Neural circuit models for evidence accumulation through choice-selective sequences

Lindsey S. Brown[1], Jounhong Ryan Cho[1], Scott S. Bolkan[1], Edward H. Nieh[1], Manuel Schottdorf[1], David W. Tank[1], Carlos D. Brody[1], Ilana B. Witten[1,*], Mark S. Goldman[2,3,*]

[1]Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA; [2]Center for Neuroscience, Department of Neurobiology, Physiology and Behavior, University of California, Davis, Davis, CA, USA; [3]Department of Ophthalmology and Vision Science, University of California, Davis, Davis, CA, USA

*Address correspondence to msgoldman@ucdavis.edu, iwitten@princeton.edu

ABSTRACT

Decision making is traditionally thought to be mediated by populations of neurons whose firing rates persistently accumulate evidence across time. However, recent decision-making experiments in rodents have observed neurons across the brain that fire sequentially as a function of spatial position or time, rather than persistently, with the subset of neurons in the sequence depending on the animal's choice. We develop two new candidate circuit models, in which evidence is encoded either in the relative firing rates of two competing chains of neurons or in the network location of a stereotyped pattern ("bump") of neural activity. Encoded evidence is then faithfully transferred between neuronal populations representing different positions or times. Neural recordings from four different brain regions during a decision-making task showed that, during the evidence accumulation period, different brain regions displayed tuning curves consistent with different candidate models for evidence accumulation. This work provides mechanistic models and potential neural substrates for how graded-value information may be precisely accumulated within and transferred between neural populations, a set of computations fundamental to many cognitive operations.

INTRODUCTION

The accumulation of evidence for different alternatives is thought to be a fundamental cognitive operation. As such, neural recordings have been performed in many tasks designed to probe how evidence is accumulated to make a decision, including tactile[1,2], visual[3–11], olfactory[12], auditory[13,14], motor[15], and value discrimination tasks[16–20]. Experiments using these tasks have been conducted with the goal of understanding how the neural circuitry within and across different brain regions contributes to the accumulation of evidence, subsequent maintenance of this information in working memory, and ultimate commitment to a decision based on this evidence [21].

Many models have been proposed to describe the decision making process. A major class of models for two-choice forced alternative decisions is the drift diffusion model[22–24], in which evidence is accumulated with a noisy drift process, and decisions are made based on when the resulting accumulation reaches some threshold. The drift-diffusion model reaches the fastest decision for any given desired accuracy, making this model optimal. Many circuit-based models have been shown to be equivalent to the drift-diffusion model[25]. Such circuit-based models include mutual inhibition models[26], feedforward inhibition models[27], and pooled inhibition models[28]. These basic circuit models have been extended to explain a number of different decision-making experiments with recordings in different regions where neurons show persistent activity[29–34].

The mechanistic essence of the circuit-based models is the accumulation of evidence-encoding signals along a low-dimensional attractor representing the decision variable. Thus, these models predict that the activity of single neurons ramps up or down with evidence throughout a decision-making task[35,36]. Such ramping activity during decision-making tasks has been observed in a number of different brain regions in different paradigms, including in the lateral intraparietal area (LIP), medial temporal area (MT), and ventral intraparietal area (VIP) of monkeys in a random dot motion discrimination task[37–40], in the dorsal premotor cortex of monkeys during discrimination tasks[41,42], and in the posterior parietal cortex (PPC) and frontal orienting fields (FOF) of rats in an auditory evidence accumulation task[14,43]. These classes of neural circuit models that accumulate evidence through persistent, ramping activity successfully capture much observed experimental data.

However, a mounting body of evidence from large scale recordings during a range of decision-making tasks suggests that, in many decision-making contexts, neurons across cortical and subcortical regions are not persistently active. Instead, neurons fire transiently and sequentially in a choice-specific manner. This occurs in decision-making tasks both with navigation[44–49] and without navigation[50–56], including in evidence accumulation tasks[44,49]. These data challenge existing models of evidence accumulation by contradicting a core tenet of previous models: the presence of persistent, rather than sequential, neural activity. For a model to accumulate evidence through sequential neural firing, it must perform two key computations. First, as in traditional models based on persistent activity, it must accumulate evidence received at each position. Second, unlike traditional models, it must transfer information between neurons at sequential positions.

Here, we develop two classes of neural circuit models that perform these computations. In each model class, the transfer of accumulated evidence between neurons encoding sequential positions is accomplished via a position-modulated gating signal. The key difference

in the models is in their encoding of evidence, which is either in the relative *amplitude* of activity in two competing chains of neurons or in the network *location* of a stereotypically shaped, unimodal pattern ('bump') of neuronal firing. We tested these differentiating evidence tuning predictions in a dataset of 14,247 imaged neurons, collected across 4 cortical and subcortical regions in mice performing an evidence accumulation task in virtual reality. This analysis revealed that while seemingly similar choice-selective sequences are present in this task across these regions, the neurons within the sequences differ in their evidence tuning properties. In the neocortex, neuronal tuning curves mainly exhibit a monotonically increasing or decreasing encoding of evidence, consistent with the competing chains model class, whereas in the hippocampus neurons tend to have narrow, non-monotonic tuning curves, consistent with the bump model. Thus, our results reveal different encodings of accumulated evidence across brain regions and suggest that different regions could use distinct circuit mechanisms to form their evidence representations.

RESULTS

Traditional Integrator Models Fail to Explain Choice-Selective Sequences in the Accumulating Towers Task

The models presented below are motivated by a large dataset (n= 14,247 neurons; N=26 mice) of newly acquired calcium imaging data (see Methods) from anterior cingulate cortex (ACC) and dorsomedial striatum (DMS), as well as recently published imaging data from hippocampus (HPC)[49] and retrosplenial cortex (RSC)[44]. These data were all obtained during the same navigation-based, accumulation of evidence task in virtual reality ('accumulating towers task', Fig. 1A). In this task ([57]; see Methods), mice navigate a T-maze with Poisson counts of visual cues in the form of towers appearing to each side. After a delay region with no cues, the mouse reaches the arms of the maze and is rewarded for turning to the side with more cues (Fig. 1A). To make the correct decision, the mouse must integrate the number of cues and hold this information in working memory during the delay region. Mice are able to perform this task with good accuracy that increases with the magnitude of the difference between the number of cues to each side[57].

Across brain regions, neurons fire sequentially rather than persistently (Fig. 1B-E; Fig. S1). Such sequences are present throughout the task, beginning in the cue region and continuing through the delay region. Moreover, many of the neurons in these sequences are choice-selective, such that neurons preferentially fire depending upon whether the animal will ultimately choose to turn left or right (Fig. 1B-E).

Previous models of evidence accumulation do not produce the observed choice-selective sequences. This includes traditional circuit models that neurally instantiate the drift diffusion model (Fig. 2A-B), as well as bump-based integrators (Fig. 1C-D) commonly used to accumulate head velocity signals into a representation of head direction. Further, canonical sequence models such as synfire chains do not accumulate evidence [58–60] or accumulate external signals but are not choice-selective[61,62]. This motivates our construction of new, sequence-based evidence-accumulation models.

Chains-Based Models

We first present a class of models in which two chains of neurons represent the current position of the animal in the location along the chains and represent accumulated evidence in the relative amplitudes of their firing activity. Here, we consider a model with mutual inhibition between neurons at the same position in the chains (Fig. 3A), but other chains-based models are possible (see Supplementary Text for a model with uncoupled, rather than mutually inhibiting chains). The temporal dynamics of the firing rate $r_{i,L}$ of the ith neuron in the left (L) side chain are given by the equation

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [br_{i,L} + cr_{i-1,L} - er_{i,R} + f\mathbf{1}_{\text{left}}(t) + P_i(t) - T]^+ \quad (1)$$

and an analogous equation describes the firing rates $r_{i,R}$ of neurons in the right (R) side chain. Here, a is the exponential decay rate of activity in the absence of input, b is the weight of self-excitation (Fig. 3A, green self-loops), c is the strength of feedforward synaptic connections from the previous neuron (green feedforward connections), and e is the strength of the inhibitory connection from the neuron at the same position in the opposite chain (brown connections). The external inputs to be integrated by the network consist of brief pulses of the form $\mathbf{1}_{left}(t)$, which is 1 if there is a left cue within 0.5 cm of the current position, and project to the chain with synaptic strength f. $P_i(t)$ is the position-gating signal, and T is the firing threshold. The position-gating signal $P_i(t)$ controls the progression of activity along the chain by raising the neuron's input above the firing threshold T, and is zero for all neurons except the pair of neurons (one from each chain) corresponding to the position of the animal in the environment. The firing rate of neurons not at the active position decay exponentially with time constant $1/a$. See Methods for complete details and parameterization of the model.

At a given position, this model has the same structure as the traditional mutual inhibition model[26], and this motif is repeated across different positions to form two chains. This motif accumulates and maintains evidence through positive feedback within and between the two neural populations receiving the active position gating signal. Competition is mediated by mutually inhibitory positive feedback between the chains, such that increases of activity in one chain results in corresponding decreases of activity in the other chain. We note that the accumulation process saturates at an upper bound when neurons in the non-dominant chain reach zero firing. This is because, when one chain is at zero firing rate, the disinhibitory component of positive feedback between the chains is lost, so that further increases in evidence cannot be maintained.

Only the neurons receiving the position gating signal actively represent and accumulate evidence. When the animal moves to a new position, the position gating signal activates the next pair of neurons in the chain, and evidence is transferred to the newly activated pair of neurons by the feedforward synaptic connections. Repeating this process as the animal navigates down the track results in a sequence of cells firing, with each successive pair of activated neurons representing the accumulated evidence in the difference of their firing rates.

We simulated the model on the accumulating towers task (see Methods), demonstrating that this model performs the task perfectly in the absence of noise for appropriately tuned parameters (Fig. 3B). Previous analysis of the accumulating towers task and other evidence accumulation tasks suggest that noise in an animal's performance is dominated by sensory noise, so that only a fraction of input cues are integrated[13,57]. When we correspondingly

introduce input noise into our model (see Methods), the performance decreases in a manner similar to that observed in animals, with a more gradual increase in performance with increasing magnitude of the difference between cues (Fig. 3B). The neurons in the model exhibit choice-selective (or, more precisely, 'accumulated-evidence-selective') sequential activity, with neurons in each chain showing greater activity for the side of the evidence that it receives (Fig. 3C). Thus, accumulated evidence is encoded monotonically (up to saturation) in the relative firing rates of the two chains in this class of models (Fig. 3D).

Other architectures of chains-based models are possible, including a model where competition between opposing inputs occurs without inhibition between the chains (Fig. S2A). At a given position, this model has a 'push-pull' feedforward arrangement of evidence inputs, with cues favoring one side exciting one chain and inhibiting the opposing chain, as in traditional decision-making models based on uncoupled competing accumulators [27]. This motif is then repeated to form two chains. As with the model based on mutually inhibiting chains of neurons, this model is able to perform the task accurately (Fig. S2B) and produce choice-selective sequences (Fig. S2C) in which individual neurons have monotonic tuning curves to evidence (Fig. S2D).

Models with different architectures instantiate different hypotheses that could be distinguished by optogenetic perturbations (Fig. 3E, Fig. S2E). To demonstrate these differing predictions, we simulate the delivery of an excitatory input to a neuron (or similarly tuned neuronal population) while the animal is at the neuron's active position (Fig. 3E). We do this for a trial with no external cues so that the accumulated evidence prior to perturbation is zero. Due to the gating of neural activity by the forward traveling position signal, both models predict that only cells tuned to later positions will be affected by the optogenetic perturbation. For the mutually inhibitory chains model, subsequent neurons in the same chain will have increased activity due to this feedforward excitation while subsequent neurons in the opposing chain will have decreased activity due to the mutual inhibition between the chains (Fig. 3F). By contrast, for the models based on uncoupled chains in which the anti-correlated activity of the two chains is solely due to opposing external cue inputs (Fig. S2), an excitatory perturbation in one chain increases firing only within the same chain but has no impact on the other chain (Fig. S2F).

Another difference between the mutually inhibitory and uncoupled chains models is that the model based on mutual inhibition is more robust to the tuning of the position gating. Specifically, the mutual inhibition model requires only that P_i be greater than or equal to T at the active position, whereas the uncoupled chains model requires that P_i be exactly equal to T to avoid continual growth of neural activity from temporally integrating the position-gating signal (Supplementary Text; [63]).

Taken together, these chains-based models represent a class of models in which position-gating gives rise to choice-selective sequences that encode graded evidence signals monotonically in the neuronal firing rates at each position.

Position-Gated Bump Attractors

We next consider a class of models in which evidence is encoded in the location of a stereotypical, unimodal pattern ("bump") of neural activity in the population. Such bump attractor models have been used to describe the neural circuitry that computes head direction in the rodent[64–69] and the fly[70–76], where the location of the bump corresponds to the animal's heading.

Similarly, two-dimensional bump attractor models have been proposed to describe path integration in the hippocampal place cell system[77–79]. In these models, motion is temporally integrated to determine head direction or to perform path integration, and analogously our models temporally integrate evidence from the visual cues. A single dimensional bump model that only accumulates evidence does not produce sequences like those observed in the neural data because the bump would recur at the same location whenever the accumulated evidence is at the same level and would remain at a fixed location throughout the delay period when the evidence level does not change (Fig. 2D). Thus, we modify these traditional bump attractor models to have separate dimensions for encoding position and evidence (Fig. 4).

The model consists of layers of recurrently connected neurons, with each layer corresponding to a given position and neurons within the layers representing the accumulated evidence at this position. The activity $r_{i,j}$ of a neuron at position i and evidence level j evolves according to

$$\frac{dr_{i,j}}{dt} = -ar_{i,j} + \Theta(P_i(t) - T)\left[F\left(\sum_k W_{j,k} r_{i,k} + br_{i-1,j} + c\left(I_{i,j+1,L}(t) + I_{i,j-1,R}(t)\right)\right)\right] \quad (2)$$

where a is the exponential decay rate of activity in the absence of input, Θ is the Heaviside function that implements the position gating and equals 1 if the position gating signal $P_i(t)$ is greater than or equal to the value of the threshold T and 0 otherwise, and F is the neuronal output nonlinearity. Within a layer, $W_{j,k}$ is the strength of the connection from the neuron representing evidence level k onto the neuron encoding evidence level j. $W_{j,k}$ is excitatory for connections of a neuron onto itself and its immediate neighbors (Fig. 4A, green recurrent connections) and inhibitory from all other neurons in that layer (Fig. 4A, brown recurrent connections). b is the synaptic strength of feedforward excitatory connections between layers of neurons encoding neighboring positions, and we assume for simplicity that these connections are only between neurons encoding the same evidence level (Fig. 4A, green feedforward connections). $I_{i,j+1,L}$ and $I_{i,j-1,R}$ are the firing rates of the 'shifter neurons' (Fig. 4A, orange and purple circles) that shift the bump left or right in response to external cue inputs (Fig. 4A, left and right tower inputs). The left shift (orange) neuron is activated when jointly receiving input from evidence accumulating neuron $j+1$ and a left-side external cue. In this manner, if the peak of the bump of neuronal activity was previously at $j+1$, this peak activity then (through a synaptic connection of strength c from the shifter neuron) drives the neuron at evidence level j. Mathematically, the firing rates of the left, and corresponding right, shifter neurons are given by

$$I_{i,j,L} = r_{i,j}\mathbf{1}_{left}(t) \quad (3)$$

$$I_{i,j,R} = r_{i,j}\mathbf{1}_{right}(t) \quad (4),$$

where $\mathbf{1}_{left/right}(t)$ equals 1 if there is a left or right cue respectively within 0.5 cm of the current position. For the complete model specification, see Methods.

Within the active position, accumulated evidence is represented and maintained in the location of a bump of activity along the evidence axis. As in traditional bump attractor models[80,81], the bump is maintained in the absence of input by local symmetric excitation while long range inhibition prevents the bump from diffusing. Asymmetric input connections cause the bump to shift left or right in response to a left or right cue respectively[65,69], as described above. This shift of the localized bump of activity occurs despite the external cue signal being global (Fig. 4A, tower inputs). Mechanistically, the ability to shift a localized bump of activity with a

global external signal is achieved via the multiplicative gating of the external input by the activity of the evidence-encoding neurons (equations (3) and (4)) so that the global external input signal only activates the shifter neurons close to the location of the current bump. Alternatively, this gating could be accomplished with additive inputs plus thresholds[76,80].

As the animal progresses forward in the maze, the active position changes so that information must be transferred to the next layer of cells, requiring both a change in the position gate and feedforward connections between layers. Here, position gating is controlled by multiplication by the Heaviside function (in contrast to the additive threshold in the chains based models), with activity decaying exponentially away from the active position with time constant *1/a.* Feedforward excitation between the layers then transfers information to neurons at the next position. We have modeled these feedforward connections as only projecting between neurons of the same evidence level, but this projection could be broader as long as the feedforward input is symmetric and centered around the current evidence level.

We simulate this model for the same set of trials as the chains-based models. As in that class of models, for appropriately tuned parameters, the position-gated bump attractor also performs the task perfectly in the absence of noise, and performance worsens when input noise is introduced (Fig. 4B). The simulated neural activity also produces choice-selective sequences (Fig. 4C). As seen by the downward curvature of the sequences in Figure 4C, the sequences are marked by the feature that more neurons respond at late positions than respond at early positions in the sequence. This is because the random walk of the bump location as evidence is accumulated leads to a greater range of evidence levels being observed at later positions in the trial; thus, on average across trials, there are more different neurons active at later positions, with each neuron corresponding to a different possible evidence level.

The most straightforward difference between the position-gated bump attractor and the chains-based models is in the shape of their evidence tuning curves. Due to the structure of the network connectivity, the neurons in the position-gated bump attractor are non-monotonically tuned to specific evidence levels at their active position (Fig. 4D), in contrast to the monotonic tuning curves of the chains-based models (Fig. 3D).

To further differentiate the position-gated bump attractor model from the chains-based models, we consider predictions for the effects of optogenetic excitation of a single cell during a trial with no cues presented (Fig. 4E). Here, we consider the regime where the stimulus is sufficiently large to cause the bump to move to the stimulated location. Due to the similar feedforward connectivity patterns between positions in both model classes, stimulation of a cell active at one position only affects the activity of cells tuned to later positions, as in the chains-based model. However, unlike the chains-based models, we see that rather than impacting every neuron at a later position equally, the effects are sparse and differ in magnitude across neurons (Fig. 4F). These differences are because the stimulation has caused the bump to move to a new evidence level, which is then sustained at subsequent positions. Thus, neurons with a similar evidence tuning to the cell that was stimulated increase in activity, while those tuned to where the bump of activity would be located in the absence of stimulation (at 0 evidence) decrease their activity since the bump has moved to the new location (Fig. 4G). We note that for smaller excitatory stimuli that do not shift the bump center to the site of stimulation, and dependent upon the exact details of excitatory and inhibitory connections, smaller shifts in the bump either toward or away from the stimulated location are possible[82–85], but the

perturbations still maintain the properties that changes are sparse and that changes within the same choice-preferring population as the stimulated neuron may be bidirectional.

Neural Activity in Different Brain Regions Correlates with Different Classes of Models

Above, we presented two novel circuit models for accumulating evidence through sequences (competing chains model: Fig. 3; position-gated bump attractor model: Fig. 4). While both models produce neurons that are sequentially tuned to position, the models predict different evidence coding schemes. The chains-based models make two core predictions: first, evidence should be monotonically encoded and, second, this monotonic encoding should be broadly graded across a wide range of evidence levels. The bump attractor model likewise makes two core predictions: first, evidence should be encoded with relatively narrow, unimodal tuning and, second, the peaks of these tuning curves should be broadly distributed across the entire evidence range. We looked for signatures of these different evidence encoding schemes in recordings from ACC, RSC, DMS and HPC during the accumulating towers task (Fig. 1A).

Our data analysis proceeded in three stages. First, we identified neurons that were tuned to evidence, and extracted the peak position tuning of the neurons (Fig. S3B). Second, we fit the evidence tuning of neurons around the location of their peak position tuning (Fig. S3C). This enabled us to both identify whether neurons were more monotonic-like or unimodal-like in their tuning and also, for unimodal tuning, to identify whether the peaks of the tuning curves were broadly distributed across the range of evidence levels. Third, as presented in the following section, we assessed whether the cells with monotonic evidence tuning curves had broadly graded tuning to evidence, consistent with evidence encoding, versus more binary-like tuning that would be consistent with a 'binary choice-like' representation. To do this, we both examined the width of the sigmoidal tuning to evidence within and pooled across the population for neurons tuned to different positions along the maze, and also fit an explicit linear model of each neuron's firing to a combination of evidence and a binary choice variable.

