 6; Supplementary discussion).

If these geometric objects are task-specific, rather than animal-specific, there should be a high degree of similarity across animals performing the same task<sup>12</sup>. To test this hypothesis, we trained a model to predict position and evidence from manifold coordinates in one mouse and used the model to decode these variables in another mouse, after aligning their manifolds in the 5-dimensional embedding space (Fig. 3h, i). We found that the majority of the geometric structure was shared across brains (Fig. 3j).

### Sequential neural activity encodes behavior

If the manifold is a good representation of hippocampal neural activity, then each trial in the accumulating towers task has a corresponding trajectory within the manifold, leading to the emergence of trial-specific sequences of active cells. To detect sequences, we identified pairs of cells that consistently fired one after the other without any restrictions on the time/place in the maze that each cell fired (Extended Data Fig. 7a) and termed a pair of cells as a “doublet” if one cell fired after the other significantly more often than in a shuffled dataset, where each neuron’s activity was circularly-shifted within each trial (Fig. 4a, Extended Data Fig. 7b). To test whether these doublets appear more often than could be expected from independently behaving choice and position selective cells, we shuffled the trial IDs of each cell independently within left and right choice trials to remove pairwise correlations while preserving the place/side structure seen in Fig. 2a (Extended Data Fig. 8). The number of trials in which doublets appeared was significantly greater than in the shuffled dataset (Fig. 4b, Extended Data Fig. 9a). Furthermore, given the mostly unidirectional trajectories of the task in conceptual E×Y space (Extended Data Fig. 8a, b), we found that doublets were asymmetric (Fig. 4c, Extended Data Fig. 9b).

Next, we used the latent dimensions from the 5-dimensional embedding of the manifold to reconstruct the neural activity of all cells and extracted doublets from this reconstructed data. Even though doublets are very rare (on average, a given doublet is only active in  $3.6 \pm 0.01\%$  SEM of trials;  $n=16088$  doublets), the manifold predicted the presence of doublets with a  $0.87 \pm 0.02$  SEM true positive rate and  $0.14 \pm 0.01$  SEM false positive rate ( $n=7$  animals; Extended Data Fig. 9d). Furthermore, we found that the manifold could predict the precise timing of doublet events as well - the correlation between the timing of a doublet and the distance traversed on the manifold was significantly greater than the correlation in a shuffled dataset where manifold path lengths were taken from different trials with the same time interval (Fig. 4d, e; two-tailed Wilcoxon signed rank test,  $n=7$  animals,  $*p=0.031$ ).

Since the manifold encodes sequential activity well and given that behavioral variables are geometrically represented on the manifold (Fig. 3f, g), we would expect sequences to encode information about animal behavior, i.e. the animal’s upcoming choice. First, we identified doublets that were significantly choice-predictive by comparing the probability that the animal turns left or right in trials in which a doublet appears to the same probability in a shuffled dataset where choices in each trial were shuffled. Next, we found that these

choice-predictive doublets were significantly more predictive than the same doublets drawn from the shuffled dataset in which trial IDs were shuffled (Fig. 4f, Extended Data Fig. 9c). Taken together, these sequences are informative beyond independently behaving cells, suggesting population activity that is consistent with movement along the low-dimensional manifold.

## Discussion

By combining large-scale calcium imaging with a behavioral task where animals accumulate abstract evidence during navigation, we show how the coordinated activity of neurons in the dorsal CA1 region of the hippocampus gives rise to a task-specific geometric representation of a cognitive process. The neural population manifests this geometric representation by having firing fields within a low-dimensional nonlinear manifold, along which key task variables, both continuous and discrete, have an orderly arrangement. Previous rodent studies have shown the existence of low-dimensional manifolds in the hippocampus representing spatial position<sup>12,13</sup>, and fMRI studies in humans have shown that more abstract variables, such as social structures<sup>10,11</sup>, can be decoded from the hippocampus. One possibility was that different sets of hippocampal neurons could have encoded these variables separately, similar to the specialized coding of sensory, motor, and cognitive variables by VTA dopamine neurons in the same task<sup>16</sup>. However, we found that the majority of task-responsive neurons encoded position and evidence jointly (Fig. 2), leading to population dynamics that also reflect this joint neural code (Fig. 3 and 4).

The formation of a conjoined geometric representation of physical and abstract task variables, within neural manifolds in the hippocampus, could serve as a common organizing principle across two roles of the hippocampus, i.e. storing declarative memory and generating spatial/cognitive maps, that have historically been studied separately<sup>21,26,40</sup>. Low-dimensional manifolds could serve as the substrate on which relational networks for both declarative and spatial memories are stored<sup>27</sup>. In addition, our work suggests that the fast replay sequences seen in human nonspatial tasks<sup>9</sup> could be organized by the geometric structure of the neural manifold, analogous to the process by which neural sequences during ongoing behavior are evoked from trajectories through the manifold (Fig. 4). Finally, recent computational work has focused on how representations of knowledge in a reinforcement learning<sup>40</sup> or predictive coding<sup>27</sup> context can be used to guide behavior. There are intriguing parallels between the latent structure identified in these models and the latent variable structure we have uncovered in our studies. However, future work is required to provide a quantitative understanding of how our experimental results relate to these learning models.

## Methods

### Animals and stereotaxic surgery.

All procedures performed in this study were approved by the Institutional Animal Care and Use Committee at Princeton University and were performed in accordance with the Guide for the Care and Use of Laboratory Animals (National Research Council, 2011). Male and female mice aged 2 – 18 months expressing GCaMP6f were used for chronic expression of the calcium indicator.

- n=5 triple transgenic crosses expressing GCaMP6f under the CaMK11 $\alpha$  promoter from Ai93-D;CaMKII $\alpha$ -tTA [Igs7<sup>tm93.1</sup>(tetO-GCaMP6f)Hze Tg(Camk2a-tTA)1Mmay/J, Jackson Laboratories, stock# 024108] and Emx1-IRES-Cre [B6.129S2-Emx1<sup>tm1</sup>(cre)Krl/J, Jackson Laboratories, stock# 005628], also referred to as Ai93xEMX1
- n=10 Thy1-GCaMP6f [C57BL/6J-Tg(Thy1-GCaMP6f)GP5.3Dkim/J, Jackson Laboratories, stock# 028280], also referred to as GP5.3

Behaviorally, no differences have been observed in Ai93xEMX1 and GP5.3 animals<sup>14</sup>. In terms of calcium imaging, Ai93xEMX1 animals have higher expression levels of GCaMP6f than GP5.3 animals and therefore higher signal-to-noise ratios (SNR), resulting in different activity thresholds used to identify calcium events (described below). Some mice were used in multiple behavioral experiments that we analyzed, i.e. the one-side cues task is a training stage in the shaping procedure for the accumulating towers task. For all analyses and statistics other than those in Figure 1b, c (described below), one imaging session for each animal was selected based on the animal's performance in the task, the number of cells identified by the automated cell-finding algorithm, the amount of noise in the F/F signal, and the quality of motion correction.

Mice underwent surgical procedures as previously described<sup>29,41</sup> in order to acquire optical access to the hippocampus. Surgery was performed on mice under aseptic conditions, and body temperature was maintained with a heating pad (Harvard Apparatus). Mice were anesthetized with isoflurane (2.5% for induction, 1–1.5% for maintenance) and given a preoperative dose of meloxicam subcutaneously for analgesia (1mg/kg) and a postoperative dose 24 hours later. After asepsis, the skull was exposed, and the periosteum was removed.

A custom lightweight titanium headplate was attached to the skull with adhesive cement (C&B Metabond; Parkell). A craniotomy in the left hemisphere centered over the cornu ammonis 1 (CA1) (mediolateral (ML): -1.8 mm from the midline, anteroposterior (AP): 2.0 mm posterior from bregma) was made using a pneumatic drill. A small volume of overlying cortical tissue was aspirated to expose the external capsule; superficial fibers were then removed until the alveus became visible. A thin layer of Kwik-Sil (WPI) was injected into the resected area, and a metal cannula (316 S/S Hypo Tub 12T GA. 0.1080/0.1100" OD x 0.0890/0.0930" ID x 0.060" long; cut and deburred) with a coverglass (2.5mm diameter, Erie Scientific) attached to the bottom (NOA81 adhesive, Norland) was implanted on top of the Kwik-Sil, so that the Kwik-Sil served as a stabilizing medium between the glass and brain tissue. Another layer of adhesive cement was added to hold the cannula to the skull and the headplate. Mice were allowed to recover for at least five days before starting water restriction for behavioral training. Mice were extensively handled during the restriction process to familiarize them to experimenters. Mice were allotted daily volumes of 1–2 mL of liquid per day, delivered either during behavioral sessions or supplemented after sessions. Animals were examined daily to ensure that there were no signs of dehydration and that body mass of at least 80% of the initial value was maintained.

## Behavioral training.

The mice were trained to perform the accumulating towers task in a virtual reality (VR) environment, as previously described<sup>14–16,42</sup>. In short, mice were headfixed so that they could run comfortably on an 8-inch Styrofoam® ball suspended by compressed air. Ball movements were measured with optical flow sensors (ADNS3080) via an Arduino Due, and the VR environment was projected onto a coated Styrofoam® screen (~270° horizontal and ~80° vertical visual field) using a DLP projector (Mitsubishi). The virtual environment was generated using ViRMEn software<sup>28</sup>. Rewards were delivered by a solenoid valve (NResearch Inc.), controlled by a NI-DAQ card (PCI-6229, National Instruments). This VR system has been used previously<sup>14–16,42</sup> and was designed, by choice of material and size of the spherical treadmill, to minimize the amount of effort to turn the floating ball, such that the moment of inertia of a mouse pushing back the ball ( $2.78 \times 10^{-4} \text{ kg} \cdot \text{m}^2$ ) is comparable to the moment of inertia of a mouse pushing itself ( $2.68 \times 10^{-4} \text{ kg} \cdot \text{m}^2$ ).

Mice were trained to run down a 330 cm virtual T-maze (30 cm start region, 200 cm cue region, and 100 cm delay region). As mice ran through the cue region, tall, high-contrast visual cues (towers, 6 cm tall and 2 cm wide) were shown along either wall. After the delay period, mice were rewarded with a liquid reward for turning into the arm on the side where more towers were presented (4–8  $\mu\text{L}$  of 10% v/v sweet condensed milk or 10% w/v sucrose). Rewarded trials were followed by a 3 s ITI, and error trials were followed by an audio error cue and a 12 s ITI. When rewards or error cues were delivered, the visual display froze for the first second after which the display was then blacked-out. Average trial length for the seven experimental mice was  $6.3 \pm 0.8 \text{ s}$  SEM (cue:  $2.4 \pm 0.4 \text{ s}$  SEM; delay:  $1.9 \pm 0.2 \text{ s}$  SEM).

Tower positions were drawn randomly from spatial Poisson processes with means of 7.7 and 2.3 towers/m on the rewarded and non-rewarded sides, respectively. Towers were transient, appearing when animals were 10 cm away from their locations and disappeared after 200 ms. Each session started with at least 10 trials of a visually-guided version of the task as warm-up before proceeding to the main task. Behavioral sessions lasted  $48:16 \pm 03:44$  SEM (mm:ss; n=7 animals). For analyses, trials in which animals turned around 180° or backtracked to before halfway in the delay region were not included. Detailed methods for the shaping procedures involved in training mice to perform the task, as well as performance and behavioral analyses have been previously published<sup>14</sup>.

A different set of mice learned a simplified version of this task (“alternation task”), where no towers were presented in the T-maze. In one version of the alternation task (n=2; Extended Data Fig. 1), the walls were textured differently along the long stem, and large distal cues were added, as previously described<sup>41</sup>. The maze itself was also slightly longer (340 cm instead of 300 cm). In a second version (n=7; Extended Data Fig. 6), the maze was identical to the accumulating towers task, except no towers were ever shown. In both cases, animals simply needed to alternate between left and right turns to be rewarded. Visual guides were also present in the arm where the reward would be located.



## Two-photon cellular resolution calcium imaging.

The 2-photon calcium imaging setup was identical to a previously published design<sup>15</sup>. Two-photon illumination was achieved with a Ti:Sapphire laser (Chameleon Vision II, Coherent) operating at 920nm. Fluorescence was acquired using a 40× 0.8 NA objective (LUMPLFLN40X/W, Olympus) and GaAsP PMTs (H10770PA-40, Hamamatsu) after passing through a dichroic (FF670-SDi01, Semrock), an IR filter (FF01-720sp, Semrock), reflected by a second dichroic (FF562-Di03, Semrock) and passing through a final bandpass filter (FF01-520/60, Semrock). The PMT output signal was amplified (Variable High Speed Current Amplifier; #59-179, Edmund Optics) and digitized (PXIe-7961R FlexRIO, National Instrument). The microscope was controlled by ScanImage (Vidrio Technologies) software using additional analog output units (PXIe-6341, National Instruments) for the laser power control and the scanners control. Double-distilled water was used as the immersion medium for the objective. Average beam power measured at the front of the objective was 60–160 mW. The region between the objective and imaging site was shielded from external sources of light using a black rubber tube. Horizontal scans of the laser were achieved using a resonant galvanometer (Thorlabs). Typical fields of view measured approximately 500×500 μm, and data was acquired at 30Hz.

