Non-canonical attractor dynamics underlie perceptual decision-making

Thomas Zhihao Luo*#, Timothy Doyeon Kim*#, Diksha Gupta, Adrian G. Bondy, Charles D. Kopec, Verity A. Elliot, Brian DePasquale, Carlos D. Brody#

*equal contribution
#Correspondence should be addressed to: T.Z.L. zhihaol@princeton.edu, T.D.K. tdkim@princeton.edu, or C.D.B. brody@princeton.edu

Abstract

Perceptual decision-making is the process by which an animal uses sensory stimuli to choose an action or mental proposition. This process is thought to be mediated by neurons organized as attractor networks[1,2]. However, whether attractor dynamics underlie decision behavior and the complex neuronal responses remains unclear. Here we use an unsupervised, deep learning-based method to discover decision-related dynamics from the simultaneous activity of neurons in frontal cortex and striatum of rats while they accumulate pulsatile auditory evidence. We show that contrary to prevailing hypotheses, perceptual choices emerge from the dynamics driven by sensory inputs that are not aligned to discrete attractors in the input-independent dynamics. Input-driven and -independent dynamics differ in strength across the decision state space, resulting in the input-driven dynamics playing a dominant role in evidence integration, while input-independent dynamics playing a principal role in decision commitment. An extension of the classic drift-diffusion hypothesis[3] to approximate the non-canonical attractor dynamics precisely predicts the internal decision commitment time and captures diverse and complex single-neuron temporal profiles, such as ramping and stepping[4–6]. It also captures choice behavior and trial-averaged curved trajectories[7–9] and reveals distinctions between brain regions. Thus, non-canonical attractor dynamics inferred from unsupervised discovery conceptually extend a classic hypothesis and parsimoniously account for multiple neural and behavioral phenomena.

Theories of attractor dynamics have been successful at capturing multiple brain functions[10], including motor planning[11] and neural representations of space[12]. Attractors are a set of states toward which a system tends to evolve, from a variety of starting positions. In these theories, the computations of a brain function are carried out by the temporal evolution, or the dynamics, of the system. Experimental findings support the idea that the brain uses systems with attractor states for computations underlying working memory[11] and navigation[12]. These theories often focus on the low-dimensional nature of neural population activity[7,8,13,14], and account for the responses across a large number of neurons using a dynamical system model whose variable has only a few dimensions[15,12,16].

Attractor models have also been hypothesized to underlie perceptual decision-making: the process through which noisy sensory stimuli are categorized to select an action or mental proposition. In these hypotheses, the dynamics of a low dimensional variable carry out the computations needed in decision formation[1,7,17,18], such as integrating sensory evidence and committing to a choice. While some experimental evidence favor a role of attractors in perceptual decisions[11,19,20], the actual systems-level dynamics underlying decision-making has been elusive. Knowledge of these dynamics would directly test the current prevailing attractor hypotheses, provide fundamental constraints on neural circuit models, and account for the often complex temporal profiles of neural activities.

A separate line of work involves deep learning tools for capturing the low-dimensional component of neural activity[21,22]. In this approach, the spike trains of simultaneously recorded neurons are modeled to encode a “latent” variable that is shared across neurons. The temporal evolution of this variable is inferred using artificial neural networks.

To combine both lines of work, we used a novel method (Kim et al., 2023) that estimates, from the spike trains of simultaneously recorded neurons, the dynamics of a low-dimensional variable z. This is given by:

$$\dot{z} = F(z, \boldsymbol{u}) + \boldsymbol{\eta} \qquad (1)$$

Applied to perceptual decisions, z represents the brain's decision state at a given time (**Fig**. 1a-c). The instantaneous change of the decision variable, or its dynamics, is given by $\dot{z}$, which depends on z itself, external inputs $\boldsymbol{u}$, and noise $\boldsymbol{\eta}$.

Intrinsic and input dynamics differentiate hypotheses

The deterministic dynamics F (that is, in the absence of $\boldsymbol{\eta}$) is useful for distinguishing among attractor hypotheses of decision-making. F can be dissected into two components: the intrinsic (i.e., input-independent) dynamics and the input-driven dynamics. Intrinsic dynamics occur even in the absence of sensory inputs $\boldsymbol{u}$ (i.e., $F(z,\mathbf{0})$; **Fig**. 1d). Changes in z driven solely by sensory inputs $\boldsymbol{u}$ can be isolated from the intrinsic dynamics as $F(z, \boldsymbol{u}) - F(z, \boldsymbol{0})$ (**Fig**. 1e).

Different intrinsic dynamics are postulated by the currently prevailing attractor hypotheses of decision-making (**Fig**. 1f-h; Extended Data Fig. 1). Whereas bistable attractors are proposed in one hypothesis (**Fig**. 1f), a line attractor is postulated by the classic drift-diffusion model (DDM)[23,24,25] hypothesis (**Fig**. 1g) and also in a hypothesis inspired by recurrent neural networks (RNN) trained to make perceptual choices (**Fig**. 1h). In both the bistable distractor and DDM hypotheses, but not in the RNN line attractor hypothesis, the input dynamics are aligned to the attractors in the intrinsic dynamics. Because these canonical hypotheses were designed to explain only a subset of the phenomena observed in decision-making experiments, it may well be that a broader range of experimental observations is captured by non-canonical dynamics. As an example, the intrinsic dynamics may contain discrete attractors that are not aligned to the input dynamics (**Fig**. 1i).

To dissociate between intrinsic and input dynamics, we trained rats to perform a task in which it listened to randomly timed auditory pulses played from its left and right and reported the side where more pulses were played (**Fig**. 1a). The stochastic pulse trains allow us to sample neural responses time-locked to pulses, which are useful for inferring the input dynamics, and also the neural activity in the intervals between pulses, which is useful for inferring the intrinsic dynamics. Expert rats are highly sensitive to minute differences in the auditory pulse number (**Fig**. 1b; Extended Data Fig. 2a), and the behavioral strategy of rats in this task is typically well captured by gradual integration[26,27].

While the rats performed this task, we recorded six frontal cortical and striatal regions with chronically implanted Neuropixels probes (**Fig**. 1j-k; Extended Data Fig. 2b). The frontal orienting fields (FOF) and anterior dorsal striatum (dStr) are known to be causally necessary for this task and are interconnected[28–30]. The dorsomedial frontal cortex (dmFC) is a major anatomical input to dStr[31], as confirmed by our retrograde tracing (Extended Data Fig. 2c) and is also causally necessary for the task (Extended data Fig. 1d). The dmFC is interconnected with the medial prefrontal cortex (mPFC), and less densely, the FOF and the primary motor cortex (M1)[32], and provides inputs to the anterior ventral striatum (vStr)[31].

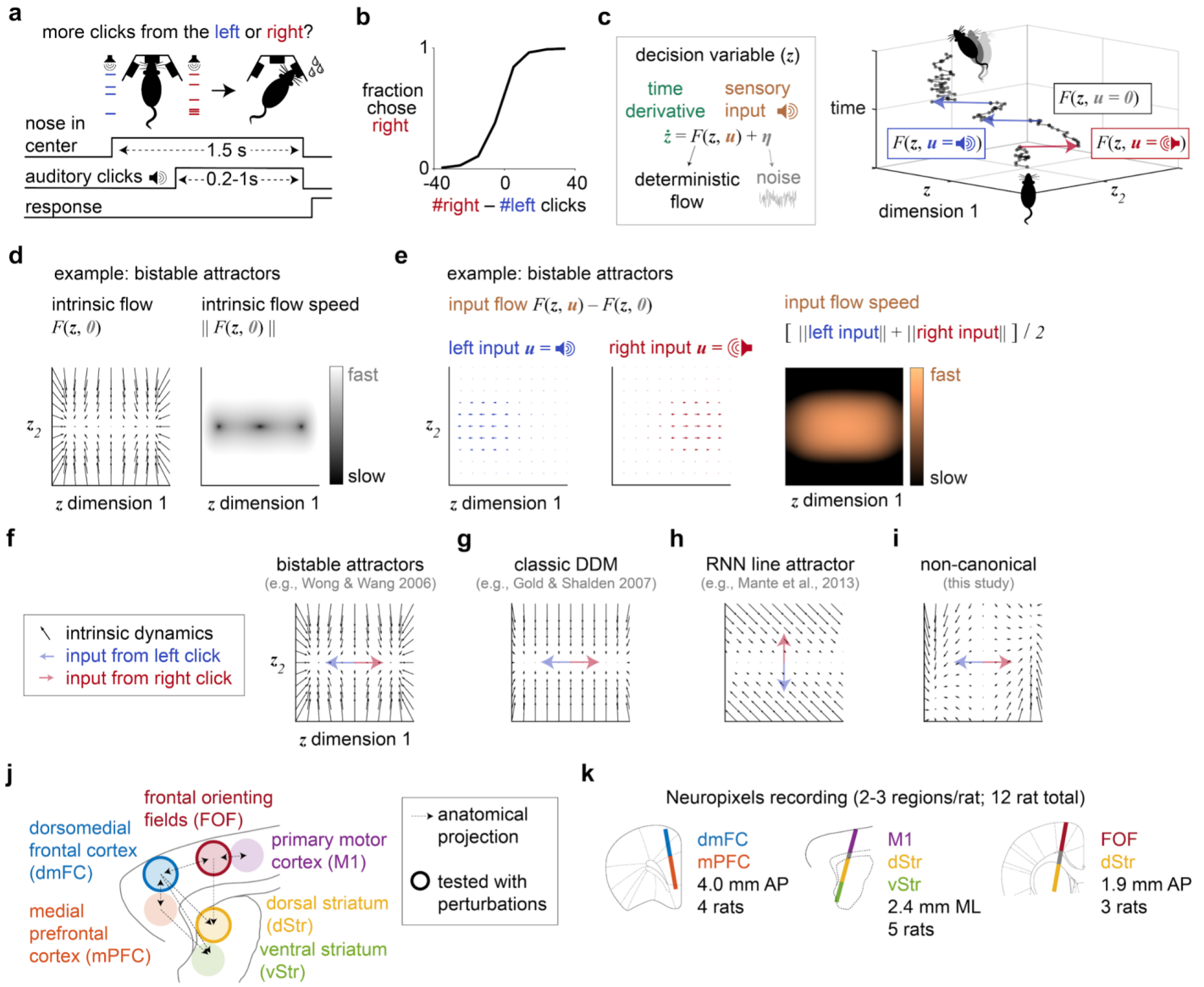


Figure 1. Attractor models of decision-making were tested by recording from rat frontal cortex and striatum. **a,** Rats were trained to integrate auditory pulsatile evidence. While keeping its head stationary, the rat listened to randomly timed clicks played from loudspeakers on its left and right. At the end of the stimulus, the rat turns to the side with more clicks for a water reward. **b**, Behavioral performance in an example recording session. **c**, The decision process is modeled as a dynamical system. Right: the blue, red, and black lines represent the change in the decision variable in the presence of left, right, or no click, respectively. **d**, Intrinsic dynamics illustrated using the bistable attractors hypothesis. In the flow field (top right), the arrow at each value of decision variable z indicates how the instantaneous change depends on z itself. The arrow's orientation represents the direction of the change, and its size the speed, which is also quantified using a heat map. **e**, Changes in z driven solely by the external sensory inputs. **f,** Bistable attractors hypothesis of decision-making, with the direction of the input flows. **g**, The drift-diffusion model (DDM) hypothesis, implemented with a line attractor. **h**, Recurrent neural networks (RNN) can be trained to make perceptual decisions using a line attractor that is not aligned to the input dynamics. **i**, An example of non-canonical attractor dynamics. **j**, Six interconnected frontal cortical and striatal regions are examined here. **k**, Neuropixels recordings (318±147 neurons/session/probe, mean±STD) from twelve rats.

Unsupervised discovery reveals non-canonical attractor dynamics

To test the canonical attractor hypotheses and to discover potentially non-canonical dynamics, a flexible and interpretable method is needed. Therefore, we used a novel deep learning method that infers low-dimensional stochastic dynamics from population spiking data: Flow field Inference from Neural Data using deep Recurrent networks (FINDR; Kim et al., 2023). FINDR uses artificial neural networks as function approximators to infer the decision-related neural dynamics (Eq. 1; **Fig**. 1c). Specifically, the decision-related dynamics are approximated with gated neural stochastic differential equations (**Fig**. 2a). Our approach separately accounts for the decision-independent, deterministic but time-varying baseline firing rate for each neuron so that FINDR infers decision-related latent variables that are directly focused on the choice formation process (**Fig**. 2b). Current deep learning-based tools for analyzing neural activity assume that the dynamics shared across neurons (i.e., the latent dynamics) are high-dimensional and deterministic[21,33]. In contrast, FINDR infers low-dimensional and stochastic latent dynamics, allowing the discovery of flow-fields that are interpretable. These flow-fields of FINDR can be used to distinguish between the competing dynamical systems hypotheses of perceptual decision-making (Extended Data Fig. 3).

To illustrate the results from FINDR, we focus on a representative recording session (**Fig**. 2c-h) and show results for other sessions in Extended Data Fig. 6. As validation, we confirmed that FINDR can well predict the heterogeneous single-trial firing rates of individual neurons, as well as the complex dynamics in their peristimulus time histograms (PSTH) given an evidence strength (Extended Data Fig. 4). We found the number of dimensions in the latent dynamics that best captures our data to be two (Extended Data Fig. 5).

Surprisingly, the configuration of intrinsic and input flows inferred by FINDR is not consistent with any of the canonical attractor hypotheses. The intrinsic flow consists of bistable attractors and an unstable steady point in between the two attractors (**Fig**. 2c-d). The axis formed by the bistable attactors in the observed dynamics is not aligned to the input flow (**Fig**. 2e-f). The absence of a line attractor contrasts with the classic DDM and RNN hypotheses, and the non-alignment in the direction of the intrinsic and input flows differs from the bistable attractor hypothesis.

Moreover, in the observed dynamics, the speed of the input flow substantially exceeds the speed of the intrinsic flow for much of the state space, except close to the periphery (**Fig**. 2g). Because of the larger speed of the input flow in much of the state space around the origin, evidence integration depends little on the intrinsic flow and is instead dominated by the input flow. In other words, although there is a non-zero intrinsic flow, on the scale of the inputs the intrinsic flow can be thought of as approximated by a plane attractor, upon which direction of movement is determined mainly by the inputs. The decision variable thus initially evolves along the direction of the input. When it reaches the periphery, where the intrinsic flow increases in speed, the input flow is no longer dominant, and the decision variable veers toward one of the bistable attractors (**Fig**. 2h). The presence of the bistable attractors obviate invoking a separate mechanism for decision commitment and maintenance.

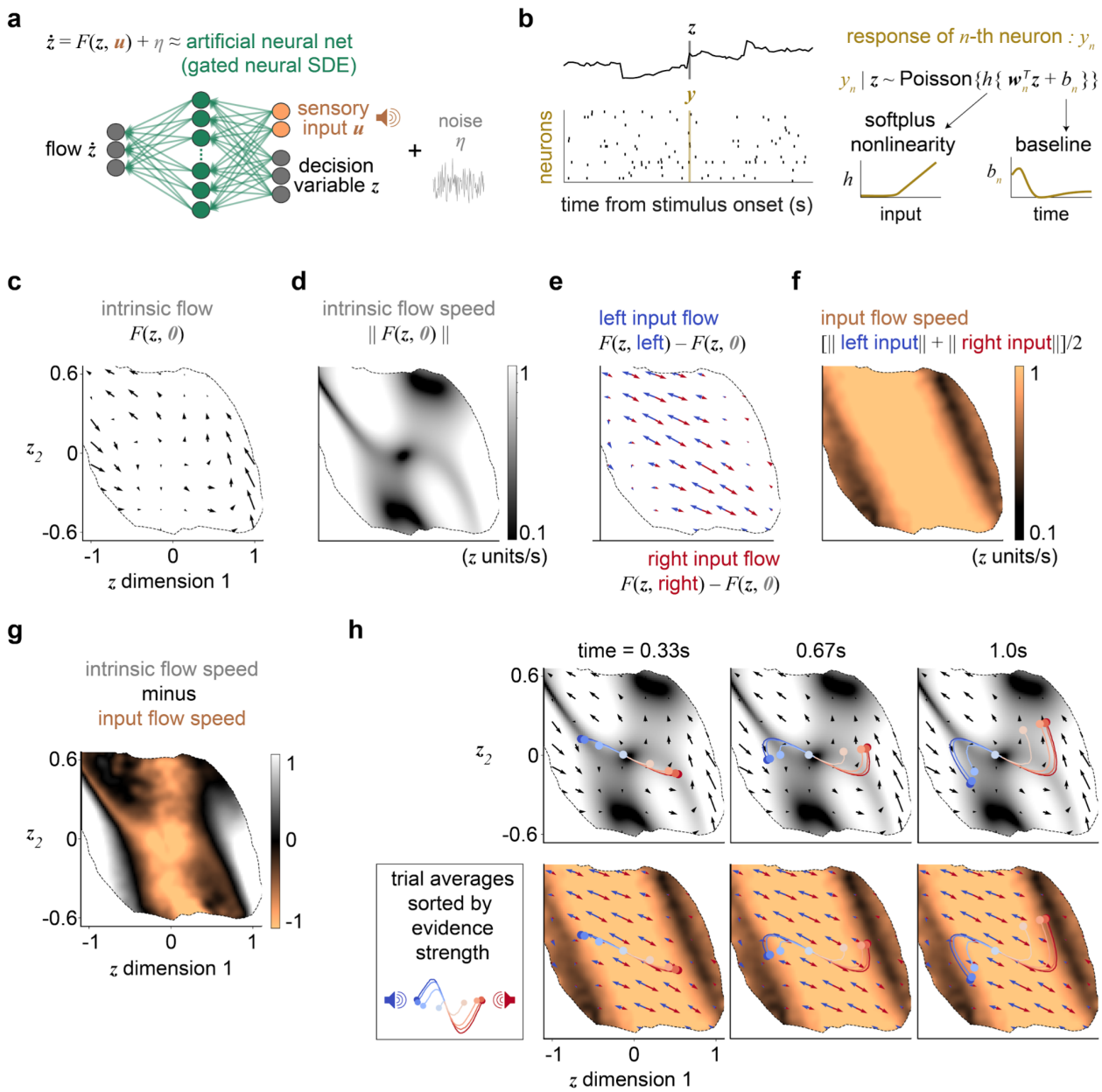


Figure 2. Unsupervised discovery reveals non-canonical attractor dynamics. **a**, The decision-related dynamics are approximated with an artificial neural network (gated neural stochastic differential equations; gnSDE) and inferred using the method "Flow-field Inference from Neural Data using deep Recurrent networks" (FINDR). The deterministic component *F* is approximated using a gated feedforward network, and stochasticity ***η*** is modeled as a Gaussian with diagonal covariance. **b**, The parameters of the gnSDE are fit so that the value of the decision variable z can best capture the spike trains of simultaneously recorded neurons. Conditioned on the decision variable at each time step, the spiking response of each neuron at that time is modeled as a Poisson random variable. A softplus nonlinearity is used to approximate the threshold-linear frequency-current curve observed in cortical neurons during awake behavior. A baseline temporal function is learned for each neuron to account for the decision-irrelevant component of the neuron's response. **c-h**, Flow field inferred from a representative recording session from 67 choice-selective neurons in dmFC and mPFC. At each time point, only the well-sampled subregion of the state space (the portion occupied by at least 100 of 5000 simulated trial trajectories) is shown. **c**, Intrinsic flow. **d**, Speed of the intrinsic flow. **e**, Input flow of either a left or right click. **f**, Speed of the input flow. **g**, Difference in the speed between intrinsic and input flow. **h**, Trial averages overlaid with either the intrinsic flows or the input flows.

Simplified model of non-canonical dynamics

We reasoned that the key features of the non-canonical dynamics that we observed through FINDR should be capturable in a simplified model with fewer parameters, and therefore potentially greater statistical power when estimating those parameters. Such a simplified model, if successful, should be able to account for a broad range of empirical observations. For simplicity, we interpret the non-canonical dynamics observed through FINDR to consist of two distinct regimes, one for evidence accumulation, and the other for choice maintenance. The evidence accumulation regime consists of the subregion of the state space close to the origin, where the inputs play a dominant role and drive the neural trajectories along a line that is monotonically related to accumulated evidence (**Fig**. 3a). In contrast, the choice maintenance regime consists of the subregions near the discrete attractors, where contributions from inputs diminish, and intrinsic dynamics turn the neural trajectories toward a discrete attractor to commit to a categorical choice and to maintain the memory of the choice.

Separate regimes for evidence accumulation and choice maintenance are explicitly implemented in the classic DDM (**Fig**. 3a). In the classic DDM, inputs and noise are accumulated through a one-dimensional (1-D) decision variable until it reaches an absorbing bound, after which the variable becomes fixed and insensitive to noise or input (**Fig**. 3b). The bounds correspond to the regime of choice maintenance, and the state space between the bounds corresponds to the regime of evidence accumulation. The classic DDM has simple 1-D dynamics, has normative mathematical properties[34,35], and is widely used to interpret behavior[36] and neural responses[3] during perceptual decisions (even for tasks for which the stimulus duration is determined by the environment[26,29,37] as used here). The classic DDM's conceptual similarity to the non-canonical dynamics, simplicity, normative properties, and prevalence motivate us to use it as the basis of a simplified model.

The classic DDM posits that the same dimension, i.e., the same mode of neural activity, underlies evidence accumulation and choice maintenance. However, in the non-canonical dynamics, accumulation and maintenance are implemented along separate neural modes. To approximate the multi-modality, we extend the classic DDM to allow neurons to encode the decision variable in a different way before and after the bound is reached. Whereas the classic DDM embeds evidence accumulation and choice maintenance in the same neural mode, we approximate the non-canonical dynamics by embedding the two states into separate neural modes.