Assessing a neuron's tuning to evidence requires jointly fitting position and evidence to account for the correlated, nonuniform sampling of these two variables (Fig. S3A-B). For example, cells tuned to the beginning of the maze are only active when a small range of evidence is presented. Thus, if such neurons were directly fit to evidence tuning alone, rather than jointly fit to evidence and position, they would appear to be tuned to small absolute evidence levels even if they had no tuning to evidence at all. Likewise, because the evidence levels sampled at later parts of the maze were most concentrated around moderate-to-large absolute values (Fig. S3A), cells untuned to evidence but tuned to later positions in the maze would artifactually appear to have a bimodal tuning to evidence. To obtain fits to our tuning curve data that could capture many different tuning curve shapes across the population, we modeled the joint tuning for position and evidence as a product of a Gaussian function of position and a Gaussian function of evidence (Fig. S3B, Fig. 5C,D,G,H,K,L, Fig. S4C,D; see Methods). The Gaussian in position intuitively captures the non-monotonic, sequential position tuning. The Gaussian in evidence, although non-monotonic in its mathematical form, allowed us to identify not only non-monotonic but also monotonic evidence tuning, because in the latter case the best fit Gaussian has a peak near the extreme of the evidence range (i.e., normalized peaks near +/- 1; Fig. S3D). Broad monotonic ramps like those of the chains-based models also have a large standard deviation relative to the observed evidence range. To visualize responses

at a given position in the maze, we also fit responses around the peak position tuning of each neuron to both a Gaussian or a logistic function (Fig. S3C) and plotted the best fit function (Fig. 5B,F,J, Fig. S4B).

Jointly fitting position and evidence tuning across the 14,247 neurons revealed that 56% of all neurons in ACC, 49% in RSC, 25% in HPC, and 40% in DMS had significant evidence tuning (statistical comparison to pseudosessions; see Methods). In analyzing cells with significant evidence tuning in each region, we saw distinguishing patterns emerge in the nature of the evidence tuning (Fig. 5, Fig. S4).

In ACC, many individual cells exhibited primarily monotonic and relatively broad evidence tuning (Fig. 5A-B, Fig. S6). Similar to these example neurons, among the population of significantly evidence-tuned cells, peak evidence tuning was towards the extreme of the observed evidence range, consistent with monotonic tuning (Fig. 5C-D). Moreover, the standard deviation of the evidence fits was relatively large (Fig. 5C-D; standard deviation normalized based on the evidence range observed at the peak position; note warm colors in Fig. 5C). The presence of monotonic and relatively broad evidence tuning is most consistent with predictions of the chains-based models (Fig. 3D, Fig. S3E-F), although many neurons were more sigmoidally (rather than linearly) tuned to evidence compared to the models (Fig. 5B). The degree of such sigmoidal tuning, and its variation across different positions of the maze, is treated in the following section.

In contrast, HPC cells showed primarily non-monotonic and narrow evidence tuning (Fig. 5E-F, Fig. S7), which tiled the range of observed evidence (Fig. 5F-H). These narrow tuning curves occurred throughout the cue and delay region (dark blue cells in Fig. 5G). The presence of non-monotonic, narrow tuning curves that tile evidence space are consistent with predictions of the bump model (Fig. 4D, Fig. S3G-H).

In RSC, the majority of cells showed monotonic patterns (Fig. 5I-J, Fig. S8) like those observed in ACC, while a smaller population showed more narrow tuning curves like, but typically somewhat wider than, those observed in HPC (example 2 in Fig. 5I, examples 3, 9, 16, 22, 28, and 34 in Fig. S8). These observations are seen in the population summary of evidence and position tuning (Fig. 5K-L), where many cells show both wide evidence tuning with peaks at the extreme (deep red) and narrow intermediate evidence tuning (dark blue). RSC also has a particularly large population of cells that, when fit to a Gaussian, have intermediate widths and evidence peaks not fully at the extremes (Fig. 5K, green). Such cells typically exhibit a narrow sigmoidal tuning curve such as that shown in example 3.

DMS was dominated by intermediate-width monotonic evidence tuning, but with some narrowly tuned cells (Fig. S4, Fig. S9). However, DMS contained relatively few evidence-modulated cells during the cue region, and instead most evidence-modulated cells had position tuning in the delay region of the maze and narrow sigmoidal tuning curves (Fig. S4C,E,F).

Progression in Relative Amount of Graded Evidence Versus Binary Choice Tuning Across the Maze in ACC and RSC

While the monotonicity of the evidence tuning observed in ACC and RSC is consistent with the chains-based models, the models not only predict monotonic tuning, but also require graded tuning to accumulate evidence within the cue region (Fig. 3D). Monotonically tuned

neurons that are instead sensitive to the sign but not the amount of accumulated evidence do not directly support our model, but may instead represent binary choice neurons found downstream of the accumulator (Fig. 6A, schematic 'choice-tuned' cell). Thus, we examined whether the regions dominated by monotonic tuning for evidence contained graded representations of evidence versus binary choice coding, first by analyzing population-averaged responses, and next by using a linear encoding model of activity based on evidence and choice (Fig. 6B).

The population-averaged position-evidence responses in both ACC and RSC show graded increases in activity, as in the chains-based models, during the early portion of the cue period (Fig. 6C,E, dark blue traces; population responses based on averaging all evidence-modulated neurons, see Methods). With progression along the maze, this graded evidence tuning sharpens, transitioning to a step function by the start of the delay period (magenta and purple traces) that resembles choice rather than graded evidence tuning. In ACC (and more subtly in RSC), population responses also increase later in the maze, due to the fact that there are more evidence-modulated neurons in the delay region (Fig. S10A,B).

To further dissociate graded evidence versus binary choice representations, we fit linear encoding models to individual cells at each position bin, using evidence and the current best choice (i.e., the sign of the evidence) as regressors (see Methods). At each position, we calculated the fraction of active neurons (see Methods) whose firing rate was significantly modulated by the linear evidence or the binary choice variable, while accounting for the contribution to neural activity of the other variable. In both ACC and RSC, we observed neurons with significant linear encoding of evidence beyond that explained by choice during the cue, but not during the delay period. The fraction of such neurons was relatively level during the cue period (crimson traces, Fig. 6D,F). By contrast, the fraction of significant choice encoding neurons (beyond that explained by evidence) increases throughout the maze, with large fractions of active neurons encoding choice in the delay period (lime traces, Fig. 6D,F). Together, this implies that linear evidence encoding in the cue region transitions to binary choice encoding in the delay period, similar to our findings from the population-averaged responses (Fig. 6C,E).

DISCUSSION

This work proposes a mechanistic model for the widespread observation of choice-selective sequences during evidence accumulation (Fig. 1). Since traditional models of evidence accumulation based on persistent activity fail to explain this observation (Fig. 2), we present two new classes of circuit models: a "competing chains" model (Fig. 3) and a "position-gated bump attractor" model (Fig. 4). These two model classes represent two different evidence accumulation schemes, in which evidence is encoded either in the amplitude of the neuronal firing rates or in the set of neurons that fire with stereotyped amplitude, and predict different effects of targeted perturbations. We tested the evidence tuning predictions in imaging data across multiple cortical and subcortical brain regions, demonstrating that although different regions show seemingly similar choice-selective sequences, their underlying evidence representations are distinct (Fig. 5, Fig. 6). Specifically, cortical regions (ACC and RSC) had primarily monotonic evidence tuning curves in the early cue region, consistent with the chains-based models, while the HPC showed non-monotonic evidence tuning with broadly

distributed tuning curve peaks throughout the maze, consistent with the position-gated bump attractor.

Transfer of Accumulated Evidence Across Positions is a Fundamental Cognitive Operation

Although our models were tested on data from a specific task, at the essence of accumulating evidence through sequences is the ability to both integrate evidence and transfer information between different populations of neurons. These two operations apply quite generally to many decision-making contexts. For example, position in our models should be interpreted broadly as a sequentially progressing state of the animal (e.g., it could represent time), so these models apply to navigation or non-navigation tasks. Moreover, perceptual decision making tasks may require transferring accumulated evidence across neuronal populations when the population that needs this information to carry out an action changes (e.g., changing a focal point while maintaining evidence for two saccadic choice targets[86]). While models for evidence integration have been explored extensively and form the core of our models' computations within a position, traditional models do not transfer graded information between positions, a novel feature in our models.

To achieve the transfer of information between positions, our models use two different gating mechanisms, an additive threshold demonstrated in the chains-based models and a multiplicative threshold demonstrated in the position-gated bump attractor, although either mechanism could be used in either model. Thus, we illustrate two possible mechanisms for controlling information flow between positions. While not explicitly modeled here, such spatio-temporal gating could result from inputs from the basal ganglia[87–90] or place cells in the hippocampus. Other models have also considered how time may be accumulated to provide a timing signal[91]. The models presented here could be expanded to include explicit models of such spatio-temporal signals.

Despite Similar Choice-Selective Sequences, Evidence Coding Differs Across Regions in a Manner Resembling Predictions of Each Model Class

Despite similar choice-selective sequences, we found that evidence coding schemes across regions were surprisingly distinct. The cortical regions primarily had monotonic encoding of evidence, as in the chains-based models, whereas the hippocampus had primarily non-monotonic encoding as in the bump model. This result emphasizes the importance of examining how variables are represented, not just if they are represented, to compare computational mechanisms and functions across regions[92].

The distinct evidence tuning properties we observed in hippocampus, ACC, RSC, and DMS support current ideas of their function in decision-making. The non-monotonic tuning curves in the hippocampus allow for a direct readout of the evidence based on the location of the active bump. This form of coding is more computationally intensive in the number of neurons required but could provide more refined information about the current state of the animal, supporting a joint cognitive map of position and evidence[48,93,94], as suggested by previous analysis[49]. In contrast, the monotonic encoding of evidence in ACC, RSC, and DMS does not allow a direct readout of evidence from the identity of the active neurons. Instead, it may be ideally suited for making a unitary decision based on the comparison of activity along the two

chains, and is similar to the graded coding that has been reported for value or confidence in ACC[95–101]. Interestingly, RSC, while dominated by monotonic tuning curves, did exhibit a population of narrowly tuned cells. This may reflect that the hippocampus acts as a major input area to RSC, while RSC receives from and projects to other areas of cortex[102–105], providing a possible suggestion as to how information about accumulated evidence is passed and transformed between brain regions. One possibility is that there is a single evidence-accumulation mechanism carried out in one set of brain regions, and other regions read out and transform this evidence into different tuning curve shapes. An alternative possibility is that both evidence accumulation mechanisms are used, but in different brain regions and for different purposes. Finally, the fact that DMS shows mainly binary choice tuning only in the delay region immediately before the turn (Fig. S4) aligns with work suggesting the role of the striatum in action selection[106–108], consistent with recordings in a different working memory task [55].

In regions with primarily monotonic evidence tuning (ACC, RSC), we observed that this tuning was graded early in the maze and sharpened to more choice-like tuning later in the cue period and throughout the delay (Fig. 6). While the graded representations of evidence during the early cue period in ACC and RSC is consistent with the chains-based model of evidence accumulation, our models do not explain the transition towards choice tuning later in the trial. This transition could reflect a change in dynamics as a result of decision commitment[109,110] or increasing urgency[111,112], and be produced in our model by increasing the strength of inhibition between the chains at later positions. The observation of a transition across the population from graded evidence coding to binary choice coding is reminiscent of the conclusions from a subpopulation of cells in another accumulation of evidence task in rats[110].

We applied our models to interpret neural activity in biological circuits, but the motifs and predictions of these models could also be tested for in trained recurrent neural network models. This could shed light on how less explicitly structured networks perform evidence accumulation[113]. A similar approach has been taken in the study of the path integration circuitry of the entorhinal cortex, where core foundational motifs of connectivity of the style presented here were revealed within the complex circuitry of the trained network[114]. Such studies also may reveal conditions favoring the emergence of evidence accumulation through sequences as opposed to traditional persistent activity, as has been analyzed for other cognitive operations [115]. Thus, tuning curve analysis in both biological and artificial networks may help to determine how these circuits perform the task.

Functionally Targeted Optogenetic Perturbations Could Provide Causal Tests of Predicted Mechanisms of Evidence Accumulation

We were able to compare evidence tuning in new and existing datasets to those predicted by the models. However, unlike targeted perturbations, correlational analyses do not provide causal evidence for the underlying computations in a region. For example, they cannot distinguish neurons involved in accumulating evidence from those that simply readout evidence accumulated elsewhere. Further, within our set of models, they do not differentiate between the varying chains-based models (Fig. 3F, Fig. S2F). Previous experiments have used optogenetic perturbations to identify brain regions that affect evidence accumulation at different times in the accumulating towers task[116,117], but perturbation experiments that selectively target neurons with

specific coding properties[118–120] remain to be performed during evidence accumulation. Perturbations in the oculomotor neural integrator, where functionally distinct populations are conveniently anatomically separated, have successfully tested model predictions about the dynamics of the integrator[121–123].

Targeted perturbations could provide insight into circuit mechanisms, and more broadly whether the evidence tuning observed in each region results from local evidence accumulation versus a downstream reflection of evidence accumulated in another region[14,124]. For example, evidence may be encoded in a traditional, non-sequential manner at some location and then combined at a downstream readout with a sequential representation of position to give the observed neural activity. If so, then an experimental prediction would be that activating choice-selective, evidence-encoding neurons during the cue region would not increase the firing rates of neurons tuned to later positions nor create a behavioral choice bias.

Summary

Understanding differences in evidence accumulation computations across brain regions is critical to advancing our understanding of how decision making is coordinated across the brain. The recent observation of choice-selective, sequential neural firing activity during decision-making challenges current models of evidence accumulation based on persistent or ramping neural activity. Here, we computationally demonstrate two distinct mechanisms for accumulating evidence with choice-selective sequences. Analyzing neural activity data across four brain regions, we found that evidence is encoded differently both across different brain regions and across different time periods within an evidence-accumulation task. Perturbation experiments suggested by our models could further elucidate how seemingly similar task representations may underlie distinct contributions to decision making.

FIGURES

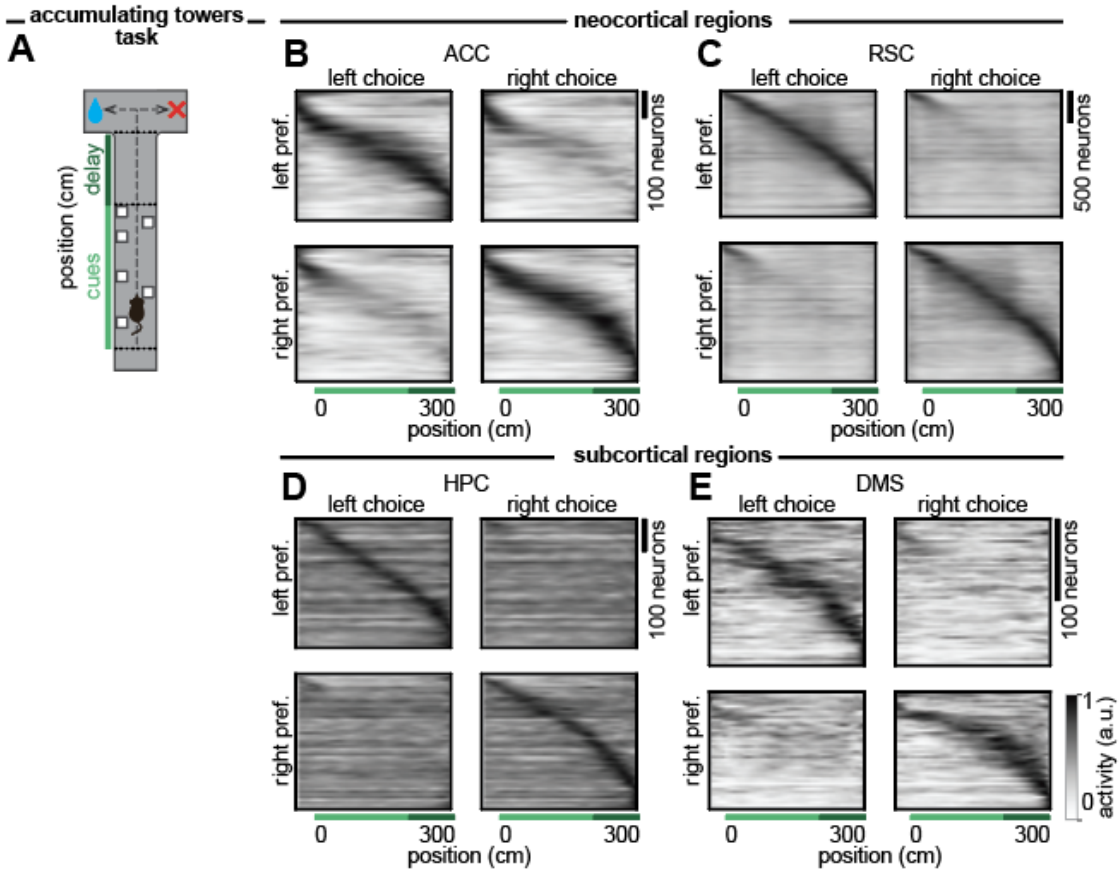

Figure 1. Choice-selective sequences are observed across brain regions in a navigation-based, accumulation-of-evidence task. (A) Schematic of a behavioral task in which mice navigate a T-maze in virtual reality. Visual cues (towers, white rectangles) are presented to either side of the maze during the cue period (light green, 200 cm long). At the end of the delay period (dark green, 100 cm long), mice are rewarded for turning to the side with more towers. **(B)** Each row shows the peak-normalized estimated firing rate at each position in the maze of a neuron with significant evidence tuning (see Methods) recorded during this task from anterior cingulate cortex (ACC, n = 964 neurons, see Methods), averaged across trials when the animal turned left (left choice, left column) or right (right choice, right column). Neurons were assigned as left or right choice preferring based on their evidence-selectivity (see Methods) and ordered based on the position of peak activity. **(C)** As in (B) but for retrosplenial cortex (RSC, n = 4190 neurons, data from [125]). **(D)** As in (B) but for hippocampus (HPC, n = 791 neurons, data from [49]). **(E)** As in (B) but for dorsomedial striatum (DMS, n = 322 neurons, see Methods). Sequences separated by even and odd trials to demonstrate cross-validation are shown in Figure S1.

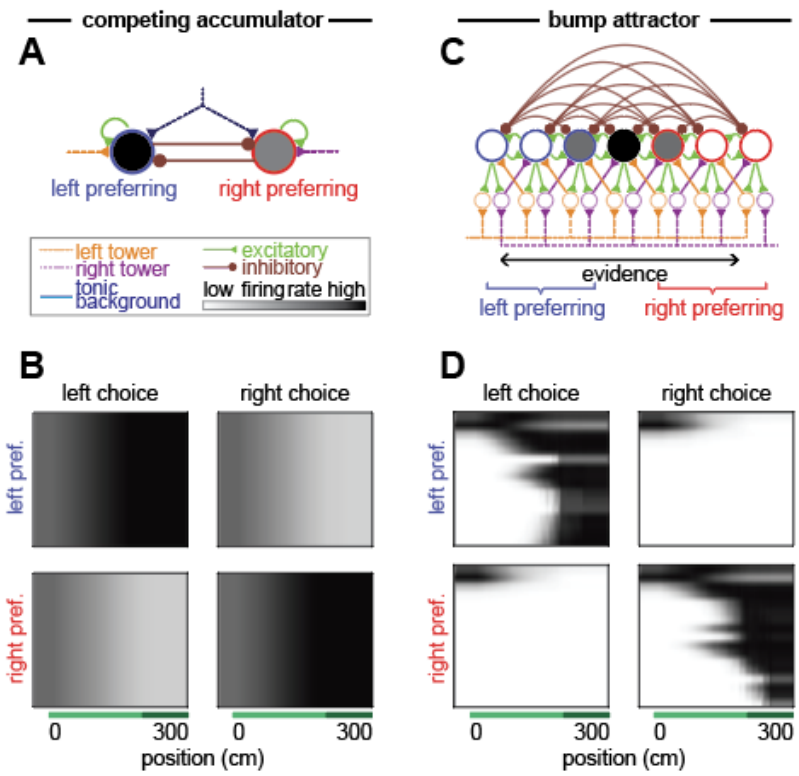

Figure 2. Choice-selective sequences are not predicted by traditional integrator models. (A) Schematic of a traditional mutually inhibitory integrator model for decision making. **(B)** Average peak-normalized, z-scored simulated activity of neurons in the model during task performance, divided into the left and right population and sorted by the peak of mean firing activity. **(C-D)** As in (A-B), but for a traditional bump attractor model.

