COGNITIVE MAP AS REFERENCE PATH MEASURE: THE HIPPOCAMPUS AS A SCHRÖDINGER BRIDGE SAMPLER

Anonymous authors
Paper under double-blind review

ABSTRACT

A cognitive map enables animals to recall past journeys, plan future routes, and imagine unexperienced paths. How a single neural circuit generates such diverse sequential experiences from learned knowledge remains a fundamental open question. We propose that a cognitive map defines a reference path measure, a probability distribution over sequence spaces, and that the hippocampus constructs goal-directed episodes by solving a Schrödinger bridge problem under endpoint constraints. In this framework, present cues and terminal evidence modulate sequence generation through a continuation field that evaluates path reachability to target outcomes. We demonstrate that this computation is naturally realized by hippocampal circuit dynamics: recurrent relaxation computes the continuation field, dendritic gain modulates assembly transitions, and inhibitory competition governs event completion. Furthermore, offline replay supplies internal endpoint expectations, enabling contrastive calibration of boundary representations against waking experience. Across network simulations and large-scale hippocampal electrophysiological recordings (HC-3, HC-11, and DANDI 000978), our model recovers exact path laws, reveals reusable temporal structure, accurately predicts complete-event trajectories and durations under boundary conditioning, and accounts for experience-dependent shifts in sleep replay. This work unifies attractor dynamics with conditional generative modeling, providing a shared computational foundation for memory retrieval, prospective planning, and offline consolidation.

1 INTRODUCTION

An animal can use its knowledge of a place to recall a journey, find a route to a goal, or imagine a path it has never travelled. A cognitive map gives these experiences a common structure (Tolman, 1948; O'Keefe & Nadel, 1978). Each use of that structure produces a particular sequence, with a beginning, a course, and an ending. How does a neural circuit generate one such episode from knowledge that supports many possible episodes?

Hippocampal activity reveals the elements of this computation. Recurrent interactions recover coherent representations from partial input (Neunuebel & Knierim, 2014; Guzman et al., 2016). These representations can be understood as attractors, population patterns toward which activity converges (Hopfield, 1982; Wills et al., 2005). Internally generated sequences move through representations of successive locations, following remembered routes or exploring possible futures (Johnson & Redish, 2007; Diba & Buzsáki, 2007; Pfeiffer & Foster, 2013). Convergence toward a location and transition to another can alternate within the brief population events known as sharp-wave ripples (Pfeiffer & Foster, 2015). Stable patterns give situations a recoverable identity; transitions give them a temporal relation.

We propose that these relations define a *reference path measure*: a probability distribution over complete sequences of represented situations. Learned transitions determine which paths the map supports and their relative probabilities. A present cue supplies information about where an episode begins; a remembered or intended outcome supplies information about how it ends. These boundary conditions select an ensemble of episodes from the reference (Figure 1). Familiar transitions can thereby support a journey that is new as a whole.

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.

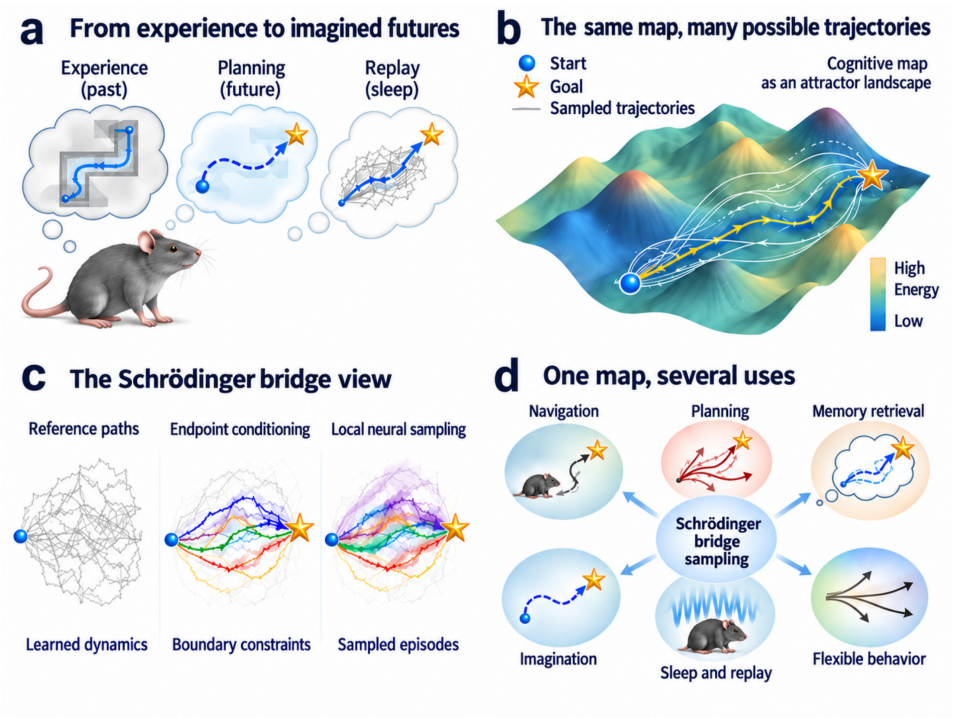

Figure 1: **A learned map supports many possible experiences.** (a) Remembering, planning, and replay use relations acquired through experience. (b) Stable patterns of population activity, called attractors, represent situations; transitions connect them into paths. (c) Endpoint information selects a path ensemble that local neural sampling can generate. (d) Different sources of this information give the computation different cognitive uses. The landscape and trajectories are conceptual illustrations; the path law and circuit dynamics are developed in Sections 2 and 3.

This inference is a Schrödinger bridge problem: find a distribution of paths that agrees with boundary observations while making the smallest change to a known law of motion (Schrödinger, 1931; Léonard, 2014). We formulate it for episodes whose lengths are also generated. A *continuation field* sums the support for compatible endings reachable from each state. It makes a distant outcome relevant at each intervening choice, coordinating where an episode starts, which states it visits, and when it finishes.

We derive a neural implementation of this computation. Recurrent relaxation propagates terminal input through the map, extending the navigation dynamics of He et al. (2026b) to complete sequence generation. Dendritic gain modulates candidate assemblies, and inhibitory competition selects continuation or completion. Replay supplies samples of the endpoints this circuit generates; their difference from experienced endpoints calibrates subsequent generation. Network simulations establish the resulting path law, and hippocampal recordings test its temporal structure, conditional event predictions, and learning consequences. The theory connects attractor dynamics to conditional generative modeling (De Bortoli et al., 2021; Liu et al., 2023) through a common account of how learned knowledge becomes an individual experience. Appendix B situates this computation relative to accounts of cognitive maps, replay, and planning (B.1–B.5) and to Schrödinger bridge methods in machine learning (B.6).

2 From Learned Dynamics to an Episode

2.1 The Schrödinger Bridge: Inferring Paths from Their Boundaries

Consider particles diffusing through a landscape. Their ordinary motion is known, and their distributions have been observed at two times. The observations leave many possible histories. Schrödinger's

problem asks which ensemble of paths explains both observations while making the smallest change to the known process (Schrödinger, 1931; Léonard, 2014). The answer retains what the dynamics already tell us about the intervening motion. A cognitive map faces the same inference: a present situation and a possible outcome constrain an episode whose course remains to be constructed.

We first specify the learned process. On a finite map in which all states are mutually reachable, let P_{ij} be the probability of moving from represented situation i to j, and π_i its stationary frequency. A situation can include position, heading, and context. A reference episode starts according to π. At each decision it continues with probability $\gamma \in (0, 1)$ and otherwise finishes at its current state. These probabilities define a *reference path measure* R over sequences and their lengths. An event with K transitions contains $K + 1$ states; a neural clock determines how quickly they are expressed.

For desired distributions of beginnings and endings, μ_0 and μ_T, the bridge minimizes relative entropy between the generated and reference path laws:

$$\mathsf{Q}^\star = \operatorname*{arg\,min}_{\mathsf{Q}:\, q_0=\mu_0,\, q_T=\mu_T} D_{\mathrm{KL}}(\mathsf{Q}\|\mathsf{R}). \tag{1}$$

Here q_0, q_T are the generated endpoint distributions. The terminal constraint applies where each episode finishes, so its length is part of the solution. For positive targets, the solution takes the form

$$\mathsf{Q}_{a,b}(x_{0:K}, K) = \frac{\pi_{x_0} a_{x_0}}{Z} \left(\prod_{k=0}^{K-1} \gamma P_{x_k x_{k+1}} \right) b_{x_K}. \tag{2}$$

The factors a_i, b_i express compatibility with initial and terminal information; Z sums the weights of all paths. Learning adjusts these factors until the generated marginals match the targets. Cues that supply compatibility directly also define a conditional ensemble through this equation.

The map determines how an episode can unfold. The boundary factors determine how often its beginnings and endings are recruited. Once both endpoints are fixed, their factors cancel from comparisons between paths: the relative probabilities of routes and lengths are those of the reference. This gives a precise meaning to reusing knowledge for a new purpose. Familiar transitions can support a new complete journey, while their learned probabilities continue to constrain its course (Appendix A.2.1).

2.2 The Doob transform: making an endpoint condition local

At a junction, a successor is useful according to the completions accessible beyond it. Let h_i sum the compatibility of those completions, weighted by the probabilities of the paths that reach them. A path can finish at i, contributing b_i, or take a step and reconsider completion:

$$h_i = b_i + \gamma \sum_j P_{ij} h_j, \qquad h = (I - \gamma P)^{-1} b. \tag{3}$$

This *continuation field* propagates terminal evidence through the map. Several probable routes contribute more than a narrow, rarely used connection. Each additional step attenuates its contribution by γ. Thus the field combines learned accessibility, available alternatives, and the extent of an episode in a single quantity.

It provides the three local decisions needed to generate that episode:

$$\begin{aligned} \Pr(X_0 = i) &= \pi_i a_i h_i / Z, \\ \Pr(\text{stop} \mid i) &= b_i / h_i, \\ \Pr(X_{k+1} = j \mid i, \text{continue}) &= P_{ij} h_j / (Ph)_i. \end{aligned} \tag{4}$$

Initiation combines the starting cue with access to compatible endings. Propagation favors successors through which those endings remain available. Completion compares ending here with the further paths the map supports. The same state can be an intermediate point in one event and its ending in another. Increasing support for a remote ending can therefore prolong an event even when local terminal evidence is unchanged.

These rules are the Doob h-transform (Doob, 1957). The joint probability of continuing to j is $\gamma P_{ij} h_j / h_i$. Multiplication along a path cancels its intermediate field ratios; initiation and completion

supply the remaining factors in Eq. (2). Local choices therefore recover the whole path distribution exactly. The operator $(I - \gamma P)^{-1}$ is the successor representation, which counts expected discounted visits (Dayan, 1993; Stachenfeld et al., 2017). Applied to terminal input, this predictive structure becomes a field for generating a particular episode.

3 A NEURAL CIRCUIT FOR CONDITIONAL SEQUENCE GENERATION

3.1 RECURRENT RELAXATION COMPUTES THE CONTINUATION FIELD

Within an attractor map, noise, adaptation, and recurrent interactions allow activity to move among stable population patterns (Pfeiffer & Foster, 2015; Ji et al., 2026). Their transition probabilities define P. The continuation field gives each represented situation influence according to the episodes that can unfold from it.

For diffusion in a potential U with noise variance $2D$ per unit time, the transform adds drift $2D\nabla \log h$. The effective potential becomes $U - 2D \log h$. Terminal input thus changes the flow through established states by propagating through their learned relations. The same calculation determines the stopping rate (Appendix A.5). This provides a physical interpretation of conditional generation within an attractor landscape.

We implement the field with the recurrent population dynamics of He et al. (2026b):

$$\tau_u \dot{r}_i = -r_i + g_s \sum_j S_{ij}\phi(r_j) - g_{\mathrm{inh}}\phi(r_i) + I_i^{\mathrm{bg}} + I_i^{\mathrm{EC}} + I_i^{\mathrm{goal}}. \tag{5}$$

Here r_i is activity, S_{ij} is excitation from j to i, and ϕ is a saturating response. Background and entorhinal currents establish the baseline; terminal evidence supplies the goal current. Recurrent excitation propagates this input, while leak and local inhibition attenuate it.

For a reversible map, reciprocal weights $S = \Pi^{1/2} P \Pi^{-1/2}$, with $\Pi = \mathrm{diag}(\pi)$, encode the transitions. Writing u_i for activity above baseline and ρ for the common response slope there, linearization gives

$$\tau_h \dot{u}_i = -u_i + \gamma \sum_j S_{ij} u_j + \beta_i, \qquad u_i^* = \sqrt{\pi_i}\, h_i, \quad \beta_i = \sqrt{\pi_i}\, b_i. \tag{6}$$

The effective parameters are $\gamma = g_s\rho/(1 + g_{\mathrm{inh}}\rho)$, $\tau_h = \tau_u/(1 + g_{\mathrm{inh}}\rho)$, and $\beta_i = I_i^{\mathrm{goal}}/(1 + g_{\mathrm{inh}}\rho)$. For $0 < \gamma < 1$, the response converges to the continuation field in neural coordinates. With one goal, its normalized shape equals the clamped navigation field derived by He et al. (2026b). Distributed current supplies evidence for several endings. Appendix A.6.1 gives the coordinate conversion, relaxation times, and nonlinear error.

The represented outcome is supplied before the episode unfolds. Relaxation carries its influence toward possible predecessors; subsequent sampling follows a path forward. Reciprocal connections support both operations. A maintained field population or input compartment preserves the alternatives while a changing assembly represents the current state.

3.2 DENDRITIC GAIN AND COMPETITION GENERATE A COMPLETE EVENT

Given active assembly i, candidate j receives recurrent input S_{ji} multiplied by its maintained field activity u_j. A completion population receives the terminal current associated with i. After field relaxation, their activation rates are

$$\lambda_{ij} = c_i \gamma S_{ji} u_j, \qquad \lambda_{i\dagger} = c_i \beta_i, \qquad \sum_j \lambda_{ij} + \lambda_{i\dagger} = c_i u_i, \tag{7}$$

where $\dagger$ denotes completion and $c_i > 0$ sets the clock. Independent exponential activation times select each alternative in proportion to its rate. Detailed balance gives $S_{ji}u_j = \sqrt{\pi_i} P_{ij} h_j$; the shared factor cancels in competition, yielding Eq. (4). Initiation uses rates proportional to $\sqrt{\pi_i} a_i u_i$. Each winner recruits inhibition that suppresses its competitors and resets the next decision.

CA3 recurrence and nonlinear dendritic integration provide substrates for the recurrent drive and its modulation (Guzman et al., 2016; Makara & Magee, 2013; Brandalise et al., 2016). Entorhinal

representations supply contextual and remote task information (Yamamoto & Tonegawa, 2017; Aery Jones et al., 2026). In the proposed completion pathway, CA1 combines current CA3 activity with terminal input and recruits inhibition to end the event. Figures 7 and 8 relate these operations to their candidate circuit organization.

In the linear regime, field relaxation and activation races implement the exact path law; the appendix bounds the error introduced by saturation. Common rescaling of the competing rates changes physical duration while preserving state order and stopping probabilities. Uniform amplification of terminal input likewise scales the field and completion rates together. The spatial pattern of input determines content, while common gain can regulate its speed. Ripple circuitry provides a setting for coordinating rapid expression through this clock (Buzsáki, 2015; Ecker et al., 2022).

3.3 Replay supplies the expectation needed for learning

Experience supplies the beginnings and endings that the circuit should generate. Replay samples those its present dynamics generate. Their frequency difference adjusts the boundary weights. With $\theta_i = \log a_i$, $\eta_j = \log b_j$, and learning rate α,

$$\mathbb{E}[\Delta\theta_i] = \alpha\big(\mu_0(i) - q_0(i)\big), \qquad \mathbb{E}[\Delta\eta_j] = \alpha\big(\mu_T(j) - q_T(j)\big). \tag{8}$$

An experienced endpoint increases its weight; a generated endpoint decreases it. More frequently generated endings receive more negative updates. Replay thereby estimates the competition among all possible episodes through the episodes actually expressed.

The rule ascends a concave objective whose optimum matches the endpoint distributions. Appendix A.7 gives the objective, convergence conditions, and its relation to iterative proportional fitting. Ripple-associated synaptic depression and inhibitory regulation during sleep provide candidate mechanisms for the replay-dependent decrease (Norimoto et al., 2018; Karaba et al., 2024).

4 Experiments

The theory connects four levels of explanation. A recurrent circuit computes a field, the field selects episodes from a learned reference, and the endpoints of those episodes supply an expectation for learning. Our experiments follow these connections. Experiment 1 asks whether local circuit operations produce the specified distribution over whole paths. Experiment 2 asks whether hippocampal population activity contains the temporal relations that boundary conditioning would reuse. Experiment 3 supplies behavioral boundaries and tests their contribution to recorded events. Experiment 4 asks whether the difference between experienced and generated endpoints predicts how sleep sequences change after experience. These comparisons move from exact calculations on known maps to population recordings, with a distinct observable at each step. Appendices C–F give the complete protocols, controls, and results.

4.1 Local neural operations generate complete paths

A terminal cue refers to the eventual outcome of an episode. The circuit has to make that information useful before the outcome is reached. We test this computation on eight maps, including a grid, a narrow passage between two regions, a line with shortcuts, and connected communities. Each map receives five families of initial and terminal evidence. Recurrent relaxation computes the continuation field, and successive local competitions choose either another state or completion. An exact finite-state calculation provides the probability law for comparison, including paths of every length and repeated visits (Figure 2).

Across five maps, the recurrent field has a mean relative error of 2.74×10^{-8}. With the analytical field, the local rules recover the complete path law in all 40 map–evidence combinations: absolute KL residuals are at most 5.62×10^{-11} nats. Sampling 520,000 episodes per condition gives a mean endpoint-joint total variation error of 0.00916.

The operation removals explain this agreement. Eliminating the field raises path divergence to 3.36 nats on the grid and 3.29 at the bottleneck. Keeping its transition bias while imposing constant stopping gives divergences of 0.182 and 0.226 nats. The same estimate of possible continuations therefore coordinates route choice and completion. In a compressed population representation,

broader endpoint evidence reduces path error in 23 of 24 matched maps at approximately 35% of full dimension. The representational demand depends on how precisely an episode is specified. Finally, learning from endpoint-frequency differences recovers the target coupling, with joint total variation errors of 0.000168–0.003246 after 9,000 updates on grid and modular maps.

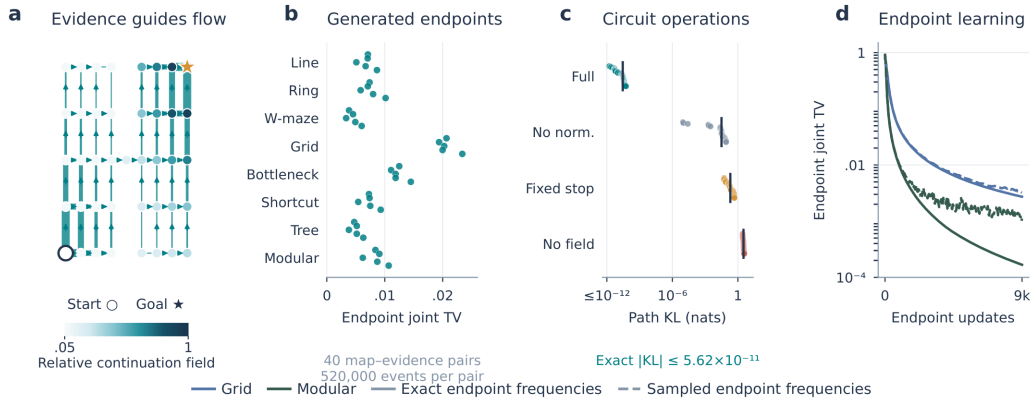

Figure 2: **Local neural operations generate complete paths and learn their boundaries.** (a) Continuation field and expected probability flow on a bottleneck map. (b) Sampled endpoint error across eight maps and five evidence families, with 520,000 events per condition. (c) Complete-path divergence for 40 intact conditions and 24 realizations per operation removal; bars mark medians and numerical residuals use a 10^{-12} display floor. (d) Endpoint learning on grid and modular maps using exact or sampled frequencies. Appendix C defines the conditions and errors.

4.2 Population sequences contain temporal relations that can be reused

For paths sharing a beginning and an ending, boundary factors cancel in their relative probabilities. The reference determines which intervening orders are favored. This property suggests a direct recording analysis: hold the endpoints and visited states fixed, and ask whether their observed order agrees with transitions learned from other events.

We identify recurring population states by clustering 20 ms spike-count vectors. HC-3 supplies repeated hippocampal recordings; HC-11 brackets novel-maze exploration with sleep; DANDI 000978 follows W-track learning (Mizuseki et al., 2014; Grosmark & Buzsáki, 2016; Shin & Jadhav, 2024). Training data define the state representation and transition matrix; later events test their predictions. The sequential readout updates its prediction after every state. The comparison readout uses the same map to predict expected visits from the initial state, then samples these visits independently. Across 27 HC-3 sessions, the sequential readout gains 0.0733 nats/state, with positive effects in 26 sessions. Thus the map's transition relations improve prediction beyond its start-conditioned visitation frequencies.

HC-11 tests reuse through PRE sleep, maze exploration, and POST sleep in eight sessions from four animals. The PRE reference improves over the future-visitation readout in all three epochs, by 0.0550, 0.0938, and 0.0541 nats/state. We then collapse consecutive repeats and compare each sequence with uniformly sampled legal orders preserving both endpoints and every state's count (Figure 3). The PRE reference favors observed task order by 0.00267 nats/transition in HC-11 and 0.00689 in DANDI. In DANDI, on the same collapsed sequences, prediction against the independent-visitation readout favors that readout by 0.1142 nats/state. These measurements locate the reusable information precisely: the fitted reference retains a preference among interior orders under shared boundaries, while overall predictive performance also depends on epoch and representation.

4.3 Behavioral boundaries organize represented content and event extent

We next ask how contextual evidence changes the use of a spatial map. Movement supplies place fields and transition probabilities in six HC-3 recordings from three rats and one HC-6 W-maze

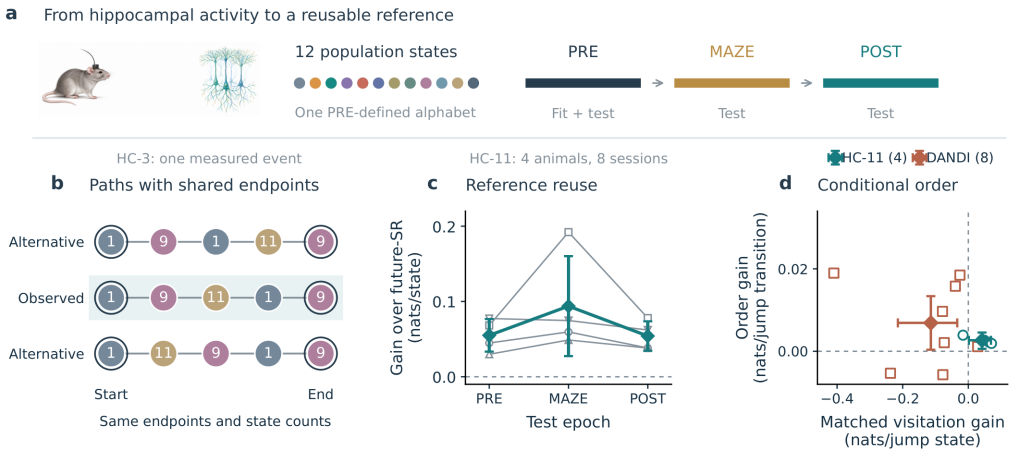

Figure 3: **A learned reference retains temporal structure across episodes.** (a) Recording and evaluation protocol (schematic): a PRE-defined alphabet is held fixed across PRE, MAZE, and POST. (b) Three legal orders of one held-out HC-3 event share endpoints and state counts; shading marks the observation. (c) HC-11 PRE-reference gains across epochs, showing four animals and their mean $\pm$ standard deviation. (d) Matched visitation gain and endpoint-conditioned order gain during running, paired within animal. Diamonds give equal-animal means; bars show marginal hierarchical 95% intervals. Appendix D defines the comparisons.

recording. Population bursts during immobility provide the events. Current location supplies the initial cue. A nearby traversal supplies its previous origin or subsequently observed destination as terminal evidence. Events without an eligible traversal receive uniform boundaries. This design evaluates represented episodes conditional on known behavioral circumstances.

The complete-event score includes the decoded sequence and its number of bins, so initiation, progression, and completion all contribute. The successor-value control receives the same continuation field and, at unit gain, chooses the same next state conditional on continuing. Its initiation and stopping rules differ. In 156 held-out HC-3 events, the bridge gains 0.0319 nats/bin over this control, with a hierarchical 95% interval of $[0.0019, 0.0842]$. In the 46 events with explicit behavioral boundaries, the gain is 0.1367 $[0.0516, 0.2371]$ (Figure 4).

How much of this effect concerns the extent of an episode? After calibrating reference survival to training-event lengths, removing field input lowers the HC-3 score by 0.2572 nats/bin; the duration term accounts for 91.2% of the difference. Here duration counts decoded 20 ms bins; this measurement fixes the observation clock and tests event extent. A complementary analysis fixes duration and sums over latent paths to predict the spikes themselves. Ripple-overlapping events yield a complete-path gain of 0.0239 $[-0.0076, 0.0673]$ and a conditional-spike gain of -0.0168 $[-0.0367, 0.0007]$ nats/bin. Complete-event probability and fixed-duration neural prediction thus resolve different consequences of conditioning. Spatial-resolution analyses and additional cohorts examine both targets in Appendix E.

4.4 Endpoint expectations predict changes after experience

The learning rule compares experienced endpoints with those generated by the current map. HC-11 provides sleep before and after novel-maze exploration; DANDI provides repeated sleep–behavior–sleep cycles during W-track learning. We fit endpoint weights to PRE events, subtract their predicted endpoint frequencies from waking frequencies, and use the resulting residual to adjust the weights while preserving the reference transitions.

A positive local score means that a small residual update improves POST likelihood more than held-out PRE likelihood. The score is positive in three of four HC-11 animals and seven of eight DANDI animals. Both endpoint populations contribute, and matching event lengths preserves the

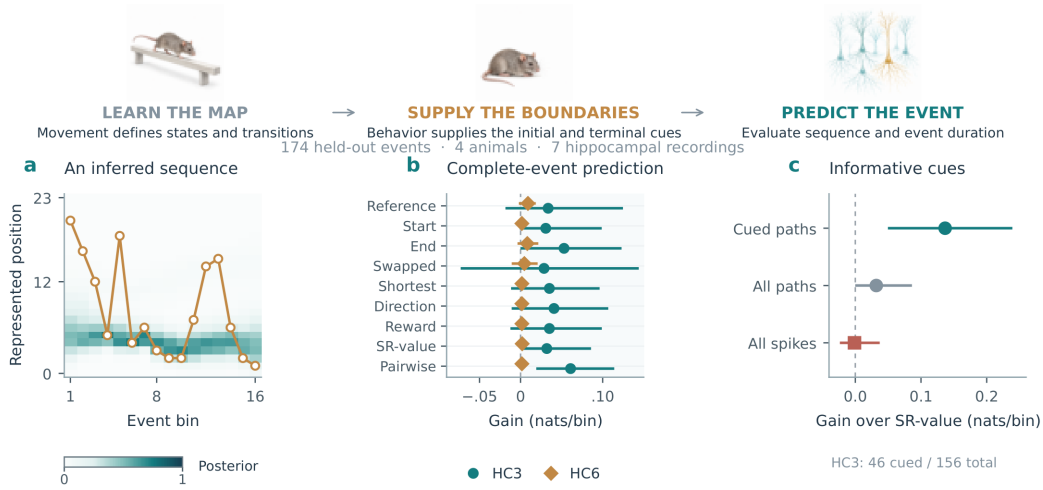

Figure 4: **Boundary information organizes complete hippocampal events.** The task strip links movement, behavioral cues, and event prediction. (a) A held-out HC-3 event: 73 spikes from 14 cells over 310 ms, selected by highest spike count among forward-cued test events. Shading shows the model posterior; the line is the separate binwise decode. (b) Complete-event gains over nine comparators. (c) HC-3 gains over SR-value for 46 cued paths, all 156 paths, and spikes at fixed duration. Bars show 95% intervals, hierarchical for HC-3 and event-bootstrap for HC-6. Appendix E gives selection and scoring details.

cohort effects. Selecting the step size from other animals gives mean held-out POST gains of 0.0359 nats/event in HC-11 and 0.00274 nats/burst in DANDI (Figure 5).

The components clarify the computation. Waking frequencies promote experienced endpoints; the model expectation reduces the influence of endpoints already frequent in preceding sleep. This second component supplies most of the DANDI effect, where alternative waking targets also predict positive change. Within HC-11 POST sleep, endpoint distributions continue to move without a monotonic approach to waking frequencies. The evidence links experience to a predictive redistribution of episodes and identifies the network's own endpoint expectation as a substantial contributor to that redistribution.

5 Discussion

Attractor dynamics recover a represented situation from partial input. Conditioning the reference path measure recovers possible episodes from partial information about their boundaries. The continuation field connects these scales of completion: represented endings propagate through the map and guide transitions among situations. A possible future influences present neural activity through the current representation of that future.