In what we will call the multi-mode drift-diffusion model (MMDDM), the decision variable behaves the same as in the classic DDM (**Fig**. 3c). For simplicity, the decision variable has only three parameters (Extended Data Fig. 7a-b) and incorporates hyperparameters to address sensory adaptation (Extended Data Fig. 7c)[38,39]. The only difference between the MMDDM and our implementation of the classic DDM is in the neurons' encoding in the decision-variable. In the MMDDM, for each neuron, two scalar weights, w_{EA} and w_{CM}, specify its encoding of the accumulator during the evidence accumulation state and the choice maintenance state, respectively. When the accumulator has not yet reached the bound, all simultaneously recorded neurons encode the decision variable through their own w_{EA}, and when the bound is reached, their own w_{CM}. The classic DDM is constrained for w_{EA} and w_{CM} to be the same. Because neurons multiplex both decision-related and unrelated signals[40,41], for both the MMDDM and classic DDM, we incorporate spike history and various baseline changes (Extended Data Fig. 7d-g). All parameters are learned simultaneously by jointly fitting to all spike trains and choices.

While the classic DDM captures ramp-like neuronal temporal profile (Extended Data Fig. 8), the MMDDM accounts for a broader range of neuronal profiles (**Fig**. 3d-h). For the vast majority of recording sessions, the data are better fit by the MMDDM than the classic DDM (**Fig**. 3i). The goodness-of-fit of the neuronal temporal profiles is improved by the MMDDM (**Fig**. 3j). We confirm that the choices can be well predicted (**Fig**. 3k; Extended Data Fig. 7i), that the flow-fields inferred from real spike trains match the

flow-fields inferred from the spike trains simulated by the MMDDM (**Fig**. 3l), and other validation analyses (Extended Data Fig. 7j-n).

The state transition from evidence accumulation to choice maintenance in the MMDDM, and a consequent switch from w_{EA} to w_{CM} is reminiscent of the switch in neural modes encoding a decision variable to those encoding an action plan, observed in[42]. The MMDDM, as a formal model, provides the further advantage of allowing precise inference of each trial's moment of the switch (the time of internal commitment to a decision), based on the recorded neural activity. If the MMDDM indeed correctly infers the time of internal decision commitment, then it should follow that the contribution of the auditory click inputs to the behavioral choice should cease abruptly at this estimated decision commitment time. We tested this prediction by measuring the time-varying weight of the stimulus fluctuations on the behavioral choice, referred to as the psychophysical kernel[43–45]. We developed a psychophysical kernel model of behavioral choice and fit it using the input fluctuations aligned to each trial's time of internal decision commitment, as inferred by the MMDDM (Methods). Consistent with a choice maintenance state, the psychophysical weight of stimulus fluctuations abruptly diminishes to zero after the MMDDM-inferred time of commitment (**Fig**. 3m; Extended Data Fig. 9).

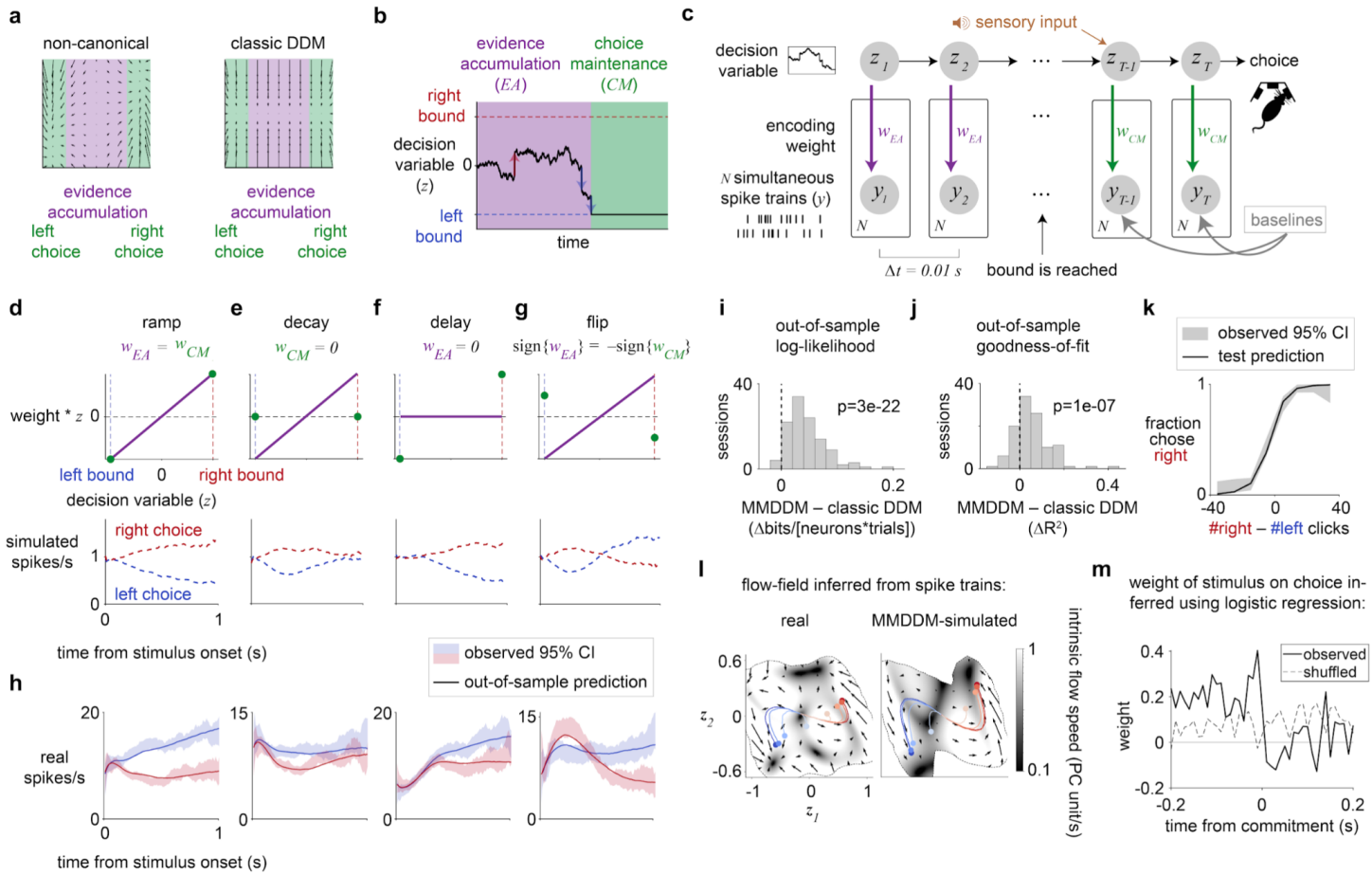


Figure 3. Multi-mode drift-diffusion model (MMDDM): a simplified model of non-canonical dynamics. **a**, The state space of both non-canonical dynamics and the classic DDM can be partitioned into an evidence accumulation regime and a choice maintenance regime. **b**, The choice maintenance regime of the classic DDM corresponds to the decision variable being at an absorbing bound. **c**, Directed graph of the MMDDM for a trial with T time steps and N simultaneously recorded neurons. **d**, The ramp temporal profile in a neuron's PSTH can be generated by setting w_{EA} and w_{CM} to be the same. **e**, The decay profile is simulated by setting w_{CM} to zero because as time passes, it is more and more likely to reach the bound and for the encoding to be mediated by w_{CM} rather than by w_{EA}. When w_{CM} is 0, there can be no selectivity. **f**, A delay profile is generated

by setting w_{EA} to zero because the probability of reaching the bound is zero for some time after stimulus onset. **g**, The flip profile is produced by setting w_{EA} and w_{CM} to have opposite signs. **h**, For actual neurons, the MMDDM captures the diversity in choice-related temporal profiles. **i**, Behavioral choices are well predicted. **j**, The MMDDM has a higher out-of-sample likelihood than the classic DDM. **k**, The MMDDM achieves a higher goodness-of-fit (coefficient-of-determination, R^2) of the choice-conditioned PSTH's than the classic DDM. **l**, The flow field inferred from real spike trains is confirmed to be similar to that inferred from MMDDM-simulated spike trains. **m**, Consistent with a choice maintenance state, the psychophysical weight of the stimulus on the behavioral choice decreases to near zero after time of decision commitment.

Abrupt and gradual changes at decision commitment

A puzzling phenomenon observed during perceptual decision-making is the diversity of ramp-to-bound and discrete step-like dynamics across choice-selective neurons[4–6]. While some neurons show a ramp-to-bound temporal profile, some others exhibit approximately discrete step-like dynamics, and still others are between a ramp and a step. An explanation for this diversity has been elusive. Whether a neuron shows signature of ramp-to-bound, step, or something in between might be explained by a rapid reorganization in population dynamics at the time of decision commitment–a proposal conceptually similar to the change in neural activity modes from motor preparation to motor execution[46]. Not all neurons would be equally coupled to this change: a neuron similarly engaged in evidence accumulation and choice maintenance would exhibit a ramp-to-bound profile, whereas a neuron more strongly engaged in maintenance would show a steep–almost discontinuous–step. Moreover, neurons that are more strongly engaged in accumulation would reflect a ramp-and-decline profile. Can the changes in neuronal responses around the time of decision commitment explain the continuum of ramping and stepping profiles?

We find evidence for these predictions when we grouped neurons by whether they are more, less, or similarly engaged in evidence accumulation relative to choice maintenance (Methods; **Fig**. 4a-d; Extended Data Fig. 10a). The peri-commitment neural response time histogram (PCTH) averaged across neurons similarly engaged in accumulation and maintenance shows a ramp-to-bound profile. The PCTH resembles a step for the neurons more engaged in maintenance, and shows a ramp-and-decline profile for neurons more engaged in accumulation. Even without grouping neurons, we find support for these predictions in the principal component analysis (PCA) on the PCTH's (Methods; **Fig**. 4e). The first three principal components correspond to the ramp-to-bound, step, and ramp-and-decline profiles. These results show that concurrent gradual and abrupt changes in neural responses around the time of decision commitment explain the continuum of ramp-to-bound and discrete step-like profiles.

Abrupt changes at decision commitment appear inconsistent with a phenomenon that is observed in many studies of decision-making: smoothly curved trial-averaged trajectories in low-dimensional neural state space[7–9]. Similar phenomenon in our data: the trial-averaged trajectories for left and right choices do not bifurcate along a straight line, but rather along curved arcs (**Fig**. 4f). This phenomenon can be explained if the neural responses are not aligned to the time of decision commitment, but instead to the stimulus onset, then the trial-averaged trajectories are expected to be smooth rather than sharply angled. The smoothness is similar to the shuffling of the commitment times when computing the PCTH. Importantly, this phenomenon can be well captured by the out-of-sample predictions of MMDDM (**Fig**. 4g; Extended Data Fig. 11), but not the single-mode DDM (**Fig**. 4h). These results indicate that the MMDDM, a simplified model of the non-canonical dynamics, can well capture the widespread observation of smoothly curved trial-averaged trajectories.

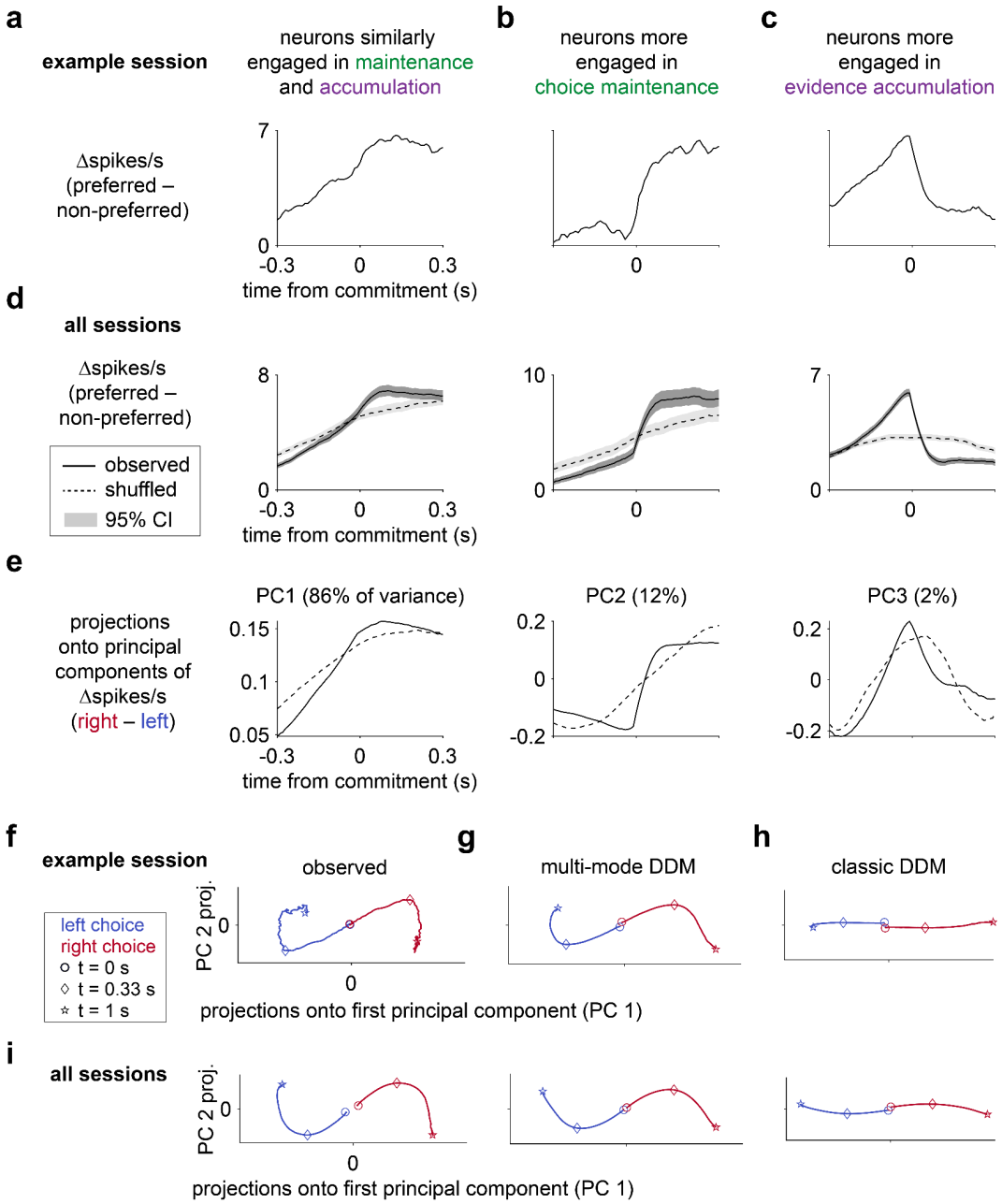


Figure 4. Model explains the diversity of ramping and stepping dynamics and trial-averaged curved trajectories. **a**, The peri-commitment time histogram (PCTH) averaged across neurons that are similarly engaged in choice maintenance and evidence accumulation show a ramp-to-bound profile. “Preferred” refers to the behavioral choice for which a neuron emitted more spikes. Example session. N=23 neurons. **b**, In the same session, the PCTH averaged across 10 neurons that are more strongly engaged in maintenance has an abrupt, step-like profile. **c**, The ramp-and-decline profile characterizes the PCTH averaged across the 43 neurons from the same session that are more strongly engaged in evidence accumulation. **d**, Across sessions, the ramp-to-bound, step-like, and ramp-and-decline profile characterize the PCTH of neurons similarly engaged in maintenance and accumulation (1,116), more engaged in maintenance (414), and more engaged in accumulation (1,529), respectively. **e**, The ramp-to-bound, step-like, and ramp-and-decline profile are observed in the first three principal components of the PCTH’s. **f**, Curved trial-averaged trajectories. **g**, Out-of-sample predictions of the MMDDM. **h**, Out-of-sample predictions of the classic DDM fail to account the trial-averaged trajectories. **i**, Results are similar when neurons are pooled across sessions.

Dynamics across brain regions

We observed non-canonical dynamics across the multiple frontal cortical and striatal areas: The MMDDM better captures neural responses in each brain region than the classic DDM, than a bistable attractor model, and also a leaky integrator model (**Fig**. 5a; Extended Data Fig. 12a-c). Other signatures of non-canonical dynamics, including peri-commitment neural changes (Extended Data Fig. 10a) and curved trial-averaged responses (Extended Data Fig. 11a) can be observed within each brain region. These results indicate that non-canonical dynamics is general across frontal cortex and striatum.

However, differences might nonetheless exist across brain regions. In particular, when the time-varying choice selectivity (Extended Data Fig. 8i) is averaged across neurons, and the latency to the peak is identified: the latency is the shortest for mPFC and the longest for FOF (**Fig**. 5b). This difference in the latency to peak across brain regions may be explained in the context of separate neural modes for evidence accumulation and choice maintenance: neurons that are more strongly engaged in evidence accumulation have a shorter latency to peak selectivity than neurons that are more strongly engaged in choice maintenance. The latencies to peak suggest these frontal brain regions differ in their relative engagement in evidence accumulation and choice maintenance.

To test this possibility, we examined the encoding weights in MMDDM, and for each neuron, computed a scalar index that compares its relative engagement in the two processes, by taking the difference between the absolute values of the w_{EA} and w_{CM} and normalizing by the sum (**Fig**. 5c). A neuron with an index near 1 is engaged only in evidence accumulation (and shows a decay), whereas a neuron with an index near -1 is engaged only in choice maintenance (and exhibits a delay). A neuron with a classic ramp profile must have an index near zero, which corresponds to equal levels of engagement. The distribution of the indices was unimodal, indicating the engagement in accumulation and maintenance is distributed rather than clustered among neurons (Extended Data Fig. 12d-e). The engagement indices show that brain regions differ in their relative engagement in evidence accumulation: mPFC, dmFC, and dStr are more strongly engaged in evidence accumulation than the regions vStr, M1, and FOF. The rank order of the brain regions by the engagement index is similar to the rank order by time to peak latency.

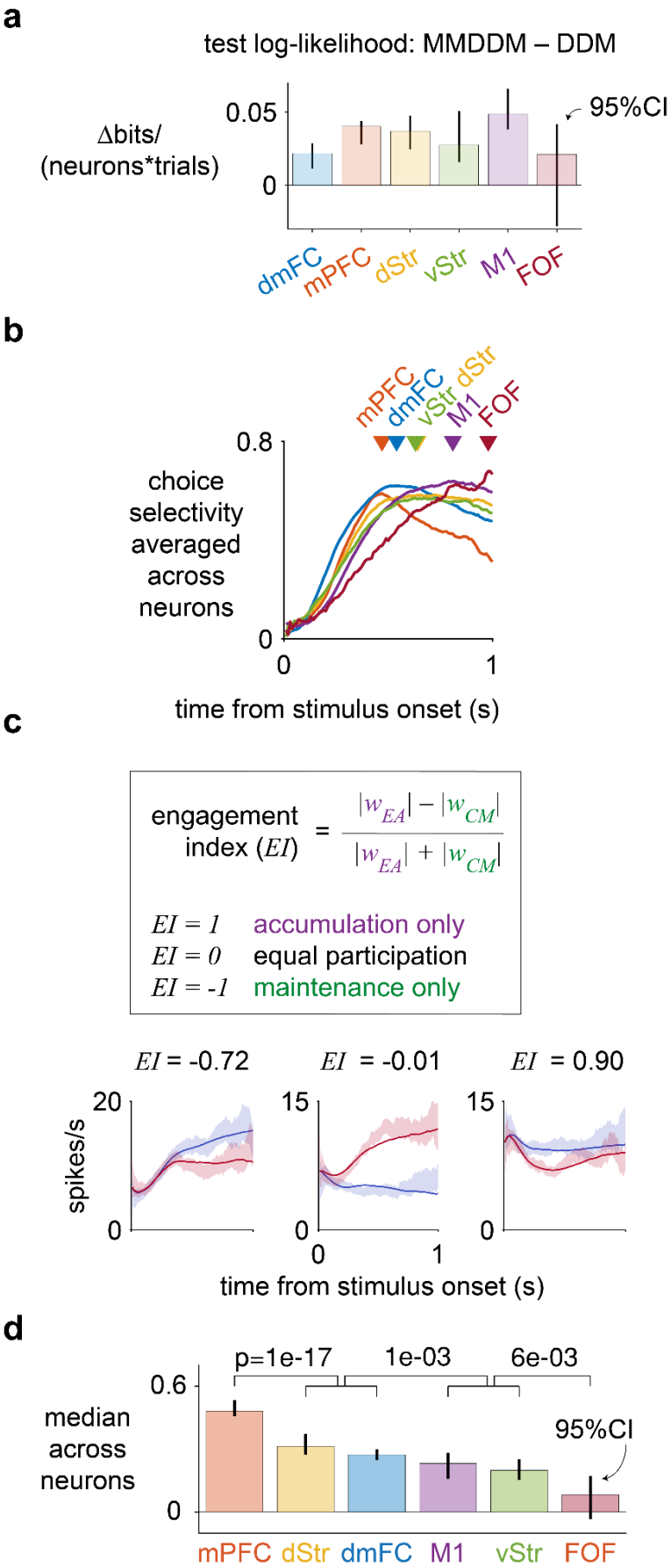


Figure 5. Non-canonical dynamics across frontal areas. **a**, MMDDM better capture the data. Error bar indicates 95% bootstrapped confidence intervals across sessions. N =29 dmFC sessions, 29 mPFC, 86 dStr, 74 vStr, 75 M1, and 7 FOF. **b**, The choice-selective neuronal dynamics averaged across neurons in each brain region show that mPFC neurons are most choice-selective near the beginning, while FOF neurons are most choice-selective toward the end. **c**, Using fits from MMDDM, a scalar index quantifies each neuron's relative level of engagement between evidence accumulation and choice maintenance. **d**, A gradient of relative engagement in evidence accumulation and choice maintenance across frontal brain regions. The rank order of the median engagement index well matches the rank of the latency to peak choice selectivity. Error bar indicates 95% bootstrapped confidence intervals across neurons.