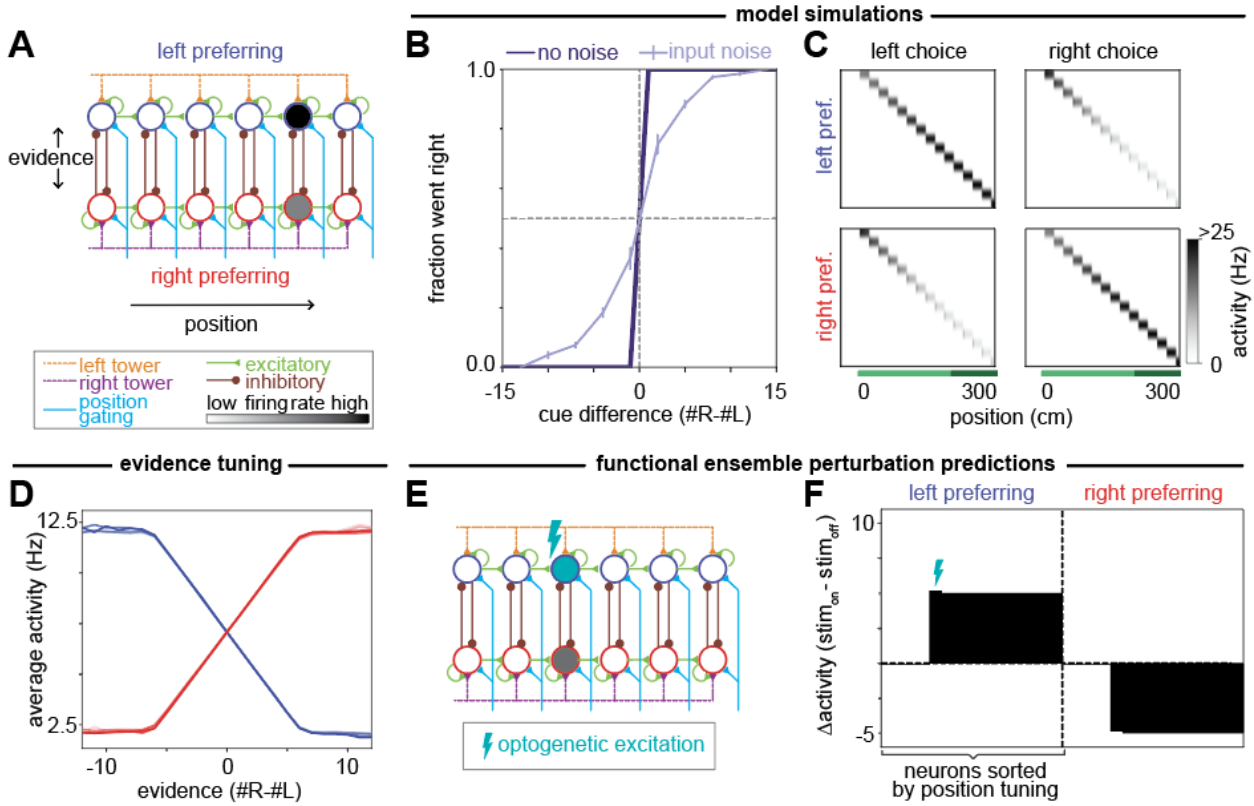

Figure 3. Mutually inhibitory competing chains model of evidence accumulation through sequences. (A) Schematic of neural circuit model for the mutually inhibitory competing chains model, showing excitatory (green) and inhibitory (brown) connections between neurons (circles) as well as external inputs to the circuit from the left (orange) and right (purple) towers and a position gating signal (cyan). **(B)** Psychometric data for the model simulated on 1000 trials describing how often the amplitude of the final neuron in the left chain was greater than that of the final neuron in the right chain for cases where the model was simulated with (light purple) and without (dark purple) noise in the input. Error bars show standard error of the mean (s.e.m.). **(C)** Each row shows the non-normalized amplitude of a model neuron at each position in the maze, averaged across simulated trials without input noise when the greater final amplitude was in the left (left choice, left column) or right (right choice, right column) chain. Neurons were divided based on their choice-selectivity (equivalent to the left and right chains in this model; see Methods) and ordered based on the position of peak activity (equivalent to position in the chain). **(D)** Tuning curves of individual neurons to evidence, defined by the average activity for different evidence levels when the neurons are activated by the position-gating signal, for left-preferring (blue) and right-preferring (red) cells. **(E)** Schematic of a single neuron perturbation experiment where optogenetic excitation is applied to a single neuron in the left chain. **(F)** Simulated changes of the firing rates of all neurons in the absence of cues when a single cell (denoted with the lightning bolt as in (E)) is optogenetically stimulated.

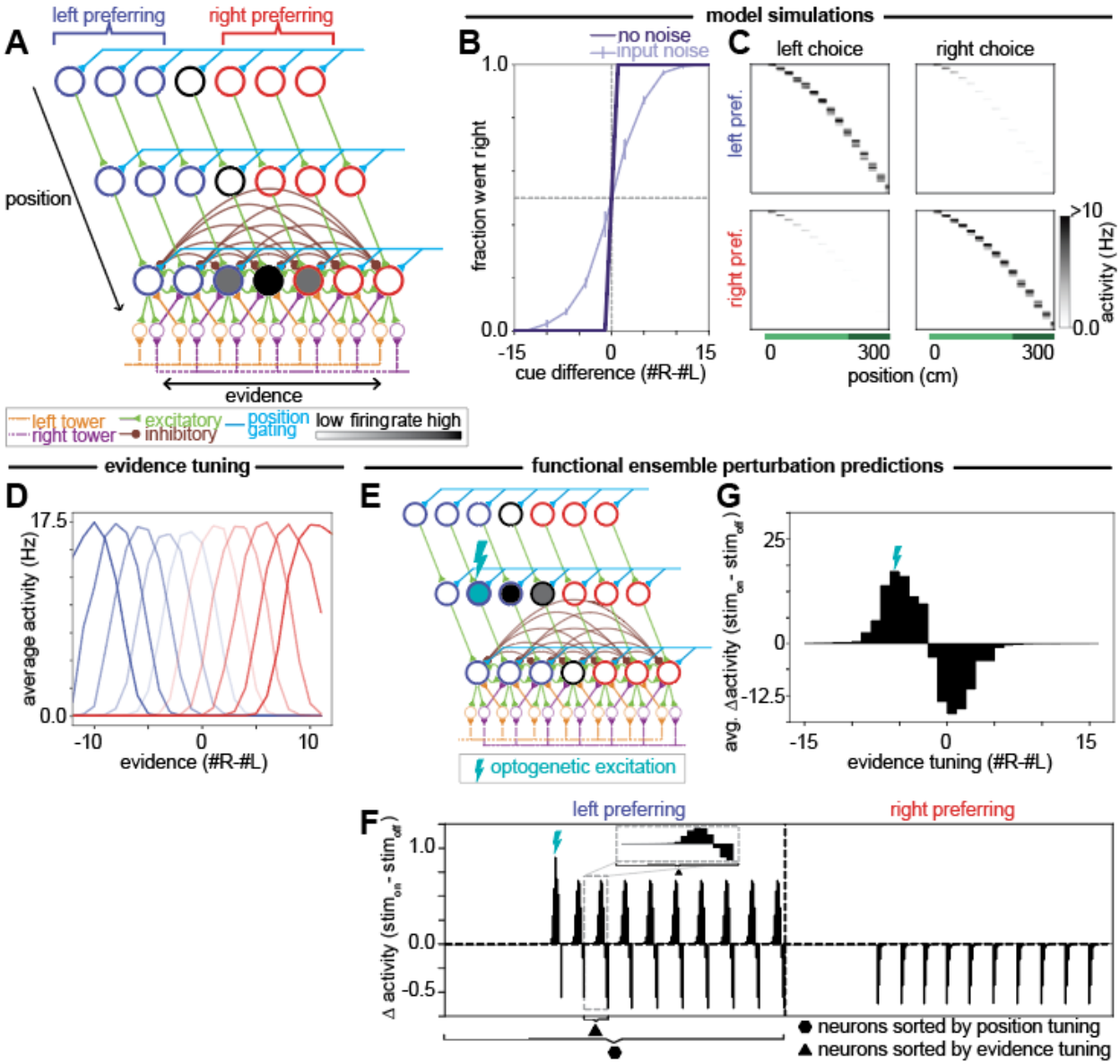

Figure 4. Position-gated bump attractor model of accumulation of evidence through sequences. **(A)** Schematic of neural circuit model for the position-gated bump attractor model, showing excitatory (green) and inhibitory (brown) connections between neurons (circles) as well as external inputs to the circuit from the left (orange) and right (purple) towers and a position gating signal (cyan). Note that excitatory and inhibitory connections within a given position are only shown for the bottom row, but these same connections are present at every position. **(B)** Psychometric data for 1000 simulated trials, indicating how often the index of the maximally firing active neuron in the last set of positions was to the left of the center neuron for cases with (light purple) and without (dark purple) noise in the input. Error bars show s.e.m. **(C)** Each row shows the non-normalized amplitude of a model neuron at each position in the maze, averaged across simulated trials without input noise when the greater final amplitude was in the left (left choice, left column) or right (right choice, right column) chain. Neurons were divided based on their choice-selectivity (see Methods) and ordered based on the position of peak activity. **(D)** Tuning curves of a subset of individual neurons to evidence, calculated when the neurons are activated by the position-gating signal, for left-preferring (blue) and right-preferring (red) cells. **(E)** Schematic of a single

neuron perturbation experiment in which optogenetic excitation is applied to a single neuron in the left preferring population. **(F)** Simulated changes of the firing rates of all neurons in the absence of cues when a single cell (denoted with the lightning bolt as in (E)) is optogenetically stimulated. **(G)** Average change in firing rate for neurons tuned to different evidence levels.

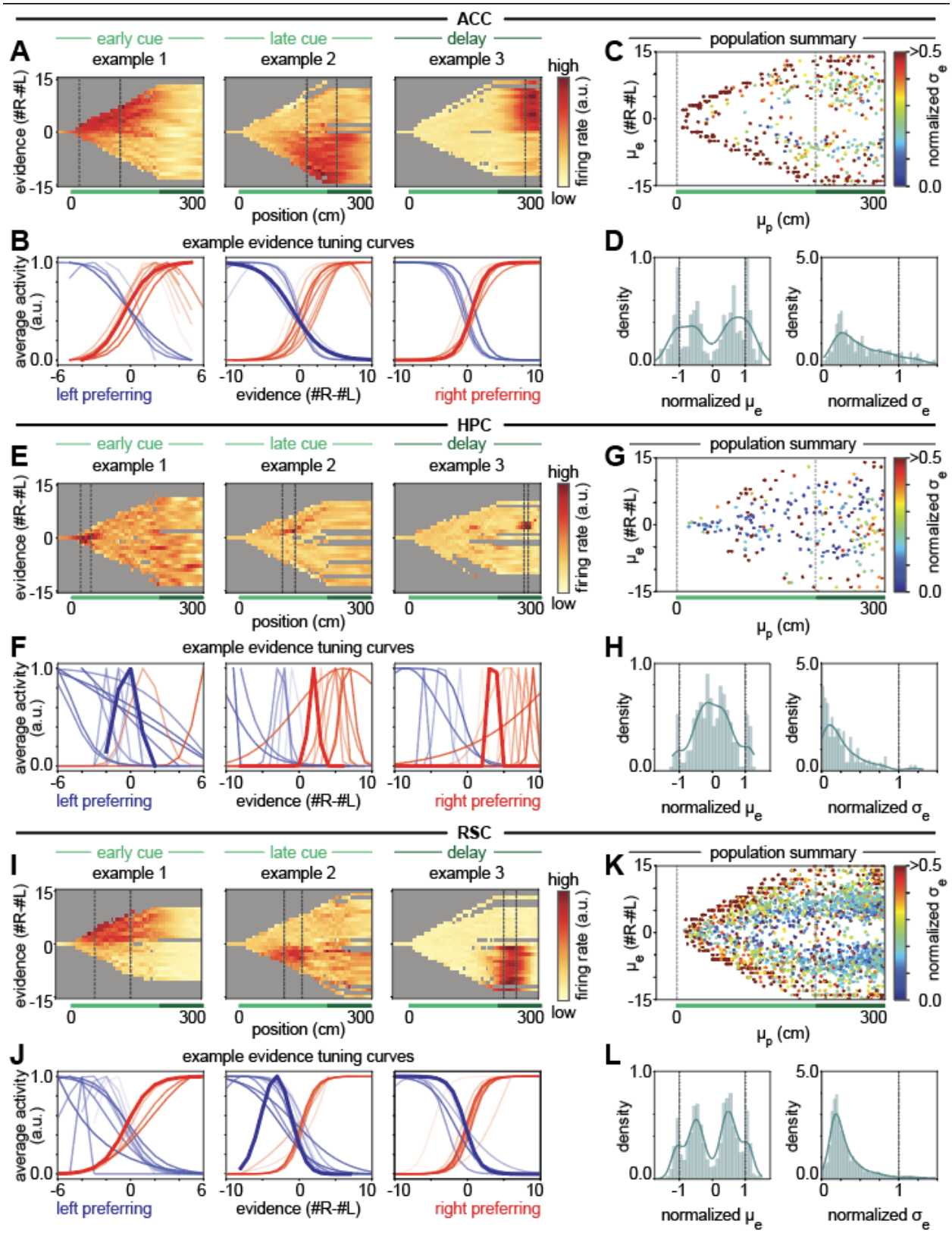

Figure 5. Different brain regions differ in the monotonicity of their evidence tuning. (A) Heatmaps showing the average firing in position-by-evidence bins for example individual neurons in ACC with mean

position in the early cue (left), late cue (middle), or delay (right) region of the maze. Gray bins denote regions for which there were fewer than 2 samples during the session. **(B)** Example ACC evidence tuning curves fit to the region of the neuron's peak activity (defined as observations within 0.5 fit position standard deviation of the neuron's fit position; see Methods) for a collection of neurons with mean position in the early cue (left), late cue (middle), or delay (right) region of the maze. Red coloring indicates neurons classified as right preferring, and blue indicates left preferring. Bold lines correspond to the examples in (A), where the neuron's region of peak activity is the region between the dashed vertical lines. Compare to raw average evidence tuning in Figure S5. **(C)** Scatter plot showing the location of the fit mean position (μ_p) and fit mean evidence (μ_e) with color indicating the normalized fit evidence standard deviation (σ_e) (see Methods) of the 80% of neurons recorded in ACC with the best fit between the neural data and the model predictions where narrowly tuned cells are only plotted if they pass non-outlier criteria (see Methods). Dashed lines indicate the boundaries of the cue period. Compare to distributions for all cells in Figure S10. **(D)** Histograms showing the normalized fit mean evidence (left) and the normalized fit evidence standard deviation (right) for the neurons shown in (C). Solid blue lines show the kernel density estimate. **(E-H)** Same as (A-D) but for neurons recorded in HPC. **(I-L)** Same as (A-D) but for neurons recorded in RSC.

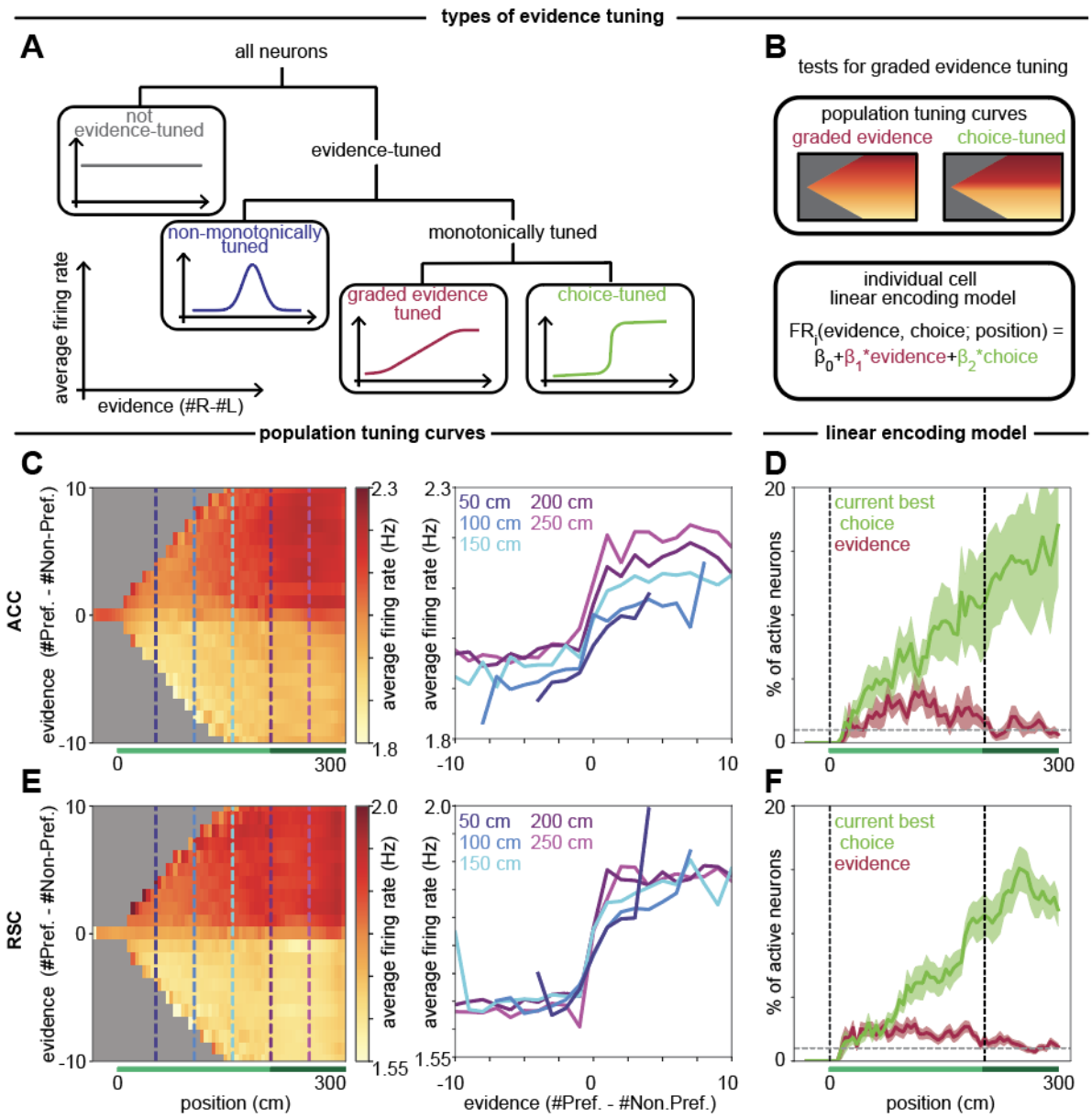

Figure 6. Progression from graded evidence to binary choice tuning across the maze in ACC and RSC. (A) Illustration of different possible types of evidence tuning, including both graded and choice-selective monotonically tuned neurons. **(B)** Schematics of two methods used to test for graded evidence tuning within monotonic regions: population averages (top) and a linear encoding model (bottom). **(C)** Left: Heatmap showing average firing rates across all evidence-tuned neurons in ACC in position by preferred-evidence bin. Gray indicates bins that were not sampled more than twice on any session. Right: Cross-sections of the heatmap at points in the cue period (50 cm, dark blue; 100 cm, medium blue; 150 cm, cyan) and delay period (200 cm, purple; 250 cm, magenta), showing average firing rate across evidence-tuned neurons as a function of preferred evidence. **(D)** For each position, the average fraction of active neurons in each session in ACC (defined as neurons with fit position mean within one position standard deviation of the location) at each location where the linear evidence encoding model (B, bottom) had significant coefficients for choice (lime) and evidence (crimson), based on the F-statistic. Error bars indicate s.e.m. across sessions. The horizontal dashed gray line indicates chance level (1%). **(E-F)** Same as (C-D) but for RSC.

METHODS

The Accumulating Towers Task

The accumulating towers task is behavioral paradigm, performed in virtual reality, for studying accumulation-of-evidence based perceptual decision making. The complete details of the task and experimental setup can be found in [57]. Here, we briefly summarize the key features of the task. Head-fixed mice run on an 8-inch styrofoam ball within a virtual environment, generated by ViRMEn [126], projected onto a surrounding screen. The virtual environment has a T-maze structure. In the first 30 cm (the precue region), there are no visual cues. In the next 200 cm (cue region), visual cues appear as towers on both sides, with the total count of towers on each side determined by a Poisson distribution of different means. Towers are positioned randomly with an enforced minimal distance between them. A tower appears when the mouse is within 10 cm of its position and disappears 200 ms later. Following the cue region, there is a 100 cm delay region during which the mouse must remember which side was presented with more towers. At the end of the delay region, the mouse receives a sweetened liquid reward for turning to the side with more cues. Water restriction and shaping protocols are also described in [57]. Recording sessions were approximately 1.0-1.5 h in duration, consisting of approximately 150-250 trials per session.

