EVIDENCE ACCUMULATION IN SYNAPTIC EFFICACY: ACTIVITY-SILENT DECISION-MAKING WITH SHORT-TERM PLASTICITY

Anonymous authors
Paper under double-blind review

ABSTRACT

Decision-making requires accumulation of evidence over time, yet exactly how neural circuits implement this computation remains debated. Classical models posit that evidence accumulation is realized by ramping neural activity, whereas recent experimental findings suggest an alternative mechanism in which neural activity remains largely silent before abruptly transitioning into an active decision state. This mechanism leaves a fundamental question unanswered: how is evidence accumulated during the neural activity-silent period? To address this question, we trained excitatory–inhibitory recurrent neural networks (RNNs) with short-term synaptic plasticity (STP) on a reaction-time decision task. We found that after training, the networks learned to leverage STP to make decisions in an activity-silent manner. Choice-tuning analysis revealed a clear functional organization of the network, and connection-pruning analysis uncovered the coupling between choice-selective excitatory and inhibitory neural groups. Guided by the architecture of trained RNNs, we constructed a minimal neural circuit model, in which evidence accumulates in facilitated synaptic efficacy that progressively reshapes the circuit's phase portrait, eventually driving the network state across the stability boundary and triggering a rapid transition to a choice state. We further compared activity before a choice between the theoretical STP and Static models to examine the potential energy benefit of activity-silent decision-making. Together, our results suggest STP as a mechanism for evidence accumulation in activity-silent decision-making and offer insights into the design of energy-efficient artificial neural networks.

1 INTRODUCTION

Decision-making often requires accumulating sensory evidence before making a choice. Early experimental and theoretical studies suggested that neurons representing competing choices accumulate evidence through a gradual increase in firing rate, known as ramping (Roitman & Shadlen, 2002; Wang, 2002; Wong & Wang, 2006). However, more recent studies have identified a different activity pattern, known as stepping, in which firing changes abruptly during a decision (Latimer et al., 2015; Zoltowski et al., 2019). The activity of neurons representing the eventual choice remains low after stimulus onset and then rises rapidly before the choice is reported. This pattern resembles activity-silent working memory, in which information is retained with little sustained neuronal activity (Rose et al., 2016; Wolff et al., 2017). We refer to the low-activity interval followed by a rapid choice-related rise as *activity-silent decision dynamics*. Ramping-based accounts link evidence accumulation to rising neuronal activity, so activity-silent decision dynamics raise two questions: how does the circuit accumulate evidence while choice-related activity remains low, and what dynamics keep neuronal activity low after stimulus onset and then cause it to rise rapidly?

Training artificial neural networks on behavioral tasks has become an established approach to studying neural computation (Mante et al., 2013; Song et al., 2016): optimizing a task objective can produce an activity pattern without prescribing the internal activity trajectories (Cueva & Wei, 2018; Banino et al., 2018), and analysis of the trained network then identifies the neurons and connections that generate the pattern and support behavior (Roach et al., 2023; Chu et al., 2026). We use this

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

approach to investigate whether task optimization produces activity-silent decision dynamics and how the resulting circuit performs evidence accumulation.

We hypothesize that short-term synaptic plasticity (STP) supports evidence accumulation during activity-silent decision dynamics. STP is a widespread property of biological synapses in which recent presynaptic activity transiently increases or decreases synaptic efficacy, often through changes in neurotransmitter release (Zucker & Regehr, 2002; Regehr, 2012). Previous models have shown that STP can support working memory and stimulus discrimination (Mongillo et al., 2008; Masse et al., 2019; Bourjaily & Miller, 2012; Ballintyn et al., 2019). Here, we test whether a recurrent neural network (RNN) trained for decision-making learns to recruit STP while choice-related neuronal activity remains low. We introduce STP into an excitatory–inhibitory (E–I) RNN with separate excitatory (E) and inhibitory (I) neurons and train it on a motion-discrimination task (Song et al., 2016). Each presynaptic neuron has a learnable *STP strength*, a_j, that controls the contribution of STP to recurrent transmission; STP strength is optimized during training and remains fixed within each trial, whereas synaptic efficacy changes within the trial. We compare RNNs with STP against Static networks trained without STP.

We found that activity-silent decision dynamics emerge consistently across trained RNNs with STP, are more prevalent than in Static networks, and are accompanied by increased STP strength across initialization conditions. The transition time, defined as the onset of the rapid rise in neuronal activity, varies across trials, so averaging neuronal activity across trials produces ramping, consistent with experimental observations (Latimer et al., 2015). Analysis of neuronal activity and connection pruning identifies a clear choice-selective E–I circuit. Excitatory neurons reinforce the same choice and recruit inhibitory neurons with the same choice preference; those inhibitory neurons suppress excitatory neurons with the opposite preference. Neurons with higher STP strength are predominantly choice-selective, and silencing them impairs performance more than silencing an equal number of low-STP neurons with matched E/I composition. These results locate strong STP recruitment within the circuit that supports the decision.

Guided by the trained RNNs, we construct a minimal model of two choice-selective E–I groups with self-excitation within each group and mutual inhibition between groups. As synaptic efficacy increases, the corresponding Static network passes from a tristable regime, in which a stable low-activity state coexists with two stable choice states, to a bistable regime in which the low-activity state becomes a saddle point (Wong & Wang, 2006). In the STP model, synaptic efficacy evolves over time: its initially low value keeps neuronal activity low after stimulus onset while evidence accumulates in synaptic efficacy, and once efficacy crosses a critical value, the low-activity state loses stability and activity in the winning group rises rapidly, explaining both how evidence accumulates during low neuronal activity and why it ends in a rapid transition. Comparison with a Static ramping model illustrates the potential energy benefit of keeping activity low before a choice.

Our contributions are threefold. We establish a task-trained neural network model that exhibits activity-silent decision dynamics and identify its choice-selective E–I circuit. We propose a theoretical model in which evidence accumulates in synaptic efficacy and STP drives a transition from tristability to bistability by destabilizing the low-activity state. We connect this account to an activity-cost comparison relevant to the design of energy-efficient AI decision systems.

2 Task design and model architecture

We trained excitatory–inhibitory RNNs on the reaction-time motion-discrimination task of Song et al. (2016) (Figure 1b). Two inputs provide noisy evidence for leftward and rightward motion, and a third input signals stimulus presentation. The network reports its choice through one of two outputs. Higher motion coherence increases the difference between the mean evidence inputs, making the direction easier to identify. Networks are trained to minimize the mean squared error between the two outputs and piecewise-constant choice targets during the supervised response interval (Equation 9). Appendix A.1 details the task and behavioral measures.

Each RNN contains 80 excitatory neurons and 20 inhibitory neurons, represented by nonnegative rate units. Excitatory neurons make nonnegative recurrent connections, and inhibitory neurons make nonpositive recurrent connections. Two linear choice outputs receive input from the excitatory neu-

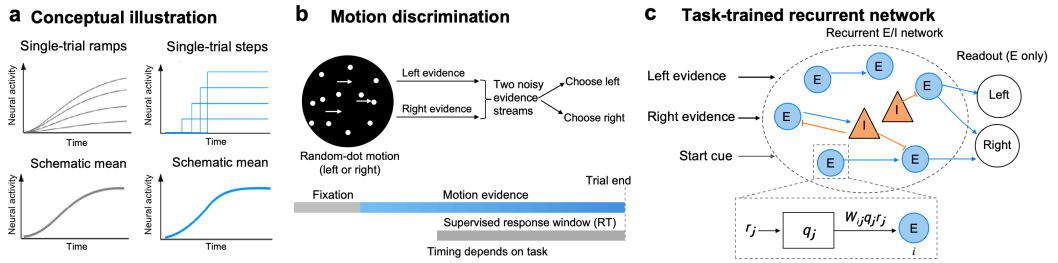

Figure 1: Task and network overview. **a,** Schematic ramping and activity-silent dynamics. Variation in transition times across trials can produce a gradual trial average. **b,** Reaction-time motion discrimination. The network receives noisy leftward and rightward motion evidence and reports a choice through one of two outputs. Motion coherence controls the difference between the mean evidence inputs. **c,** Excitatory–inhibitory RNN. STP modulates recurrent transmission, and the two choice outputs receive input from excitatory neurons. Connectivity is schematic. The STP strength a_j is learned during training. Synaptic utilization u_j and the available-resource fraction x_j evolve within each trial. Appendices A.1 and A.2 specify the task and network implementation.

rons (Figure 1c). Neuronal activity follows

$$\tau_r \frac{dh_i}{dt} = -h_i + \sum_j W_{ij} q_j^{\text{eff}}(t) r_j(t) + \sum_k W_{ik}^{\text{in}} v_k(t) + b_i + \xi_i(t), \qquad r_i(t) = [h_i(t)]_+, \tag{1}$$

where h_i is the state of neuron i, r_i is its activity in arbitrary units, and $[h_i]_+ = \max(0, h_i)$ is the rectified-linear activation function. The parameter τ_r is the neuronal time constant. The weight W_{ij} connects presynaptic neuron j to neuron i, W_{ik}^{in} connects external input v_k to neuron i, and b_i and ξ_i are the bias and recurrent noise. The factor q_j^{eff} scales all recurrent outputs from neuron j.

STP changes recurrent transmission through two variables for each presynaptic neuron: synaptic utilization u_j and the available-resource fraction x_j. They evolve according to

$$\frac{du_j}{dt} = \frac{U - u_j}{\tau_u} + s_r U(1 - u_j) r_j, \tag{2}$$

$$\frac{dx_j}{dt} = \frac{1 - x_j}{\tau_x} - s_r u_j x_j r_j. \tag{3}$$

Here, U is baseline synaptic utilization, τ_u and τ_x are the recovery time constants for u_j and x_j, and s_r converts neuronal activity to the rate used in the STP equations. The activity-dependent terms increase utilization and deplete available resources. The synaptic efficacy normalized by its resting value is $u_j x_j / U$.

We compare two conditions. Static networks use $q_j^{\text{eff}} = 1$. In RNNs with STP, the learnable STP strength a_j determines how strongly normalized synaptic efficacy changes recurrent transmission:

$$a_j = \sigma(\theta_j / T_g), \qquad q_j^{\text{eff}}(t) = 1 - a_j + a_j \left[\frac{u_j(t) x_j(t)}{U} \right]. \tag{4}$$

The logistic function σ maps the trainable parameter θ_j to an STP strength between zero and one. The parameter T_g sets the scale of this mapping. The limits $a_j = 0$ and $a_j = 1$ correspond to static transmission and full normalized synaptic efficacy, respectively. Training adjusts a_j together with the network weights, while u_j and x_j evolve within each trial. Both conditions use the same neuronal dynamics, task, and readout. Appendix A.2 gives parameter values, initialization, numerical integration, and equivalence checks. Appendix A.3 describes training and checkpoint selection, and Appendix A.4 summarizes the paired experimental configuration.

3 Activity-silent decision dynamics emerge in RNNs with STP

We found that activity-silent decision dynamics emerged more frequently in RNNs with STP. Training reduced task loss and improved task performance (Figure 2a). After training, among choice-tuned neurons, activity-silent dynamics were more prevalent in RNNs with STP ($52.2 \pm 16.2\%$) than

in Static RNNs ($36.2 \pm 7.0\%$; mean $\pm$ standard deviation (SD) across networks), shifting the population toward the high prevalence of step-like single-trial dynamics reported in choice-selective LIP neurons (78%; Latimer et al., 2015). Selected single-trial examples further illustrate this contrast, showing ramping versus activity-silent responses under identical sensory input and noise (Figure 2b; Appendix A.5).

We next asked whether task optimization alone leads networks to exploit STP. Comparing STP strength before and after training showed increases in subsets of neurons across low, intermediate, and high initialization conditions, with ten networks in each condition (Figure 2c; Appendix A.2): the networks learned to use stronger STP while learning the task.