Discussion

Whether, and which, attractor dynamics govern the emergence of a perceptual choice has been long debated[1,7,10]. Here we show that neural activity in multiple brain regions causally involved in perceptual decision-making are governed by attractor dynamics, but the input and intrinsic dynamics of the attractors differ from previous attractor hypotheses. Sensory inputs are integrated in the absence of a line attractor. Discrete attractors implement the commitment to a choice option and maintain the choice in short-term memory. The axis of integration is not aligned with the axis spanned by the discrete attractors. While the attractor dynamics observed here differs from theoretical proposals, it provides a parsimonious explanation of experimental findings from multiple species. Across primates and rodents, sensory inputs and choice are represented in separate neural dimensions[7–9,47–49]. Across time, neither sensory responses nor the neural dimensions for optimal decoding of the choice are fixed[8,47,49]. These phenomena, along

with other observations, including single neuron dynamics, ramping and stepping profiles, and choice behavior, are well captured by the attractor dynamics observed here.

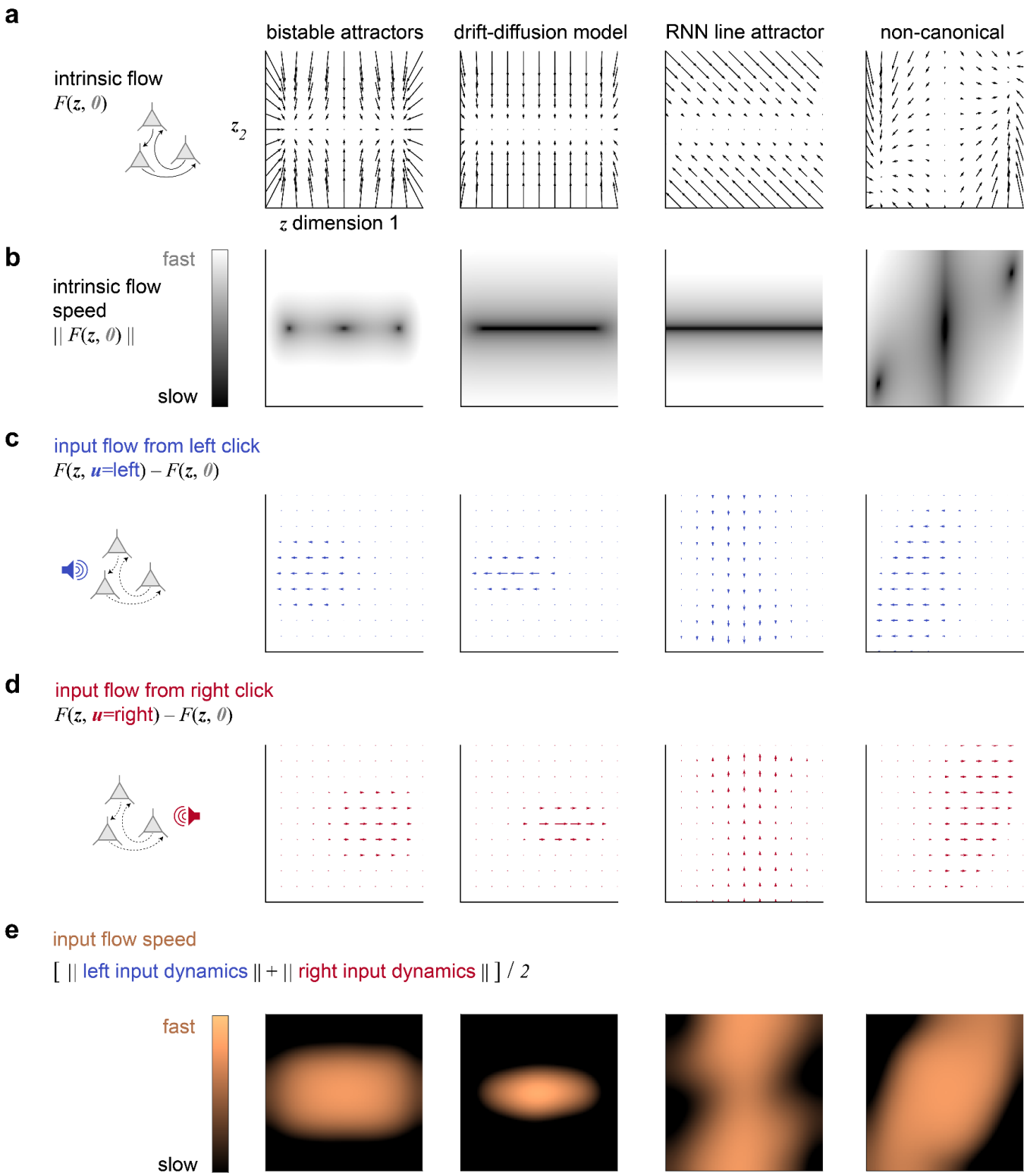
How do decisions end? In many studies, an animal learns over many months to couple the termination of their deliberation with the initiation of a motor response, so that the moment when the animal initiates its response is used to operationally define the time when it commits to a choice[50,51]. Yet, committing to a choice and initiating a response are distinct processes[52], and it has been pointed out that the many-month training conflates neural activity underlying decision commitment and response preparation[53]. Here we show that the precise time of decision commitment on individual trials can be inferred without training an animal to couple their decision commitment and movement initiation. The peri-commitment neural responses observed here contrasts sharply with the ramp-and-burst neural responses observed in animals trained to couple their decision commitment with response initiation[51]. The distinction between perceptual decision and movement preparation, and their interaction[54,55], can be addressed in future studies using decision dynamics inferred from neural activity and the animal's pose estimated from videos.

The attractor dynamics in the neural state space observed here provides a constraint for network models of perceptual decision-making, which consist of recurrently connected nodes that represent different brain regions or classes of neurons. Because the inference of the interaction between neural nodes based on correlational methods alone is susceptible to systematic bias[56], elucidating the attractor model at the level of recurrent connections will require combining transient perturbations and dense electrophysiological recording.

Moreover, the dynamics observed here provide hypotheses for studying decisions driven by evidence of other modalities, such as retrieval from long-term memory[57] or subjective value[58], across multiple contexts, or reported by multiple motor effectors. Decisions involving different modalities of evidence[59], or across different contexts[7], may involve different input dynamics to integrate evidence of different modalities or in different contexts. In contrast, when the modality of the evidence is fixed but the motor effectors for reporting the choice vary[60], different intrinsic dynamics may explain the neural responses for different effectors.

Extended Data Figures

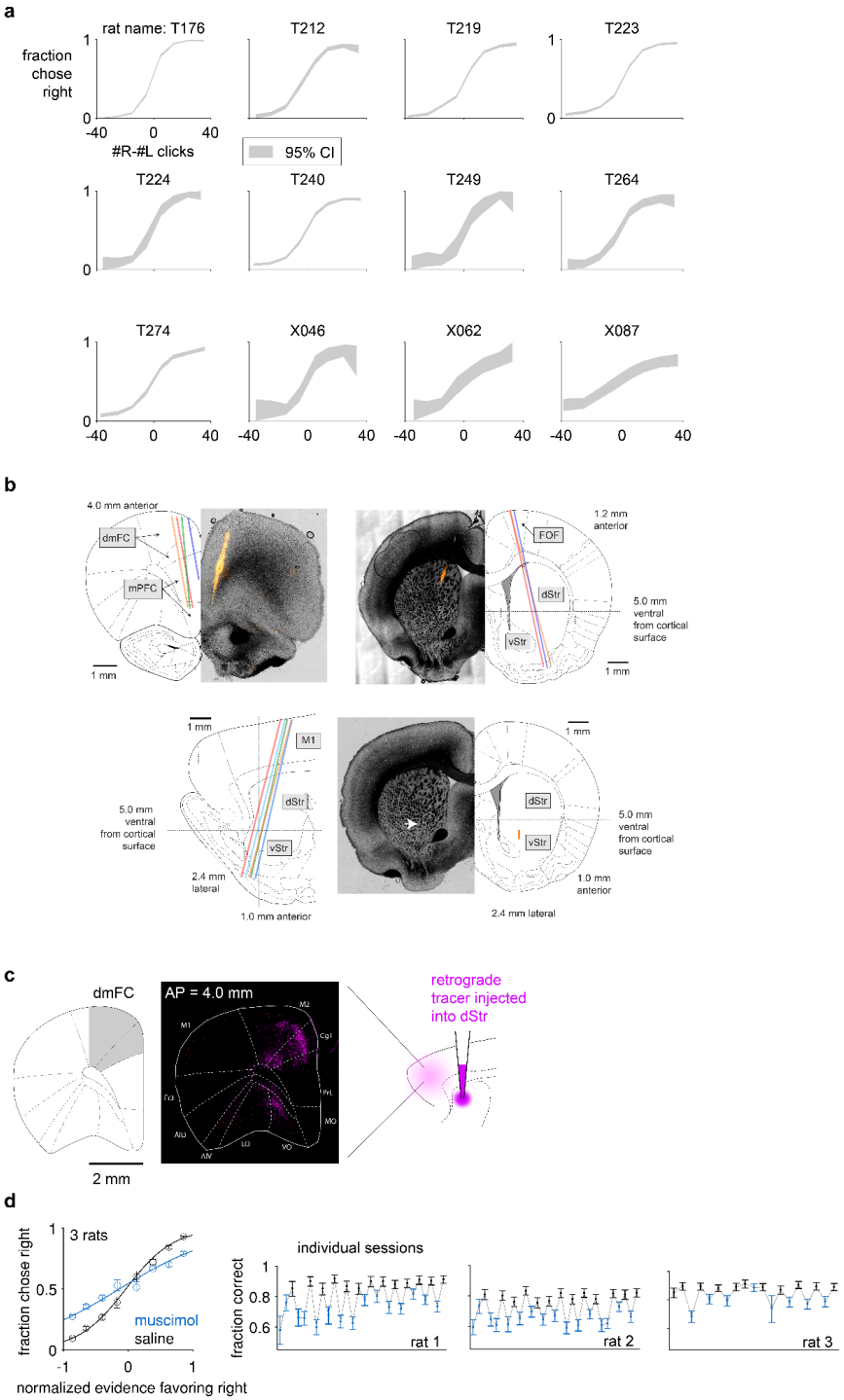
Extended Data Figure 1



Extended Data Figure 1. Attractor hypotheses of perceptual decision-making. In these hypotheses, the decision process is represented by the state of a dynamical system, which we refer to as the "decision variable

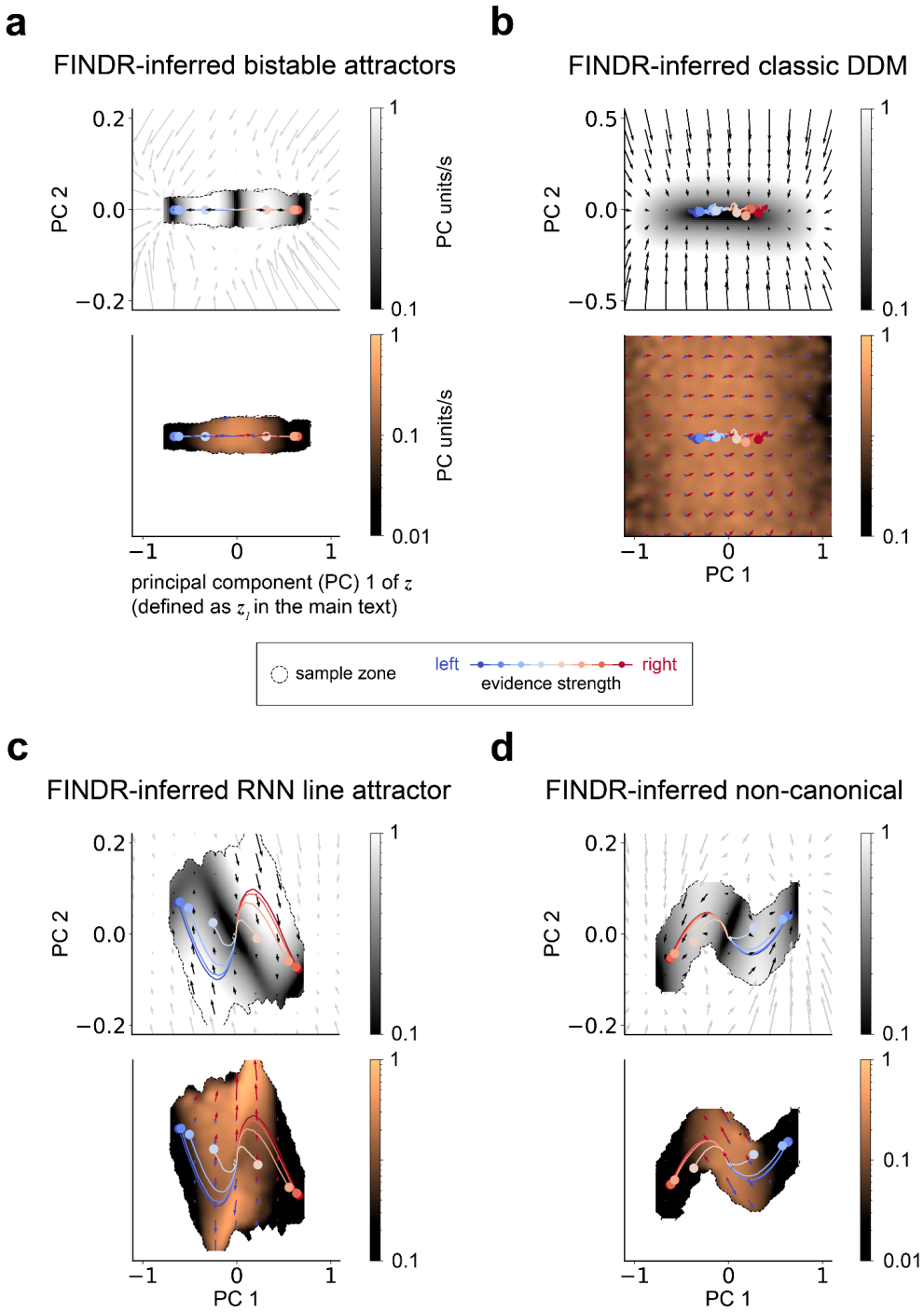
(z)" and is depicted as two-dimensional here but may have fewer or more dimensions. An attractor is a set of states for which the dynamical system tends to move toward, from a variety of starting states. When the system is in an attractor, small perturbations away from the attractor tend to return the system toward the attractor. An attractor can implement the commitment to a choice and the maintenance of the choice in working memory. **a**, In all these hypotheses, the attractors are implemented by the intrinsic dynamics, which corresponds to the deterministic dynamics *F* in the absence of inputs and depends only on z itself. In the bistable attractor hypothesis, there are two discrete attractors, each of which corresponds to a choice alternative. In the classic DDM hypothesis, the intrinsic dynamics form not only two discrete attractors but also a line attractor in between. The intervening line attractor allows an analog memory of the accumulated evidence when noise is relatively small. In the RNN line attractor hypothesis, the intrinsic dynamics form a line attractor, and a separate readout mechanism is necessary for the commitment to a discrete choice. Finally, in the non-canonical hypothesis, the intrinsic dynamics form two discrete attractors. **b**, The intrinsic flow speed is the magnitude of the intrinsic dynamics. A dark region corresponds to a steady state, which can be an attractor, repeller, or saddle point. In the bistable attractor hypothesis, the left and right steady states are each centered on an attractor, and the middle is a saddle point. In both the classic DDM and the RNN line attractor hypothesis, the steady states correspond to attractors. Finally, in the non-canonical hypothesis, the left and right steady states are attractors, and the middle a repeller. **c-d**, Input dynamics corresponding to a left and right auditory pulse, respectively. Here we show the "effective" input flow, which is normalized by $p(\boldsymbol{u}|z)$ to account for the statistics of our stimuli. Whereas in the bistable attractor and classic DDM hypothesis, the inputs are aligned to the attractors, in the RNN line attractor and non-canonical hypotheses, the inputs are not aligned. **e**, The input flow speed is the average of the magnitude of the left input dynamics and the magnitude of the right input dynamics.

Extended Data Figure 2



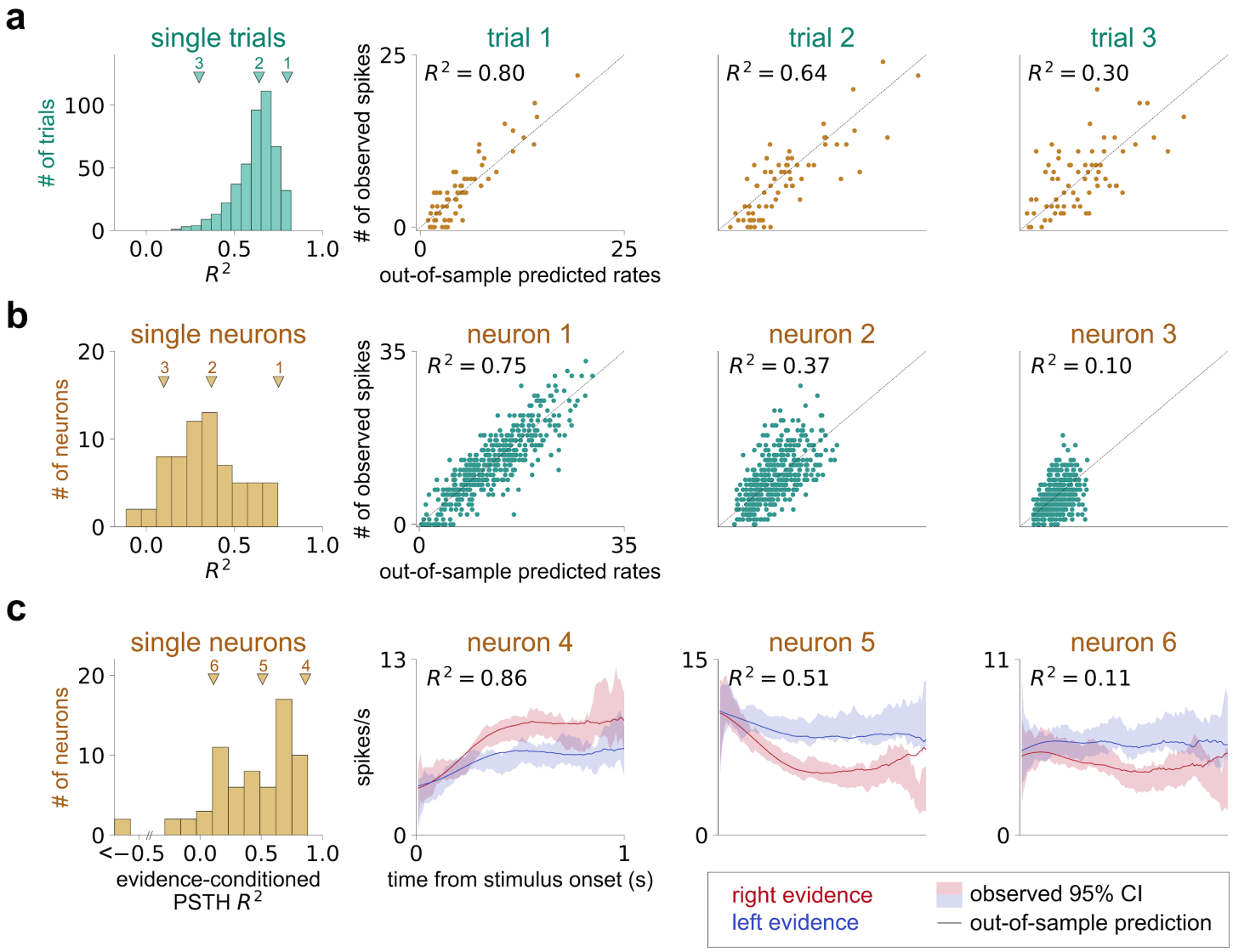
Extended Data Figure 2. Behavioral performance, histological slices, anatomical tracing, and the causal necessity of dmFC. **a**, Psychometric functions of each of the twelve rats recorded aggregated across recording sessions. **b**, Histological images of probe tracks. Each color indicates a probe chronically implanted in a rat. **c**, Dorsomedial frontal cortex (dmFC) provides a major input to the anterior dorsal striatum (dStr). **d**, dmFC is causally necessary for the auditory decision-making task studied here.