In addition to the accumulation of evidence trials described above, sessions also included some warm-up maze trials (cues on one side only, no distractors, tall visual guides at the end of the maze) at the beginning of the session and interspersed in the session if the animal's performance dropped below some threshold, as described in [57]. Since these trials do not require evidence accumulation, we did not include them in our analyses.

Neural Recordings

Animal procedures were conducted in accordance with standards from the National Institutes of Health and under the approval of the Princeton University Institutional Animal Care and Use Committee.

The head-fixed setup of the accumulating towers task allows for neural imaging throughout the task. Our neural datasets consist of calcium fluorescence data previously recorded from retrosplenial cortex (RSC, 8 animals, 41 sessions (29 sessions from layers 2/3 and 12 sessions from layer 5), 8579 total neurons) [44] and hippocampus (HPC, 7 animals, 7 sessions, 3144 total neurons) [49], as well as previously unpublished recordings from anterior cingulate cortex (ACC, 3 animals, 7 sessions, 1720 total neurons) and dorsomedial striatum (DMS, 8 animals, 11 sessions, 804 total neurons). See below for details of ACC and DMS recordings.

Animals and Surgery for ACC Recordings

Three female mice of CaMKIIa-tTA (JAX 007004) crossed with tetO-GCaMP6s [127] (JAX 024742) were used for two-photon imaging of ACC. These crossed lines show robust and stable expression of GCaMP6s in a subset of CaMKIIa-positive, glutamatergic neurons in the cortex. All mice were between 12-16 weeks when they underwent sterile stereotaxic surgery under isoflurane anesthesia (4% for induction, 1-1.5% for maintenance). The skull was exposed by making a vertical cut with a scalpel. Two craniotomy holes were made, one over ACC (anteroposterior axis: 0.8 mm, mediolateral axis: -0.5 mm, relative to bregma, diameter of ~1.2

mm) and another over DMS (anteroposterior axis: 0.75 mm, mediolateral axis: -1.1 mm, relative to bregma, diameter of ~0.3 mm). Adeno-associated virus (AAV; AAVretro2 serotype) encoding mCherry under CaMKIIa promoter (diluted to 2.5 x 10^{12} genome copies/mL, packaged at Princeton Vector Core) was injected to the DMS (anteroposterior axis: 0.75 mm, mediolateral axis: -1.1 mm, dorsoventral axis: -3.1 and -2.8 mm, relative to bregma) to provide information about a projection target and to be used for better motion correction. A total of 300 nL of AAV was infused at each site along the dorsoventral axis, at a rate of 30 nL/min. After injection, the needle was held in the same place for an additional 10 minutes. The needle was slowly withdrawn over 10 minutes to prevent backflow. After injections, the dura below the craniotomy was carefully removed within the ACC craniotomy site, and the cortical tissue above ACC was carefully aspirated up to 300 μm below the surface. Then we implanted a 1.0-mm diameter GRIN lens (4.3 mm length, Inscopix, 1050-002184) into the ACC, down to 1.5 mm in the dorsoventral axis (relative to bregma), using a 3D-printed custom lens holder. After implant, a small amount of diluted metabond (Parkell) was applied to affix the lens to the skull using a 1-mL syringe and 25-gauge needle. After 10 minutes, the lens holder grip was loosened while the lens was observed through the stereoscope to make certain that there was no movement of the lens. After applying a more diluted metabond around the lens, a titanium headplate was positioned over the skull using a custom tool and aligned parallel to the stereotaxic frame using an angle meter. The headplate was affixed to the skull using a metabond. Finally, a titanium ring to hold the immersion medium was placed on the headplate, with lens in the middle, and then affixed to the headplate using dental cement blackened with carbon. After the dental cement was fully cured, mice were unmounted from the stereotaxic frame and protective head caps were placed. Recovery of the mice was monitored for ~2 hours. After one or two weeks of recovery, mice started water-restriction, and after 5 days behavioral training for the accumulating towers task began, as described in [57]. Histology showed that GRIN lenses sampled mostly deep layers (layers 5 and 6) in two animals, while in one mouse a GRIN lens was medially implanted to cover both superficial (layers 2 and 3) and deep layers. For the data analysis, we did not make distinctions between neurons in different layers but instead pooled neurons from all layers together.

Animals and Surgery for DMS Recordings

D1R-Cre (n = 3, female, EY262Gsat, MMRRC-UCD), A2a-Cre (n = 3, male, KG139Gsat, MMRRC-UCD), and D2R-Cre (n = 2, male, ER44Gsat, MMRRC-UCD) mice were used for DMS two-photon imaging. Surgical procedures were identical to those described above for the ACC except for the following details. Approximately one week prior to GRIN lens implantation, a Cre-dependent AAV expressing GCaMP6f under either the CAG promoter (AAV5-CAG-DIO-GCaMP6f, Addgene, diluted to 3.5 x 10^{12} genome copies/mL) or synapsin promoter (AAV1-Syn-DIO-GCaMP6f, diluted to 3.8 x 10^{12} genome copies/mL) was delivered (500ul per injection site; 200 ul/min) to the DMS (anteroposterior axis: 0.7 mm, mediolateral axis: 1.15 and 1.85 mm, dorsoventral axis: -3 mm, relative to bregma) and the surgical incision was closed using monofilament non-absorbable sutures (Johnson and Johnson). This allowed for expression of GCaMP6f in D1R- or D2R-expressing medium spiny neurons according to the transgenic mouse line used, and facilitated the settling of brain tissue prior to GRIN lens implantation. Following recovery (5-10 days), sutures were removed and the viral injection

craniotomy was expanded and dura removed to accommodate a 1.0-mm diameter GRIN lens with a prism (4.3 mm length, Inscopix, 1050-004601). The lens was implanted with prism window facing posterior and just anterior to viral injection sites (anteroposterior axis: 0.75 mm, mediolateral axis: 1.35 mm, dorsoventral axis: -3.25 mm, relative to bregma). Fixation of the lens to the skull, as well as the custom-printed headplate, water immersion ring, and protective covering were identical to that described above for ACC. Across all DMS targeted GCaMP6f imaging mice we obtained three distinct fields of view from one mouse, two distinct fields of view from a second mouse, and the remaining mice contributed single fields of view (8 animals, 11 sessions, 804 total neurons).

Optical Imaging and Data Acquisition for ACC and DMS Recordings

Imaging was performed using a custom-built two-photon microscope compatible with virtual reality (VR). The microscope was equipped with two high-power femtosecond lasers (Alcor, Spark Laser), with fixed wavelengths for 920 and 1064 nm. The scanning units used a 5-mm galvanometer and a 8kHz resonant scanning mirror (Novanta). The collected photons were passed to a dichroic mirror that separates near-infrared light from the visible light and then split into two channels by a dichroic mirror (FF562-Di03, Semrock). The light for respective green and red channels was filtered using a bandpass filter (FF01-520/60 and FF02-607/70, Semrock), and then detected using GaAsP photomultiplier tubes (1077PA-40, Hamamatsu). The signal from the photomultiplier tubes was amplified using a high-speed current amplifier (59-179, Edmund). Output of the amplifier was sent to a data acquisition system (National Instruments). Black rubber tubing was attached around the objective (Zeiss, W N-Achroplan 20x/0.5 M27, 20x magnification, 0.5 NA) to shield the light from the virtual reality screen. The objective was aligned to the implanted GRIN lens within the implanted titanium headring. Double-distilled water was filled inside the headring and used as the immersion medium. The amount of laser power at the objective was 30~40 mW for ACC and 40~50 mW for DMS. Control of the microscope and imaging acquisition was performed with the ScanImage software (MBF Bioscience), on a separate computer from the VR-running one. Images were acquired at 30 Hz with a size of 512 x 512 pixels. Synchronization between behavioral logs and acquired images was achieved by sending behavioral information from the VR controlling computer to the scanning computer via an I2C serial bus each time the ViRMEn environment was refreshed: this information about sessions, trials, and frames was stored in the header of TIFF image files.

Data Processing for ACC and DMS Recordings

Both ACC and DMS imaging data were processed with Suite2p. For ACC recordings in which mCherry signals are present, rigid motion correction was performed with that channel and then applied to the functional green channel with GCaMP6s, with the following parameter settings: nimg_init = 1000, smooth_sigma = 1.15, maxregshift = 0.1. For DMS recordings, the same motion correction was performed, but with the green channel. Region-of-interests (ROI) corresponding to individual neurons were detected, and calcium signals were extracted and neuropil-corrected (neucoeff = 0.7). Subsequently, an experimenter-mediated curation was performed on these ROIs to discriminate putative neurons from noise. Neuropil-corrected fluorescence from these curated neurons were exported to MATLAB, and the $\Delta F/F$ was

calculated, taking baseline fluorescence as the mode of the fluorescence within 3 minute long windows (the same baseline as applied in the V1, RSC, and HPC datasets).

Spike Inference

In order to compare neural recordings across datasets with different calcium indicators, we performed spike inference on our data based on a recently developed and validated algorithm that has only two free parameters, the average estimated firing rate, and the calcium decay rate [128,129]. For each ΔF/F dataset, we assumed a baseline firing rate of 2 Hz (but our results were not sensitively dependent on this specific choice) and set the calcium decay rate according to the values fit by [128] for each calcium indicator (0.9885 for GCaMP6s (ACC), 0.9704588 for GCaMP6f (all other regions)). From these inferred spikes, we inferred a continuous firing rate by smoothing with a gaussian window function (window length = 1 s, standard deviation = 0.25 s).

Position Binning of Neural and Evidence Data

To compare neural data along maze positions and evidence levels, we resampled the neural (continuous firing rate as calculated above) and behavioral (cumulative evidence up to that point) data in position bins. Specifically, firing rates were averaged based on the position of the animal within 5 cm bins from the start of the trial (-30 cm) to the end of the delay region (300 cm) for each trial. Cumulative evidence for each bin was defined by the number of right minus the number of left cues observed before the start of the position bin.

Trial Selection Criteria

We discarded a few trials from analysis using the following criteria. As mentioned above, any trials from the warm-up maze were excluded, as these trials do not require evidence accumulation. Additionally, trials with excessive travel were discarded (3.2% of trials in ACC sessions, 8.0% of trials in DMS sessions), which were identified as trials where mice showed a large view angle deviation (more than 90 degrees, either left or right) before the end of the delay region.

Determining Choice Selectivity of Neurons

To identify choice-selective neurons (Fig. 1B-E), we first jointly fit the position and evidence tuning of each neuron in our dataset as described in Data Analysis below. We then considered only cells that had significant evidence tuning. To produce the sequence plots, cells were first sorted as right preferring if the mean of the neuron's evidence tuning was greater than 0 and left preferring otherwise and then sorted by their position mean. Using this sorting, we then took the average activity of each neuron at different positions for correct left and correct right trials. The resulting traces for each neuron were then peak normalized.

Parameterization of the Mutually Inhibitory Chains-Based Models

As in equation (1), the activity of the neuron at position *i* in the left chain can be described as

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [br_{i,L} + cr_{i-1,L} - er_{i,R} + f\mathbf{1}_{\text{left}}(t) + P_i(t) - T]^+$$

A neuron's active position is defined by a square pulse over a range of positions, where *p(t)* is the position of the animal at time *t,*

$$P_i(t) = \begin{cases} T + X(t) & i \leq \frac{p(t)}{P_0} < (i+1) \\ 0 & \text{otherwise} \end{cases}$$

The external input is of the form

$$\mathbf{1}_{\text{left}}(t) = \begin{cases} 1 & \min_{\ell \in L} |t - \ell| < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\mathbf{1}_{\text{right}}(t) = \begin{cases} 1 & \min_{R \in \mathcal{R}} |t - R| < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

where L is the set of times of left cues and $\mathcal{R}$ is the set of times of right cues.

In our simulations, we use parameters a =1 s^{-1}, b = 0.2 s^{-1}, c = 1 s^{-1}, e =0.8 s^{-1}, f = 2 s^{-1}, T = 300 Hz/s, and $X(t)$ = 20 Hz/s. (See Supplementary Text for necessary conditions on the parameters.) We take P_0 = 20 cm, which divides the total length of the maze (330 cm) into 17 different sections of active positions. Therefore, in these models we simulate 17 neurons in each chain for a total of 34 neurons.

Parameterization of the Position-Gated Bump Attractor

In Equation (2), we defined the evolution of the firing rate for a neuron at position *i* and evidence *j* as

$$\frac{dr_{i,j}}{dt} = -ar_{i,j} + \Theta(P_i(t) - T)\left[F\left(\sum_k W_{j,k} r_{i,k} + br_{i-1,j} + c\left(I_{i,j+1,L}(t) + I_{i,j-1,R}(t)\right)\right)\right].$$

The Heaviside function that acts as a multiplicative position gate is defined as

$$\Theta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

so that the term in square brackets only affects the activity of the neuron at the active position. A neuron's active position is defined by a square pulse over a range of positions, where *p(t)* is the position of the animal at time *t,*

$$P_i(t) = \begin{cases} T + X(t) & i \leq \frac{p(t)}{P_0} < (i+1) \\ 0 & \text{otherwise} \end{cases}.$$

The neuronal nonlinearity is

$$F(x) = \frac{q(1 + \tanh(x))}{2}.$$

The strengths of the synaptic connections between neurons were chosen to be symmetric with excitatory connections between neurons close in evidence and inhibitory connections between neurons farther in evidence. This was done by assigning each evidence level an angle, evenly spaced between 0 and 2π, and setting

$$W_{j,k} = \omega_0(\cos(\theta_j - \theta_k) + \omega_1).$$

We note that, since evidence is not periodic, a sufficiently large number of evidence levels were used such that substantial interactions do not occur between the nominally periodically connected endpoints.

When an input is present, asymmetric connections cause the active evidence neurons to shift towards the evidence. This is modeled through shifter neurons (Fig. 4A, purple and orange small circles) that are activated by the coincidence of external cue signals and neural activity from the current location of the bump by

$$I_{i,j,L}(t) = r_{i,j}\mathbf{1}_{\text{left}}(t)$$

and

$$I_{i,j,R}(t) = r_{i,j}\mathbf{1}_{\text{right}}(t)$$

where

$$\mathbf{1}_{\text{left}}(t) = \begin{cases} 1 & \min_{\ell \in L} |t - \ell| < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\mathbf{1}_{\text{right}}(t) = \begin{cases} 1 & \min_{R \in \mathcal{R}} |t - R| < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

where L is the set of times of left cues and $\mathcal{R}$ is the set of times of right cues.

As in the chains-based models, the position-gating of each layer of neurons is activated over a length 20 cm (P_0 = 20), again giving 17 different active positions. For the neurons to span the range of evidence of the experiments, we use 35 neurons in the evidence dimension, corresponding to evidence levels of -17 to 17. Thus we simulate a total of 595 neurons.

The parameters used in the simulation are: a = 1.1 s^{-1}, b = 0.088 s^{-1}, T = 300 Hz/s, $X(t)$=0 Hz/s, q = 25, ω_0 = 0.12 s^{-1}, ω_1 = -0.9, and c = 0.2 s^{-1}. Neurons $r_{0,-1}$, $r_{0,0}$, and $r_{0,1}$ are initialized to 0.6 activity level, and all other neurons are initialized to 0 activity.

Simulation of Differential Equations

Differential equations for each model (equations (1) and (2) and (S2)) were simulated in Python using scipy.odeint, which uses the lsoda algorithm, with a maximum stepsize of 1 ms (or equivalently 0.5 cm when assuming a constant velocity of 50 cm/s) to find the activity levels of the neurons in each model for points between -30 and 300 cm in 0.1 cm increments. In all simulations, we assume the animal travels at a constant velocity of 50 cm/s through the maze so that time and position are interchangeable. We simulated the neural activity for 1000 trials with cue positions derived from 1000 experimental trials from the behavioral and neural recording data. For the chains-based models, we say that the animal makes a left choice if the activity of the last neuron in the left chain is greater than the activity of the last neuron in the right chain. For the bump model, we say that the animal makes a left choice if the index k of the neuron with maximal firing in the last layer is less than 0.5J, where J is the total number of neurons in the layer. If the maximal firing occurs exactly at 0.5J, we determine the animal's decision by randomly choosing between the two alternatives with equal probability.

Simulations with Input Noise

For each model, we simulated 25 sessions with input noise (Figs. 3B, 4B, S2B). For each session, we selected 150 trials drawn randomly without replacement from the 1000 experimental trials that we used for simulation of the model. Let *L* and *R* be the set of left cues and right cues, respectively, in the original trial. For each cue *x* in *L*, we drew *z*, a random uniform number between 0 and 1, and if $z<0.33$, we added *x* to the set *L*'. We repeated the same process for each cue in *R*, yielding two reduced sets of cues when input noise is present, *L*' and *R*', corresponding to the animal ignoring or not perceiving 67% of the cues. We then simulated the model with the input terms $\mathbf{1}_{left}(t)$ and $\mathbf{1}_{right}(t)$ determined according to the reduced cue sets. We then calculated the psychometric curves (as described in Psychometric Curves below) with the difference in cues based on the original cue sets, but whether the animal turns left is based on the simulation of the trial with the reduced cue sets.

Evidence Tuning Curves for Simulated Model Data

In order to find the evidence tuning curves (Figs. 3D, 4D, S2D), we averaged the activity of each neuron binned by position and accumulated evidence for each of the 1000 simulated trials. Position bins were of size 0.1 cm, ranging from -30 to 300 cm. Evidence bins were of size 1 tower, ranging from -15 to 15 towers. Bins that had fewer than 2 samples were not included in further analysis and not plotted. For a given neuron, we averaged across evidence levels to find the position with maximum firing. We then plotted the average firing rate versus evidence level at this position.

Simulated Single Cell Perturbations

Single cell perturbations were simulated in the same manner as no-perturbation trials but in the absence of any cues. A subthreshold excitatory term $O_i(t)$ was added within the thresholded dynamics of each model (i.e. within the square brackets of equation (1) or within the function F of equation (2)) to simulate the effects of stimulating an individual cell, where $O_i(t) = 1$ if neuron *i* is the neuron being stimulated and 0 otherwise.

Psychometric Curves

For each session, we sorted trials based on the difference in the final number of cues at the end of the maze. We binned these trial differences with bin size of 3 towers, ranging from -14 to 14 towers. Within each bin, we then found the fraction of trials in each bin for which a left choice was made (Fig. 3B, Fig. 4B, Fig. S2B).

Data Analysis

Determining Position and Evidence Selectivity

In our neural recording data, we sought to identify the set of cells whose firing encoded accumulated evidence. As described in the main text, due to the nonuniform sampling of position and evidence and fact that neurons are only active at some positions, we found that robustly identifying evidence selectivity required that we jointly identify the position range over which a neuron is active and the evidence selectivity within this set of positions. To do so, we assumed the firing rate of each neuron had the form,

$$FR(p,e) = aP(p)E(e) + b$$

where *FR(p,e)* is the firing rate of the neuron at position *p* and evidence *e*, and *P* and *E* were described by Gaussians,

$$P(p) = \exp\left(\frac{-(p-\mu_p)^2}{2\sigma_p^2}\right)$$

and

$$E(e) = \exp\left(\frac{-(e-\mu_e)^2}{2\sigma_e^2}\right).$$

The above functional form, which assumes evidence and position tuning can be approximated by independent Gaussians, was not used as a detailed descriptor of neuronal firing (for less constrained fits around a given position, see Neural Evidence Tuning Curves Around the Most Active Position). Rather, we found that it was effective for its purpose of identifying evidence-modulated neurons and determining if the mean evidence selectivity favors left or right evidence (as used to sort neurons by evidence-selectivity in Fig. 1B-E). Here, the Gaussian was allowed to extend beyond the range of observed evidence values, so that it can accommodate both monotonic and non-monotonic tuning curves (i.e., if the mean evidence occurs at the extremes of the evidence range, the evidence tuning will be monotonic throughout the range).

For each neuron, we normalized the firing rate observations (binned as in Position Binning of Neural and Evidence Data) to the range [0, 1]. We then fit the parameters of the above model using an iterative fitting procedure, alternating fitting μ_p, σ_p, a, and b with fitting μ_e, σ_e, a, and b with a nonlinear fitting method (scipy curvefit, which employs the trust region reflective algorithm, with bounds $-50 \leq \mu_p \leq 350$, $0 \leq \sigma_p \leq 200$, $e_{low}-1 \leq \mu_e \leq e_{high}+1$, $0 \leq \sigma_e \leq 30$, $0 \leq a \leq 10$, and $0 \leq b \leq 1$). The bounds on μ_e are defined by $e_{low} = \min(E')$ and $e_{high} = \max(E')$, where E' is the set of all evidence levels observed by the neuron within 5cm of μ_p. These bounds are updated on each iteration based on the current fit of μ_p. Due to the iterative nature of our fitting procedure, after the final position update, μ_e may no longer satisfy these bounds, and we do not include such cells when analyzing the distributions of the normalized evidence parameters.