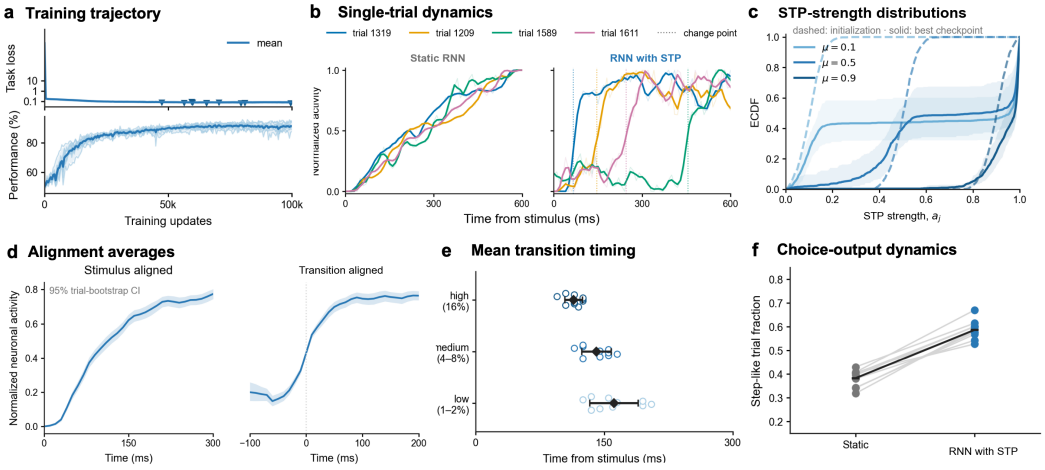

Figure 2: Learning and activity-silent decision dynamics. **a,** Validation loss and performance during training for the ten RNNs with STP. **b,** Matched single-trial examples showing ramping in the Static RNN and activity-silent decision dynamics in the RNN with STP under the same sensory input and noise. **c,** STP-strength distributions before (dashed) and after training (solid). Training increases STP strength in subsets of neurons under each initialization condition. **d,** The same trials produce a gradual ramp when aligned to stimulus onset and a rapid rise when aligned to their fitted transition times. **e,** Transitions occur later under weaker sensory evidence (within-network median transition times of representative units). **f,** Fraction of choice-output trajectories classified as step-like in Static networks and RNNs with STP. Panels a, e, and f summarize the ten primary networks; panel c summarizes ten networks per initialization condition; panels b and d show single-network examples. STP-strength initialization, display conventions, trial selection, and the ramp–step Bayesian information criterion (BIC) comparison are detailed in Appendices A.2, A.3, and A.5.

Because experiments report both single-trial stepping and stimulus-aligned ramping (Latimer et al., 2015), we asked whether transition-time variability across trials reproduces their coexistence in the trained networks. Transition times varied across trials, so stimulus-aligned averaging produced a gradual ramp whereas transition-aligned averaging recovered the abrupt rise (Figure 2d), and stronger evidence advanced the transition whereas weaker evidence prolonged the low-activity period (Figure 2e).

We then tested whether the distinction holds at the population level, using the choice output (the target readout minus the nontarget readout on correct-response trials). Step-like and ramp-like trajectories coexisted in both conditions, but step-like trajectories were more prevalent in RNNs with STP for every training seed (0.587 ± 0.040 versus 0.383 ± 0.036 across networks, mean$\pm$SD; Figure 2f). Fitting procedures and numerical summaries are given in Appendices A.3 and A.5.

4 Choice-selective populations support self-excitation and mutual inhibition

We next analyze the activity and connectivity of the trained networks to understand the mechanism underlying their activity-silent decision-making, first identifying the neurons representing each

choice, then analyzing the recurrent connections among these populations, and finally examining the role of neurons with high STP strength. The connection analyses use *connection pruning*, which removes selected connections by setting their weights to zero and measures the resulting change in task accuracy.

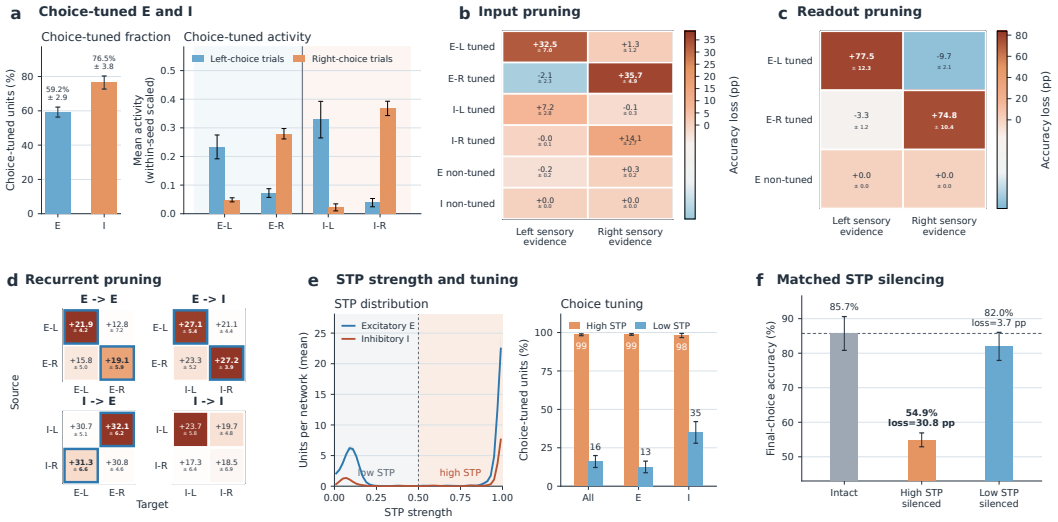

Figure 3: Choice tuning, connection pruning, and neuronal silencing in RNNs with STP. **a,** Fractions of choice-selective excitatory and inhibitory neurons and their population activity. **b,** Accuracy loss after removing a sensory-input channel's connections to each population. **c,** Accuracy loss for each evidence direction after disconnecting an excitatory population from both outputs. **d,** Accuracy loss after removing all recurrent connections from one choice-selective population to another; outlines mark selected pathways with large mean effects. **e,** Distributions of learned STP strength and choice-selective fractions among neurons with high or low STP strength. **f,** Accuracy after silencing neurons with high or low STP strength, with neuron count and E/I composition matched between conditions. Error bars and matrix labels report the mean ± standard error of the mean (SEM) across ten networks. Appendix Sections A.6–A.8 give the analysis procedures.

4.1 Choice tuning, sensory inputs, and readout of choice-selective populations

We first identified choice-selective neurons, whose activity differed significantly between correct left and right choices (Appendix A.6). We found that both the excitatory and inhibitory populations contained substantial fractions of neurons preferring each choice ($59.2 \pm 2.9\%$ of excitatory and $76.5 \pm 3.8\%$ of inhibitory neurons, mean±SEM; Figure 3a). We denote the left- and right-preferring excitatory populations by E_L and E_R, and the inhibitory populations by I_L and I_R.

We next asked whether these populations differed in their inputs and outputs, first testing whether each population depended on sensory evidence for its preferred choice. We found that left-preferring populations depended more strongly on the left-evidence input, while right-preferring populations depended more strongly on the right-evidence input (Figure 3b). Removing sensory connections to nonselective populations had much smaller effects. We then tested how the populations influenced the network's output. Disconnecting E_L from both outputs reduced accuracy on left-evidence trials and increased accuracy on right-evidence trials; disconnecting E_R produced the reverse pattern (Figure 3c). Disconnecting nonselective excitatory neurons had little effect. These results show that each choice-selective population preferentially receives evidence for its preferred choice and that each choice-selective excitatory population contributes to the output favoring that choice.

4.2 Recurrent connections supporting self-excitation and mutual inhibition

We then examined how recurrent connections allow these populations to interact (Appendix A.7). Our connection-pruning analysis identified a pattern of self-excitation and mutual inhibition (Fig-

ure 3d). Removing excitation within a choice-selective population impaired accuracy more than removing excitation between the two populations. Removing an excitatory population's projections to inhibitory neurons with the same choice preference caused a larger accuracy loss than removing its projections to inhibitory neurons with the opposite preference. Removing inhibitory projections to the opposite-choice excitatory population caused large accuracy losses as well. Together, these effects support an organization in which each excitatory population reinforces its own activity and recruits inhibitory neurons that suppress the competing excitatory population.

The excitatory-to-inhibitory and inhibitory-to-excitatory pathways resemble connections observed in mouse posterior parietal cortex: excitatory neurons preferentially connect to inhibitory neurons with the same choice preference, and inhibitory neurons preferentially connect to excitatory neurons with the opposite preference (Kuan et al., 2024).

4.3 Neurons with higher STP strength contribute more to decisions

Because every presynaptic neuron carries its own learnable STP strength, training could in principle assign any distribution of strengths. As Figure 2c showed, training increased STP strength in only a subset of neurons, splitting the learned strengths into a high group converging to values near one and a low group maintaining weak STP (Figure 3e). At the minimum-validation-loss checkpoints, high-STP neurons constituted $56.1 \pm 10.5\%$ of each network (mean$\pm$SD). Separating neurons into high-STP and low-STP groups (Appendix A.8), we found that almost all high-STP neurons were choice-selective, whereas choice selectivity was much less common among low-STP neurons (Figure 3e).

To test whether this association corresponded to a difference in behavioral contribution, we silenced high-STP or low-STP neurons while matching the number of silenced neurons and their E/I composition. Silencing high-STP neurons reduced accuracy to near chance, whereas silencing low-STP neurons had only a small effect (85.7% intact, 54.9% after high-STP silencing, and 82.0% after low-STP silencing; Figure 3f). Thus, neurons with higher learned STP strength contributed more to accurate decisions.

5 A theoretical model of activity-silent decision dynamics

5.1 Model

The trained RNNs suggest two ingredients for activity-silent decision dynamics: competition between two choice-selective E–I groups and activity-dependent changes in synaptic efficacy. We combine these ingredients in a minimal model with four neuronal populations, E_L, E_R, I_L, and I_R (Figure 4a). Each excitatory population excites itself and recruits the inhibitory population with the same choice preference, which suppresses the excitatory population representing the other choice and reproduces the self-excitation and mutual inhibition identified in the trained RNNs. This model formalizes the biological hypothesis, motivated but not directly demonstrated by the trained RNNs, that evidence accumulates in synaptic efficacy. Stimulus history is integrated in the synaptic state by construction, so the model demonstrates that the two ingredients suffice for activity-silent decisions.

For each choice $p \in \{L, R\}$, $r_{E,p}$ and $r_{I,p}$ denote the mean firing rates (in Hz) of the excitatory and inhibitory populations representing that choice, respectively. The index $\bar{p}$ denotes the opposite choice. These firing rates evolve according to

$$\tau_E \dot{r}_{E,p} = -r_{E,p} + \phi_E(J_{EE} q_p r_{E,p} - J_{IE} r_{I,\bar{p}} + B_E + e_p + \eta_{E,p}), \tag{5}$$

$$\tau_I \dot{r}_{I,p} = -r_{I,p} + \phi_I(J_{EI} q_p r_{E,p} + B_I + \eta_{I,p}). \tag{6}$$

Here, J_{EE}, J_{EI}, and J_{IE} are positive connection strengths; B_E and B_I are background inputs; e_p is sensory evidence; and $\eta_{E,p}$ and $\eta_{I,p}$ are input noise. The activation functions ϕ_E and ϕ_I convert input into firing rate. To simplify simulation and analysis, we apply STP only to excitatory outputs, although adding STP to inhibitory outputs does not qualitatively change the results (Appendix B). The factor $q_p = \alpha u_p x_p / U$ represents the combined effect of STP on excitatory synaptic transmission, where u_p is synaptic utilization and x_p is the fraction of available synaptic resources. Each excitatory population has its own u_p and x_p, and the resulting q_p scales both its self-excitation and

its output to the same-choice inhibitory population. The synaptic variables evolve according to

$$\dot{u}_p = \frac{U - u_p}{\tau_u} + U(1 - u_p) r_{E,p}, \tag{7}$$

$$\dot{x}_p = \frac{1 - x_p}{\tau_x} - u_p x_p r_{E,p}. \tag{8}$$

These equations describe activity-dependent facilitation and resource depletion (Masse et al., 2019). The parameters U, τ_u, and τ_x set baseline utilization and recovery times. The coefficient $\alpha = 0.35$ scales excitatory synaptic transmission. Appendix B specifies the activation functions, parameters, noise, and numerical implementation.