This account gives memory and imagination a shared operation. When the represented state retains episodic context, remembered boundaries select paths associated with a particular past experience. Broader cues permit combinations of transitions acquired on different occasions. An intended outcome makes these combinations available for planning and evaluation (Mattar & Daw, 2018; Jensen et al., 2024). The dependence of imagined scenes on the hippocampus (Hassabis et al., 2007) and the recombination of familiar elements during human planning sequences (He et al., 2026a) support this connection. Models of relational learning and generative memory explain how the represented situations acquire structure and content (Whittington et al., 2020; Raju et al., 2024; Spens & Burgess, 2024); the bridge specifies how their succession forms a complete episode.

The distinction between knowledge and current evidence has an observable consequence. Changing boundary conditions redistributes endpoint pairs while preserving relative route and length probabilities within each pair. Changing the map can alter those conditional probabilities. Replay that adapts to relocated barriers despite stable place fields illustrates why the identity of represented

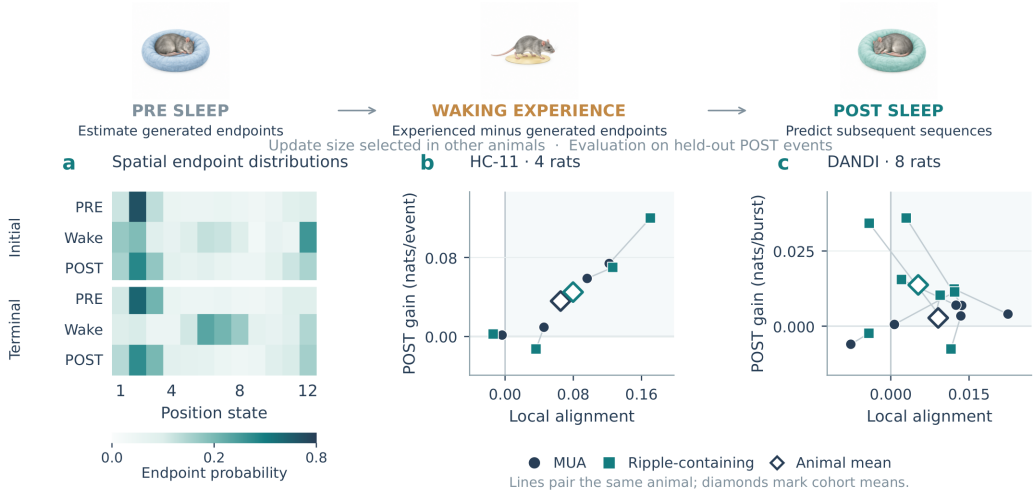

Figure 5: **Endpoint expectations predict changes from PRE to POST sleep.** The task strip shows sleep, experience, and subsequent sleep. (a) Initial and terminal distributions in one HC-11 session, on a square-root probability scale. (b,c) Local alignment and finite-update POST gains for every animal. Lines pair burst and ripple-containing estimates; open diamonds give equal-animal means. Other animals select update sizes. Local alignment is the POST-minus-PRE gain per unit update; finite gains use nats/event for HC-11 and nats/burst for DANDI (Appendix F).

places and their temporal relations warrant separate measurements (Widloski & Foster, 2022). A changed goal acts through boundary evidence; a changed barrier acts through the available transitions. These manipulations can produce similar changes in average replay direction while making different predictions after the endpoints have been matched.

Completion is part of the represented experience. At the same location, activity can finish one episode and continue another because the accessible endings differ. Increasing support for a remote ending reduces the probability of stopping earlier along its possible routes. The theory links this change to the modulation of successor choice through one measurable field (Appendix A.9). Human neurons that respond to event boundaries provide a setting in which to examine such dependence on the surrounding episode (Zheng et al., 2022). The physiological clock supplies a further distinction: scaling all competing activation rates changes an episode's speed while preserving its sequence law. Replay across several ripples, at behavioral speeds, or without detectable ripples motivates measuring represented extent and neural timing separately (Davidson et al., 2009; Denovellis et al., 2021; Widloski & Foster, 2025).

During sleep, internally supplied cues can organize episodes through the same dynamics. Pre-existing sequences can acquire spatial meaning when their assemblies become associated with a new environment (Dragoi & Tonegawa, 2011; Zhou et al., 2025). Sleep substates that favor recent or older memories provide settings for changing the evidence that guides generation (Chang et al., 2025). This also suggests how internally supplied boundaries could organize dream episodes with associated sensory content, a hypothesis motivated by reduced episodic detail in dreams after hippocampal damage (Spanò et al., 2020).

Replay also exposes the circuit's own expectations. A frequently generated ending contributes frequently to the negative term in learning; experience supplies the positive term. Their difference adjusts the relative availability of future episodes. Ripple-associated synaptic depression and inhibitory regulation of sleep reactivation provide candidate substrates for this comparison (Norimoto et al., 2018; Karaba et al., 2024). The observed redistribution of sleep endpoints supports this computational role, with contributions that vary across recording cohorts. At agreement with experienced frequencies, the expected update vanishes while sequence generation continues. Replay can therefore remain active after calibration: its samples continue to express the map, support downstream learning, and reveal the expectations that later experience will revise.

REPRODUCIBILITY STATEMENT

The accompanying supplementary material provides simulation and analysis code, experiment configurations, numerical outputs, model and event caches, and figure-generation scripts. Readers should first consult the bilingual reports for Experiments 1–4 as the primary supplementary document, accessible via `supplementary_reports/README.md`. Reviewers can begin with `README.md` for environment setup and follow each experiment's `README_REVIEWER.md` for the inputs, execution steps, and outputs. `DATA_ACCESS.md` identifies the public datasets, session selections, and data-access procedures. Appendices A–F present the mathematical derivations, data definitions, event-selection rules, model comparisons, parameter settings, and experimental protocols. These materials guide reviewers through the simulations, recording-based analyses, and reconstruction of the manuscript figures.

AI USE STATEMENT

Generative AI tools were used to assist with writing and language polishing, retrieve and discover related literature, and support code execution. Image-generation tools were used to create schematic illustrations. The authors take responsibility for the final text, references, analyses, and figures.

REFERENCES

Emily A. Aery Jones, Isabel I. C. Low, Frances S. Cho, and Lisa M. Giocomo. Entorhinal cortex represents task-relevant remote locations independently of CA1. *Nature Neuroscience*, 29(5): 1181–1190, 2026. doi: 10.1038/s41593-026-02232-0. URL https://www.nature.com/articles/s41593-026-02232-0.

R. Ellen Ambrose, Brad E. Pfeiffer, and David J. Foster. Reverse replay of hippocampal place cells is uniquely modulated by changing reward. *Neuron*, 91(5):1124–1136, 2016. doi: 10.1016/j.neuron.2016.07.047.

Jacob J. W. Bakermans, Joseph Warren, James C. R. Whittington, and Timothy E. J. Behrens. Constructing future behavior in the hippocampal formation through composition and replay. *Nature Neuroscience*, 28(5):1061–1072, 2025. doi: 10.1038/s41593-025-01908-3. URL https://www.nature.com/articles/s41593-025-01908-3.

Pietro Berkes, Gergő Orbán, Máté Lengyel, and József Fiser. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. *Science*, 331(6013):83–87, 2011. doi: 10.1126/science.1195870. URL https://doi.org/10.1126/science.1195870.

Katie C. Bittner, Aaron D. Milstein, Christine Grienberger, Sandro Romani, and Jeffrey C. Magee. Behavioral time scale synaptic plasticity underlies CA1 place fields. *Science*, 357(6355):1033–1036, 2017. doi: 10.1126/science.aan3846.

Federico Brandalise, Stefano Carta, Fritjof Helmchen, John Lisman, and Urs Gerber. Dendritic NMDA spikes are necessary for timing-dependent associative LTP in CA3 pyramidal cells. *Nature Communications*, 7:13480, 2016. doi: 10.1038/ncomms13480.

György Buzsáki. Hippocampal sharp wave-ripple: A cognitive biomarker for episodic memory and planning. *Hippocampus*, 25(10):1073–1188, 2015. doi: 10.1002/hipo.22488.

Josue Casco-Rodriguez, Nanda H. Krishna, and Richard G. Baraniuk. Leakage and second-order dynamics improve hippocampal RNN replay. *arXiv preprint arXiv:2602.18401*, 2026. doi: 10.48550/arXiv.2602.18401. URL https://arxiv.org/abs/2602.18401.

Sarthak Chandra, Sugandha Sharma, Rishidev Chaudhuri, and Ila Fiete. Episodic and associative memory from spatial scaffolds in the hippocampus. *Nature*, 638(8051):739–751, 2025. doi: 10.1038/s41586-024-08392-y. URL https://www.nature.com/articles/s41586-024-08392-y.

Hongyu Chang, Wenbo Tang, Annabella M. Wulf, Thokozile Nyasulu, Madison E. Wolf, Antonio Fernandez-Ruiz, and Azahara Oliva. Sleep microstructure organizes memory replay. *Nature*, 637 (8048):1161–1169, 2025. doi: 10.1038/s41586-024-08340-w. URL https://www.nature.com/articles/s41586-024-08340-w.

Yongxin Chen, Tryphon T. Georgiou, and Michele Pavon. The most likely evolution of diffusing and vanishing particles: Schrödinger bridges with unbalanced marginals. *SIAM Journal on Control and Optimization*, 60(4):2016–2039, 2022. doi: 10.1137/21M1447672. URL https://epubs.siam.org/doi/10.1137/21M1447672.

Yusi Chen, Huanqiu Zhang, Mia Cameron, and Terrence Sejnowski. Predictive sequence learning in the hippocampal formation. *Neuron*, 112(15):2645–2658.e4, 2024. doi: 10.1016/j.neuron.2024.05.024.

Imre Csiszár. I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, 3(1):146–158, 1975. doi: 10.1214/aop/1176996454. URL https://doi.org/10.1214/aop/1176996454.

Thomas J. Davidson, Fabian Kloosterman, and Matthew A. Wilson. Hippocampal replay of extended experience. *Neuron*, 63(4):497–507, 2009. doi: 10.1016/j.neuron.2009.07.027.

Peter Dayan. Improving generalization for temporal difference learning: The successor representation. *Neural Computation*, 5(4):613–624, 1993. doi: 10.1162/neco.1993.5.4.613.

Valentin De Bortoli, James Thornton, Jeremy Heng, and Arnaud Doucet. Diffusion Schrödinger bridge with applications to score-based generative modeling. In *Advances in Neural Information Processing Systems*, volume 34, pp. 17695–17709, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/940392f5f32a7ade1cc201767cf83e31-Abstract.html.

Luisa de Vivo, Michele Bellesi, William Marshall, Eric A. Bushong, Mark H. Ellisman, Giulio Tononi, and Chiara Cirelli. Ultrastructural evidence for synaptic scaling across the wake/sleep cycle. *Science*, 355(6324):507–510, 2017. doi: 10.1126/science.aah5982.

Eric L. Denovellis, Anna K. Gillespie, Michael E. Coulter, Marielena Sosa, Jason E. Chung, Uri T. Eden, and Loren M. Frank. Hippocampal replay of experience at real-world speeds. *eLife*, 10: e64505, 2021. doi: 10.7554/eLife.64505.

Kamran Diba and György Buzsáki. Forward and reverse hippocampal place-cell sequences during ripples. *Nature Neuroscience*, 10:1241–1242, 2007. doi: 10.1038/nn1961.

Nicolas Diekmann and Sen Cheng. A model of hippocampal replay driven by experience and environmental structure facilitates spatial learning. *eLife*, 12:e82301, 2023. doi: 10.7554/eLife.82301.

Graham H. Diering, Raja S. Nirujogi, Richard H. Roth, Paul F. Worley, Akhilesh Pandey, and Richard L. Huganir. Homer1a drives homeostatic scaling-down of excitatory synapses during sleep. *Science*, 355(6324):511–515, 2017. doi: 10.1126/science.aai8355.

Joseph L. Doob. Conditional Brownian motion and the boundary limits of harmonic functions. *Bulletin de la Société Mathématique de France*, 85:431–458, 1957. doi: 10.24033/bsmf.1494.

George Dragoi and Susumu Tonegawa. Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature*, 469:397–401, 2011. doi: 10.1038/nature09633.

András Ecker, Bence Bagi, Eszter Vértes, Orsolya Steinbach-Németh, Mária R. Karlócai, Orsolya I. Papp, István Miklós, Norbert Hájos, Tamás F. Freund, Attila I. Gulyás, and Szabolcs Káli. Hippocampal sharp wave-ripples and the associated sequence replay emerge from structured synaptic interactions in a network model of area CA3. *eLife*, 11:e71850, 2022. doi: 10.7554/eLife.71850.

Asmaa Eldesoukey and Tryphon T. Georgiou. Schrödinger's control and estimation paradigm with spatio-temporal distributions on graphs, 2024. URL https://arxiv.org/abs/2312.05679. Version 2.

David J. Foster and Matthew A. Wilson. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440:680–683, 2006. doi: 10.1038/nature04587.

Dileep George, Rajeev V. Rikhye, Nishad Gothoskar, J. Swaroop Guntupalli, Antoine Dedieu, and Miguel Lázaro-Gredilla. Clone-structured graph representations enable flexible learning and vicarious evaluation of cognitive maps. *Nature Communications*, 12:2392, 2021. doi: 10.1038/s41467-021-22559-5.

Gabrielle Girardeau, Karim Benchenane, Sidney I. Wiener, György Buzsáki, and Michaël B. Zugaro. Selective suppression of hippocampal ripples impairs spatial memory. *Nature Neuroscience*, 12: 1222–1223, 2009. doi: 10.1038/nn.2384.

Joaquin Gonzalez and Mihaly Voroslakos. High-density extracellular recordings (Neuropixels/SiNAPS) of hippocampal-cortical dynamics during spatial behavior. DANDI Archive, 2026. URL https://dandiarchive.org/dandiset/001695/0.260319.2023. Data set DANDI:001695, version 0.260319.2023.

Andres D. Grosmark and György Buzsáki. Diversity in neural firing dynamics supports both rigid and learned hippocampal sequences. *Science*, 351(6280):1440–1443, 2016. doi: 10.1126/science.aad1935.

Andres D. Grosmark, J. Long, and György Buzsáki. Recordings from hippocampal area CA1, PRE, during and POST novel spatial learning. CRCNS.org, 2016. URL https://doi.org/10.6080/K0862DC5. Data set HC-11.

Anoopum S. Gupta, Matthijs A. A. van der Meer, David S. Touretzky, and A. David Redish. Hippocampal replay is not a simple function of experience. *Neuron*, 65(5):695–705, 2010. doi: 10.1016/j.neuron.2010.01.034.

Segundo Jose Guzman, Alois Schlögl, Michael Frotscher, and Peter Jonas. Synaptic mechanisms of pattern completion in the hippocampal CA3 network. *Science*, 353(6304):1117–1123, 2016. doi: 10.1126/science.aaf1836.

Demis Hassabis, Dharshan Kumaran, Seralynne D. Vann, and Eleanor A. Maguire. Patients with hippocampal amnesia cannot imagine new experiences. *Proceedings of the National Academy of Sciences*, 104(5):1726–1731, 2007. doi: 10.1073/pnas.0610561104. URL https://doi.org/10.1073/pnas.0610561104.

Li He, Xiongfei Wang, Jinbo Zhang, Zhibing Xiao, Xiangyu Hu, Philipp Schwartenbeck, Jacob Bakermans, Tim Behrens, and Yunzhe Liu. Human hippocampal ripples coordinate planning sequences and compositional representations in neocortex. *Nature Neuroscience*, 29(7):1711–1721, 2026a. doi: 10.1038/s41593-026-02291-3. URL https://www.nature.com/articles/s41593-026-02291-3.

Yuhang He, Junfeng Zuo, Tianhao Chu, and Si Wu. Planning as dynamics relaxation: Hippocampal recurrent network realizes optimal goal-directed navigation, 2026b. URL https://arxiv.org/abs/2609.13219.

John J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982. doi: 10.1073/pnas.79.8.2554.

Hiroshi T. Ito, Sheng-Jia Zhang, Menno P. Witter, Edvard I. Moser, and May-Britt Moser. A prefrontal–thalamo–hippocampal circuit for goal-directed spatial navigation. *Nature*, 522(7554): 50–55, 2015. doi: 10.1038/nature14396. URL https://www.nature.com/articles/nature14396.

Kristopher T. Jensen, Guillaume Hennequin, and Marcelo G. Mattar. A recurrent network model of planning explains hippocampal replay and human behavior. *Nature Neuroscience*, 27(7): 1340–1348, 2024. doi: 10.1038/s41593-024-01675-7. URL https://www.nature.com/articles/s41593-024-01675-7.

Zilong Ji, Tianhao Chu, Xingsi Dong, Changmin Yu, Daniel Bush, Neil Burgess, and Si Wu. Dynamical modulation of hippocampal replay through firing rate adaptation. *Nature Communications*, 17: 1282, 2026. doi: 10.1038/s41467-025-68042-3.

Adam Johnson and A. David Redish. Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *Journal of Neuroscience*, 27(45):12176–12189, 2007. doi: 10.1523/JNEUROSCI.3761-07.2007. URL https://www.jneurosci.org/content/27/45/12176.

Lindsay A. Karaba, Heath L. Robinson, Ryan E. Harvey, Weiwei Chen, Antonio Fernandez-Ruiz, and Azahara Oliva. A hippocampal circuit mechanism to balance memory reactivation during sleep. *Science*, 385(6710):738–743, 2024. doi: 10.1126/science.ado5708. URL https://www.science.org/doi/10.1126/science.ado5708.

Mattias Karlsson, Margaret Carr, and Loren M. Frank. Simultaneous extracellular recordings from hippocampal areas CA1 and CA3 (or MEC and CA1) from rats performing an alternation task in two W-shaped tracks that are geometrically identically but visually distinct. CRCNS.org, HC-6 dataset, 2015. URL https://doi.org/10.6080/K0NK3BZJ.

Beomsu Kim, Gihyun Kwon, Kwanyoung Kim, and Jong Chul Ye. Unpaired image-to-image translation via neural Schrödinger bridge. In *International Conference on Learning Representations*, 2024. URL https://arxiv.org/abs/2305.15086.

Noémi Kis, Balázs Lükő, Judit Herédi, Ádám Magó, Bela Erlinghagen, Mahboubeh Ahmadi, Snezana Raus Balind, Mátyás Irás, Balázs B. Ujfalussy, and Judit K. Makara. Cholinergic regulation of dendritic Ca^{2+} spikes controls firing mode of hippocampal CA3 pyramidal neurons. *Proceedings of the National Academy of Sciences*, 121(46):e2321501121, 2024. doi: 10.1073/pnas.2321501121.

Chongxi Lai, Shinsuke Tanaka, Timothy D. Harris, and Albert K. Lee. Volitional activation of remote place representations with a hippocampal brain–machine interface. *Science*, 382(6670):566–573, 2023. doi: 10.1126/science.adh5206. URL https://www.science.org/doi/10.1126/science.adh5206.

Christian Léonard. A survey of the Schrödinger problem and some of its connections with optimal transport. *Discrete and Continuous Dynamical Systems*, 34(4):1533–1574, 2014. doi: 10.3934/dcds.2014.34.1533.

Hui Lin, Yukun Yang, Rong Zhao, Giovanni Pezzulo, and Wolfgang Maass. Neural sampling from cognitive maps enables goal-directed imagination and planning. *Nature Machine Intelligence*, 8: 1045–1065, 2026. doi: 10.1038/s42256-026-01254-4.

Guan-Horng Liu, Arash Vahdat, De-An Huang, Evangelos Theodorou, Weili Nie, and Anima Anandkumar. I^2SB: Image-to-image Schrödinger bridge. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 22042–22062. PMLR, 2023. URL https://proceedings.mlr.press/v202/liu23ai.html.

Yunzhe Liu, Raymond J. Dolan, Zeb Kurth-Nelson, and Timothy E. J. Behrens. Human replay spontaneously reorganizes experience. *Cell*, 178(3):640–652.e14, 2019. doi: 10.1016/j.cell.2019.06.012. URL https://doi.org/10.1016/j.cell.2019.06.012.

Judit K. Makara and Jeffrey C. Magee. Variable dendritic integration in hippocampal CA3 pyramidal neurons. *Neuron*, 80(6):1438–1450, 2013. doi: 10.1016/j.neuron.2013.10.033.

Caitlin S. Mallory, John Widloski, and David J. Foster. The time course and organization of hippocampal replay. *Science*, 387(6733):541–548, 2025. doi: 10.1126/science.ads4760. URL https://www.science.org/doi/10.1126/science.ads4760.

Marcelo G. Mattar and Nathaniel D. Daw. Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, 21:1609–1617, 2018. doi: 10.1038/s41593-018-0232-z.

Daniel C. McNamee, Kimberly L. Stachenfeld, Matthew M. Botvinick, and Samuel J. Gershman. Flexible modulation of sequence generation in the entorhinal–hippocampal system. *Nature Neuroscience*, 24:851–862, 2021. doi: 10.1038/s41593-021-00831-7.

Kenji Mizuseki, Anton Sirota, Eva Pastalkova, Kamran Diba, and György Buzsáki. Multiple single unit recordings from different rat hippocampal and entorhinal regions while the animals were performing multiple behavioral tasks. CRCNS.org, 2013. URL https://doi.org/10.6080/K09G5JRZ. Data set HC-3.

Kenji Mizuseki, Kamran Diba, Eva Pastalkova, Jeff Teeters, Anton Sirota, and György Buzsáki. Neurosharing: large-scale data sets (spike, LFP) recorded from the hippocampal-entorhinal system in behaving rats. *F1000Research*, 3:98, 2014. doi: 10.12688/f1000research.3895.2. Version 2.

Joshua P. Neunuebel and James J. Knierim. CA3 retrieves coherent representations from degraded input: Direct evidence for CA3 pattern completion and dentate gyrus pattern separation. *Neuron*, 81(2):416–427, 2014. doi: 10.1016/j.neuron.2013.11.017.

Hiroaki Norimoto, Kenichi Makino, Mengxuan Gao, Yu Shikano, Kazuki Okamoto, Tomoe Ishikawa, Takuya Sasaki, Hiroyuki Hioki, Shigeyoshi Fujisawa, and Yuji Ikegaya. Hippocampal ripples down-regulate synapses. *Science*, 359(6383):1524–1527, 2018. doi: 10.1126/science.aao0702. URL https://www.science.org/doi/10.1126/science.aao0702.

John O'Keefe and Lynn Nadel. *The Hippocampus as a Cognitive Map*. Clarendon Press, Oxford, 1978. ISBN 9780198572060.

Brad E. Pfeiffer and David J. Foster. Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497:74–79, 2013. doi: 10.1038/nature12112.

Brad E. Pfeiffer and David J. Foster. Autoassociative dynamics in the generation of sequences of hippocampal place cells. *Science*, 349(6244):180–183, 2015. doi: 10.1126/science.aaa9633. URL https://www.science.org/doi/10.1126/science.aaa9633.

Payam Piray and Nathaniel D. Daw. Linear reinforcement learning in planning, grid fields, and cognitive control. *Nature Communications*, 12:4942, 2021. doi: 10.1038/s41467-021-25123-3.

Payam Piray and Nathaniel D. Daw. Reconciling flexibility and efficiency: Medial entorhinal cortex represents a compositional cognitive map. *Nature Communications*, 16:7444, 2025. doi: 10.1038/s41467-025-62733-7.

Rajkumar Vasudeva Raju, J. Swaroop Guntupalli, Guangyao Zhou, Carter Wendelken, Miguel Lázaro-Gredilla, and Dileep George. Space is a latent sequence: A theory of the hippocampus. *Science Advances*, 10(31):eadm8470, 2024. doi: 10.1126/sciadv.adm8470.

P. Dylan Rich, Hua-Peng Liaw, and Albert K. Lee. Large environments reveal the statistical structure governing hippocampal representations. *Science*, 345(6198):814–817, 2014. doi: 10.1126/science.1255635. URL https://doi.org/10.1126/science.1255635.

P. Dylan Rich, Hua-Peng Liaw, and Albert K. Lee. Extracellular CA1 activity from rats during exploration of a novel 48m track. CRCNS.org, 2022. URL https://doi.org/10.6080/K0BV7DT2.

Herbert Robbins and David Siegmund. A convergence theorem for non negative almost supermartingales and some applications. In Jagdish S. Rustagi (ed.), *Optimizing Methods in Statistics*, pp. 233–257. Academic Press, 1971. doi: 10.1016/B978-0-12-604550-5.50015-8. URL https://www.sciencedirect.com/science/article/abs/pii/B9780126045505500158.

Sandro Romani and Misha Tsodyks. Short-term plasticity based network model of place cells dynamics. *Hippocampus*, 25(1):94–105, 2015. doi: 10.1002/hipo.22355. URL https://onlinelibrary.wiley.com/doi/full/10.1002/hipo.22355.

Erwin Schrödinger. Über die Umkehrung der Naturgesetze. *Sitzungsberichte der Preussischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse*, pp. 144–153, 1931. URL https://doi.org/10.1140/epjh/s13129-021-00032-7. English translation: Chetrite, Muratore-Ginanneschi and Schwieger, European Physical Journal H 46, 28 (2021).

Nicolas W. Schuck and Yael Niv. Sequential replay of nonspatial task states in the human hippocampus. *Science*, 364(6447):eaaw5181, 2019. doi: 10.1126/science.aaw5181. URL https://doi.org/10.1126/science.aaw5181.

Yuyang Shi, Valentin De Bortoli, Andrew Campbell, and Arnaud Doucet. Diffusion Schrödinger bridge matching. In *Advances in Neural Information Processing Systems*, volume 36, pp. 62183–62223. Curran Associates, Inc., 2023. doi: 10.52202/075280-2717. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/c428adf74782c2092d254329b6b02482-Abstract-Conference.html.

Justin Shin. Odor-place spatial association task. DANDI Archive, 2025. URL https://dandiarchive.org/dandiset/001539/0.250815.1203. Data set DANDI:001539, version 0.250815.1203.

Justin Shin and Shantanu Jadhav. Single day W-track learning. DANDI Archive, 2024. URL https://doi.org/10.48324/dandi.000978/0.240511.0307. Data set DANDI:000978, version 0.240511.0307.

Justin Shin and Shantanu P. Jadhav. Novel-familiar-novel WTrack (CA1-PFC). DANDI Archive, 2025. URL https://dandiarchive.org/dandiset/000447/0.250724.2224. Data set DANDI:000447, version 0.250724.2224.

Justin D. Shin, Wenbo Tang, and Shantanu P. Jadhav. Dynamics of awake hippocampal-prefrontal replay for spatial learning and memory-guided decision making. *Neuron*, 104(6):1110–1125.e7, 2019. doi: 10.1016/j.neuron.2019.09.012. URL https://doi.org/10.1016/j.neuron.2019.09.012.

Delia Silva, Ting Feng, and David J. Foster. Trajectory events across hippocampal place cells require previous experience. *Nature Neuroscience*, 18(12):1772–1779, 2015. doi: 10.1038/nn.4151. URL https://www.nature.com/articles/nn.4151.

Richard Sinkhorn and Paul Knopp. Concerning nonnegative matrices and doubly stochastic matrices. *Pacific Journal of Mathematics*, 21(2):343–348, 1967. doi: 10.2140/pjm.1967.21.343. URL https://msp.org/pjm/1967/21-2/p14.xhtml.

Vignesh Ram Somnath, Matteo Pariset, Ya-Ping Hsieh, Maria Rodriguez Martinez, Andreas Krause, and Charlotte Bunne. Aligned diffusion Schrödinger bridges. In *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pp. 1985–1995. PMLR, 2023. URL https://proceedings.mlr.press/v216/somnath23a.html.

Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=PxTIG12RRHS.

Goffredina Spanò, Gloria Pizzamiglio, Cornelia McCormick, Ian A. Clark, Sara De Felice, Thomas D. Miller, Jamie O. Edgin, Clive R. Rosenthal, and Eleanor A. Maguire. Dreaming with hippocampal damage. *eLife*, 9:e56211, 2020. doi: 10.7554/eLife.56211. URL https://elifesciences.org/articles/56211.

Eleanor Spens and Neil Burgess. A generative model of memory construction and consolidation. *Nature Human Behaviour*, 8:526–543, 2024. doi: 10.1038/s41562-023-01799-z.

Kimberly L. Stachenfeld, Matthew M. Botvinick, and Samuel J. Gershman. The hippocampus as a predictive map. *Nature Neuroscience*, 20:1643–1653, 2017. doi: 10.1038/nn.4650.

Federico Stella, Peter Baracskay, Joseph O'Neill, and Jozsef Csicsvari. Hippocampal reactivation of random trajectories resembling Brownian diffusion. *Neuron*, 102(2):450–461.e7, 2019. doi: 10.1016/j.neuron.2019.01.052.

Haowen Su, Xian Li, Savannah Born, Christopher J. Honey, Janice Chen, and Hongmi Lee. Neural dynamics of spontaneous memory recall and future thinking in the continuous flow of thoughts. *Nature Communications*, 16:6433, 2025. doi: 10.1038/s41467-025-61807-w. URL https://doi.org/10.1038/s41467-025-61807-w.

Xuan Su, Jiaming Song, Chenlin Meng, and Stefano Ermon. Dual diffusion implicit bridges for image-to-image translation. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2203.08382.

Claire A. Symanski, John H. Bladon, Emi T. Kullberg, Paul Miller, and Shantanu P. Jadhav. Rhythmic coordination and ensemble dynamics in the hippocampal-prefrontal network during odor-place associative memory and decision making. *eLife*, 11:e79545, 2022. doi: 10.7554/eLife.79545. URL https://elifesciences.org/articles/79545.