## Data processing and cell identification.

All imaging data was corrected for nonrigid brain motion via custom Matlab code based on a technique similar to NoRMCorre, where the image is divided into multiple overlapping patches and a rigid translation is estimated for each patch and frame by aligning against a template<sup>37</sup>. The set of translations are then upsampled to create a smooth motion field that is applied to a set of smaller overlapping patches, and the registered frame is then used to update the template by calculating a running mean of past registered frames.

After correcting for motion, fluorescence traces (downsampled to 15 Hz) corresponding to individual cells were extracted using a constrained non-negative matrix factorization algorithm (CNMF)<sup>36</sup>. Initialization of the spatial components for CNMF was done as previously published, as was classification of identified components into cell-like and non-cell-like categories<sup>15</sup>. Automated classification was followed by manual re-classification of a subset of components and artifact rejection. F/F for each cell was calculated using the modal value of fluorescence in 3-minute long windows as baseline fluorescence. An important note is that CNMF can only identify cells with calcium activity during the imaging session, hence total cell numbers reported are for active cells. Cells that were silent for the entire imaging session are not included here.

## Psychometric curves.

Psychometric curves (Fig. 1b) were plotted using methods described previously<sup>14</sup>.

In brief, psychometric curves were fit using a 4-parameter sigmoid,

$p_R(\Delta) = p_0 + B \left[ 1 + e^{-(\Delta - \Delta_0)/\lambda} \right]^{-1}$ , where  $\Delta$  is the difference between the number of right and left towers. The binomial confidence interval was calculated using Jeffrey's method<sup>14,15</sup>.

### Logistic regression analysis.

regression (Fig. 1c) was performed using methods described previously<sup>14</sup>. In brief, we modeled the animals' choices in each trial with logistic regression where the factors are the evidence (# right towers - # left towers) in five equally-sized regions in the cue period.

For both the psychometric curves (Fig. 1b) and the logistic regression analysis (Fig. 1c), all sessions in which animals (n=7 animals) performed above 60% correct were included (n=109 total sessions).

### Mutual information analysis.

For each cell, we evaluated a mutual information metric defined previously<sup>38</sup>,

$$I = \int \lambda(x) \log_2 \frac{\lambda(x)}{\lambda} p(x) dx, \text{ where } I \text{ is the mutual information rate of the cell in bits per second,}$$

$x$  is the mouse's spatial location,  $\lambda(x)$  is the cell's mean F/F at location  $x$ ,  $p(x)$  is the mouse's probability density of occupying location  $x$ , and  $\lambda = \int \lambda(x)p(x) dx$  is the overall

mean F/F of the cell.

To obtain  $\lambda(x)$ , we first denoised F/F by smoothing with a Gaussian filter with a length of 5 bins and thresholded the result so that values less than 2 robust standard deviations across the time series were set to 0.  $\lambda(x)$  was then calculated bin-wise by collecting all smoothed and thresholded F/F values in their respective bins across the entire session and taking the mean.  $\lambda(x)$  was then smoothed by convolution with a Gaussian filter with a length of 5 bins and a standard deviation of 1 bin.  $p(x)$  was calculated similarly by counting the number of frames that the animal spent in each bin across trials and normalized to have a sum of 1.

For position data, 10 cm bins from 0 cm to 300 cm were used. For evidence data, 31 bins (-15 to 15 #R - #L towers) were used. For multidimensional spaces where we randomized one of the dimensions (i.e.  $R_E \times Y$  and  $E \times R_Y$  in Fig. 2e), the randomized variables ( $R_E$  or  $R_Y$ ) were created by uniform random sampling with replacement from the joint distribution of discrete evidence (E) and position (Y) values. More specifically, for the  $R_E \times Y$  space, where Y is the non-randomized dimension, we first found the distribution of E values present in the data for each Y value. This created 30 separate E distributions with respect to Y. The  $R_E$  value for each frame was generated by randomly sampling from the sole E distribution that corresponded to the non-randomized Y value for that frame. This procedure was performed to control for the non-uniformity of the joint E×Y distribution in which specific combinations of E and Y values can have greatly different probabilities. A similar procedure was followed for the  $E \times R_Y$  analysis.

To determine significance, each cell's mutual information value was compared against the mean mutual information value of a shuffled dataset (100 shuffles), where each cell's F/F was circularly shifted by a random interval within each trial - disrupting the relationship between position and neural activity, while maintaining neural activity patterns. Only cells that had mutual information values greater than 2 standard deviations above the average mutual information of the shuffle distribution were considered statistically significant. Cells with statistically significant mutual information between neural activity and position in left

choice trials, but not right choice trials were categorized as left-choice preferring, while cells with statistically significant mutual information between neural activity and position in right choice trials, but not left choice trials were categorized as right-choice preferring. Those that were significant for both left and right choice trials were categorized as non-preferring. Similar tests were done for mutual information between neural activity and evidence, with the addition that cells where training and testing sets were not correlated (described below) were rejected.

For one-dimensional sequence plots (Fig. 2a, b, Extended Data Fig. 1c, and Extended Data Fig. 3),  $\lambda(x)$  were sorted and normalized based on their peak mean  $F/F$  values. For the cross-validation procedure for evidence fields (Extended Data Fig. 1c), trials were ranked based on the cell's maximum  $F/F$  value in a given trial. Odd-ranked and even-ranked trials were assigned to the training and testing sets, respectively.  $\lambda(x)$  was recalculated on the training and testing sets and smoothed as described above. Only cells with significantly correlated  $\lambda(x)$  between the training and testing sets ( $p < 0.05$ ) were used in the sequence plots. The training set was sorted based on peak mean  $F/F$  values and plotted. This same sorting index was then applied to plotting the testing set.

For Extended Data Fig. 2d, cells were considered to encode both evidence and position if they had significant mutual information in  $E \times Y$  space, as described above. Of the remaining cells, cells were considered to encode only position if they were significant in  $R_E \times Y$  space (16%) and only evidence if they were significant in  $E \times R_Y$  space (6%). For Extended Data Fig. 2e, distributions of mutual information in  $R_E \times Y$  and  $E \times R_Y$  space were calculated from 50 different shuffles, where either E or Y were shuffled. Of the  $E \times Y$  cells described above, 89.9% had mutual information values in  $E \times Y$  space greater than 2 standard deviations above both shuffled distributions. 9.8% had mutual information values that were greater than only the  $E \times R_Y$  distribution, and 0.3% had mutual information values that were greater than only the  $R_E \times Y$  distribution.

### Counting the number of place fields.

To estimate the number of place fields in  $E \times Y$  space, we followed a heuristic to count peaks derived from previous studies<sup>43,44</sup>. Using the neural activity maps for each neuron in  $E \times Y$  space (Fig. 2c, d and Extended Data Fig. 2a) obtained from methods described above (see Mutual information analysis), we considered all bins that surpassed 2 standard deviations above the shuffled mean as candidate place fields in the  $E \times Y$  space. We then joined all bins with adjacent significant bins, and if a connected component exceeded  $3 \times 3 = 9$  bins, we counted the connected component as a place field. The distribution of the place field counts is shown in Extended Data Fig. 2g. Note that a very small number of cells (3%,  $n=31/917$  cells) had significant firing fields above the shuffled control that were smaller than 9 bins. These appear in the histogram as “0”. Cells had approximately  $1.7 \pm 0.3$  SEM firing fields, with 53% ( $n=490/917$ ) of cells having more than one firing field.

### Manifold inference from neural dynamics.

To infer latent dimensions from neural dynamics, we adopted the procedure developed by Low, Lewallen and colleagues<sup>13</sup> for calcium imaging data. We first smoothed the raw

F/F traces with an 11-bin gaussian filter and thresholded at  $4\sigma$ , where we estimated the robust standard deviation ( $\sigma$ ) across the entire time series, but individually for every neuron. We restricted our analysis to cells that had at least one transient in the recording session, and imaging frames that had at least one active cell, as well as the portion of the maze represented by 0–300 cm (cue and delay periods). We then followed the procedure by Low, Lewallen, et al. to calculate distances between pairs of population activity vectors, extracting a set of latent variables from these distances with multidimensional scaling, and learning a map between latent space and network activity with local linear embedding (LLE).

Briefly, we first learn a generative model of transition probabilities from population activity in the training dataset  $s(t) = [s_1(t), \dots, s_N(t)]$  of  $N$  neurons at time  $0 < t < T$ , to the activity  $s(t + \Delta t)$  using the random forest method developed by Low, Lewallen, et al. with a few modifications. First, when splitting the neural state-space into regions using a set of hyperplanes organized in a decision tree, we assessed 20 random hyperplane orientations at every node of the tree and selected the orientation which best split the data. This improved performance with the large numbers of neurons typically encountered in calcium imaging. Second, we set the minimum number of leaves in each random tree to 500. Third, to define transitions, we considered all states  $\Delta t = 67$  ms apart (one frame at a 15 Hz frame rate). Fourth, we fit manifolds to all data points, not only a subset of landmarks. All other hyperparameters were chosen as in Low, Lewallen, et al.<sup>13</sup>. The random forest model provides us with a set of transition probabilities  $p(s(t + \Delta t) | s(t))$  that can be translated into a local distance  $\delta(s(t + \Delta t), s(t))$  under a Diffusion approximation, where the transition probability  $p$  decreases with distance  $\delta$  as  $p \sim \exp(-\delta^2)$ . Similar to isomap<sup>45</sup>, we then calculated the global distance between two states as the length of the shortest path from one to the other via any intermediate, connected states. The pairwise geodesic distances of  $I$  points  $\rho(i, j)$ , where  $0 < i, j \leq I$ , then yields a matrix of size  $I \times I$  that was embedded using multidimensional scaling with Sammon's nonlinear mapping. This yielded latent variables to describe population data. The mapping from latent space to neural activity and back was then achieved with local-linear embedding<sup>13</sup>.

### Manifold inference on video files.

To construct a low-dimensional representation of the task itself, we applied the algorithm described above to the visual input that the mice received in a typical experimental session, more specifically to the blue channel across all rgb pixels in each frame of the video files displaying the animals' field of view. This corresponds to a vectorized time series of  $1792 \times 1088 = 1,949,696$  pixels as a function of time. To make this analysis computationally viable, we first downsampled the videos  $17\times$  from the original  $1792 \times 1088$ , restricted our analysis to trials shorter than 30 s and frames with positions between 0 and 350 cm, and simplified the hyperparameters, in comparison to the analysis of neural data, by using only two random hyperplane orientations and 1000 landmarks. All other parameters were identical to the analysis of neural data. The results are shown in Extended Data Fig. 5, where Extended Data Fig. 5a shows the mean luminance of the blue channel, after averaging across all pixels.

### Dimensionality estimation.

To estimate the dimensionality of the latent manifold, we analyzed the geometric properties of the geodesic distance matrix  $\rho(i,j)$ . We specifically studied the statistics of nearest neighbor distances. Suppose that the neural states were confined to a two-dimensional sheet in high-dimensional neural state space. Within the sheet, counting the cumulative number of points  $N$  within distance  $r$  will increase quadratically with distance  $r$ , as more points on the sheet will fall within the neighborhood, thus recovering the two-dimensional sheet structure. Using this variation of the correlation dimension that can also be used for complex attractor geometries<sup>12,39</sup>, we found a wide range of values where the number of points scaled like a power law.

We fit this power law by minimizing the quadratic error to the model function  $N(r) = cr^d$ , where  $N$  is the total number of neighbors,  $r$  is the distance, and  $c$  and  $d$  are fit parameters. We fit this function over three orders of magnitude, from  $10^3 < N < 10^6$ . The average across the seven mice yielded  $d=5.4$  [4.8, 6.0] (95% bootstrapped confidence intervals). These numbers are consistent with a  $d \sim 4$ - to 6- dimensional manifold, embedded in a  $\sim 450$ -dimensional neural state space (Fig. 3a). For the illustrations in Fig. 3a and Extended Data Fig. 3a, we normalized the distance by the average length of a trial along the manifold for each animal.

### Reconstructing neural data from embedded manifolds.

To assess the quality of the dimensionality reduction performed with MIND, we measured how well the neural data can be reconstructed from the  $d$  latent variables after embedding the manifold into  $d$  dimensions (Extended Data Fig. 4). This provides us with an estimate of the minimum number of dimensions required for the reconstruction quality to saturate. This number should be comparable to the intrinsic dimensionality of the manifold, and thus provided us with a separate measurement of the manifolds' dimensionality.

### Measuring how well the coordinated activity of neurons is predicted by the manifold.

To this end, we held out a random trial, fit a manifold to the remaining data, and embedded this manifold into two to seven dimensions using the methods described above. After fitting the manifold on the training data, we first projected the held-out trial onto the manifold to obtain  $d$  coordinates for every time point and then reconstructed neural activity from these  $d$  numbers in the test dataset using LLE<sup>13</sup>. We then thresholded the LLE estimate in order to capture the thresholding nonlinearity of calcium imaging. The thresholding cutoff was estimated from the training data for best reconstruction. To assess the similarity between the raw data and the reconstruction, we then measured the correlation coefficient between the reconstructed neural data and the real data. Note that these data are a vectorized time-series of the form neurons  $\times$  time. To perform an element-wise comparison, we concatenate all columns into a single vector and calculate the correlation coefficient. This number was averaged across the 10 held-out trials, i.e. the decoding index, and the process was repeated for all seven animals (Extended Data Fig. 4a, b). The data shown in Fig. 3c is the mean  $\pm$  SEM for the seven animals. In Fig. 3b, raw  $F/F$  and reconstructed  $F/F$  traces have been smoothed with an 11-bin gaussian filter and thresholded at  $4\sigma$ . For the reconstructed

F/F traces, baseline subtraction prior to smoothing and thresholding was accomplished by subtracting the mean of the reconstructed activity of each cell from the activity of each cell.