We then tested for significance of the evidence tuning fit using the pseudosession method [130]. For each session of firing rate and position data, we generated new evidence for each trial in the session by resampling trials from actual sessions, and applied the iterative fitting procedure to the new pseudosession evidence and position data. We repeated this process for 50 pseudosessions and calculated the correlation coefficient between the true firing rates and the model predictions for each session. We then used a one sample t-test to compare the fit of the model, as measured by the correlation coefficient, in the true session to the fit of the null distribution from the pseudosession, and only considered sessions for which the fit was significantly better ($p<0.05$) than the null distribution for further analysis.

In order to compare the parameterization of the evidence tuning across positions (Fig. 5D,H,L, Fig. S4D), we normalized μ_e and σ_e to account for the different observed evidence ranges at different positions. Let $\mathcal{E}$ be the set of all evidence values observed by the neuron at positions within 5 cm (one position bin) of μ_p. Define $\mathcal{E}^+ = \{e \in \mathcal{E} | e > 0\}$, i.e. the set of positive observed evidence levels, and $\mathcal{E}^- = \{e \in \mathcal{E} | e < 0\}$. For $\mu_e < 0$, the normalized mean is given by

$$\bar{\mu}_e = \frac{\mu_e}{\max_{e \in \mathcal{E}^-}(|e|)},$$

and for $\mu_e > 0$, the normalized mean is given by

$$\bar{\mu}_e = \frac{\mu_e}{\max_{e \in \mathcal{E}^+}(e)}.$$

The normalized evidence standard deviation is given by

$$\bar{\sigma}_e = \frac{\sigma_e}{\max(\mathcal{E}) - \min(\mathcal{E})}.$$

For visualization purposes (Fig. 5, Fig. S4), we plot only the 80% of significantly evidence-selective neurons in each region with the best fit. For robustness to single large observations making a neuron appear narrowly tuned, for cells with $\sigma_e < 3$, we identified the nearest position-evidence bin to (μ_p, μ_e). If there were fewer than 3 observations in this bin or if there was a single outlier in this bin (defined by one observation greater than 1.5 times the interquartile range), the cell was excluded from plotting. For visualization of all significantly evidence-selective cells without any further criteria applied, see Figure S10.

Neural Evidence Tuning Curves Around the Most Active Position

To generate the 1-D evidence tuning curves around the most active position (Fig. 5B,F,J, Fig. S4B) for each neuron with significant evidence tuning, we identified the subset of data within $0.5\sigma_p$ of μ_p. (Note that if this range extended less than 0 cm or greater than 300 cm, only the data within 0-300 cm was used.) We then fit the data within this range to two functional forms, a Gaussian and a logistic function, and plotted the curve that better fit the true firing rate data. The Gaussian evidence tuning curve fit took the form

$$FR(e) = a \exp\left(\frac{(e-\mu)^2}{\sigma^2}\right) + c$$

where *FR(e)* denotes the firing rate as a function of the evidence *e*. The fit was performed with a nonlinear fitting method, scipy curvefit, which employs the trust region reflective algorithm, with bounds $m < \mu < M$, $0 < \sigma < 30$, $0 < a < 10$, $0 < c < 1$, where *m* is the minimum and *M* is the maximum observed evidence within the active position. The logistic evidence tuning curve fit took the form

$$FR(e) = \frac{a}{(1 + \exp(-k(e - x_0))} + c$$

with a nonlinear fitting method (scipy curvefit, which employs the trust region reflective algorithm, with bounds $-1 < k < 1$, $-15 < x_0 < 15$, $0 < a < 10$, $0 < c < 1$). To determine the curve with better fit, we calculated the correlation of the predicted firing rates from each fit curve with the true firing rates. We note that each form of the tuning curve has the same number of fit parameters, so there was no need to correct for parameter numbers when comparing model predictions.

For comparison to our fit tuning curves, we also plotted the raw evidence tuning curves around the most active position (Fig. S5). These were computed by taking the average of the firing rates around the most active position at each evidence level. We note that such locally fit or raw evidence tuning curves around a given position required a reasonable assessment of the peak position tuning, as was obtained in our joint position-evidence fits described above.

However, once the most active positions were obtained, the fits no longer depended on requiring the evidence to have the same functional form at every position or, in the case of the raw fits, on a particular functional form. Thus, we found that the fits could provide a more quantitatively precise visualization of the evidence tuning around the most active position and, for the cells best fit by logistic functions, a better quantitative assessment of the width of the sigmoid. However, beyond this small quantitative difference, the fits typically did not differ substantially from the joint Gaussian fits, and required fitting to two separate functional forms, so we used the joint Gaussian fits for the population-wide assessment of tuning in Figure 5.

Population-Averaged Tuning Curves

In Figure 6C,E, we calculated the population-averaged response across all significantly evidence-selective neurons as follows. For each neuron, we calculated the average firing rate in each position-evidence bin. If a bin had fewer than 3 observations for a particular neuron, that neuron was not included in the population average for that bin. To combine data for neurons that preferred positive with data from neurons that preferred negative evidence, for each neuron, we defined its preferred-evidence as the sign of the neuron's μ_e value and remapped the evidence axis to be in terms of preferred-evidence rather than absolute evidence (by multiplying the evidence values by -1 if the preferred tuning was left instead of right). We then took the average in each (position, preferred evidence) bin across all neurons.

Linear Encoding Model

For each neuron at each position bin, we fit a linear encoding model (Fig. 6D,F, Fig. S4F) to describe the firing rate of the *i*th neuron

$$FR_i(e, c) = \beta_0 + \beta_1 e + \beta_2 \mathrm{sign}(e)$$

as a function of evidence (*e*) and the current best choice (i.e., the sign of the evidence), using ordinary least squares regression. For each coefficient, we calculated the F-statistic, and used the F-test to determine whether a coefficient was significant at a 0.01 significance level. At each position, for each session, we found the fraction of neurons with a significant coefficient of the total neurons that are active at that position. For this analysis, a neuron was defined as active at position p if $|p - \mu_p| < \sigma_p$, based on the results of the position-evidence fitting procedure described above.

Data Availability

Code for the simulation of all models and the data analysis is available upon request.

Conflicts of Interest

The authors report no conflicts of interest.

Acknowledgements

We thank Sue Ann Koay for her useful discussions about the retrosplenial cortex dataset. We thank Shizhe Chen, Tim Hanks, Ben Lankow, and Avinash Baidya for useful discussions about this project. Funding was provided by NIH T32MH065214 (LSB), F32MH132179 (LSB), NIH-NINDS BRAIN Initiative 5U19NS104648 (LSB, DWT, CBD, IBW, MSG) and

1U19NS132720 (LSB, DWT, CBD, IBW, MSG), K99DA053388 (EHN), Sloan Swartz Foundation (LSB), a C.V. Starr Fellowship (JRC, MS), and a Burroughs Wellcome Fund's CASI award (MS).

REFERENCES

1. Mountcastle, V. B., Steinmetz, M. A. & Romo, R. Frequency discrimination in the sense of flutter: psychophysical measurements correlated with postcentral events in behaving monkeys. *J. Neurosci.* **10**, 3032–3044 (1990).
2. Luna, R., Hernández, A., Brody, C. D. & Romo, R. Neural codes for perceptual discrimination in primary somatosensory cortex. *Nat. Neurosci.* **8**, 1210–1219 (2005).
3. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J. Neurosci.* **12**, 4745–4765 (1992).
4. Britten, K. H., Shadlen, M. N., Newsome, W. T. & Movshon, J. A. Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* **10**, 1157–1169 (1993).
5. Shadlen, M. N., Britten, K. H., Newsome, W. T. & Movshon, J. A. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J. Neurosci.* **16**, 1486–1510 (1996).
6. Britten, K. H. & Van Wezel, R. J. A. Area MST and heading perception in macaque monkeys. *Cereb. Cortex* **12**, 692–701 (2002).
7. Heuer, H. W. & Britten, K. H. Optic flow signals in extrastriate area MST: comparison of perceptual and neuronal sensitivity. *J. Neurophysiol.* **91**, 1314–1326 (2004).
8. DeAngelis, G. C. & Uka, T. Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J. Neurophysiol.* **89**, 1094–1111 (2003).
9. Afraz, S.-R., Kiani, R. & Esteky, H. Microstimulation of inferotemporal cortex influences face categorization. *Nature* **442**, 692–695 (2006).
10. Heekeren, H. R., Marrett, S., Bandettini, P. A. & Ungerleider, L. G. A general mechanism for perceptual decision-making in the human brain. *Nature* **431**, 859–862 (2004).

11. Philiastides, M. G. & Sajda, P. Temporal characterization of the neural correlates of perceptual decision making in the human brain. *Cereb. Cortex* **16**, 509–518 (2006).
12. Uchida, N. & Mainen, Z. F. Speed and accuracy of olfactory discrimination in the rat. *Nat. Neurosci.* **6**, 1224–1229 (2003).
13. Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95–98 (2013).
14. Boyd-Meredith, J. T., Piet, A. T., Dennis, E. J., El Hady, A. & Brody, C. D. Stable choice coding in rat frontal orienting fields across model-predicted changes of mind. *Nat. Commun.* **13**, 3235 (2022).
15. Carpenter, R. H. S. Movements of the eyes, 2nd rev. & enlarged ed. *nd rev. & enlarged ed* **2**, 593 (1988).
16. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Choosing the greater of two goods: neural currencies for valuation and decision making. *Nat. Rev. Neurosci.* **6**, 363–375 (2005).
17. Pearson, J. M., Hayden, B. Y., Raghavachari, S. & Platt, M. L. Neurons in posterior cingulate cortex signal exploratory decisions in a dynamic multioption choice task. *Curr. Biol.* **19**, 1532–1537 (2009).
18. Sugrue, L. P., Corrado, G. S. & Newsome, W. T. Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782–1787 (2004).
19. Dorris, M. C. & Glimcher, P. W. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* **44**, 365–378 (2004).
20. Costa, V. D., Mitz, A. R. & Averbeck, B. B. Subcortical Substrates of Explore-Exploit Decisions in Primates. *Neuron* **103**, 533–545.e5 (2019).
21. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
22. Laming, D. R. J. *Information Theory of Choice-reaction Times*. (Academic P., 1968).
23. Ratcliff, R. A theory of memory retrieval. *Psychol. Rev.* **85**, 59–108 (1978).

24. Stone, M. Models for choice-reaction time. *Psychometrika* **25**, 251–260 (1960).

25. Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113**, 700–765 (2006).

26. Usher, M. & McClelland, J. L. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* **108**, 550–592 (2001).

27. Ditterich, J., Mazurek, M. E. & Shadlen, M. N. Microstimulation of visual cortex affects the speed of perceptual decisions. *Nat. Neurosci.* **6**, 891–898 (2003).

28. Wang, X.-J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).

29. Machens, C. K., Romo, R. & Brody, C. D. Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* **307**, 1121–1124 (2005).

30. Lo, C.-C. & Wang, X.-J. Cortico–basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat. Neurosci.* **9**, 956–963 (2006).

31. Engel, T. A. & Wang, X.-J. Same or different? A neural circuit mechanism of similarity-based pattern match decision making. *J. Neurosci.* **31**, 6982–6996 (2011).

32. Roach, J. P., Churchland, A. K. & Engel, T. A. Choice selective inhibition drives stability and competition in decision circuits. *Nat. Commun.* **14**, 147 (2023).

33. Wong, K.-F. & Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* **26**, 1314–1328 (2006).

34. Lam, N. H. *et al.* Effects of Altered Excitation-Inhibition Balance on Decision Making in a Cortical Circuit Model. *J. Neurosci.* **42**, 1035–1053 (2022).

35. Goldman, M. S., Compte, A. & Wang, X.-J. Neural integrator models. *Encyclopedia of neuroscience* 165–178 (2009).

36. Gold, J. I. & Shadlen, M. N. Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* **5**, 10–16 (2001).

37. Doya, K. *Bayesian Brain: Probabilistic Approaches to Neural Coding*. (MIT Press, 2007).

38. Roitman, J. D. & Shadlen, M. N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* **22**, 9475–9489 (2002).

39. Cook, E. P. & Maunsell, J. H. R. Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat. Neurosci.* **5**, 985–994 (2002).

40. Shadlen, M. N. & Newsome, W. T. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *J. Neurophysiol.* **86**, 1916–1936 (2001).

41. Chandrasekaran, C., Peixoto, D., Newsome, W. T. & Shenoy, K. V. Laminar differences in decision-related neural activity in dorsal premotor cortex. *Nat. Commun.* **8**, 614 (2017).

42. Peixoto, D. *et al.* Decoding and perturbing decision states in real time. *Nature* **591**, 604–609 (2021).

43. Hanks, T. D. *et al.* Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015).

44. Koay, S. A., Charles, A. S., Thiberge, S. Y., Brody, C. D. & Tank, D. W. Sequential and efficient neural-population coding of complex task information. *Neuron* **110**, 328–349.e11 (2022).

45. Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62–68 (2012).

46. Krumin, M., Lee, J. J., Harris, K. D. & Carandini, M. Decision and navigation in mouse parietal cortex. *Elife* **7**, (2018).

47. Gage, G. J., Stoetzner, C. R., Wiltschko, A. B. & Berke, J. D. Selective activation of striatal fast-spiking interneurons during choice execution. *Neuron* **67**, 466–479 (2010).

48. MacDonald, C. J., Lepage, K. Q., Eden, U. T. & Eichenbaum, H. Hippocampal ‘Time Cells’ Bridge the Gap in Memory for Discontiguous Events. *Neuron* **71**, 737–749 (2011).

49. Nieh, E. H. *et al.* Geometry of abstract learned knowledge in the hippocampus. *Nature* **595**, 80–84 (2021).

50. Siniscalchi, M. J., Wang, H. & Kwan, A. C. Enhanced Population Coding for Rewarded Choices in the Medial Frontal Cortex of the Mouse. *Cereb. Cortex* **29**, 4090–4106 (2019).
51. Cox, J. *et al.* A neural substrate of sex-dependent modulation of motivation. *Nat. Neurosci.* **26**, 274–284 (2023).
52. Parker, N. F. *et al.* Choice-selective sequences dominate in cortical relative to thalamic inputs to NAc to support reinforcement learning. *Cell Rep.* **39**, 110756 (2022).
53. Terada, S., Sakurai, Y., Nakahara, H. & Fujisawa, S. Temporal and Rate Coding for Discrete Event Sequences in the Hippocampus. *Neuron* **94**, 1248–1262.e4 (2017).
54. Pastalkova, E., Itskov, V., Amarasingham, A. & Buzsáki, G. Internally generated cell assembly sequences in the rat hippocampus. *Science* **321**, 1322–1327 (2008).
55. Akhlaghpour, H. *et al.* Dissociated sequential activity and stimulus encoding in the dorsomedial striatum during spatial working memory. *Elife* **5**, (2016).
56. Balewski, Z. Z., Elston, T. W., Knudsen, E. B. & Wallis, J. D. Value dynamics affect choice preparation during decision-making. *Nat. Neurosci.* **26**, 1575–1583 (2023).
57. Pinto, L. *et al.* An Accumulation-of-Evidence Task Using Visual Pulses for Mice Navigating in Virtual Reality. *Front. Behav. Neurosci.* **12**, 36 (2018).
58. Aviel, Y., Pavlov, E., Abeles, M. & Horn, D. Synfire chain in a balanced network. *Neurocomputing* **44-46**, 285–292 (2002).
59. Diesmann, M., Gewaltig, M. O. & Aertsen, A. Stable propagation of synchronous spiking in cortical neural networks. *Nature* **402**, 529–533 (1999).
60. Hayon, G., Abeles, M. & Lehmann, D. A model for representing the dynamics of a system of synfire chains. *J. Comput. Neurosci.* **18**, 41–53 (2005).
61. Goldman, M. S. Memory without feedback in a neural network. *Neuron* **61**, 621–634 (2009).
62. Lim, S. & Goldman, M. S. Noise tolerance of attractor and feedforward memory models. *Neural Comput.* **24**, 332–390 (2012).
63. Cannon, S. C., Robinson, D. A. & Shamma, S. A proposed neural network for the integrator

of the oculomotor system. *Biol. Cybern.* **49**, 127–136 (1983).

64. Skaggs, W. E., Knierim, J. J., Kudrimoti, H. S. & McNaughton, B. L. A model of the neural basis of the rat's sense of direction. *Adv. Neural Inf. Process. Syst.* **7**, 173–180 (1995).
65. Redish, A., Elga, A. & Touretzky, D. A coupled attractor model of the rodent head direction system. *Network* **7**, 671–685 (1996).
66. Zhang, K. Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: a theory. *J. Neurosci.* **16**, 2112–2126 (1996).
67. Clark, B. J. & Taube, J. S. Vestibular and attractor network basis of the head direction cell signal in subcortical circuits. *Front. Neural Circuits* **6**, 7 (2012).
68. Ajabi, Z., Keinath, A. T., Wei, X.-X. & Brandon, M. P. Population dynamics of head-direction neurons during drift and reorientation. *Nature* **615**, 892–899 (2023).
69. Song, P. & Wang, X.-J. Angular Path Integration by Moving 'Hill of Activity': A Spiking Neuron Model without Recurrent Excitation of the Head-Direction System. *J. Neurosci.* **25**, 1002–1014 (2005).
70. Hulse, B. K. & Jayaraman, V. Mechanisms Underlying the Neural Computation of Head Direction. *Annu. Rev. Neurosci.* **43**, 31–54 (2020).
71. Turner-Evans, D. B. *et al.* The neuroanatomical ultrastructure and function of a biological ring attractor. *Neuron* **109**, 1582 (2021).
72. Turner-Evans, D. *et al.* Angular velocity integration in a fly heading circuit. *Elife* **6**, (2017).
73. Kim, S. S., Hermundstad, A. M., Romani, S., Abbott, L. F. & Jayaraman, V. Generation of stable heading representations in diverse visual scenes. *Nature* **576**, 126–131 (2019).
74. Cope, A. J., Sabo, C., Vasilaki, E., Barron, A. B. & Marshall, J. A. R. A computational model of the integration of landmarks and motion in the insect central complex. *PLoS One* **12**, e0172325 (2017).
75. Green, J. *et al.* A neural circuit architecture for angular integration in Drosophila. *Nature* **546**, 101–106 (2017).

76. Kim, S. S., Rouault, H., Druckmann, S. & Jayaraman, V. Ring attractor dynamics in the Drosophila central brain. *Science* **356**, 849–853 (2017).

77. McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I. & Moser, M.-B. Path integration and the neural basis of the 'cognitive map'. *Nat. Rev. Neurosci.* **7**, 663–678 (2006).

78. Samsonovich, A. & McNaughton, B. L. Path integration and cognitive mapping in a continuous attractor neural network model. *J. Neurosci.* **17**, 5900–5920 (1997).

79. Conklin, J. & Eliasmith, C. A controlled attractor network model of path integration in the rat. *J. Comput. Neurosci.* **18**, 183–203 (2005).

80. Ben-Yishai, R., Bar-Or, R. L. & Sompolinsky, H. Theory of orientation tuning in visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3844–3848 (1995).

81. Amari, S. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977).

82. Qi, Y. & Gong, P. Dynamic patterns in a two-dimensional neural field with refractoriness. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **92**, 022702 (2015).

83. Krishnan, N., Poll, D. B. & Kilpatrick, Z. P. Synaptic efficacy shapes resource limitations in working memory. *J. Comput. Neurosci.* **44**, 273–295 (2018).

84. Ma, H., Qi, Y., Gong, P., Lu, W. & Feng, J. Dynamics of bump attractors in neural circuits with emergent spatial correlations. *arXiv [physics.bio-ph]* (2022).

85. Burak, Y. & Fiete, I. R. Fundamental limits on persistent activity in networks of noisy neurons. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 17645–17650 (2012).

86. So, N. & Shadlen, M. N. Decision formation in parietal cortex transcends a fixed frame of reference. *Neuron* **110**, 3206–3215.e5 (2022).

87. Paton, J. J. & Buonomano, D. V. The Neural Basis of Timing: Distributed Mechanisms for Diverse Functions. *Neuron* **98**, 687–705 (2018).

88. Monteiro, T. *et al.* Manipulation of striatal population dynamics using temperature warps judgment of time. *bioRxiv* 2020.08.24.251827 (2020) doi:10.1101/2020.08.24.251827.

89. O'Reilly, J. X., Mesulam, M. M. & Nobre, A. C. The cerebellum predicts the timing of perceptual events. *J. Neurosci.* **28**, 2252–2260 (2008).

90. Rodrigues, F. S., Monteiro, T., Motiwala, A. & Paton, J. J. Dorsolateral striatum encodes a temporal basis for the organization of behavior. *bioRxiv* 2023.11.13.566826 (2023) doi:10.1101/2023.11.13.566826.