Wenbo Tang, Justin D. Shin, and Shantanu P. Jadhav. Geometric transformation of cognitive maps for generalization across hippocampal-prefrontal circuits. *Cell Reports*, 42(3):112246, 2023. doi: 10.1016/j.celrep.2023.112246. URL https://doi.org/10.1016/j.celrep.2023.112246.

Emanuel Todorov. Efficient computation of optimal actions. *Proceedings of the National Academy of Sciences*, 106(28):11478–11483, 2009. doi: 10.1073/pnas.0710743106.

Edward C. Tolman. Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208, 1948. doi: 10.1037/h0061626.

Balázs B. Ujfalussy and Gergő Orbán. Sampling motion trajectories during hippocampal theta sequences. *eLife*, 11:e74058, 2022. doi: 10.7554/eLife.74058. URL https://elifesciences.org/articles/74058.

Abraham Z. Vollan, Richard J. Gardner, May-Britt Moser, and Edvard I. Moser. Left–right-alternating theta sweeps in entorhinal–hippocampal maps of space. *Nature*, 639:995–1005, 2025. doi: 10.1038/s41586-024-08527-1. URL https://www.nature.com/articles/s41586-024-08527-1.

James C. R. Whittington, Timothy H. Muller, Shirley Mark, Guifen Chen, Caswell Barry, Neil Burgess, and Timothy E. J. Behrens. The Tolman–Eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell*, 183(5):1249–1263.e23, 2020. doi: 10.1016/j.cell.2020.10.024.

John Widloski and David J. Foster. Flexible rerouting of hippocampal replay sequences around changing barriers in the absence of global place field remapping. *Neuron*, 110(9):1547–1558.e8, 2022. doi: 10.1016/j.neuron.2022.02.002. URL https://doi.org/10.1016/j.neuron.2022.02.002.

John Widloski and David J. Foster. Replay without sharp wave ripples in a spatial memory task. *Nature Communications*, 16:10287, 2025. doi: 10.1038/s41467-025-65181-5. URL https://www.nature.com/articles/s41467-025-65181-5.

Andrew M. Wikenheiser and A. David Redish. Hippocampal theta sequences reflect current goals. *Nature Neuroscience*, 18(2):289–294, 2015. doi: 10.1038/nn.3909. URL https://www.nature.com/articles/nn.3909.

Tom J. Wills, Colin Lever, Francesca Cacucci, Neil Burgess, and John O'Keefe. Attractor dynamics in the hippocampal representation of the local environment. *Science*, 308(5723):873–876, 2005. doi: 10.1126/science.1108905.

Jun Yamamoto and Susumu Tonegawa. Direct medial entorhinal cortex input to hippocampal CA1 is crucial for extended quiet awake replay. *Neuron*, 96(1):217–227.e4, 2017. doi: 10.1016/j.neuron.2017.09.017.

Jie Zheng, Andrea G. P. Schjetnan, Mar Yebra, Bernard A. Gomes, Clayton P. Mosher, Suneil K. Kalia, Taufik A. Valiante, Adam N. Mamelak, Gabriel Kreiman, and Ueli Rutishauser. Neurons detect cognitive boundaries to structure episodic memories in humans. *Nature Neuroscience*, 25(3):358–368, 2022. doi: 10.1038/s41593-022-01020-w. URL https://www.nature.com/articles/s41593-022-01020-w.

Yuchen Zhou, Jeremie Sibille, and George Dragoi. Generative emergence of non-local representations in the hippocampus. *Nature Communications*, 16:8012, 2025. doi: 10.1038/s41467-025-63346-w. URL https://www.nature.com/articles/s41467-025-63346-w.

Zhenglong Zhou, Michael J. Kahana, and Anna C. Schapiro. A unifying account of replay as context-driven memory reactivation. *eLife*, 13:RP99931, 2026. doi: 10.7554/eLife.99931. URL https://elifesciences.org/articles/99931.

Ipshita Zutshi and György Buzsáki. Hippocampal sharp-wave ripples and their spike assembly content are regulated by the medial entorhinal cortex. *Current Biology*, 33(17):3648–3659.e4, 2023. doi: 10.1016/j.cub.2023.07.039.

Part I

Theoretical details

The derivation follows the computation of an episode. A learned map first defines a probability law over paths and their points of completion. Endpoint information selects a new ensemble, a continuation field makes that selection local, and neural dynamics compute and sample the field. Learning then adjusts the endpoint factors using experience and replay.

Reading guide. Reference paths and the bridge problem establish the statistical model. Local sampling, the control interpretation, and event statistics give the path probabilities and their consequences. Attractor dynamics connects the construction to population activity. Recurrent field computation, including the developing response and arrival and completion at a goal, and neural competition give the circuit derivation. Endpoint learning gives the objective, updates, and convergence proof. Approximation bounds, quantitative predictions, and hippocampal computations develop its further implications. Methods and results for Experiments 1–4 appear in Part II (p. 48).

CONTENTS

A The path model and its consequences **19**

A.1 Reference dynamics and random sequence length 19

A.2 Endpoint evidence and the Schrödinger bridge 20

A.2.1 Prescribed endpoint distributions 21

A.3 The local sampling rule 22

A.4 Occupancy, probability flow, and sequence length 24

A.5 Attractor maps and changes in replay direction 26

A.6 Computing and sampling the field in a circuit 29

A.6.1 From recurrent neural dynamics to the continuation field 29

A.6.2 Dendritic gain, competition, and completion 34

A.7 Learning the endpoint factors from wake and replay 36

A.8 Approximation errors and extensions 37

A.9 Quantitative predictions from the same construction 38

A.10 Interpretation across hippocampal computations 40

B Cognitive maps, sequence generation, and hippocampal function **42**

B.1 From represented situations to possible experience 42

B.2 Constructing and completing an episode 42

B.3 Selection, direction, and the timing of replay 44

B.4 Circuit operations that link local choices to complete sequences 45

B.5 How sleep changes future replay 46

B.6 Schrödinger bridges, control, and generative models 46

A THE PATH MODEL AND ITS CONSEQUENCES

Knowledge of a starting point and a destination leaves many possible journeys. The map supplies their relative probabilities through the transitions learned from experience. The present task supplies information about their beginnings and endings. The theory describes how these two sources of information jointly determine an ensemble of complete episodes, and how a neural circuit draws individual episodes from that ensemble.

A path records the order of represented states and the point of completion. Its number of transitions is random. The reference process assigns a probability to each such path, including its length. A circuit supplies the physical time taken by each local decision. We keep sequence index and physical time explicit throughout, so that the same law of state succession can describe an episode expressed at different speeds.

Table 1: Quantities used in the derivation. State functions and probability distributions use column vectors unless an index is displayed.

Symbol	Meaning
P_{ij}	Reference probability of a transition from represented state i to j.
π_i	Stationary frequency of state i in the reference transition process.
γ	Probability that a reference event continues at a decision.
K	Number of transitions before completion; an event contains $K+1$ states.
a_i, b_i	Compatibility of a state with initial and terminal information.
$G = (I - \gamma P)^{-1}$	Expected discounted visits under the reference transitions.
$h = Gb$	Continuation field: the total weight of compatible completions.
q_0, q_T	Distributions of the initial and terminal states of generated events.
T_{ij}, s_i	Probability of continuing to j, or completing, at state i.
F_{ij}	Next-state probability conditional on continuation.
n_i, J_{ij}	Expected visits to i and transitions from i to j per event.

A.1 REFERENCE DYNAMICS AND RANDOM SEQUENCE LENGTH

Let $\mathcal{X} = \{1, \dots, N\}$ be a finite set of represented states. A state can include location, heading, context, or the identity of an active assembly. Let P_{ij} be the probability of moving from i to j, with $P\mathbf{1} = \mathbf{1}$. We assume that P is irreducible, and write its positive stationary distribution as a column vector π, so that $P^{\mathsf{T}}\pi = \pi$. Irreducibility ensures that every state can be reached from every other state. Periodicity does not affect the resolvent or the bridge results below.

A sequence contains K transitions and $K+1$ states. Its sample space is the disjoint union

$$\Omega = \coprod_{K=0}^{\infty} \mathcal{X}^{K+1}. \tag{9}$$

For $0 < \gamma < 1$, the reference probability of $\omega = (x_0, \dots, x_K)$ is

$$\mathsf{R}(\omega) = (1-\gamma)\pi_{x_0} \prod_{k=0}^{K-1} \gamma P_{x_k x_{k+1}}. \tag{10}$$

Empty products equal one. Thus $K = 0$ describes a sequence that terminates in its initial state. Summing over states at fixed K gives $(1-\gamma)\gamma^K$. Summing this geometric series gives $\mathsf{R}(\Omega) = 1$. Consequently,

$$\mathsf{R}(K = k) = (1-\gamma)\gamma^k, \qquad \mathbb{E}_{\mathsf{R}} K = \frac{\gamma}{1-\gamma}. \tag{11}$$

Here a step records one sampled transition between represented states; P_{ii} permits persistence of the current representation. The reference can be constructed by adjoining a completed state †. From a live state i, it moves to j with probability γP_{ij} and to † with probability $1-\gamma$. The number of live-to-live transitions before the first visit to † is K. Completion is therefore a stopping event generated by the process. A dwell-time model assigns physical time to these decisions.

The matrix

$$G = (I - \gamma P)^{-1} = \sum_{k=0}^{\infty} \gamma^k P^k \tag{12}$$

exists because $\|\gamma P\|_\infty = \gamma < 1$. Its rows sum to $(1-\gamma)^{-1}$. Every entry is positive under irreducibility, because some power of P connects any chosen pair of states. The joint distribution of the two endpoints is

$$C_{ij} := \mathsf{R}(X_0 = i, X_K = j) = (1-\gamma)\pi_i G_{ij}. \tag{13}$$

Both marginals of C equal π. Stationarity supplies a convenient reference with no preferred starting time. All conditioning and sampling formulas also hold with an arbitrary positive initial distribution in place of π.

Relation to recurrent connectivity. For nonnegative weights W_{ij} with positive row sums $d_i = \sum_j W_{ij}$, define $P_{ij} = W_{ij}/d_i$. If W is symmetric and its graph is connected, then

$$\pi_i = \frac{d_i}{\sum_j d_j}, \qquad \pi_i P_{ij} = \pi_j P_{ji}. \tag{14}$$

The second identity is detailed balance. It follows by substituting $P_{ij} = W_{ij}/d_i$ and using $W_{ij} = W_{ji}$. Directed experience can be represented with an asymmetric P; the bridge construction itself does not require symmetry. Assembly excitability and assembly size may help determine an empirical starting distribution. Equating that distribution with the stationary law is a modeling choice that can be checked from spontaneous sequences.

Relation to the successor representation. Let $N_j(\omega) = \sum_{k=0}^{K} \mathbf{1}\{X_k = j\}$ count visits to state j, including the initial and terminal states. Nonnegative summation gives

$$\mathbb{E}_{\mathsf{R}}[N_j \mid X_0 = i] = \sum_{k \geq 0} \mathsf{R}(K \geq k, X_k = j \mid X_0 = i) \tag{15}$$

$$= \sum_{k \geq 0} \gamma^k (P^k)_{ij} = G_{ij}. \tag{16}$$

Thus the successor representation (Dayan, 1993) is an expected-occupancy statistic of the reference sequence process. For an exact finite-state matrix and a known γ, the relation is invertible: $P = \gamma^{-1}(I - G^{-1})$. The path formulation makes the joint sequence law, its stopping rule, and its conditioning operations explicit. Compressed or approximate occupancy representations can then be studied as approximations to this generative description.

A.2 Endpoint evidence and the Schrödinger bridge

Let $a_i > 0$ and $b_i > 0$ measure the compatibility of state i with the starting and terminal evidence. They are positive factors, without a sum-to-one requirement. Define

$$h = Gb, \qquad Z = \sum_i \pi_i a_i h_i, \qquad \mathsf{Q}_{a,b}(\omega) = \frac{\pi_{x_0} a_{x_0}}{Z} \left(\prod_{k=0}^{K-1} \gamma P_{x_k x_{k+1}} \right) b_{x_K}. \tag{17}$$

The normalizer for $a(X_0)b(X_K)\mathsf{R}$ is $(1-\gamma)Z$. The common stopping factor in R therefore cancels from the normalized expression for $\mathsf{Q}_{a,b}$. The factor a_i expresses starting compatibility, while b_i expresses terminal compatibility. Their effect on observed endpoint frequencies also depends on the paths the map makes available between them.

If the two observations have likelihoods proportional to $a(X_0)$ and $b(X_K)$ and are conditionally independent given the path, Bayes' rule gives (17). More generally, write

$$V(\omega) = \log a_{x_0} + \log b_{x_K}, \qquad A = \mathbb{E}_{\mathsf{R}} e^V = (1-\gamma)Z. \tag{18}$$

For any path law Q absolutely continuous with respect to R,

$$D_{\mathrm{KL}}(\mathsf{Q}\|\mathsf{R}) - \mathbb{E}_{\mathsf{Q}} V = D_{\mathrm{KL}}(\mathsf{Q}\|\mathsf{Q}_{a,b}) - \log A. \tag{19}$$

This follows directly by inserting $\log(\mathsf{Q}_{a,b}/\mathsf{R}) = V - \log A$. Nonnegativity of relative entropy makes $\mathsf{Q}_{a,b}$ the unique minimizer of the left-hand side. Endpoint evidence therefore specifies a tradeoff between retaining the reference process and assigning weight to compatible endpoints.

A.2.1 PRESCRIBED ENDPOINT DISTRIBUTIONS

Schrödinger's original question concerned a population of diffusing particles whose distributions were observed at two times (Schrödinger, 1931; Léonard, 2014). The reference dynamics specify how the particles ordinarily move. The observations specify how often different beginnings and endings occur. The inference problem asks which ensemble of intervening paths best reconciles them. Relative entropy measures the change from the reference ensemble, giving a precise statistical meaning to retaining the dynamics already specified.

Here the two distributions, μ_0 and μ_T, describe frequencies of starting and finishing states across neural events. The Schrödinger bridge is the path law with those marginals that minimizes $D_{\mathrm{KL}}(\mathsf{Q}\|\mathsf{R})$. A single observed starting state is obtained by concentrating μ_0 on that state. The terminal index K belongs to the sampled path, so the second constraint is imposed at each event's completion. We use T in the subscript μ_T to designate this terminal marginal; the event has no externally fixed transition count. This choice allows the same inference problem to determine sequence content and completion. Schrödinger problems with random first-arrival constraints on graphs provide a broader setting for inference at task completion (Eldesoukey & Georgiou, 2024). Here termination is available at every represented state, and its time distribution follows from the reference process and the fitted endpoint factors.

The relative-entropy criterion has a statistical origin. Imagine independently drawing many paths from the reference and observing only their endpoint frequencies. On a finite partition of path space, a histogram with frequencies q_ℓ and reference probabilities r_ℓ has multinomial probability. Stirling's formula gives its leading logarithm as $-M\sum_\ell q_\ell \log(q_\ell/r_\ell)$ for M paths, up to terms smaller than M. Among histograms compatible with the observations, the least relative entropy therefore identifies the most probable ensemble as the sample grows (Léonard, 2014). This interpretation connects the bridge to inference about a population of events: the observed boundaries select an ensemble from learned sequence statistics.

The reduction from paths to endpoints has a short proof. Let $\Gamma_{ij} = \mathsf{Q}(X_0 = i, X_K = j)$, and let Q^{ij} and R^{ij} denote conditional path laws given those endpoints. Factoring each joint probability into an endpoint probability and a conditional path probability gives the relative-entropy chain rule

$$D_{\mathrm{KL}}(\mathsf{Q}\|\mathsf{R}) = D_{\mathrm{KL}}(\Gamma\|C) + \sum_{ij}\Gamma_{ij} D_{\mathrm{KL}}(\mathsf{Q}^{ij}\|\mathsf{R}^{ij}). \tag{20}$$

For a fixed endpoint coupling Γ, the second term is minimized by $\mathsf{Q}^{ij} = \mathsf{R}^{ij}$. The remaining problem is finite-dimensional:

$$\min_{\Gamma\geq 0}\sum_{ij}\Gamma_{ij}\log\frac{\Gamma_{ij}}{C_{ij}}, \quad \sum_j \Gamma_{ij} = \mu_0(i), \quad \sum_i \Gamma_{ij} = \mu_T(j). \tag{21}$$

For strictly positive C, μ_0, μ_T, the feasible set is compact and contains a positive coupling. Strict convexity gives a unique minimizer. The derivative of $x\log x$ diverges to $-\infty$ as $x \downarrow 0$, so the optimum is positive: a small mixture with a positive feasible coupling would otherwise decrease the objective. The Lagrange equations then imply

$$\Gamma^\star_{ij} = \frac{a_i C_{ij} b_j}{\sum_{uv} a_u C_{uv} b_v} = \frac{a_i \pi_i G_{ij} b_j}{Z} \tag{22}$$

for positive factors a, b. Combining this coupling with R^{ij} yields (17). The factors are selected so that its marginals equal the prescribed marginals.

Proposition 1 (Minimal change of the path law)**.** *For a finite irreducible reference chain, $0 < \gamma < 1$, and positive endpoint marginals, the Schrödinger bridge exists and is unique. It belongs to the family $\mathsf{Q}_{a,b}$, and its conditional distribution of routes and lengths given the two endpoints equals that of R.*

Proof. Existence, uniqueness, and the factorized coupling follow from (21) and (22). Equation (20) then gives the unique conditional path laws. Equivalently, for every feasible Q,

$$D_{\mathrm{KL}}(\mathsf{Q}\|\mathsf{R}) = D_{\mathrm{KL}}(\mathsf{Q}\|\mathsf{Q}^\star) + D_{\mathrm{KL}}(\mathsf{Q}^\star\|\mathsf{R}). \tag{23}$$

To verify this identity, subtract the first term on the right and average $\log(\mathsf{Q}^\star/\mathsf{R}) = \log a(X_0) + \log b(X_K) - \log A$ under Q. Its expectation depends only on the two prescribed marginals, and therefore equals its expectation under $\mathsf{Q}^\star$. □

This property expresses precisely how a task reuses the map. Boundary information changes which pairs of states are connected across events. For each pair, the reference process supplies the relative probabilities of the available routes. For two supported paths ω, ω' with the same endpoints,

$$\frac{\mathsf{Q}_{a,b}(\omega)}{\mathsf{Q}_{a,b}(\omega')} = \frac{\mathsf{R}(\omega)}{\mathsf{R}(\omega')}. \tag{24}$$

This equality also includes paths of different lengths. A task that changes the relative desirability of routes with identical endpoints can be represented by an additional path observation, an intermediate constraint, or a changed reference process.

Any positive a, b define a bridge for the marginals that they induce. Matching an externally specified pair of marginals requires solving for those factors. When some target probabilities are zero, the same construction applies on the supported endpoint sets, provided that the corresponding endpoint coupling is feasible.

Association between beginnings and endings. Equation (22) also specifies how the two endpoints remain associated. For any four states with positive endpoint probabilities,

$$\frac{\Gamma_{ij}\Gamma_{kl}}{\Gamma_{il}\Gamma_{kj}} = \frac{C_{ij}C_{kl}}{C_{il}C_{kj}}. \tag{25}$$

All four boundary factors cancel. Endpoint calibration therefore changes marginal frequencies while retaining these association ratios from the reference. The map determines which beginnings and endings tend to belong together; boundary factors control how frequently each is recruited. Paired endpoints provide a way to test this consequence after fitting the marginal frequencies.

A.3 The local sampling rule

The field h_i sums the weights of all future continuations from state i. Equivalently,

$$h_i = \sum_{n\geq 0} \gamma^n \mathbb{E}_i[b(X_n)], \qquad (1-\gamma)h_i = \mathbb{E}_{\mathsf{R}}[b(X_K) \mid X_0 = i]. \tag{26}$$

Its first term describes finishing at the present state. The remaining terms describe taking one step and reconsidering the same question:

$$h_i = b_i + \gamma \sum_j P_{ij} h_j. \tag{27}$$

The inverse in (12) proves uniqueness. Positivity follows from $b > 0$ and the nonnegative series. In particular,

$$\frac{\min_i b_i}{1-\gamma} \leq h_i \leq \frac{\max_i b_i}{1-\gamma}. \tag{28}$$

Define the initial distribution, probability of a transition, and probability of termination by

$$q_0(i) = \frac{\pi_i a_i h_i}{Z}, \tag{29}$$

$$T_{ij} = \frac{\gamma P_{ij} h_j}{h_i}, \qquad s_i = \frac{b_i}{h_i}. \tag{30}$$

Here T_{ij} includes the probability that the event continues. Equation (27) gives $s_i + \sum_j T_{ij} = 1$. Conditional on continuation, the next-state distribution is

$$F_{ij} := \Pr(X_{k+1} = j \mid X_k = i, K \geq k+1) = \frac{P_{ij} h_j}{(Ph)_i}. \tag{31}$$

The transformation $\gamma P \mapsto T$ is the Doob h-transform of a process that can terminate (Doob, 1957). In this setting h satisfies an equation with a positive source b. It is excessive for γP, since $\gamma P h \leq h$. The remaining probability at each state is exactly the stopping probability.

Proposition 2 (Exact local sampling)**.** *An event initialized with* (29) *and evolved with* (30) *terminates almost surely and has path law* $\mathsf{Q}_{a,b}$.

Proof. The probability of any finite path produced by these rules is

$$q_0(x_0)\prod_{k=0}^{K-1} T_{x_k x_{k+1}} s_{x_K} = \frac{\pi_{x_0} a_{x_0} h_{x_0}}{Z} \prod_{k=0}^{K-1} \frac{\gamma P_{x_k x_{k+1}} h_{x_{k+1}}}{h_{x_k}} \frac{b_{x_K}}{h_{x_K}} \tag{32}$$

$$= \mathsf{Q}_{a,b}(\omega). \tag{33}$$

The factors h cancel successively along the path. Since (17) sums to one over finite paths, the sampler assigns no probability to infinite events. A direct bound follows from (28): $s_i \geq (1-\gamma)\min b/\max b > 0$. □

The recursion and the sampler answer successive parts of the computation. Computing h distributes information about a terminal cue through the learned transitions. Once that field is present, a sequence can be generated through local comparisons. A state is favored when its possible continuations collectively have high terminal compatibility. Each route contributes according to its reference probability, survival, and ending. Several likely continuations can therefore increase a state's influence, while a bottleneck can reduce it. The field retains alternatives throughout this calculation; the local race selects one continuation when the sequence is expressed.

The local rule can also be read directly as Bayes' rule. Under the reference, the alternatives at i are completion with probability $1-\gamma$ and continuation to j with probability γP_{ij}. Completion receives likelihood b_i. Continuation to j receives the expected likelihood of its eventual ending, $(1-\gamma)h_j$. Their normalized posterior weights are exactly s_i and T_{ij}. The future constraint is represented as present input. Its influence on the current decision is the conditional expectation computed through the map.

Completion is one of the alternatives at every visit. Even at a state with strong terminal evidence, the event can continue when further compatible sequences retain substantial weight. Repeated visits are therefore possible before the event ends. The model assigns a probability to the point of completion together with the sequence that precedes it.

Which endings contribute to the field. Starting from state i, the transformed event terminates at j with probability

$$B_{ij} := \mathsf{Q}(X_K = j \mid X_0 = i) = \frac{G_{ij} b_j}{h_i}, \qquad \frac{\partial \log h_i}{\partial \log b_j} = B_{ij}. \tag{34}$$

The first equality follows by summing all supported paths from i to j; the second follows by differentiating $h_i = \sum_j G_{ij} b_j$. Every row of B sums to one. The same formula applies to the remaining part of a live event at i, by the Markov property of the sampler. Thus the response of a state's field to one possible ending equals that ending's probability within the state's available completions. A remote terminal cue exerts a strong influence where the map supplies substantial access to it. Multiplying every b_j by the same constant multiplies every h_i by that constant and leaves all local probabilities unchanged.

Relative entropy as a cost of changing dynamics. The reference process offers the local alternatives $(1-\gamma, \gamma P_{i1}, \ldots, \gamma P_{iN})$ at each state. Consider another Markov sampler with local probabilities $(s_i, T_{i1}, \ldots, T_{iN})$, almost-sure termination, and finite expected visit counts n_i. Assume it uses transitions supported by P. Taking the logarithm of the ratio of its path probability to (10), then averaging, gives

$$D_{\mathrm{KL}}(\mathsf{Q}\|\mathsf{R}) = D_{\mathrm{KL}}(q_0\|\pi) \tag{35}$$

$$+ \sum_i n_i \left[s_i \log \frac{s_i}{1-\gamma} + \sum_j T_{ij} \log \frac{T_{ij}}{\gamma P_{ij}} \right]. \tag{36}$$

Each visit contributes the relative entropy between the available local choices. The bridge therefore minimizes the aggregate change in initiation, continuation, and termination probabilities required by the endpoint distributions. Equation (19) supplies the corresponding formulation with terminal evidence. This connects the path model to stochastic control with relative-entropy costs and linear equations for transformed value functions (Todorov, 2009).

A value equation that includes completion. The control interpretation follows directly from the local alternatives. Write $H = (1-\gamma)h$, so H_i is the expected terminal compatibility under the reference starting at i, and define $V_i = -\log H_i$. At state i, let $p_\dagger$ be a proposed stopping probability and p_j a proposed probability of continuing to j, with $p_\dagger + \sum_j p_j = 1$. The reference probabilities for these alternatives are $r_{i\dagger} = 1-\gamma$ and $r_{ij} = \gamma P_{ij}$. The field recursion is equivalent to

$$V_i = \min_p \left\{ D_{\mathrm{KL}}(p\|r_i) - p_\dagger \log b_i + \sum_j p_j V_j \right\}. \tag{37}$$

Only alternatives supported by r_i enter the minimization. To prove the identity, assign costs $c_\dagger = -\log b_i$ and $c_j = V_j$. For $A_i = \sum_\ell r_{i\ell} e^{-c_\ell}$ and $p^*_{i\ell} = r_{i\ell} e^{-c_\ell}/A_i$, direct substitution gives

$$D_{\mathrm{KL}}(p\|r_i) + \sum_\ell p_\ell c_\ell = D_{\mathrm{KL}}(p\|p^*_i) - \log A_i. \tag{38}$$

The minimum is $-\log A_i$. Since $A_i = (1-\gamma)b_i + \gamma \sum_j P_{ij} H_j = H_i$, the minimizing probabilities are precisely s_i and T_{ij}. Completion has terminal cost $-\log b_i$; continuation has the expected remaining cost. Both pay the same statistical cost for changing the reference choice. The exponential change of variables makes this joint decision linear. The learned reference and its survival probability specify the relative weight of short, long, common, and unusual episodes. Scaling every b_i by the same constant shifts every V_i by a constant and leaves the choices unchanged.

A.4 Occupancy, probability flow, and sequence length

We use column vectors for state functions and distributions. Let $f_0 = \pi \odot a$, where $\odot$ denotes elementwise multiplication, and define

$$f_k = (\gamma P^{\mathsf{T}})^k f_0, \qquad f = \sum_{k\geq 0} f_k = G^{\mathsf{T}}(\pi \odot a). \tag{39}$$

The transpose expresses forward propagation of a distribution. Backward propagation of a state function uses P, as in $h = Gb$.

Summing the weights of all prefixes that end at i after k transitions gives $f_k(i)$. Summing all suffixes that can follow that prefix gives h_i. Therefore,

$$\mathsf{Q}(X_k = i, K \geq k) = \frac{f_k(i) h_i}{Z}, \tag{40}$$

$$\mathsf{Q}(X_k = i, K = k) = \frac{f_k(i) b_i}{Z}, \tag{41}$$

$$\mathsf{Q}(X_k = i, K \geq k+1) = \frac{f_k(i)(h_i - b_i)}{Z}. \tag{42}$$

The first expression is a subprobability distribution whose mass is $\mathsf{Q}(K \geq k)$. Dividing by that mass yields the state distribution conditional on reaching step k. These distinctions matter when comparing events with different durations.

Summing over k gives the terminal marginal and the expected visit count:

$$q_T(i) = \frac{b_i f_i}{Z}, \qquad n_i := \mathbb{E}_{\mathsf{Q}} N_i = \frac{f_i h_i}{Z}. \tag{43}$$

The expected number of transitions from i to j is

$$J_{ij} = \frac{\gamma f_i P_{ij} h_j}{Z}. \tag{44}$$

The two recursions for f and h give the flow identities

$$n_i = q_0(i) + \sum_j J_{ji} = q_T(i) + \sum_j J_{ij}. \tag{45}$$

Each visit begins either with an event initiation or an incoming transition, and ends either with termination or an outgoing transition. Thus a population occupancy pattern combines how often a

state is reached with how much further sequence weight it supports. Its normalization $n_i / \sum_j n_j$ describes the distribution obtained by pooling all state visits across events. It differs from the distribution obtained by giving each event equal weight after normalizing its individual occupancy.

Subtracting the incoming and outgoing balances gives

$$\sum_j (J_{ij} - J_{ji}) = q_0(i) - q_T(i). \tag{46}$$

Initiation and completion thus act as sources and sinks of probability across events. This relation holds for reversible and directed maps alike. Even symmetric learned connectivity can support net traffic from starting regions to terminal regions because the two boundaries recruit and remove activity in different places. Replay direction can consequently change with endpoint assignment while the recurrent map is fixed.