5.2 Activity-silent dynamics and their mechanism

Our simulations show that the model produces activity-silent decision dynamics on individual trials. After stimulus onset, both excitatory populations initially show low activity. One population then rises rapidly, recruits its same-choice inhibitory population, and suppresses the opposite excitatory population (Figure 4b). The transition time varies across trials (Figure 4c). Averaging after stimulus alignment therefore produces a gradual rise, whereas averaging the same trials after transition alignment preserves the rapid rise (Figure 4d).

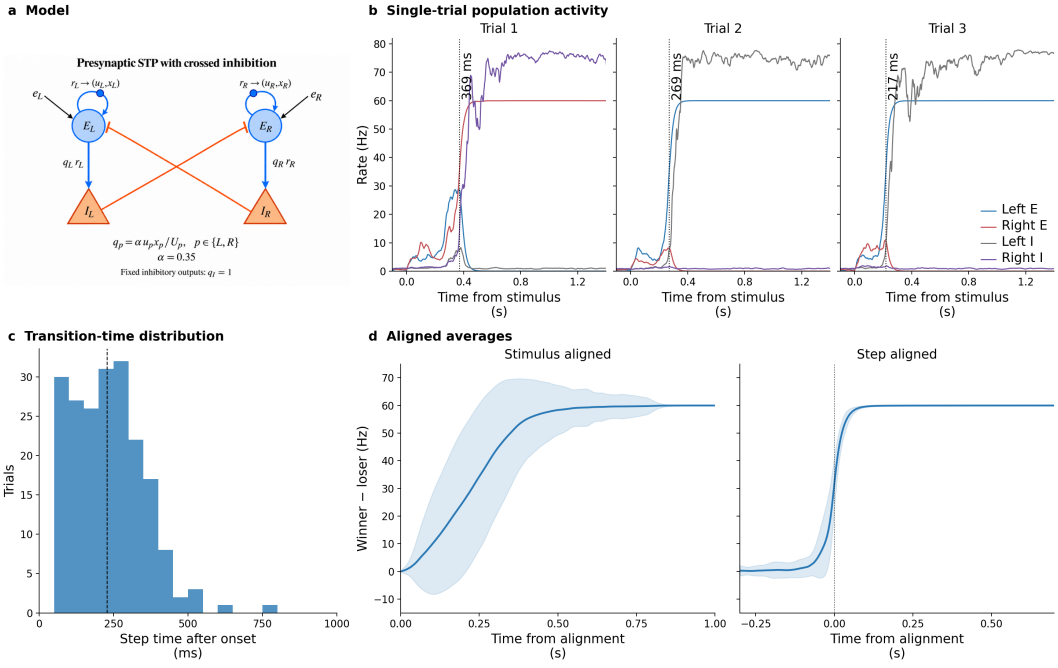

Figure 4: Activity-silent decision dynamics in the four-population model. **a,** Each excitatory population excites itself and its same-choice inhibitory population, which inhibits the opposite excitatory population. STP acts on excitatory outputs; inhibitory outputs are static. **b,** The first three example trials show the activity of all four populations. Dashed lines mark transition times obtained by fitting two activity levels to the winning excitatory population. **c,** Distribution of detected transition times; all 200 trials met the detection criteria. **d,** The same 200 trials averaged after alignment to stimulus onset or transition time; curves and shading show the mean and trial standard deviation (SD) of winner-minus-loser excitatory activity.

To explain the delayed transition, we examine the Static network with $q_L = q_R = q$ held fixed (Wong & Wang, 2006). Increasing synaptic efficacy strengthens excitatory self-excitation and recruitment of the inhibitory populations. We calculate the equilibrium states and their stability to determine how these changes affect competition between the two choice groups (Figure 5a,b).

Our stability analysis shows that, at intermediate synaptic efficacy, a stable low-activity state coexists with two stable choice states, forming a tristable regime. Further increases in synaptic efficacy

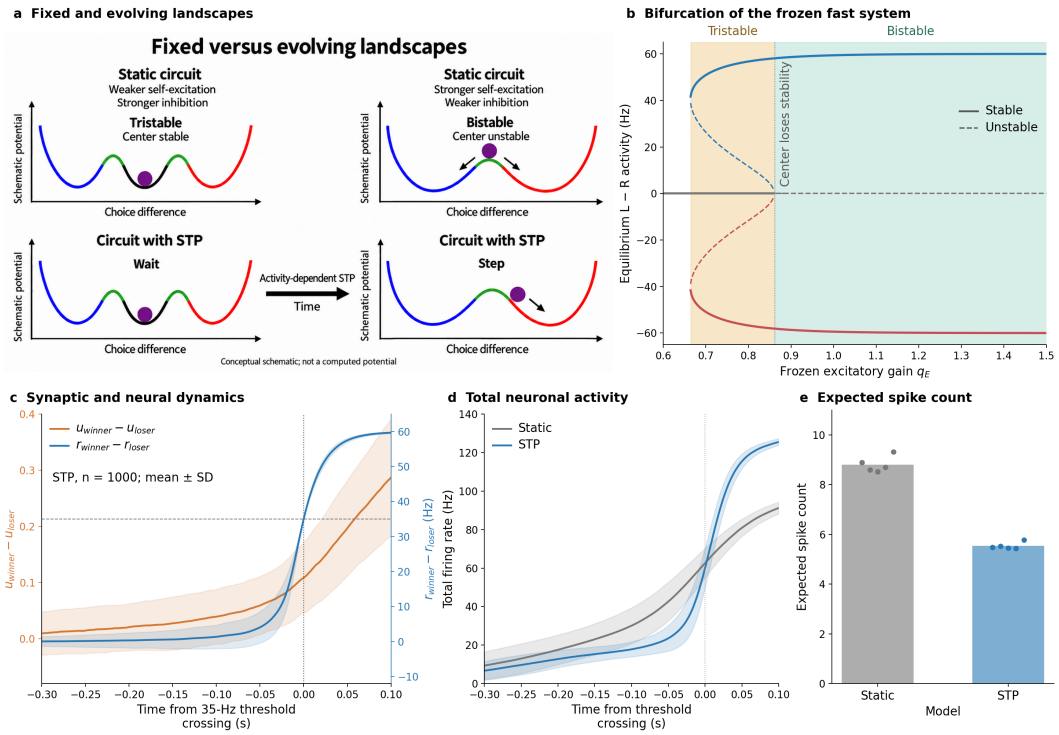

Figure 5: Synaptic mechanisms and expected spike count. **a,** Conceptual phase portraits illustrating the transition from tristability to bistability as the central low-activity state loses stability; the curves are schematic, not computed from the model. **b,** Equilibrium excitatory activity differences with $q_L = q_R = q$ held fixed over the displayed range $q = 0.6$–1.5. Solid and dashed lines denote stable and unstable equilibria, determined from the four-population Jacobian. The pale orange and pale green backgrounds mark the tristable and bistable regimes, respectively. The sampled boundaries are approximately $q = 0.664$ and $q = 0.862$; the unshaded interval to the left is monostable. **c,** Synaptic-utilization and excitatory-activity differences in the STP model, aligned to the first 35-Hz activity-difference threshold crossing (1000 trials; mean and trial SD). **d,** Sum of the firing rates of all four populations, aligned to the same decision threshold (1000 trials per condition; mean and trial SD). **e,** Expected spike count obtained by integrating the total firing rate from stimulus onset to threshold crossing, with one effective rate unit per population. Bars show trial means; points show five noise-batch means. Panels **d** and **e** use an adjusted STP variant whose parameters are set so that its psychometric curve, which measures accuracy across evidence levels, matches that of the Static model (Figure S4); all other parameters, including the Static time constant, are unchanged. Appendix B specifies the adjustment and protocols.

destabilize the low-activity state, which becomes a saddle point, while the two stable choice states remain. This transition from tristability to bistability explains why a small activity difference can remain suppressed and then grow rapidly. In the STP model, sensory input changes synaptic utilization and available resources through neuronal activity. The resulting changes in synaptic efficacy move the neuronal dynamics toward this loss of stability: simulated synaptic gain crosses the corresponding frozen-synapse stability boundary before the rapid neuronal transition (Figure S5), allowing the winning population to rise rapidly. Appendix B gives the equilibrium and stability calculations.

Our simulations show that synaptic utilization begins to differ between the winning and losing populations before the rapid neuronal transition. When trials are aligned to the decision threshold, the difference in synaptic utilization between the winning and losing populations develops while their activity difference remains small (Figure 5c). These dynamics support the proposed role of synaptic states in retaining input history during the low-activity period: a linear decoder on the utilization and resource differences recovers evidence direction above chance throughout this period (Appendix B).

We next examine how the delayed rise in neuronal activity affects the expected spike count. For this comparison, parameters of the STP model are adjusted so that its psychometric curve, which measures choice accuracy across evidence levels, matches that of the Static model (Figure S4); all other parameters, including the Static excitatory time constant, are unchanged. We measure total neuronal activity as the sum of the firing rates of all four populations. The STP model maintains lower total activity before the decision threshold than the Static model (Figure 5d), and integrating this total firing rate from stimulus onset to threshold crossing gives fewer expected spikes before a choice, present in every noise batch (Figure 5e), supporting a reduction in spike-related energy expenditure. Appendix B specifies the model adjustment, rate-unit convention, protocols, and numerical results.

6 Discussion and conclusion

In summary, we trained excitatory–inhibitory RNNs with learnable STP on a reaction-time decision task and found that they developed activity-silent decision dynamics: choice-related activity remained low after stimulus onset and rose abruptly at transition times that varied across trials, so that stimulus-aligned trial averages reproduced ramping. Learning increased STP strength in a subset of neurons that were almost exclusively choice-selective and contributed most to accurate decisions. Connection pruning revealed a choice-selective E–I circuit of self-excitation and mutual inhibition, and a four-population model built on this circuit explained activity-silent decision-making as a bifurcation: evidence accumulates in facilitated synaptic efficacy until the low-activity state loses stability and the network transitions rapidly to a choice state, with fewer expected spikes before a choice than in a calibrated Static ramping model.

Analysis of the trained RNNs identifies connections that can implement this competition, resembling the choice-dependent connections measured in posterior parietal cortex (Kuan et al., 2024) and relating to models in which inhibitory connectivity controls decision dynamics (Roach et al., 2023); this mechanism complements experimental evidence for changes in neuronal dynamics during decision-making (Luo et al., 2025).

Why did task optimization produce activity-silent decision dynamics even though nothing in the objective prescribed them? The objective contained no term rewarding low activity or energy saving, yet gradient descent on the task objective discovered a mixed regime in which neurons with strong STP carry much of the decision computation. A network equipped with STP naturally favors storing short-timescale information in synaptic efficacy rather than in sustained neural activity: synaptic efficacy evolves on recovery time constants that are slow relative to neuronal dynamics and therefore well suited to retaining input history over hundreds of milliseconds, whereas rapidly fluctuating neural activity is better suited to fast operations such as the abrupt commitment to a choice. This division of labor is consistent with earlier STP network models of working memory, in which trained networks likewise stored information primarily in synaptic states (Mongillo et al., 2008; Masse et al., 2019; Kozachkov et al., 2022).

Delaying strong neuronal activity may additionally reduce the metabolic cost of making a choice: spikes and postsynaptic currents require energy to restore ionic gradients (Attwell & Laughlin, 2001; Howarth et al., 2012), and this cost constrains sustained cortical activity (Lennie, 2003). The STP model's lower total firing rate and smaller expected spike count before a choice provide a proxy for lower spike-related energy expenditure at psychometrically matched accuracy, though a complete metabolic comparison would additionally require accounting for synaptic transmission and vesicle cycling (Rangaraju et al., 2014).