Extended Data Figure 3



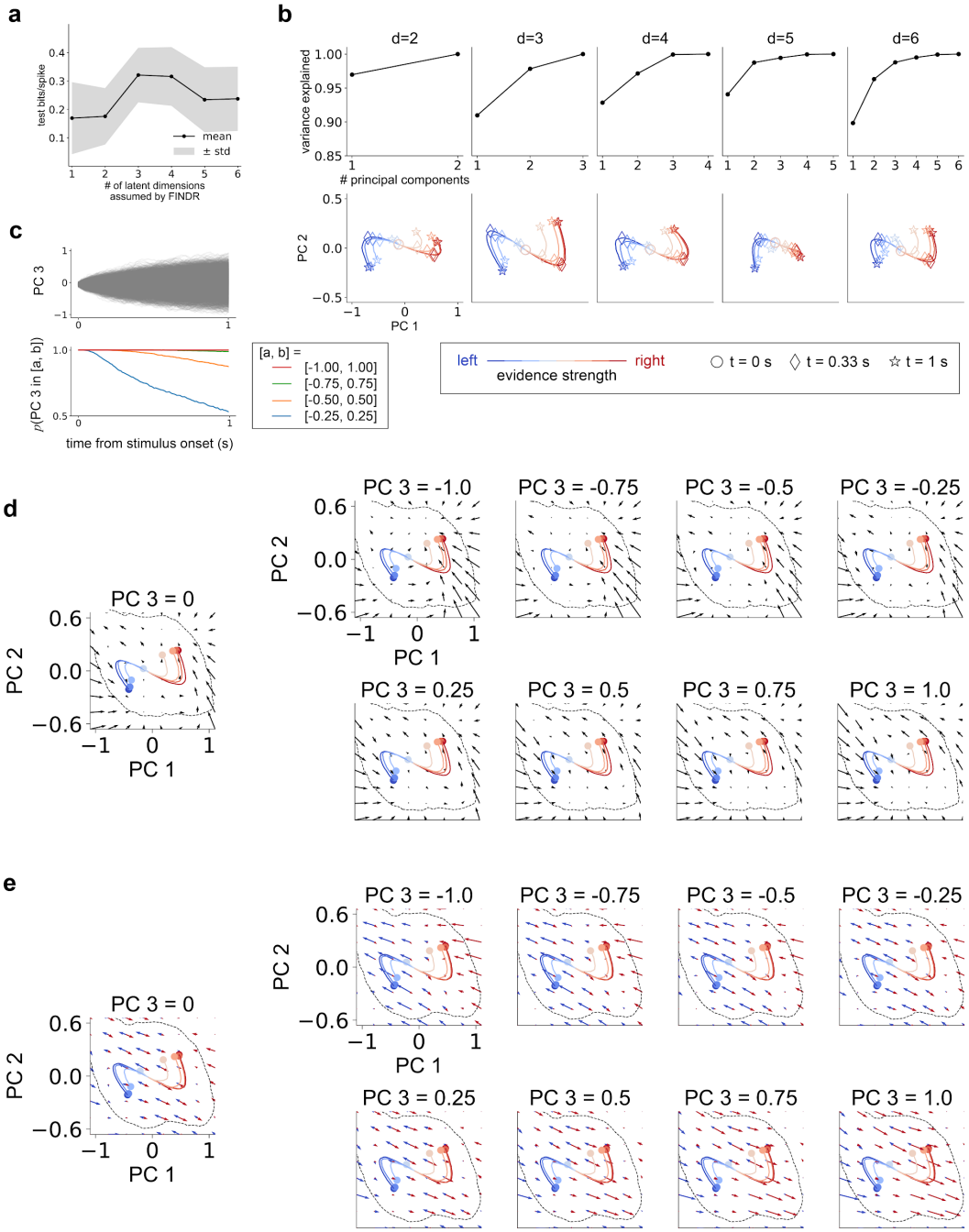
Extended Data Figure 3. FINDR can be used to distinguish between the competing dynamical systems hypotheses of perceptual decision-making. **a**, We simulated spikes that follow the bistable attractor dynamics in Extended Data Fig. 1 to create a synthetic dataset with the number of trials, number of neurons, and firing rates in the range observed in our datasets. Then, we fit FINDR to this synthetic dataset from random initial parameters. The intrinsic and input dynamics inferred by FINDR qualitatively match the bistable attractors hypothesis. **b-d**, FINDR-inferred dynamics qualitatively match the dynamics in **Fig**. 1g-i and Extended Data Fig. 1.

Extended Data Figure 4



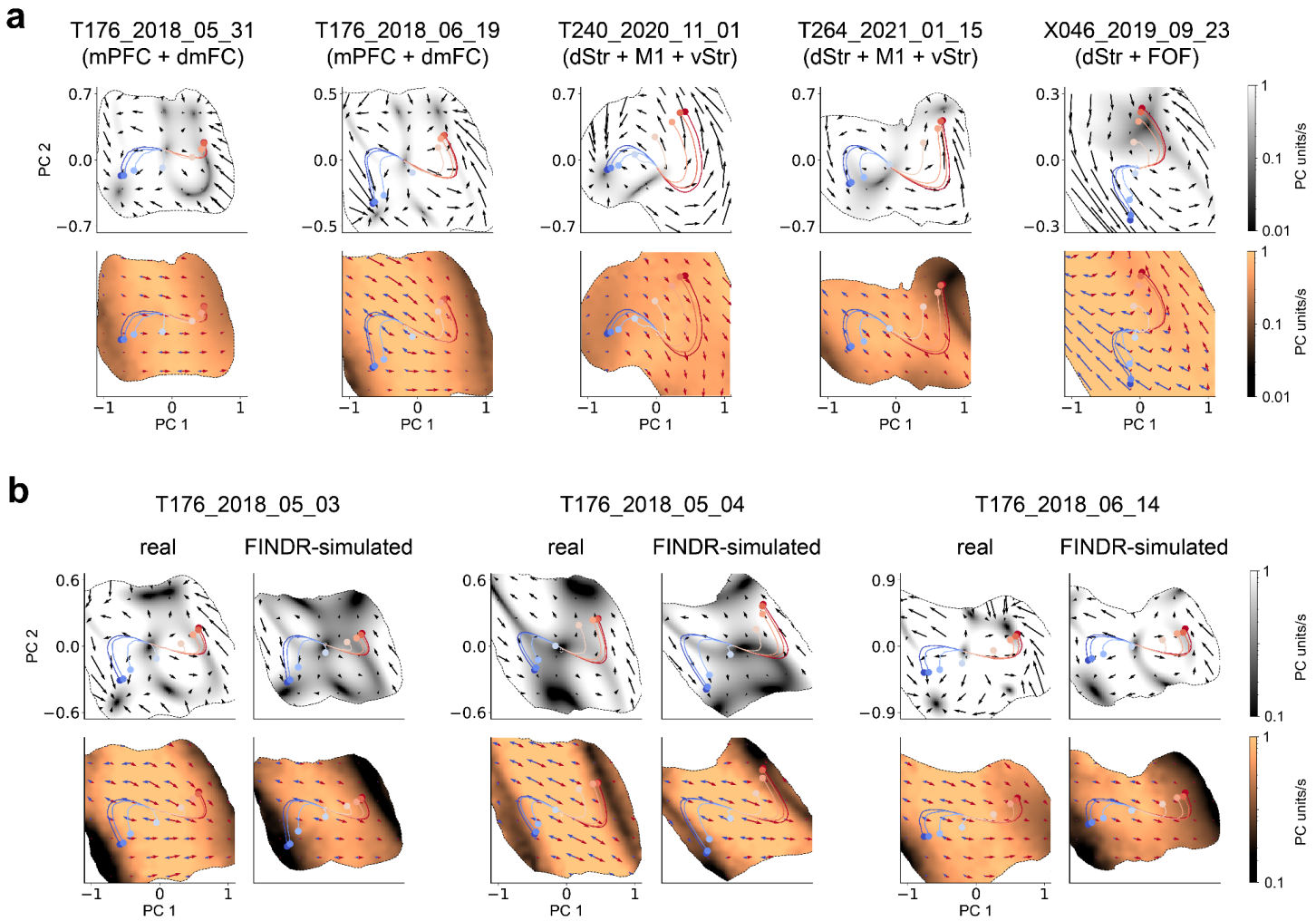
Extended Data Figure 4. FINDR can well capture the neural responses. **a-b**, FINDR captures the underlying firing rates of the single-trial responses of individual neurons from mPFC and dmFC from a representative session. **c**, FINDR captures the complex trial-averaged dynamics of individual neurons in mPFC and dmFC as can be seen in the peristimulus time histograms (PSTH). The goodness-of-fit is measured using the coefficient of determination (R^2).

Extended Data Figure 5



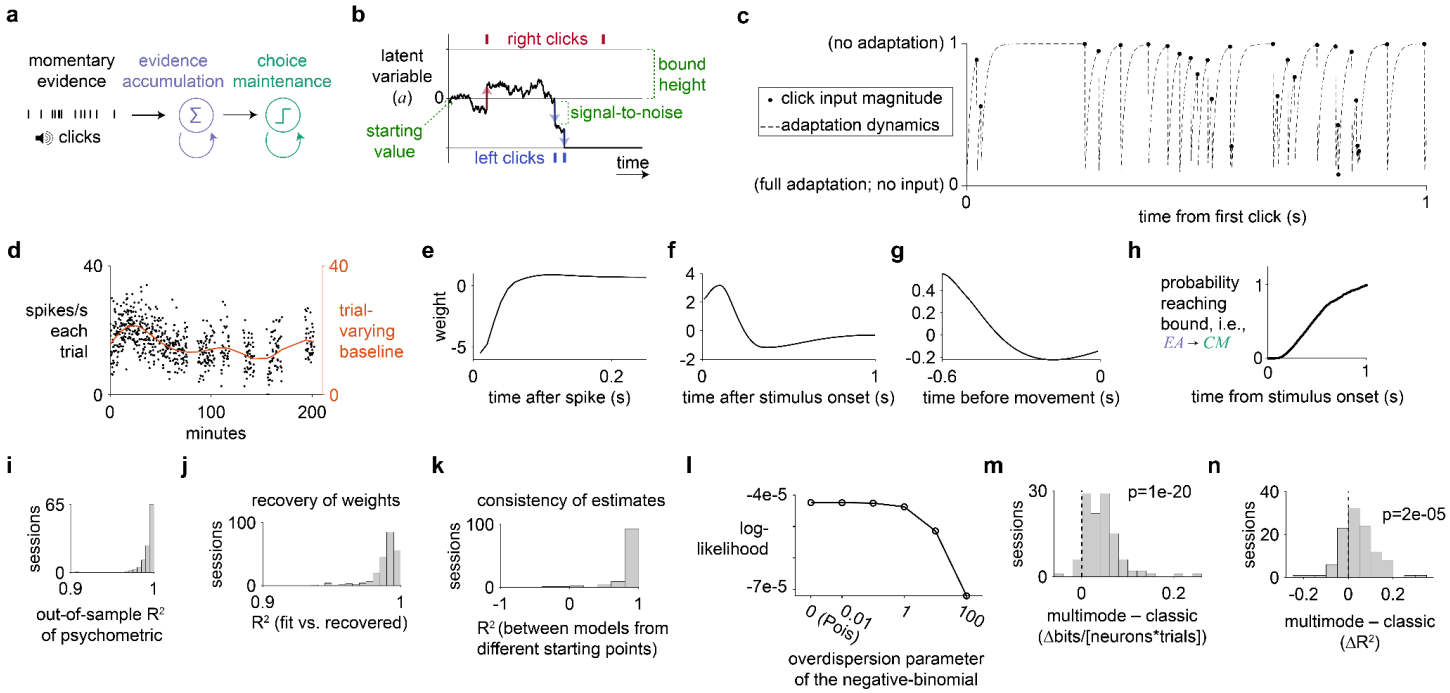
Extended Data Figure 5. FINDR reveals 2-dimensional decision-making dynamics. **a**, Across different FINDR models with dimensions (d) ranging from 1 to 6, the bits/spike of held-out trials maximizes for d=3. Trials were partitioned into 5 folds, and the standard deviation was computed across the 5 partitions. **b**, For FINDR models with different latent dimensions, more than 95% of the variance is captured by the first two principal components (PC's) for d=3-6. For models with two or more dimensions, the trajectories projected onto the first two dimensions were qualitatively similar. **c**, For the FINDR model with d=3, PC 3 takes a value between -0.25 and 0.25 for more than 50% of the trials, and does not go beyond -1 or 1. **d-e**, For the FINDR model with d=3, when we project intrinsic (**d**) and input (**e**) flow fields onto the first two PCs, the projected 2-d flow fields are similar across different PC 3 values. We show the flow fields with PC 3 = 0 in **Fig**. 2. Inside the dashed lines indicate the well-sampled subregion of the state space (sample zone).

Extended Data Figure 6



Extended Data Figure 6. Non-canonical dynamics across sessions, animals, and frontal brain regions. **a**, Non-aligned input and intrinsic flows across different recording sessions, animals, and brain regions. **b**, FINDR can reliably recover the inferred input and intrinsic flows. After fitting FINDR to a dataset, the model parameters are used to simulate a synthetic dataset using the exact same set of sensory stimuli in the real dataset and containing the same number of neurons and trials. From new initial parameter values, FINDR was fit to the simulated data to infer the “FINDR-simulated” flow fields.

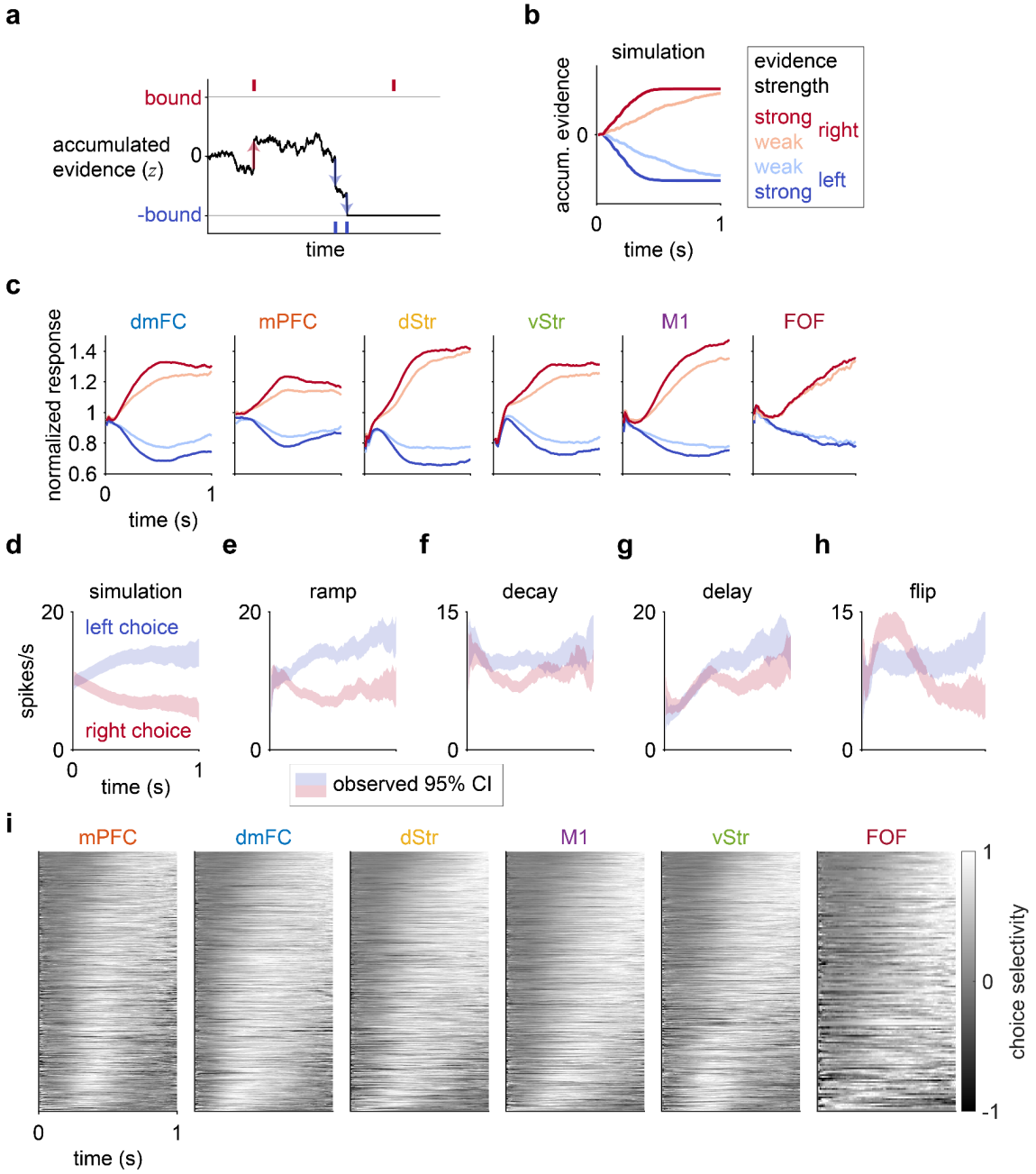
Extended Data Figure 7



Extended Data Figure 7. Multi-mode drift-diffusion model (MMDDM). **a**, The discovered motifs are consistent with two-stage, sequential processing: momentary evidence is first integrated over time until a fixed threshold is reached, at which point decision commitment occurs and the choice is maintained in memory. **b**, The three parameters that are fit in MMDDM consist of the bound height, the starting value, and the signal-to-noise of each momentary input. **c**, The magnitude of the input after sensory adaptation of each click in a simulated Poisson auditory click train. Based on previous findings, the adaptation strength (φ) is fixed to 0.001, and the post-adaptation recovery rate (k) to 100. The generative click rate is 40 hz, as in the behavioral task. **d**, The slowly trial-varying baseline function, parametrized by smooth temporal basis functions, for an example neuron. The baseline is added before the nonlinear rectification. **e**, The spike history filter of the same neuron. **f**, The post-stimulus filter of the neuron. This filter does not depend on the content of the click train and only depends on the timing of the first click. **g**, The kernel of the same neuron to account for movement anticipation. The kernel does not depend on the actual choice of the animal. **h**, The cumulative probability of the latent variable reaching decision commitment over time in a simulated trial. **i**, The psychometric function is well captured across sessions. **j**, After fitting the model to each recording session, the learned parameters are used to simulate a data set, using the same number of trials and the same auditory click trains. The simulations are used to fit a new model, the recovery model, starting from randomized parameter values. The encoding weights of the accumulated evidence of the recovery model are compared against the weights used for the simulation (which were learned by fitting to the data) using the coefficient-of-determination metric. **k**, Consistency in the encoding weights between the training models during five-fold cross-validation. For each session, a coefficient-of-determination was computed for each pair of training models (10 pairs), and the median is included in the histogram. **l**, Whereas the Poisson distribution requires the mean to be the same as the variance, the negative binomial distribution is a count response model that allows the variance to be larger than the mean μ, with an additional parameter α, the overdispersion parameter, that specifies the variance to be equal to $\mu+\alpha\mu^2$. When the overdispersion parameter is zero, the distribution is equivalent to a Poisson. Fitting the data to varying values of the overdispersion parameter shows that log-likelihood is maximized with a Poisson distribution for the conditional spike count response. Similarly when the overdispersion parameter was learned from the data, the best fit values were all close to zero. **m**, Sensory adaptation is not critical to the

improvement in fit by dynamic encoding. Even without modeling sensory adaptation–by setting $\varphi=1$ and $k=0$, such that every click has the same input magnitude–the out-of-sample log-likelihood is reliably improved by dynamic encoding. **n**, The out-of-sample goodness-of-fit of the PSTH's is also reliably improved even in the absence of sensory adaptation.

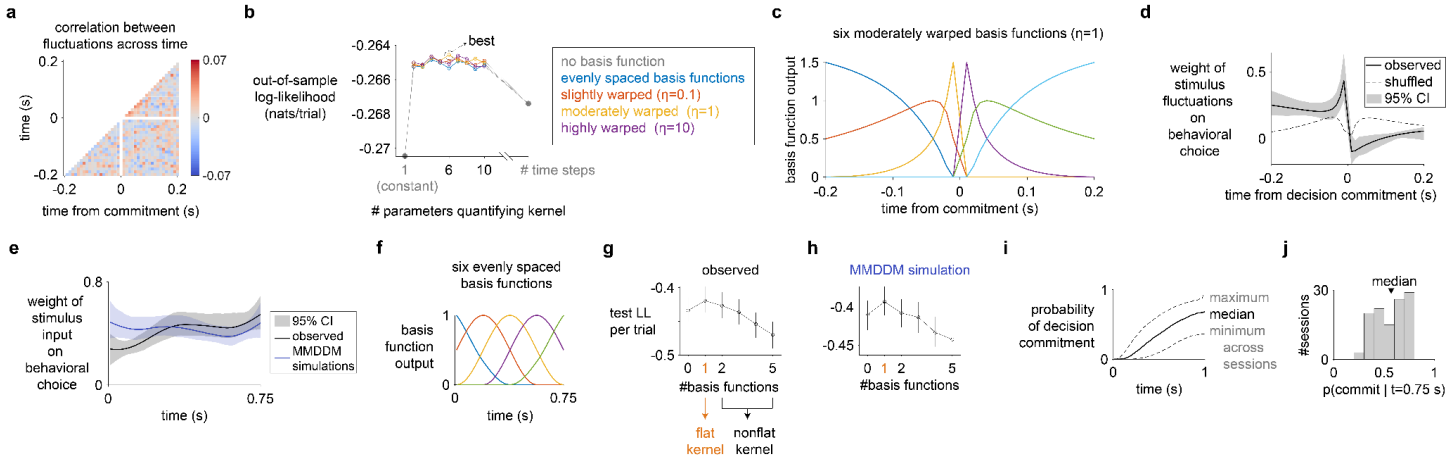
Extended Data Figure 8



Extended Data Figure 8. The time-varying choice selectivity of individual neurons are heterogeneous and are inconsistent with an one-dimensional embedding of an integration-to-bound process. **a**, A computation hypothesized to underlie many decisions is that noisy inputs are accumulated over time through a latent variable (z) until the accumulator variable reaches a fixed bound, which triggers the commitment to a choice. In simulations of bounded integration, evidence accumulates quickly when the evidence strength is strong and more slowly when the strength is weak. **b**, Aggregate responses in each of frontal cortical and striatal regions. Only choice-selective neurons are included; only spikes when the animal remained in fixation (i.e., kept its nose in the center port) are counted; and error trials are excluded. N = 1324 (dmFC), 1076 (mPFC), 1289 (dStr), 714 (vStr), 822 (M1), 163 (FOF). **c**, The responses of a simulated neuron coupled to an integration-to-bound process shows the classic ramping dynamics. Shading indicates the bootstrapped 95% confidence interval of the trial-mean of the filtered response. **d**, A neuron with a classic ramp profile. **e**, A

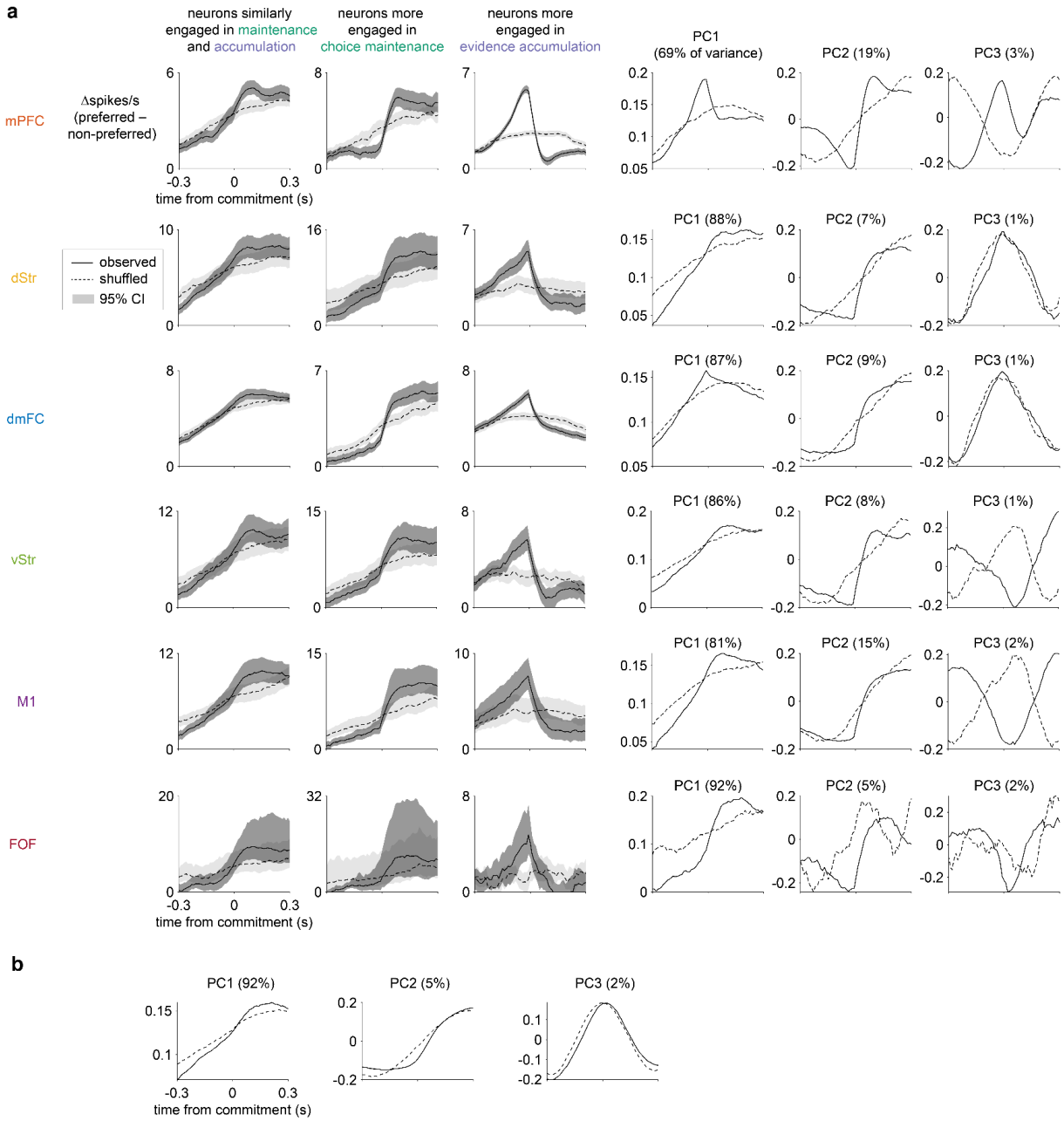
neuron recorded from the session with choice selectivity that decays over time. **f**, A neuron exhibiting a substantial delay in its choice-selective neuronal dynamics. **g**, A neuron whose choice selectivity flips in sign. (**b-h**) Only spikes before movement are counted. **f**, The diversity of the temporal profile of the choice selectivity of individual neurons are inconsistent with a one-dimensional dynamical process.