Length distribution and response to survival. Summing at fixed K gives

$$\mathsf{Q}(K = k) = \frac{\gamma^k (\pi \odot a)^\mathsf{T} P^k b}{Z}. \tag{47}$$

For fixed P, π, a, b, let $t = \log \gamma < 0$. Differentiating the convergent power series on compact subintervals of $0 < \gamma < 1$ gives

$$\frac{\partial \log Z}{\partial t} = \mathbb{E}_\mathsf{Q} K, \qquad \frac{\partial \mathbb{E}_\mathsf{Q} K}{\partial t} = \mathrm{Var}_\mathsf{Q}(K). \tag{48}$$

The first identity follows by bringing down a factor k from γ^k. The second follows by differentiating the normalized distribution, whose score is $K - \mathbb{E}K$. The same argument gives $\partial_t \mathbb{E}A = \mathrm{Cov}(A, K)$ for any bounded path statistic A independent of γ.

With a flat terminal factor $b = c\mathbf{1}$, one has $h = c\mathbf{1}/(1-\gamma)$, $s_i = 1 - \gamma$, and $F = P$. With both factors flat, $\mathsf{Q} = \mathsf{R}$. A localized terminal factor changes both the paths followed and where they stop. Its effect on duration depends on the starting cue and on the reference routes, as expressed by (47).

Long sequences under a fixed reference. Endpoint information changes the distribution of transition counts through both initiation and local stopping. An exact property describes its effect on long sequences. Write $D_h = \mathrm{diag}(h)$. The matrix of transitions before termination obeys

$$T = \gamma D_h^{-1} P D_h, \qquad T^k = \gamma^k D_h^{-1} P^k D_h. \tag{49}$$

This is a similarity transformation, so the eigenvalues of T equal those of γP, and its spectral radius is γ. More directly, summing all event prefixes of length k gives

$$\mathsf{Q}(K \geq k) = \frac{\gamma^k (\pi \odot a)^\mathsf{T} P^k h}{Z}. \tag{50}$$

Since a stochastic matrix averages its argument, the last factor is bounded between $\min_i h_i \sum_i \pi_i a_i / Z$ and $\max_i h_i \sum_i \pi_i a_i / Z$, both finite and positive. Consequently,

$$\lim_{k \to \infty} \frac{1}{k} \log \mathsf{Q}(K \geq k) = \log \gamma. \tag{51}$$

For an aperiodic P, the prefactor in (50) also converges, because $P^k h \to \mathbf{1}\pi^\mathsf{T} h$. Irreducibility alone suffices for the displayed exponential rate. Thus changing positive endpoint factors can shift the mean, the mode, and the finite-length profile while retaining this asymptotic rate. A change in the tail rate, measured in transition counts under a fixed state representation, identifies a change beyond endpoint factors alone. Physical duration additionally depends on the transition clock. The continuing chain F, obtained by renormalizing each live row, is a different operator and can have a different spectrum.

Sampling at fixed duration. If an analysis conditions on $K = L$, its terminal field depends on the number of remaining transitions: $h_k = P^{L-k} b$. The local transition at index $k < L$ is $P_{ij} h_{k+1}(j) / h_k(i)$, and the event ends at L. The initial distribution is proportional to $\pi_i a_i h_0(i)$. The factor γ^L cancels when conditioning on duration. This gives the appropriate comparison for datasets whose event durations have been fixed by construction.

Analytical illustration in Figure 6. The illustration uses the bottleneck graph constructed by the Experiment 1 reference code with seed 20260524: 41 states in two rooms joined by a corridor, with 64 undirected non-self edges. The constructor supplies the transition matrix P and its stationary distribution π. Writing z_i for the plotted coordinates, the start state minimizes $z_{i,1} + 0.2z_{i,2}$ and the goal state maximizes it. The endpoint factors are $a_i = 10^{-8} + \exp(-\|z_i - z_{\text{start}}\|^2/0.5)$ and $b_i = 10^{-8} + \exp(-\|z_i - z_{\text{goal}}\|^2/0.5)$, with survival $\gamma = 0.93$. Matrix inversion gives G, h, f, and Z. Thus the figure evaluates the same path law at the level of map traffic, local choices, and whole-event duration.

Panels a and b compare flat terminal input, $b = \mathbf{1}$, with localized goal input while holding the start factor and reference dynamics fixed. Node area increases with expected visit count on a shared scale with a visibility floor, and goal-map node color represents $h_i / \max_j h_j$ on a logarithmic scale. Edge width is an affine function of $J_{ij} + J_{ji}$ from Equation (44), and arrows indicate the sign of $J_{ij} - J_{ji}$. Every state and non-self edge is retained. Self-transitions contribute to all probability calculations.

Panels d and e follow the position represented during an ongoing sequence. Let $m_k^{\mathsf{T}} = (\pi \odot a)^{\mathsf{T}}(\gamma P)^k$. The probability that the sequence has reached state i after k transitions and remains available to continue or finish there is $m_k(i)h_i/Z$. Conditioning on the event reaching that index gives

$$\Pr(X_k = i \mid K \geq k) = \frac{m_k(i)h_i}{\sum_j m_k(j)h_j}. \tag{52}$$

The heatmaps sum this probability over all states with the same horizontal coordinate, producing nine position bins at each index $k = 0, \ldots, 80$. Each column sums to one. The corridor bin contains one state and each room column contains five; occupancy in the corridor measures the probability of representing it at that index, while map edge traffic measures how often a route crosses it. Both panels share a square-root color normalization of probability. White curves show the exact conditional mean horizontal position at each transition index. Goal input shifts the evolving position distribution toward the far room through the same transitions displayed above.

Panel c displays paired probabilities for all four available outcomes at the goal: moving left, moving down, remaining in the state, and stopping. For a flat terminal factor, these probabilities are γP_{ij} and $1 - \gamma$. For the goal factor, they are $\gamma P_{ij} h_j / h_i$ and b_i / h_i. In this example, goal conditioning changes the stopping probability there from 0.070 to 0.434.

Panel f evaluates Equation (47) for $k = 0, \ldots, 100$, keeping the same start factor in both conditions. A flat terminal factor gives the geometric length law $(1 - \gamma)\gamma^k$. The mean numbers of transitions are 13.286 for flat terminal input and 39.943 for goal input; the probabilities beyond the displayed range are 0.000656 and 0.010177. These are deterministic calculations over paths of every length. The source tables retain the state factors, field, visit counts, edge traffic, local probabilities, state and position distributions at each index, and displayed length probabilities.

A.5 Attractor maps and changes in replay direction

An attractor model describes how recurrent activity settles into stable or slowly changing population configurations (Hopfield, 1982; Wills et al., 2005). Such a configuration supplies a reproducible representation of a situation. Noise, adaptation, and recurrent interactions permit transitions among configurations; their probabilities define P at the resolution of the model. Stability gives the represented states their persistence, and transitions give the map its sequential structure. A useful state representation retains the information needed to predict these transitions.

The continuation field gives this established map a task-dependent use. It assigns weight to a configuration according to the complete paths that can follow it. Dendritic gain changes the probabilities of leaving for its successors, and the completion input determines whether the episode ends there. The resulting change in apparent attraction follows from the reachable endings throughout the map. Reversible networks and gradient diffusions make this relation explicit as a change in a potential; the transition construction also applies to directed maps.

Reversible networks. Suppose detailed balance (14) holds. The chain F that records transitions conditional on continuation is reversible with respect to

$$m_i = \frac{\pi_i h_i (Ph)_i}{\sum_j \pi_j h_j (Ph)_j}. \tag{53}$$

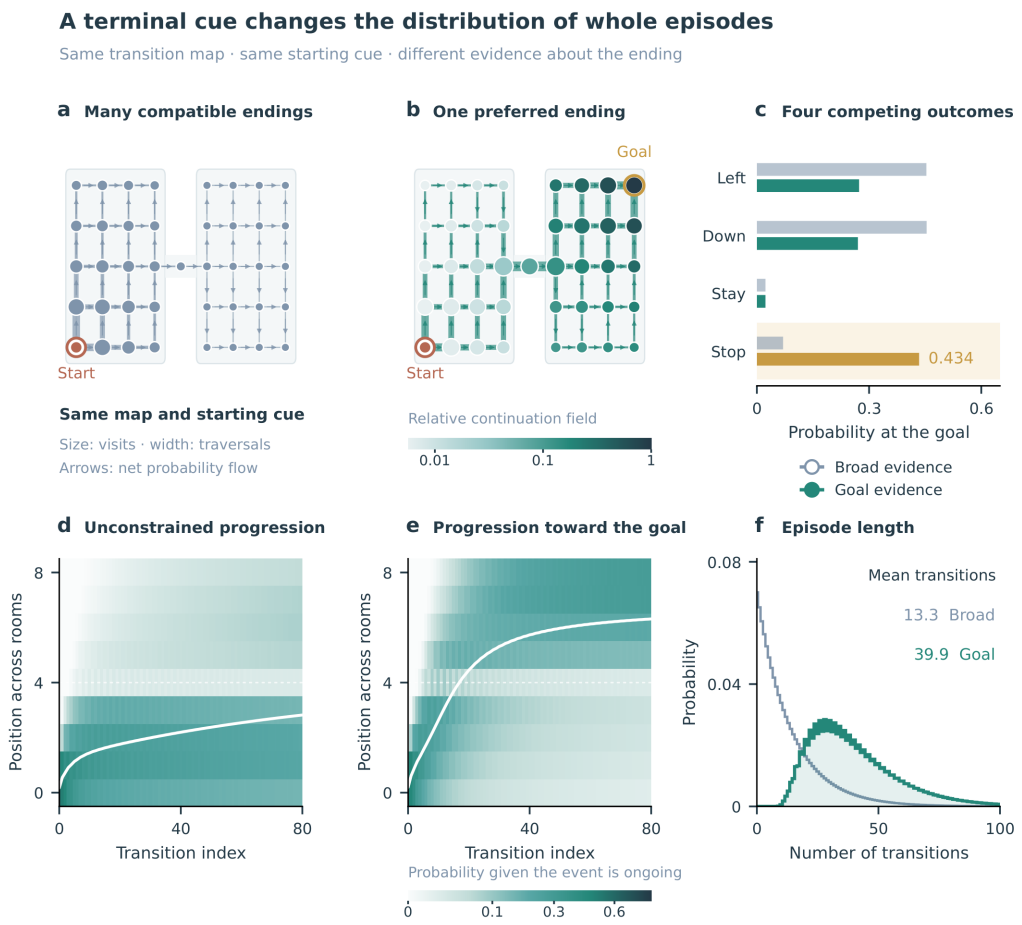

Figure 6: **Endpoint input organizes the route, its progression, and its completion.** Exact calculations on a 41-state map with a fixed start cue and survival probability $\gamma = 0.93$. (a,b) Broad and localized terminal input act on the same learned transitions. Edge widths show expected traversals on a shared scale; arrows show net flow. Node area grows with expected visits using a common scale and a visibility floor. Goal-map node color shows relative access to the endpoint on a logarithmic scale. Every state and non-self edge is displayed. (c) All four available outcomes at the goal enter the local competition. Gray bars show broad terminal input; teal bars show goal input, with completion highlighted in gold. Terminal input changes the probability of stopping there from 0.070 to 0.434. (d,e) The position represented at each transition index, conditional on the event reaching that index. All states contribute to the nine horizontal position bins. The common color scale uses square-root normalization of probability; each column sums to one. White curves show the conditional mean position. (f) The same local probabilities determine sequence length. Means include every possible length; steps show the exact probability masses through 100 transitions. Shading marks the goal distribution. All values are deterministic calculations.

Indeed, $m_i F_{ij}$ is proportional to $\pi_i P_{ij} h_i h_j$, which is symmetric in i, j. The quantity $-\log m_i$ gives a potential for the stationary distribution of this continuing chain, up to an additive constant. Relative to $-\log \pi_i$, its change contains $-\log h_i - \log(Ph)_i$. Where stopping is rare, $\gamma(Ph)_i/h_i = 1 - s_i$ is close to one, so this change approaches $-2\log h_i$ up to a constant.

The distribution m describes repeated application of the conditional continuation rule. Recorded events also have beginnings and endings, giving occupancy $n_i = f_i h_i / Z$. The same task input can therefore alter the continuing dynamics and the distribution of activity pooled over complete events in different ways. An inferred potential describes the former; measured visits during events depend jointly on initiation, propagation, and completion.

Continuous attractor dynamics. Consider a diffusion with generator

$$\mathcal{L}\varphi = -\nabla U \cdot \nabla\varphi + D\Delta\varphi, \qquad D > 0, \tag{54}$$

and an independent exponential stopping time with rate $\kappa > 0$. Here U is a potential for the reference drift. Assume the reference diffusion is nonexplosive and g is a positive bounded terminal likelihood, with sufficient smoothness for the derivatives below. Define $H(x) = \mathbb{E}_x[g(X_T)]$. Resolving this expectation by the stopping time gives

$$H(x) = \int_0^\infty \kappa e^{-\kappa t} \mathbb{E}_x[g(X_t)]\, dt, \qquad (\kappa - \mathcal{L})H = \kappa g. \tag{55}$$

To obtain the second identity, use $\partial_t \mathbb{E}_x[g(X_t)] = \mathcal{L}\mathbb{E}_x[g(X_t)]$ and integrate by parts. The boundary term at zero is $-\kappa g$, and boundedness of g makes the term at infinity vanish; the remaining integral is κH. Thus $\mathcal{L}H = \kappa(H - g)$. This is the continuous counterpart of $(1-\gamma)h$: terminal compatibility is averaged over every possible completion time. The transformed generator acting on a smooth function φ is

$$\widetilde{\mathcal{L}}\varphi = H^{-1}\mathcal{L}(H\varphi) - H^{-1}\varphi\mathcal{L}H \tag{56}$$

$$= \mathcal{L}\varphi + 2D\nabla \log H \cdot \nabla\varphi. \tag{57}$$

The product rule proves the second equality. To include termination explicitly, apply the same transformation to the generator $\mathcal{L} - \kappa$ of the reference before completion:

$$H^{-1}(\mathcal{L} - \kappa)(H\varphi) = \widetilde{\mathcal{L}}\varphi - \frac{\kappa g}{H}\varphi. \tag{58}$$

The negative term removes probability at the transformed stopping rate $\kappa g/H$. Both the change in drift and the change in stopping therefore follow from the same field. The reference drift becomes the gradient drift of

$$U_{\text{task}} = U - 2D\log H. \tag{59}$$

This calculation gives a precise meaning to modulation of an attractor map. Recurrent dynamics determine the reference potential and its accessible routes. Terminal input is propagated through those dynamics to obtain H. The added drift follows the gradient of the logarithm of this propagated compatibility. A basin can gain influence because paths from it have useful completions, even when the terminal cue is concentrated elsewhere. The same operation can reshape attraction basins and transition rates throughout the landscape. Equation (58) completes the construction by assigning a stopping rate to the activity moving through it. For a general nonreversible process, (30) gives the change in transition probabilities directly.

Time reversal and endpoint exchange. Let $\omega^{\leftarrow} = (x_K, \dots, x_0)$. Under detailed balance, multiplication of the edge identities gives

$$\pi_{x_0} \prod_{k=0}^{K-1} P_{x_k x_{k+1}} = \pi_{x_K} \prod_{k=0}^{K-1} P_{x_{k+1} x_k}. \tag{60}$$

Consequently, $\mathsf{R}(\omega) = \mathsf{R}(\omega^{\leftarrow})$ and

$$\mathsf{Q}_{a,b}(\omega^{\leftarrow}) = \mathsf{Q}_{b,a}(\omega), \qquad \frac{\mathsf{Q}_{a,b}(\omega)}{\mathsf{Q}_{a,b}(\omega^{\leftarrow})} = \frac{a_{x_0} b_{x_K}}{a_{x_K} b_{x_0}}. \tag{61}$$

The partition functions agree because $\mathrm{diag}(\pi)G$ is symmetric. Exchanging the two boundary factors therefore reverses the entire path distribution in a reversible map. For a directed map, reversal additionally involves the reversed reference kernel $P^{\leftarrow}_{ij} = \pi_j P_{ji}/\pi_i$. This identifies separately the contributions of learned directional structure and boundary assignment to replay direction.

A.6 Computing and sampling the field in a circuit

The circuit has two coupled forms of activity. A maintained population response represents the compatibility of possible continuations. A changing assembly pattern represents the state of the episode currently being generated. Terminal input establishes the first through recurrent relaxation. The second advances through stochastic competition, with each candidate weighted by the maintained response it receives. Completion enters the same competition as a possible continuation. Figure 7 presents the circuit organization, and Figure 8 gives its local rates and learning signals.

The derivation proceeds from the neural rate equation to the continuation field, then from physical synaptic weights and population activities to the probabilities of complete paths. A reciprocal recurrent map supports both operations. Maintaining the field while assemblies change requires a separation of their activity variables, which can be supplied by distinct populations or input compartments within the circuit.

A.6.1 From recurrent neural dynamics to the continuation field

The field equation follows from an explicit recurrent population model. We use the rate dynamics of He et al. (2026b), whose goal-clamped response computes a navigation field:

$$\tau_u \dot{r}_i = -r_i + g_s \sum_j S_{ij}\phi(r_j) - g_{\mathrm{inh}}\phi(r_i) + I_i^{\mathrm{bg}} + I_i^{\mathrm{EC}} + I_i^{\mathrm{goal}}. \tag{62}$$

Here $S_{ij} \geq 0$ is excitation from neuron or assembly j to i, local inhibition provides negative feedback, and ϕ is an increasing saturating response. Background and entorhinal currents establish the baseline r^0. Hold these currents fixed while the terminal input is represented. A distributed I^{goal} can then express several possible endings. Subtracting the baseline removes its fixed contribution and reveals how terminal input propagates through the same recurrent connections.

Write $u_i = r_i - r_i^0$ and expand

$$\phi(r_i^0 + u_i) - \phi(r_i^0) = \rho_i u_i + R_i(u_i), \qquad \rho_i = \phi'(r_i^0). \tag{63}$$

If $|\phi''| \leq M$ throughout the activity range, Taylor's theorem gives $|R_i(u_i)| \leq M|u_i|^2/2$. Subtracting the baseline equation from (62) yields

$$\tau_u \dot{u}_i = -(1 + g_{\mathrm{inh}}\rho_i)u_i + g_s \sum_j S_{ij}\rho_j u_j + I_i^{\mathrm{goal}} + g_s \sum_j S_{ij} R_j(u_j) - g_{\mathrm{inh}} R_i(u_i). \tag{64}$$

Thus recurrence propagates the cue, and leak together with local inhibition attenuates its propagation.

For a common local slope $\rho_i = \rho > 0$, define

$$\ell = 1 + g_{\mathrm{inh}}\rho, \qquad \tau_h = \frac{\tau_u}{\ell}, \qquad \gamma = \frac{g_s\rho}{\ell}, \qquad \beta = \frac{I^{\mathrm{goal}}}{\ell}. \tag{65}$$

The first-order response is $\tau_h \dot{u} = -u + \gamma S u + \beta$. Division by ℓ rescales both the time constant and the terminal current. To relate this equation to the transition kernel, assume that P is reversible and let $\Pi = \mathrm{diag}(\pi)$. Choose the recurrent weights

$$S = \Pi^{1/2} P \Pi^{-1/2}, \qquad S = S^{\mathsf{T}}. \tag{66}$$

Detailed balance gives $S_{ij} = \sqrt{\pi_i/\pi_j}\, P_{ij} = S_{ji}$. For example, symmetric transition strengths W with degrees d_i give $P = D^{-1}W$ and $S = D^{-1/2} W D^{-1/2}$, where $D = \mathrm{diag}(d)$. Thus reciprocal weights can encode a transition process whose state frequencies are nonuniform. The coordinate factors account for these frequencies.

Set $v = \Pi^{-1/2}u$ and $b = \Pi^{-1/2}\beta$. Multiplying the response equation by $\Pi^{-1/2}$ gives

$$\tau_h \dot{v} = -v + b + \gamma P v. \tag{67}$$

For $0 < \gamma < 1$, the unique equilibrium is $v = h$, or $u_i = \sqrt{\pi_i} h_i$ in neural coordinates. Neural activity therefore carries the continuation field with a fixed statewise gain. An activation function that is affine over the operating range gives this equality exactly. For a smooth saturating response, the curvature terms in (64) quantify the departure from it. For $e = v - h$ in the linear response dynamics,

$$e(t) = e^{-t/\tau_h} e^{\gamma P t/\tau_h} e(0), \qquad \|e(t)\|_\infty \leq e^{-(1-\gamma)t/\tau_h} \|e(0)\|_\infty. \tag{68}$$

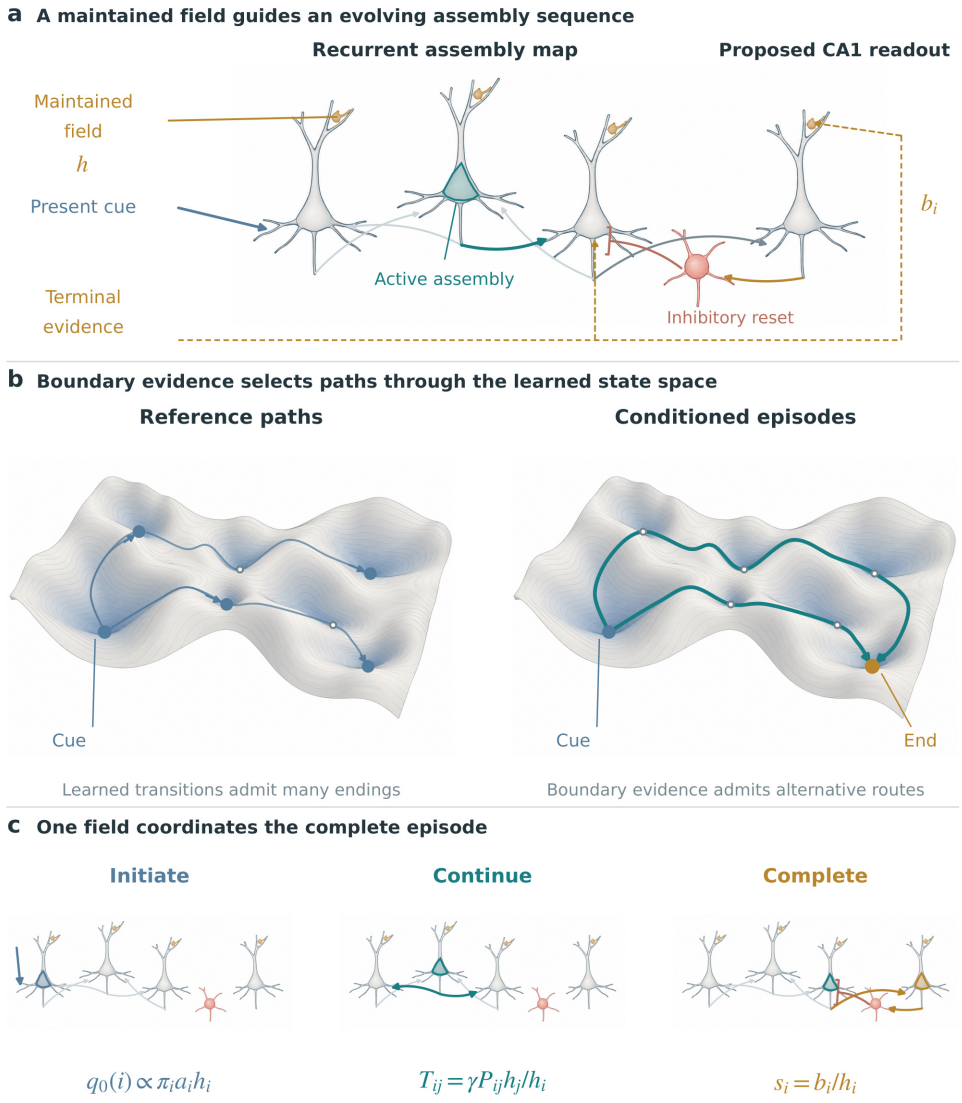

Figure 7: **A continuation field organizes complete episodes within a recurrent cognitive map.** (a) Recurrent populations carry a maintained continuation field and a changing assembly sequence. Gold apical inputs on recurrent cells represent the field h available to candidate assemblies; the teal soma marks current episode activity. The present cue recruits an initial assembly. Terminal evidence enters the field computation and provides input b_i to the proposed CA1 completion readout, whose activation recruits inhibitory reset. Dashed connections indicate the model's terminal inputs. (b) The same conceptual landscape is shown at identical scale and registration. Its basins represent stable population states, and its accessible routes represent learned transitions. Reference paths from a common cue can end at several states. Conditioning on a compatible ending favors an ensemble of routes to that ending, illustrated by two complete teal paths. Both panels depict the same learned state space; the comparison concerns path selection. Terrain height and line width have no quantitative interpretation. (c) Three views of the same circuit show initiation, competition between possible continuations, and completion through the readout and inhibition. The maintained field remains available across these operations. The equations give their probabilities. Cellular motifs represent populations; the diagrams illustrate computational organization without anatomical scale.

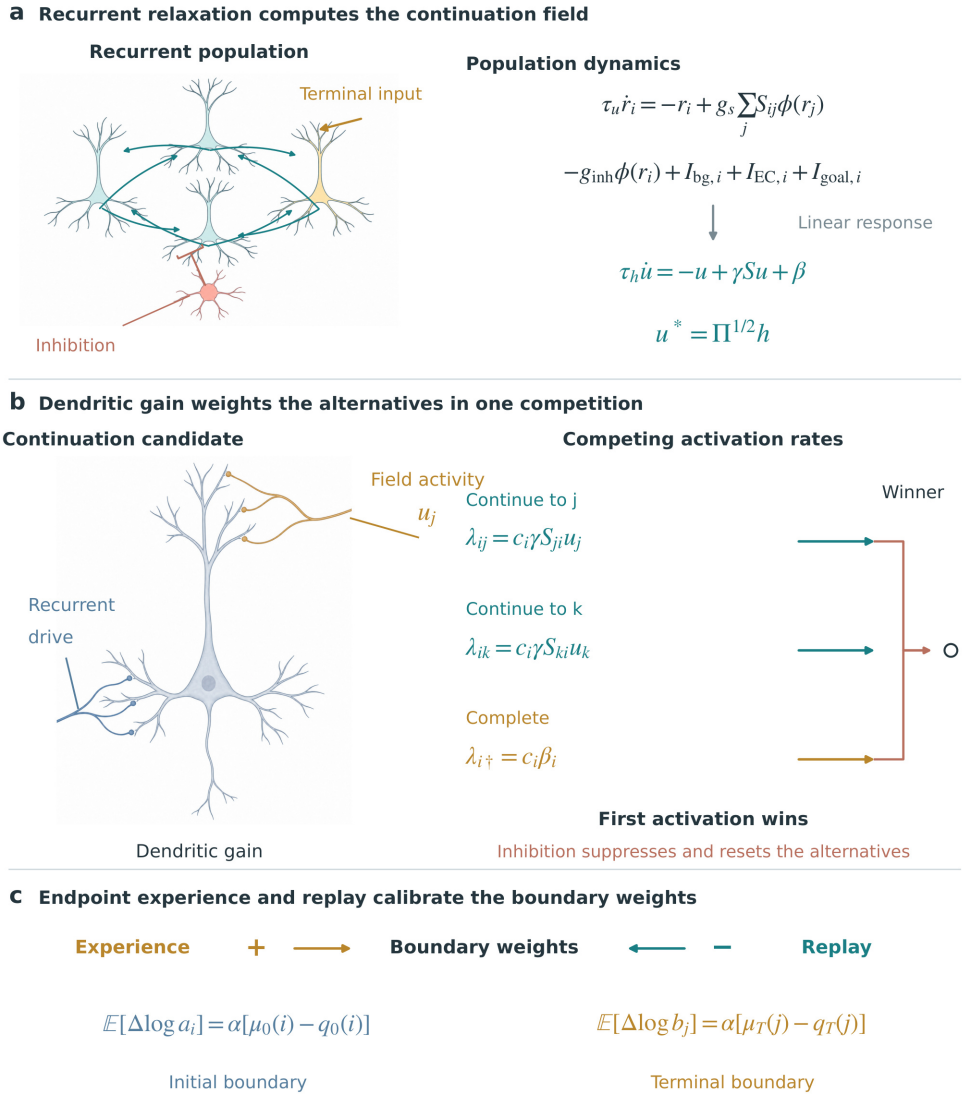

Figure 8: **Neural dynamics compute the field, sample an episode, and calibrate its boundaries.** (a) The population rate equation of He et al. (2026b) propagates terminal input through recurrent excitation regulated by leak and inhibition. The amber cell illustrates a population receiving terminal input. Editable connections indicate recurrent excitation and a representative inhibitory contact. The matched linear response of the reciprocal recurrent map computes the continuation field in neural coordinates, $u^* = \Pi^{1/2} h$. (b) A continuation candidate receives learned recurrent drive through basal afferents and maintained field activity through apical afferents. Their product determines its activation rate. Two possible continuations and a separate completion channel enter the same exponential activation race. The first activation selects the outcome and recruits inhibition, which resets the alternatives. The maintained field remains available across successive decisions. Common gain changes the waiting time while preserving outcome probabilities. Recurrent connectivity and dendritic integration supply physiological ingredients (Guzman et al., 2016; Makara & Magee, 2013; Brandalise et al., 2016). (c) Experienced and replayed endpoints make opposing contributions to the initial and terminal boundary updates. Replay supplies the endpoint frequencies expected under the current circuit. Cellular illustrations show candidate implementations without anatomical scale; the equations specify the model's operations.