These observations also speak to artificial intelligence. Test-time-training architectures incorporate STP-like mechanisms by updating part of the network during inference (Ba et al., 2016; Sun et al., 2020), yet the computational advantages of storing moment-to-moment information in such fast synaptic states, rather than in recurrent neuron states, have received little systematic discussion. Our results provide a biological case study: when a network possesses STP, gradient-based optimization on a decision task alone shifts information storage from firing rates toward synaptic efficacy and supports accurate decisions with fewer expected spikes, suggesting STP-style mechanisms as a design principle for how adaptive networks allocate short-term memory between synapses and neuronal activity.

REPRODUCIBILITY STATEMENT

Appendix A specifies the task, RNN dynamics, numerical implementation, optimization, checkpoint selection, and analysis procedures used for the reported network results. Appendix B gives the parameters, simulation procedure, transition detection, stability calculations, and expected-spike-count analysis for the theoretical model. The corresponding sections identify the trial sets, training seeds, checkpoints, and aggregation units used for each figure where applicable.

AI USE STATEMENT

OpenAI Codex was used for research execution, including inspecting implementations, developing model and analysis code, configuring and monitoring experiments, interpreting recorded outputs, and preparing scientific figures. It was also used to retrieve and summarize related literature, edit and polish the writing and mathematical presentation. The methods and experiments resulting from these uses are described in the paper. The authors reviewed AI-assisted code and outputs, checked experimental claims against the underlying records, and verified cited sources before inclusion. The authors take responsibility for the design, execution, interpretation, and final content of the work.

REFERENCES

David Attwell and Simon B Laughlin. An energy budget for signaling in the grey matter of the brain. *Journal of Cerebral Blood Flow & Metabolism*, 21(10):1133–1145, 2001.

Jimmy Ba, Geoffrey E Hinton, Volodymyr Mnih, Joel Z Leibo, and Catalin Ionescu. Using fast weights to attend to the recent past. *Advances in neural information processing systems*, 29, 2016.

Benjamin Ballintyn, Benjamin Shlaer, and Paul Miller. Spatiotemporal discrimination in attractor networks with short-term synaptic plasticity. *Journal of computational neuroscience*, 46(3):279–297, 2019.

Andrea Banino, Caswell Barry, Benigno Uria, Charles Blundell, Timothy Lillicrap, Piotr Mirowski, Alexander Pritzel, Martin J Chadwick, Thomas Degris, Joseph Modayil, et al. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, 2018.

Mark A Bourjaily and Paul Miller. Dynamic afferent synapses to decision-making networks improve performance in tasks requiring stimulus associations and discriminations. *Journal of neurophysiology*, 108(2):513–527, 2012.

Tianhao Chu, Yuling Wu, Neil Burgess, Zilong Ji, and Si Wu. Unfolding the black box of recurrent neural networks for path integration. *Advances in Neural Information Processing Systems*, 38: 101975–102008, 2026.

Christopher J Cueva and Xue-Xin Wei. Emergence of grid-like representations by training recurrent neural networks to perform spatial localization. *arXiv preprint arXiv:1803.07770*, 2018.

Clare Howarth, Padraig Gleeson, and David Attwell. Updated energy budgets for neural computation in the neocortex and cerebellum. *Journal of Cerebral Blood Flow & Metabolism*, 32(7): 1222–1232, 2012.

Leo Kozachkov, John Tauber, Mikael Lundqvist, Scott L Brincat, Jean-Jacques Slotine, and Earl K Miller. Robust and brain-like working memory through short-term synaptic plasticity. *PLoS computational biology*, 18(12):e1010776, 2022.

Aaron T Kuan, Giulio Bondanelli, Laura N Driscoll, Julie Han, Minsu Kim, David GC Hildebrand, Brett J Graham, Daniel E Wilson, Logan A Thomas, Stefano Panzeri, et al. Synaptic wiring motifs in posterior parietal cortex support decision-making. *Nature*, 627(8003):367–373, 2024.

Kenneth W Latimer, Jacob L Yates, Miriam LR Meister, Alexander C Huk, and Jonathan W Pillow. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science*, 349(6244):184–187, 2015.

Peter Lennie. The cost of cortical computation. *Current biology*, 13(6):493–497, 2003.

Thomas Zhihao Luo, Timothy Doyeon Kim, Diksha Gupta, Adrian G Bondy, Charles D Kopec, Verity A Elliott, Brian DePasquale, and Carlos D Brody. Transitions in dynamical regime and neural mode during perceptual decisions. *Nature*, 646(8087):1156–1166, 2025.

Valerio Mante, David Sussillo, Krishna V Shenoy, and William T Newsome. Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, 503(7474):78–84, 2013.

Nicolas Y Masse, Guangyu R Yang, H Francis Song, Xiao-Jing Wang, and David J Freedman. Circuit mechanisms for the maintenance and manipulation of information in working memory. *Nature neuroscience*, 22(7):1159–1167, 2019.

Gianluigi Mongillo, Omri Barak, and Misha Tsodyks. Synaptic theory of working memory. *Science*, 319(5869):1543–1546, 2008.

Vidhya Rangaraju, Nathaniel Calloway, and Timothy A Ryan. Activity-driven local atp synthesis is required for synaptic function. *Cell*, 156(4):825–835, 2014.

Wade G Regehr. Short-term presynaptic plasticity. *Cold Spring Harbor perspectives in biology*, 4 (7):a005702, 2012.

James P Roach, Anne K Churchland, and Tatiana A Engel. Choice selective inhibition drives stability and competition in decision circuits. *Nature communications*, 14(1):147, 2023.

Jamie D Roitman and Michael N Shadlen. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *The Journal of neuroscience*, 22(21): 9475–9489, 2002.

Nathan S Rose, Joshua J LaRocque, Adam C Riggall, Olivia Gosseries, Michael J Starrett, Emma E Meyering, and Bradley R Postle. Reactivation of latent working memories with transcranial magnetic stimulation. *Science*, 354(6316):1136–1139, 2016.

H Francis Song, Guangyu R Yang, and Xiao-Jing Wang. Training excitatory-inhibitory recurrent neural networks for cognitive tasks: a simple and flexible framework. *PLoS computational biology*, 12(2):e1004792, 2016.

Yu Sun, Xiaolong Wang, Zhuang Liu, John Miller, Alexei Efros, and Moritz Hardt. Test-time training with self-supervision for generalization under distribution shifts. In *International conference on machine learning*, pp. 9229–9248. PMLR, 2020.

Xiao-Jing Wang. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, 36(5):955–968, 2002.

Michael J Wolff, Janina Jochim, Elkan G Akyürek, and Mark G Stokes. Dynamic hidden states underlying working-memory-guided behavior. *Nature neuroscience*, 20(6):864–871, 2017.

Kong-Fatt Wong and Xiao-Jing Wang. A recurrent network mechanism of time integration in perceptual decisions. *The Journal of neuroscience*, 26(4):1314–1328, 2006.

David M Zoltowski, Kenneth W Latimer, Jacob L Yates, Alexander C Huk, and Jonathan W Pillow. Discrete stepping and nonlinear ramping dynamics underlie spiking responses of lip neurons during decision-making. *Neuron*, 102(6):1249–1258, 2019.

Robert S Zucker and Wade G Regehr. Short-term synaptic plasticity. *Annual review of physiology*, 64(1):355–405, 2002.

APPENDIX CONTENTS

A. Implementation and functional analysis of the RNN model 12
A.1 Trial generation and behavioral readout 12
A.2 Network dynamics and numerical implementation 13
A.3 Optimization and checkpoint selection 14
A.4 Configuration summary 16
A.5 Ramp–step BIC comparison 16
A.6 Choice tuning and functional population assignment 18
A.7 Sensory, readout, and recurrent connection pruning 20
A.8 STP-strength distributions and matched neuronal silencing 20
B. Implementation and stability analysis of the theoretical model 22
B.1 Simulation and parameterization 22
B.2 Transition detection and stability analysis 22
B.3 Expected-spike-count comparison 24

A IMPLEMENTATION AND FUNCTIONAL ANALYSIS OF THE RNN MODEL

The methods follow the order of the main text: task and network specification, optimization and checkpoint selection, paired-configuration summary, temporal-dynamics analysis, functional population assignment, connection pruning, and neuronal silencing.

A.1 TRIAL GENERATION AND BEHAVIORAL READOUT

The reaction-time task has ten nonzero conditions formed by coherence $c \in \{1, 2, 4, 8, 16\}\%$ and two latent directions. Gradient minibatches sample coherence and direction uniformly on non-catch trials. The official task sets the gradient-trial catch probability to $5/11$; validation cycles over the ten nonzero conditions and one catch condition. Each 1100-trial validation set therefore contains 100 trials from every condition. Zero-coherence trials are used only for behavioral evaluation. Gradient, validation, and parameter-initialization random streams are separate.

The clean task inputs are zero during fixation. During the stimulus, the target evidence channel receives $(1 + 3.2c/100)/2$, the other evidence channel receives $(1 - 3.2c/100)/2$, and the third channel receives 1. The simulator adds a baseline of 0.2 and Gaussian input noise with time-step-dependent scaling to all three channels and rectifies the resulting inputs. The latent direction, rather than the noisy realized inputs, defines the target choice. On catch trials all three clean task inputs remain zero for 2000 ms.

The two linear readouts are trained against piecewise-constant targets. On a non-catch training or validation trial, both targets equal 0.2 during the 100-ms fixation period. The output mask is zero during the first 300 ms after stimulus onset, so this interval contributes no task loss. During the remaining 500-ms response interval, the target is 1.2 for the output corresponding to the latent direction and 0.2 for the other output. On a catch trial, both targets remain at 0.2 and both outputs are supervised throughout the 2000-ms trial. Padding beyond the end of a trial has zero mask weight.

Training and validation non-catch trials therefore last 900 ms: 100 ms of fixation followed by 800 ms of motion evidence. The held-out simulations used for Figures 2 and 3 follow the official test-time protocol instead, with 300 ms of fixation and 1500 ms of continuously presented motion evidence. This duration change affects only evaluation; it does not alter the learned parameters or the 300-ms onset of response supervision used during training. Table 1 summarizes the training, validation, and evaluation protocols.

The two linear readouts are $z_o(t) = \sum_i W^{\mathrm{out}}_{oi} r_i(t)$; only excitatory neurons have nonzero output connections. Let z_{tbo} and y_{tbo} denote the predicted and target values for time point t, trial b, and output $o \in \{1, 2\}$, and let $m_{tbo} \in \{0, 1\}$ denote the output mask. The task loss is the masked mean squared error

$$\mathcal{L}_{\mathrm{task}} = \frac{\sum_{t,b,o} m_{tbo}(z_{tbo} - y_{tbo})^2}{\sum_{t,b,o} m_{tbo}}. \tag{9}$$

Table 1: Trial protocols used for training and the reported analyses. All held-out simulations use a 0.5-ms integration step and save outputs every 10 ms.

Use	Fixation	Stimulus	Trials per network	Role
Gradient training	100 ms	800 ms	minibatches of 20	Parameter optimization; catch probability $5/11$
Validation	100 ms	800 ms	1100 per check	Checkpoint selection from ten evidence conditions and one catch condition
Figure 2	300 ms	1500 ms	2000	Temporal-shape and transition-time analyses
Figure 3a,e	300 ms	1500 ms	600	Choice tuning at 16% coherence; panel e combines these labels with checkpoint STP strengths
Figure 3b–d,f	300 ms	1500 ms	16,500 per condition	Accuracy under pruning or silencing

The normalization is over supervised output elements rather than first averaging within each trial or condition. The reported root-mean-squared error is $\sqrt{\mathcal{L}_{\text{task}}}$.

Training and validation performance is the percentage of non-catch trials for which the larger output at trial end identifies the target choice. Threshold-based reaction time is an evaluation measure rather than the training objective. Outputs are sampled every 10 ms, and the response is the first saved sample at which the larger of the two linear outputs is strictly greater than 1.0; its argmax gives the reported choice. Thus the first crossing wins across samples, and the larger output resolves a same-sample double crossing. A trial with no crossing before its end is a timeout. Reaction time is the crossing time minus stimulus-onset time. Analyses specify whether accuracy includes timeouts; latency summaries are restricted to the stated responded or correct-responded trials.