Extended Data Figure 9



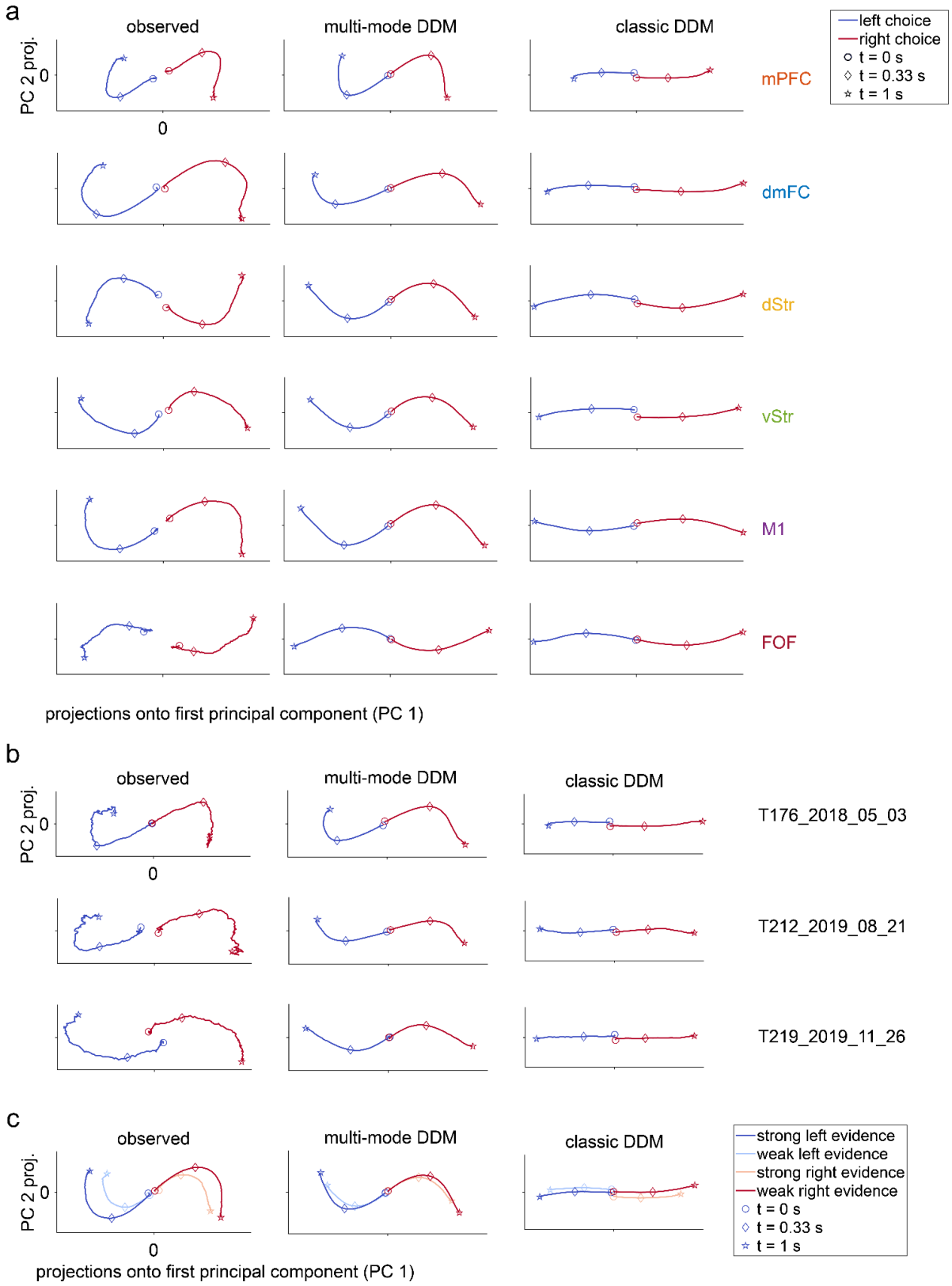
Extended Data Figure 9. Psychophysical kernel model. **a**, For the inferred weights of the stimulus fluctuations to be interpretable, the click input fluctuations must not be strongly correlated across time steps. On each time step on each trial, the fluctuation in auditory click input was computed by counting the observed difference in right and left clicks at that time step, and then subtracting from it the expected difference given the random processes used to generate the stimulus. The input fluctuations at time step of t=0 s were excluded because they are strongly correlated with the input fluctuations before decision commitment and strongly anti-correlated with input fluctuations after commitment. **b**, To determine the time resolution of the kernel that best captures the weight of the input fluctuations, 10-fold cross-validation was performed to compare kernels quantified by different numbers of parameters and types of basis functions. The kernel with the lowest temporal resolution is a constant, represented by a single parameter, implying that fluctuations across time have the same weight. At the highest time resolution, the kernel can be parametrized by a separate weight for each time step. At intermediate time resolution, the kernel is parametrized by basis functions that span the temporal window. The basis functions can be evenly spaced across the temporal window, or stretched such that time near t=0 s is represented with higher resolution and time far from t=0 s with lower resolution. The most likely model had six moderately stretched (η=1) basis functions. **c**, The optimal model's set of six moderately stretched (η=1) basis functions. **d**, The psychophysical kernel inferred by the optimal model is consistent with the hypothesis of a choice maintenance state: the psychophysical weight of the stimulus fluctuations on the behavioral choice decreases to and remains around zero after time of decision commitment. **e**, Mean stimulus onset-aligned psychophysical kernel across sessions, estimated using a model with five temporal basis functions. Shading indicates the 95% bootstrapped confidence interval of the mean. For each session, 10-fold cross-validation was performed on fitting the psychophysical kernel model to the data, and ten estimated kernels are averaged. Then, the kernels are averaged across sessions. **f**, The onset-aligned psychophysical kernel is parametrized by five evenly spaced radial basis functions. **g**, Cross-validated model comparison shows that a temporally flat psychophysical kernel is most likely given the observed data. **h**, Similarly, given the simulated choices generated by the multi-mode drift-diffusion models (MMDDM) fit to the data, the out-of-sample log-likelihood is maximized by assuming a flat kernel. **i**, The approximately flat psychophysical kernel inferred from MMDDM-simulated choices is consistent with the MMDDM's prediction of the probability of decision commitment given the stimulus: throughout the trial, the probability of decision commitment is relatively low, and at no point in the trial is decision commitment an absolute certainty. **j**, At t=0.75 s, the window used to compute the psychophysical kernel, the median probability of decision commitment across sessions is 0.57.

Extended Data Figure 10



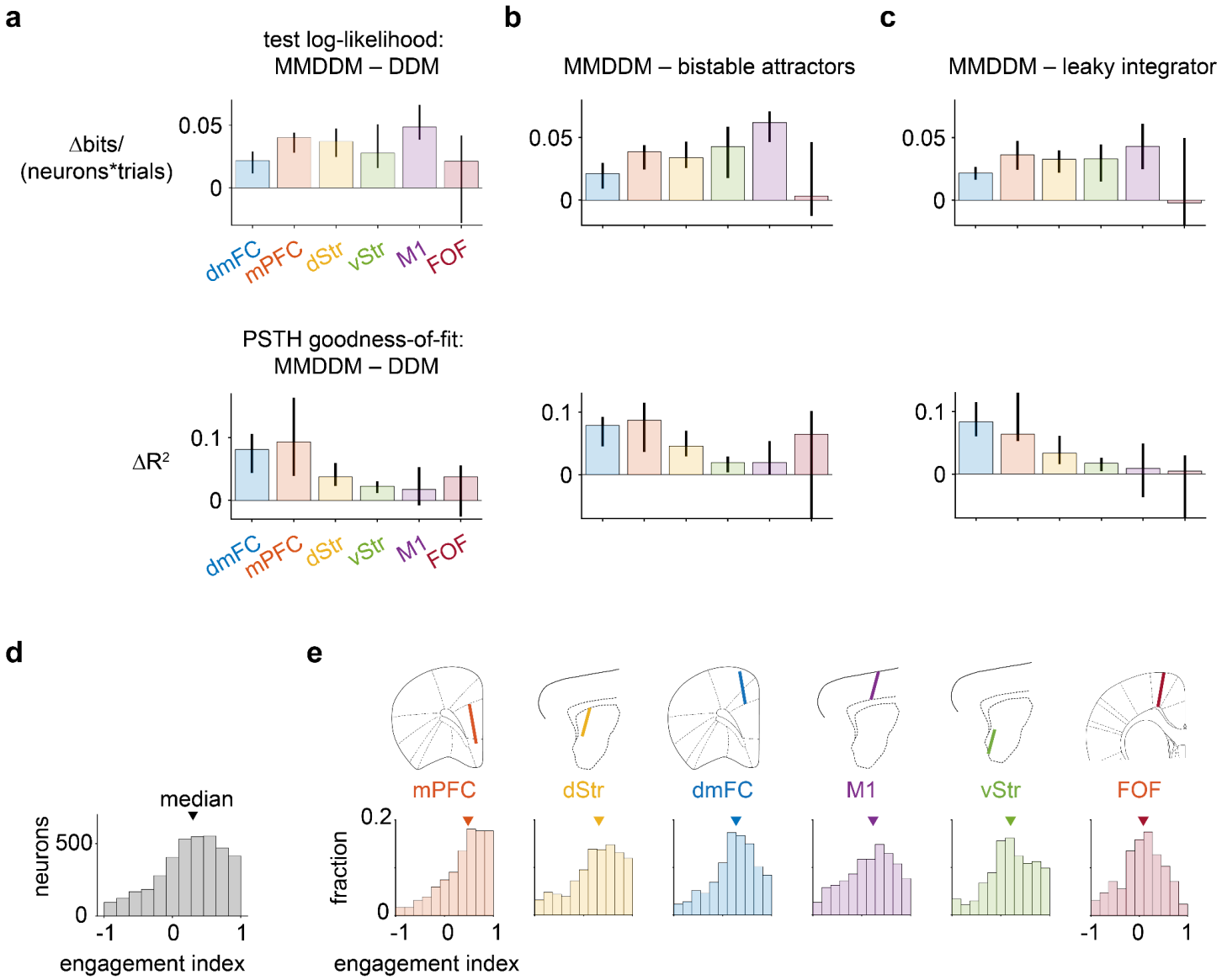
Extended Data Figure 10. The ramp-to-bound, step-like, and ramp-and-decline profiles can be observed in individual brain regions (panel **a**), and the ramp-to-bound, step-like profiles can be observed in neurons not choice-selective enough to be included in model analysis (panel **b**).

Extended Data Figure 11



Extended Data Figure 11. The multi-mode drift-diffusion model (MMDDM) captures curved population trajectories when the analysis is performed on data from **a**, individual brain regions; **b**, individual sessions; and **c**, on trials conditions that depend on not only the choice but also the evidence.

Extended Data Figure 12



Extended Data Figure 12. Dynamics across brain regions. **a**, Top, comparison of the out-of-sample log-likelihood between the MMDDM and the classic DDM. Bottom, comparison of the out-of-sample R^2 of the PSTH across brain regions. Error bar indicates 95% bootstrapped confidence intervals across sessions. **b**, Comparison between MMDDM and bistable attractor model, which is implemented identically to the classic DDM except that the feedback parameter of the latent decision variable is not constrained to be zero but allowed to vary between 0 and 5 and is fit to the data. **c**, Comparison between MMDDM and bistable attractor model, which is implemented identically to the classic DDM except that the feedback parameter of the latent decision variable is not constrained to be zero but allowed to vary between 0 and -5 and is fit to the data. **d**, The median engagement index is not centered at zero, indicating that frontal cortical and striatal neurons are more strongly engaged in evidence accumulation than in choice maintenance. **e**, The differences between the brain regions are apparent in their distribution: whereas the medial prefrontal cortex (mPFC)'s distribution is centered near 1, the distribution for frontal cortical orienting fields (FOF) neurons is centered at zero.

Acknowledgments

We thank Julie Charlton, Long Ding, Joshua Gold, John Maunsell, and Jonathan Pillow for their suggestions and comments. We also thank Jessica Morrison, Klaus Osario, Jovanna Teran, and Emily Valance for technical assistance. This work was supported by grants from NIH F32 MH115416, NIH R01MH108358, and by the Howard Hughes Medical Institute.

References

1. Wong, K.-F. & Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* **26**, 1314–1328 (2006).
2. Dubreuil, A., Valente, A., Beiran, M., Mastrogiuseppe, F. & Ostojic, S. The role of population structure in computations through neural dynamics. *Nat. Neurosci.* **25**, 783–794 (2022).
3. Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
4. Bollimunta, A., Totten, D. & Ditterich, J. Neural dynamics of choice: single-trial analysis of decision-related activity in parietal cortex. *J. Neurosci.* **32**, 12684–12701 (2012).
5. Latimer, K. W., Yates, J. L., Meister, M. L. R., Huk, A. C. & Pillow, J. W. NEURONAL MODELING. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science* **349**, 184–187 (2015).
6. Zoltowski, D. M., Latimer, K. W., Yates, J. L., Huk, A. C. & Pillow, J. W. Discrete Stepping and Nonlinear Ramping Dynamics Underlie Spiking Responses of LIP Neurons during Decision-Making. *Neuron* **102**, 1249–1258.e10 (2019).
7. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
8. Aoi, M. C., Mante, V. & Pillow, J. W. Prefrontal cortex exhibits multidimensional dynamic encoding during decision-making. *Nat. Neurosci.* **23**, 1410–1420 (2020).
9. Yao, J. D. *et al.* Transformation of acoustic information to sensory decision variables in the parietal cortex. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2212120120 (2023).
10. Khona, M. & Fiete, I. R. Attractor and integrator networks in the brain. *Nat. Rev. Neurosci.* **23**, 744–766

(2022).

11. Inagaki, H. K., Fontolan, L., Romani, S. & Svoboda, K. Discrete attractor dynamics underlies persistent activity in the frontal cortex. *Nature* **566**, 212–217 (2019).
12. Gardner, R. J. *et al.* Toroidal topology of population activity in grid cells. *Nature* **602**, 123–128 (2022).
13. Briggman, K. L., Abarbanel, H. D. I. & Kristan, W. B., Jr. Optical imaging of neuronal populations during decision-making. *Science* **307**, 896–901 (2005).
14. Nieh, E. H. *et al.* Geometry of abstract learned knowledge in the hippocampus. *Nature* **595**, 80–84 (2021).
15. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
16. Chaudhuri, R., Gerçek, B., Pandey, B., Peyrache, A. & Fiete, I. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. *Nat. Neurosci.* **22**, 1512–1520 (2019).
17. Wang, X.-J. Decision making in recurrent neuronal circuits. *Neuron* **60**, 215–234 (2008).
18. 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).
19. Finkelstein, A. *et al.* Attractor dynamics gate cortical information flow during decision-making. *Nat. Neurosci.* **24**, 843–850 (2021).
20. Wimmer, K., Nykamp, D. Q., Constantinidis, C. & Compte, A. Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat. Neurosci.* **17**, 431–439 (2014).
21. Pandarinath, C. *et al.* Inferring single-trial neural population dynamics using sequential auto-encoders. *Nat. Methods* **15**, 805–815 (2018).
22. Schneider, S., Lee, J. H. & Mathis, M. W. Learnable latent embeddings for joint behavioural and neural analysis. *Nature* **617**, 360–368 (2023).
23. Roxin, A. & Ledberg, A. Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.* **4**, e1000046 (2008).
24. Usher, M. & McClelland, J. L. The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* **108**, 550–592 (2001).
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. Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95–98 (2013).
27. Hyafil, A. *et al.* Temporal integration is a robust feature of perceptual decisions. *Elife* **12**, (2023).
28. Erlich, J. C., Brunton, B. W., Duan, C. A., Hanks, T. D. & Brody, C. D. Distinct effects of prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat. *Elife* **4**, (2015).
29. Hanks, T. D. *et al.* Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* **520**, 220–223 (2015).
30. Yartsev, M. M., Hanks, T. D., Yoon, A. M. & Brody, C. D. Causal contribution and dynamical encoding in the striatum during evidence accumulation. *Elife* **7**, (2018).
31. Hunnicutt, B. J. *et al.* A comprehensive excitatory input map of the striatum reveals novel functional organization. *Elife* **5**, (2016).
32. Anastasiades, P. G. & Carter, A. G. Circuit organization of the rodent medial prefrontal cortex. *Trends Neurosci.* (2021) doi:10.1016/j.tins.2021.03.006.
33. Sussillo, D., Jozefowicz, R., Abbott, L. F. & Pandarinath, C. LFADS - Latent Factor Analysis via Dynamical Systems. *arXiv [cs.LG]* (2016).
34. Wald, A. & Wolfowitz, J. Bayes Solutions of Sequential Decision Problems. *Proc. Natl. Acad. Sci. U. S. A.* **35**, 99–102 (1949).
35. Kira, S., Yang, T. & Shadlen, M. N. A neural implementation of Wald's sequential probability ratio test. *Neuron* **85**, 861–873 (2015).
36. Ratcliff, R. & McKoon, G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* **20**, 873–922 (2008).
37. Kiani, R., Hanks, T. D. & Shadlen, M. N. Bounded integration in parietal cortex underlies decisions even when viewing duration is dictated by the environment. *J. Neurosci.* **28**, 3017–3029 (2008).
38. Yin, T. C. Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat. *J. Neurosci.* **14**, 5170–5186 (1994).

39. Brown, A. D., Stecker, G. C. & Tollin, D. J. The precedence effect in sound localization. *J. Assoc. Res. Otolaryngol.* **16**, 1–28 (2015).

40. Meister, M. L. R., Hennig, J. A. & Huk, A. C. Signal multiplexing and single-neuron computations in lateral intraparietal area during decision-making. *J. Neurosci.* **33**, 2254–2267 (2013).

41. Park, I. M., Meister, M. L. R., Huk, A. C. & Pillow, J. W. Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nat. Neurosci.* **17**, 1395–1403 (2014).

42. Okazawa, G., Hatch, C. E., Mancoo, A., Machens, C. K. & Kiani, R. Representational geometry of perceptual decisions in the monkey parietal cortex. *Cell* **184**, 3748–3761.e18 (2021).

43. Neri, P., Parker, A. J. & Blakemore, C. Probing the human stereoscopic system with reverse correlation. *Nature* **401**, 695–698 (1999).

44. Okazawa, G., Sha, L., Purcell, B. A. & Kiani, R. Psychophysical reverse correlation reflects both sensory and decision-making processes. *Nat. Commun.* **9**, 3479 (2018).

45. Odoemene, O., Pisupati, S., Nguyen, H. & Churchland, A. K. Visual Evidence Accumulation Guides Decision-Making in Unrestrained Mice. *J. Neurosci.* **38**, 10143–10155 (2018).