The bound follows by expanding the matrix exponential and using $\|P^k\|_\infty = 1$. Starting from $v(0) \geq 0$, the response remains nonnegative: the propagator has nonnegative entries, as does the integrated response to b. The field can therefore be represented by increases above baseline. Increasing γ extends the reference sequence distribution and slows the least attenuated relaxation mode. Longer temporal integration through the map consequently comes with a longer field establishment time.

How the cue spreads during relaxation. The transient response also has a path interpretation. Suppose the terminal current is switched on at time zero and $v(0) = 0$. Variation of constants gives

$$v(t) = \int_0^{t/\tau_h} e^{-s} e^{\gamma P s} b \, ds = \sum_{k \geq 0} w_k(t) \gamma^k P^k b, \tag{69}$$

$$w_k(t) = 1 - e^{-t/\tau_h} \sum_{m=0}^{k} \frac{(t/\tau_h)^m}{m!}. \tag{70}$$

To obtain the second expression, expand $e^{\gamma P s}$ in powers of P and integrate each term. Nonnegativity permits exchanging the integral and the series. The integral of $e^{-s} s^k/k!$ is the displayed w_k. Thus $w_k(t)$ is the probability that a Poisson variable with mean t/τ_h exceeds k. Short paths contribute early; longer paths contribute as recurrent propagation develops. As $t \to \infty$, every w_k tends to one and the sum approaches Gb. This describes the establishment of the field, whose activity simultaneously represents many continuations. The subsequently sampled sequence is drawn through the separate competition dynamics below. A rapid change of terminal input therefore predicts a developing population response before its full effect on distant successor choices is available.

Once established, the field can guide a faster succession of assemblies while terminal input remains stable. More generally, suppose the input varies with time and define its instantaneous equilibrium $h(t) = Gb(t)$. The error $e(t) = v(t) - h(t)$ satisfies $\tau_h \dot{e} = -(I - \gamma P)e - \tau_h \dot{h}$. Variation of constants and (68) give

$$\|e(t)\|_\infty \leq e^{-(1-\gamma)t/\tau_h} \|e(0)\|_\infty + \frac{\tau_h}{1-\gamma} \sup_{0 \leq s \leq t} \|\dot{h}(s)\|_\infty. \tag{71}$$

This states the required timescale relation in measurable quantities. A changing baseline additionally contributes $-\tau_u \dot{r}^0$ before decoding, which can be propagated by the same linear response.

A single recalled goal. The relation to a clamped navigation field is particularly direct. Let e_g be the indicator vector for a terminal state g and take $b = e_g$, understood as the limit of positive terminal factors. Then $h_i = G_{ig} > 0$. For the first hitting time $\tau_g = \inf\{k \geq 0 : X_k = g\}$ under P, define $z_i^g = \mathbb{E}_i[\gamma^{\tau_g}]$. Splitting the discounted number of visits to g at its first visit and applying the Markov property gives

$$G_{ig} = \mathbb{E}_i[\gamma^{\tau_g}] G_{gg}, \qquad z_i^g = \frac{G_{ig}}{G_{gg}}, \qquad z_g^g = 1. \tag{72}$$

The normalized field therefore solves $z_i^g = \gamma \sum_j P_{ij} z_j^g$ at every $i \neq g$. A circuit clamped at the goal to $u_g = \bar{u}_g$ has

$$u_i = \bar{u}_g \sqrt{\frac{\pi_i}{\pi_g}} \, z_i^g, \tag{73}$$

which is the linear network relation established by He et al. (2026b). This relation identifies the same field shape in a clamped circuit and in the source-driven circuit used here. The event model also specifies what happens when the sampled state reaches the goal. With $b = e_g$, each visit to g completes the event with probability $s_g = 1/G_{gg}$; all other states have $s_i = 0$. An event that continues makes an excursion and returns to the goal before completing. The factor G_{gg} counts the reference weight of these returns. Equivalently, the unit-clamped field $h = z^g$ has source $b = e_g/G_{gg}$, since $(I - \gamma P) z^g = e_g/G_{gg}$. The source and field determine completion together. Their common scaling leaves the whole path law unchanged.

The distinction between arrival and completion can be resolved explicitly. Let $\tau_g^+ = \inf\{k \geq 1 : X_k = g\}$ be the first return under P, and write $r_g = \mathbb{E}_g[\gamma^{\tau_g^+}]$. Splitting discounted visits into the initial visit and visits after the first return gives $G_{gg} = 1 + r_g G_{gg}$. Therefore

$$r_g = 1 - G_{gg}^{-1}, \qquad \Pr_{\mathsf{Q}}(N_g = m \mid X_0 = i) = G_{gg}^{-1}(1 - G_{gg}^{-1})^{m-1}, \quad m \geq 1. \tag{74}$$

Here N_g counts visits to g before completion, and the expression holds for every starting state i. The sampler reaches g almost surely, and each visit offers the same stopping probability; the strong Markov property gives the geometric law. In particular, $\mathbb{E}_{\mathsf{Q}}[N_g \mid X_0 = i] = G_{gg}$. Before the first arrival, telescoping the field ratios weights a reference path by γ^{τ_g}/z_i^g. This is the discounted first-arrival law associated with the clamped desirability field. The complete episode then includes the return excursions allowed by the map. Thus goal accessibility controls the initial approach, and recurrent return structure also controls the extent of the completed event.

Direction of field propagation. Field computation carries information about successors back to their possible predecessors: a reference transition $i \to j$ contributes $P_{ij}h_j$ to the field at i. Sequence generation subsequently propagates activity from i to j. In a symmetric recurrent graph, the same reciprocal weights support both operations, and $(Ph)_i = d_i^{-1}\sum_j W_{ij}h_j$ can be implemented by weighted input followed by state-specific normalization. For a directed map, the circuit additionally needs feedback pathways or a separately learned operator that supplies this propagation of terminal information. The orientation is determined by the conditional expectation represented by h.

A compact population representation. A lower-dimensional implementation uses a decoder $B \in \mathbb{R}^{N\times m}$ and dynamics $\tau\dot{u} = -u + e + \gamma Au$, with $m < N$. At equilibrium, $\widehat{h} = B(I - \gamma A)^{-1}e$. Exact implementation on the represented subspace follows from

$$Be = b, \qquad BA = PB, \qquad \rho(\gamma A) < 1. \tag{75}$$

Indeed, multiplying the equilibrium equation by B gives $(I - \gamma P)\widehat{h} = b$. For approximate alignment, the residual is

$$r := b + \gamma P\widehat{h} - \widehat{h} = (b - Be) + \gamma(PB - BA)u. \tag{76}$$

This equation separates approximation of the terminal cue from mismatch in recurrent propagation. A useful reduced implementation represents the terminal fields required by the task family. Equation (75) specifies alignment of the recurrent action on those fields, including their spatial structure and response to input.

Nonlinear integration. Suppose a circuit is linearized around an operating point and its exact steady state satisfies

$$\widehat{h} = b + \gamma P\widehat{h} + \delta(\widehat{h}), \tag{77}$$

where δ collects the omitted nonlinear terms and input mismatch. Then $\widehat{h} - h = G\delta(\widehat{h})$ and

$$\|\widehat{h} - h\|_\infty \le \frac{\|\delta(\widehat{h})\|_\infty}{1 - \gamma}. \tag{78}$$

For (62), the nonlinear term in decoded field coordinates is

$$\delta(v) = \frac{1}{\ell}\Pi^{-1/2}\big(g_s S - g_{\mathrm{inh}} I\big)R(\Pi^{1/2}v). \tag{79}$$

Let $\pi_{\min} = \min_i \pi_i$ and $\pi_{\max} = \max_i \pi_i$. Combining the Taylor bound with the matrix norm gives

$$\|\delta(v)\|_\infty \le \frac{M\pi_{\max}\big(g_s\|S\|_\infty + g_{\mathrm{inh}}\big)}{2\ell\sqrt{\pi_{\min}}}\|v\|_\infty^2. \tag{80}$$

This bound quantifies the field error while the activity remains within the specified operating range. If the local Lipschitz constant of δ on an invariant region is less than $1 - \gamma$, the steady-state map $v \mapsto b + \gamma Pv + \delta(v)$ is a contraction there and has a unique stable equilibrium. For heterogeneous slopes, let $D_\rho = \mathrm{diag}(\rho_i)$ and $D_\ell = I + g_{\mathrm{inh}}D_\rho$. The first-order steady-state operator becomes $K = g_s D_\ell^{-1} S D_\rho$, with source $D_\ell^{-1} I^{\mathrm{goal}}$. The positive resolvent $(I - K)^{-1}$ exists when $\rho(K) < 1$. Thus gain heterogeneity changes the effective propagation law; when its decoded rows have sums below one, it gives the state-dependent survival construction in Appendix A.8.

A controlled operating regime for a saturating circuit. The scale freedom of the path law gives a useful way to control nonlinear error. Let terminal input have the form $I_\epsilon^{\mathrm{goal}} = \epsilon \ell \Pi^{1/2} b$, where $b > 0$ fixes its shape and $\epsilon > 0$ fixes its amplitude. The decoded equilibrium then obeys

$$v_\epsilon = \epsilon b + \gamma P v_\epsilon + \delta(v_\epsilon), \qquad \delta(0) = 0, \quad D\delta(0) = 0. \tag{81}$$

For a twice continuously differentiable activation function, the derivative with respect to v of the steady-state equation at $(v, \epsilon) = (0, 0)$ is $I - \gamma P$, which is invertible. The implicit function theorem therefore gives a nearby equilibrium with

$$v_\epsilon = \epsilon h + O(\epsilon^2), \qquad u_\epsilon = \epsilon \Pi^{1/2} h + O(\epsilon^2). \tag{82}$$

The baseline linearization has eigenvalues with negative real parts, so this equilibrium is locally asymptotically stable for sufficiently small ϵ. Since $\min_i h_i > 0$, its field is also positive for sufficiently small ϵ.

Dividing the equilibrium response by its common amplitude gives $v_\epsilon/\epsilon = h + O(\epsilon)$. The neural competition below normalizes all scores locally, so it performs this common rescaling implicitly. Its normalized probabilities approach the exact sampler as ϵ decreases. The positive stopping probabilities remain bounded away from zero in a neighborhood of the limit. Finite-path probabilities therefore converge with a common geometric tail bound, giving convergence of the complete path law in total variation. An affine activation function over the active range gives the exact law at finite amplitude; a smooth saturating response gives this controlled approximation. The common rate gain can compensate for amplitude when a fixed physical clock is desired.

A.6.2 Dendritic gain, competition, and completion

Probabilistic competition. At state i, assign rates $\lambda_{ij} = c_i \gamma P_{ij} h_j$ to continuation candidates and $\lambda_{i\dagger} = c_i b_i$ to termination, with $c_i > 0$. Each candidate is an activation event in a Poisson process while the current state and the field are held fixed. The first activation recruits inhibition that suppresses the remaining candidates and resets the competition. The independent activation times are exponential, so the probability that candidate j occurs first is

$$\int_0^\infty \lambda_{ij} e^{-t(\lambda_{i\dagger} + \sum_l \lambda_{il})}\, dt = \frac{\lambda_{ij}}{\lambda_{i\dagger} + \sum_l \lambda_{il}} = T_{ij}. \tag{83}$$

The analogous integral for termination gives s_i. The same competition with initial rates proportional to $\pi_i a_i h_i$ yields q_0. These stochastic activation laws, gain products, and inhibitory resets define the neural sampling model. Each is local to the current assembly and its candidates. Their combination yields the complete path distribution through Proposition 2.

The role of recurrent normalization becomes explicit when synaptic strengths are represented by W, with $d_i = \sum_j W_{ij}$ and $P_{ij} = W_{ij}/d_i$. Equivalent candidate scores are $\gamma W_{ij} h_j$ for continuation and $d_i b_i$ for completion. Their total is $d_i h_i$, which gives the same T and s after normalization. The factor d_i cancels when comparing two successors. Its relative contribution to continuation and completion remains relevant to the complete event. This explains why recurrent gain normalization affects sequence endings even when conditional successor probabilities are preserved.

The symmetric neural coordinate in (66) admits a direct implementation using the field activity u. At current assembly i, let recurrent input to candidate j be proportional to the physical outgoing weight S_{ji}. Multiply it by that candidate's maintained field activity u_j. Use the scaled terminal current β_i for the completion candidate. Detailed balance gives

$$\lambda_{ij}^{\mathrm{neural}} = c\gamma S_{ji} u_j = c\sqrt{\pi_i}\, \gamma P_{ij} h_j, \tag{84}$$

$$\lambda_{i\dagger}^{\mathrm{neural}} = c\beta_i = c\sqrt{\pi_i}\, b_i, \tag{85}$$

$$\lambda_{i\dagger}^{\mathrm{neural}} + \sum_j \lambda_{ij}^{\mathrm{neural}} = c u_i = c\sqrt{\pi_i}\, h_i. \tag{86}$$

Every score contains the same factor $\sqrt{\pi_i}$ at the current state. Competition cancels it and gives exactly T_{ij} and s_i. The physical recurrent weights, maintained field activity, and terminal current thus suffice for sampling. The circuit obtains the field ratios by competition among activation events. Initiation uses scores proportional to $\sqrt{\pi_i} a_i u_i = \pi_i a_i h_i$. A continuation winner recruits the next assembly and resets the competition; a completion winner terminates the event.

In the saturating regime of (82), use u_ϵ and $\beta_\epsilon = \epsilon\Pi^{1/2}b$ in these same rates. Each nonzero continuation score is its ideal value times $\epsilon[1+O(\epsilon)]$, and the completion score is its ideal value times ϵ. Normalization yields the stated convergence of sequence probabilities. Choosing the common rate gain c/ϵ also makes the waiting-time law converge to the ideal circuit clock.

Physical duration. Multiplying every competing rate at a state by the same positive number preserves its choice probabilities and changes its waiting time. In decoded coordinates the total decision rate is $\Lambda_i = c_i h_i$; in (86) it is $\Lambda_i = cu_i$. Conditional on a live state, waiting time and the identity of the winning candidate are independent. Let D be the elapsed time after the initial state has been chosen, including the final completion decision. Its expected duration is

$$\mathbb{E}D = \sum_i \frac{n_i}{\Lambda_i}. \tag{87}$$

Each of the expected n_i visits to state i contributes a mean waiting time $1/\Lambda_i$. The full duration law follows from the same first-step decomposition. For $z \geq 0$, let $\psi_i(z) = \mathbb{E}_i[e^{-zD}]$ and $D_z = \mathrm{diag}(\Lambda_i/(\Lambda_i + z))$. Then

$$\psi(z) = D_z\big(s + T\psi(z)\big), \qquad \mathbb{E}[e^{-zD}] = q_0^{\mathsf{T}}(I - D_z T)^{-1} D_z s. \tag{88}$$

The inverse exists because $D_z T$ is dominated entrywise by the transient matrix T. At $z = 0$ the transform equals one; differentiation gives (87). Uniformly scaling terminal input and the resulting field changes these waiting times when the common rate gain is fixed, while preserving state order and transition counts. A gain change restricted to the completion channel changes both content and length. With a fixed sampling interval Δt, duration is instead represented by $(K+1)\Delta t$; the length law in (47) then gives its distribution directly.

Temporal gating across bursts. The same separation holds when all candidates share a time-dependent gate. During a visit to state i, let candidate j have hazard $g_i(t)\lambda_{ij}$, including the completion candidate among the values of j. The nonnegative gate g_i is externally specified and shared by every candidate. Write $A_i(t) = \int_0^t g_i(v)\,dv$ and $\Lambda_i = \sum_j \lambda_{ij}$. If $A_i(t) \to \infty$, then

$$\Pr(j \text{ wins} \mid i) = \int_0^\infty g_i(t)\lambda_{ij} e^{-\Lambda_i A_i(t)}\,dt = \frac{\lambda_{ij}}{\Lambda_i}. \tag{89}$$

The substitution $v = A_i(t)$ gives the last equality. Intervals with $g_i = 0$ pause the competition. If the active represented state, terminal evidence, and continuation field persist, resuming the gate continues the same episode with its original choice probabilities. Its physical duration changes, and its path law is preserved. This provides a possible relation between an extended represented sequence and several oscillatory bursts. A reset of the active state instead invokes a new initiation distribution. Candidate-specific gating or changes in the maintained field can change the path law; their effects are measured through the relative hazards.

Multiplicative gain and endpoint readout. The candidate score $P_{ij}h_j$ makes the effect of reference input depend on the postsynaptic field. Nonlinear integration and dendritic spikes in CA3 provide relevant physiological operations (Makara & Magee, 2013; Brandalise et al., 2016). Their role in the proposed computation can be assessed by measuring whether conditional choice probabilities follow the product over the operating range. The termination pathway responds to the current state's terminal compatibility b_i and competes with the summed continuation score. Assigning these operations to particular projections specifies where their inputs, gain interactions, and consequences for sequence statistics can be measured.

Represented states and memory. An exact coarse representation of microscopic dynamics is Markov when transition probabilities into each represented state are the same for every microscopic configuration assigned to the current state. Approximate separation of rapid internal relaxation and slower state changes can make this condition useful over the event timescale. If a location alone does not predict its next state, the represented state can include heading, recent context, or other relevant memory. The path construction then uses the enlarged transition kernel. This preserves a clear relation between the information retained by a state and the sequences the model can generate.

A.7 Learning the endpoint factors from wake and replay

The local sampler produces the path law associated with its current factors. Learning can adjust those factors so that the frequency of its endpoints agrees with experience. Let $\theta_i = \log a_i$ and $\eta_i = \log b_i$. With $M = \mathrm{diag}(\pi)G$, define the concave objective

$$\mathcal{J}(\theta, \eta) = \mu_0^{\mathsf{T}}\theta + \mu_T^{\mathsf{T}}\eta - \log \sum_{ij} e^{\theta_i} M_{ij} e^{\eta_j}. \tag{90}$$

Using C in place of M changes this objective by a constant. The two marginals and gradients are

$$q_0(i) = \frac{a_i \pi_i (Gb)_i}{Z}, \qquad \frac{\partial \mathcal{J}}{\partial \theta_i} = \mu_0(i) - q_0(i), \tag{91}$$

$$q_T(j) = \frac{b_j [G^{\mathsf{T}}(\pi \odot a)]_j}{Z}, \qquad \frac{\partial \mathcal{J}}{\partial \eta_j} = \mu_T(j) - q_T(j). \tag{92}$$

The derivative of the logarithmic normalizer is the corresponding model endpoint frequency. A wake sample from the target marginal supplies a positive endpoint indicator. A replay sample supplies a negative endpoint indicator. Their difference is an unbiased estimate of the gradient when the two samples have the stated distributions.

More explicitly, for experience endpoints $(I^{\mathrm{w}}, J^{\mathrm{w}})$ and replay endpoints $(I^{\mathrm{r}}, J^{\mathrm{r}})$, use

$$\theta_{n+1,i} = \theta_{n,i} + \alpha_n \big[\mathbf{1}\{I^{\mathrm{w}} = i\} - \mathbf{1}\{I^{\mathrm{r}} = i\}\big], \tag{93}$$

$$\eta_{n+1,j} = \eta_{n,j} + \alpha_n \big[\mathbf{1}\{J^{\mathrm{w}} = j\} - \mathbf{1}\{J^{\mathrm{r}} = j\}\big]. \tag{94}$$

Each update concerns the states participating at an event boundary. Correlation between the two endpoints of one event changes the covariance of the update and preserves its mean. For m independent replay events, the variance of an estimated starting frequency is $q_0(i)(1 - q_0(i))/m$, with an analogous expression for terminal frequencies. Correlated replay events require the corresponding covariance terms in the batch variance.

In synaptic variables, a small log-weight update gives $\Delta a_i \simeq \alpha_n a_i(\mu_0(i) - q_0(i))$, and similarly for b. This is a hypothesis of multiplicative plasticity tied to event boundaries. Uniform multiplication of all starting or all terminal weights leaves the normalized path law invariant. Calibration follows from changes linked to the particular states used as endpoints: experience strengthens their contribution, while replay supplies the frequency-dependent negative term. The rule specifies both the identity of the updated weights and the statistics carried by the update.

Curvature and parameter freedom. For any perturbation (u, v), differentiation once more yields

$$-(u, v)^{\mathsf{T}} \nabla^2 \mathcal{J}\,(u, v) = \mathrm{Var}_{\Gamma_{a,b}}(u_{X_0} + v_{X_K}) \geq 0. \tag{95}$$

Since $\Gamma_{a,b} > 0$ on every pair, the variance vanishes exactly when u and v are each constant. Multiplying a by any positive constant and multiplying b by any positive constant both leave $\mathsf{Q}_{a,b}$ unchanged after normalization. We remove these two freedoms by setting $\sum_i \theta_i = \sum_j \eta_j = 0$. On this subspace the objective is strictly concave. Its unique maximizer corresponds to the bridge of Proposition 1.

Alternating marginal fitting. The exact row and column corrections are

$$a_i \leftarrow a_i \frac{\mu_0(i)}{q_0(i)}, \qquad b_j \leftarrow b_j \frac{\mu_T(j)}{q_T(j)}, \tag{96}$$

with q_T recomputed after the first update. For example, the first update multiplies each row of the normalized endpoint coupling by $\mu_0(i)/q_0(i)$; its row sum becomes $\mu_0(i)$ and the total mass remains one. These are iterative proportional fitting steps (Csiszár, 1975; Sinkhorn & Knopp, 1967). They solve the same endpoint problem as the stochastic gradient rule. Their individual steps are different: the exact correction uses a log frequency ratio, whereas (94) averages indicator differences.

Convergence of the sampled update. Hold P, π, γ fixed. Assume positive targets and conditionally unbiased experience and replay endpoint samples with bounded second moments. Let $\alpha_n > 0$, $\sum_n \alpha_n = \infty$, and $\sum_n \alpha_n^2 < \infty$. Assume that the centered update stays in a compact convex subset of the gauge subspace containing the optimum in its relative interior, or apply Euclidean projection onto such a set after each update. These conditions specify a stationary task and a decreasing learning rate under which sampling fluctuations average out.

Proposition 3 (Endpoint calibration)**.** *Under these assumptions, the centered stochastic update converges almost surely to the maximizing factors of* (90)*. Its path law converges in total variation to the Schrödinger bridge with endpoint marginals* (μ_0, μ_T)*.*

Proof. Write $w = (\theta, \eta)$, let $w^\star$ be the centered optimum, and denote the unbiased sampled gradient by $\widehat{g}_n$. Continuity of (95) and compactness give a positive lower curvature bound λ on the gauge subspace. Thus

$$(w - w^\star)^\mathsf{T} \nabla \mathcal{J}(w) \le -\lambda \|w - w^\star\|^2. \tag{97}$$

Projection onto a convex set containing $w^\star$ cannot increase distance to $w^\star$. Conditional expectation of the squared-distance update therefore gives

$$\mathbb{E}[\|w_{n+1} - w^\star\|^2 \mid \mathcal{F}_n] \le (1 - 2\lambda\alpha_n)\|w_n - w^\star\|^2 + B\alpha_n^2 \tag{98}$$

for a finite second-moment bound B. The almost-supermartingale convergence lemma (Robbins & Siegmund, 1971) implies that the squared distance converges and that $\sum_n \alpha_n \|w_n - w^\star\|^2 < \infty$ almost surely. Since $\sum_n \alpha_n = \infty$, the limit is zero. Endpoint couplings depend continuously on finite parameters, so $\Gamma_n \to \Gamma^\star$. All corresponding path laws share the same conditional distributions R^{ij}. Hence

$$\|\mathsf{Q}_n - \mathsf{Q}^\star\|_{\mathrm{TV}} = \tfrac{1}{2} \sum_{ij} |\Gamma_{n,ij} - \Gamma^\star_{ij}| \longrightarrow 0. \tag{99}$$

□

This result concerns calibration on a fixed reference map. If the map changes slowly, the same rule can track moving endpoint solutions when their drift is small relative to the effective learning rate. A constant learning rate supports tracking with persistent sampling fluctuations. Learning restricted factors, such as $\theta_i = w^\mathsf{T}\phi(i)$, replaces complete marginal matching by matching the corresponding endpoint features: the gradient becomes $\mathbb{E}_{\mu_0}\phi - \mathbb{E}_{q_0}\phi$.

A.8 Approximation errors and extensions

Residual control of the field. For any positive approximate field $\widehat{h}$, define $r = b + \gamma P\widehat{h} - \widehat{h}$. The identity $\widehat{h} - h = -Gr$ gives

$$\|\widehat{h} - h\|_\infty \le \frac{\|r\|_\infty}{1 - \gamma}. \tag{100}$$

For a truncated recursion $h^{(L)} = \sum_{k=0}^{L} (\gamma P)^k b$,

$$\|h - h^{(L)}\|_\infty \le \frac{\gamma^{L+1}\|b\|_\infty}{1 - \gamma}. \tag{101}$$

Both bounds follow from the nonnegative geometric series. They connect a computable recursion error to the representational accuracy required by the sampler.

A sampler with an approximate field. A local competition normalizes its candidate scores. With $\widehat{h}$, the total score at state i is $S_i = b_i + \gamma(P\widehat{h})_i = \widehat{h}_i + r_i$. Accordingly, use

$$\widehat{T}_{ij} = \frac{\gamma P_{ij}\widehat{h}_j}{S_i}, \quad \widehat{s}_i = \frac{b_i}{S_i}, \quad \widehat{q}_0(i) = \frac{\pi_i a_i \widehat{h}_i}{\widehat{Z}}, \qquad \widehat{Z} = \sum_i \pi_i a_i \widehat{h}_i. \tag{102}$$

The generated path law satisfies the exact identity

$$\frac{\widehat{\mathsf{Q}}(\omega)}{\mathsf{Q}(\omega)} = \frac{Z}{\widehat{Z}} \prod_{k=0}^{K} \frac{\widehat{h}_{x_k}}{\widehat{h}_{x_k} + r_{x_k}}. \tag{103}$$

To derive it, multiply the local probabilities and cancel successive numerator factors $\widehat{h}$. Thus an inconsistent field changes path weights through the states visited, and its effect can accumulate with sequence length. If $|r_i| \le \epsilon \widehat{h}_i$ with $0 < \epsilon < 1$, each logarithmic factor in (103) has absolute value at most $-\log(1-\epsilon)$.

Errors in competing scores. Suppose every nonzero candidate score, including initiation and termination, is multiplied by a factor in $[1-\epsilon, 1+\epsilon]$. After normalization, the ratio of any perturbed local probability to its intended value lies between $(1-\epsilon)/(1+\epsilon)$ and its reciprocal. Hence the local relative entropy is at most

$$c_\epsilon = \log \frac{1+\epsilon}{1-\epsilon}. \tag{104}$$

The chain rule for a stopped process gives

$$D_{\mathrm{KL}}(\mathsf{Q} \| \widehat{\mathsf{Q}}) \le c_\epsilon \big(\mathbb{E}_{\mathsf{Q}} K + 2 \big). \tag{105}$$

There is one initiation choice and $K+1$ local choices including the final stop. The bound assumes common support and the stated multiplicative control at each visited state. Pinsker's inequality then bounds total variation by the square root of half the right-hand side. This is a sufficient error bound; direct comparisons can be sharper.

Errors in the transition map and state representation. For another row-stochastic kernel $\widehat{P}$, with the same b and γ, let $\widehat{h} = (I - \gamma \widehat{P})^{-1} b$. The resolvent identity gives

$$\|\widehat{h} - h\|_\infty \le \frac{\gamma \|\widehat{P} - P\|_\infty \|b\|_\infty}{(1-\gamma)^2}. \tag{106}$$

If a microscopic process has history-dependent coarse transitions within total variation ϵ of the proposed kernel at every step, maximal coupling controls the discrepancy through the number of transitions observed. For a fixed L, the probability of a discrepancy before step L is at most $L\epsilon$. Extending the bound to a stopped process adds its stopping and initiation discrepancies and averages the per-step error over the reference event length.

State-dependent survival. Let $\widetilde{P}_{ij} = \gamma_i P_{ij}$ with $0 \le \gamma_i < 1$, and let $t_i = 1 - \gamma_i$. The normalized reference is $\pi_{x_0} \prod_{k<K} \widetilde{P}_{x_k x_{k+1}} t_{x_K}$. With endpoint factors a, b, define

$$H = (I - \widetilde{P})^{-1} (t \odot b). \tag{107}$$

The transformed initial law is proportional to $\pi_i a_i H_i$, the transition probability is $\widetilde{P}_{ij} H_j / H_i$, and the stopping probability is $t_i b_i / H_i$. Their sum equals one by the defining equation for H. When $\gamma_i = \gamma$, one has $H = (1-\gamma)h$, which recovers the main model. Thus heterogeneous stopping has the same local structure, with the reference terminal factor retained at each state.

Intermediate observations. When a task provides evidence at successive steps of a fixed-length sequence, let $\ell_k(i) > 0$ be its likelihood at state i and index $k = 0, \dots, L-1$. Include evidence at the final state in b. The path weight is then proportional to $\pi_{x_0} a_{x_0} b_{x_L} \prod_{k<L} [\ell_k(x_k) P_{x_k x_{k+1}}]$. Resolving the suffix weights gives $h_L = b$ and $h_k(i) = \ell_k(i) \sum_j P_{ij} h_{k+1}(j)$. The initial law is proportional to $\pi_i a_i h_0(i)$, and the transition at $k < L$ is $P_{ij} h_{k+1}(j) / \sum_l P_{il} h_{k+1}(l)$. These are the forward and backward recursions used in sequence inference. With fixed L, the reference survival factor γ^L is common to all paths and cancels. The endpoint model concerns the case in which evidence specifies the beginning and completion of an event whose length is sampled.