### Measuring how well the activity of individual neurons is predicted by the manifold.

This analysis is similar to the one above but tailored to quantify the predictive power of the manifold on a single-cell level (Extended Data Fig. 4c). To this end, we removed 1 test-neuron from all the  $N$  cells in the data and used MIND to fit a manifold to the remaining  $N-1$  training-neurons<sup>46</sup>. We then used gaussian process regression to learn a map  $g(x)$  from manifold coordinates  $x$  to the test neuron's activity in 80% of trials. We used a squared exponential kernel function to specify the covariance, making these fits smooth and differentiable, as expected for a response similar to a firing field. In the remaining 20% of trials, we evaluated  $g(x)$  and measured the correlation coefficient between the predicted and observed data of the test-neuron. This was repeated 4 more times for 5-fold cross-validation and the correlation coefficient over the five folds was averaged. This value was calculated for 10 randomly chosen neurons from the 25 most active neurons in each animal and averaged, i.e. the decoding index (Extended Data Fig. 4e). In Extended Data Fig. 4d, reconstructed F/F traces were baseline-subtracted, smoothed, and thresholded identically to the procedure mentioned above for Fig. 3b.

### Comparing MIND with PCA.

To compare this nonlinear dimensionality reduction technique with a linear method, we also calculated the decoding index (cross-validated correlation coefficient between predicted and observed data in a held-out trial) using principal component analysis (PCA). To this end, we removed a held-out trial from the data, calculated the principal components (PCs) of the remaining data and identified the  $d$  PCs with greatest coefficients in the training data. We then projected the held-out trial onto these  $d$  PCs and used the obtained coefficients to project back into neural state space. The similarity of the observed held-out trial and the reconstruction from PCA was assessed with the correlation coefficient and averaged across 10 random held-out trials. To reach the same mean cross-validated decoding index as MIND for manifolds embedded in  $d=4, 5$  and 6 dimensions, PCA required  $d=29, 40$ , and 47 principal components, respectively.

### Decoding position and evidence from the manifold and neural activity.

We used gaussian process regression to learn a function from latent space or neural activity (selecting only the top 10% of cells with highest mutual information for position or evidence to limit overfitting) to position and evidence. Other nonlinear regression methods like LLE yielded similar results, while linear decoding methods generally failed. Fig. 3g shows the correlation coefficients between the position and evidence values in each animal's behavioral session predicted from the learned regression model (trained on 80% of trials, applied to the test dataset of 20% of trials, and repeated for 5-folds) and true position and evidence values (averaged over the 5-folds), i.e. the decoding index. For visualizing evidence and position (Fig. 3f), as well as luminance (Extended Data Fig. 5b) and view angle (Extended Data Fig. 6a), we smoothed across the 20 nearest neighbors in latent space.

Similar methods were used in analyses shown in Extended Data Fig. 6. To assess whether knowledge of variable X adds to how well variable Y predicts the manifold, we decoded manifold dimensions (as described above) using both X and Y as inputs or X and shuffled-Y. To assess whether correlated and orthogonal components of X and Y could both be decoded, i.e. linearly regressing out Y from X, we used PCA on variables X and Y and decoded both PC1 and PC2 from the manifold dimensions. For evaluating the accuracy of the decoding for binary variables, such as the upcoming choice, the choice in the previous trial, and whether the previous trial was correct, we averaged the prediction from the gaussian process regression across the trial to come up with a single value, which was binarized as a single prediction, i.e. left or right choice, and compared it to the true value, i.e. whether the animal makes a left or right choice in the trial.

### Hyperalignment procedure.

Hyperalignment across two animals was performed as follows. We first fit the neural data of subject *A* with MIND, to obtain a set of  $T$   $d$ -dimensional latents  $x_A^t$ . We then perform GPR to learn a map from the  $d$ -dimensional latents to a behavioral variable  $e_A^t = \text{GPR}(x_A^t)$ . This is the same analysis as earlier in Fig. 3. Next, we perform MIND on the data of subject *B*. This yields a different set of  $d$ -dimensional latent vectors  $x_B^t$ . From these latents, we predict the behavior of subject *B* using the GPR trained on mouse *A* and a 5-dimensional rotation matrix  $R$  with  $e_B^t = \text{GPR}(R x_B^t)$ . The rotation matrix was calculated from a five-dimensional representation of the special orthogonal group of degree 5 [SO(5)] so that  $R = \prod_{i=1}^{10} \text{expm}(g_i c_i)$ . Here,  $\text{expm}()$  indicates the matrix-exponential of  $g_i$ , the ten generators of SO(5), multiplied with a scalar angular parameter  $c_i$ . These parameters were cross-validated by optimizing on the first half of the data and decoding of position and evidence assessed on the second half. For each animal, we decoded position and evidence using the hyperaligned 5-dimensional manifolds of the six other manifolds. In Fig. 3i, we show the maximum decoding that can be done across the six other animals for each animal, compared to the cross-validated GPR (Fig. 3g) for 5-dimensional embeddings of the manifold in the same animal. Means were then calculated across the seven animals. We estimated the contribution of shared geometry for each animal in terms of fractional variance explained by dividing the  $r^2$  of position and evidence decoding obtained with hyperalignment by the  $r^2$  of the best decoding that could be done with either method.

### Task trajectories.

To visualize the sequential patterns of the task (Extended Data Fig. 8), we first extracted “task trajectories”, i.e. smooth spline interpolations of the specific trajectory through E×Y space experienced over trials. The task trajectories for single trials in a behavioral session are plotted as thin lines in Extended Data Fig. 8a, together with fits across all left or right trials (thick lines). In addition, we also visualized task trajectories as a flow field, where we binned E×Y space into 10 cm and 1 tower bins and calculated the trial-averaged gradient in the position and evidence directions for every bin. The resulting gradient-matrices were then individually smoothed by convolution with a Gaussian filter with a length of 5 bins and a standard deviation of 1 bin. Every other bin was plotted as arrows centered on the respective bin and pointing to the average direction of the gradient (Extended Data Fig. 8b).

### Identification and analysis of sequences (“doublets” and “triplets”).

A pair of cells was classified as a doublet if the number of trials in which the first cell had a transient event before the second cell was greater than 2 standard deviations above the mean of the same value obtained from a shuffled dataset (100 times) where neural activity was circularly shifted in each trial. Doublets that appeared in fewer than 3 trials were removed. A transient event was defined as any time  $F/F$  (smoothed with a gaussian filter with a length of 5 bins) for that cell was greater than a threshold equal to 11 (Ai93xEMX1) or 5 (GP5.3) times the robust standard deviation across the entire imaging session. Different thresholds for event detection were used for the two animal strains due to the difference in signal-to-noise ratios. Triplets were constructed by simply combining doublets without allowing the same cell to appear twice, i.e. a cell cannot be the first cell and the third cell in the triplet, and tested using the same significance test as was used for doublets.

Even if two place cells had activity that was completely independent, we would still expect, by chance, that they fire in the same trials for a subset of trials, i.e. two place cells with fields at 100 cm and 200 cm that are each active in 100 random trials in a session with 200 trials would, on average, show up together in 50 random trials. However, if these two cells appeared in all 100 trials together, it would be unlikely that their activity was independent. To test whether doublets appeared more often than chance, trial IDs of each cell were independently shuffled, so that relationships between cells were disrupted without affecting the neural activity of each cell (Extended Data Fig. 8), and then we searched this shuffled dataset for the doublets again to determine the number of instances a doublet would show up if the activity of the two cells were independent ( $n=100$  shuffles).

A doublet was determined to be choice-predictive if the probability that the animal was going to turn right in trials in which the given doublet occurred was greater than 2 standard deviations above or below the mean probability of a right turn after shuffling the choices for each trial ( $n=1000$  shuffles). The same assessment was made to determine choice-predictiveness in triplets. Once choice-predictive doublets and triplets were identified, we compared the predictiveness of real doublet events to events obtained from datasets in which trial IDs were shuffled ( $n=100$  shuffles).

### Comparison of sequences and the predictions from the manifold.

To show the manifold’s predictive power for sequences, we used the manifold to reconstruct  $F/F$  of each cell in each imaging session with LLE (described above). We then detected doublet events from this reconstructed data and compared the trials in which doublet events were found against the real data to generate the true positive rate (TPR) and false positive rate (FPR) for doublet events in each animal.

More specifically, we constructed a boolean array  $B_{data}$  of size  $N_{cells} \times N_{cells} \times N_{trials}$  indicating the presence or absence of a doublet in a specific trial. We populated this array with the doublet-finding algorithm described above and the observed calcium data. This constitutes ground truth. We then reconstructed all neural activity from the latent dimensions of the 5-dimensional embedding of the manifold. This activity was then thresholded at an activity level  $\theta$ , and we considered only transients that exceed this threshold. By definition,



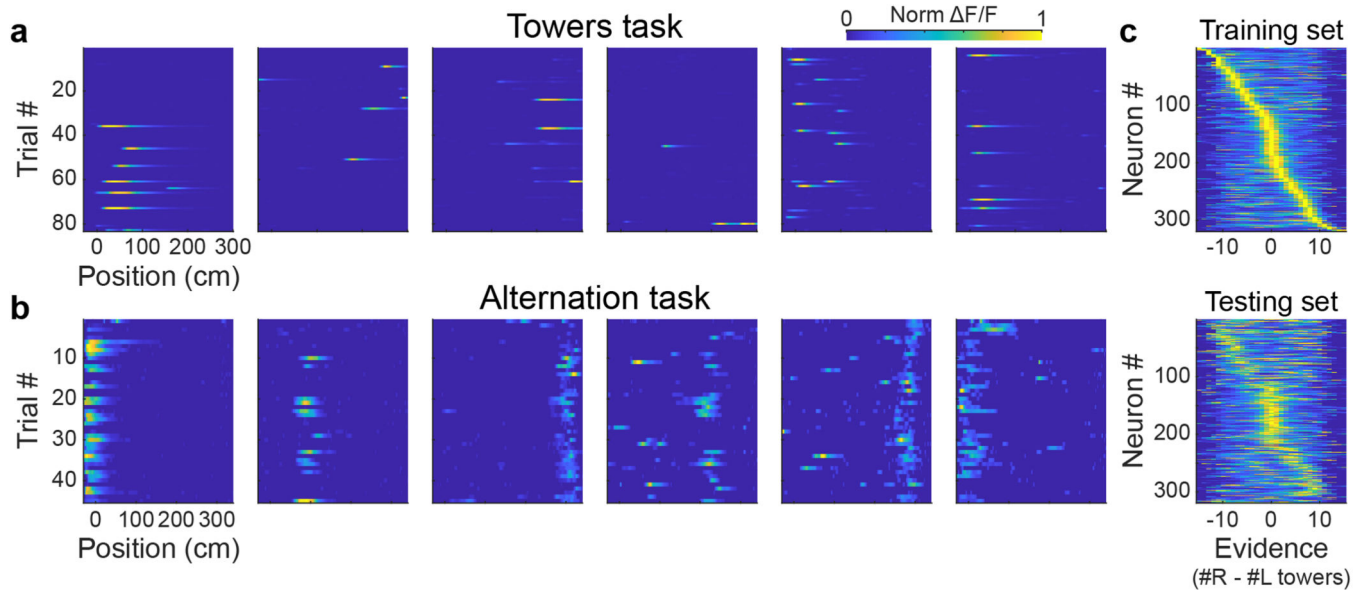
this data is the manifold prediction. We identified doublet events in this surrogate data with the same algorithm to construct a boolean array  $B_{prediction}$ . Comparing this prediction with the ground truth, we can count the number of true positives (“1” in both the ground truth and the surrogate array), false positives (“1” in the surrogate array, “0” in the ground truth), false negatives (“0” in the surrogate array, “1” in the ground truth) and true negatives (“0” in both the ground truth and the surrogate arrays). True positive rate (TPR) was defined as  $TP / (TP + FN)$ , where  $TP$  is the number of true positives, and  $FN$  is the number of false negatives. False positive rate (FPR) was defined as  $FP / (FP + TN)$ , where  $FP$  is the number of false positives, and  $TN$  is the number of true negatives. We then scanned across  $\theta$  (1 to 100, in intervals of 5) to construct a receiver operating characteristic (ROC) curve (Extended Data Fig. 9d) and calculated the distance  $d$  between the point (0, 1) on the upper left-hand corner of ROC space and any point on the ROC curve  $d^2 = (1 - TPR)^2 + FPR^2$ . We chose the threshold  $\theta$  such that this distance was minimal to identify a point of best discriminant capacity. The values of TPR and FPR reported in the main text are averages across these points for all seven animals.

We next calculated the predictive power of the manifold for the exact timing of a doublet. For all doublets, we measured the length of the trajectory between the first cell’s firing and the second cell’s firing on the manifold in each trial when the doublet was active. This length, plotted against the time between the sequentially active cells, is shown in Fig. 4e. To test whether the observed correlation of time elapsed and distance on the manifold was significantly greater than the correlation between time elapsed and any distance on the manifold, we compared the observed correlation to the correlation coefficients obtained from comparing time elapsed in a trial with manifold distances over the same time interval obtained from a different trial. We averaged the correlations across 100 random trajectories obtained from other trials and over all doublets for each animal and performed a two-tailed Wilcoxon signed rank test on the average real and random correlation values of the mice ( $n=7$ ) to test if real correlation values were significantly greater than the random correlation values.