46. Inagaki, H. K. *et al.* Neural Algorithms and Circuits for Motor Planning. *Annu. Rev. Neurosci.* **45**, 249–271 (2022).

47. Yates, J. L., Park, I. M., Katz, L. N., Pillow, J. W. & Huk, A. C. Functional dissection of signal and noise in MT and LIP during decision-making. *Nat. Neurosci.* **20**, 1285–1292 (2017).

48. Charlton, J. A. & Goris, R. Abstract deliberation by visuomotor neurons in prefrontal cortex. *bioRxiv* 2022.12.06.519340 (2022) doi:10.1101/2022.12.06.519340.

49. Scott, B. B. *et al.* Fronto-parietal Cortical Circuits Encode Accumulated Evidence with a Diversity of Timescales. *Neuron* **95**, 385–398.e5 (2017).

50. 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).

51. Stine, G. M., Trautmann, E. M., Jeurissen, D. & Shadlen, M. N. A neural mechanism for terminating decisions. *Neuron* (2023) doi:10.1016/j.neuron.2023.05.028.

52. Hernández-Navarro, L., Hermoso-Mendizabal, A., Duque, D., de la Rocha, J. & Hyafil, A. Proactive and

reactive accumulation-to-bound processes compete during perceptual decisions. *Nat. Commun.* **12**, 7148 (2021).

53. Huk, A. C., Katz, L. N. & Yates, J. L. The Role of the Lateral Intraparietal Area in (the Study of) Decision Making. *Annu. Rev. Neurosci.* **40**, 349–372 (2017).
54. Joo, S. J., Katz, L. N. & Huk, A. C. Decision-related perturbations of decision-irrelevant eye movements. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 1925–1930 (2016).
55. Harris, L. R. *et al.* How our body influences our perception of the world. *Front. Psychol.* **6**, 819 (2015).
56. Das, A. & Fiete, I. R. Systematic errors in connectivity inferred from activity in strongly recurrent networks. *Nat. Neurosci.* **23**, 1286–1296 (2020).
57. Ratcliff, R. A theory of memory retrieval. *Psychol. Rev.* **85**, 59–108 (1978).
58. Padoa-Schioppa, C. & Conen, K. E. Orbitofrontal Cortex: A Neural Circuit for Economic Decisions. *Neuron* **96**, 736–754 (2017).
59. Raposo, D., Kaufman, M. T. & Churchland, A. K. A category-free neural population supports evolving demands during decision-making. *Nat. Neurosci.* **17**, 1784–1792 (2014).
60. de Lafuente, V., Jazayeri, M. & Shadlen, M. N. Representation of accumulating evidence for a decision in two parietal areas. *J. Neurosci.* **35**, 4306–4318 (2015).

Methods

Contents

1 Subjects **1**

2 Behavioral task **2**

3 Electrophysiological recording **2**

4 Neuronal choice selectivity **2**

5 Inferring the flow of latent decision variables **2**

6 Intrinsic and input dynamics **3**

7 Multi-mode drift-diffusion model (MMDDM) **4**

8 Peri-stimulus time histogram (PSTH) **5**

9 Trial-averaged trajectories in neural state space **5**

10 Prediction of the time of decision commitment **5**

11 Peri-commitment time histograms (PCTH) **6**

12 Psychophysical kernel model (PKM) **6**

13 Muscimol inactivation **6**

14 Retrograde tracing **6**

15 Histology **6**

16 Statistical tests **7**

1 Subjects

The animal procedures described in this study were approved by the Princeton University Institutional Animal Care and Use Committee and were carried out in accordance with National Institutes of Health standards. All subjects were adult male Long-Evans rats (Taconic, NY) that were pair housed in Technoplast cages and were kept in a 12 hr reversed light-dark cycle. All training and testing procedures were conducted during the dark cycle. Rats had free access to food but had restricted water access. The amount of water that the rats obtained daily was at least 3% of their body weight.

2 Behavioral task

Rats performed the behavioral task in custom-made training enclosures (Island Motion, NY) placed inside sound- and light-attenuated chambers (IAC Acoustics, Naperville, IL). Each enclosure consisted of three straight walls and one curved wall in which three nose ports were embedded (one in the center and one on each side). Each nose port also contained one light-emitting diode (LED) that was used to deliver visual stimuli, and the front of the nose port was equipped with an infrared (IR) beam to detect the entrance of the rat's nose into the port. A loudspeaker was mounted above each of the side ports and was used to present auditory stimuli. Each of the side ports also contained a silicone tube that was used for water reward delivery, with the amount of water controlled by valve opening time.

Rats performed an auditory discrimination task in which optimal performance required the gradual accumulation of auditory clicks [1]. At the start of each trial, rats inserted their nose in the central port and maintained this placement for 1.5 s ("fixation period"). After a variable delay of 0.5-1.3 s, two trains of randomly timed auditory clicks were presented simultaneously, one from the left and one from the right speaker. At the beginning of each click train, a click was played simultaneously from the left and the right speaker a ("stereoclick") Regardless of onset time, the click trains terminated at the end of the fixation period, resulting in stimuli whose duration varied from 0.2-1 s. The train of clicks from each speaker was generated by an underlying Poisson process, with different mean rates for each side. The combined mean click rate was fixed at 40 Hz, and trial difficulty was manipulated by varying the ratio of the generative click rate between the two sides. The generative click rate ratio varied from 39:1 clicks/s (easiest) to 26:14 (most difficult). At the end of the fixation period, rats could orient towards the nose port on the side where more clicks were played and obtain a water reward.

Psychometric functions were computed by dividing trials into eight similarly sized groups according to the total difference in the right and left clicks, and for each group, computing the fraction of trials ending in a right choice. The confidence interval of the fraction of right response was computed using the Clopper-Pearson method.

3 Electrophysiological recording

Neurons were recorded using chronically implanted Neuropixels 1.0 probes that are recoverable after the experiment [4]. In each of four animals, a probe was implanted at 4.0 mm anterior to Bregma, 1.0 mm lateral, for a distance of 4.2 mm, and at an angle of 10 degrees relative to the sagittal plane intersecting the insertion site, such that the probe tip was more medial than the probe base. In each of five other animals, a probe was implanted to target primary motor cortex, dorsal striatum, and ventral striatum, at the site 1.0 mm anterior, 2.4 mm lateral, for a distance of 8.4 mm, and at an angle of 15 degrees relative to the coronal plane intersecting the insertion site, such that the probe tip was more anterior than the probe base. In each of three final set of rats (Gupta et al., in preparation), a probe was implanted to target the frontal orienting fields and anterior dorsal striatum at 1.9 mm anterior, 1.3 mm lateral, for a distance of 7.4 mm, and at angle of -10 degree relative to the sagittal plane intersecting the insertion site, such that the probe tip was more lateral than the probe base. Spikes were sorted into clusters using Kilosort2 [6], and clusters were manually curated. Different electrodes were recorded from across different days.

4 Neuronal choice selectivity

Only neurons that meet a pre-selected threshold for choice selective are included for analysis. For each neuron, choice selectivity is measured using the area under the receiver operating characteristic (aUROC) indexing how well an ideal observer can classify between a left- and right-choice trial based on the spike counts of the neuron. Spikes were counted in four non-overlapping time windows (0.01-0.21 s after stimulus onset, 0.21-0.4, 0.41-0.6, and 0.61-0.9), and an auROC was computed for each time window. A neuron with an auROC < 0.42 or > 0.58 for any of the these windows is deemed choice selective and included for analysis.

5 Inferring the flow of latent decision variables

Detailed description are provided in [3]. Briefly, to infer flow fields from the neural population spike trains, we used a sequential variational autoencoder (SVAE) called Flow-field inference from neural data using deep recurrent networks (FINDR).

At each time step ($\Delta t = 0.01s$), the firing rates of N simultaneously recorded neurons $\boldsymbol{r}_t$ are given by

$$\boldsymbol{r}_t = \text{softplus}(\boldsymbol{W} z_t + \boldsymbol{b}_t) \tag{1}$$

where softplus is a function approximating the firing rate-synaptic current relationship (f-I curve) of neurons, z_t is a D-dimensional vector representing the latent decision variable, $\boldsymbol{W}$ a $N \times D$ matrix representing the encoding weights, $\boldsymbol{b}_t$ a N-dimensional vector representing the the decision-irrelevant baseline input. Conditioned on the value of the latent decision variable, the observed spike train response of each neuron at each time step is modeled as a Poisson random variable.

The discretized time derivative (or the flow) of the decision variable is given by

$$\frac{z_{t+\Delta t} - z_t}{\Delta t} = F(z_t, \boldsymbol{u}_t) \tag{2}$$

where the input $\boldsymbol{u}_t$ is a two-dimensional vector representing the number of left and right clicks played on a time step. We obtain $z_{t+\Delta t}$ from z_t and $\boldsymbol{u}_t$ using

$$z_{t+\Delta t} = z_t + \Delta t F(z_t, \boldsymbol{u}_t) + \eta_t \tag{3}$$

where $\eta_t \sim \mathcal{N}(0, \Delta t \Sigma)$. Σ is a D-dimensional diagonal matrix. F is parametrized by a gated feedforward neural network, and Eq. (3) is an Euler-discretized gated neural stochastic differential equations (gnSDE) [2]. FINDR learns $\boldsymbol{W}$, $\boldsymbol{b}$, Σ, and the parameters of the neural network representing F, by maximizing the approximate lower bound of the log-likelihood of the simultaneously recorded spike trains.

During optimization, FINDR learns an encoding function which transforms spiking neural data and experimental stimuli into low-dimensional dynamical trajectories, and a decoding function that learns to reconstruct the spiking neural data from the low-dimensional trajectories. In addition, FINDR simultaneously infers the flow field in the latent state space that best describes the learned low-dimensional trajectories. During training, FINDR uses stochastic gradient descent (SGD) to minimize both the reconstruction of spiking neural data from low-dimensional dynamical trajectories, and the distance between the distribution of the inferred trajectories and the distribution of trajectories generated from the flow field.

After optimization, because the same $\boldsymbol{r}_t$ can be given by different values of $\boldsymbol{W}$ and z_t, to uniquely identify the latent dynamics, we define

$$\tilde{z}_t = S V^\top z_t \tag{4}$$

where S is a $D \times D$ diagonal matrix containing the singular values of the inferred matrix W and V a $D \times D$ matrix containing the right singular vectors. To obtain meaningful axes for the transformed latent space, we generate 5000 different trajectories of $\tilde{z}$, and perform principal component analysis (PCA) on the trajectories. The principal components (PCs) were used to define the dimensions of the decision variable $\tilde{z}$. In all of our analyses, the latent trajectories and flow fields inferred by FINDR are in the transformed latent space.

6 Intrinsic and input dynamics

Our dynamical system can be written as

$$\dot{z} = F(z, \boldsymbol{u}), \tag{5}$$

where we separate the components of the dynamics $\dot{z}$ into the intrinsic, or input-independent, component:

$$\dot{z}_{\text{intrinsic}} = F(z, \boldsymbol{0}) \tag{6}$$

and the input-driven component:

$$\dot{z}_{\text{input}} = F(z, \boldsymbol{u}) - F(z, \boldsymbol{0}). \tag{7}$$

Note that $\dot{z} = \dot{z}_{\text{intrinsic}} + \dot{z}_{\text{input}}$. If the sensory stimulus $\boldsymbol{u}$ is presented continuously throughout the trial, Eq. (7) indeed describes the direction and magnitude of input (i.e., input vector) at a particular state z. For example, in the standard random dot motion (RDM) task, the coherence of the dots (or the evidence strength) may be represented as a 1-dimensional variable $\boldsymbol{u} \in [-1, 1]$, where -1 indicates 100% coherence leftward motion and 1 indicates 100% coherence rightward motion. However, if $\boldsymbol{u}$ is discrete and presented at random times in the trial, as in our task, $\boldsymbol{u}$ may be represented as a 2-dimensional variable where $\boldsymbol{u} = \boldsymbol{0} = [0; 0]$ when there is no click, $\boldsymbol{u} = [1; 0]$ when there is a

left click, and $\boldsymbol{u} = [0; 1]$ when there is a right click. The crucial difference between the RDM task and our task is that in the RDM task, $\boldsymbol{u}$ stays fixed throughout the trial and encodes evidence strength, while in our task, $\boldsymbol{u}$ changes throughout the trial depending on whether there was a click or not, and does not itself encode evidence strength. What this means is that, for example, in our task, $\dot{z} = F(z, [1; 0]) - F(z, [0; 0])$ is similar to presenting a strong leftward evidence continuously throughout the trial (similar to giving a 100% coherence leftward motion in the RDM task). To correct for the fact that the value that $\boldsymbol{u}$ can take discretely and stochastically changes across time steps in a trial, and thus $p(\boldsymbol{u}|z)$ is different across the state space of z, we multiply $p(\boldsymbol{u}|z)$ to Eq. (7):

$$\dot{z}_{\text{input}} = p(\boldsymbol{u}|z)[F(z, \boldsymbol{u}) - F(z, \mathbf{0})]. \tag{8}$$

More specifically,

$$\begin{aligned} \dot{z}_{\text{left}} &= p(\boldsymbol{u} = [1; 0]|z)[F(z, \boldsymbol{u} = [1; 0]) - F(z, \boldsymbol{u} = [0; 0])], \\ \dot{z}_{\text{right}} &= p(\boldsymbol{u} = [0; 1]|z)[F(z, \boldsymbol{u} = [0; 1]) - F(z, \boldsymbol{u} = [0; 0])]. \end{aligned} \tag{9}$$

Note that we have separated the dynamics $\dot{z}$ into three distinct components – intrinsic dynamics, leftward input dynamics and rightward input dynamics. Adding Eq. (6) and Eq. (9), i.e., $\dot{z} = \dot{z}_{\text{intrinsic}} + \dot{z}_{\text{left}} + \dot{z}_{\text{right}}$, we have

$$\dot{z} = p(\boldsymbol{u} = [0; 0]|z)F(z, \boldsymbol{u} = [0; 0]) + p(\boldsymbol{u} = [1; 0]|z)F(z, \boldsymbol{u} = [1; 0]) + p(\boldsymbol{u} = [0; 1]|z)F(z, \boldsymbol{u} = [0; 1]). \tag{10}$$

Therefore, $\dot{z}$ gives us the overall dynamics $F(z, \boldsymbol{u})$ with an "effective" $\boldsymbol{u}$ that normalizes across the three possible values that $\boldsymbol{u}$ can take in a given z: $[0; 0]$, $[1; 0]$, $[0; 1]$.

To estimate $p(\boldsymbol{u}|z)$, we generate click trains for 5000 trials in a way that is similar to how clicks are generated for the task done by our rats. Then, we simulate 5000 latent trajectories from F and the generated click trains. We then bin the state space of z and ask, for a single bin, how many times the latent trajectories cross that bin in total and how many of the latent trajectories, when crossing that bin had $\boldsymbol{u} = [1; 0]$ or $\boldsymbol{u} = [0; 1]$. That is, we estimate $p(\boldsymbol{u} = [1; 0]|z)$ with $\frac{\text{\# of latent states with } \boldsymbol{u}=[1;0] \text{ in the bin that covers } z}{\text{\# of latent states in the bin that covers } z}$. For Fig. 2 in the main text, because z is 2-dimensional, we use bins of 8-by-8 that cover the state space traversed by the 5000 latent trajectories, and weigh the flow arrows of the input dynamics with the estimated $p(\boldsymbol{u}|z)$. Similarly, for the background shading that quantifies the speed of input dynamics in Fig. 2, we use bins of 100-by-100 to estimate $p(\boldsymbol{u}|z)$, and apply a Gaussian filter with $\sigma = 2$ to smooth the histogram. A similar procedure was done in Extended Data Fig. 3 and 6 to estimate $p(\boldsymbol{u}|z)$.

7 Multi-mode drift-diffusion model (MMDDM)

A detailed account is provided in the Supplementary Information. Briefly, the MMDDM is an autoregressive, input-output hidden-Markov model (HMM) whose emissions consists of the spike train responses of simultaneously recorded neurons and the behavioral choice. The latent variable, the accumulator, has transition dynamics that are modelled as a Orstein-Uhlenbeck process [1] to express drift-diffusion dynamics. Unlike typical HMM's, the transition matrix is not fixed across time steps but varies depending on whether an sensory input has occurred. The transition matrix at each time step is computed by using the Fokker-Planck equations to obtain the conditional forward propagation distributions and then discretizing the forward distributions. The transition matrix (as well as the prior probability vector of the latent accumulator variable) is specified by only three free parameters–bound height, initial value, and the signal-to-noise ratio of the input–that are learned separately for each recording session.

The emissions of the MMDDM are the spike train responses at each time step of the all simultaneously recorded neurons and the behavioral choice. On each time step, given the value of the accumulator, the conditional spike train response is modelled as Poisson generalized linear models (GLM). The GLM allows us to mitigate potential omitted variable bias by incorporating nuisance variables as input to the neural spike trains. These nuisance variables include the spike history and choice-independent baseline changes aligned to stimulus onset and movement onset. Slow fluctuations in baseline across trials are also incorporated. The spike train response at each time step depends on the accumulator value on only that time step and no other time step.

Each neuron relates to the accumulator through two scalar weights, w_{EA} and w_{CM} , that specifies the encoding of the accumulator during the evidence integration phase and the choice maintenance phase, respectively. When the accumulator has not yet reached the bound, all simultaneously recorded neurons encode the accumulator through their own w_{EA}, and when the bound is reached, their own w_{CM}. Separate w_{EA} and w_{CM} are learned for each neuron.

The behavioral choice depends on the accumulator value on only the last time step of the trial. Given the value of the accumulator at the last time step of the trial, the conditional behavioral choice is modelled as a Bernoulli random variable. The probability of a right choice is the sum of the probabilities of the accumulator $a > 0$.

All parameters, including the three parameters of the accumulator, as well as the eleven parameters private to each neuron, are learned simultaneously by jointly fitting to all spike trains and choices. To fit the model, the analytic gradient of the log-likelihood is computed using the forward-backward algorithm. Models were fit using five-fold cross-validation.

The engagement index (EI) was computed for each neuron on the basis of its w_{EA} and w_{CM}: $EI \equiv (|w_{EA}| - |w_{CM}|)/(|w_{EA}| + |w_{CM}|)$. The index is based on the widely used modulation index used in analyses in systems neuroscience. It ranges from -1 to 1. A neuron with an EI near -1 encodes the accumulator only during choice maintenance; a neuron with an EI near 0 encodes the accumulator similarly during evidence accumulation and choice maintenance; and an EI near 1 indicates encoding only during evidence accumulation.

8 Peri-stimulus time histogram (PSTH)

Spike times were binned at 0.01 s and included spikes up to one second after the onset of the auditory stimulus (click trains) until one second after the stimulus onset, or until when the animal removed its nose the central port, whichever came first. The time-varying firing rate of of each neuron in each group of trials (i.e., task condition) was estimated with a peristimulus time histogram (PSTH), which was computed by convolving the spike train on each trial with a causal gaussian linear filter with a standard deviation of 0.1 s and a width of 0.3 s and averaging across trials. The confidence interval of a PSTH was computed by bootstrapping across trials. The normalized PSTH of each neuron in each task condition was computed by dividing the PSTH by the mean firing rate of that neuron across all time steps across all trials. Trials were grouped by the animal's choice and/or the generative click rate ratio. A "preferred" task condition was defined for each neuron as the group of trials with the behavioral choice when the neuron responded more strongly, and a "null" task condition was defined as the trials associated with the other choice. The aggregate PSTHs across neurons was computed by averaging across the PSTH for either the preferred or null task condition. The time-varying choice selectivity of each neuron was computed by subtracting the null PSTH from the preferred PSTH, and then dividing the difference by the maximum. When the choice dynamics of multiple neurons are shown, neurons are sorted by the center of mass of its absolute value of the choice selectivity.

9 Trial-averaged trajectories in neural state space

To measure the trial-averaged dynamics in neural state space, principal component analysis (PCA) was performed on a data matrix made by concatenating the peri-stimulus time histograms. The data matrix X has dimensions $T \times C$-by-N, where T is the number of time steps ($T = 100$), C is the number of task conditions ($C = 2$ for choice-conditioned PSTHs, $C = 4$ for PSTHs conditioned on both choice and evidence strength), and N is the number of neurons. The mean across rows is subtracted from X, and singular value decomposition is performed: $USV^\top = X$. The principal axes correspond to the columns of the right singular matrix V, and the projections of the original data matrix X onto the principal axes correspond to the left singular matrix (U) multiplied by S, the rectangular diagonal matrix of singular values. The first two columns of the projections US are plotted as the trajectories in neural state space.

10 Prediction of the time of decision commitment

The time step when decision commitment occurred is selected to be when the posterior probability of the accumulator at either the left or the right bound, given the learned MMDDM parameters, click times, spike trains, and behavioral choice, to be greater than 0.8. Results were similar for other thresholds, and the particular threshold of 0.8 was chosen to balance between the accuracy of the prediciton and the number of trials for which commitment is predicted to have occurred. Under definition, commitment occurred on 34.6% of the trials. The MMDDM parameters were learned using training data in five-fold cross-validation.

11 Peri-commitment time histograms (PCTH)

On trials for which a time of decision commitment could be detected (see section 10), the spike train are aligned to the predicted time of commitment and then averaged across those trials. The trial-average is then filtered with a causal gaussian kernel with a standard deviation of 0.05s. The PCTH were averaged across three groups of neurons: 1) neuorns that are similar engaged in evidence accumulation and choice maintenance; 2) neurons more strongly engaged in evidence accumulation; and 3) neurons more strongly engaged in choice maintenance. Each neuron was assigned to one of these groups according to its engagement index (EI; see section 7). Neurons with $-\frac{1}{3} \leq EI < \frac{1}{3}$ are considered to be similarly engaged in evidence accumulation and choice maintenance; neurons with $EI \geq \frac{1}{3}$ are considered to be more strongly engaged in evidence accumulation; and those with $EI < -\frac{1}{3}$ are considered to be more strongly engaged in choice maintenance. Principal component analysis on the PCTH's are performed as described in section 9.