91. Zemlianova, K., Bose, A. & Rinzel, J. A biophysical counting mechanism for keeping time. *Biol. Cybern.* **116**, 205–218 (2022).

92. Salinas, E. How behavioral constraints may determine optimal sensory representations. *PLoS Biol.* **4**, e387 (2006).

93. O'Keefe, J. & Nadel, L. *The hippocampus as a cognitive map*. (Oxford university press, 1978).

94. Kumaran, D. & Maguire, E. A. The human hippocampus: cognitive maps or relational memory? *J. Neurosci.* **25**, 7254–7259 (2005).

95. Shenhav, A., Botvinick, M. M. & Cohen, J. D. The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013).

96. Alexander, W. H. & Brown, J. W. Medial prefrontal cortex as an action-outcome predictor. *Nat. Neurosci.* **14**, 1338–1344 (2011).

97. Amiez, C., Joseph, J. P. & Procyk, E. Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex* **16**, 1040–1055 (2006).

98. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).

99. Brown, J. W. & Braver, T. S. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* **307**, 1118–1121 (2005).

100. Cai, X. & Padoa-Schioppa, C. Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *J. Neurosci.* **32**, 3791–3808 (2012).

101. Carter, C. S. *et al.* Anterior cingulate cortex, error detection, and the online monitoring of

performance. *Science* **280**, 747–749 (1998).

102. Wyss, J. M. & Van Groen, T. Connections between the retrosplenial cortex and the hippocampal formation in the rat: a review. *Hippocampus* **2**, 1–11 (1992).

103. Sherrill, K. R. *et al.* Hippocampus and retrosplenial cortex combine path integration signals for successful navigation. *J. Neurosci.* **33**, 19304–19313 (2013).

104. Smith, D. M., Barredo, J. & Mizumori, S. J. Y. Complimentary roles of the hippocampus and retrosplenial cortex in behavioral context discrimination. *Hippocampus* **22**, 1121–1133 (2012).

105. Vann, S. D. & Albasser, M. M. Hippocampal, retrosplenial, and prefrontal hypoactivity in a model of diencephalic amnesia: Evidence towards an interdependent subcortical-cortical memory network. *Hippocampus* **19**, 1090–1102 (2009).

106. Kim, H., Lee, D. & Jung, M. W. Signals for previous goal choice persist in the dorsomedial, but not dorsolateral striatum of rats. *J. Neurosci.* **33**, 52–63 (2013).

107. Braun, S. & Hauber, W. The dorsomedial striatum mediates flexible choice behavior in spatial tasks. *Behav. Brain Res.* **220**, 288–293 (2011).

108. Bradfield, L. A. & Balleine, B. W. Thalamic Control of Dorsomedial Striatum Regulates Internal State to Guide Goal-Directed Action Selection. *J. Neurosci.* **37**, 3721–3733 (2017).

109. Johnson, B., Verma, R., Sun, M. & Hanks, T. D. Characterization of decision commitment rule alterations during an auditory change detection task. *J. Neurophysiol.* **118**, 2526–2536 (2017).

110. Luo, T. Z. *et al.* Transitions in dynamical regime and neural mode underlie perceptual decision-making. *bioRxiv* (2023) doi:10.1101/2023.10.15.562427.

111. Cisek, P., Puskas, G. A. & El-Murr, S. Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* **29**, 11560–11571 (2009).

112. Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N. & Pouget, A. The cost of accumulating evidence in perceptual decision making. *J. Neurosci.* **32**, 3612–3628

(2012).

113. Langdon, C. & Engel, T. A. Latent circuit inference from heterogeneous neural responses during cognitive tasks. *bioRxiv* 2022.01.23.477431 (2022) doi:10.1101/2022.01.23.477431.

114. Sorscher, B., Mel, G. C., Ocko, S. A., Giocomo, L. M. & Ganguli, S. A unified theory for the computational and mechanistic origins of grid cells. *Neuron* **111**, 121–137.e13 (2023).

115. Orhan, A. E. & Ma, W. J. Publisher Correction: A diverse range of factors affect the nature of neural representations underlying short-term memory. *Nat. Neurosci.* **22**, 505 (2019).

116. Bolkan, S. S. *et al.* Opponent control of behavior by dorsomedial striatal pathways depends on task demands and internal state. *Nat. Neurosci.* **25**, 345–357 (2022).

117. Pinto, L. *et al.* Task-Dependent Changes in the Large-Scale Dynamics and Necessity of Cortical Regions. *Neuron* **104**, 810–824.e9 (2019).

118. Rickgauer, J. P., Deisseroth, K. & Tank, D. W. Simultaneous cellular-resolution optical perturbation and imaging of place cell firing fields. *Nat. Neurosci.* **17**, 1816–1824 (2014).

119. Pégard, N. C. *et al.* Three-dimensional scanless holographic optogenetics with temporal focusing (3D-SHOT). *Nat. Commun.* **8**, 1228 (2017).

120. Adesnik, H. & Abdeladim, L. Probing neural codes with two-photon holographic optogenetics. *Nat. Neurosci.* **24**, 1356–1366 (2021).

121. Aksay, E. *et al.* Functional dissection of circuitry in a neural integrator. *Nat. Neurosci.* **10**, 494–504 (2007).

122. Fisher, D., Olasagasti, I., Tank, D. W., Aksay, E. R. F. & Goldman, M. S. A modeling framework for deriving the structural and functional architecture of a short-term memory microcircuit. *Neuron* **79**, 987–1000 (2013).

123. Gonçalves, P. J., Arrenberg, A. B., Hablitzel, B., Baier, H. & Machens, C. K. Optogenetic perturbations reveal the dynamics of an oculomotor integrator. *Front. Neural Circuits* **8**, 10 (2014).

124. Seung, H. S. How the brain keeps the eyes still. *Proc. Natl. Acad. Sci. U. S. A.* **93**,

13339–13344 (1996).

125. Koay, S. A. Sequential and efficient neural-population coding of complex task information. (2021) doi:10.5061/dryad.cvdncjt53.

126. Aronov, D. & Tank, D. W. Engagement of neural circuits underlying 2D spatial navigation in a rodent virtual reality system. *Neuron* **84**, 442–456 (2014).

127. Wekselblatt, J. B., Flister, E. D., Piscopo, D. M. & Niell, C. M. Large-scale imaging of cortical dynamics during sensory perception and behavior. *J. Neurophysiol.* **115**, 2852–2866 (2016).

128. Jewell, S. W., Hocking, T. D., Fearnhead, P. & Witten, D. M. Fast nonconvex deconvolution of calcium imaging data. *Biostatistics* **21**, 709–726 (2020).

129. Fleming, W., Jewell, S., Engelhard, B., Witten, D. M. & Witten, I. B. Inferring spikes from calcium imaging in dopamine neurons. *PLoS One* **16**, e0252345 (2021).

130. Harris, K. D. Nonsense correlations in neuroscience. *bioRxiv* 2020.11.29.402719 (2021) doi:10.1101/2020.11.29.402719.

SUPPLEMENTARY INFORMATION

SUPPLEMENTARY FIGURES

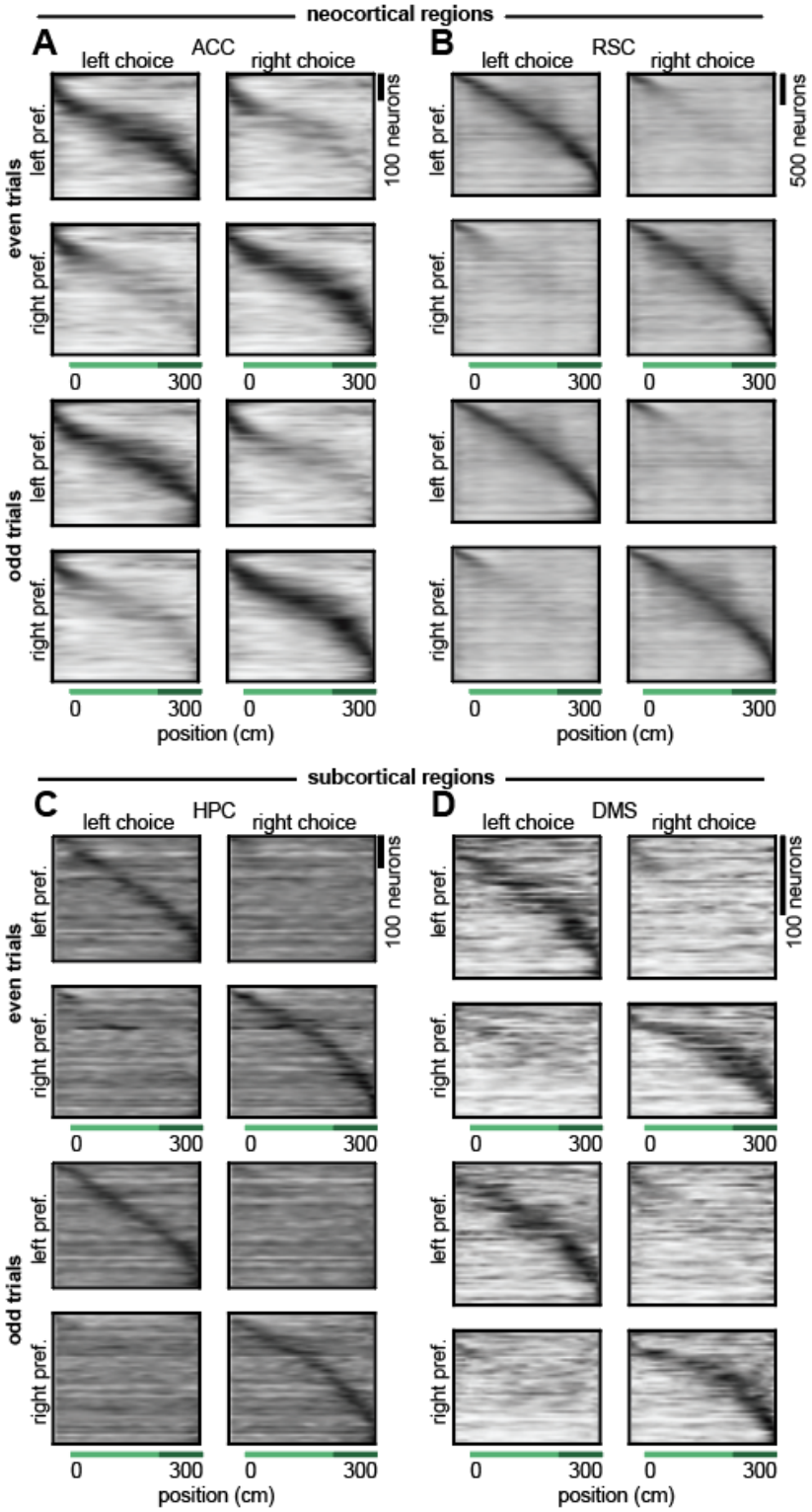

Figure S1. Choice-selective sequences across brain regions during the accumulating towers task appear similar on even and odd trials. (A) Each row shows the peak-normalized estimated firing rate

(based on the odd trials) at each position in the maze of a neuron with significant evidence tuning (see Methods) recorded during the accumulating towers task from ACC (n=964 neurons), averaged across even trials (top) or odd trials (bottom) when the animal turned left (left choice, left column) or right (right choice, right column). Neurons were divided based on their choice-selectivity (see Methods) and ordered based on the position of peak activity. **(B-D)** Same as (A) but for RSC (n = 4190 neurons; B), HPC (n = 791 neurons; C), and DMS (n = 322 neurons; D).

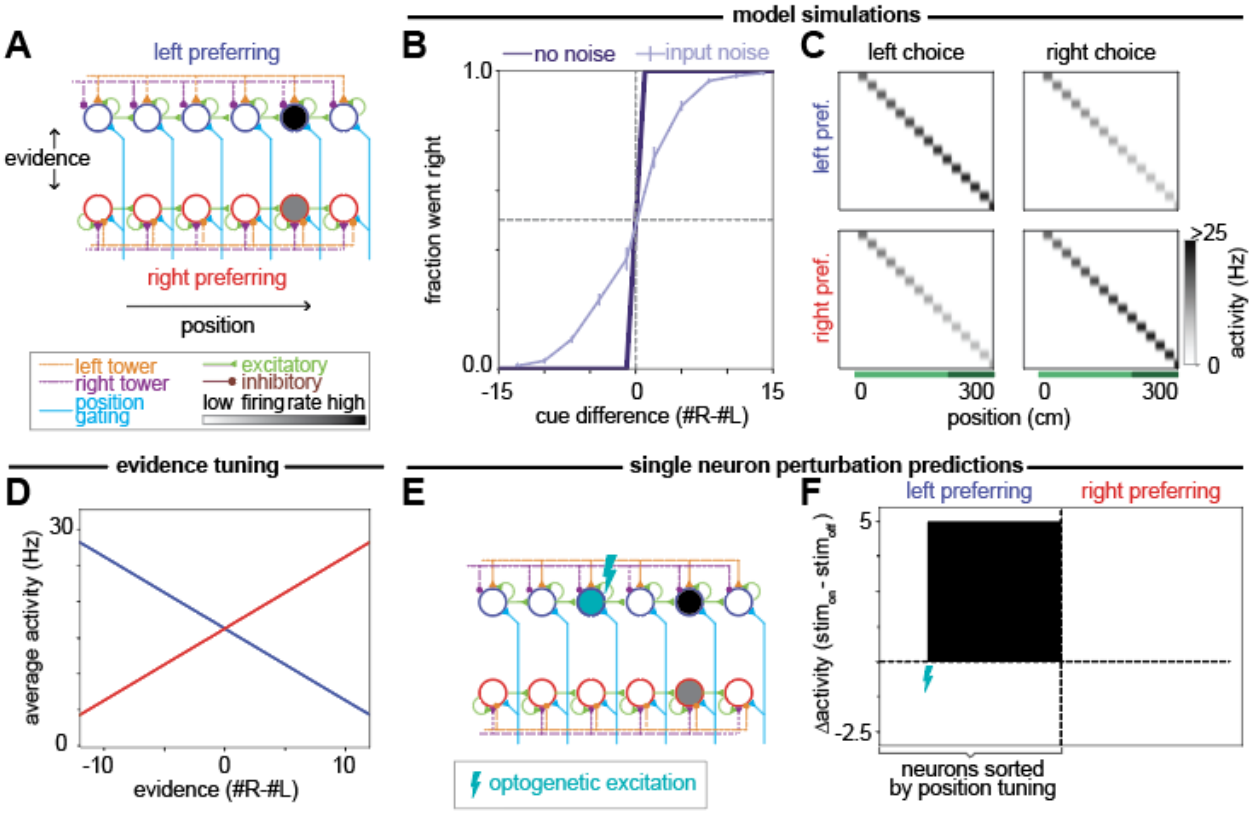

Figure S2. Uncoupled chains model of evidence accumulation through sequences. (A) Schematic of neural circuit architecture for the uncoupled chains model, showing excitatory (green) connections between neurons (circles) as well as inputs to the circuit from the external left (orange) and right (purple) towers and a position gating signal (cyan). **(B)** Psychometric data for model simulated trials describing how often the amplitude of the final neuron in the left chain was greater than that of the final neuron in the right chain for cases in which the model was simulated with (light purple) and without (dark purple) noise in the input. Error bars show s.e.m. **(C)** Each row shows the non-normalized amplitude of a model neuron at each position in the maze, averaged across simulated trials without input noise when the greater final amplitude was in the left (left choice, left column) or right (right choice, right column) chain. Neurons were divided based on their choice-selectivity (see Methods) and ordered based on the position of peak activity. **(D)** Tuning curves of individual neurons to evidence by finding the average activity for different evidence levels at the neuron's peak position, for left-preferring (blue) and right-preferring (red) cells. **(E)** Schematic of a single neuron perturbation experiment in which optogenetic excitation is applied to a single neuron in the left chain. **(F)** Simulated changes of the firing rates of all neurons in the absence of cues when a single cell (denoted with the laser as in E) is optogenetically stimulated.

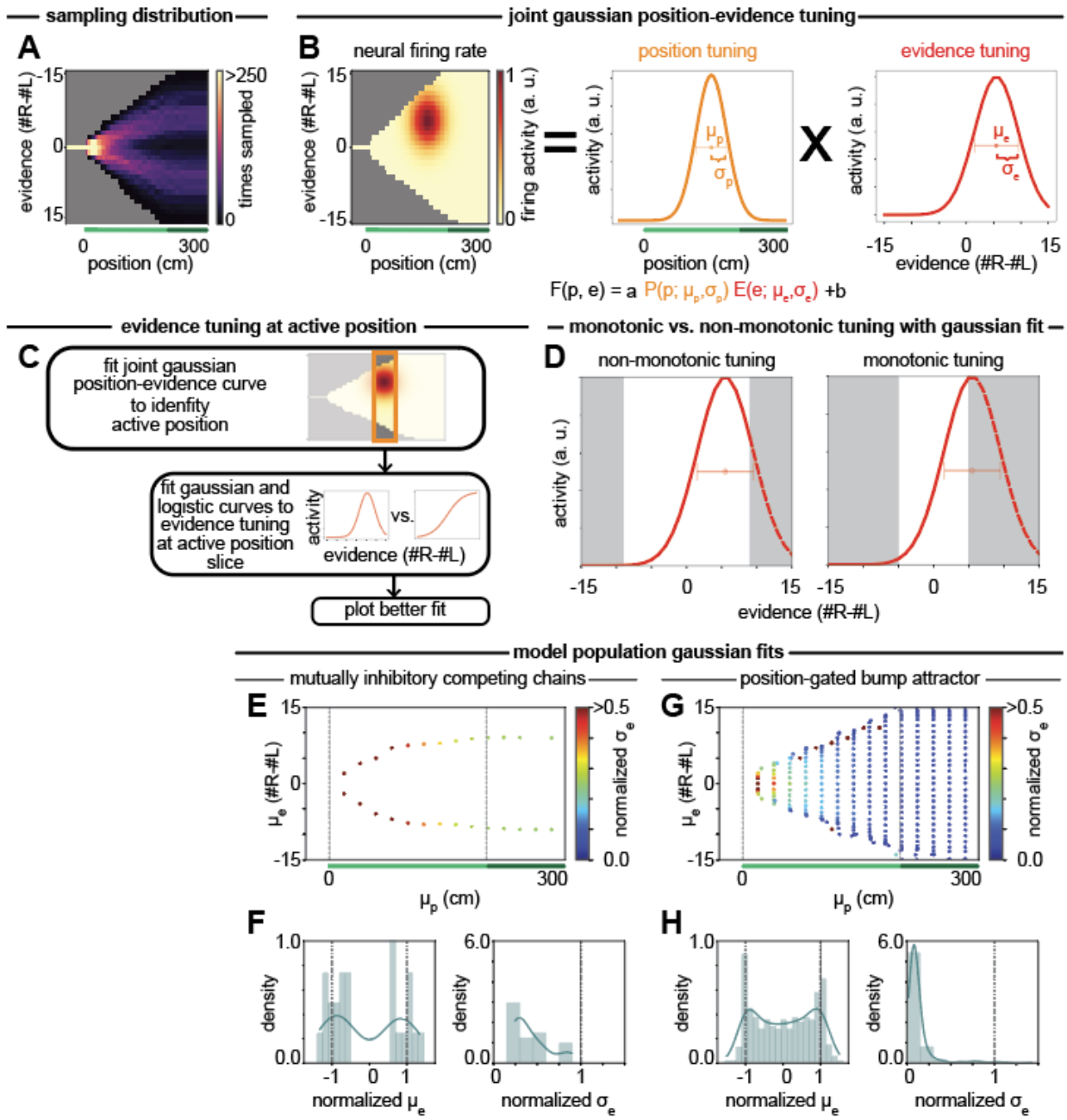

Figure S3. Joint position-evidence tuning curve fitting procedure to compare monotonicity in simulated and recorded neurons. (A) Number of times each position-evidence bin is sampled within the 1000 empirical trials used to simulate the models. The statistics of cue counts, distribution, and positioning used to generate trials make it such that at different positions, different evidence levels are not sampled uniformly. For example, cells tuned to the beginning of the maze only see a small range of small evidence levels at the positions where they are most active; such a cell may not appear to respond to high levels of evidence because its position tuning is never active at these extreme evidence levels. **(B)** Fitting firing rates to a joint position-evidence Gaussian to identify μ_p, σ_p, μ_e, and σ_e. **(C)** Schematic of procedure for plotting evidence tuning curves by first identifying the active position region and then comparing the fit between a Gaussian and logistic curve. **(D)** Examples of Gaussian tuning to produce non-monotonic tuning (left) and monotonic tuning (right), with the monotonicity of the curve dependent on

the observed evidence range (gray regions indicate non-observed evidence levels). **(E)** Scatter plot showing the location of the fit mean position (μ_p) and fit mean evidence (μ_e) with color indicating the normalized fit evidence standard deviation (σ_e) (see Methods) of the simulated data for neurons in the mutually inhibitory competing chains model. Dashed lines indicate the boundaries of the cue period. **(F)** Histograms showing the normalized fit mean evidence (left) and the normalized fit evidence standard deviation (right) for the neurons shown in (E). **(G-H)** Same as for (E-F) but for simulated neurons from the position-gated bump attractor.