On the matched held-out streams, both conditions showed the expected dependence on sensory evidence: final-choice accuracy increased and median reaction time decreased with motion coherence (Figure S3). Accuracy is the final-output argmax over all attempted trials; reaction-time summaries include correct threshold responses only. Curves and error bars first summarize trials within each network and then report the mean and SEM across the ten independently trained networks.

A.2 Network dynamics and numerical implementation

Equations 1–4 define the neural and synaptic dynamics. We use $\tau_r = 100$ ms, $U = 0.15$, $\tau_u = 1.5$ s, $\tau_x = 0.2$ s, $s_r = 0.1$ s^{-1}/a.u., $u_j(0) = U$, $x_j(0) = 1$, and $T_g = 0.1$ (Masse et al., 2019). Activity is expressed in arbitrary units, so s_r is a numerical activity-to-rate conversion rather than a physiological calibration. The initial conditions give $u_j(0)x_j(0)/U = 1$ and therefore preserve resting recurrent strength. The variables (a_j, u_j, x_j) are presynaptic-unit-specific and shared by all recurrent projections leaving unit j; STP is not applied to the input or readout weights.

The implementation combines two public code bases at fixed revisions:

- the task, E–I rate network, objective, optimizer, and long-term parameter initialization follow Pycog (Song et al., 2016) at revision `7c5d54b` (`https://github.com/frsong/pycog`);
- the utilization and resource-depletion updates are adapted from the code accompanying Masse et al. (2019), revision `6c3c486` (`https://github.com/nmasse/Short-term-plasticity-RNN`); and
- the normalized efficacy, learnable STP strength, truncated-Gaussian initialization of STP strength, and checkpoint snapshots are additions made within the Pycog training framework. No task, neuronal, readout, or long-term-weight component is imported from the STP source network.

The initial STP strengths are sampled from a normal distribution with mean μ and truncated to the open unit interval. We use $\mu \in \{0.1, 0.5, 0.9\}$ for the initialization comparison and $\mu = 0.1$ for the

primary multiseed analyses:

$$a_j(0) \sim \mathcal{N}(\mu, 0.05^2), \qquad 0 < a_j(0) < 1, \qquad \theta_j(0) = T_g \operatorname{logit} a_j(0). \tag{10}$$

They are learned jointly with the long-term weights; U, τ_u, τ_x, s_r, and the initial synaptic states remain fixed. The model uses no STP-strength regularizer. Unit-gain and depletion-off checks recover the corresponding static and facilitation-only computations to numerical precision.

For comparison with a Static network, resting effective recurrent strength is distinguished from the unmodulated weight matrix. When synaptic efficacy is normalized by its resting value, weight initialization and regularization use the same convention. The E/I constraint fixes outgoing signs; dynamical excitation–inhibition balance is not inferred from this constraint alone.

All 100 recurrent units receive the three task inputs. Recurrent connectivity contains every directed connection except self-connections; Dale's law fixes the sign of each unit's outgoing recurrent weights. Input weights are nonnegative, and only the 80 excitatory units project to the two outputs. Recurrent weights are initialized from the official gamma distribution and rescaled to spectral radius 1.5; input and output weights use the official uniform initialization. Long-term weights (input, recurrent, and output) and the initial neural state are trainable. The initial state begins at $h_i(0) = 0.1$ for every unit, whereas recurrent and output biases remain fixed at zero.

Input noise has simulator parameter $\sigma_{\text{in}} = 0.01$ and time constant 100 ms; recurrent noise has parameter $\sigma_{\text{rec}} = 0.15$. The simulator rescales both stochastic inputs with the integration step. Sensory inputs are rectified after adding the constant input level and input noise. Neural states are reset to the learned initial condition at each trial. Synaptic states are reset to $u_j(0) = U$ and $x_j(0) = 1$ for every presynaptic neuron j.

The code evaluates the continuous dynamics using the forward Euler method. In the discrete equations below, Δt and all time constants are expressed in milliseconds, and $\Delta t_s = \Delta t/1000$ is the time step in seconds. The superscript n denotes the integration step, and $\Pi_{[0,1]}$ clips its argument to $[0, 1]$. One step is

$$x_j^{n+1} = \Pi_{[0,1]}\left[x_j^n + \frac{\Delta t}{\tau_x}(1 - x_j^n) - \Delta t_s s_r u_j^n x_j^n r_j^n\right], \tag{11}$$

$$u_j^{n+1} = \Pi_{[0,1]}\left[u_j^n + \frac{\Delta t}{\tau_u}(U - u_j^n) + \Delta t_s s_r U(1 - u_j^n) r_j^n\right], \tag{12}$$

$$q_j^{\text{eff},n+1} = 1 + a_j\left(\frac{u_j^{n+1} x_j^{n+1}}{U} - 1\right), \tag{13}$$

$$h_i^{n+1} = \left(1 - \frac{\Delta t}{\tau_r}\right) h_i^n + \frac{\Delta t}{\tau_r}\left[\sum_j W_{ij} q_j^{\text{eff},n+1} r_j^n + \sum_k W_{ik}^{\text{in}} v_k^n + b_i + \xi_i^n\right],$$

$$r_i^{n+1} = [h_i^{n+1}]_+. \tag{14}$$

Thus the previous presynaptic rate first updates both synaptic states; their updated normalized product then modulates the recurrent drive, and the neural state is advanced last. Training uses $\Delta t = 20$ ms. Held-out evaluation uses $\Delta t = 0.5$ ms and saves activity and outputs every 10 ms while preserving the simulator's time-step-dependent noise scaling.

A.3 Optimization and checkpoint selection

Minibatches contain 20 trials, and Equation 9 is evaluated jointly over all supervised elements in each minibatch. Optimization uses float32 stochastic gradient descent (SGD) with a base learning rate of 0.01 and global gradient-norm clipping at 1. The update includes the original vanishing-gradient correction $\lambda_\Omega \nabla\Omega$, with $\lambda_\Omega = 2$. All explicit L1 weight, L2 weight, activity-rate, and STP-strength penalties are zero. The logged task loss and regularized objective are therefore identical; the separately logged Ω value contributes to the parameter update through its gradient but is not added to the displayed objective.

Both conditions train for exactly 100,000 updates and validate every 500 updates on 1100 trials. Training is divided into 0–50k and 50–100k stages; the continuation preserves the first-stage best

checkpoint, and final selection compares it with every continuation checkpoint. Checkpoints are selected by the minimum pure validation task loss over the complete 0–100k record rather than by root-mean-squared error, validation performance, or threshold-based reaction time. Specifically, if $\mathcal{K}_s$ contains the preserved first-stage best checkpoint and every available continuation checkpoint for seed s, the analyzed update is

$$\hat{k}_s = \arg\min_{k \in \mathcal{K}_s} \mathcal{L}^{\mathrm{val}}_{\mathrm{task}}(s, k). \tag{15}$$

Static networks and RNNs with STP are selected by the same rule and from the same training budget.

For the RNNs with STP in Figures 2 and 3, only the raw STP-strength-parameter update is multiplied by 400; all other parameters retain the base learning rate. In the complete seed-201 training log, sampled every 500 updates from 500 to 100,000, the median raw θ-gradient norm was 7.39×10^{-5} and the median full-parameter gradient norm was 9.64×10^{-2}. Their median ratio was 6.86×10^{-4} (interquartile range, 4.08×10^{-4}–1.28×10^{-3}), placing the unscaled STP-strength-parameter gradient norm approximately three orders of magnitude below the full gradient norm.

A same-initialization, 100,000-update calibration in seed 201 compared multipliers of 200, 400, 500, 600, and 1000. Their minimum validation task losses were 0.06797, 0.06489, 0.06962, 0.07486, and 0.07722, respectively. All five runs remained finite, and 400 gave the lowest validation loss in this calibration, so this multiplier was fixed before the ten-network analysis. The gradient diagnostic uses the full parameter norm rather than an isolated long-term-weight gradient, and the learning-rate sweep is a single-seed calibration.

RNNs with STP reached a validation performance of $91.9 \pm 2.2\%$ at the selected checkpoints (mean±SD across ten networks), and their learning trajectories are shown in Figure 2a. Static networks reached $96.2 \pm 0.6\%$, with learning trajectories shown in Figure S1. Validation performance uses the final-output argmax on the balanced non-catch validation trials and is distinct from threshold-response measures and the held-out test protocol.

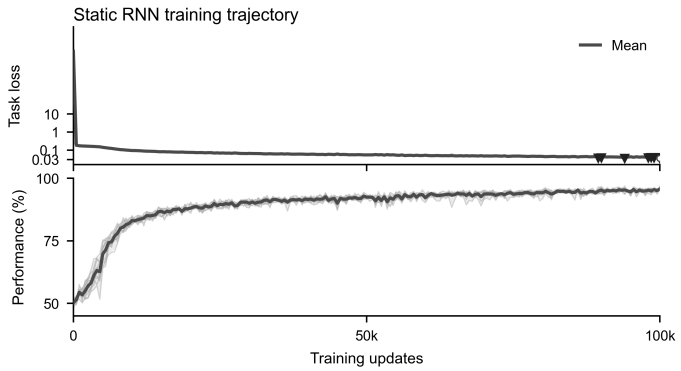

Figure S1: Static RNNs learn the reaction-time motion-discrimination task. Validation task loss and validation performance are shown over 100,000 training updates for ten independently initialized networks. Thin lines denote individual networks, the thick line denotes their mean, and shading shows the pointwise 2.5th–97.5th percentiles across networks. Downward triangles mark the minimum-validation-task-loss checkpoints; open circles mark the 100,000-update endpoints.

For training seed s, ordinary network parameters use seed s, while gradient and validation trials use fixed seeds 11 and 22, respectively. Static and STP runs with the same training seed share the ordinary initialization and task-stream seeds. For RNNs with STP, the Gaussian STP-strength sample uses seed $s + 910000$ and inverse-CDF truncation on $(0, 1)$, avoiding point masses at the boundaries. Seeds 201–210 define the ten primary STP networks paired with the ten Static networks. The initialization-sensitivity analysis contains ten additional networks initialized at mean STP strength 0.5 and ten initialized at 0.9, while retaining standard deviation 0.05, seeds 201–210, and the remaining protocol. Figure 2c therefore summarizes 30 STP trainings, ten per initialization condition; the 20 additional networks do not enter the primary paired comparisons.

For paired network-level comparisons, we use a two-sided exact paired randomization test that exhaustively flips the signs of the ten within-seed differences. Confidence intervals are percentile intervals obtained by resampling the ten paired networks 200,000 times with random seed 20260925. Trials and neurons within a network are not treated as independent replicates.

A.4 Configuration summary

The current main configuration and run inventory are summarized below. The RNN with STP is the default treatment, and the Static network is its computational control. Both retain the same reaction-time task, neural dynamics, long-term parameterization, optimizer, and training budget.

Table 2: Implemented configuration and training coverage for the paired reaction-time comparison.

Field	Static control	RNN with STP
Network size and integration	3 inputs, 100 rate units (80 E/20 I), 2 outputs; $\tau_r = 100$ ms; training $\Delta t = 20$ ms	Same; evaluation $\Delta t = 0.5$ ms and saved interval 10 ms for both conditions
Behavioral protocol	Official `rdm_rt`, including the 300-ms supervised target delay	Same
Trainable parameters	$W^{\mathrm{in}}, W, W^{\mathrm{out}}$, and initial neural state	Same, plus one STP-strength parameter θ_j per presynaptic neuron
STP constants and allocation	$q_j^{\mathrm{eff}} = 1$	Facilitation and depression on all 100 recurrent presynaptic units; $U = 0.15$, $\tau_u = 1500$ ms, $\tau_x = 200$ ms, $s_r = 0.1$, $T_g = 0.1$
Optimization and regularization	Minibatch SGD, lr 0.01, batch 20, $\lambda_\Omega = 2$, global norm limit 1	Same
Training budget and seed count	100,000 updates; seeds 201–210	Same; STP-strength learning-rate multiplier 400
Selected checkpoint	Minimum validation task loss over 0–100,000 updates	Same
Validation performance	$96.2 \pm 0.6\%$	$91.9 \pm 2.2\%$
Current coverage	10 reaction-time networks	10 primary reaction-time networks at mean initial STP strength 0.1, plus 10 additional sensitivity networks at each mean of 0.5 and 0.9 (30 STP networks in Figure 2c)

A.5 Ramp–step BIC comparison

A.5.1 Evaluation set and fitted models

Figure 2 is evaluated on 200 trials in each of the ten nonzero coherence–direction conditions (2000 trials per network), using random seed 100 and the 300-ms-fixation/1500-ms-stimulus test protocol. A trial batch is generated once and replayed through the paired Static RNN and RNN with STP, so their input and recurrent-noise realizations are matched; resetting the evaluation seed also gives every training seed the same external trial stream. Behavioral summaries include all attempted trials, whereas the shape analysis admits only trials with a correct threshold response and a complete requested window. The denominator of each displayed trajectory fraction is the number of eligible fitted trials in that network.