### Statistical Tests.

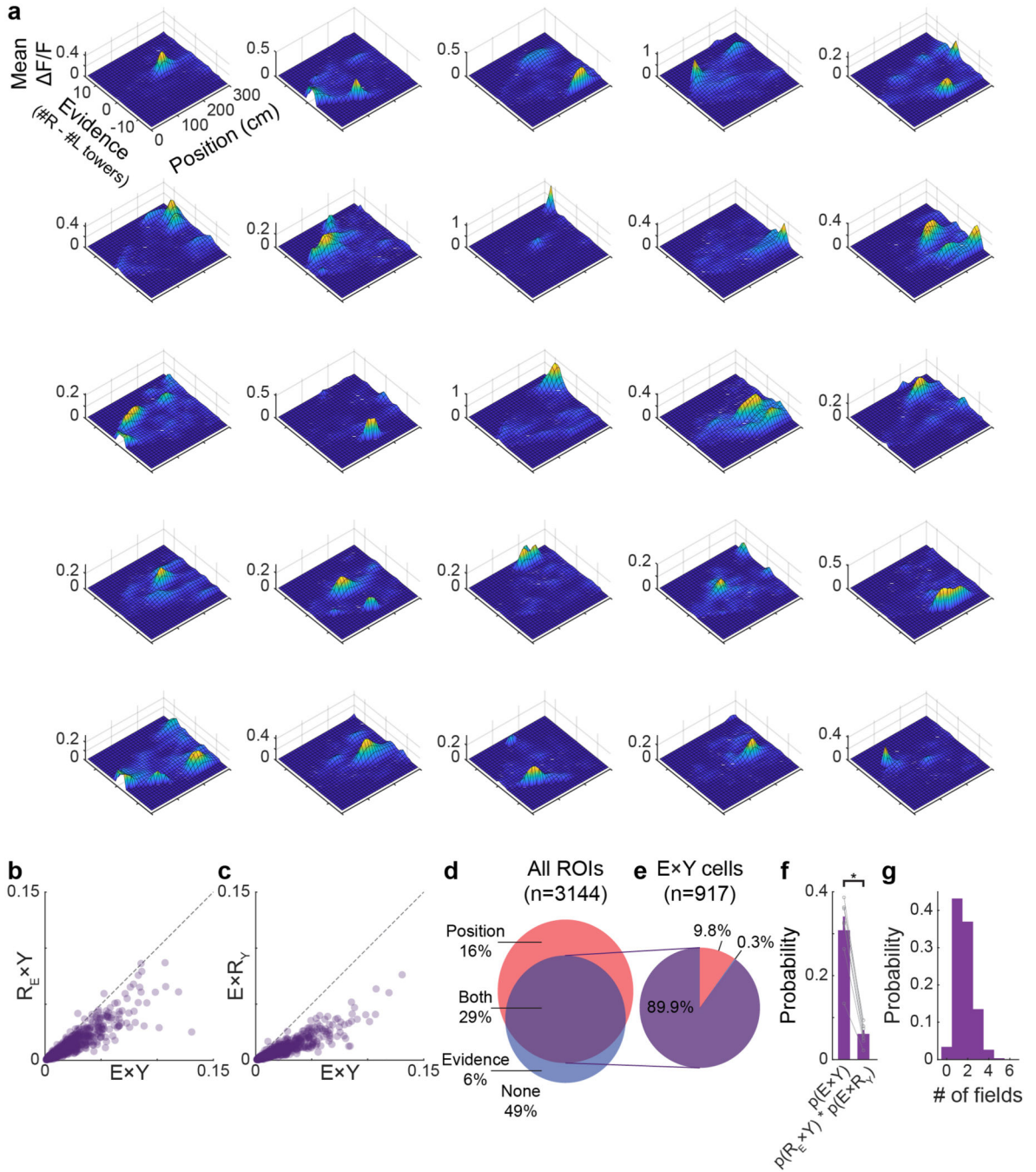
All statistical tests were performed with Matlab (2015b, 2018a, 2018b, and 2020a; Mathworks Inc). Bonferroni correction of p-values was performed by multiplying the unadjusted p-value by the number of multiple comparisons made. In cases where the corrected p-value exceeded 1.0, we reported the value as 1.0.

## Extended Data



**Extended Data Figure 1. Characterization of CA1 neural variability in the accumulating towers task.**

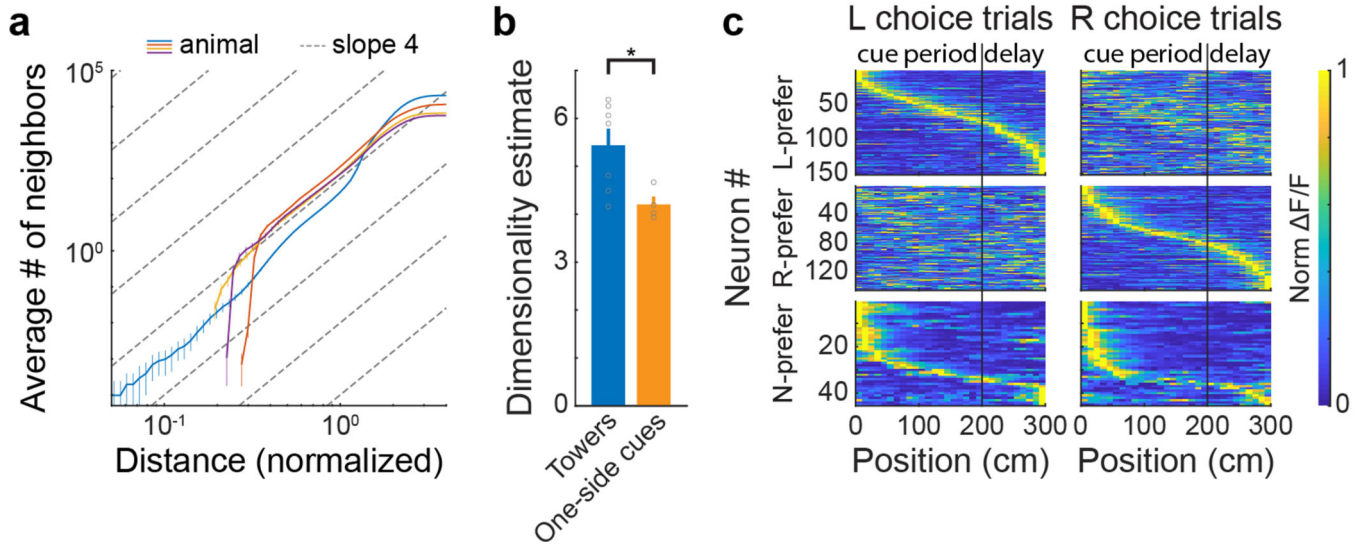
**a**, Each heatmap represents one neuron and the trial-by-trial activity of that neuron in the towers task for left-choice trials. Each row in each heatmap is the  $\Delta F/F$  (normalized within each session) of the neuron in that trial. **b**, Same as in **a**, but for the alternation task. Note that the single trial activity appears more variable in the towers task and more reliable in the alternation task, consistent with the results that evidence is also being represented by neurons in the towers task. **c**, Neural activity ( $\Delta F/F$  normalized within each neuron) of cells significantly encoding evidence, sorted by activity in half the trials (left), and plotted using the same sorting in the other half of the trials (right).



**Extended Data Figure 2. Place fields in evidence-by-position (E×Y) space.**

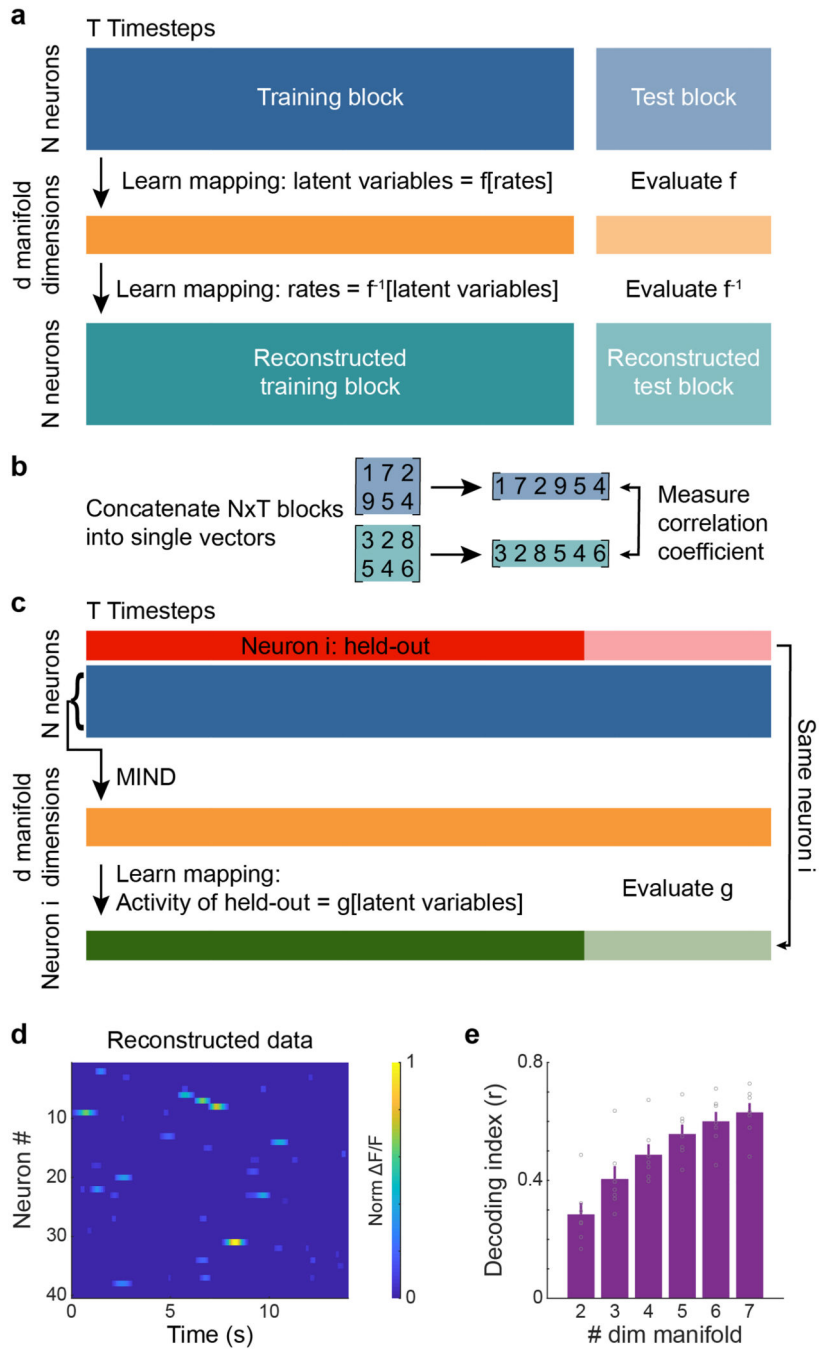
**a**, Each heatmap shows the mean  $\Delta F/F$  of a neuron with significant mutual information in E×Y space. **b**, Scatterplot of the mutual information in  $R_E \times Y$  space vs E×Y space for each cell with significant information in E×Y space ( $n=917$  neurons).  $R_E$  is randomized evidence. **c**, Same as in **b**, but for  $E \times R_Y$  space vs E×Y space.  $R_Y$  is randomized position. **d**, 29% of imaged neurons had significant mutual information in E×Y space, while 16% had significant mutual information for position only and 6% had significant mutual information for evidence only. **e**, Of the cells with significant mutual information in E×Y space, 89.9%

had significantly more information in E×Y space than just place or evidence information alone, while 9.8% could not be differentiated from place cells, and 0.3% could not be differentiated from evidence cells (see Methods). **f**, The probability of a cell having significant mutual information in E×Y space is significantly greater than the joint probability of a cell being a place cell and a cell being an evidence cell (two-tailed Wilcoxon signed rank test,  $n=7$  animals,  $*p=0.016$ ; error bars: mean±SEM). **g**, Cells with significant mutual information in E×Y space had  $1.7\pm 0.03$  SEM firing fields ( $n=917$  cells).



**Extended Data Figure 3. Dimensionality of an earlier training stage.**

During the training of the towers task, animals proceed through various stages of training. In one of these training stages, animals perform a task virtually identical to the towers task, except that visual cues only show up on one side of the maze. **a**, The intrinsic dimensionality of the one-side cues task is  $\sim 4.2$  [4.0, 4.5] ( $n=4$  animals, bracket values represent 95% bootstrapped confidence interval; error bars: mean±95% bootstrapped confidence intervals for each animal). **b**, Intrinsic dimensionality of the one-side cues task is significantly lower than the dimensionality of the towers task (two-tailed Wilcoxon rank sum test,  $n=7$  towers task animals and  $n=4$  one-side cues task animals,  $*p=0.042$ ; error bars: mean±SEM). **c**, Choice-specific place cell sequences in the one-side cues task, similar to Fig. 2a. Sequences are divided into left-choice (top row), right-choice (middle row) and non- (bottom row) preferring cells. Data is split between left-choice trials (left column) and right-choice trials (right column). Cells are shown in the same order within each row group. F/F was normalized within each neuron.