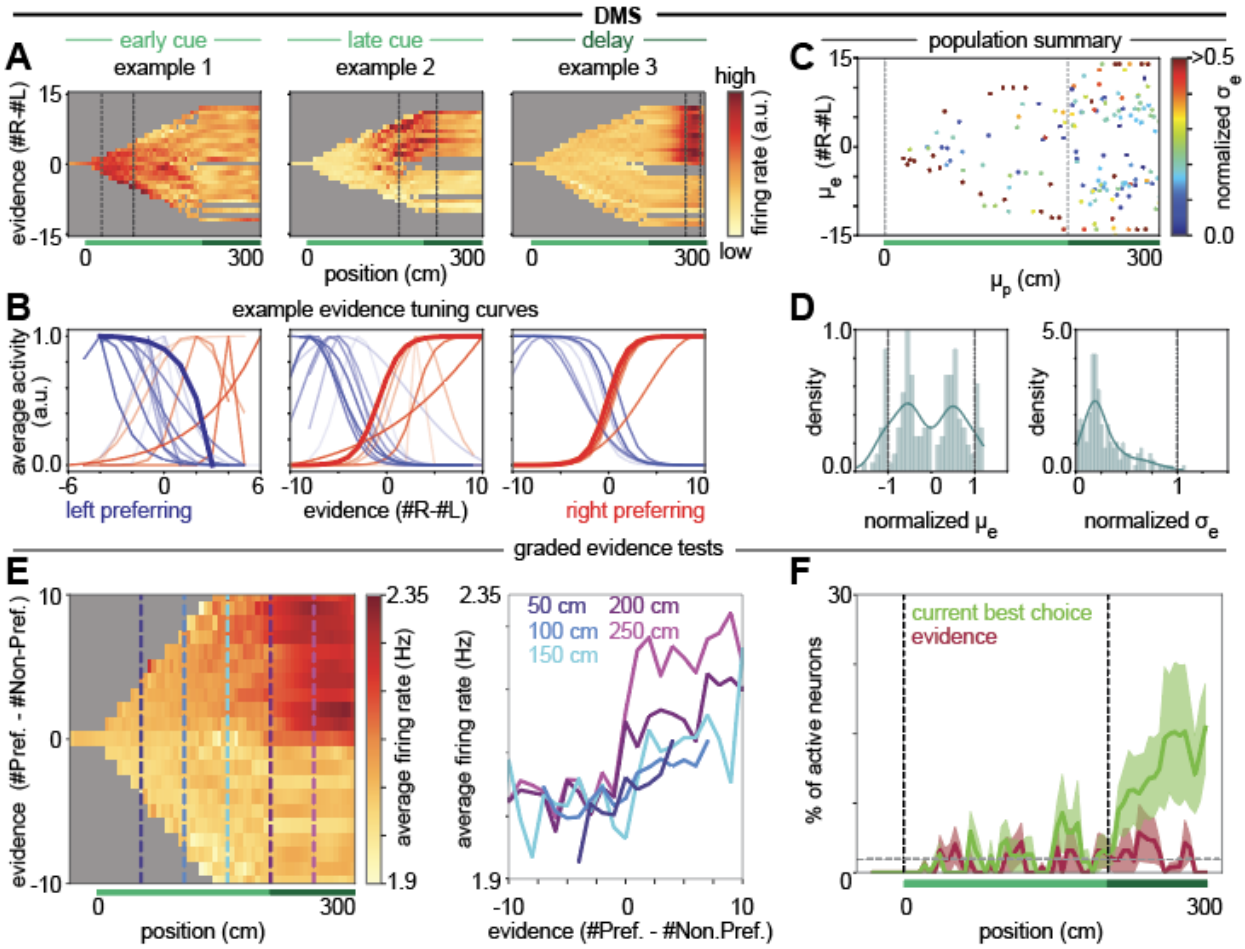

Figure S4. DMS exhibits mainly monotonic evidence tuning dominated by neurons with position tuning in the delay region. (A) Heatmaps showing the average firing in position by evidence bins for example individual neurons in DMS with mean position in the early cue (left), late cue (middle), or delay (right) region of the maze. Gray bins denote regions for which there were fewer than 2 samples during the session. **(B)** Example DMS evidence tuning curves fit to the region of the neuron's peak activity (defined as observations within 0.5 position standard deviation of the neuron's peak position tuning; see Methods) for a collection of neurons with mean position in the early cue (left), late cue (middle), or delay (right) region of the maze. Red coloring indicates neurons classified as right preferring, and blue indicates left preferring. Bold lines correspond to the examples in (A), where the neuron's region of peak activity is the region between the dashed vertical lines. **(C)** Scatter plot showing the location of the fit mean position (μ_p) and fit mean evidence (μ_e) with color indicating the normalized fit evidence standard deviation (σ_e) (see Methods) of the 80% of neurons recorded in DMS with the best fit between the neural data and the model predictions. Dashed lines indicate the boundaries of the cue period. **(D)** Histograms showing the normalized fit mean evidence (left) and the normalized fit evidence standard deviation (right) for the neurons shown in (C). **(E)** Left: Heatmap showing average firing rates across all evidence-tuned neurons in position by preferred-evidence bin. Gray indicates bins that were not sampled more than once on any session. Right: Cross-sections of the heatmap at a point in the early cue period (50 cm, light green), late cue period (150 cm, medium green), and delay period (250 cm, dark green), showing average firing rate across evidence-tuned neurons as a function of preferred evidence. **(F)** For each position, the fraction of all active neurons (defined as neurons with fit position mean within one position standard

deviation of the location) at each location where the linear evidence encoding model had significant coefficients for choice (crimson) and evidence (lime), based on the F-statistic.

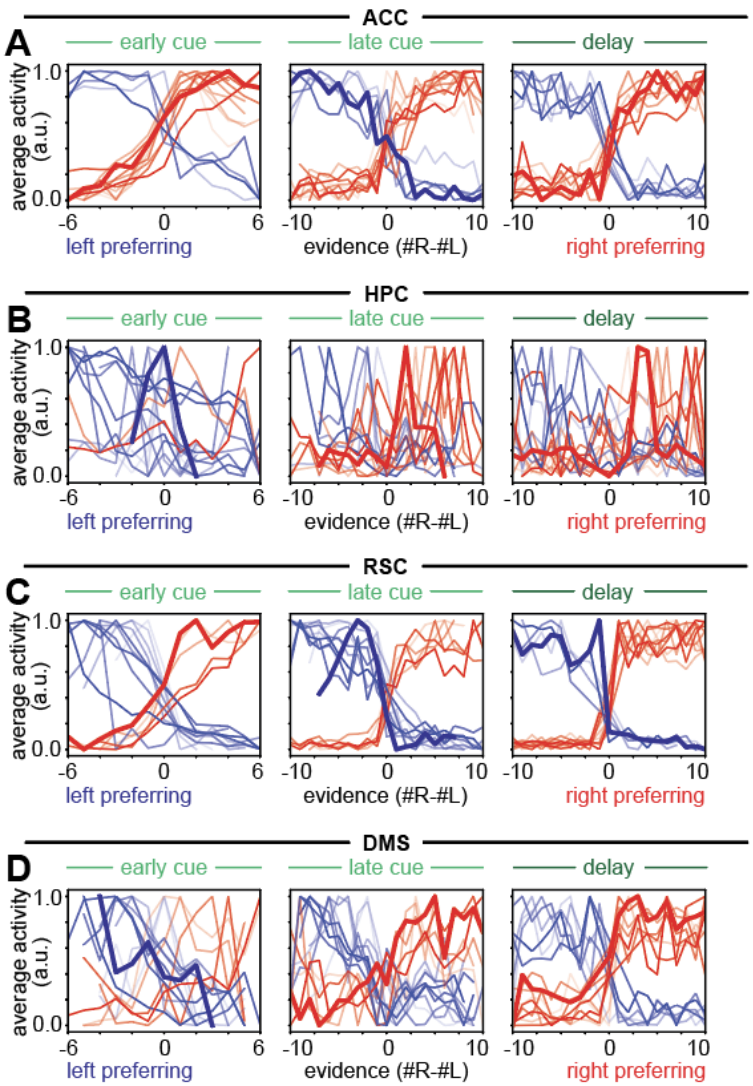

Figure S5. Raw tuning curves demonstrate the same patterns as fit tuning curves throughout the brain. (A) Average activity of example ACC neurons at different evidence levels in the region of the neuron's peak activity (defined as observations within 0.5 fit position standard deviation of the neuron's fit position; see Methods) for a collection of neurons with mean position in the early cue (left), late cue (middle), or delay (right) region of the maze. Red coloring indicates neurons classified as right preferring, and blue indicates left preferring. Bold lines correspond to the examples plotted in Figure 5. **(B)** Same as (A) but for HPC. **(C)** Same as (A) but for RSC. **(D)** Same as (A) but for DMS, with the examples from Figure S4.

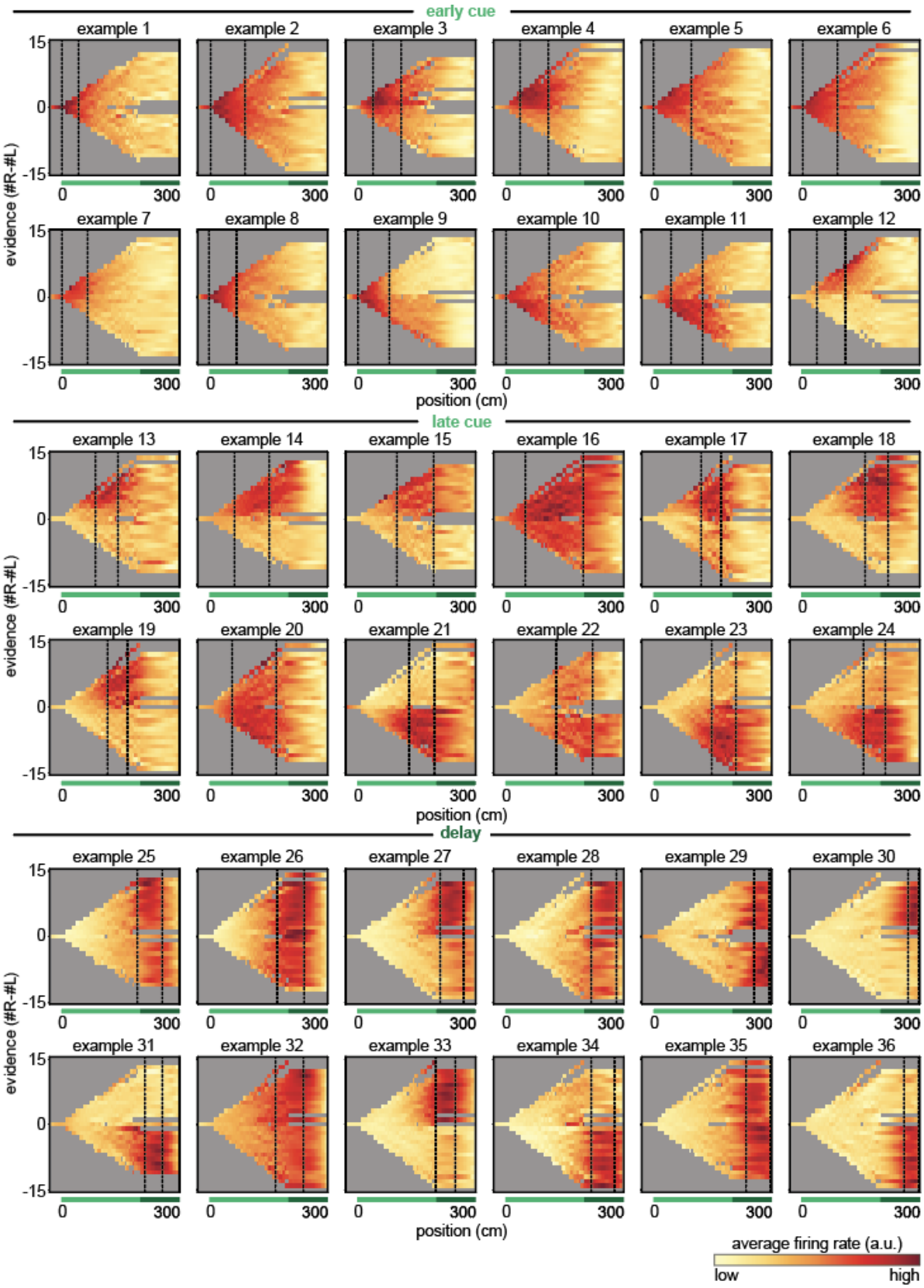

Figure S6. Example evidence-tuned neurons in ACC. Heatmaps of average normalized firing activity

within position by evidence bins of 36 example individual neurons from ACC with significant evidence tuning (see Methods). Gray denotes position-evidence bins where there were fewer than 2 observations. A representative selection of neurons was made from neurons with mean position in the early cue (0-100 cm), late cue (100-200 cm), and delay (200-300 cm) regions of the maze.

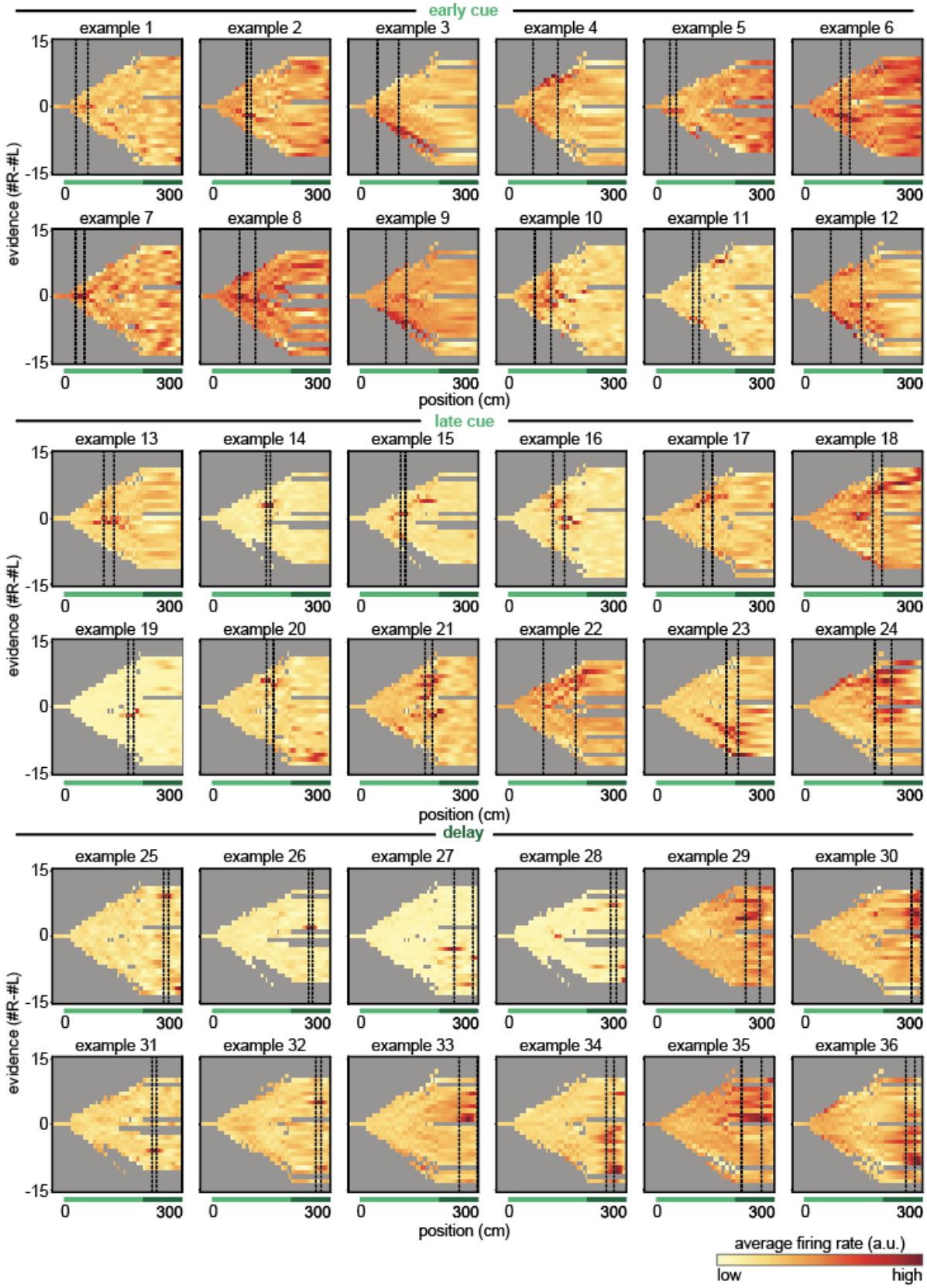

Figure S7. Example evidence-tuned neurons in HPC. Same as in Figure S6 but for neurons from HPC.

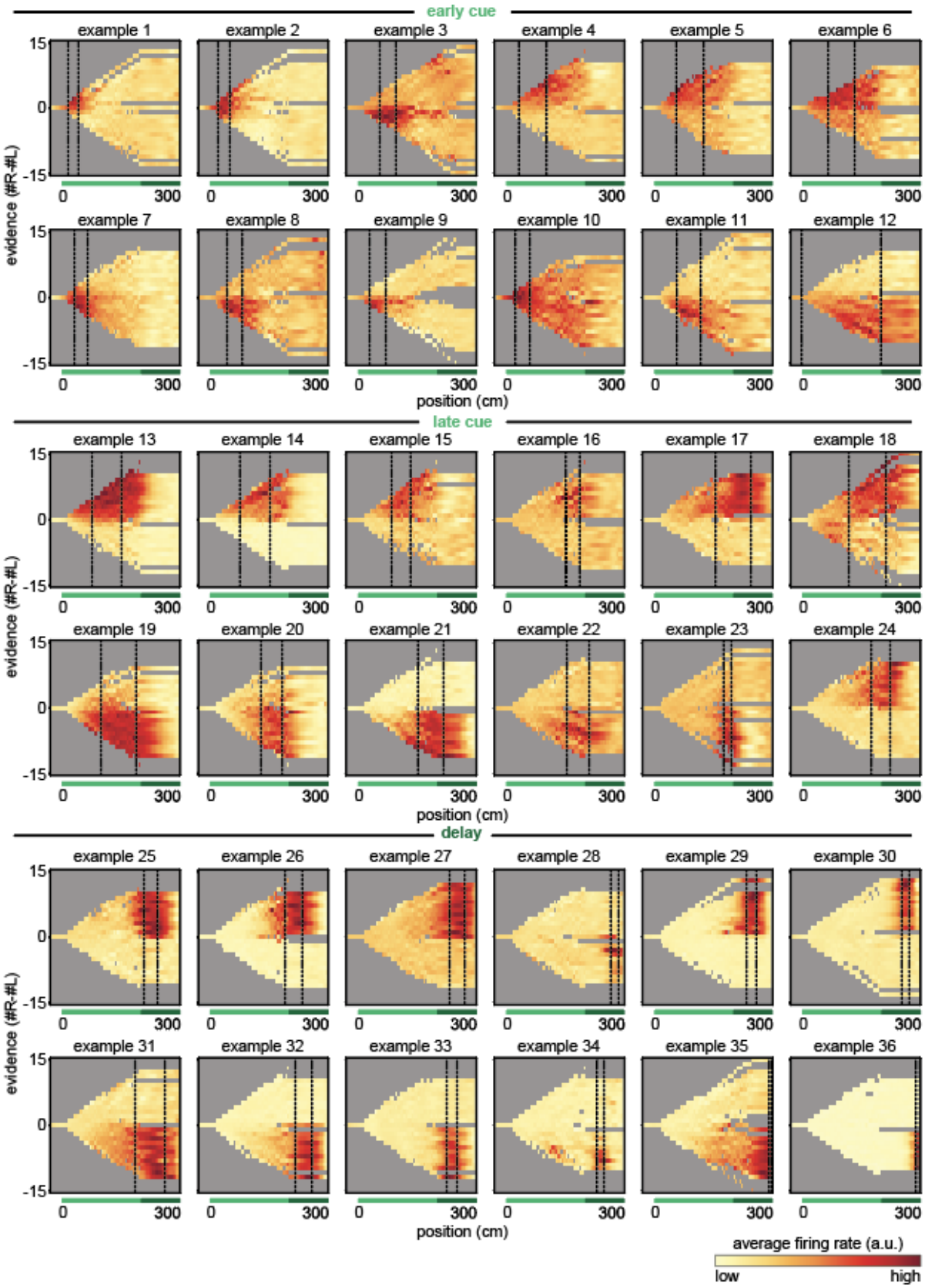

Figure S8. Example evidence-tuned neurons in RSC. Same as in Figure S6 but for neurons from RSC.