The network-level comparison in Figure 2f uses the target readout minus the nontarget readout in the response-aligned interval from -400 to 0 ms. Representative-neuron analyses in Figure 2b,d,e use the interval from 0 to 600 ms after stimulus onset. Both are sampled every 10 ms and fitted without smoothing.

For samples $(t_\ell, y_\ell)_{\ell=1}^n$, the ramp model is $\hat{y}_\ell = \beta_0 + \beta_1 t_\ell$. The step model has one change point k and two constant levels,

$$\hat{y}_\ell = \begin{cases} \mu_1, & \ell \le k, \\ \mu_2, & \ell > k, \end{cases} \tag{16}$$

where every candidate leaves at least $\max(3, \lceil 0.1n \rceil)$ samples on each side; the candidate with minimum residual sum of squares is retained. Under a common independent-Gaussian residual approximation,

$$\mathrm{BIC}_m = n \log(\mathrm{RSS}_m/n) + p_m \log n, \qquad p_{\mathrm{ramp}} = 2, \quad p_{\mathrm{step}} = 3. \tag{17}$$

We define $\Delta\mathrm{BIC} = \mathrm{BIC}_{\mathrm{ramp}} - \mathrm{BIC}_{\mathrm{step}}$ and label trials with $\Delta\mathrm{BIC} > 2$ as step-like, those with $\Delta\mathrm{BIC} < -2$ as ramp-like, and the remainder as indeterminate. This is an operational comparison of two simple trajectory families under an independent-Gaussian residual approximation; temporal autocorrelation limits a literal generative interpretation.

Each trained network is the independent unit for condition-level comparisons. Trial fractions are first computed within each network, and paired differences then compare matched Static and STP seeds. Trials and units within a network are not treated as independent network replicates.

A.5.2 Neuron-level prevalence and pre-transition activity

We complemented the choice-output analysis with a split-sample neuronal analysis. Even-numbered trials determined each neuron's preferred choice from its mean activity during the final 100 ms of the 0–600-ms fitting window. Odd-numbered correct threshold-response trials in that preferred direction were used for classification. We retained neurons with at least 20 assessment trials and peak activity of at least 0.2 on at least half of those trials. A retained neuron was classified as activity-silent when at least half of its active trials favored an upward step ($\Delta\mathrm{BIC} > 2$ and $\mu_2 > \mu_1$) and its median normalized pre-transition deviation across those upward-step trials was at most 0.2. For trial i, this deviation was

$$d_{\mathrm{pre},i} = \frac{|\mu_{1,i} - r_{0:50,i}|}{r_{\mathrm{peak},i} - r_{0:50,i}}, \tag{18}$$

where $r_{0:50,i}$ is mean activity during the first 50 ms after stimulus onset. This permits a positive resting rate but requires the fitted pre-transition level to remain close to the trial's own initial activity.

Across the ten RNNs with STP, $32.9 \pm 9.7\%$ of recurrent neurons met this joint shape-and-pre-transition-activity criterion (mean±SD), compared with $18.1 \pm 4.7\%$ in the ten paired Static RNNs. The paired mean increase was 0.148 (95% bootstrap confidence interval, 0.089–0.210; exact paired-randomization $p = 0.004$). The corresponding excitatory and inhibitory fractions were $34.8 \pm 9.1\%$ and $25.5 \pm 15.9\%$ in RNNs with STP, and $16.8 \pm 3.8\%$ and $23.5 \pm 11.6\%$ in Static RNNs. The choice-output trajectory fraction in Figure 2f instead quantifies trajectory shape alone using $\Delta\mathrm{BIC} > 2$; it does not apply the neuron-level pre-transition-activity criterion.

The main-text prevalence restricts the denominator to choice-tuned neurons. Let A_s be the activity-silent neurons identified above and C_s the choice-tuned neurons in network s, classified by the procedure in Appendix A.6. Choice tuning was evaluated independently at the selected checkpoint of each Static and STP network using the same high-coherence trials, analysis window, and Welch-test/BH-FDR rule. For each network, we computed

$$P_s^{\mathrm{silent|choice}} = \frac{|A_s \cap C_s|}{|C_s|}. \tag{19}$$

The ten network-level fractions were then averaged with equal weight. They were $52.2 \pm 16.2\%$ for RNNs with STP and $36.2 \pm 7.0\%$ for Static RNNs (mean±SD). The paired mean increase was 0.160 (95% bootstrap confidence interval, 0.068–0.249; exact paired-randomization $p = 0.016$). Conversely, $97.6 \pm 4.4\%$ of activity-silent neurons in RNNs with STP and $98.2 \pm 3.5\%$ in Static RNNs were choice tuned.

For the neurons meeting the joint criterion, mean activity during the final 50 ms before the fitted transition was $41.2 \pm 2.2\%$ of the raw trial peak. After subtracting each trial's initial activity and normalizing by its initial-to-peak range, the pre-transition deviation was $23.8 \pm 0.9\%$. These summaries are distinct from the fitted-level threshold in Equation 18 and provide a complementary description of the low-activity period.

A.5.3 Representative units and display conventions

Representative units for Figures 2d,e are selected on a separate discovery set containing 200 trials per direction at 16% coherence. Selection is restricted to excitatory units whose absolute readout-weight difference is at least the E-unit median, whose peak activity reaches 0.2 on at least half of

valid preferred-direction trials, and whose stimulus-aligned trials are classified as step-like at least half of the time. The unit closest to the eligible subgroup median in step fraction and peak activity is retained. Static representative units use the analogous majority-ramp criterion. These units are then evaluated on the independent 2000-trial stream described above.

Figure S2 illustrates the scoring rule on the four paired seed-210 trials shown in Figure 2b. Static unit 68 is paired with STP units 52, 32, 57, and 32 on trial IDs 1319, 1209, 1589, and 1611. The STP units passed the independent discovery screen; one clear paired example was then chosen from each of four fixed transition-time windows (50–100, 120–180, 430–480, and 220–270 ms). These contrastive examples illustrate the fitted models and do not enter the population prevalence estimate.

For Figure 2b, each trajectory is scaled by its 5th and 95th activity percentiles and clipped to $[0, 1]$. A centered 30-ms moving average is used only for the colored display line; the raw trace and all fits remain unsmoothed. Figure 2d uses 171 correct, preferred-direction trials from the representative STP unit at 16% coherence, restricted to step-classified trials with change points between 50 and 450 ms. The same percentile-normalized trials are averaged after alignment to stimulus onset or fitted change point. Bands are 95% percentile intervals from 5000 trial-level bootstrap resamples and quantify within-network uncertainty. Because fitted change points define the second alignment, its sharpening is descriptive. Figure 2e groups within-network median transition times as low (1–2%), medium (4–8%), or high (16%) coherence and reports the mean and SD across networks. Figure 2c averages network-level empirical cumulative distribution functions (ECDFs) of the 100 STP strengths and shows pointwise 2.5th–97.5th percentiles across networks.

Across networks, the mean within-network median transition time decreased from 162 ± 28 ms at low coherence to 115 ± 10 ms at high coherence (mean±SD). Applying the same change-point procedure to the independently selected majority-ramp Static representative units gave 272 ± 98 ms at low coherence, 233 ± 88 ms at medium coherence, and 183 ± 61 ms at high coherence. Nine Static networks contributed a low-coherence median because one had no eligible low-coherence step-classified trial; all ten contributed to the other groups. Thus evidence-dependent transition timing was not unique to the STP condition, whereas activity-silent, step-like choice-output trajectories, operationally defined by $\Delta\mathrm{BIC} > 2$, were more prevalent with STP. Their fractions were 0.587 ± 0.040 for RNNs with STP and 0.383 ± 0.036 for Static networks. The paired mean increase was 0.204 (95% bootstrap confidence interval, 0.177–0.231; exact paired-randomization $p = 0.002$). The corresponding ramp-like fractions were 0.350 ± 0.037 and 0.579 ± 0.033, and indeterminate fractions were 0.063 ± 0.004 and 0.038 ± 0.008.

A.6 Choice tuning and functional population assignment

Figure 3 uses ten RNNs with STP, all with mean initial STP strength 0.1 and training seeds 201–210. In these runs, an STP-strength learning-rate multiplier of 400 means that the raw parameter θ_j is updated at 400 times the base learning rate, while all other trainable parameters retain the base rate of 0.01. Each network is evaluated using its minimum-validation-task-loss checkpoint over its 100,000-update run. The multiseed analysis obtains the selected update from the checkpoint and applies that network's tuning labels to all subsequent pruning and STP-strength analyses. These labels are estimated independently of training and of the perturbation evaluation stream. Each trained network is one independent replicate, and Figure 3 reports network means and SEM.

For each network, 300 left-evidence and 300 right-evidence trials are generated at 16% coherence, with 300 ms of fixation and 1500 ms of stimulus presentation. Opposite-direction trials are generated in pairs sharing the same input and recurrent noise realizations; the pair seeds are drawn using random seed 20260916. Simulation uses a 0.5-ms integration step and saves activity every 10 ms. A trial enters the tuning analysis if its first output threshold crossing (strictly above 1.0) indicates the correct direction. Rates are averaged over absolute trial time 1400–1800 ms, inclusive, giving 41 saved samples per trial. This is the final 400 ms of the trial, corresponding to 1100–1500 ms after motion onset while evidence remains present.

For each of the 100 neurons, the implementation applies a two-sided Welch test to the trial-level window means in the two conditions, with Benjamini–Hochberg correction across all neurons within that network. We denote the adjusted p value for neuron j by p_j^{adj}. Neurons with $p_j^{\mathrm{adj}} < 0.05$ are classified as choice-selective; the sign of the left-minus-right mean defines preference. E/I identity