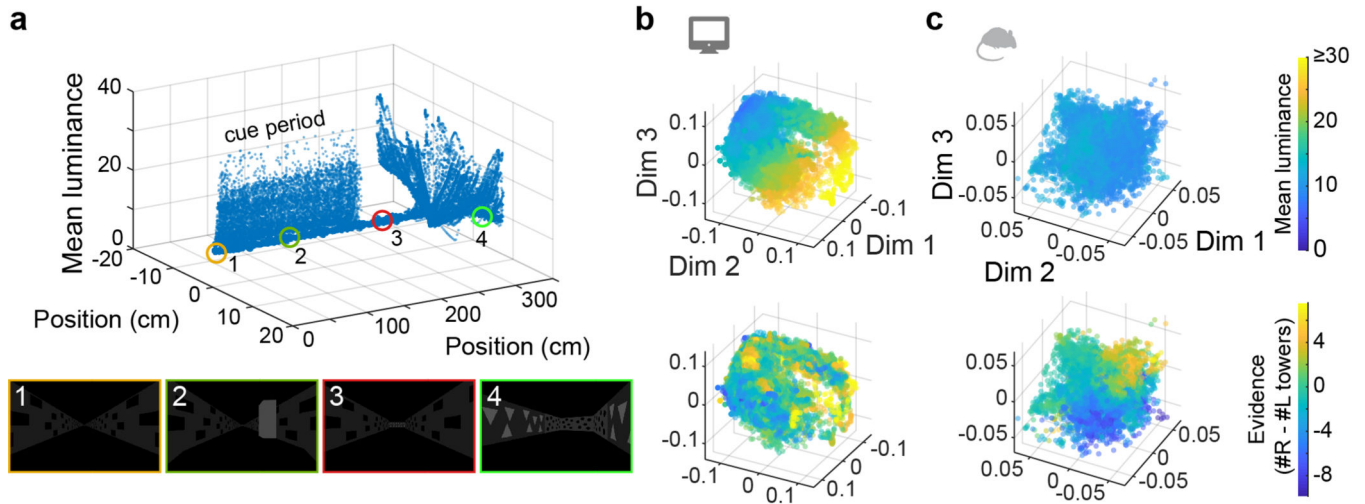


**Extended Data Figure 4. Cross-validation methods and results demonstrating how neural activity from single neurons is captured by coordinated population activity.**

**a**, Illustration of the cross-validation method to calculate the decoding index in Fig. 3c. Data is split for training (solid colors) and testing (shaded colors). With the training data, a map is obtained from  $F/F$  to latent dimensions and back. This map is evaluated on the test data.

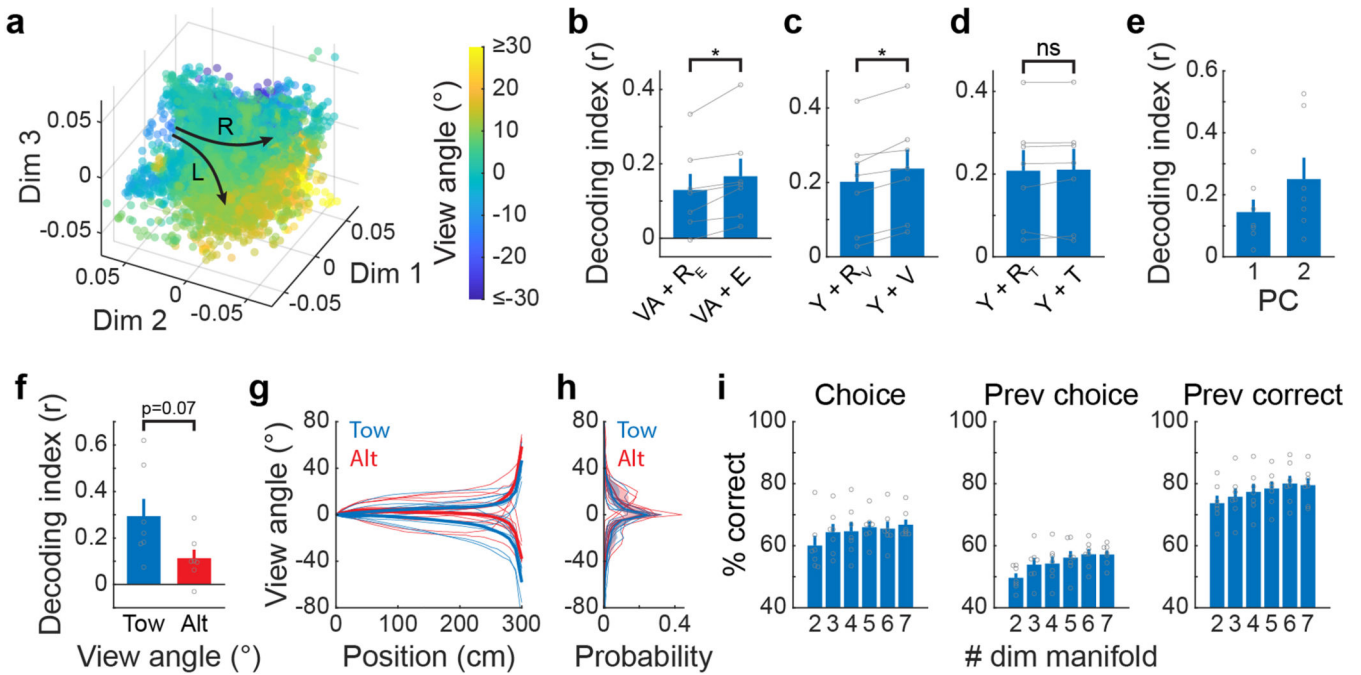
**b**, To assess the performance of the map, we concatenate the neuron  $\times$  time data in the test block and reconstructed test block into two vectors and calculate the correlation coefficient from the elementwise pairwise comparison of the vectors. The correlation coefficient was

averaged across 10 individually held-out trials to yield the decoding index. **c**, Illustration of a similar analysis where the activity of a single cell is decoded from a manifold fit to the rest of the neural population. One neuron (red) is removed before using MIND to obtain a set of latents. Next, in the training data (solid green), a map is calculated from the manifold to the held-out neuron's activity. The map is then used to predict the test data (shaded green). The correlation coefficient is calculated as in **b** and averaged over 5-folds as the decoding index. **d**, Example of neural activity from 40 individually reconstructed neurons, where activity of each neuron was decoded from the 5-dimensional manifold fit to the other cells following procedures in **c** (comparable to Fig. 3b, where the method in panels **a** and **b** was used).  $F/F$  is normalized to the maximum  $F/F$  in the window shown. **e**, Cross-validated correlation coefficients between activity of individual neurons in the real and reconstructed data, where reconstruction was accomplished with  $d$ -dimensional embeddings of the neural manifold. Decoding index is the correlation coefficient between the predicted and real  $F/F$  of the held-out ROIs ( $n=7$  animals; error bars: mean $\pm$ SEM).



**Extended Data Figure 5. Task manifold and neural manifold encode different variables.**

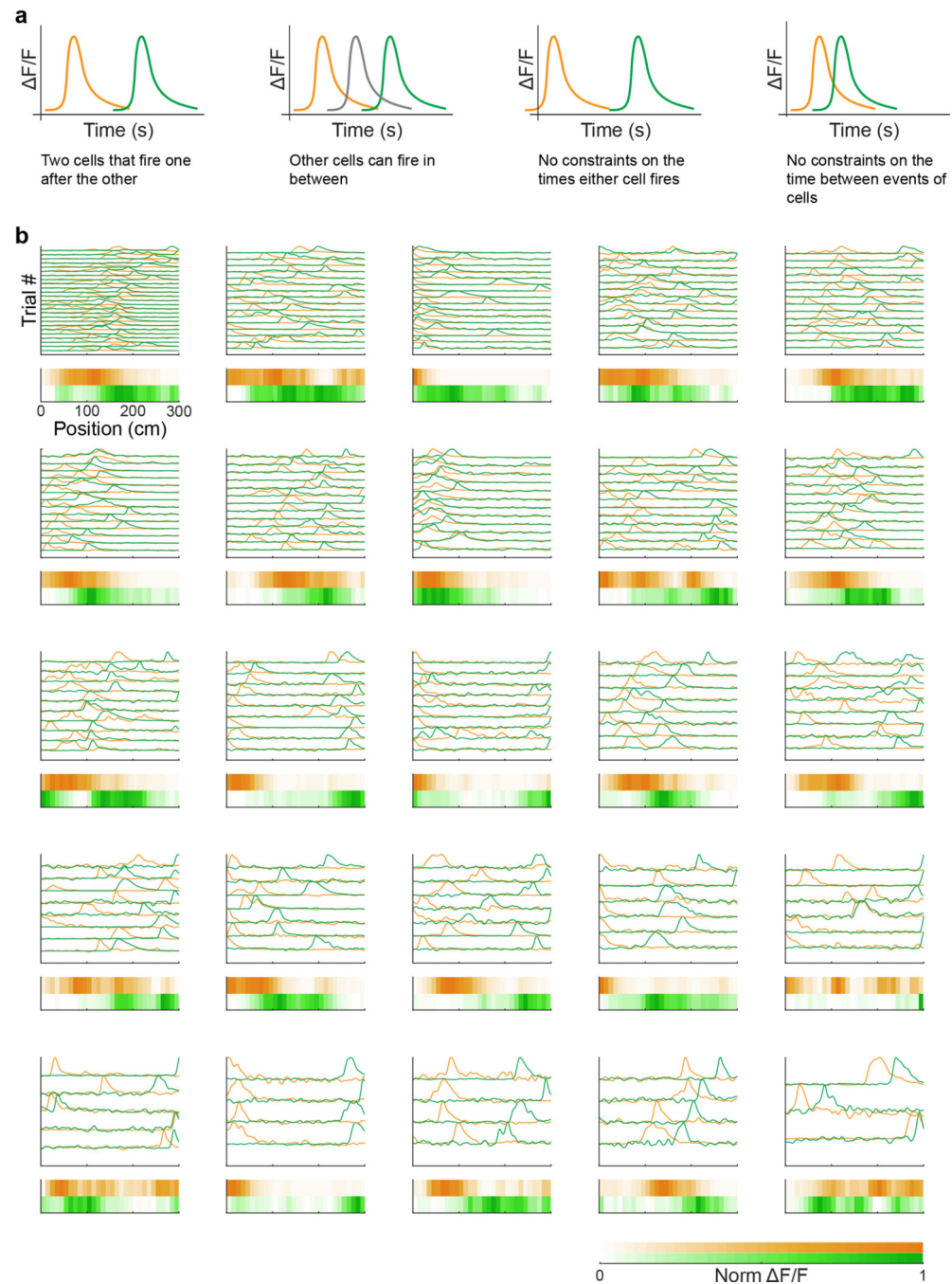
**a**, The visual space of the accumulating towers task across a representative session. Shown is the mean luminance of the virtual reality visual field as a function of position in the T-maze. Four example frames are shown below. Note the high variability of luminance during the cue period, where bright towers are randomly presented on the left and right walls. **b**, Performing dimensionality reduction on the pixels' time series in the raw video stream using MIND reveals a low-dimensional manifold, reflecting the visual sensory structure of the accumulating towers task. Plotting luminance (top) and evidence (bottom) on the manifold reveals that luminance is represented as a smooth gradient, whereas evidence requires memory and is thus absent on the task manifold. **c**, same as in **b**, but showing the neural manifold obtained from the animal that ran this session (Fig. 3f). Notice the absence of a luminance representation, but the emergence of evidence.



**Extended Data Figure 6. Decoding other variables from the neural manifold.**

**a**, Similar to Fig. 3f, view angle is plotted as color on the 3-dimensional embedding of the manifold. **b**, The 5 latent variables of the neural manifold embedded in 5-dimensional space are better predicted by gaussian process regression from view angle and evidence values than from view angle and shuffled evidence values (two-tailed Wilcoxon signed rank test,  $n=7$  animals,  $*p=0.016$ ; error bars: mean $\pm$ SEM). Decoding index is the correlation coefficient between the predicted manifold values and true manifold values, averaged over the 5 dimensions of the manifold. **c**, Same as in **b**, but for decoding manifold values using position and velocity. The addition of velocity to position information significantly improves the decoding of manifold values (two-tailed Wilcoxon signed rank test,  $n=7$  animals,  $*p=0.016$ ; error bars: mean $\pm$ SEM). **d**, Same as in **b**, but for decoding using position and time. The addition of time information does not significantly increase how well manifold values are decoded (two-tailed Wilcoxon signed rank test,  $n=7$  animals,  $^{ns}p=0.30$ ; error bars: mean $\pm$ SEM). **e**, We used PCA to separate the correlated and orthogonal dimensions between evidence and view angle and decoded both PC1 (correlated) and PC2 (orthogonal) from the neural manifold embedded in 5-dimensional space ( $n=7$  animals; error bars: mean $\pm$ SEM). Decoding index is the correlation coefficient between the predicted PC and true PC values. **f**, View angle is better decoded from the neural manifold (5-dimensional embedding) in the towers task (“Tow”), when evidence is also present, than in the alternation task (“Alt”) when evidence is not present (two-tailed Wilcoxon rank sum test,  $n=7$  towers task animals and  $n=7$  alternation task animals,  $p=0.07$ ; error bars: mean $\pm$ SEM). Decoding index is the correlation coefficient between the predicted and true view angle values. **g**, Average view angle trajectories, separated between left- and right-choice trials, for the towers task ( $n=7$ ; blue/thin) and the alternation task ( $n=7$ ; red/thin) animals. Thick lines represent averages across animals. **h**, Average view angle values in the towers task ( $n=7$ ; blue/thin) and the alternation task ( $n=7$ ; red/thin) over all trials. Thick lines and shaded area: mean $\pm$ 95%

bootstrapped confidence interval. **i**, Accuracy in predicting the upcoming choice (left), the animal's choice in the previous trial (center), and whether the previous trial was rewarded (right) from  $d$ -dimensional embeddings of the neural manifold ( $n=7$  animals; error bars:  $\text{mean} \pm \text{SEM}$ ).