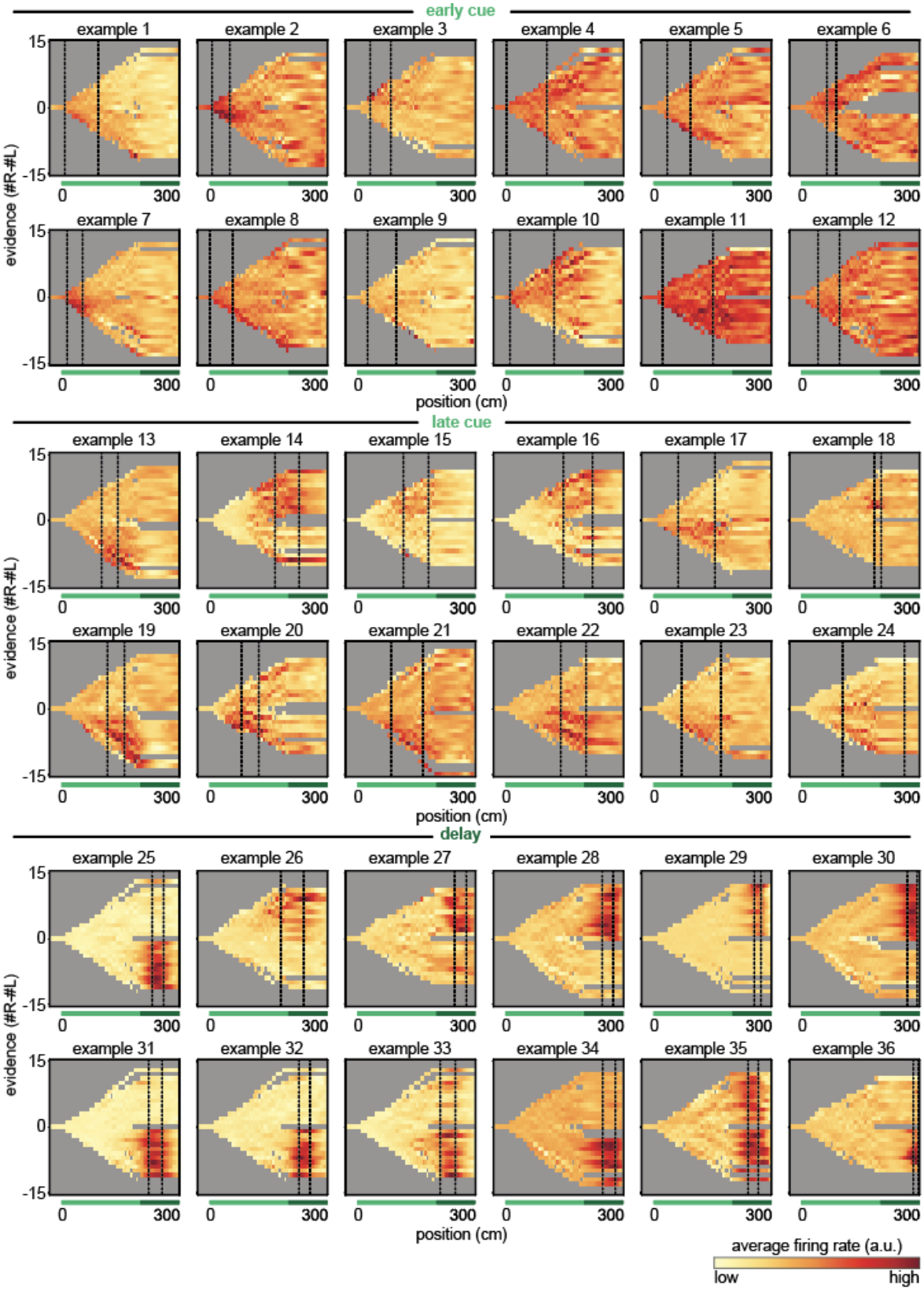

Figure S9. Example evidence-tuned neurons in DMS. Same as in Figure S6 but for neurons from DMS.

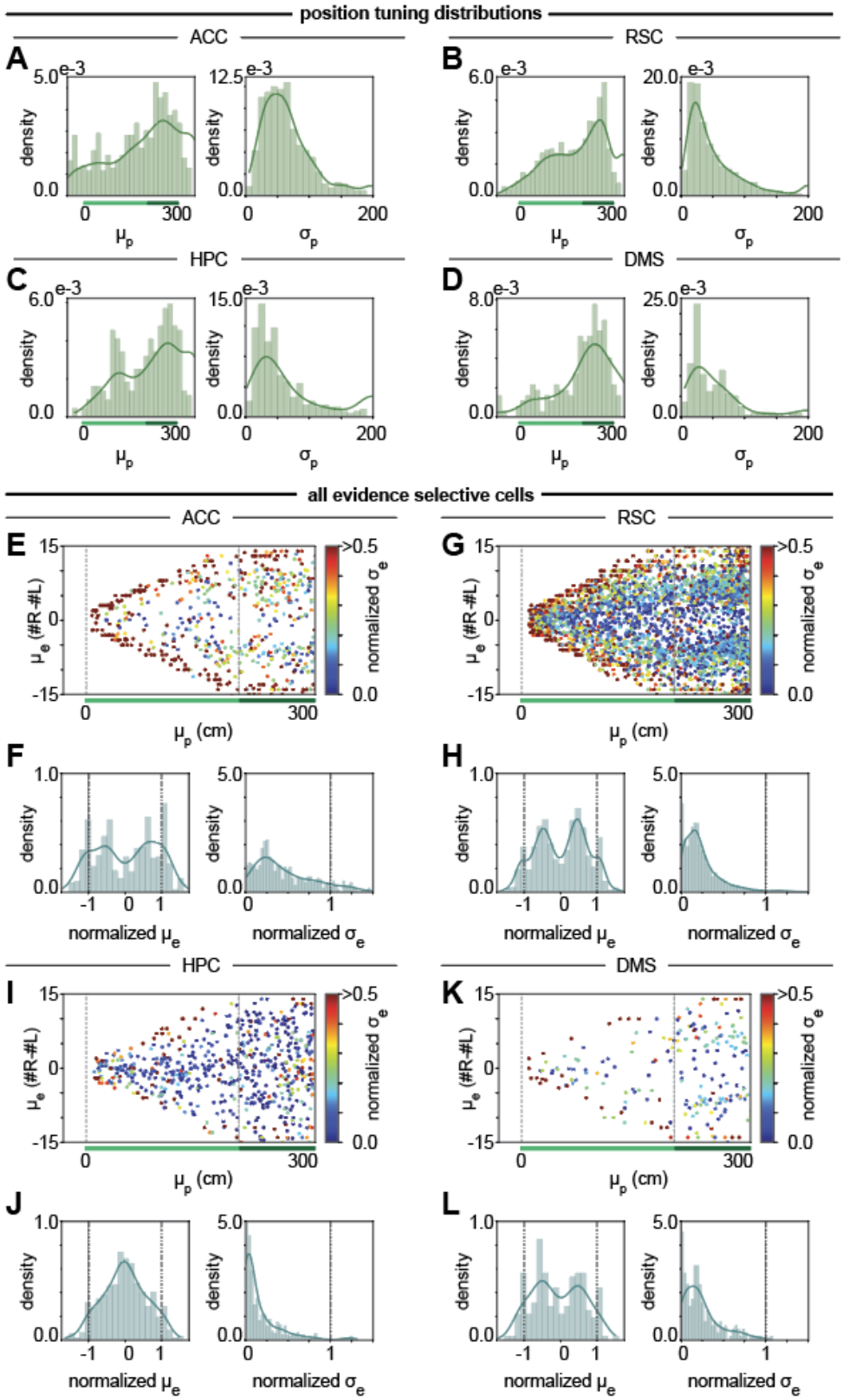

Figure S10. Supplementary distribution plots. (A) Left: Distribution of μ_p for the cells in ACC plotted in Figure 5. Right: Distribution of σ_p for the same cells. **(B-D)** Same as (A) but for cells in RSC (B), HPC (C),

and DMS (D) where DMS cells are plotted in Figure S4. **(E)** Scatter of fit position mean vs. fit evidence mean colored by normalized σ_e for all significantly evidence-tuned cells in ACC. **(F)** Left: Distribution of normalized μ_e for the cells in (E). Right: Distribution of normalized σ_e for the cells in (E). **(G-H)** Same as (E-F) but for RSC. Note that, by comparison to Figure 5, the narrowly tuned cells in RSC (deep blue) tend to have poorer fits. **(I-J)** Same as (E-F) but for HPC. **(K-L)** Same as (E-F) but for DMS.

SUPPLEMENTARY TEXT

The chains-based models represent evidence in the activity of two chains of neurons, one that increases its firing in response to right evidence and one that increases its firing in response to left evidence. In principle, these chains could be completely independent, with each chain integrating the cues only to its respective side, but this would result in chains that only increase in amplitude across the trial, inconsistent with observed data. Instead, we consider architectures with ipsilateral excitation and contralateral inhibition, where the contralateral inhibition comes either from external inputs or from the other chain. Specifically, we consider: (i) a model where the chains are uncoupled but towers appearing to the ipsilateral side are excitatory and towers appearing to the contralateral side are inhibitory (the uncoupled chains model, Fig. S2A) and (ii) a model where each chain receives input only from the ipsilateral side and competes with the other chain through mutual inhibition within the same position (the mutually inhibitory competing chains model, Fig. 3A).

Basic Features of Chains-Based Models

The neurons in each model can be represented by the same basic set of differential equations, where the firing rate of the ith neuron in the left chain, $r_{i,L}$, evolves according to

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [br_{i,L} + cr_{i-1,L} - er_{i,R} + f\mathbf{1}_{left}(t) - g\mathbf{1}_{right}(t) + P_i(t) - T]^+ \quad (S1)$$

and similarly for the firing rate of neurons in the right chain. The various terms on the right side of the equation are described in the main text (see text below Eq. (1)), although here we add a term $g\mathbf{1}_{right}(t)$ that allows for inhibitory external inputs from cues on the contralateral side. The external cue inputs are given by $\mathbf{1}_{left}(t)$ for the left inputs and $\mathbf{1}_{right}(t)$ for the right inputs, which take the form

$$\mathbf{1}_{\text{left}}(t) = \begin{cases} 1 & \min_{\ell \in L} |t - \ell| < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\mathbf{1}_{\text{right}}(t) = \begin{cases} 1 & \min_{R \in \mathcal{R}} |t - R| < 0.5 \\ 0 & \text{otherwise} \end{cases}$$

where L is the set of times of left cues and $\mathcal{R}$ is the set of times of right cues.

The position signal is given by

$$P_i(t) = \begin{cases} T + X(t) & i \leq \frac{p(t)}{P_0} < (i+1) \\ 0 & \text{otherwise} \end{cases},$$

where $X(t) \geq 0$ and $p(t)$ is the position of the animal along the maze, so that each neuron receives an active position signal for length P_0, and only one neuron in each chain receives an active signal at any time.

For both models, we require the chains to exhibit a number of properties in order to accurately accumulate evidence, so that at the end of a trial, the chain corresponding to the side with more cues will have greater amplitude. In particular, we look to achieve the following properties: (i) Each neuron is only active around a specific position, tiling space, with activity decaying away when the simulated animal leaves this position. (ii) External cue inputs are

integrated linearly. (iii) In the absence of external cue inputs, when at a fixed position, the difference in neural activity of the two chains remains constant, without the sum growing unboundedly. (iv) In the absence of external cue inputs, the difference in amplitude of the two chains is preserved across positions. We analyze the necessary conditions on the model parameters to achieve each of these properties.

Uncoupled Chains Model

For the case of the uncoupled chains model (Fig. S2), we take *e*=0, giving us

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [br_{i,L} + cr_{i-1,L} + f\mathbf{1}_{left}(t) - g\mathbf{1}_{right}(t) + P_i(t) - T]^+ \quad (S2)$$

We consider below the conditions needed for this model to achieve the desired properties.

Neurons Active Only Around a Specific Position

Activity only occurring around a specific position when the simulated animal is navigating down the maze is achieved through the position-gating mechanism. We set the threshold *T* of the neuron sufficiently high that the term in square brackets will only affect the firing rate of the neuron in the presence of the position gating input $P_i(t)$. Otherwise, neural activity decays exponentially. This leads to the transient, sequential responses.

External Cues Integrated Linearly

In order for each cue to be integrated linearly, first, each cue should be weighted equally in its contribution to the final activity of the chain. This is achieved by giving each cue the same form of input in $\mathbf{1}_{left}(t)$ and $\mathbf{1}_{right}(t)$.

Second, this signal must be fully integrated by the chain. Since the external cue input term resides in the square brackets, only the neurons with the active position gating signal can integrate this signal. Without loss of generality, we consider the left-side chain. For the left-side neuron with an active position gating signal,

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [br_{i,L} + cr_{i-1,L} + f\mathbf{1}_{left}(t) - g\mathbf{1}_{right}(t) + X(t)]^+.$$

For this neuron to integrate its inputs, we need the term inside square brackets to be positive. We can see that this is guaranteed if $br_{i,L} + cr_{i-1,L}$ is greater than $-g$. We show below that the difference in activity of the chains is preserved across positions in the absence of external cues when $b=c=a$, so that $r_{i,L}+r_{i-1,L} = C$ for some constant C. Thus, by initializing the activity at the start of the trial, $r_{0,L}$, to be sufficiently large that the activity of the lower firing rate chain does not drop below g/a, these inputs will always be integrated.

We note that the argument above does not require the cues to have the square pulse specified here. Rather, it only requires that each cue has the same form.

Difference in Activity Between Chains Remains Constant at a Fixed Position

When the animal maintains a fixed position, the activity of the neurons at the preceding position decay away, so that in the absence of external cue inputs, we have

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [br_{i,L} + P_i(t) - T]^+.$$

If we take $X(t)=0$, we have $P_i(t) = T$ at the active position, and thus

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [br_{i,L}]^+ = -ar_{i,L} + br_{i,L}.$$

Thus, when *a*=*b*, each chain maintains constant activity, both preserving the difference between the chains and keeping the sum constant.

If instead $X(t) > 0$, this signal will be integrated, causing both chains to increase in amplitude. Such a signal could be designed to act as an urgency signal, but must be carefully crafted to prevent the chains from growing too rapidly.

Difference in Amplitude Preserved Across Positions

Preserving the amplitude of the chain across positions is a special case of preserving the difference across positions. To preserve amplitude across positions in the absence of cues, the next neuron in the chain must integrate the activity of the previous neuron at the same rate at which it decays. At the active position in the absence of external cue inputs, and assuming $X(t)=0$, we have

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [ar_{i,L} + cr_{i-1,L}]^+$$

$$\frac{dr_{i,L}}{dt} = cr_{i-1,L}$$

while we also have

$$\frac{dr_{i-1,L}}{dt} = -ar_{i-1,L}$$

To preserve the total amplitude of activity,

$$-\frac{dr_{i-1,L}}{dt} = \frac{dr_{i,L}}{dt},$$

giving us the condition *a*=*c*.

We note that neuron *i* will only fully integrate the activity of the previous position in the infinite time limit. Due to the transient nature of the position-gating signal, we have that the final amplitude of the ith neuron as a function of the position signal duration P_0 will be

$$r_{i,L}(P_0) = \int_0^{P_0} ar_{i-1,L}dt = \int_0^{P_0} ar_{i-1,L}(0)e^{-at}dt = r_{i-1,L}(0)\left(1 - e^{-aP_0}\right).$$

For *a* and P_0 sufficiently large, the difference in amplitude between positions will be exponentially small.

Parameterization of the Uncoupled Chains Model

In our simulations, in addition to taking $X(t)$ = 0 Hz/s, we set a =1 s^{-1}, b =1 s^{-1}, c = 1 s^{-1}, f = 1 s^{-1}, g = 1 s^{-1}, P_0 = 20 cm, and T = 300 Hz/s. Neurons $r_{0,L}$ and $r_{0,R}$ are initialized to 16.25 Hz activity level. We note that the neurons must be initialized to a sufficiently high non-zero activity level such that an external input to the contralateral side can be negatively integrated into the activity.

Mutually Inhibitory Competing Chains Model

In this model, the chains compete through mutual inhibition, rather than receiving opposing inputs. For the mutually inhibitory competing chains model, we take g=0 in Equation S1, giving

$$\frac{dr_{i,L}}{dt} = -ar_{i,L} + [br_{i,L} + cr_{i-1,L} - er_{i,R} + f\mathbf{1}_{left}(t) + P_i(t) - T]^+ .$$

We now turn to the necessary conditions for this model to achieve the desired properties.

Neurons Active Only Around a Specific Position

Activity only occurring around a specific position when the simulated animal is navigating down the maze is achieved through the same position gating mechanism as before, with T much larger than the other terms in square brackets so that the neuron will only be active when the position gate $P_i(t)$ is active.

External Cues Integrated Linearly

Each cue should be weighted equally in its contribution to the final activity of the chain. This is achieved by giving each cue the same form of input in $\mathbf{1}_{left}(t)$ and $\mathbf{1}_{right}(t)$. Moreover, the weights of mutual inhibition and self-excitation must be tuned such that the system of equations is a perfect integrator when the term in square brackets is above threshold. The position gating ensures that this can be true for at most two neurons, both at the same position. Specifically, when the term in square brackets is above threshold, we have

$$\frac{d}{dt}\begin{pmatrix} r_{i,L} \\ r_{i,R} \end{pmatrix} = \begin{pmatrix} b-a & -e \\ -e & b-a \end{pmatrix}\begin{pmatrix} r_{i,L} \\ r_{i,R} \end{pmatrix} + \begin{pmatrix} cr_{i-1,L} + f\mathbf{1}_{left}(t) + X(t) \\ cr_{i-1,R} + f\mathbf{1}_{right}(t) + X(t) \end{pmatrix},$$

which has eigenvalues (b-a-e) for the common mode eigenvector, (1, 1), and (b-a+e) for the difference mode eigenvector, (1, -1). For a perfect integrator that accumulates evidence in the difference of firing between the two chains, we require that the eigenvalue of the difference mode eigenvector is zero, giving $e = a - b$. We note that, for the assumption that both neurons are above threshold to be true, we require $X(t)+br_{i,L}+cr_{i-1,L} - er_{i,R} > 0$ and $X(t)+br_{i,R}+cr_{i-1,R} - er_{i,L} > 0$. This can be achieved by setting $X(t) = I_{ext}$ for some sufficiently large constant I_{ext}, and integration will only be perfect in the range that this is true. This common external input to both chains serves as a background about which integration occurs at each position in the chain.

Difference in Activity Between Chains Remains Constant at a Fixed Position

Holding the difference in activity between the neurons in the chain constant at a constant position is a natural consequence of having a perfect integrator for the difference mode eigenvector. Furthermore, because the common mode eigenvector is associated with a negative (decay-associated) eigenvalue, the sum of activity will not grow uncontrollably.

Difference in Amplitude Preserved Across Positions

To appropriately represent the difference in accumulated evidence between the chains, we require the difference $\Delta_i = r_{i,L} - r_{i,R}$ between the chains to be preserved between positions. This difference evolves over time according to

$$\frac{d\Delta_i}{dt} = -ar_{i,L} + br_{i,L} + cr_{i-1,L} - er_{i,R} - \left(-ar_{i,R} + br_{i,R} + cr_{i-1,R} - er_{i,L}\right)$$

$$\frac{d\Delta_i}{dt} = (-a + b + e)r_{i,L} + (a - b - e)r_{i,R} + c(r_{i-1,L} - r_{i-1,R})$$

$$\frac{d\Delta_i}{dt} = (-a + b + e)\Delta_i + c\Delta_{i-1}$$

From our conditions on perfect integration, we have (-a+b+e)=0, giving us

$$\frac{d\Delta_i}{dt} = c\Delta_{i-1}.$$

We also have that

$$\frac{d\Delta_{i-1}}{dt} = -ar_{i-1,L} + ar_{i-1,R} = -a\Delta_{i-1}$$

For the total difference in chains to be preserved, we require

$$-\frac{d\Delta_{i-1}}{dt} = \frac{d\Delta_i}{dt},$$

giving the condition *a*=*c*.

As in the case of the uncoupled chains model, due to the finite time of integration, we have that

$$\Delta_i(P_0) = \Delta_{i-1}(0)\left(1 - e^{-aP_0}\right)$$

so that the difference between positions will be exponentially small for *a* and P_0 sufficiently large.