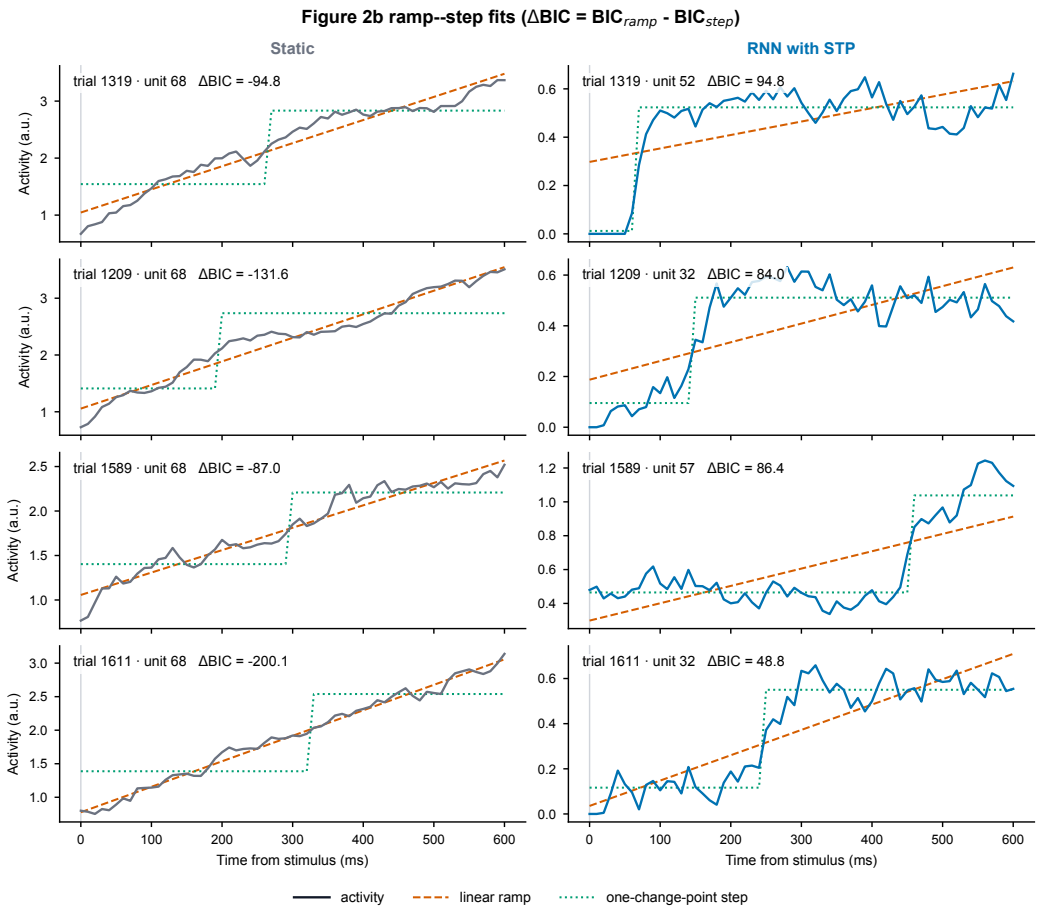

Figure S2: Representative fits for the ramp–step BIC comparison. Rows are the same four paired seed-210 trials shown in Figure 2b; columns show the Static and STP networks. Solid curves show raw unsmoothed activity from 0 to 600 ms after stimulus onset, dashed orange curves show fitted ramps, and dotted green curves show fitted one-change-point steps. Unit identifiers and ΔBIC values are given in each panel. Positive ΔBIC favors the step model and negative values favor the ramp model. These contrastive examples illustrate scoring rather than population prevalence.

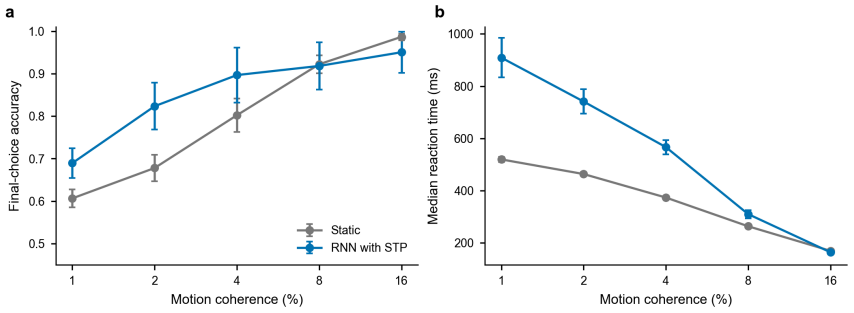

Figure S3: Held-out behavior across motion coherence. **a,** Final-choice accuracy over all attempted nonzero-coherence trials. **b,** Median reaction time on correct threshold-response trials. Points and error bars show the mean and SEM across ten paired networks; each network contributes 400 trials per coherence, equally divided between the two directions.

then gives E_L, E_R, I_L, and I_R, and nonselective neurons remain in separate E and I groups. No effect-size or slope criterion enters this classification.

The Welch test treats the retained conditions as two samples although their noise realizations are paired. Figure 3a summarizes the same trials used for classification. Before cross-network aggregation, each network's group means are divided by its maximum unit-level condition mean, without subtracting a minimum; empty groups are omitted. The labels therefore describe late-trial selectivity when evidence direction and choice coincide, rather than separating sensory and choice coding. Across networks, $59.2 \pm 2.9\%$ of excitatory neurons and $76.5 \pm 3.8\%$ of inhibitory neurons were choice-selective (mean±SEM).

A.7 Sensory, readout, and recurrent connection pruning

The pruning experiments hold the checkpoint and functional group assignments fixed and set the specified weight block to zero during forward evaluation, without retraining. Within a network, all variants share the same inputs and recurrent noise. Each evaluation uses random seed 100, with 1500 trials for each of the ten nonzero coherence–direction conditions ($c = 1, 2, 4, 8, 16\%$ in both directions), plus 1500 zero-coherence trials. Thus 16,500 trials are simulated per variant; overall accuracy is the fraction of correct final-output argmax choices among the 15,000 nonzero-coherence trials. Threshold timeouts are retained in this denominator and judged by the final output. Simulation uses a 0.5-ms step and 10-ms output sampling.

Let $A_s^{(0)}$ and $A_s^{(v)}$ be intact and pruned accuracies, expressed as fractions, in training seed s. Heatmaps show the loss

$$D_s^{(v)} = 100\big(A_s^{(0)} - A_s^{(v)}\big), \tag{20}$$

in percentage points. Positive values indicate impaired performance and negative values indicate improvement.

The three connection-pruning analyses are:

1. **Sensory input (Figure 3b).** Set $W^{\text{in}}_{G,k} = 0$ for each of the six populations G and each sensory channel $k \in \{0, 1\}$, giving 12 pruned conditions plus the intact condition. The cue channel remains connected. Because the complete input weight is removed, this operation deletes both the coherence-dependent signal and the noisy baseline-offset contribution carried by that channel.
2. **Choice readout (Figure 3c).** Disconnect each of the three E populations from both outputs in turn by setting $W^{\text{out}}_{:,G} = 0$; direct I-to-output weights are structurally absent. Effects are evaluated separately on 7500 left-evidence and 7500 right-evidence trials and referenced to intact accuracy on the same evidence side. This pruning analysis therefore identifies a source population, not a single output channel.
3. **Recurrent pathways (Figure 3d).** For every ordered pair of the four tuned populations, remove the connections from source G to target H by setting the complete block $W_{H,G} = 0$, giving 16 pruned conditions plus intact. Within-group blocks are included, nonselective neurons remain active, and blocks involving nonselective groups are not tested. Empty blocks produce a recorded effect of zero.

Each 2×2 E/I matrix in Figure 3d is color-scaled independently to emphasize its internal rank structure; color intensity is therefore not comparable across matrices, and cross-matrix effect magnitudes must be read from the printed mean ± SEM values. Outlines mark the largest across-network mean loss among the three off-diagonal targets for each source and, separately, the two largest mean effects in the E-to-E submatrix. These rankings are descriptive and are not accompanied by across-network hypothesis tests. The removed-connection count and the total removed weight magnitude are not matched across blocks, so the panel quantifies dependence on whole pathways rather than per-synapse importance or anatomical connection probability.

A.8 STP-strength distributions and matched neuronal silencing

Across networks, 98.6% of high-STP neurons were choice-selective, compared with 16.1% of low-STP neurons (Figure 3e). Accuracy was 85.7% in intact networks, 54.9% after high-STP neurons

were silenced, and 82.0% after low-STP neurons were silenced (Figure 3f). Accuracy after low-STP silencing exceeded that after high-STP silencing by 27.1 percentage points on average (95% bootstrap confidence interval, 17.2–36.0; exact paired-randomization $p = 0.004$).

STP strength in Figure 3e denotes the learned a_j. During evaluation, a_j remains constant, while normalized synaptic efficacy $u_j x_j / U$ and the recurrent transmission factor q_j^{eff} evolve within each trial. STP strengths are histogrammed separately for E and I neurons in bins of width 0.02. For display, each network's bin counts are smoothed with a Gaussian kernel of standard deviation one bin and radius three bins, with edge-value padding, then summarized as network means and SEM. The ordinate is neurons per network, not a probability density. The two peaks describe the displayed distributions; no formal test of bimodality is used. High-STP membership is $a_j \geq 0.5$ and low-STP membership is $a_j < 0.5$. At the selected checkpoints, high-STP neurons accounted for $56.1 \pm 10.5\%$ of all neurons, $53.8 \pm 10.0\%$ of E neurons, and $65.5 \pm 15.4\%$ of I neurons (mean±SD across networks). Choice-selective fractions use every member of the corresponding threshold-defined group in each network, and are then averaged over networks with equal weight. The choice-selective fractions were 98.7% versus 12.5% for high- versus low-STP E neurons and 98.1% versus 35.0% for the corresponding I groups. The high-versus-low contrast was insensitive to the exact STP-strength cutoff: with cutoffs of 0.4, 0.5, and 0.6, the across-network choice-selective fractions were respectively 98.6% versus 15.9%, 98.6% versus 16.1%, and 98.5% versus 17.1%. This sensitivity analysis changes only the grouping used in Figure 3e; the matched silencing subsets in Figure 3f retain the primary cutoff of 0.5 and their deterministic extreme-STP-strength selection.

Figure 3f uses matched subsets of those groups. Within each network and each cell type $c \in \{E, I\}$, let $n_c = \min(N_{c,\text{high}}, N_{c,\text{low}})$. The silencing comparison selects the n_c largest STP strengths from the high group and the n_c smallest STP strengths from the low group. Combining the E and I subsets matches both total count and E/I composition. Selection is deterministic rather than random sampling; the individual panel reports a rounded mean of 40 silenced neurons per network in each condition.

For a selected set S, silencing is imposed throughout the trial as

$$r_j^n \leftarrow 0, \qquad j \in S, \quad n = 0, 1, \ldots. \tag{21}$$

This assignment is applied at initialization and after every neural-state update. Recurrent and readout contributions from selected neurons are therefore removed, while the neural state equations and STP updates continue to run. The analysis evaluates intact networks and networks with the matched high- or low-STP subsets silenced, using the same behavioral stream and accuracy definition as pruning. This comparison tests the contribution of neurons selected by STP strength, including their learned connectivity and choice tuning. The paired-randomization test uses the within-network accuracy difference between low-STP and high-STP silencing conditions and follows the procedure specified in Appendix A.3.

B IMPLEMENTATION AND STABILITY ANALYSIS OF THE THEORETICAL MODEL

The simulations use the four-population model defined in Equations 5–8. Its parameters are specified below. The displayed Static comparison and the equal-time-constant comparison are reported separately; neither changes the trained RNN configuration.

B.1 SIMULATION AND PARAMETERIZATION

For cell class $a \in \{E, I\}$, the activation function is

$$\phi_a(z) = \frac{R_a}{1 + \exp[-(z - \theta_a)/\beta_a]}, \qquad \phi_a'(z) = \frac{\phi_a(z)}{\beta_a}\left(1 - \frac{\phi_a(z)}{R_a}\right). \tag{22}$$

Firing rates are in Hz and time is in seconds. Inputs use the model's dimensionless input scale, so J_{EE}, J_{EI}, and J_{IE} convert rates into that scale. Parameter values are listed in Table 3.

Table 3: Parameters of the four-population theoretical model.

Parameter	Value
J_{EE}, J_{EI}, J_{IE}	$0.00495,\ 0.003,\ 0.00405$
B_E, B_I	$0.20,\ 0.22$
R_E, R_I	$60,\ 80$ Hz
θ_E, θ_I; β_E, β_I	$0.4,\ 0.4$; $0.04,\ 0.04$
τ_E, τ_I (STP and equal-time-constant Static)	$20,\ 5$ ms
τ_E (Static in Figure 5d,e)	54.8383 ms
U, α	$0.05,\ 0.35$
τ_u, τ_x	$1500,\ 50$ ms
Input-noise correlation time	20 ms
Input-noise stationary SD, E and I	$0.015,\ 0.005$
Utilization and resource noise amplitudes	$1.4,\ 0.2\ \mathrm{s}^{-1/2}$
Integration and recording intervals	$1,\ 2$ ms
Fixation and stimulus duration	$0.3,\ 2.5$ s

During the stimulus, $e_L = 0.1 + 0.01c$ and $e_R = 0.1 - 0.01c$, where positive c favors left. Evidence inputs are zero during fixation. Each population receives Ornstein–Uhlenbeck input noise with the correlation time and stationary SD in Table 3. The excitatory noise terms are independent; the two inhibitory populations receive the same external noise realization in the displayed simulations. Their neuronal activities remain separate variables.