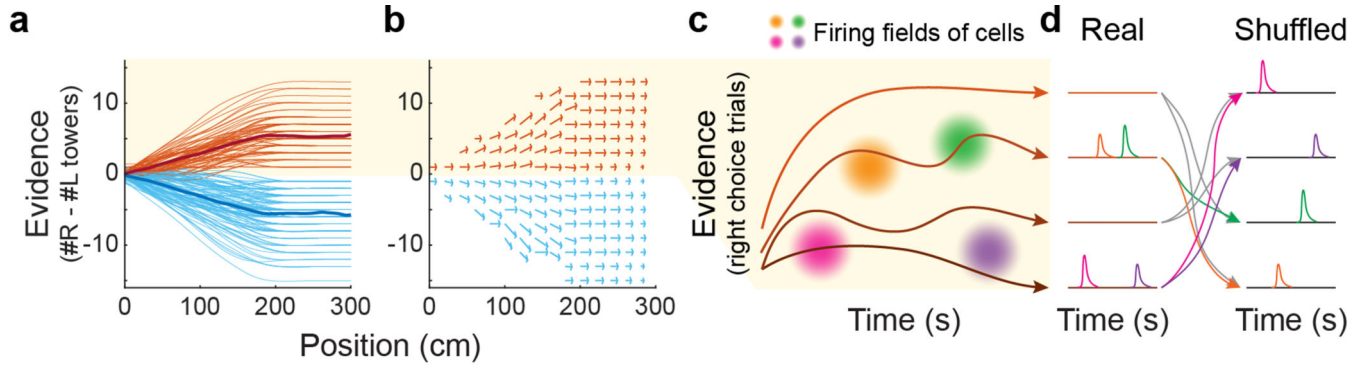


**Extended Data Figure 7. Examples of sequences in CA1 neural activity.**

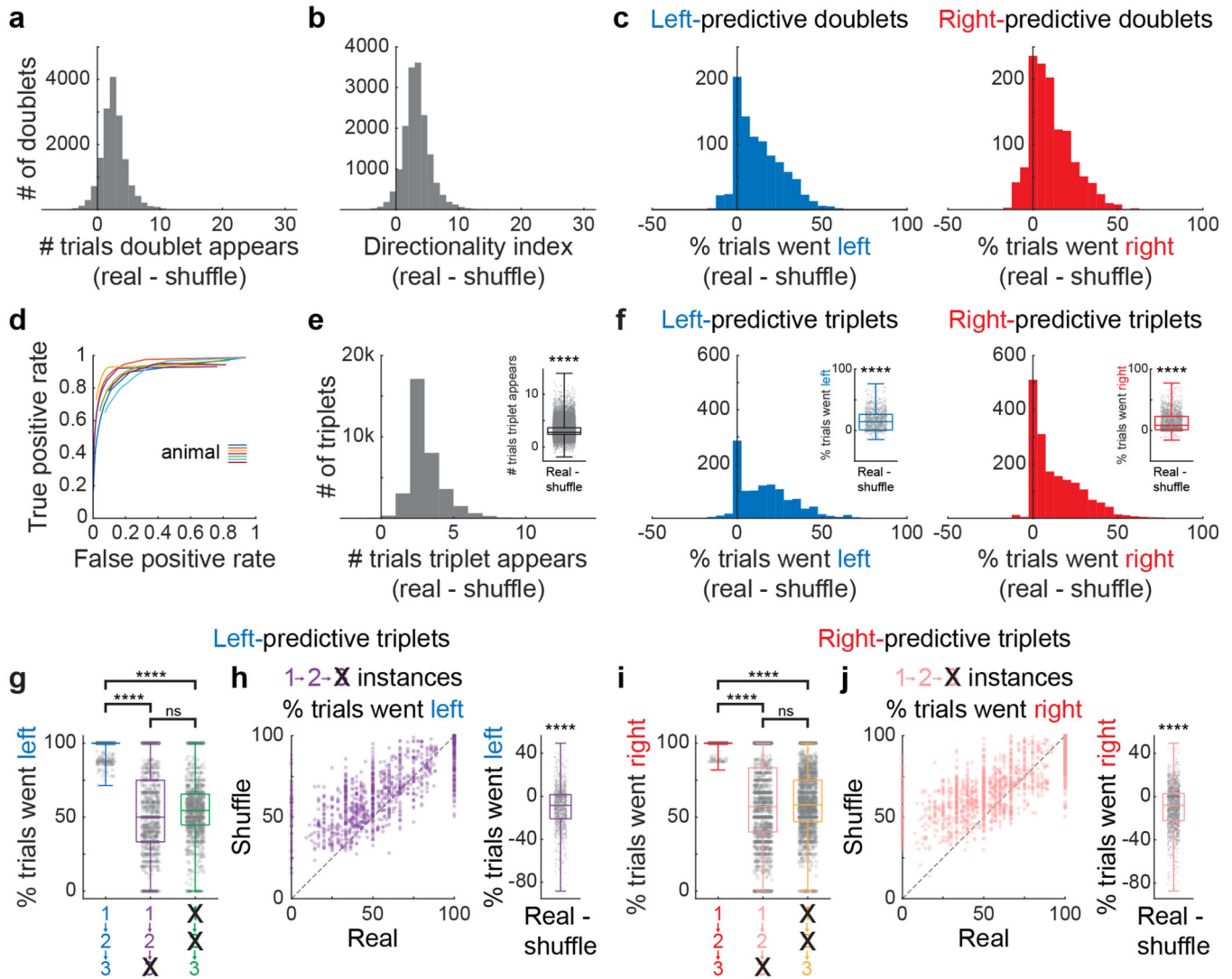
**a**, Schematic to describe how “doublets” were defined. Orange and green are calcium traces of the cells under consideration. Grey is the calcium trace of a third cell. **b**, 25 examples of



doublets in a single session from one animal. Each panel shows traces for trials in which the doublet was present. Orange traces are the neural activity from the first cell in the doublet, while green traces are the neural activity from the second cell in the doublet. Heatmaps represent the normalized neural activity of each cell across all trials in the session.



**Extended Data Figure 8. Neural activity generated by trajectories through the task.**  
**a**, Trajectories through evidence and position in one session of the task. Each thin line represents a fit with a cubic spline to a single trial, while thick lines represent fits over all trials in which the animal was supposed to turn left or right. **b**, Shown is the average change of position and evidence over time across trials in a single session for a set of representative states in evidence and position space. **c**, Conceptual diagram showing four trajectories through the neural manifold in right choice trials. Two different doublets are activated because the trajectories pass through their firing fields. **d**, Shuffling trial IDs within right choice trials will disrupt doublet activity while maintaining trial-averaged place and choice preferences of each cell.



**Extended Data Figure 9. Choice-predictive sequences in CA1 neural activity.**

**a**, Distribution of the values in Fig. 4b. **b**, Distribution of the values in Fig. 4c. **c**, Distribution of the values in Fig. 4f. **d**, Receiver operating characteristic (ROC) curves for sequential activity predicted from the 5-dimensional embedding of the manifold compared to sequential activity in real data ( $n=7$  animals). **e**, Similar to **a**, but for triplets. Inset shows that triplets are significantly more likely to appear in the real data than in the shuffled dataset where trial IDs were shuffled (two-tailed paired t-test,  $n=34737$  triplets, \*\*\*\* $p<0.0001$ ). **f**, Similar to **c**, but for triplets, showing that left- and right-choice predictive triplets from real data are more predictive than triplets obtained from the shuffled dataset where trial IDs were shuffled (left inset: left-predictive, two-tailed paired t-test,  $n=1135$  triplets, real vs shuffle: \*\*\*\* $p<0.0001$ ; right inset: right-predictive, two-tailed paired t-test,  $n=1755$  triplets, real vs shuffle: \*\*\*\* $p<0.0001$ ). **g**, Left-choice predictive triplets are significantly more predictive than instances where the first two cells in the triplet fire, but the third does not, or when the third cell fires alone (two-tailed paired t-tests, Bonferroni corrected,  $n=1135$  triplets,  $1\rightarrow 2\rightarrow 3$  vs  $1\rightarrow 2\rightarrow$ not 3: \*\*\*\* $p<0.0001$ ;  $1\rightarrow 2\rightarrow 3$  vs not  $1\rightarrow$ not 2 $\rightarrow$ 3: \*\*\*\* $p<0.0001$ ;

1→2→not 3 vs not 1→not 2→3:  $^{ns}p=0.78$ ). **h**, Importantly, for left-choice predictive triplets, in trials where cells 1 and 2 fire, but cell 3 does not, significantly more trials end with the animal turning right than the same instances in the shuffled dataset (right panel: two-tailed paired t-test,  $n=1135$  triplets, real vs shuffle: \*\*\*\* $p<0.0001$ ). **i**, Same as **g**, but for right-choice predictive triplets (two-tailed paired t-tests, Bonferroni corrected,  $n=1755$  triplets, 1→2→3 vs 1→2→not 3: \*\*\*\* $p<0.0001$ ; 1→2→3 vs not 1→not 2→3: \*\*\*\* $p<0.0001$ ; 1→2→not 3 vs not 1→not 2→3:  $^{ns}p=1.0$ ). **j**, Same as in **h**, but for right-choice predictive triplets (right panel: two-tailed paired t-test,  $n=1755$  triplets, real vs shuffle: \*\*\*\* $p<0.0001$ ). For boxplots, boundaries: 25th/75th percentiles, midline: median, whiskers: min/max.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements.

We thank A. Song and S. Thiberge for assistance with 2-photon imaging, S. Stein and S. Baptista for technical support with animal training, M. Ioffe for providing code, and E.M. Diamanti and B.E. Engelhard for discussions. This work was supported by the NIH grants U01NS090541, U19NS104648, and F32MH119749.

## Data availability

The datasets from this study are available from the corresponding authors on reasonable request.

## References

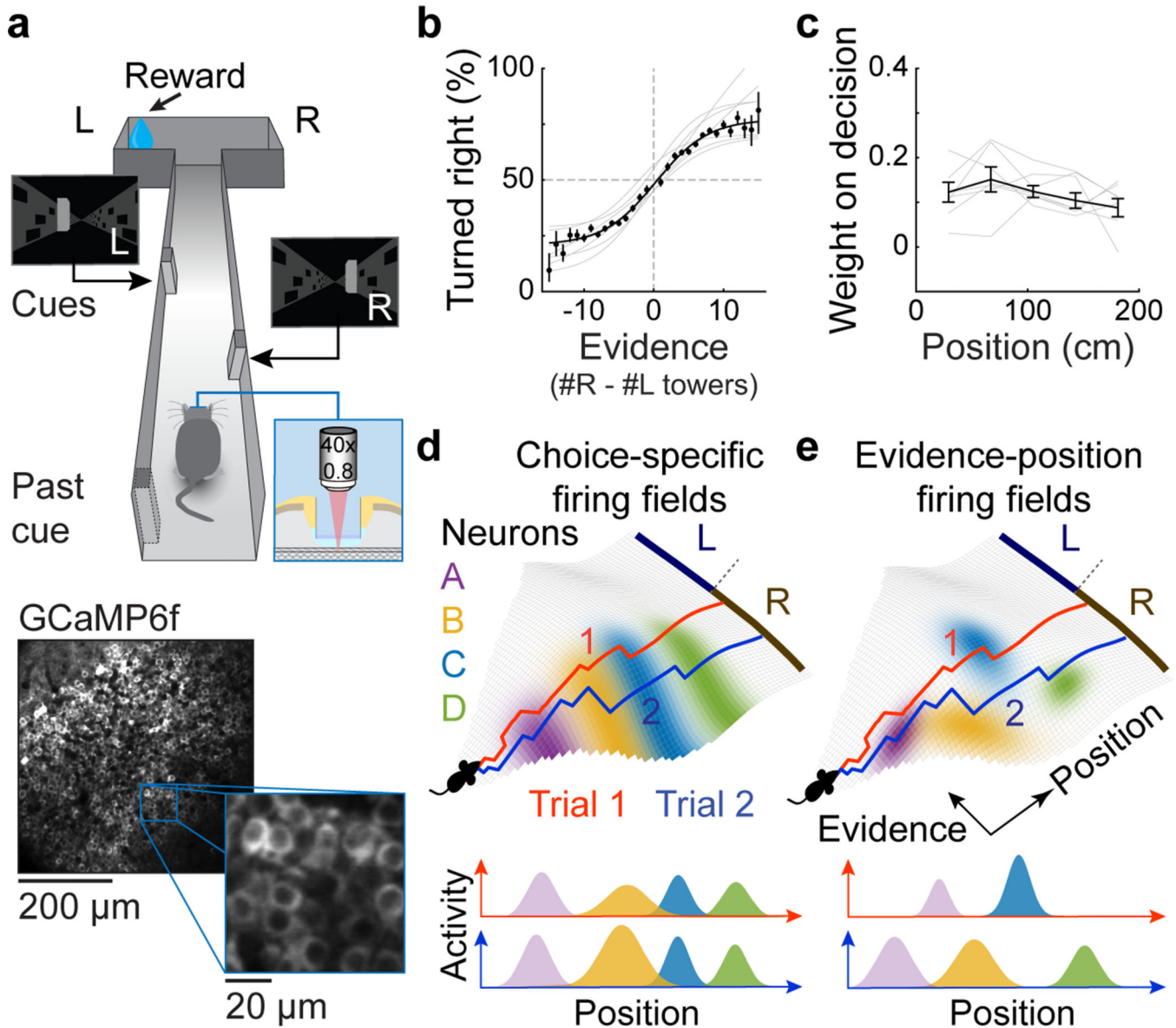
- O'Keefe J. & Dostrovsky J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34, 171–175 (1971). [PubMed: 5124915]
- Frank LM, Brown EN & Wilson M. Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron* 27, 169–178 (2000). [PubMed: 10939340]
- Wood ER, Dudchenko PA, Robitsek RJ & Eichenbaum H. Hippocampal Neurons Encode Information about Different Types of Memory Episodes Occurring in the Same Location. *Neuron* 27, 623–633 (2000). [PubMed: 11055443]
- Eichenbaum H, Kuperstein M, Fagan A. & Nagode J. Cue-sampling and goal-approach correlates of hippocampal unit activity in rats performing an odor-discrimination task. *J. Neurosci* 7, 716–732 (1987). [PubMed: 3559709]
- Herzog LE et al. Interaction of Taste and Place Coding in the Hippocampus. *J. Neurosci* 39, 3057–3069 (2019). [PubMed: 30777885]
- Aronov D, Nevers R. & Tank DW Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature* 543, 719–722 (2017). [PubMed: 28358077]
- Taxidis J. et al. Differential Emergence and Stability of Sensory and Temporal Representations in Context-Specific Hippocampal Sequences. *Neuron* 108, 984–998.e9 (2020). [PubMed: 32949502]
- O'Keefe J. & Nadel L. *The hippocampus as a cognitive map.* (Clarendon Press, 1978).
- Schuck NW & Niv Y. Sequential replay of nonspatial task states in the human hippocampus. *Science* 364, eaaw5181 (2019).
- Tavares RM et al. A Map for Social Navigation in the Human Brain. *Neuron* 87, 231–243 (2015). [PubMed: 26139376]