For this analysis, we focused on only the 65/115 sessions for which MMDDM improves the R2 of the PSTH's and the inferred encoding weights are reliable across cross-validation folds ($R^2 > 0.9$). From this subset of sessions, there were 1,116 neurons similarly engaged in maintenance and accumulation, 414 neurons that are more engaged in maintenance, and 1,529 neurons more engaged in accumulation.

To compute the shuffled PCTH, the predicted times of commitment are shuffled among only the trials on which commitment were detected. If the randomly assigned commitment time extends beyond the length of the trial, then the time of commitment is assigned to be the last time step of that trial.

12 Psychophysical kernel model (PKM)

We extended the logistic model [5] for inferring the psychophysical kernel to include a lapse parameter and also temporal basis functions to parametrize the kernel. Cross-validated model comparison was performed to select the optimal type and number of temporal basis functions. A detailed account of the model and the inference method is provided in the Supplemental Information.

13 Muscimol inactivation

Infusion cannulas (Invivo1) were implanted bilaterally over dorsomedial frontal cortex (4.0 mm AP, 1.2 mm ML). After the animal recovered from surgery, the animal is anesthetized and on alternate days, a 600 nL solution of either only or 150 nG of muscimol was infused in each hemisphere. Half an hour after the animal wakes up from anesthesia, the animal is allowed to perform the behavioral task.

14 Retrograde tracing

To characterize the anatomical inputs into dorsal striatum, 50 nL of cholera toxin subunit B conjugate (ThermoFisher Scientific) was injected into dorsal striatum at 1.9 mm AP, 2.4 ML, 3.5 mm below the cortical surface. The animal was perfused seven days after surgery.

15 Histology

The rat was fully anesthetized with 0.4 mL ketamine (100 mg/ml) and 0.2 mL xylazine (100 mg/ml) IP, followed by transcardial perfusion of 100 mL saline (0.9% NaCl, 0.3x PBS, pH 7.0, 0.05 mL heparin 10,000 USP units/mL), and finally transcardial perfusion of 250 mL 10% formalin neutral buffered solution (Sigma HT501128). The brain was removed and post fixed in 10% formalin solution for a minimum period of 7 days. 100 micrometer sections were prepared on a Leica VT1200S vibratome, and mounted on Superfrost Pus glass slides (Fisher) with Fluoromount-G (Southern Biotech) mounting solution and glass cover slips. Images were acquired on a Hamtasu NanoZoomer under 4x magnification.

16 Statistical tests

All confidence intervals were computed using the bias corrected and accelerated percentile method [7]. P-values comparing medians were computed using a two-sided Wilcoxon rank sum test, which tests the null hypothesis that two independent samples are from continuous distributions with equal medians, against the alternative that they are not.

References

[1] B. W. Brunton, M. M. Botvinick, and C. D. Brody. Rats and humans can optimally accumulate evidence for decision-making. *Science*, 340(6128):95–98, 2013.

[2] Timothy Doyeon Kim, Tankut Can, and Kamesh Krishnamurthy. Trainability, Expressivity and Interpretability in Gated Neural ODEs. *Proceedings of the 40th International Conference on Machine Learning*, 2023.

[3] Timothy Doyeon Kim, Thomas Zhihao Luo, Tankut Can, Kamesh Krishnamurthy, Jonathan W. Pillow, and Carlos D. Brody. Flow field inference from neural data using deep recurrent networks. *In preparation*.

[4] Thomas Zhihao Luo, Adrian Gopnik Bondy, Diksha Gupta, Verity Alexander Elliott, Charles D. Kopec, and Carlos D. Brody. An approach for long-term, multi-probe neuropixels recordings in unrestrained rats. *Elife*, 2020.

[5] O. Odoemene, S. Pisupati, H. Nguyen, and A. K. Churchland. Visual evidence accumulation guides decision-making in unrestrained mice. *Journal of Neuroscience*, 2018.

[6] M. Pachitariu, N. A. Steinmetz, S. N. Kadir, M. Carandini, and K. D Harris. Fast and accurate spike sorting of high-channel count probes with kilosort. *Advances in neural information processing systems*, 29, 2016.

[7] Ryan J. Tibshirani and Bradley Efron. *An Introduction to the Bootstrap*. Chapman Hall/CRC, 1994.

Supplementary information

Contents

1 Simulations of the canonical and non-canonical attractor hypotheses **1**

2 Multi-mode drift-diffusion model **2**
2.1 Latent accumulator variable . . . 2
2.1.1 Transition probability . . . 3
2.1.2 Prior probability . . . 3
2.1.3 Discretization of time and accumulated evidence . . . 4
2.1.4 Sensory adaptation . . . 5
2.1.5 Latency . . . 7
2.2 Spike train response . . . 7
2.2.1 Accumulator-dependent predictor . . . 7
2.2.2 Accumulator-independent, within-trial varying predictors . . . 7
2.2.3 Accumulator-independent, across-trial varying predictor . . . 8
2.3 Behavioral choice . . . 8
2.4 Inference . . . 8
2.4.1 Gradient of the log-likelihood . . . 9

3 Psychophysical kernel model **9**
3.1 Model definition . . . 10
3.2 Inference . . . 10
3.3 Data used to fit the model . . . 11

4 Basis functions **11**
4.1 Time warping . . . 11

1 Simulations of the canonical and non-canonical attractor hypotheses

The dynamical equations for each of the attractor hypotheses used to generate the flow fields in Fig. 1 and Extended Data Fig. 1, and synthetic datasets in Extended Data Fig. 3 are as follows.

Bistable attractors:

$$\begin{aligned} dz_1 &= 10z_1(0.7 + z_1)(0.7 - z_1)dt + cudt + dW \\ dz_2 &= -10z_2 dt + dW \end{aligned} \tag{1}$$

Classic DDM:

$$\begin{aligned} dz_1 &= \begin{cases} cudt + 0.5dW, & z_1 \in (-0.7, 0.7) \\ 10z_1(0.7 - z_1)(0.7 + z_1)dt + 0.5dW, & z_1 \notin (-0.7, 0.7) \end{cases} \\ dz_2 &= -30z_2 + 0.5dW \end{aligned} \tag{2}$$

RNN line attractor:

$$\begin{aligned} dz_1 &= 5z_2 + dW \\ dz_2 &= -5z_2 dt + cudt + dW \end{aligned} \tag{3}$$

Non-canonical:

$$\begin{aligned} dz_1 &= 5z_1(0.85 + z_1)(0.85 - z_1)dt + cudt + dW \\ dz_2 &= 5(0.5|z_1| + 0.1)(z_1 - 1.2z_2) + dW \end{aligned} \tag{4}$$

Here, u at each time point can take a value from the set $\{-1, 0, 1\}$, where -1 indicates a leftward click, 0 indicates no click, and 1 indicates a rightward click. $c = 1/10$ for RNN line attractor, and $c = 1/40$ for others. Similar to the task done by our rats, for each trial, γ was drawn from the set $\{-3.5, -2.5, -1.5, -0.5, 0.5, 1.5, 2.5, 3.5\}$. We generated left and right clicks with rates $40/(1 + \exp(\gamma))$ and $40/(1 + \exp(-\gamma))$, respectively. The total duration of each trial was $1s$, and the duration of the click stimulus was randomly chosen to be one of $0.5s$, $0.7s$, and $0.9s$ for each trial. We generated 500 trajectories from each of Eq. (1–4), and assumed that the initial condition $\boldsymbol{z}(0) = \mathbf{0}$. Each trajectory corresponds to a trial. The spikes were generated from

$$\begin{aligned} C_{ij} &\sim 1 + 5\mathcal{N}(0, 1) \\ r_i &= \text{softplus}\left(\sum_j C_{ij} z_j + 5\right) \\ x_i &\sim \text{Poisson}(\Delta t \cdot r_i) \end{aligned} \tag{5}$$

where $C \in \mathbb{R}^{80\times 2}$, and x_i is the number of spikes in a time bin with width $\Delta t = 0.01s$. Therefore, a total of 80 neurons and 500 trials were simulated for each dataset.

2 Multi-mode drift-diffusion model

2.1 Latent accumulator variable

The multi-mode drift-dffusion model supposes that a single, one-dimensional stochastic dynamical process to the choice-related dynamics of all simultaneously recorded neurons across multiple brain regions, as well as the animal's behavioral choice. Following previous studies[3, 4], the decision process is modelled as a one dimensional bounded drift-diffusion process:

$$da = \begin{cases} 0, & a \in \{-B, B\} \\ (\lambda a + m)dt + \sigma_a W, & a \in (-B, B) \end{cases} \tag{6}$$

When a is at either the left bound $-B$ or the right B, its value does not change. The bound height parameter B is learned from the data.

The hyperparameter λ quantifies the consistent drift in a. For $\lambda < 0$, the integration is leaky, a drifts toward 0, and early inputs affect a less than later inputs. For $\lambda > 0$, a positively feedbacks to itself, a is impulsive, and early inputs affect a more than later inputs. The time constant of a is $1/\lambda$. In the case of $\lambda = 0$, a is an approximate line attractor. Based on the results from the flow-field inference, λ is set to be 0.

The momentary input $m(t)$ at each time step is given by by the difference between the input from the right and left clicks:

$$m(t) = c_R(t) - c_L(t) \tag{7}$$

The input from each click is given by

$$c(t) = \delta(\boldsymbol{\tau} - t) \cdot \eta \cdot C(t) \tag{8}$$

where $\delta(x)$ is the Direct delta function, which has a value of 0 everywhere except for $\delta(x = 0)$, which occurs when x is equal to one of the elements in the set $\boldsymbol{\tau}$ containing the times of either the left or right clicks. For $\tau \in \boldsymbol{\tau}$

$$\int_{t-dt/2}^{t+dt/2} \delta(\tau - t)dx = 1 \tag{9}$$

The momentary input is corrupted by multiplicative noise η, a sample from the Gaussian distribution:

$$\eta \sim \mathcal{N}(1, \sigma_s) \tag{10}$$

The parameter σ_s quantifies the signal-to-noise ratio of the sensory input and is optimized during model fitting. To account for sensory adaptation, the momentary input is depressed by preceding clicks, and its magnitude is multiplicatively scaled by $C(t)$ (see below).

In the absence of a consistent drift ($\lambda = 0$) and no momentary input, the value of a drifts stochastically due to the diffusion noise $\sigma_a W(t)$, which is a Weiner process scaled by the hyperparameter σ_a:

$$W(t) - W(t - dt) \sim \mathcal{N}(0, dt) \tag{11}$$

Following previous findings [3], diffusion noise is fixed to be small by setting $\sigma_a = 1$.

2.1.1 Transition probability

The solution to the Fokker-Planck equations in Eq. 6 provides the transition probability of the latent variable:

$$p(a_t \mid a_{t-1}) = \begin{cases} 1, & a_{t-1} = a_t = -B \\ 1, & a_{t-1} = a_t = B \\ f(a_t \mid \mu_t, \sigma_t^2), & (a_{t-1}, a_t) \in (-B, B) \\ \int_{-\infty}^{-B} f(x \mid \mu_t, \sigma_t^2)\, dx, & a_t = -B, a_{t-1} \in (-B, B) \\ \int_{B}^{\infty} f(x \mid \mu_t, \sigma_t^2)\, dx, & a_t = B, a_{t-1} \in (-B, B) \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

If the latent variable is at either of the two absorbing bounds in the previous time step, then the latent variable must retain the same value for the current time step. If both the previous and present values are within the bounds, the probability is given by a Gaussian distribution with mean μ_t and variance σ_t^2:

$$f(x \mid \mu_t, \sigma_t^2) \sim \mathcal{N}(\mu_t, \sigma_t^2) \tag{13}$$

$$\mu_t = e^{\lambda dt}\left(a_{t-1} + \frac{m_t}{\lambda dt}\right) - \frac{m_t}{\lambda dt} \tag{14}$$

$$\sigma_t^2 = \sigma_a^2 + \sigma_s^2 |c_t| \tag{15}$$

where c_t is the magnitude of momentary input from either a left or a right click. In the limit of λ toward 0, the mean μ_t is given by

$$\lim_{\lambda \to 0} \mu_t = a_{t-1} e^{\lambda dt} + m_t \left(1 - e^{-\lambda dt}\right) \tag{16}$$

2.1.2 Prior probability

The prior probability of the latent variable is multi-mode distribution consisting a mode given by a Gaussian distribution with mean μ_1 and variance σ_a^2,

$$f(x \mid \mu_1, \sigma_1^2) \sim \mathcal{N}(\mu_1, \sigma_1^2) \tag{17}$$

a left mode at the lower bound, and a third mode at the upper bound:

$$p(a_1) = \begin{cases} f(a_1 \mid \mu_1, \sigma_1^2), & B > a_1 > -B \\ \int_{-\infty}^{-B} f(x \mid \mu_1, \sigma_1^2)\, dx, & a_1 = -B \\ \int_{B}^{\infty} f(x \mid \mu_1, \sigma_1^2)\, dx, & a_1 = B \end{cases} \tag{18}$$

The mean of the Gaussian distribution is specified by the parameter $\mu_1 \in (-5, 5)$, and the variance is a hyperparameter fixed to $\sigma_1^2 = 1$.

2.1.3 Discretization of time and accumulated evidence

Similar to previous studies[3, 4], the model is implemented by discretizing both time and the latent variable a. Time is discretized to $\Delta t = 0.01$s, and the accumulated evidence a is discretized to $\Xi = 53$ values equally spaced in $[-B, B]$. The i-th bin ξ_i is given by

$$\xi_i = \frac{2i - \Xi - 1}{\Xi - 1} B \tag{19}$$

The difference between successive bins is

$$\Delta\xi = \frac{2B}{\Xi - 1} \tag{20}$$

Under this discretization scheme, the probability $p(a)$ at each time step is represented by a Ξ-element probability vector $\boldsymbol{p}$ whose first element indicates the probability of a at the left bound, the last element the probability at the right bound, and the intervening elements correspond to the approximate probability at each bin. By definition of being of a probability vector, the elements of $\boldsymbol{p}$ always sum to 1. For each accumulator value a, the probability $p(a)$ is divided between the two closest bins, weighted by the relative distance between a and the two bins.

Prior probability vector To illustrate the discretization in detail, below we demonstrate the computation of the prior probability vector. Each element of the probability vector is computed as a definite integral of the normal distribution in Eq. 18. For the first element of the prior probablity vector, which quantifies the probability of a at the left bound:

$$\boldsymbol{p}_{i=1} = \int_{-\infty}^{\xi_1} f(x; \mu_1, \sigma_1^2)\, dx + \frac{1}{\Delta\xi}\left[\int_{\xi_1}^{\xi_2} f(x; \mu_1, \sigma_1^2)\,(\xi_2 - x)\, dx\right] \tag{21}$$

The definite integral is equal to

$$\boldsymbol{p}_{i=1} = \Phi(z_1) + \frac{\sigma_1}{\Delta\xi}\left\{z_2\left[\Phi(z_2) - \Phi(z_1)\right] + \phi(z_2) - \phi(z_1)\right\} \tag{22}$$

where z_i is the z-score

$$z_i = \frac{\xi_i - \mu_1}{\sigma_1}$$

and $\Phi(x)$ and ϕ are the standard normal cumulative distribution function and standard normal probability density function, respectively. The mean μ_1 and the standard deviation σ_1 are parameters of the model. For the second to second-to-last elements of the probability vector, i.e., $i \in \{2, ...\Xi - 1\}$:

$$\boldsymbol{p}_i = \frac{\sigma_1}{\Delta\xi}\left\{z_{i+1}\left[\Phi(z_{i+1}) - \Phi(z_i)\right] + z_{i-1}\left[\Phi(z_{i-1}) - \Phi(z_i)\right] + \phi(z_{i+1}) + \phi(z_{i-1}) - 2\phi(z_i)\right\} \tag{23}$$

Finally, for the last element:

$$\boldsymbol{p}_\Xi = 1 - \Phi(z_\Xi) + \frac{\sigma_1}{\Delta\xi}\left\{z_{\Xi-1}\left[\Phi(\xi_{\Xi-1}) - \Phi(\xi_\Xi)\right] + \phi(z_{\Xi-1}) - \phi(z_\Xi)\right\} \tag{24}$$

Stochastic matrix The transition probability of the latent variable a in Eq. 12 is represented as a square $\Xi \times \Xi$ stochastic matrix:

$$A_t \approx p(a_t \mid a_{t-1}) \tag{25}$$

Each column of the matrix corresponds to the a probability vector corresponding to the conditional probability given the value of a on the preceding time step:

$$A_t[:, j] \approx p(a_t \mid a_{t-1} = \xi_j) \tag{26}$$

Each element of the transition matrix is the conditional probability

$$A_t[i, j] \approx p(a_t = \xi_i \mid a_{t-1} = \xi_j) \tag{27}$$

The first column of the transition matrix represents the conditional transition probability given that the latent variable a was at the bound on the previous time step. By the definition of the absorbing bounds, for all time step t, the first column of the transition matrix has a value at the first row element and zero elsewhere.

$$A[i,1] = \begin{cases} 1, & i = 1 \\ 0, & \text{otherwise} \end{cases} \tag{28}$$

Similarly, the last column of has a value of 1 at the last element and zero elsewhere:

$$A[i,\Xi] = \begin{cases} 1, & i = \Xi \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

Each of the intervening columns is computed similarly to the prior probability vector, but with a different gaussian mean μ_j and variance σ^2. For $j \in \{2, ...\Xi - 1\}$. Both μ_j and σ^2 vary across time steps, but the subscript t is omitted for clarity.

$$A_t[i,j] = \begin{cases} \Phi(z_1) + \frac{\sigma_t}{\Delta\xi}\Big(z_2\left[\Phi(z_2) - \Phi(z_1)\right] + \phi(z_2) - \phi(z_1)\Big), & i = 1 \\ \frac{\sigma_t}{\Delta\xi}\Big(z_{i+1}\left[\Phi(z_{i+1}) - \Phi(z_i)\right] + z_{i-1}\left[\Phi(z_{i-1}) - \Phi(z_i)\right] + \phi(z_{i+1}) + \phi(z_{i-1}) - 2\phi(z_i)\Big), & \Xi > i > 1 \\ 1 - \Phi(z_\Xi) + \frac{\sigma_t}{\Delta\xi}\Big(z_{\Xi-1}\left[\Phi(\xi_{\Xi-1}) - \Phi(\xi_\Xi)\right] + \phi(z_{\Xi-1}) - \phi(z_\Xi)\Big), & i = \Xi \end{cases} \tag{30}$$

where

$$z_i = \frac{\xi_i - \mu_j}{\sigma}$$

The mean of the gaussian for each gaussian is given by

$$\mu_j = e^{\lambda\Delta t}\left(\xi_j + \frac{m}{\lambda\Delta t}\right) - \frac{m}{\lambda\Delta t} \tag{31}$$

where the momentary input m is the sum of the difference between the sum of the right and left click inputs:

$$m = \sum_{\tau\in[t-\Delta t,t)} c_R(\tau) - c_L(\tau) \tag{32}$$

The variance σ^2 does not vary across the columns of the transition matrix and a sum of the stimulus-independent diffusion noise and the noise emanating from each click

$$\sigma^2 = \sigma_a^2\Delta t + \sigma_s^2 \sum_{\tau\in[t-1,t-\Delta t)} c_R(\tau) + c_L(\tau) \tag{33}$$

Because the momentary click input varies across time step, and because both μ_j and σ^2 depend on the momentary input,

2.1.4 Sensory adaptation

Sensory adaptation is conceptualized as nearly full depression immediately after a click and gradually less and less depression after the preceding click. This idea is modelled with the ordinary differential equation parametrizing the gain C of the momentary input

$$\frac{dC}{dt} = C \cdot (\phi - 1) \cdot \delta(\boldsymbol{\tau}_{L,R} - t) + k(1 - C) \tag{34}$$

with initial condition

$$C(t = 0) = 1 \tag{35}$$

At the occurrence of the first click on each trial, which is always the stereoclick, the gain of the momentary input is given by

$$C(dt) = \phi$$

In accordance with previous findings [3], the extensive reports of the behavioral and neural phenomena assocated with the precedence effect[2], the hyperparameter ϕ is fixed to be 0.001. After the occurrence of a click, in the gap of time before the subsequent click, the gain recovers toward 1 with the change rate k. We specify the change to be fast by setting $k = 100$.