A.9 Quantitative predictions from the same construction

The following relations can be evaluated with a reference kernel estimated independently of the events used to test them.

Changes in successor odds. For two available next states j, l from state i,

$$\log \frac{F_{ij}}{F_{il}} = \log \frac{P_{ij}}{P_{il}} + \log h_j - \log h_l. \tag{108}$$

Changing the terminal cue from b to b' changes the log odds by $\log(h'_j/h_j) - \log(h'_l/h_l)$. This change is shared across predecessor states that offer the same two successors. It is a stronger test than an increase in movement toward a goal.

Coupled changes in movement and stopping. The same field that enters choice also determines

$$\frac{s_i}{1 - s_i} = \frac{b_i}{\gamma(Ph)_i}. \tag{109}$$

The same field and terminal evidence therefore predict successor choices and completion. Joint fits of sequence content, endpoint frequencies, and length distributions test this coupling. Uniformly rescaling a or b leaves all normalized predictions unchanged. Within this model, changes in event statistics arise through statewise contrast, reference propagation, survival, or the relative gains of competing channels.

The response to a selective change in one possible ending is particularly informative. Recall from (34) that $B_{i\ell} = G_{i\ell} b_\ell / h_i$ is the probability of eventually completing at ℓ from current state i. Differentiating the stopping rule and the successor odds gives

$$\frac{\partial \log s_i}{\partial \log b_\ell} = \mathbf{1}\{i = \ell\} - B_{i\ell}, \qquad \frac{\partial}{\partial \log b_\ell} \log \frac{F_{ij}}{F_{ik}} = B_{j\ell} - B_{k\ell}. \tag{110}$$

These relations hold for positive terminal factors and available successors. Increasing the evidence for a remote ending reduces stopping at the present state in proportion to that ending's accessibility. At the emphasized ending itself, local stopping increases by a logarithmic sensitivity $1 - B_{\ell\ell}$. The same change favors one successor over another according to their difference in eventual access to that ending. A selective terminal perturbation consequently predicts three linked measurements: field modulation, changes in successor odds, and changes in local completion. Their relation follows from the same transition map and supplies a test of the circuit computation without separate response coefficients for the three effects.

Endpoint exchange and route composition. In a reversible reference map, exchanging start and terminal factors predicts the reversed path distribution through (61). With the two endpoints held fixed, changing only their factors preserves route probabilities through (24). These results separate boundary selection from changes in the learned transition structure.

Experience-dependent calibration. The expected log-weight change of an endpoint state is proportional to its excess frequency in experience relative to replay. Overrepresented replay endpoints should consequently receive stronger negative updates under the proposed plasticity rule. The relevant signal is endpoint identity and frequency. Overall activity downscaling can coexist with this rule, while endpoint selectivity determines its calibration effect.

Online occupancy and phase-resolved sequences. The fields $f_k \odot h$ predict state occupancy among events that reach step k. Mapping successive propagation steps to phases of an oscillatory cycle supplies a possible mechanism for ordered sweeps. As the starting cue approaches a represented location along the reference transitions, fewer propagation steps are needed to reach it. A model relating propagation latency to oscillatory phase then translates this change into a prediction for phase precession. State and step statistics can first be estimated from sequences; the timing model determines how those statistics appear within an oscillatory cycle.

Spontaneous sequences and new route combinations. Flat boundary factors recover the reference process. Sharpening either factor changes the subset of reference sequences expressed in a given context. A Markov map can also generate a route assembled from familiar transitions even when that complete route has never been experienced. Its probability is determined by the learned transition product and stopping factor. For preplay in a new environment, the relevant comparison follows the assignment of earlier assembly sequences to their later spatial or contextual meaning.

A.10 INTERPRETATION ACROSS HIPPOCAMPAL COMPUTATIONS

The same model variables have different assignments in different cognitive tasks. These assignments connect the mathematical predictions to experiments. The reference process represents established relations among situations; the starting factor specifies the conditions from which a sequence is generated; the terminal factor specifies the conditions under which it counts as complete. The survival parameter controls the reference distribution of transition counts. This description makes the role of a task boundary explicit even when its represented state is nonspatial.

Memory and imagination. During retrieval, a cue can concentrate the starting factor on a set of assemblies associated with an experience. A broad terminal factor permits multiple completions, while a specific question or contextual cue can concentrate the expected endpoint. Sampling then generates complete sequences compatible with both the cue and the learned transitions. When different experiences share represented states, their transitions can recombine. The resulting sequence can be coherent under the learned map even when that complete sequence has never occurred. This supplies a computational account of constructive retrieval and a source of memory errors within the same mechanism.

Imagining a future event uses the same operation with a prospective interpretation of the terminal factor. A destination or desired situation first changes b; propagation through the reference map produces h; local choices favor states that support that prospective completion. Retrieval and imagination use evidence with different meanings and can recruit sequences with different histories. The sampling rule accommodates both familiar routes and new combinations of learned transitions. Context retained in the represented state preserves the distinctions between particular episodes and determines which transitions can be combined during retrieval.

Generation from an informative reference. Schrödinger bridges also provide a construction for generative models. Diffusion Schrödinger Bridge uses iterative proportional fitting to learn a stochastic process joining a tractable distribution and a data distribution (De Bortoli et al., 2021). Image-to-Image Schrödinger Bridge applies related diffusions to restoration, using the structure retained in a degraded image to constrain reconstruction (Liu et al., 2023). These examples show how a reference process and endpoint information can jointly specify generated content. In the hippocampal account, the reference consists of learned transitions among represented situations, and a generated path expresses their possible succession. In an image model, the path indexes stages of an algorithm producing an image. The mathematical connection concerns generation under constraints on a reference process; the represented states and the meaning of each transition are specified by the application.

Forward, reverse, and nonlocal replay. For prospective replay, the starting factor may encode the current state and the terminal factor a possible destination, an assignment motivated by goal-directed hippocampal sequences (Pfeiffer & Foster, 2013). For retrospective replay, the starting factor may instead encode the recently reached outcome and the terminal factor a preceding context, connecting the model to reverse replay and its reward dependence (Foster & Wilson, 2006; Diba & Buzsáki, 2007; Ambrose et al., 2016). In a reversible reference map, the exact endpoint-exchange identity predicts reversed sequence statistics. In an asymmetric map, the reference transition asymmetry contributes additional directionality. Experiments that manipulate starting evidence and prospective evidence independently can measure these two contributions.

Nonlocal replay follows when internally maintained context concentrates the factors on states away from the animal's current location. The decoded position of an event can therefore differ from the source of its initiating signal, consistent with the range of trajectories expressed during replay (Gupta et al., 2010). Preplay can be investigated by estimating the spontaneous reference sequences before a new experience and then measuring how new sensory assignments associate those sequences with an environment (Dragoi & Tonegawa, 2011). The theory predicts reference sampling under flat factors; the relation between an internally represented sequence and a future spatial route also depends on that later assignment.

Sleep and recent experience. Reduced sensory constraint during sleep permits internally determined boundary factors to play a larger role. Recent experience can influence sampling by changing

the factors, the starting accessibility distribution, or the recurrent map. These possibilities have different signatures. A factor change preserves the route distribution conditional on endpoints; a map change can modify that conditional distribution. If a change in physiological state increases effective survival, (48) predicts longer sequences for fixed factors, together with slower field relaxation. Temporal compression alone can arise through faster common race rates while preserving the distribution of transition counts.

These relations organize a study of sleep sequences around their initiation, internal transitions, and physical duration. NREM replay supplies direct measurements of these quantities. A study of dream construction could align the same sequence statistics with prior experience, sleep stage, and subjective report, testing how a change in boundary input affects the episodes that are generated and experienced.

Goals, context, and online population activity. The occupancy identity $n_i = f_i h_i / Z$ expresses how activity depends jointly on the starting context and possible completion. A goal can change firing in states that support routes to it, including states shared by several routes. Context-dependent activity at the same physical location follows when the represented state distinguishes the relevant past or intended future. This provides a natural representation of splitter activity: spatially overlapping assemblies can occupy different states in the map and receive different forward or terminal weights.

For a fixed initial cue, replacing b by b' changes occupancy according to $n_i'/n_i = (h_i'/h_i)(Z/Z')$. A measured change in the field therefore makes a quantitative prediction for the distribution of population visits. The position of the largest change depends on both accessibility from the start and compatibility with the goal. An observation model translates these visit statistics into firing rates through the membership of each assembly and the time spent in each state.

Sequence completion and plasticity. A terminal event supplies a useful learning signal because it identifies where an internally generated sequence has ended. Behavioral time scale synaptic plasticity in CA1 provides an example of a dendritic event producing synaptic changes over an extended temporal window (Bittner et al., 2017). In a circuit implementation with a completion-sensitive dendritic event, a preceding eligibility trace could associate recently active inputs with that completion. If an eligibility trace at time t has the form $e_j(t) = \int_0^\infty k(u) x_j(t-u)\, du$, a completion-triggered update proportional to $e_j(t)$ preferentially modifies states represented shortly before the endpoint. The shape of the biological kernel k determines the temporal and spatial extent of the change. This connects the proposed endpoint readout to a testable sequence of neural activity and later synaptic modification, while keeping the sampling rule and the plasticity rule separately measurable.

Perturbing apical input and the stopping channel. The circuit proposal distinguishes the terminal evidence b, the propagated field h, and the interaction that combines h with recurrent input. Suppressing contrast in the propagated field makes conditional next-state choices more closely reflect P. If the same perturbation leaves terminal evidence intact, the stopping channel still carries information about completion. Disrupting terminal evidence affects both its propagated field and the local stopping score. These manipulations can consequently be distinguished by fitting successor choices and termination jointly.

NMDA-dependent integration and cholinergic modulation provide candidate interventions (Brandalise et al., 2016; Kis et al., 2024). Multiplying continuation rates by $c > 0$ while holding termination fixed gives stopping odds $b_i/[c\gamma(Ph)_i]$ and preserves relative successor choices. Joint measurements of field contrast, dendritic gain, recurrent propagation, and stopping identify the affected operation. Altered duration can be tested alongside preserved directional preferences. Scaling terminal evidence and its propagated field together leaves the path law unchanged. Apical disruption tests dependence between task cues and sequence content, while spontaneous activity estimates the reference transitions.

B COGNITIVE MAPS, SEQUENCE GENERATION, AND HIPPOCAMPAL FUNCTION

B.1 FROM REPRESENTED SITUATIONS TO POSSIBLE EXPERIENCE

A cognitive map relates each represented situation to other situations. Predictive maps describe these relations through expected future occupancy (Dayan, 1993; Stachenfeld et al., 2017). The Tolman–Eichenbaum machine learns relational structure that can be reused with new sensory associations (Whittington et al., 2020). Clone-structured graphs distinguish latent states that share an observation but have different sequential contexts (George et al., 2021; Raju et al., 2024). These accounts specify information that a path model also needs: two visits to the same location can lead to different continuations when direction, recent history, or task context differs.

This organization extends beyond physical space. During rest after a decision task, human hippocampal activity reflects the order of abstract task states (Schuck & Niv, 2019). In another study, people learned an ordering rule and then saw new objects in a scrambled order; subsequent replay followed the learned rule (Liu et al., 2019). A reference path measure expresses the regularities that make such sequences coherent. Its states carry the distinctions needed by the task, and its transitions express which represented situations can follow one another. Learning a map therefore establishes a repertoire of possible experiences, including combinations that have yet to occur.

Attractor networks explain the stability of a representation. Recurrent interactions restore a population pattern after perturbation (Hopfield, 1982; Wills et al., 2005); hippocampal CA3 supports the recovery of coherent representations from incomplete input (Neunuebel & Knierim, 2014; Guzman et al., 2016). During sharp-wave ripples, recordings have been interpreted as successive periods of convergence toward a represented location followed by transitions to another (Pfeiffer & Foster, 2015). This connects associative retrieval with sequence generation. An attractor identifies a represented situation, and the probabilities of moving among situations specify the experiences the network can generate. Architectures that associate episodic content with a spatial scaffold supply one way to give these states sensory meaning (Chandra et al., 2025).

Endpoint input changes which transitions the network expresses. Its influence at a junction depends on the continuations reachable through the map. A nearby goal behind a bottleneck can be less accessible than a more distant goal connected by several likely routes. The continuation field combines these alternatives, weighting every route by its learned probability and length. For the reversible diffusion in Appendix A.5, its logarithm determines the change in potential. The resulting landscape retains the established relations among states while changing the relative ease of moving through them. Stability gives each state a recoverable identity; the conditioned dynamics give an episode its direction.

Directed events can also arise on a reversible map. Initiation supplies probability at some states, and completion removes it at others; the ensemble of finite sequences consequently carries net flow between their endpoints. A task can thus reorganize directional replay while preserving reciprocal recurrent structure. Place-field stability also leaves room for changes in transition structure: replay reroutes around relocated barriers while place fields remain largely stable (Widloski & Foster, 2022). The model gives a test for locating the change. Endpoint selection preserves relative route and length probabilities after both endpoints are fixed. Changes in those conditional statistics identify changes in the reference dynamics or in information applied within the path.

B.2 CONSTRUCTING AND COMPLETING AN EPISODE

Episodic recall and imagination both involve organizing familiar elements into a coherent situation. Patients with hippocampal amnesia can have difficulty constructing novel experiences with a consistent spatial setting (Hassabis et al., 2007). Rats can also sustain remote place representations voluntarily in a hippocampal brain–machine interface (Lai et al., 2023). Generative models explain how memory reconstruction combines particular experiences with learned regularities (Spens & Burgess, 2024), while predictive learning provides a way to acquire hippocampal representations from temporal structure (Chen et al., 2024).

The path formulation identifies a common inference in these functions. A cue supplies evidence about a beginning, and a remembered context or desired outcome supplies evidence about a completion. Learned transitions connect them. Recall constrains this construction with information about a particular past episode. Planning constrains it with information about an intended outcome. Imagination

explores paths compatible with a partially specified situation. The generated sequence can be new as a whole because familiar transitions can be combined in an unexperienced order. Reconstructing a particular episode requires the contextual distinctions that separate it from other episodes sharing the same places or objects. Compositional models address how replay can acquire and combine the underlying structure (Bakermans et al., 2025). Human intracranial recordings associate hippocampal ripples with sequences that recombine familiar elements and with corresponding changes in prefrontal representations during inference (He et al., 2026a).

The input representing an outcome is present before the imagined sequence unfolds. Its propagation through the map computes how that outcome could be reached. A represented future can therefore organize present neural activity through an ordinary causal process: terminal input changes recurrent activity, that activity changes local competition, and successive winners construct an episode. The source of the input gives the episode its interpretation. A remembered ending supports reconstruction; an intended ending supports prospective simulation. Both draw on the same learned relations, and both retain the uncertainty among the paths those relations permit.

Completion contributes to the content of an episode. The same represented state may be an intermediate point in one sequence and its ending in another. The stopping rule expresses this difference by comparing support for completion now with support for further compatible continuations. Even when terminal evidence at the current state is held fixed, adding an accessible continuation can make stopping less likely. Event extent thus depends on both the current outcome and the remaining possibilities. Human medial temporal neurons respond to cognitive boundaries within narratives, and changes in population state at boundaries predict later recognition and temporal-order memory (Zheng et al., 2022). At a broader scale, transitions between topics during spontaneous thought recruit posterior medial regions of the default and control networks, with spatial activation patterns resembling those at external event boundaries (Su et al., 2025). These observations locate event organization at cellular and distributed network levels. The completion rule supplies a specific computational prediction: the probability of ending at a represented situation depends on the alternative completions accessible from it. Aligning decoded sequence endings with boundary responses tests this dependence while holding the local situation fixed.

Preplay concerns how an available neural sequence acquires the meaning of a later experience. Correlations between sequences before exploration and later place-cell order have been reported (Dragoi & Tonegawa, 2011); other recordings find that trajectory structure emerges with experience (Silva et al., 2015). Recent detour experiments connect rapidly emerging theta sequences to motifs expressed during preceding sleep (Zhou et al., 2025). The reference formulation accommodates pre-existing dynamics and changes learned from experience. Its account of preplay requires an explicit association between internal states and the spatial or contextual variables subsequently observed. Conditioning then determines which of these available sequences a current cue selects.

Variability across generated trajectories can express uncertainty about possible experience. Ujfalussy & Orbán (2022) compared several neural codes and found that variability across theta cycles supported sequential sampling of possible motion trajectories. Their distinction between uncertainty in the represented trajectories and uncertainty in decoding spikes also applies here. A decoder's posterior describes what the recorded activity allows an observer to infer; variation among generated episodes describes the alternatives represented by the circuit. The bridge further distinguishes uncertainty about endpoints from variation in paths joining a specified pair. Sharpening endpoint evidence changes the mixture of endpoint pairs. Among paths with the same pair, the relative route and length probabilities remain fixed. This predicts how a manipulation of cue precision should redistribute replay while preserving its conditional structure. The reported alternation between successive theta samples also motivates measuring dependence across events alongside the distribution of individual events (Ujfalussy & Orbán, 2022).

The meaning of the represented transitions determines how these samples are used. A model of movement supplies candidate routes; a spatial coordinate system can also compare locations beyond the traveled path. Recordings of entorhinal and hippocampal populations reveal alternating sweeps into surrounding locations, including inaccessible regions, with related dynamics persisting during REM sleep (Vollan et al., 2025). These observations motivate defining the reference at the level of represented situations and specifying the behavioral or internal dynamics its transitions describe. In the bridge formulation, current evidence selects among the complete sequences supported by that reference.

B.3 SELECTION, DIRECTION, AND THE TIMING OF REPLAY

Hippocampal replay includes forward and reverse sequences, paths toward remembered goals, remote trajectories, and combinations of familiar transitions (Foster & Wilson, 2006; Diba & Buzsáki, 2007; Gupta et al., 2010; Pfeiffer & Foster, 2013). Reward changes selectively affect reverse replay (Ambrose et al., 2016), while coordinated hippocampal and prefrontal replay changes with learning and the retrospective or prospective demands of a task (Shin et al., 2019). The bridge assigns this selection to endpoint evidence applied before the sequence unfolds.

Prospective sequences during ongoing behavior also depend on intended destinations. Wikenheiser & Redish (2015) found that the spatial extent of hippocampal theta sequences varied with the distance to the animal's current goal and predicted its destination. In the path model, terminal evidence changes the distribution of represented sequences and their transition counts. The physiological clock determines how this represented extent is expressed within a theta cycle or a compressed replay event. This distinction connects prospective content to current intention while retaining separate predictions about neural timing.

Models of prioritized memory access select experiences according to their expected contribution to future decisions (Mattar & Daw, 2018). Recurrent planning models relate simulated futures to choices and the time spent deliberating (Jensen et al., 2024). Context-based retrieval and models combining experience with environmental structure explain further replay statistics (Zhou et al., 2026; Diekmann & Cheng, 2023). These computations can determine which cues and outcomes receive weight. The bridge specifies the distribution of complete sequences generated from that evidence. Evaluation of those samples can then support action selection.

Several replay regimes follow from the same generator. Flat endpoint factors recover spontaneous reference sampling, consistent with the broad exploratory trajectories observed after open-field experience (Stella et al., 2019). Localized beginnings and endings favor events connecting those regions. Under a reversible reference, exchanging the endpoint factors reverses the complete path law. Adaptation and spectral modulation can change the motion and spatial extent of generated sequences (Ji et al., 2026; McNamee et al., 2021). The observed tendency to avoid recently replayed paths further indicates that recent activity can influence sampling (Mallory et al., 2025). At a resolution where such dependence matters, the represented state includes the relevant activity history or the reference changes with physiological state.

Sequence order, elapsed time, and ripple expression supply distinct measurements of the generator. Decoded replay includes trajectories with speeds close to behavior (Denovellis et al., 2021), and replay also occurs without detectable ripples (Widloski & Foster, 2025). In the circuit sampler, a common rescaling of the competing rates changes elapsed time while preserving state order and stopping probabilities. Changing the relative drive to continuation and completion changes the sequence distribution. The interaction with ripple-generating circuitry can therefore be studied through specific effects on initiation, relative competition, and the physical clock of an event.

Extended replay can remain coherent across a train of ripple events (Davidson et al., 2009). This makes the boundary of a represented episode and the boundary of an oscillatory burst separately measurable. Common temporal gating of the candidate rates can interrupt and resume an ongoing sample while preserving its path law, provided that the active state and field persist (Appendix A.6.2). Resetting initiation between bursts instead selects a new episode. Decoding across ripple intervals can distinguish these operations by asking whether the next represented state continues the preceding sequence or is drawn from a renewed starting distribution. Sequence extent, neural speed, and the grouping of bursts consequently provide complementary tests of the same circuit.

Models of adaptation and momentum address how the reference activity itself evolves. In recurrent networks trained for path integration, Casco-Rodriguez et al. (2026) show that adaptation promotes exploration while changing temporal dependence, and that momentum can accelerate the generated dynamics. Such mechanisms introduce history through additional activity variables. At a resolution that retains those variables, the bridge can condition their joint dynamics; the projected spatial trajectory can then exhibit persistence or avoidance. This connection clarifies what has to be measured when replay changes speed. A common rate change preserves transition probabilities, while a change in adaptation or momentum can change both timing and route statistics. Matching the distribution of occupied locations alone leaves these temporal relations undetermined, motivating the order-sensitive comparisons in Experiment 2.

B.4 CIRCUIT OPERATIONS THAT LINK LOCAL CHOICES TO COMPLETE SEQUENCES

The continuation field gives recurrent activity a definite computational meaning: activity at a state measures the total compatibility of the episodes that can still unfold from it. A recalled ending supplies current; learned connectivity propagates that current; leak and inhibition attenuate long chains of propagation. He et al. (2026b) derive an LMDP desirability field from hippocampal recurrent relaxation. Their equivalence holds for the linear response of a reversible network with matched gains and a clamped goal. Appendix A.6.1 relates this response to the present continuation field through the neural coordinate transformation and normalization at the goal. Dendritic gain makes this distributed field available to stochastic choices, and competition with completion generates a route with an ending. Future compatibility influences a transition through activity already present in the circuit.

Biophysical CA3 models show how structured excitation and inhibition can generate sharp waves, ripples, and associated replay (Ecker et al., 2022). The construction in Section 3 assigns a sampling computation to these interactions. Recurrent drive determines the candidate successors; the continuation field changes their relative probability according to the compatible paths available from each.

Entorhinal input provides a candidate route for the information applied to this dynamics. Perturbing direct medial entorhinal input to CA1 affects extended awake replay (Yamamoto & Tonegawa, 2017), and medial entorhinal silencing alters ripple incidence and assembly participation (Zutshi & Buzsáki, 2023). Simultaneous recordings further show representations of task-relevant remote locations in medial entorhinal cortex that can occur independently of CA1 (Aery Jones et al., 2026). Such activity could supply endpoint evidence. The model additionally predicts a distributed signal whose value at a state depends on its access to those endpoints. Its computation carries information from possible successors back toward their predecessors, requiring matched feedback in a directed map.

Prefrontal input through the thalamic nucleus reuniens supplies another route for prospective information. Lesion or optogenetic silencing of reuniens reduces trajectory-dependent firing in CA1 (Ito et al., 2015). This pathway gives the circuit proposal an anatomical basis for applying task information to hippocampal readout. Its projection specificity also guides experiments: the source of a boundary signal and the recurrent population that propagates its influence can be measured separately, together with their effects on successor choice and completion.

Nonlinear dendritic integration and NMDA spikes in CA3 provide physiological mechanisms for interactions between recurrent excitation and additional input (Makara & Magee, 2013; Brandalise et al., 2016). The proposed gain operation has a measurable form: the drive to a candidate successor is proportional to its reference drive multiplied by its continuation compatibility. The CA1 completion pathway combines the current representation with terminal input and competes with these successors. Behavioral time scale plasticity allows a CA1 dendritic event to modify synapses activated over an extended preceding interval (Bittner et al., 2017). In the proposed circuit, such an event could associate preceding activity with completion. Measuring the gain interaction, the endpoint signal, and their effects on termination would connect these circuit assignments to the path computation.

This joint account of continuation and completion provides a focused perturbation design. Suppressing field contrast should weaken prospective successor preferences. Weakening the termination channel while holding continuation drive fixed should reduce local stopping probability while preserving the relative choice among successors. Changing terminal evidence should affect both its local completion score and its propagated influence on the map. The predicted pattern across these measurements identifies which operation a perturbation changes.

The model also fixes the relative size of these effects. The sensitivity of log field activity to a log terminal cue equals the probability that an episode starting there will finish at the cued ending. Differences in that probability determine the cue's effect on successor odds, while its effect on local stopping follows from the same quantity (Eq. (110)). Recording the distributed response to a selective endpoint input can therefore predict a change in sequence statistics. This connects a measurable population signal to the behavior of complete events.

B.5 How sleep changes future replay

Sleep organizes replay according to physiological state and memory history. Disrupting sleep ripples impairs spatial memory (Girardeau et al., 2009). Within NREM sleep, recent and older memories are preferentially replayed in different pupil-defined substates, accompanied by differences in extrinsic excitation and local inhibition (Chang et al., 2025). Such organization provides a physiological setting in which different endpoint distributions can be sampled from established dynamics.

The learning rule in Section 3.3 gives those samples a specific computational role. Experience supplies the endpoint frequencies to be expressed; replay samples the frequencies currently expressed by the network. Their difference adjusts endpoint weights. An ending that occurs too often in generated events receives more negative updates, while an underrepresented ending receives a net increase. An offline sequence thus supplies a local sample of the network's own expectation. The network can estimate how changing a boundary weight affects the distribution of all its possible episodes by observing the episodes it generates.

This interpretation connects learning to the normalization of a generative model. Increasing one endpoint weight changes its frequency relative to every other ending. The replay term measures that competition without explicitly summing over complete paths. At agreement with experience, the expected parameter change vanishes while replay can continue. A related principle has empirical support in sensory cortex: spontaneous and stimulus-evoked activity become statistically more similar during development (Berkes et al., 2011). The present rule makes a specific hippocampal prediction about the adjustment of episode boundaries.

Hippocampal ripples can trigger lasting, pathway-selective synaptic depression (Norimoto et al., 2018). Sleep also produces changes in synaptic size and receptor regulation (de Vivo et al., 2017; Diering et al., 2017), and a CA2 circuit recruits inhibition that counterbalances elevated reactivation after learning (Karaba et al., 2024). These findings establish physiological routes through which replay and sleep can reduce subsequent activation. The bridge learning rule predicts a particular selectivity: the change in an endpoint weight depends on the discrepancy between its experienced and generated frequency. This can be tested alongside synaptic strengthening, cortical transfer, and regulation of overall excitability.

Dreaming extends the question to internally experienced scenes. Patients with hippocampal damage report fewer dreams and less episodic detail (Spanò et al., 2020). When sensory constraints are weak, internal cues can select an episode from learned relations among situations. Its local continuity follows from the reference, while uncertainty about its boundaries allows many possible constructions. This supplies a hypothesis for the organization of dream sequences. Relating it to subjective content calls for a description of the sensory representations associated with states and their coordination across sleep stages. In particular, the NREM events used to study replay and the REM dynamics associated with spatial sweeps provide different physiological settings for internally generated content (Chang et al., 2025; Vollan et al., 2025).

B.6 Schrödinger bridges, control, and generative models

Schrödinger's problem selects a process that matches endpoint distributions while minimizing departure from a reference (Schrödinger, 1931; Léonard, 2014); the Doob transform gives its local transitions (Doob, 1957). Related positive transforms support linearly solvable control (Todorov, 2009) and reusable planning in cognitive maps (Piray & Daw, 2021; 2025). These relations provide the mathematical setting for treating learned dynamics as a reusable generator.

Completion can itself belong to the inferred process. Eldesoukey & Georgiou (2024) formulate graph bridges with initial distributions and constraints on the time of first arrival at absorbing states. The optimal process remains Markov and can be constructed through marginal fitting. In diffusion models with losses, Chen et al. (2022) infer both transport and disappearance under a common path probability law. The present event model gives every represented state a completion alternative and conditions the resulting paths on their endpoint distributions. Its stationary continuation field then supplies a common neural signal for successor choice and stopping. These connections locate complete episode generation within inference over processes whose duration is uncertain.

Lin et al. (2026) give a complementary neural account of generating plans from a cognitive map. Their model learns embeddings of observations and actions, a forward model that predicts the state

change caused by an action, and an inverse model that maps a desired state change to candidate actions. Noisy selection of virtual actions generates diverse routes toward a target, including solutions in compositional problem spaces. Their construction explains how local predictive learning acquires a geometry useful for flexible action selection. The bridge specifies a complete event distribution once the reference dynamics are given. Its field aggregates the learned probabilities of all compatible continuations, and its local competition jointly determines successor choice and completion. The central test concerns the joint law of endpoint frequencies, routes conditional on endpoints, and event lengths, together with how replay calibrates that law. An action-based world model could supply the reference transitions, while bridge samples provide episodes for subsequent evaluation. The roles of representation learning, conditional generation, and action evaluation can thus be tested separately within a shared planning system.

Diffusion models use learned reverse dynamics to recover a data distribution from a simple source (Song et al., 2021). Diffusion Schrödinger bridges connect prescribed source and target distributions (De Bortoli et al., 2021), using the reference to specify the intervening dynamics. Bridge matching exploits the relation between the endpoint coupling and the conditional paths joining each pair (Shi et al., 2023). This relation also clarifies what the reference contributes to a cognitive episode: boundary evidence selects which beginnings and endings occur together, and learned dynamics supply their intervening sequences.