The deterministic drift is integrated by the fourth-order Runge–Kutta method. After each step, independent Gaussian increments are added to the log odds of each excitatory u_p and x_p, with SD proportional to the square root of the integration interval; this corresponds to bounded Stratonovich noise for $z \in \{u_p, x_p\}$. Inhibitory synapses have no STP or synaptic noise. Each simulation starts from the no-stimulus static neuronal equilibrium with $u_p = U$ and $x_p = 1$, followed by the fixation period. Static simulations set $q_L = q_R = 1$; this differs from freezing the STP model at its initial factor 0.35. The robustness claim in Section 5.1 was verified by rerunning the timing and decision35 protocols with STP enabled on all four populations' outputs and applying q to all four populations in the frozen-synapse scan: the tristable-onset and center-instability boundaries are unchanged (0.664 and 0.862), and behavior is comparable (accuracy 65.0% versus 66.2%, median threshold-crossing time 197 ms versus 233 ms, mean spike count 4.24 versus 4.14).

B.2 TRANSITION DETECTION AND STABILITY ANALYSIS

A descriptive transition time in Figure 4b–d is fitted to the winning excitatory trace using two constant activity levels. Both intervals must last at least 100 ms, the fitted increase must be at least 15 Hz, and the explained variance must be at least 0.5. Figure 5c–e instead uses the first crossing of $|r_{E,L} - r_{E,R}| = 35$ Hz. These are distinct alignment and response conventions. In the original model, the winner's synaptic gain crosses the center-instability boundary before the neuronal transition: over the 200 trials of the Figure 4 protocol, the winning population's $q_p = \alpha u_p x_p / U$ crosses

0.862 at a median of 137 ms, whereas the fitted activity transition occurs at a median of 228 ms, and the losing population's gain stays near the tristable onset (Figure S5).

The synaptic states also retain signed input history during the low-activity period. A linear decoder on the utilization and resource differences ($u_L - u_R$, $x_L - x_R$) recovers evidence direction above chance throughout the silent period, rising from 57.6% at 100 ms after stimulus onset to 64.4% at 400 ms (five-fold cross-validated linear discriminant analysis; 1600 trials over eight nonzero coherence levels, 200 trials per level and five seeds; chance 50%).

Psychometric curves (Figure S4), which measure choice accuracy as a function of evidence strength, use seeds 100–104 with 40 trials per coherence level, giving 2200 trials per condition. The spike-count comparison of Figure 5d,e uses an adjusted variant of the STP model in which the noise added to the synaptic states is set to zero, with all other parameters unchanged. The adjustment serves the fairness of the comparison: with synaptic-state noise, the STP model's psychometric accuracy falls well below that of the Static model, whereas the variant acts at accuracy comparable to Static, and the two curves overlap within 95% Wilson intervals at every evidence level. This behavioral equivalence, rather than a matched mean rise duration, defines the fair basis for the spike-count comparison of Figure 5e.

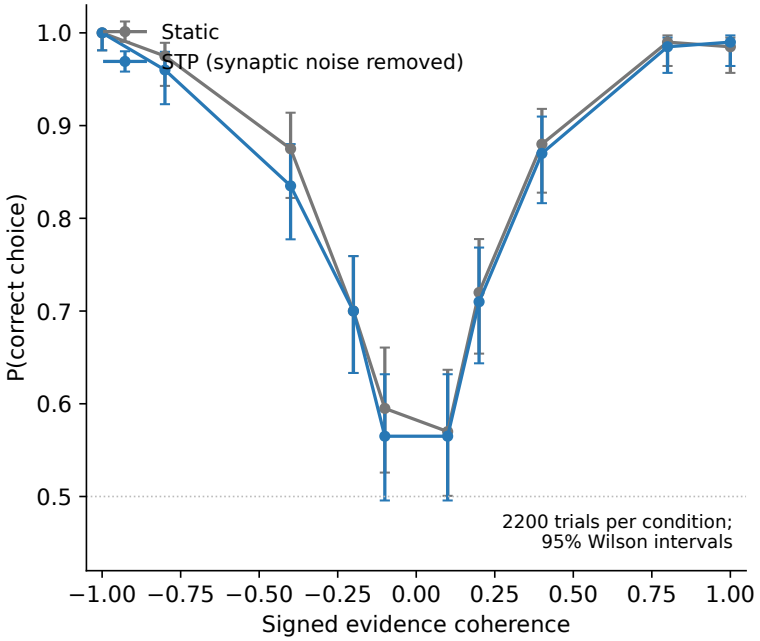

Figure S4: Psychometric comparison of the Static model and the adjusted STP variant used in Figure 5d,e (parameters adjusted for a matched psychometric curve; all other parameters unchanged). Fraction of correct choices as a function of signed evidence coherence; error bars show 95% Wilson intervals over 2200 trials per condition (11 coherence levels, 40 trials per level, five seeds). The curves overlap at every evidence level, so the two models decide with comparable accuracy; this behavioral equivalence is the basis for the fair spike-count comparison in Figure 5e.

For stability analysis, we remove noise, apply equal stimulus inputs, and hold $q_L = q_R = q$ fixed. With the ordering $(r_{E,L}, r_{E,R}, r_{I,L}, r_{I,R})$, define

$$W = \begin{pmatrix} J_{EE} & 0 & 0 & -J_{IE} \\ 0 & J_{EE} & -J_{IE} & 0 \\ J_{EI} & 0 & 0 & 0 \\ 0 & J_{EI} & 0 & 0 \end{pmatrix}, \qquad Q = \mathrm{diag}(q, q, 1, 1). \tag{23}$$

At an equilibrium $\mathbf{r}^*$, let D_ϕ be the diagonal matrix of activation-function derivatives at the four population inputs and $D_\tau = \mathrm{diag}(\tau_E, \tau_E, \tau_I, \tau_I)$. The neuronal Jacobian is

$$A = D_\tau^{-1}(-I_4 + D_\phi W Q). \tag{24}$$

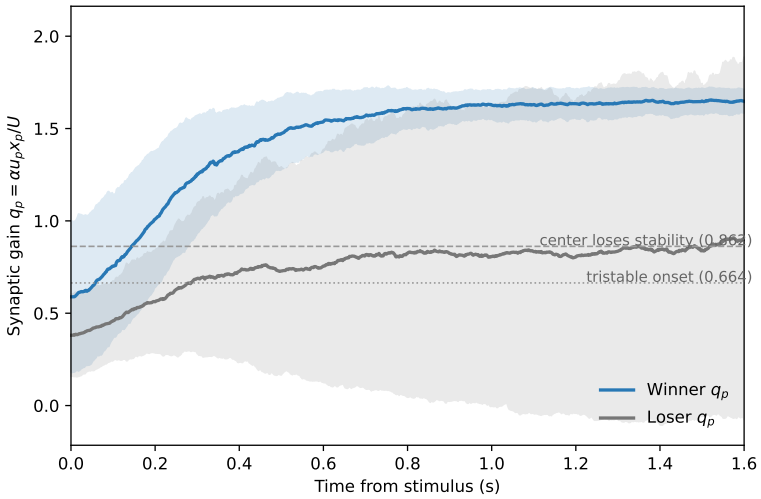

Figure S5: Synaptic gain relative to the frozen-synapse stability boundaries. Mean ± SD of $q_p = \alpha u_p x_p / U$ for the winning and losing populations across the 200 trials of the Figure 4 protocol in the original model. The winner's gain crosses the center-instability boundary (0.862) at a median of 137 ms, before the median activity transition (228 ms); the loser's gain stays near the tristable onset (0.664).

An equilibrium is stable if all four eigenvalues have negative real parts. The recorded equilibrium search spans $q = 0.35$–1.5 in increments of 0.002, combining scalar bracketing for symmetric roots with multiple initial guesses for asymmetric roots. The maximum recorded equilibrium residual is below 3.8×10^{-12}. At $q = 0.35$, the stable symmetric equilibrium has excitatory rates of approximately 5.1 Hz. The sampled data first show three stable equilibria at $q = 0.664$ and loss of stability of the symmetric equilibrium at $q = 0.862$. These values locate the regimes at the resolution of the parameter sweep.

At a symmetric equilibrium, write $g_E = \phi'_E$ and $g_I = \phi'_I$. The differences $d_E = (r_{E,L} - r_{E,R})/2$ and $d_I = (r_{I,L} - r_{I,R})/2$ obey the linearized equation

$$\frac{d}{dt}\begin{pmatrix} d_E \\ d_I \end{pmatrix} = \begin{pmatrix} (-1 + g_E J_{EE} q)/\tau_E & g_E J_{IE}/\tau_E \\ g_I J_{EI} q/\tau_I & -1/\tau_I \end{pmatrix} \begin{pmatrix} d_E \\ d_I \end{pmatrix}. \tag{25}$$

The positive off-diagonal terms express opponent inhibition: an increase in I_L relative to I_R favors E_L by suppressing E_R. A zero eigenvalue in this mode satisfies $g_E q(J_{EE} + g_I J_{IE} J_{EI}) = 1$. Both activation-function derivatives depend on the equilibrium. The common-activity mode and the other equilibria must also be checked; this fixed-synapse calculation does not establish the stability of the full system with dynamic u_p and x_p.

B.3 Expected-spike-count comparison

For Figure 5d,e, we define the total firing rate and expected spike count as

$$R_{\text{total}}(t) = r_{E,L}(t) + r_{E,R}(t) + r_{I,L}(t) + r_{I,R}(t), \tag{26}$$

$$N_{\text{spikes}} = \int_0^{T_{35}} R_{\text{total}}(t)\, dt. \tag{27}$$

Here, T_{35} is the first crossing of $|r_{E,L} - r_{E,R}| = 35$ Hz. All four rates contribute positively, including activity in the losing choice populations. Figure 5d shows the threshold-aligned total rate, and Figure 5e shows its integral over the full interval from stimulus onset to threshold crossing. The integration uses linear interpolation at the crossing time. Each condition contains five batches of 200 trials with independent noise seeds.

Each population is represented by one effective rate unit in this calculation. Because rates are in Hz and time is in seconds, their integral gives the expected spike count of the four-unit model.

For population index $a \in \{E_L, E_R, I_L, I_R\}$, let n_a denote population size and r_a its mean firing rate. Then the corresponding network spike count would be $\int \sum_a n_a r_a(t)\, dt$. We use the same unit population weights in both conditions and interpret their expected spike counts as a proxy for spike-related energy expenditure.

The Static value is $\tau_E = 54.8383$ ms, the STP value is $\tau_E = 20$ ms, and both models use $\tau_I = 5$ ms. The Static time constant was calibrated using 1000 separate trials (noise seed 510) to match the rise time of the stimulus-aligned mean excitatory activity difference of the original STP model from 12 to 48 Hz; these levels correspond to 20% and 80% of a fixed 60-Hz reference. The calibrated rise durations are approximately 226 ms for Static and 221 ms for STP in the original model. In the adjusted variant used for Figure 5d,e, the STP rise duration is approximately 145 ms, so that comparison is not rise-matched; its fairness rests instead on the overlapping psychometric curves of Figure S4. Evaluation uses seeds 410–414 for both the original model and the adjusted variant.

All evaluation trials reach the decision threshold. For the adjusted variant of Figure 5d,e, the mean expected spike counts are 8.803 for Static and 5.536 for STP, a reduction of approximately 37% in the STP model. Mean threshold-crossing times are 353 and 303 ms, and choice accuracies are 86.5% and 84.7%, respectively. Spike counts are integrated to each model's own threshold crossing; the STP model both crosses earlier and accumulates less. The overlapping psychometric curves (Figure S4) indicate that the comparison takes place at comparable accuracy. Across the five evaluation noise batches, the adjusted variant's mean spike count is lower than the Static model's in every batch (paired difference -3.27 spikes; the exact sign-flip test over five batches gives $p = 0.0625$, the smallest attainable value for five pairs; a trial-level bootstrap 95% confidence interval is $[-3.42, -3.12]$, 200,000 resamples).