11. Park SA, Miller DS, Nili H, Ranganath C. & Boorman ED Map Making: Constructing, Combining, and Inferring on Abstract Cognitive Maps. *Neuron* 107, 1226–1238.e8 (2020). [PubMed: 32702288]
12. Rubin A. et al. Revealing neural correlates of behavior without behavioral measurements. *Nat. Commun* 10, 1–14 (2019). [PubMed: 30602773]
13. Low RJ, Lewallen S, Aronov D, Nevers R. & Tank DW Probing variability in a cognitive map using manifold inference from neural dynamics. *bioRxiv* 418939 (2018) doi:10.1101/418939.
14. Pinto L. et al. An Accumulation-of-Evidence Task Using Visual Pulses for Mice Navigating in Virtual Reality. *Front. Behav. Neurosci* 12, (2018).
15. Koay SA, Thiberge S, Brody CD & Tank DW Amplitude modulations of cortical sensory responses in pulsatile evidence accumulation. *eLife* 9, e60628 (2020). [PubMed: 33263278]
16. Engelhard B. et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature* 570, 509 (2019). [PubMed: 31142844]
17. MacDonald CJ, Lepage KQ, Eden UT & Eichenbaum H. Hippocampal “time cells” bridge the gap in memory for discontinuous events. *Neuron* 71, 737–749 (2011). [PubMed: 21867888]
18. Pastalkova E, Itskov V, Amarasingham A. & Buzsáki G. Internally Generated Cell Assembly Sequences in the Rat Hippocampus. *Science* 321, 1322–1327 (2008). [PubMed: 18772431]
19. Tolman EC Cognitive maps in rats and men. *Psychol. Rev* 55, 189–208 (1948). [PubMed: 18870876]
20. Bellmund JLS, Gärdenfors P, Moser EI & Doeller CF Navigating cognition: Spatial codes for human thinking. *Science* 362, eaat6766 (2018).
21. Eichenbaum H. What Versus Where: Non-spatial Aspects of Memory Representation by the Hippocampus. in *Behavioral Neuroscience of Learning and Memory* (eds. Clark RE & Martin SJ) 101–117 (Springer International Publishing, 2018). doi:10.1007/7854\_2016\_450.
22. Constantinescu AO, O’Reilly JX & Behrens TEJ Organizing Conceptual Knowledge in Humans with a Grid-like Code. *Science* 352, 1464–1468 (2016). [PubMed: 27313047]
23. Gallego JA, Perich MG, Miller LE & Solla SA Neural Manifolds for the Control of Movement. *Neuron* 94, 978–984 (2017). [PubMed: 28595054]
24. Russo AA et al. Motor Cortex Embeds Muscle-like Commands in an Untangled Population Response. *Neuron* 97, 953–966.e8 (2018). [PubMed: 29398358]
25. Chaudhuri R, Gerçek B, Pandey B, Peyrache A. & Fiete I. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nat. Neurosci* 22, 1512–1520 (2019). [PubMed: 31406365]
26. Eichenbaum H. & Cohen NJ Can We Reconcile the Declarative Memory and Spatial Navigation Views on Hippocampal Function? *Neuron* 83, 764–770 (2014). [PubMed: 25144874]
27. Recanatesi S. et al. Predictive learning as a network mechanism for extracting low-dimensional latent space representations. *Nat. Commun* 12, 1417 (2021). [PubMed: 33658520]
28. Aronov D. & Tank DW Engagement of neural circuits underlying 2D spatial navigation in a rodent virtual reality system. *Neuron* 84, 442–456 (2014). [PubMed: 25374363]
29. Dombeck DA, Harvey CD, Tian L, Looger LL & Tank DW Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat. Neurosci* 13, 1433–1440 (2010). [PubMed: 20890294]
30. Harvey CD, Coen P. & Tank DW Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* 484, 62–68 (2012). [PubMed: 22419153]
31. Brunton BW, Botvinick MM & Brody CD Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science* 340, 95–98 (2013). [PubMed: 23559254]
32. Gold JI & Shadlen MN The neural basis of decision making. *Annu. Rev. Neurosci* 30, 535–574 (2007). [PubMed: 17600525]
33. Gill PR, Mizumori SJY & Smith DM Hippocampal episode fields develop with learning. *Hippocampus* 21, 1240–1249 (2011). [PubMed: 20665593]
34. McKenzie S. et al. Hippocampal Representation of Related and Opposing Memories Develop within Distinct, Hierarchically Organized Neural Schemas. *Neuron* 83, 202–215 (2014). [PubMed: 24910078]

35. Howard MW, Luzardo A. & Tiganj Z. Evidence Accumulation in a Laplace Domain Decision Space. *Comput. Brain Behav* 1, 237–251 (2018). [PubMed: 31131363]
36. Pnevmatikakis EA et al. Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron* 89, 285–299 (2016). [PubMed: 26774160]
37. Pnevmatikakis EA & Giovannucci A. NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data. *J. Neurosci. Methods* 291, 83–94 (2017). [PubMed: 28782629]
38. Skaggs WE, McNaughton BL & Gothard KM An Information-Theoretic Approach to Deciphering the Hippocampal Code. in *Advances in Neural Information Processing Systems 5* (eds. Hanson SJ, Cowan JD & Giles CL) 1030–1037 (Morgan-Kaufmann, 1993).
39. Grassberger P. & Procaccia I. Measuring the strangeness of strange attractors. *Phys. Nonlinear Phenom* 9, 189–208 (1983).
40. Stachenfeld KL, Botvinick MM & Gershman SJ The hippocampus as a predictive map. *Nat. Neurosci* 20, 1643–1653 (2017). [PubMed: 28967910]

## Methods references

41. Gauthier JL & Tank DW A Dedicated Population for Reward Coding in the Hippocampus. *Neuron* 99, 179–193.e7 (2018). [PubMed: 30008297]
42. Pinto L. et al. Task-Dependent Changes in the Large-Scale Dynamics and Necessity of Cortical Regions. *Neuron* 104, 810–824.e9 (2019). [PubMed: 31564591]
43. Domnisoru C, Kinkhabwala AA & Tank DW Membrane potential dynamics of grid cells. *Nature* 495, 199–204 (2013). [PubMed: 23395984]
44. Rich PD, Liaw H-P & Lee AK Place cells. Large environments reveal the statistical structure governing hippocampal representations. *Science* 345, 814–817 (2014). [PubMed: 25124440]
45. Tenenbaum JB, Silva V. de & Langford JC A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* 290, 2319–2323 (2000). [PubMed: 11125149]
46. Yu BM et al. Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. *J. Neurophysiol* 102, 614–635 (2009). [PubMed: 19357332]



**Figure 1. Imaging CA1 neural activity in mice performing the accumulating towers task.**  
**a**, Schematic of the task in which head-fixed mice navigate in a virtual reality evidence accumulation T-maze task. Insets show example views from the animals' perspective (top). While animals (n=7) perform the task, 2-photon calcium imaging records hippocampal CA1 neural activity (bottom). **b**, Psychometric curves of mice performing the towers task (grey lines: n=7 animals, black line: metamouse combining data across animals; error bars: mean  $\pm$  binomial confidence interval). **c**, Logistic regression showing that mice use evidence throughout the cue period (grey lines: n=7 animals, black line: metamouse combining data across animals; error bars: mean  $\pm$  SEM). **d**, Firing fields of right-choice selective place cells would not depend on evidence and would thus divide a joint Evidence-by-Position (E $\times$ Y) space into two halves (top). Two right choice trials would generate the same neural sequence (bottom). **e**, Alternatively, if hippocampal neurons encoded evidence jointly with position, smaller firing fields dividing up evidence would appear in E $\times$ Y space (top), and two right

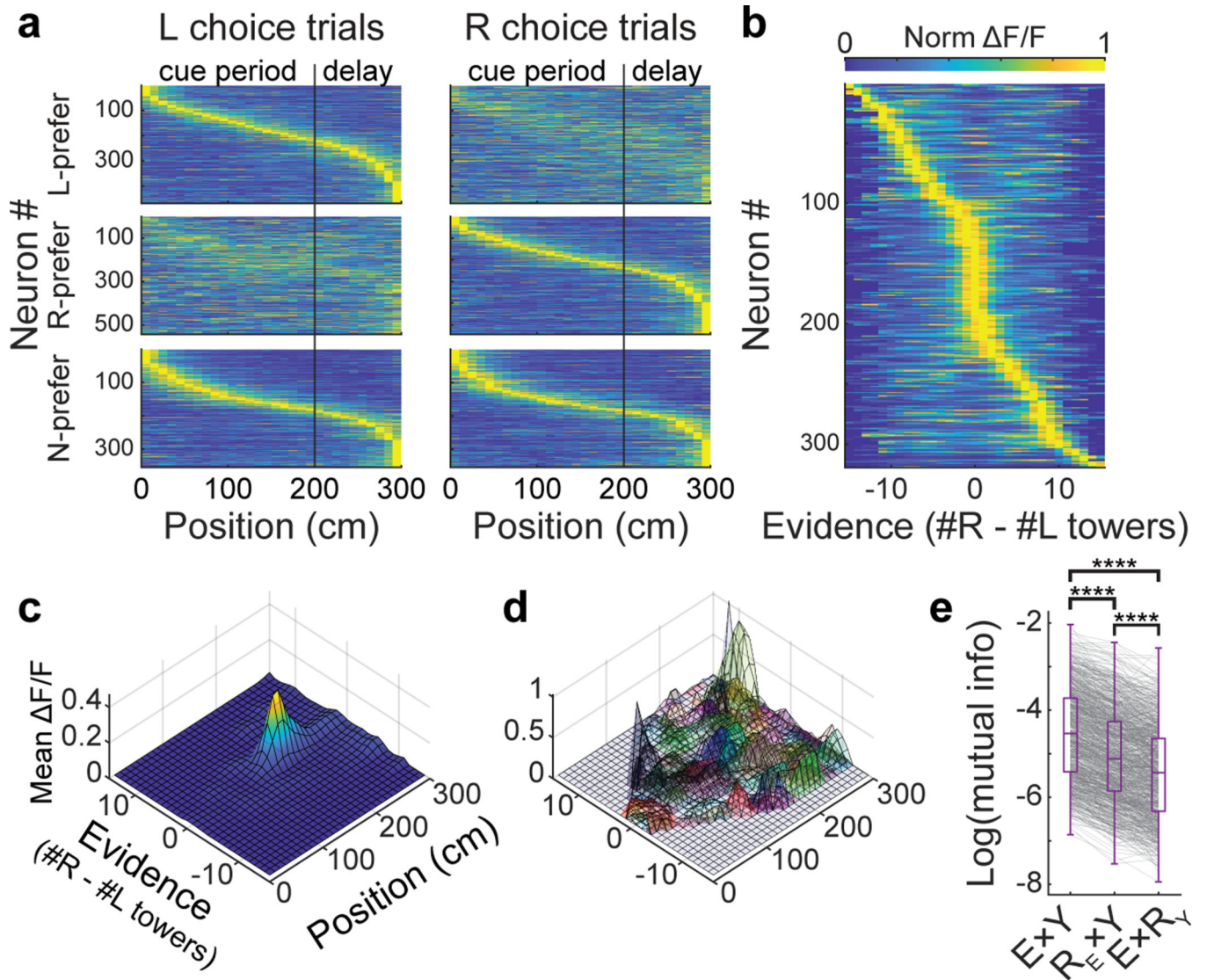
choice sequences could have different neural sequences depending on the evidence values traversed (bottom).