Image restoration gives the boundaries a concrete interpretation. In I^2SB, a degraded image provides evidence about the clean image to be reconstructed, and training uses paired boundaries to learn the connecting diffusion (Liu et al., 2023). Neural Schrödinger bridges also support translation between image domains using unpaired examples (Kim et al., 2024). Dual diffusion implicit bridges compose source and target transformations through a shared latent distribution (Su et al., 2023); aligned diffusion bridges incorporate correspondence between observations when recovering biological dynamics (Somnath et al., 2023). These constructions illustrate how learned regularities and current evidence jointly organize generation.

The role of an intermediate state gives each construction its meaning. In image restoration, the path describes stages of transforming an observation into a reconstructed image. In the hippocampal model, successive states constitute the represented episode itself. An intermediate location, a detour, and the point of completion are all parts of the generated content. The neural clock determines how quickly that content unfolds. This places recall, prospective simulation, and replay within one computation: learned relations determine the possible sequences, present evidence selects an ensemble, and local dynamics construct individual episodes.

The dynamics supplying the reference. Recurrent excitation can sustain a localized population pattern while adaptation or synaptic depression moves activity through the map. Romani & Tsodyks (2015) showed that a recurrent map with short-term synaptic depression can express both theta-associated activity and internally generated trajectories under different external inputs. Recent measurements connect avoidance of recently expressed paths to neuronal fatigue (Mallory et al., 2025). These mechanisms determine the dynamics available for conditioning. When their internal variables influence future transitions, including those variables in the represented state preserves the temporal information needed by the reference. The bridge then specifies how endpoint evidence acts on the resulting dynamics. An empirical reference should consequently be evaluated on sequence order as well as the distribution of visited states, as in Experiment 2.

The object generated by a bridge. Diffusion Schrödinger bridges learn stochastic processes connecting prescribed source and target distributions (De Bortoli et al., 2021). In image restoration, a degraded image provides information about the reconstruction; paired training examples specify how a bridge should connect the two (Liu et al., 2023). This is a useful computational precedent for combining learned regularities with incomplete evidence. In a hippocampal sequence, each intermediate state is itself part of the represented experience. Its route, ordering, and point of completion carry content that a theory of the whole event has to predict. This interpretation motivates treating event length and termination as consequences of the same conditioning operation that determines the next represented state.

Part II

Experimental details

The experiments follow the theory from local neural operations to complete episodes and learning. Experiment 1 establishes the generated path law on maps whose exact solution is known. Experiment 2 identifies temporal relations in recorded population activity that different boundary conditions can reuse. Experiment 3 evaluates the influence of behavioral boundaries on represented paths, event extent, and spike observations. Experiment 4 uses experienced and generated endpoint frequencies to predict changes in sleep sequences. Table 2 connects these questions to their data and comparisons. The linked contents below lead to the full protocols and results. Theoretical derivations appear in Part I (p. 18).

Table 2: Guide to the experimental evidence. Section names and page numbers in the contents below link directly to the corresponding methods and results.

Scientific question	Data	Primary comparison
1. Local generation	Eight finite maps	Recurrent computation and local sampling versus the exact distribution over complete paths.
2. Reusable temporal relations	HC-3, HC-11, DANDI 000978	Held-out sequence prediction and interior order with endpoints and state counts fixed.
3. Conditional episodes	HC-3, HC-6, additional DANDI cohorts	Complete-event probability and spike likelihood at fixed event length under shared behavioral boundary cues.
4. Learning from experience	HC-11, DANDI 000978; HC-31 transfer	PRE-to-POST likelihood change along the endpoint learning direction and prediction after a finite update.

CONTENTS

C Experiment 1: local sequence generation and endpoint learning 51

C.1 Reference environments and evidence 51

C.2 Local computation, sampling and error measures 52

C.3 Local implementation and mechanism results 53

C.4 Representation capacity and endpoint localization 54

C.5 Perturbations and nonlinear recurrent responses 58

C.6 Reference variation, computation time and sampling budgets 61

C.7 Learning initial and terminal distributions 63

C.8 Analytical observables and illustrative conditions 68

D Experiment 2: reusable temporal structure in hippocampal population activity 71

D.1 From a reference map to a test of sequence order 71

D.2 Data and independent sampling units 71

D.3 From spikes to a state sequence 72

D.4 Reference and comparison readouts 73

D.5 A concrete population sequence 74

D.6 Prediction, order, and model complexity in HC-3 . . . 74
D.7 A PRE reference across behavioral epochs . . . 75
D.8 State resolution and population information . . . 77
D.9 Removing dwell and matching the legal transitions . . . 77
D.10 Conditional tests of internal order . . . 80
D.11 Running and stationary task activity . . . 81
D.12 Reproducible outputs . . . 84

E Experiment 3: behavioral boundaries and complete hippocampal events 85
E.1 From movement and spikes to represented paths . . . 85
E.2 What the path distribution predicts . . . 89
E.3 Fitting and prediction on test events . . . 89
E.4 How continuation and stopping shape an event . . . 93
E.5 Spatial resolution and prediction of held-out spikes . . . 96
E.6 Independent ripple evidence . . . 102

F Experiment 4: experience and the redistribution of sleep paths 104
F.1 How experience changes the use of a learned map . . . 104
F.2 State representations, reference transitions, and behavioral endpoints . . . 105
F.3 Population events and ripple screening . . . 105
F.4 Path likelihood and the endpoint score . . . 106
F.5 Temporal cross-fitting and the local prediction . . . 107
F.6 Experience predicts the direction of endpoint reallocation . . . 108
F.7 From a learning direction to prediction of subsequent sleep . . . 109
F.8 Repeated experience stabilizes sleep as behavior becomes more efficient . . . 110
F.9 Transfer across cohorts and structure within sleep paths . . . 110
F.10 Endpoint summaries and illustrative conditions . . . 111
F.11 Computational materials . . . 112

How to read the measurements

Represented states are graph nodes in Experiment 1, clusters of simultaneous spike counts in Experiment 2, and spatial states with firing-field emissions in Experiment 3. Experiment 4 uses spatial states in HC-11 and population prototypes in DANDI. State numbers identify elements of these fitted representations.

Session estimates are combined within animal before animals are averaged in HC-11 and DANDI and in the hierarchical HC-3 analyses. Experiment 2's 27-session HC-3 summary uses recordings as its averaging unit. HC-6 and the separate HC-3 temporal characterization each contain one animal. Every interval is reported with the corresponding resampling level.

Prediction gains are log-likelihood differences on paired observations. Positive values indicate greater probability under the stated model. A gain per state or bin divides the sum of event log-likelihood differences by the total number of scored states or bins. A gain per second uses total exposure. A gain per event averages complete-event differences. These denominators remain distinct across experiments. In Experiment 1, K counts transitions, including self-transitions, so immediate

completion has $K = 0$. In Experiment 3, $L = K + 1$ counts observed bins. In Experiment 4, the path $x_{0:L}$ uses L transitions and $L + 1$ states.

The population-event definition and the ripple criterion are specified separately. Population windows provide the sequence support; a ripple analysis selects windows with independent LFP evidence while retaining that support. The event tables state whether sleep epochs are annotated NREM or designated sleep/rest.

C EXPERIMENT 1: LOCAL SEQUENCE GENERATION AND ENDPOINT LEARNING

An endpoint cue concerns what will happen after several transitions. This experiment asks whether a recurrent circuit can make that information available at each transition and at the decision to stop. The test uses known maps so that the distribution over complete paths can be calculated exactly. A narrow passage, a shortcut, and a connection between communities provide different ways for the same local computation to organize a whole episode.

Each graph supplies a fixed transition matrix. Initial and terminal evidence specify the episode being requested. Recurrent propagation computes the continuation field, after which local competition selects successive states or completion. Replacing the evidence while preserving the graph asks the same map to generate another ensemble of episodes. The comparison includes where events begin and end, which transitions they take, and how long they continue.

Three calculations separate the sources of approximation. An analytical continuation field tests whether the local sampling rules compose into the intended path law. Numerical recurrent relaxation tests whether circuit activity reaches that field. A restricted population basis tests how much the generated distribution changes when the field is represented with fewer activity patterns. Table 3 gives their shared target and distinct comparisons. Subsequent sections vary endpoint precision (Section C.4), perturb local operations (Section C.5), measure integration and sampling costs (Section C.6), and train new endpoint distributions (Section C.7).

Table 3: Computations used in Experiment 1. Every approximation is compared with the exact path law for the same reference and boundary evidence.

Calculation	What is computed	Observable comparison
Analytical local law	Solve the field and assemble initiation, transition, and stopping probabilities.	KL over all finite paths; endpoint joint; mean length.
Recurrent computation	Integrate the relaxation equation from zero activity.	Field error and the induced complete-path law.
Compressed field	Solve the projected recurrence in a basis of specified dimension.	Path and endpoint errors at matched graph, evidence, and realization.
Sampled episodes	Draw successive local choices until completion.	Endpoint and length frequencies; expected state visits.
Endpoint learning	Update log weights using experienced minus generated endpoint frequencies.	Marginal error and joint error to the bridge coupling.

C.1 REFERENCE ENVIRONMENTS AND EVIDENCE

The reference map specifies the transitions available to the circuit. Endpoint evidence specifies which beginnings and endings are supported in the present condition. Holding the map fixed while changing this evidence tests the reuse of learned dynamics. Applying the same computation to distinct graphs tests its dependence on environmental connectivity.

Graph construction. We use eight finite environments (Table 4). Every graph state has coordinates used to construct endpoint evidence. A positive entry W_{ij} denotes an allowed transition from state i to state j. The reference transition matrix is $P_{0,ij} = W_{ij} / \sum_j W_{ij}$; its stationary distribution π supplies the initial prior. Standard graph weights are symmetric before row normalization. Multiplicative Gaussian weight perturbations preserve the allowed edges. Specifically, the construction symmetrizes the initial weights, multiplies them by $\exp(\sigma Z_{ij})$ for independent standard normal Z_{ij}, symmetrizes again, and raises each diagonal entry to at least $0.02 \max_{ij} W_{ij}$. The noise scales are 0.01, 0.015, 0.018, 0.025, 0.02, 0.02, 0.02 and 0.02 for line, ring, W-maze, grid, bottleneck, shortcut, tree and modular graphs. The shortcut graph starts from a perturbed line, adds its two extra edges, and receives the additional perturbation. Graph coordinates specify evidence and display geometry; transition weights are given by the construction rules.

Table 4: Finite environments used to test local sequence generation. Node indices start at zero.

Environment	States	Connections and initial weights
Line	28	Adjacent nodes, weight 1; self weight 0.08
Ring	30	Adjacent nodes and closing edge, weight 1; self weight 0.06
W-maze	13	Four corridors at node 3, weight 1; two cross edges 0.22
Grid	63	9×7 grid; axial edges 1, diagonals 0.35; self weight 0.05
Bottleneck	41	Two 4×5 regions joined by one node; axial edges 1
Shortcut	24	Line plus edges $(4, 19)$ of weight 0.75 and $(8, 15)$ of 0.45
Tree	15	Four binary levels; parent–child edges 1; self weight 0.04
Modular	28	Four 7-node cliques; within weights uniform on $[0.55, 1)$

The W-maze corridors are 0–1–2–3, 3–4–5–6, 3–7–8–9, and 3–10–11–12, with cross edges $(5, 8)$ and $(8, 11)$. The bottleneck retains the central node at $(4, 2)$ from an otherwise removed central column of a 9×5 grid. Its self weight is 0.05. The modular graph connects consecutive modules cyclically, joining their local node 0 with weight 0.35 and local nodes 3 and 4 with weight 0.20. Its self weight is 0.05, as in the W-maze.

Endpoint evidence. Let $c_i = (x_i, y_i)$ be the coordinates of state i. Define

$$\sigma_0 = 0.35 \max\{\mathrm{std}(x), \mathrm{std}(y), 1\}, \qquad G_\sigma(i; c) = \exp[-\|c_i - c_c\|^2/(2\sigma^2)]. \tag{111}$$

The standard initial center minimizes $x_i + 0.15y_i$; the standard terminal center is the state farthest from it in Euclidean distance. Single-center evidence is $0.03 + G_\sigma$. Sharp evidence uses $\sigma = 0.45\sigma_0$, and smooth evidence uses $\sigma = \sigma_0$. Multiple-center evidence uses $\sigma = 0.8\sigma_0$, adds a second initial center at index $\lfloor N/3 \rfloor$ and a second terminal center maximizing $y_i - 0.1x_i$, and adds the background once. Reversed evidence uses $\sigma = 0.65\sigma_0$, starts at the standard terminal center, and ends at the state minimizing $\|c_i - c_{c_0}\|^2 + 0.2i$. Flat evidence sets both vectors to one. These inputs give relative support for candidate endpoints. The resulting endpoint probabilities also depend on reference paths connecting the states.

Reference and target distributions. For a path $\omega = (x_0, \ldots, x_K)$, with K counting completed transitions including self-transitions, the distributions are

$$R(\omega) = \pi_{x_0}(1 - \gamma) \prod_{t<K} \gamma P_0(x_t, x_{t+1}), \tag{112}$$

$$Q^\star(\omega) = Z^{-1}\pi_{x_0}\varepsilon_0(x_0) \left[\prod_{t<K} \gamma P_0(x_t, x_{t+1})\right] \varepsilon_T(x_K), \qquad Z = \sum_i \pi_i \varepsilon_0(i)\phi_i, \tag{113}$$

$$\phi = (I - \gamma P_0)^{-1}\varepsilon_T. \tag{114}$$

Here $P_0 = P$, $\varepsilon_0 = a$, $\varepsilon_T = b$, and $\phi = h$ in the main-text notation; g will denote a numerical approximation to the field. Immediate termination has $K = 0$. Every perturbation of the reference map or evidence defines its own matched target. Compression approximates the computation of ϕ while retaining the original state space and reference transitions.

C.2 Local computation, sampling and error measures

The continuation field summarizes the terminal support carried by all remaining paths from a state. A local transition can therefore take account of distant evidence through one number at each candidate destination. We evaluate the resulting process at two levels: its analytically implied law, which includes all possible paths, and finite samples drawn from that law.

Recurrent field. The full recurrent loop starts from zero and follows

$$u^{n+1} = u^n + \delta_{\mathrm{E}}[-u^n + \gamma P_0 u^n + \varepsilon_T], \qquad \delta_{\mathrm{E}} = \Delta t/\tau_\phi. \tag{115}$$

Full-loop implementation conditions use $\delta_{\mathrm{E}} = 0.05$ and 1,400 updates, then hold the resulting field fixed while generating sequences. Thus loop updates and sequence transitions are separate

computational steps. BrainPy implements the recurrent dynamics; JAX implements the parallel stochastic draws. The implementation calculations use double precision, BrainPy 2.7.6 and JAX 0.6.2 on GPUs. The systematic parameter scans use BrainPy 2.6.1, JAX 0.4.38 and NumPy 1.26.4. The time constant and integration step are used through their ratio and have no assigned physiological duration in this experiment.

Local probabilities. For any positive field g, the implemented initial, continuation and stopping probabilities are

$$s_i = \frac{\pi_i \varepsilon_0(i) g_i}{\sum_j \pi_j \varepsilon_0(j) g_j}, \quad T_{ij} = \frac{\gamma P_{0,ij} g_j}{\varepsilon_T(i) + \gamma (P_0 g)_i}, \quad r_i = \frac{\varepsilon_T(i)}{\varepsilon_T(i) + \gamma (P_0 g)_i}. \tag{116}$$

Here T includes the probability of continuing; $T\mathbf{1} + r = \mathbf{1}$. The sampler draws an initial state, decides whether to stop at its current state, and, on continuation, draws the next state from $T_{ij} / \sum_j T_{ij}$. At $g = \phi$, these rules give the target bridge. With an approximate field, the same competing-drive denominator keeps every local decision normalized.

Complete-path evaluation. A sequence law specified by (s, T, r) has fundamental matrix $F = (I - T)^{-1}$, expected visits $v = s^\top F$, and endpoint joint distribution $J_{ij} = s_i F_{ij} r_j$. It follows that $\mathbb{E}[K] = \sum_i v_i - 1$. These finite matrix calculations include arbitrarily long paths and repeated visits. For a comparison law $(s^\star, T^\star, r^\star)$,

$$D_{\mathrm{KL}}(Q \| Q^\star) = \sum_i s_i \log \frac{s_i}{s_i^\star} + \sum_i v_i \left[r_i \log \frac{r_i}{r_i^\star} + \sum_j T_{ij} \log \frac{T_{ij}}{T_{ij}^\star} \right]. \tag{117}$$

This identity follows by separating the path log likelihood ratio into its initial decision, all continuation decisions and the final stop, then taking their expected counts. Endpoint total variation (TV) is $\frac{1}{2} \sum_{ij} |J_{ij} - J_{ij}^\star|$. Relative field error is $\|g - \phi\|_2 / \|\phi\|_2$, and the reported Bellman residual is $\|g - \varepsilon_T - \gamma P_0 g\|_2 / \|g\|_2$. Numerical residuals near zero are recorded with their signs; the parameter-scan plots map negative KL values of magnitude below 10^{-10} to zero.

Sampling measures. Empirical endpoint and length distributions are compared with their analytical probabilities by TV. Empirical stopping probabilities use stops divided by visits; empirical conditional transitions use edge counts divided by departures from the corresponding source. Their root mean square errors are weighted by visits and departures, respectively. The successor representation is $S = (I - \gamma P_0)^{-1}$. For each fixed initial state, the sampler estimates expected visits to all states and compares the resulting matrix with S using relative Frobenius error. Initial and terminal visits, self-transitions and subsequent returns are included. Array padding after termination is excluded.

C.3 Local implementation and mechanism results

The basic implementation suite contains 140 records. Reference recovery uses eight flat-evidence conditions with $\gamma = 0.72$ and 420,000 events each. Forty analytical-field bridge conditions cross the eight graphs with five evidence types, using $\gamma = 0.76$ and 520,000 events each. Five graphs additionally compare the full loop and a fixed representation at the same sampling budget. The remaining records comprise 28 mechanism conditions, one directed-reference condition, 36 nonlinear conditions, 12 state-dependent propagation conditions, four learning conditions and one record summarizing five sequence configurations. The extended scans below contain 21,678 condition records; graph realizations, evidence types, and numerical settings define these records. The extended scans compare these controlled computational conditions with matched exact targets.

The analytical field gives a maximum absolute complete-path KL residual of 5.62×10^{-11} nats across the 40 bridge conditions. Sampled endpoint joint TV averages 0.00916 and reaches 0.02338 at its maximum. Weighted stopping and conditional-transition root mean square errors average 0.00146 and 0.00303, respectively. Across the recurrent implementation, the mean relative field error is 2.74×10^{-8} across line, ring, W-maze, grid and bottleneck graphs. The fixed representation retains 28 of 28, 30 of 30, 13 of 13, 53 of 63 and 36 of 41 dimensions, respectively. In the two compressed cases, grid and bottleneck, relative field errors are 0.01992 and 0.01145, and complete-path KL values are 0.001805 and 0.000675. The systematic rank experiment below provides the controlled comparison of representation dimension.

Which local decisions matter? At $\gamma = 0.78$, we remove the continuation field by setting $g_i = 1$, replace local stopping with the constant probability $1 - \gamma$, or replace row-normalized transition weights with $W_{ij}/\sum_j W_{ij}$. Each is compared with the intact target under sharp evidence. Removing the field gives KL 3.3619 on the grid and 3.2937 at the bottleneck; fixing the stopping probability gives 0.1822 and 0.2257; replacing normalization of outgoing weights gives 0.03017 and 0.02479. The grid's mean event length increases to 23.90 transitions when the continuation field is removed. The target mean is approximately 3.55. The continuation field consequently controls both where the sequence goes and when it ends. A state with strong terminal evidence can still support continuation when further paths carry substantial evidence. Stopping follows the balance of those alternatives. The increased event length after field removal makes this relation visible at the level of an entire sequence.

Direction and extent. Five configurations on the same line graph use 140,000 events each. Forward and reversed evidence at $\gamma = 0.76$ yield mean displacements 0.4150 and -0.4142, with mean lengths 3.1529 and 3.1700. Flat evidence at the same survival value gives mean displacement -0.0077 and length 3.1813. Increasing flat-evidence survival to 0.91 raises mean length to 10.0866, with displacement 0.0020. Directional evidence at survival 0.38 gives length 0.6110 and displacement 0.0944. The same transition map thus supports forward, reverse and weakly directed sequences. Endpoint evidence determines their directional bias, while survival changes the weight assigned to longer continuations. This reuse of one map under different boundary conditions is the computational relation underlying the examples of planning, recall and unconstrained sequence generation in the main text.

C.4 Representation capacity and endpoint localization

A finite population may represent the continuation field with fewer activity patterns than there are map states. Its accuracy then depends on the structure of the required field. Broad terminal evidence can be represented by slowly varying patterns; localized evidence may require finer variation. Propagation through the map changes both kinds of evidence before the field is read out. We vary representation dimension and evidence width together to measure this interaction, retaining the same environmental states and allowed transitions throughout.

Controlled dimension scan. Three graph realizations use seeds 202609120, 202609121 and 202609122. For every graph, all integer dimensions $k = 1, \ldots, N$ are tested with all five evidence families, giving 3,630 conditions. The basis C contains the lowest-frequency eigenvectors of the normalized graph Laplacian. It is determined from the reference connections before endpoint evidence is supplied. With $D_\pi = \mathrm{diag}(\pi)$, define

$$E = (C^\top D_\pi C)^{-1} C^\top D_\pi, \quad A = E P_0 C, \quad u = (I - \gamma A)^{-1} E \varepsilon_T, \quad g_i = \max\{(Cu)_i, 10^{-12}\}. \tag{118}$$

The compressed equilibrium is solved directly to measure representation error independently of integration error. At $k = N$, the control uses $C = E = I$ and $A = P_0$. The actual rank of every tested spectral basis is checked, giving 726 rank checks. An additional 720 conditions compare Euclidean projection at requested fractions 0.1, 0.2, 0.35, 0.5, 0.7 and 1. Requested fractions are rounded upward to an integer dimension.

All full-dimensional controls recover the target, with maximum KL 5.90×10^{-16} and maximum endpoint TV 3.06×10^{-16}. At approximately 35% capacity, mean sharp-evidence KL ranges from 0.10697 on the grid to 0.26821 on the tree. Broadening sharp evidence to the smooth condition reduces KL in 23 of the 24 matched graph realizations. The largest relative benefits occur on the line, ring and shortcut graphs. The full scan in Figure 11 shows that increasing dimension can produce local increases in error because spectral truncation, projection and positivity correction jointly determine the field presented to the sampler. The relevant representational demand is set by evidence after it has propagated through the map. Slowly varying activity patterns can carry broad evidence over many states, while a sharply localized endpoint often requires finer patterns. Connectivity and the geometry of the retained modes also shape the propagated field, explaining the variation in representational demand across maps.

Endpoint support and width. A joint scan crosses eight graphs, three weight realizations, three initial locations, two terminal locations, five evidence-width multipliers $\{0.25, 0.5, 1, 2, 4\}$, five back-

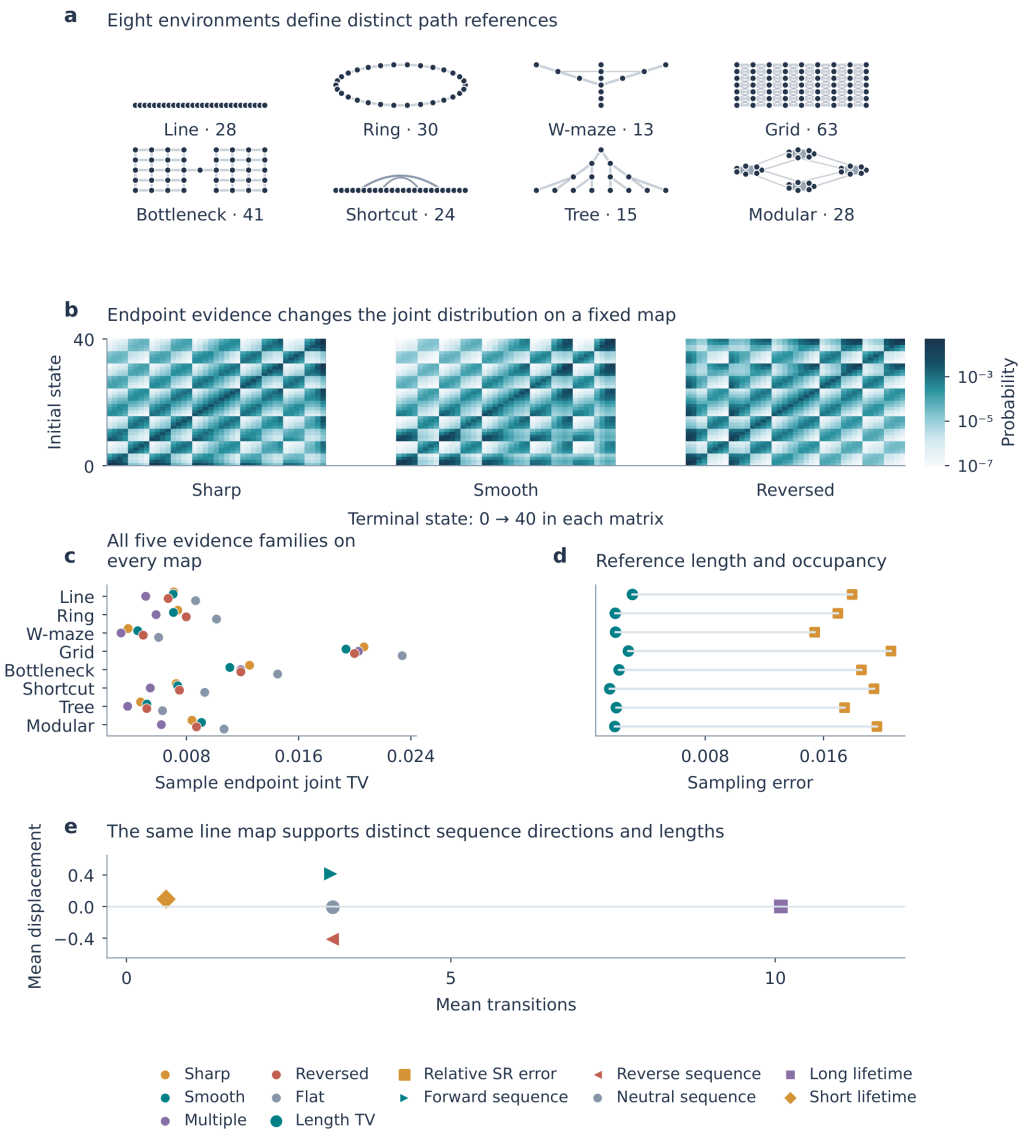

Figure 9: **Reference geometry, endpoint evidence and sampled sequences.** (a) The eight reference graphs. (b) Analytical endpoint joints for sharp, smooth and reversed bottleneck evidence, with shared state order and probability scale. (c) Sampled endpoint joint TV for every graph and evidence family, using 520,000 events per condition. (d) Length TV and relative successor-representation error, using 420,000 events and 4,000 events per initial state, respectively. (e) Mean transitions and displacement in five line-graph configurations, each using 140,000 events. The matrices are analytical; the other quantitative panels show sampling measurements.