After each click at time τ, the dynamics of C at time $\tau+\Delta t$ before any other click's arrival is computed by separately integrating over the intervals $[\tau, \tau+\epsilon]$ and $(\tau+\epsilon, \tau+\Delta t]$ for small $\epsilon \ll \Delta t$:

$$\begin{aligned}
&\int_{\tau}^{\tau+\epsilon} dC = k\int_{\tau}^{\tau+\epsilon}[1-C(t)]dt + \int_{\tau}^{\tau+\epsilon}(\phi-1)\delta(\tau-t)C(t)dt \\
&C(\tau+\epsilon) - C(\tau) \approx k[1-C(\tau+\epsilon)-1+C(\tau)]\epsilon + (\phi-1)C(\tau) \\
&C(\tau+\epsilon) \approx C(\tau) + k\epsilon[C(\tau)-C(\tau+\epsilon)] + (\phi-1)C(0) \\
&C(\tau+\epsilon) \approx C(\tau)\left[\phi + k\epsilon\left(1-\frac{C(\tau+\epsilon)}{C(\tau)}\right)\right]
\end{aligned}$$

Integrating over the second interval,

$$\begin{aligned}
&\int_{\tau+\epsilon}^{\tau+\Delta t}\frac{dC}{1-C(t)} = \int_{\tau+\epsilon}^{\tau+\Delta t} kdt + \int_{\tau+\epsilon}^{\tau+\Delta t}(\phi-1)\delta(\tau-t)\frac{C(t)}{1-C(t)}dt \\
&-\log(|1-C(\tau+\Delta t)|) + \log(|1-C(\tau+\epsilon)|) = k(\Delta t-\epsilon)
\end{aligned}$$

Using the result obtained by integrating over the first interval,

$$\begin{aligned}
&|1-C(\tau+\Delta t)| = \exp\left(\log(|1-C(\tau+\epsilon)|) - k(\Delta t-\epsilon)\right) \\
&|1-C(\tau+\Delta t)| = |1-C(\tau+\epsilon)|\exp\left(-k(\Delta t-\epsilon)\right) \\
&|1-C(\tau+\Delta t)| \approx \left|1-C(\tau)\left[\phi + k\epsilon\left(1-\frac{C(\tau+\epsilon)}{C(\tau)}\right)\right]\right|\exp\left(-k(\Delta t-\epsilon)\right)
\end{aligned}$$

In the limit of $\epsilon \to 0$,

$$|1-C(\tau+\Delta t)| = |1-\phi C(\tau+\Delta t)|\exp\left(-k\Delta t\right)$$

Therefore,

$$C(\tau+\Delta t) = \begin{cases} 1-[1-\phi C(\tau)]\exp\left(-k\Delta t\right) \\ 1+[1-\phi C(\tau)]\exp\left(-k\Delta t\right) \end{cases}$$

These two possibilities are disambiguated by the requirements

$$\begin{aligned}
&k \geq 0 \\
&\phi \in [0,1] \\
&C \in [0,1] \\
&C(0) = 1
\end{aligned}$$

Therefore

$$C(\tau+\Delta t) = 1-[1-\phi C(\tau)]\exp\left(-k\Delta t\right) \tag{36}$$

Note that if we take the derivative, we recover

$$\frac{dC}{dt} = k(1-C)$$

For click times $\boldsymbol{\tau}$, the dynamics of C at the moment of each click is computed recursively:

$$\begin{aligned}
&C(\tau_1) \equiv 1 \\
&C(\tau_2) = 1-[1-\phi]\exp\left[-k(\tau_2-\tau_1)\right] \\
&C(\tau_3) = 1-[1-\phi C(\tau_2)]\exp\left[-k(\tau_3-\tau_2)\right] \\
&\quad\ldots
\end{aligned}$$

2.1.5 Latency

The clicks have a latency of Δt=0.01 s, set based on the latency of auditory responses in rat primary auditory cortex[13]. On time step t, the accumulator variable receives input from clicks that occurred during $[t - 2\Delta t, t - \Delta t)$. Results are not affected by the exact value of the latency.

2.2 Spike train response

Conditioned on the latent accumulator variable a, each neuron's spike train response y is modelled as Poisson random variable:

$$p(y \mid a) = (\lambda \Delta t)^y \exp(-\lambda \Delta t)/y! \tag{37}$$

The spike train is binned at the time steps of $\Delta t = 0.01$ s, the same as that of the latent variable. The firing rate λ is the nonlinearly rectified output of the softplus function (f), which approximates the rectification observed in the frequency-current curve of a neuron:

$$f(x) = \log\{1 + \exp(x)\} \tag{38}$$

When the input x is very negative, the output of the softplus nonlinearity is approximately zero, and then when x is large and positive, the output is approximately x itself. The same rectification function is used for neurons.

2.2.1 Accumulator-dependent predictor

The input x to the softplus function varies on each time step t on each trial and is a linear combination of the accumulated evidence $a(t)$ and accumulator-independent predictors $\boldsymbol{u}(t)$

$$x(t) = w_a \cdot \tilde{a}(t) + u(t) \tag{39}$$

To maximize interpretability and minimize trade-offs between parameters, the value of the accumulator being encoded is independent of the bound height:

$$\tilde{a}_i \equiv \frac{a_i}{B} = \frac{2i - \Xi - 1}{\Xi - 1} \tag{40}$$

The weight of the accumulator depends on the state i of the accumulator itself:

$$w_a = \begin{cases} w_{\mathrm{CM}} & \text{if } a \text{ is at a bound, i.e., } i = 1 \text{ or } i = \Xi \\ w_{\mathrm{EA}} & \text{otherwise} \end{cases} \tag{41}$$

Each neuron has its own scalar w_{CM} and w_{EA}.

2.2.2 Accumulator-independent, within-trial varying predictors

The accumulator-independent predictors consist of four types of variables, three of which vary both within and across trials, and a fourth that varies only across trials.

Each within-trial varying predictor is aligned to an event: 1) the stereoclick, which defines the start of the trial; 2) the animal's departure from the center port, which ends a trial if it occurs before 1s from the stereoclick; and 3) a neuron's own previous spiking.

On each trial m, the timing of the i-th event is represented by a Delta function, and the predictor at time step t is the linear convolution of the Delta function with a linear filter, or kernel, that is fixed across trials.

$$u(t) = \sum_i k_i * \delta(i, m, t) \tag{42}$$

To reduce model complexity and promote smoothness, for each type of input, the kernel k_i is parametrized as the linear combination of a set of smooth temporal basis functions:

$$k_i = \Phi_i \boldsymbol{w}_i \tag{43}$$

where each Φ_i is a $T \times D$ temporal basis matrix representing D orthogonal temporal basis functions, and $\boldsymbol{w}_i$ is a D-dimensional vector of weights that are learned from the data. See section 4.

Post-stereoclick kernel The beginning of each trial and the onset of the auditory click trains is indicated by an auditory click played simultaneously by the left and right speaker, i.e., a stereoclick. The kernel k extends from 0-1.0 s after the stereoclick and is parameterized using $D = 5$ basis functions without time warping .

Pre-movement kernel This kernel extends from 0.6 s to 0.01 s before the the animal departs from the center port and is parameterized by $D = 2$ temporal basis functions without time warping.

Post-spike kernel This kernel extends from 0.01-0.25 s after each spike emitted by the same neuron (but not by other neurons) and is parametrized by $D = 3$ temporal basis functions with a time warping parameter $\eta = 1$.

2.2.3 Accumulator-independent, across-trial varying predictor

The only accumulator-independent predictor is that is fixed within each trial is a trial-varying baseline $b(m)$, parameterized as the linear combination of eight smooth temporal basis functions evaluated over the start times of each trial m. To reduce the number parameters fit in the MMDDM, the relative weights of up to $B = 8$ smooth temporal basis functions are first learned using a separate linear-gaussian model fit to the firing rates $r(m)$:

$$r(m) \equiv \frac{1}{T_m \Delta t} \sum_{t}^{T_m} y(m,t) \tag{44}$$

where T_m is the number of time steps on the m-th trial, and $\Delta t = 0.01$ s is the duration of each time step in seconds. The linear-gaussian basis function regression model is defined as

$$r \sim \mathcal{N}(\mu, \sigma) \tag{45}$$

$$\mu \sum_{i}^{B} \phi_i(\tau_m) \tag{46}$$

Note that $b(m)$ is fixed for all time steps of a trial.

2.3 Behavioral choice

The behavioral choice is modelled as Bernoulli random variable parametrized by the accumulated evidence a at the last time step of each trial

$$p(\text{choice} = \text{right}) \equiv \sum_{a>0} p(a) \tag{47}$$

2.4 Inference

By discretizing the states of the accumulator, the MMDDM has the form of an autoregressive, input-output hidden Markov model (HMM)[1]. The ”autoregressive” feature refers to the inputs from each neuron’s own spike history to capture the long-range correlations between spike train observations of the same neuron, and the ”input-output” feature points to the dependence of the spike train observations on multiple baselines that vary on different time scales.

The inference of the most likely parameters of a model of this structure given simultaneously observed spike trains can be efficiently accomplished using the expectation-maximization (EM) algorithm [5]. However, the EM algorithm exhibits slow, first-order convergence when the conditional distributions of the observations, given the latent state, are not well-separated across states [11]. For the MMDDM, a poor separation between the conditional spike train distributions is expected from the approximation of the accumulator as many discrete states. Therefore, instead of the EM algorithm, we use the closely related expectation-conjugate-gradient (ECG) algorithm, which shows superior performance for poorly separated conditional distributions[11]. The ECG algorithm computes the exact gradient of the log-likelihood of the observations by solving the expectation step of the standard forward-backward algorithm for HMM’s. The gradient can be used by any first-order optimization algorithm to maximize the log-likelihood. The first-order optimization algorithm used here is the limited-memory BFGS [7], which is a quasi-Newton method that efficiently approximates the hessian matrix using the gradients.

2.4.1 Gradient of the log-likelihood

On each trial, the observations on which the likelihood of the MMDDM parameters are computed consist the behavioral choice:

$$d = \begin{cases} 0, & \text{left choice} \\ 1, & \text{right choice} \end{cases} \tag{48}$$

and the spike trains N simultaneously recorded neurons:

$$\boldsymbol{Y} = \{y_1, y_2, ...y_N\} \tag{49}$$

$$y = [y_1, y_2, ...y_T]^\top \tag{50}$$

and the spike train of each neuron on each time step $y_{n,t}$ is the spike count of that neuron over $t, t + \Delta t$). The gradient of the log-likelihood of the simultaneously recorded spike trains $\boldsymbol{Y}$ and the choice d on each trial is given by

$$\begin{aligned} \nabla \log p(\boldsymbol{Y}, d) &= \frac{\nabla p(\boldsymbol{Y}, d)}{p(\boldsymbol{Y}, d)} \\ &= \frac{\nabla \sum_{\boldsymbol{a}} p(\boldsymbol{Y}, d, \boldsymbol{a})}{p(\boldsymbol{Y}, d)} \\ &= \frac{\sum_{\boldsymbol{a}} p(\boldsymbol{Y}, d, \boldsymbol{a}) \nabla \log p(\boldsymbol{Y}, d, \boldsymbol{a})}{p(\boldsymbol{Y}, d)} \\ &= \sum_{\boldsymbol{a}} p(\boldsymbol{a} \mid \boldsymbol{Y}, d) \nabla \log p(\boldsymbol{Y}, d, \boldsymbol{a}) \\ &= \sum_{\boldsymbol{a}} p(\boldsymbol{a} \mid \boldsymbol{Y}, d) \left[\nabla \log p(\boldsymbol{Y}, d \mid \boldsymbol{a}) + \nabla \log p(\boldsymbol{a})\right] \end{aligned} \tag{51}$$

where $\boldsymbol{a}$ is the set of the values of the latent accumulator variable on each of the T time steps of the trial:

$$\boldsymbol{a} = \{a_1, a_2, ..., a_T\}$$

and conditioning jointly over the latent variable on all time steps. However, because of the first-order Markov dependence, the gradient can be simplified further

$$\begin{aligned} \nabla \log p(\boldsymbol{Y}, d) = &\sum_{a_1} p(a_1 \mid \boldsymbol{Y}, d) \nabla \log p(a_1) + \\ &\sum_{t=2}^{T} \sum_{a_t} p(a_t, a_{t-1} \mid \boldsymbol{Y}, d) \nabla \log p(a_t \mid a_{t-1}) + \\ &\sum_{t=2}^{T} \sum_{a_t} p(a_t \mid \boldsymbol{Y}, d) \sum_{n=1}^{N} \nabla \log p(y_{n,t} \mid a_{t-1}) + \\ &\sum_{a_T} p(a_T \mid \boldsymbol{Y}, d) \nabla \log p(d \mid a_T) \end{aligned} \tag{52}$$

The posterior probabilities $p(a_t \mid \boldsymbol{Y}, d)$ and $p(a_t, a_{t-1} \mid \boldsymbol{Y}, d)$ were computed using the forward-backward algorithm, which can be used even when the spike trains have spike history inputs [5].

3 Psychophysical kernel model

To validate the results from the MMDDM, we compared the psychophysical kernel inferred from the actual behavioral choices and the kernel inferred from simulated behavioral choices of the MMDDM. The psychophysical kernel quantifies the time-varying weight of the stimulus input on the behavioral choice.

The traditional approach to infer the psychophysical kernel is to use the reverse correlation technique [6], a method highly similar to the spike-triggered average [12]. However, the traditional reverse correlation technique critically depends on the stimulus fluctuations being independent across time. The traditional approach can be biased by inputs

that are correlated in time and also biased by stimulus-independent factors that influence the decision but cannot be incorporated in the analysis [9]. Moreover, the traditional technique typically assume a single temporal resolution for the weights of the stimulus fluctuations. When the assumed temporal resolution exceeds what can inferred from the data, the inferred kernel can be highly noisy and has limited interpretability.

We develop a new model to infer the psychophysical kernel that mitigates the bias from sensory inputs with temporally-correlated stimulus fluctuations and the bias from omitted factors that impact the choice. The psychophysical kernel is parametrized using temporal basis functions, and therefore the temporal resolution of the kernel is quantified by the number of basis functions. The optimal number of basis functions for a given dataset can be identified using out-of-sample model comparison.

To mitigate the bias from temporally correlated stimulus fluctuations, we incorporate a method based on the logistic regression model presented in [8] that reparametrizes the inputs to remove the temporal correlations in the stimulus input. To mitigate the potential bias from omitted factors, a behavioral lapse parameter is incorporated. To identify the optimal temporal resolution of the psychophysical kernel that can be inferred using a given data set, the kernel is parametrized as a linear combination of smooth temporal basis functions. Cross-validated model comparison identifies the optimal temporal resolution.

3.1 Model definition

The behavioral choice d is modelled as a Bernoulli random variable. The probability of a right choice is given by

$$p(d = 1) = f(\beta_1)f(\beta_2) + [1 - f(\beta_1)]f(y) \tag{53}$$

where the function f is the logistic function

$$f(x) = \frac{1}{1 + \exp(-x)} \tag{54}$$

which monotonically maps the input $x \in (-\infty, \infty)$ to the range $(0, 1)$. The value $f(\beta_1)$ quantifies the behavioral lapse rate, or the fraction of trials in which the animal is a behavioral state that ignores the stimulus and depends on only its bias β_2. In the remaining $1 - f(\beta_1)$ of the trials, the choice depends on not only the bias β_2 but also the sensory stimulus. The linear combination of these inputs is given by

$$y = \beta_2 + \beta_3 \lambda T \Delta t + \boldsymbol{e}^\top \Phi \boldsymbol{w} \tag{55}$$

The parameter β_3 is the weight of the expected stimulus input at each time step ($\Delta t = 0.01$), which is represented as $\lambda T \Delta t$. The independent variable λ is the expected stimulus input per second given the random processes used to generate the stimuli on each trial. The hyperparameter $T = 75$ indicates the number of time steps on each trial examined in this analysis. When λ does not equal 0, the stimulus fluctuations are temporally correlated across time steps within a trial. Therefore, separate weights are learned for the expected stimulus input and the fluctuations from the expected input.

The expected input rate λ is the difference between the values of right and left input rate expected given the random process used to generate the trial, and the typical expected input rates are $\lambda \in \{-38, -34, -25, -10, 10, 25, 34, 38\}$

Fluctuations from the expected input are quantified as the excess stimulus input $\boldsymbol{e}$. The T dimensional vector $\boldsymbol{e}$ is the difference between the actual and the expected input at each time step τ

$$e(\tau) = -\lambda \Delta t + \sum_{\tau \in [t-\Delta t, t)} c_R(\tau) - \sum_{\tau \in [t-\Delta t, t)} c_L(\tau) \tag{56}$$

The weights of the excess stimulus input is given by the vector $\Phi \boldsymbol{w}$. The matrix Φ is $T \times D$. Each row of the matrix Φ corresponds to one of T time step in the trial, and each column consists of the output of one D radial basis functions, evaluated without time warping (see section 4). The weight vector $\boldsymbol{w}$ has the dimensionality D corresponding to the number of the temporal basis functions, and the values of the vector are learned from the data. When $D = 1$, Φ is constrained to be a vector whose elements are the identical, and the weights $\Phi \boldsymbol{w}$ is constrained to be a flat line.

3.2 Inference

The free parameters of the model are $\{\beta_1, \beta_2, \beta_3, \boldsymbol{w} \in \mathbb{R}^D\}$. The hyperparameter D, which quantifies the number of the temporal basis functions and therefore represents the time resolution of the psychophysical kernel, is optimized by 10-fold cross-validated model comparison. To improve consistency between cross-validation folds, the L2 regularization

weight of $\alpha = 0.01$ (i.e., a spherical gaussian prior) was place on the parameters, but we confirmed that the conclusions are the same even if no regularization was used. The eigenspectrum of the hessian matrix of the log-posterior was examined to confirm that there is no trade off between the parameters during estimation.

3.3 Data used to fit the model

To estimate the psychophysical kernel aligned to trial onset (i.e., the start of the auditory click trains), trials with a stimulus duration of at least 0.75s were included from the set of recording sessions used in other analyses.

4 Basis functions

In either the multi-mode drift-diffusion model (MMDDM) and the psychophysical kernel model (PKM), nonlinear basis functions of the independent variables are used to reduce the number of parameters needed to be learned. For a given independent variable whose values vary across T time steps on each trial, we did not learn the weight at each time step. Instead, we learn the weight of each of D basis functions of the independent variable. Because $D << T$, the number of parameters can be greatly reduced.

For a given independent variable, such as the time-varying fluctuations in auditory click input in the psychophysical kernel model, its value on a given trial with T time steps can be expressed as the T-dimensional input vector $\boldsymbol{x}$. The basis function output $\boldsymbol{y}$ of the independent variable is given by

$$\boldsymbol{y} = \Phi^\top \boldsymbol{x} \tag{57}$$

The matrix Φ has the dimensions $T \times D$: each column corresponds to an individual basis function, and each row to a time step. The basis function output y is therefore a vector of length D.

The linear combination of $\boldsymbol{y}$ is an input in a pre-rectification stage of either the MMDDM or the PKM. For example, in the MMDDM, the Poisson rate of a neuron $\boldsymbol{\lambda}$ is given by

$$\boldsymbol{\lambda} = f\left(\boldsymbol{w}_1^\top \boldsymbol{y}_1 + \boldsymbol{w}_2^\top \boldsymbol{y}_2 + ...\right) \tag{58}$$

where f is the softplus function. The linear projection weights $\boldsymbol{w}_i$ are learned for each of the i-th independent variable.

The basis function matrix Φ is pre-specified and fixed for each independent variable and for all trials. Each column of Φ consists of a radial basis function evaluated on the index τ of each time step. The radial basis function we used has the form of a raised cosine [10]. The j-th basis function is given by

$$\phi_j(\tau) = \begin{cases} 1/2 - \cos\left\{\theta_j(\tau)\right\}/2, & \theta \in [0, 2\pi] \\ 0, & \text{otherwise} \end{cases} \tag{59}$$

where the transformation from time step index to radian is given by

$$\theta_j(\tau) = \frac{[f(\tau) - f(1) - (j-3)\Delta]\pi}{2\Delta} \tag{60}$$

The function f is a monotonic time warping function described below. The value Δ is the distance between the peaks of two adjacent basis functions:

$$\Delta = \frac{f(T) - f(1)}{D - 1} \tag{61}$$

After evaluating Φ, it is transformed into a unitary matrix using singular vector decomposition.

4.1 Time warping

To parametrize the post-spike kernel in the MMDDM or the peri-commitment psychophysical kernel, we wish to parametrize the portion of the kernel closer to the event of interest with a higher density of basis functions than the portion of the kernel farther away from the event time. To implement a continuous change in the density of representation, the time step indices are monotonically transformed using the function f

$$f(\tau; \eta) = \sinh^{-1}\{\eta(\tau - \tau_0)\} \tag{62}$$

The time index when the event occurred is indicated by τ_0. The parameter η indicates the degree of time warping. The larger is the value of η, the greater the warping: i.e., the portion of the kernel closer to τ_0 is represented by more functions and can have a higher temporal resolution than the portion father away from τ_0. For $1 >> \eta > 0$, the warping is essentially absent, and $f(\tau; \eta) \approx \tau$. This can be seen using the first order Taylor approximation of $\text{asinh}^{-1}(\eta\tau)$ around zero

$$f(\eta\tau) \approx f(0) + f'(0) \cdot \eta\tau = \eta\tau$$

References

[1] C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer Science & Business Media, 2006.

[2] A. D. Brown, G. C. Stecker, and J. T. Daniel. The precedence effect in sound localization. *Journal of the Association for Research in Otolaryngology*, 16, 2015.

[3] B. W. Brunton, M. M. Botvinick, and C. D. Brody. Rats and humans can optimally accumulate evidence for decision-making. *Science*, 340(6128):95–98, 2013.

[4] B. D. DePasquale, J. W. Pillow, and C. D. Brody. Accumulated evidence inferred from neural activity accurately predicts behavioral choice. In *Computational and Systems Neuroscience (Cosyne) Abstracts*, 2019.

[5] S. Escola, A. Fontanini, D. Katz, and L. Paninski. Hidden markov models for the stimulus-response relationships of multistate neural systems. *Neural Computation*, 23, 2011.

[6] P. Neri, A. J. Parker, and C. A. Blakemore. Probing the human stereoscopic system with reverse correlation. *Nature*, 1999.

[7] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Science & Business Media, 2000.

[8] O. Odoemene, S. Pisupati, H. Nguyen, and A. K. Churchland. Visual evidence accumulation guides decision-making in unrestrained mice. *Journal of Neuroscience*, 2018.

[9] G. Okazawa, L. Sha, B. A. Purcell, and R. Kiani. Psychophysical reverse correlation reflects both sensory and decision-making processes. *Nature Communications*, 2018.

[10] I. M. Park, M. L. Mester, A. C. Huk, and J. W. Pillow. Encoding and decoding in parietal cortex during sensorimotor decision-making. *Nature Neuroscience*, 17, 2014.

[11] R. Salakhutdinov, S. Roweis, and Ghahramani Z. Optimization with em and expectation-conjugate-gradient. *Proceedings of the 20th International Conference on Machine Learning*, 2003.

[12] O. Schwartz, J. W. Pillow, N. C. Rust, and E. P. Simoncelli. Spike-triggered neural characterization. *Journal of Vision*, 2006.

[13] J. D. Yao, P. Bremen, and J. C Middlebrooks. Rat primary auditory cortex is tuned exclusively to the contralateral hemifield. *Journal of Neurophysiolology*, 110, 2013.