Author Manuscript

Author Manuscript

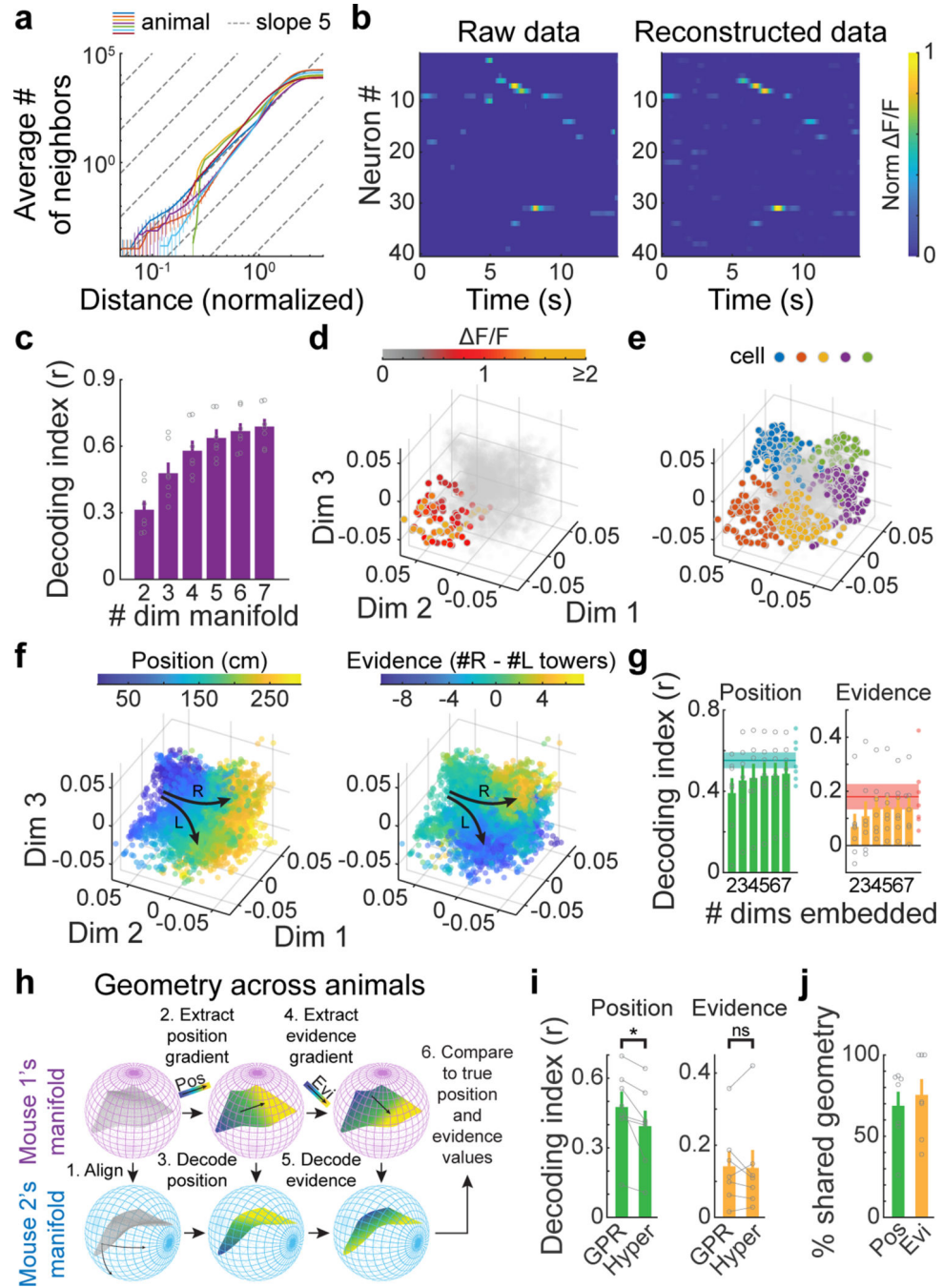
Author Manuscript

Author Manuscript



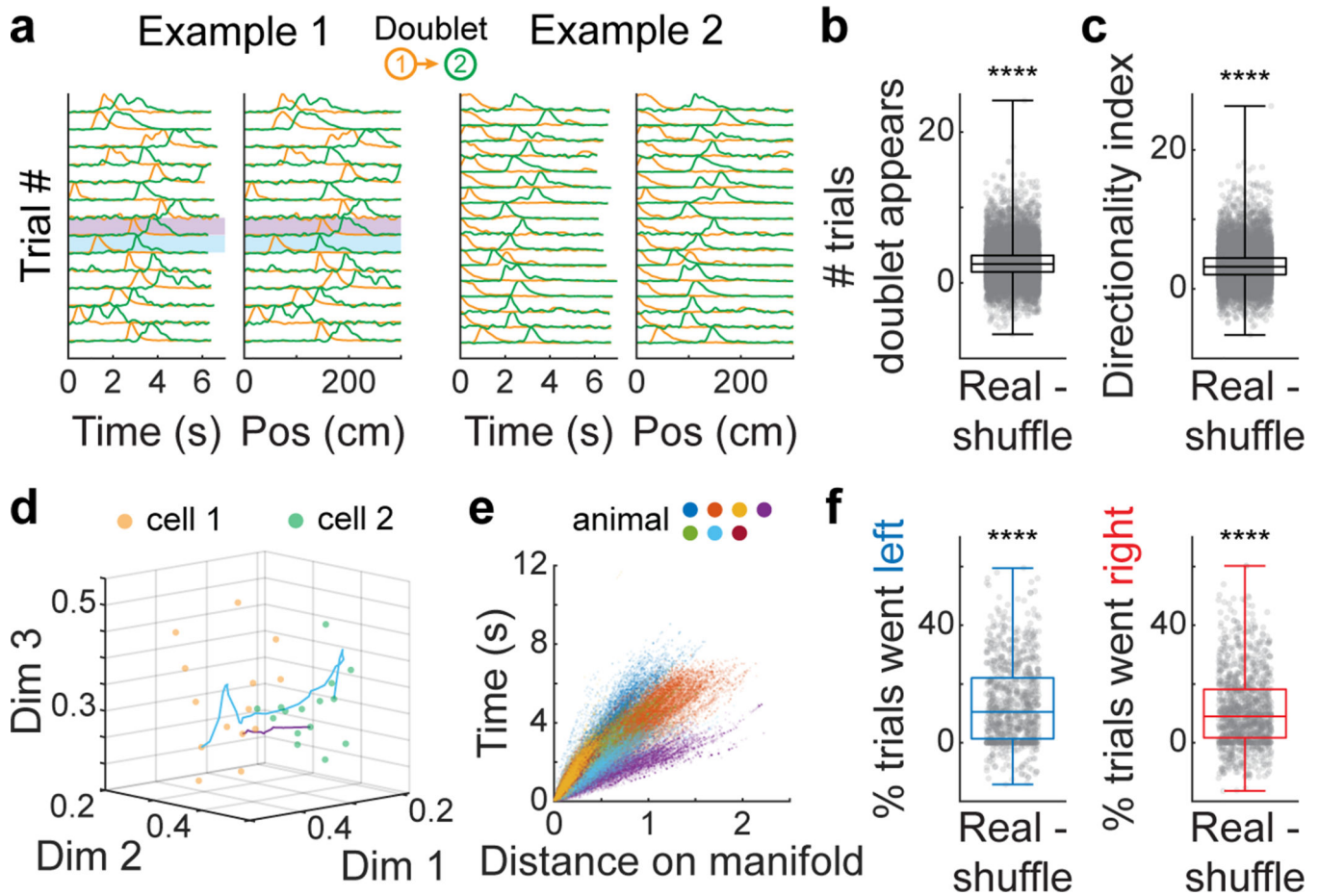
**Figure 2. CA1 neurons jointly encode position and evidence in an evidence accumulation task.** **a**, Choice-specific place cell sequences, divided into left-choice (top row), right-choice (middle row) and non- (bottom row) preferring cells. Cells are shown in the same order within each row group. F/F was normalized within each neuron. **b**, CA1 neurons have firing fields in accumulated evidence space ( $\#$  right towers -  $\#$  left towers). **c**, Example of the average neural activity of a single neuron in joint evidence-by-position ( $E \times Y$ ) space. **d**, Twenty-five neurons with significant information in  $E \times Y$  space. Each color represents one cell, and surfaces represent neural activity that exceeds 2 standard deviations (SD) above the shuffled means (Extended Data Fig. 2a). **e**, Mutual information of cells found to have significant information in  $E \times Y$  space is significantly greater than mutual information in 2D spaces where either evidence ( $R_E$ ) or position ( $R_Y$ ) has been randomized (two-tailed paired t-tests, Bonferroni corrected,  $n=917$  neurons,  $E \times Y$  vs  $R_E \times Y$ : \*\*\*\* $p < 0.0001$ ;  $E \times Y$  vs  $E \times R_Y$ : \*\*\*\* $p < 0.0001$ ;  $R_E \times Y$  vs  $E \times R_Y$ : \*\*\*\* $p < 0.0001$ ). For boxplots, boundaries: 25th/75th percentiles, midline: median, whiskers: min/max.





**Figure 3. Geometric representation of task variables on low-dimensional neural manifolds.**  
**a**, The average cumulative number of neighboring neural states as a function of the geodesic distance, plotted on a log-log axis, revealing a ~5.4 [4.8, 6.0] dimensional manifold ( $n=7$  animals, bracket values represent 95% bootstrapped confidence interval; error bars: mean $\pm$ 95% bootstrapped confidence intervals for each animal). **b**, Example of neural activity from 40 neurons (left) and the activity of those same 40 neurons reconstructed from the five latent variables obtained from embedding the manifold in a 5-dimensional Euclidean space (right).  $F/F$  is normalized to the maximum  $F/F$  in the window shown.

**c**, Reconstruction of held-out neural data from  $d$ -dimensional embeddings of the neural manifold. Decoding index is the correlation coefficient between the predicted and real F/F data in held-out trials. **d**, Each point in this plot is a location in the 3-dimensional embedding of the manifold at one time point in an imaging session. Colored points represent F/F values that are 3 standard deviations above the mean activity for one example cell. **e**, Firing field of five cells, each in a different color, plotted on the manifold. **f**, Position (left) and evidence (right) plotted as color on the 3-dimensional embedding of the manifold. Black arrows represent two hypothetical trajectories through manifold space that would traverse through position space and increasing left or right evidence values. **g**, Decoding position (left) and evidence (right) from  $d$ -dimensional embeddings of the manifold using gaussian process regression (GPR). Decoding index is the correlation coefficient between the predicted and true position or evidence values. The shaded area and line represent the mean decoding index  $\pm$ SEM using GPR on the top 10% of neurons with the highest mutual information for position or evidence as inputs. **h**, Schematic of the hyperalignment method for aligning two manifolds (see Methods). **i**, Decoding index of position and evidence for the hyperalignment, i.e. the best decoding that can be done using one of the six other manifolds, vs. decoding with GPR in the same animal for the 5-dimensional embedding of the manifolds (two-tailed Wilcoxon signed rank test,  $n=7$  animals, position:  $*p=0.016$ ; evidence:  $^{ns}p=0.81$ ). **j**, Percent of geometry shared across animals. The majority of manifold geometry ( $n=7$  animals, position:  $69\% \pm 9\%$  SEM; evidence:  $75\% \pm 10\%$  SEM) is shared between the best pairs of animals. In panels **c**, **g**, **i**, and **j**, error bars: mean $\pm$ SEM ( $n=7$  animals).



**Figure 4. Sequential activity of CA1 neurons in single trials is predictive of behavior and explained by the manifold.**

**a**, Two examples of doublets, where two neurons consistently fire one after the other. In example 1, activity does not appear to be tied to time (left) or position (right). Highlighted trials in cyan and purple are the same trials plotted in **d**. In example 2, activity in both neurons appears to be related to time/position in the trial. **b**, Doublets appear more frequently in real data than in a shuffled dataset (two-tailed paired t-test,  $n=16088$  doublets, real vs shuffle: \*\*\*\* $p<0.0001$ ). **c**, Doublets are asymmetric (two-tailed paired t-test,  $n=16088$  doublets, real vs shuffle: \*\*\*\* $p<0.0001$ ). Directionality index is defined as the number of times cell 1 fires before cell 2 in a trial minus the number of times cell 2 fires before cell 1 in a trial. **d**, Example showing how events from cell 1 (orange) and cell 2 (green) of a doublet are separated in manifold space. Cyan and purple lines each represent a trial trajectory between when cell 1 and cell 2 fire. **e**, Amount of time between when cell 1 and cell 2 fire plotted against distance in manifold space. **f**, Left- and right-choice predictive doublets (left and right panels, respectively) are significantly more predictive of upcoming choice than doublets generated from shuffled data where trial IDs were shuffled (left-predictive, two-tailed paired t-test,  $n=922$  doublets, real vs shuffle: \*\*\*\* $p<0.0001$ ; right-predictive, two-tailed paired t-test,  $n=1227$  doublets, real vs shuffle: \*\*\*\* $p<0.0001$ ). For boxplots, boundaries: 25th/75th percentiles, midline: median, whiskers: min/max.