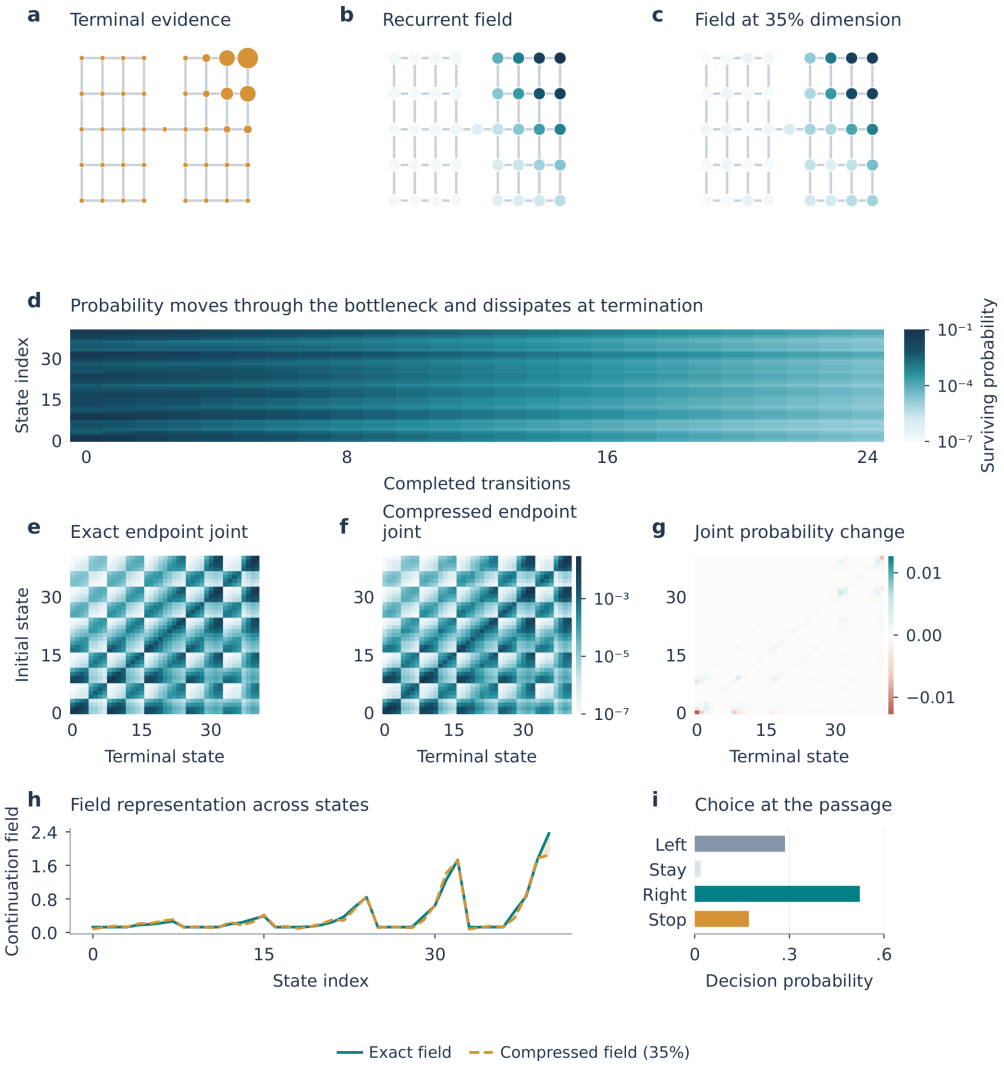

Figure 10: **The continuation field organizes probability flow and endpoint coupling.** (a–c) Terminal evidence, exact recurrent field and compressed field on the same 41-state bottleneck graph. Node area encodes evidence; field colors share a logarithmic relative scale. (d) Exact probability of occupying each state after t transitions while the sequence remains active, $s^\top T^t$. Every state is shown through transition 24. (e,f) Exact and compressed initial–terminal joints on a common probability scale. (g) Compressed minus exact joint probability. (h) Both fields by state, with their difference shaded. The analytical illustration uses seed 20260524, $\gamma = 0.76$ and 15 retained dimensions. (i) At the central passage, local continuation drives and terminal evidence compete to select left, self, right, or completion; bars give the normalized probabilities.

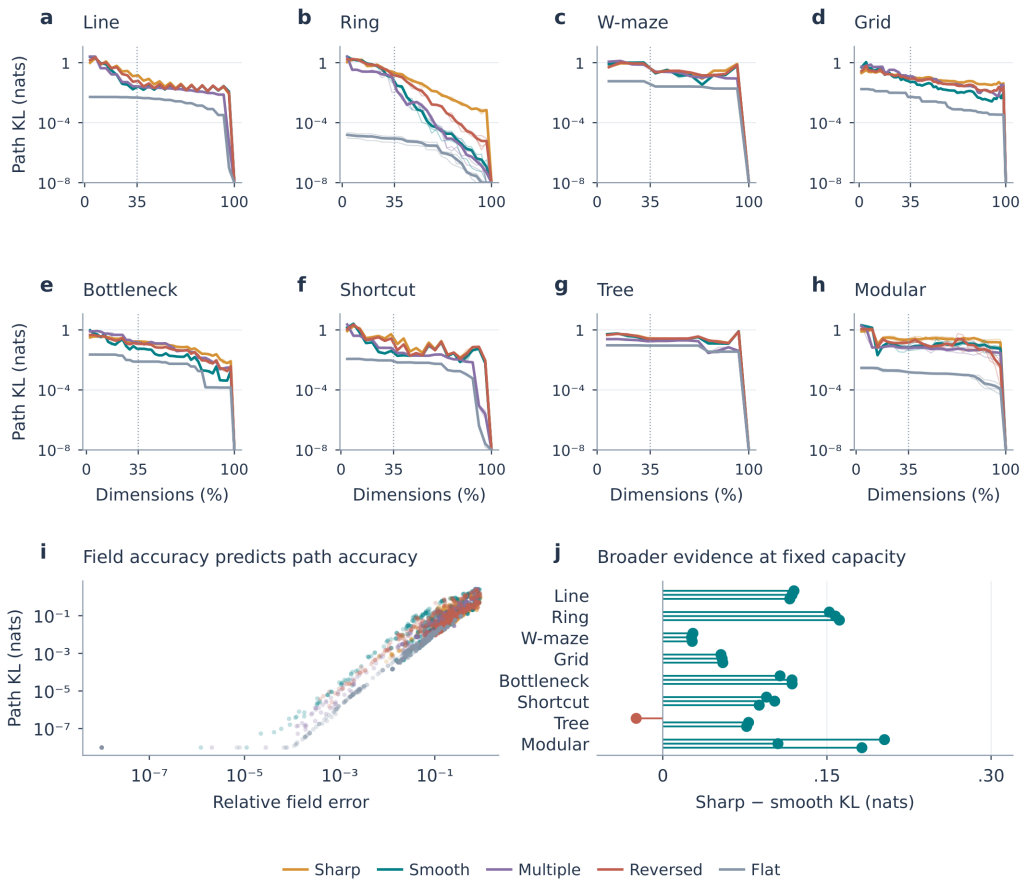

Figure 11: **Representation dimension sets sequence precision in a topology-dependent way.** (a–h) Every integer rank and evidence family on all eight graphs. Thin curves show three graph realizations; bold curves show their means. Horizontal coordinates give $100k/N$; dotted lines mark 35%. (i) Field error and path KL for all 3,630 conditions. (j) All 24 paired sharp-minus-smooth KL differences at $k = \lceil 0.35N \rceil$. Positive values indicate lower error with smooth evidence; terracotta marks the opposite sign. Display floors are 10^{-8}; source tables retain the numerical values.

grounds $\{0, 0.001, 0.01, 0.03, 0.1\}$, and three requested dimension fractions $\{0.35, 0.7, 1\}$. Initial locations are the standard center, the state closest to the coordinate centroid, and a state farthest from the standard center in graph distance. Terminal locations are an adjacent state and a graph-distant state. Width multiplies the baseline sharp width. These choices give 10,800 conditions.

At zero background, 265 of the 720 conditions at 35% dimension and 213 of the 720 at 70% dimension yield an ill-conditioned numerical evaluation. Full dimension remains evaluable in all 720 conditions. At each positive background, every condition is evaluable. At background 0.03, median KL across the 720 conditions is 0.09204 at 35% dimension, 0.02721 at 70%, and 5.36×10^{-17} at full dimension. Raising the background to 0.1 gives medians 0.03346, 0.01074 and 1.02×10^{-16}. Zero background can make the stopping system poorly conditioned when a compressed readout assigns negligible continuation support to required states. Positive diffuse evidence maintains support throughout the represented map.

A further 648 conditions vary initial and terminal widths independently over $\{0.5, 1, 2\}$ at 35% and full dimension, and vary the shared width for smooth, multiple-center and reversed evidence at 35% dimension. Another 624 conditions vary the initial-center coefficient over $\{-0.3, 0, 0.15, 0.3, 1\}$, second-terminal coefficient over $\{-0.3, 0, 0.1, 0.3, 1\}$, reversed-index penalty over $\{0, 0.05, 0.2, 0.5, 1\}$, secondary-initial index fraction over $\{0.2, 1/3, 0.5, 0.75\}$, reference-width coefficient over $\{0.175, 0.35, 0.7\}$, and coordinate-scale floor over $\{0.25, 0.5, 1, 2\}$. All 624 endpoint-construction conditions are evaluable, with path KL ranging from 0.00342 to 0.44764. Across the independent-width conditions, full-dimensional KL remains below 6.63×10^{-16}; at 35% dimension its median is 0.13806. These scans evaluate compression against the target defined by the modified evidence.

C.5 Perturbations and nonlinear recurrent responses

The complete path law contains an initial decision, a series of transition decisions and a stopping decision. Perturbing one operation at a time reveals which aspect of sequence generation it controls. We first introduce graded errors in the continuation field, transition drive and stop competition. We then change the recurrent response while keeping the desired path distribution fixed. This second manipulation measures how a nonlinear neural response affects the computation that the recurrent field is intended to perform.

Graded local perturbations. We test eight graphs and three weight realizations with sharp evidence at $\gamma = 0.78$, yielding 840 conditions. Perturbation directions are held fixed within each realization while magnitude varies. Log gain noise uses $g_i = \phi_i \exp(\sigma z_i)$, $z_i \sim \mathcal{N}(0, 1)$, at $\sigma \in \{0, 0.05, 0.1, 0.2, 0.45, 0.7\}$. Bounded field gain uses $g_i = \phi_i(1 + \delta u_i)$ and transition-drive error uses $P_{0,ij}(1+\delta U_{ij})$, where u_i, U_{ij} are uniform on $[-1, 1]$ and $\delta \in \{0, 0.05, 0.1, 0.2, 0.35, 0.5\}$. Saturation uses $g_i = \phi_i/(1 + \phi_i/c)$ at $c \in \{0.25, 0.5, 1, 2, 4, 8, 32\}$. Relative continuation gain multiplies the continuing drive by $\rho \in \{0.5, 0.7, 0.85, 1, 1.15, 1.3, 1.5\}$ while keeping terminal drive fixed. The three removals described above complete the scan.

The resulting path errors increase continuously with small gain and transition perturbations. For example, grid and bottleneck mean KL values at log gain standard deviation 0.2 are 0.06569 and 0.06454, compared with 0.35429 and 0.33728 at 0.45. Relative continuation gains 0.85 and 1.15 give grid KL values 0.008885 and 0.007346. These manipulations connect specific local quantities to observable sequence errors. A field perturbation changes the relative support for different continuations; an error in transition drive changes which continuations the map makes available; a change in continuation gain changes the competition with termination. Their effects can be distinguished by examining endpoint coupling and event length together.

State-dependent reference propagation. We compare the scalar-survival sampler with a target whose sub-stochastic kernel is $B_{ij} = 0.76P_{0,ij}(1 + \delta Z_{ij})$. Gaussian perturbations are centered under each reference row and clipped to $[-1, 1]$ before multiplication. The six magnitudes $\{0, 0.03, 0.06, 0.12, 0.2, 0.3\}$ across eight graphs and three realizations give 144 conditions. The target uses $\phi_B = (I - B)^{-1}\varepsilon_T$ and the same local rules with B replacing γP_0. We evaluate the original scalar model against this modified target and separately check the modified target against its analytical endpoint matrix. This comparison measures the cost of replacing state-dependent propagation by a common survival probability.

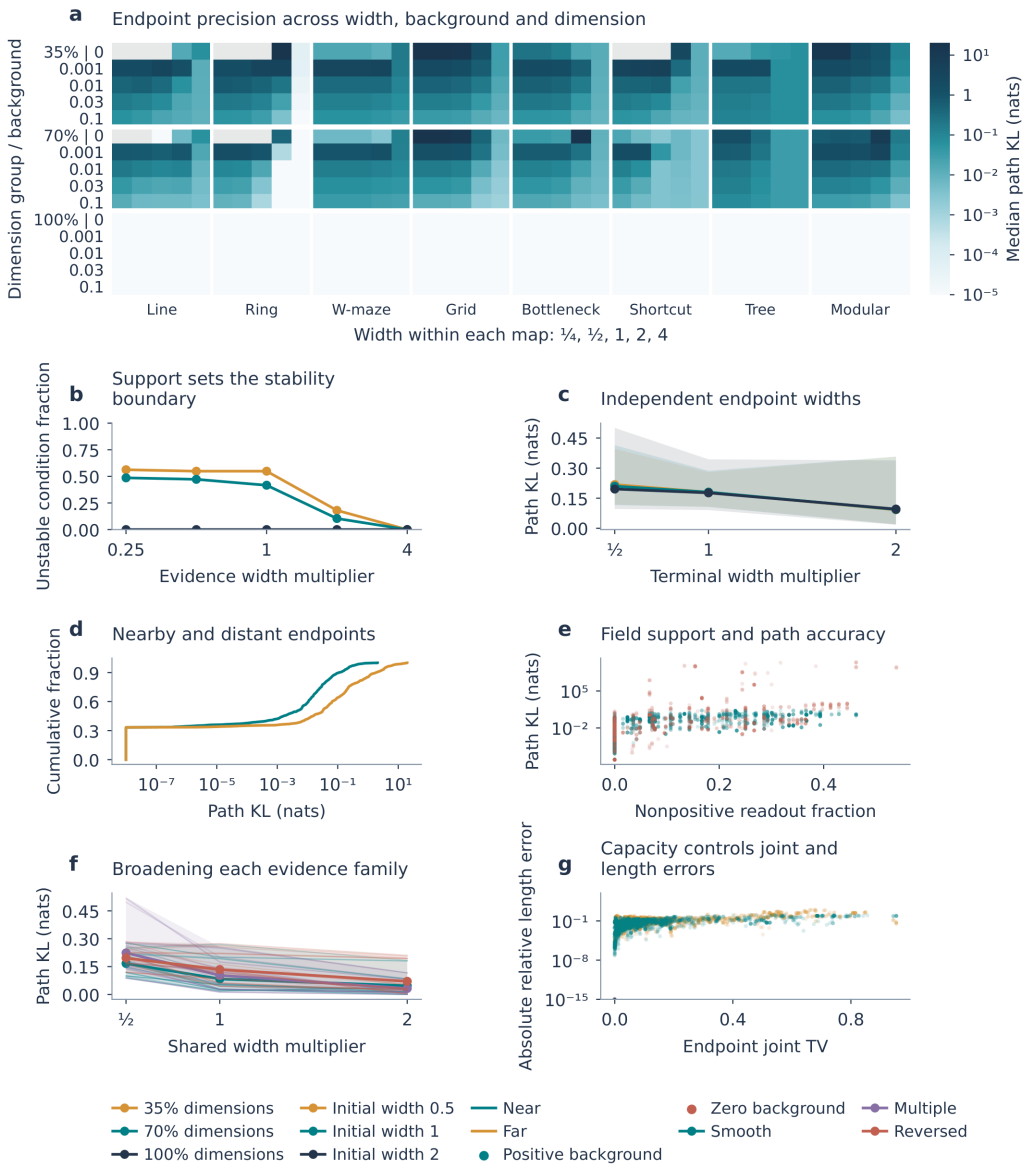

Figure 12: **Endpoint localization and support shape compression accuracy.** (a) Complete width–background atlas at 35%, 70% and full dimension. Cells give median evaluable KL; gray cells have no evaluable outcome. (b) Zero-background failure fractions by width and dimension. (c) Independent endpoint widths: means and full ranges across graph realizations. (d) Path-KL cumulative distributions for nearby and distant endpoints. (e) Path KL versus nonpositive raw-field fraction. (f) Shared-width effects: individual realizations, means and ranges. (g) Joint TV and absolute relative mean-length error for all 8,640 positive-background conditions. Detailed pooling and sample counts are specified below.

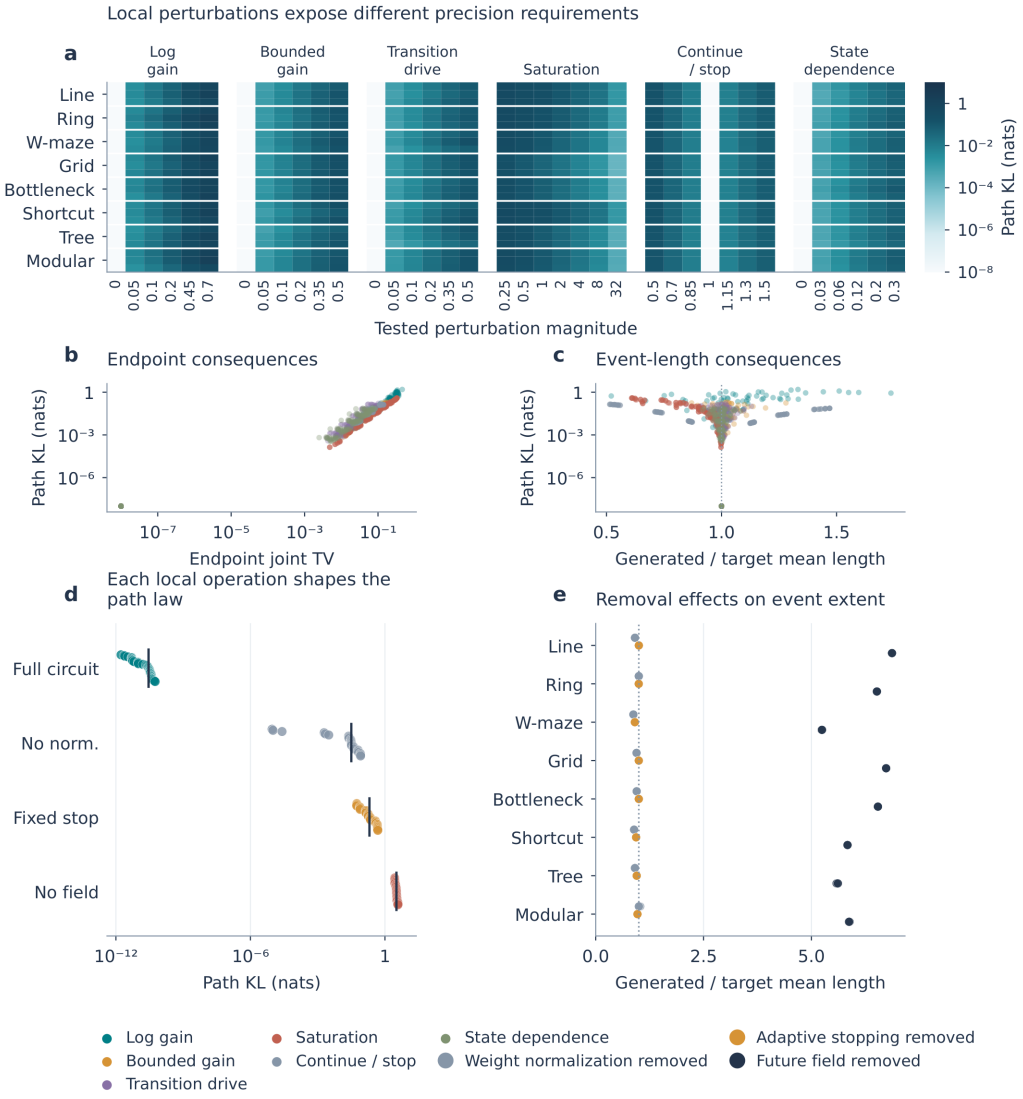

Figure 13: **Local computations control endpoint coupling and sequence extent.** (a) The perturbation atlas, with three realization rows per graph. (b,c) Path KL, endpoint joint TV, and mean-length ratio for all 912 graded conditions. (d) Complete-path divergence for the intact law and removal of weight normalization, adaptive stopping, or the continuation field. Points show 40 intact conditions and 24 realizations per removal; bars mark medians. Absolute numerical KL residuals for the intact law are displayed with a 10^{-12} floor. (e) Mean-length ratios for every removal realization. Vertical lines mark equal target and generated means. Colors in (b,c) identify perturbations and colors in (e) identify removed operations.

Nonlinear response. The recurrent response is changed to $\Phi(z) = z + \kappa(\sqrt{1+z^2} - 1)$, with dynamics $\tau_\phi \dot{u} = -u + \gamma P_0 \Phi(u) + \zeta \varepsilon_T$. At small activity, $\Phi(z) = z + \kappa z^2/2 + O(z^4)$. The input amplitude ζ scales all terminal evidence, including the background. The target bridge is invariant to this common positive scaling because its path weights and normalizer both acquire the factor ζ. Consequently, changes in the generated path law measure nonlinear distortion of an invariant computational target.

A first nonlinear comparison uses three graphs, amplitudes $\{0.025, 0.05, 0.1, 0.2\}$ and curvature values $\{0.1, 0.25, 0.5\}$, with 3,000 updates at $\delta_{\mathrm{E}} = 0.025$. The full nonlinear scan crosses line, grid and bottleneck graphs, three weight realizations, amplitudes $\{0.0125, 0.025, 0.05, 0.1, 0.2, 0.4\}$, and curvatures $\{0, 0.1, 0.25, 0.5, 0.75, 1\}$, giving 324 conditions. It uses $\gamma = 0.76$, $\delta_{\mathrm{E}} = 0.025$, relative input-normalized residual threshold 10^{-8}, and at most 30,000 updates. Activity exceeding 10^8 is recorded as numerical divergence.

Of the 324 conditions, 303 converge and 21 diverge. The line has 90 convergent and 18 divergent conditions, the grid 108 convergent conditions, and the bottleneck 105 convergent and three divergent conditions. Among convergent conditions, maximum path KL is 0.16165, 0.08190 and 0.06449, respectively. Increasing amplitude or curvature increases field distortion, while weak-input conditions approach the linear calculation. Ninety further integrations retest all 21 divergent conditions and nine moderate conditions at $\delta_{\mathrm{E}} \in \{0.0125, 0.025, 0.05\}$ with matched integrated duration 750. All 90 integrations preserve the convergence classification of the amplitude scan: the 63 retests of divergent conditions diverge, and all 27 retests of the moderate conditions converge. The observed boundary therefore persists across the three integration steps.

C.6 Reference variation, computation time and sampling budgets

A longer reference lifetime gives distant continuations more weight. It also changes how long the field takes to settle, how much information a compressed representation must retain, and how far sampled events may extend. The following experiments measure these consequences, alongside the precision gained by drawing more events.

Recurrent convergence. For line, grid and bottleneck graphs, we cross $\gamma \in \{0.38, 0.72, 0.76, 0.91, 0.97\}$, Euler ratios $\delta_{\mathrm{E}} \in \{0.025, 0.05, 0.1, 0.25, 0.5, 1\}$ and tolerances $\{10^{-6}, 10^{-8}, 10^{-10}\}$, giving 270 conditions. The stopping criterion is $\|u - \gamma P_0 u - \varepsilon_T\|_\infty / \|\varepsilon_T\|_\infty$ below tolerance, with a limit of 50,000 updates. All 270 conditions converge. At $\delta_{\mathrm{E}} = 0.05$ and tolerance 10^{-10}, the line requires 1,767 updates at $\gamma = 0.76$ and 13,724 at $\gamma = 0.97$; grid values are 1,640 and 13,165, and bottleneck values 1,673 and 13,312. Longer reference horizons increase the time needed to establish the continuation field. Twelve compressed-loop Euler checks compare four dimensions per graph against their compressed analytical equilibria. All twelve compressed-loop checks converge; their largest relative coordinate error is 1.64×10^{-9}. Seventy-two readout checks vary the positivity floor over 10^{-14} through 10^{-4}; the largest within-condition change in path KL across those floors is 0.000260.

Finite sampling. Line, grid and bottleneck graphs are sampled at budgets $\{1{,}000, 4{,}000, 16{,}000, 64{,}000\}$ with sharp and flat evidence across three graph realizations, giving 72 conditions. Events are generated one transition at a time until natural termination. A 4,096-step safety limit raises an error if any event remains active. Thirty-six additional conditions estimate each row of the successor representation from $\{100, 400, 1{,}600, 4{,}000\}$ events per fixed initial state. Figure 15 shows the reduction of endpoint and occupancy errors with sample budget.

Reference survival and finite horizons. The survival scan uses nine values $\{0.38, 0.60, 0.72, 0.74, 0.76, 0.78, 0.85, 0.91, 0.97\}$ at 35% and full dimension across the eight graphs and three realizations, giving 432 conditions. Separately, 1,152 tail calculations cross all eight graphs, nine survival values, sharp and flat evidence, and transition limits $L \in \{32, 48, 70, 72, 82, 92, 128, 256\}$. The forced-truncation probability is $s^\top T^{L+1}\mathbf{1}$, the probability that the event would continue beyond the cap. At the 82-transition cap, the largest sharp-evidence probability across graphs is 2.24×10^{-10} at survival 0.76, 6.36×10^{-4} at 0.91 and 0.11668 at 0.97. At $L = 256$ and $\gamma = 0.97$, it decreases to 6.38×10^{-4}. The required sequence horizon depends on the reference lifetime. At fixed dimension, changing this lifetime also changes

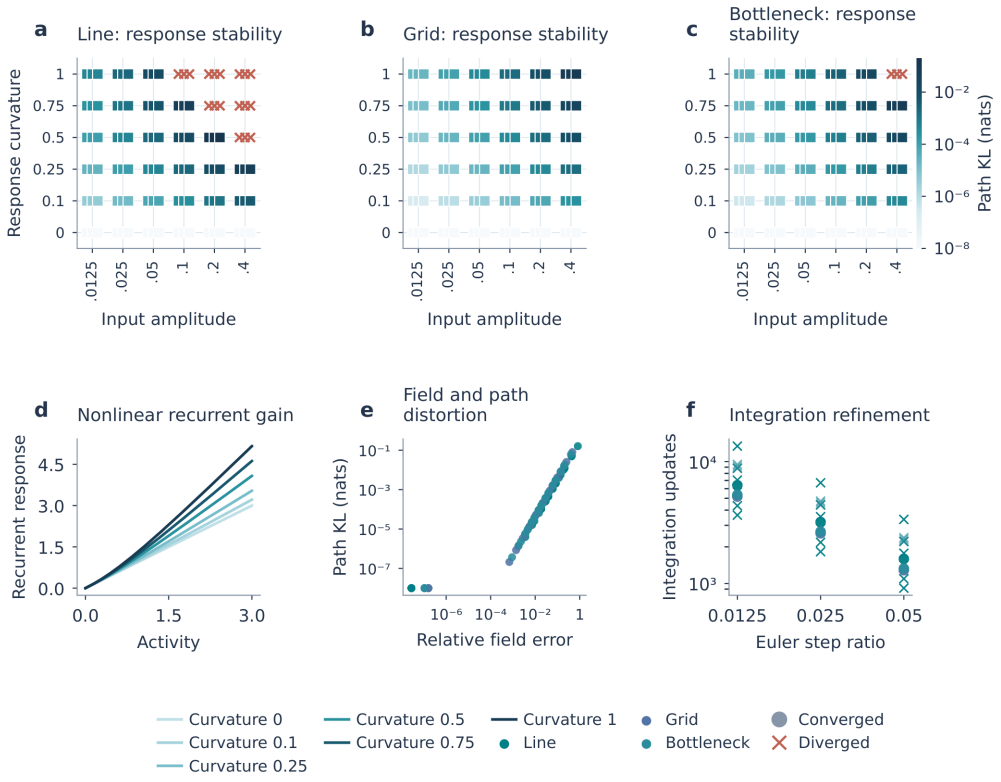

Figure 14: **Nonlinear recurrence changes field accuracy and stability.** (a–c) Every amplitude–curvature condition on three maps. Three horizontally offset marks show the individual realizations: square color encodes convergent path KL and a cross marks divergence. (d) The specified recurrent response at every tested curvature. (e) Field error and path KL for all 303 convergent conditions. (f) All 90 step-size retests, showing updates until convergence or divergence. Circles denote convergence and crosses divergence; colors identify graphs in (e,f). Common scaling of terminal evidence preserves the analytical target throughout the amplitude scan.

compression error. Repeated propagation gives slowly decaying spatial modes more relative weight, which can improve the representation of the continuation field even as recurrent settling becomes slower. The full scan retains the topology dependence of this effect: increasing survival from 0.76 to 0.97 lowers mean path KL on five of the eight maps, while the other three increase.

Directed reference maps. A directed grid at $\gamma = 0.74$ gives endpoint TV 0.01710 from sampled paths and analytical endpoint TV 1.21×10^{-16}. The systematic directionality test applies independent directional log weight perturbations of strengths $\{0, 0.05, 0.1, 0.22, 0.4, 0.8\}$ to every graph realization, yielding 144 conditions. The stationary distribution is recomputed for each matrix. The detailed-balance residual $\frac{1}{2}\sum_{ij} |\pi_i P_{ij} - \pi_j P_{ji}|$ quantifies directionality. The local rules use the resulting directed matrix directly. Across all 144 directed conditions, maximum analytical path KL is 7.04×10^{-16} and endpoint TV is at most 2.14×10^{-16}. Exact local sampling therefore extends to these directed reference maps.

Connection construction and projection. A further 864 conditions change self-loop scale over $\{0, 0.25, 1, 4\}$, special-edge scale over $\{0.25, 0.5, 1, 2\}$, additional symmetric log weight heterogeneity over $\{0, 0.01, 0.025, 0.1, 0.25\}$, and diagonal floor over $\{0, 0.01, 0.02, 0.05, 0.1\}$. Special edges are positive off-diagonal weights below 0.5, with the two designated shortcut edges used for the shortcut graph. Each modified map is evaluated at 35% and full dimension. Another 510 conditions vary construction-noise scale over $\{0, 0.25, 1, 4, 10\}$, construction diagonal floor over $\{0, 0.01, 0.02, 0.05, 0.1\}$, and the lower endpoint of within-module uniform weights over $\{0.1, 0.3, 0.55, 0.8, 1\}$. The modular parameter applies only to the modular graph. Random draws are paired across each scale manipulation. Finally, 216 Euclidean-projection controls vary the spectral rescaling threshold over $\{0.95, 0.985, 0.999\}$ at 10%, 35% and 70% dimension. Every connection-construction condition is evaluable. Full-dimensional KL remains below 6.84×10^{-16} in the connection scan and 5.33×10^{-16} in the additional construction scan. At 35% dimension the respective median KL values are 0.17523 and 0.17857. None of the 216 threshold controls activates spectral rescaling. The primary stationary-weighted projection uses neither a ridge penalty nor spectral rescaling.

C.7 Learning initial and terminal distributions

Motivation and target. Endpoint constraints specify how often sequences should begin and end at each state. The circuit can compare these target frequencies with its own generated frequencies. This experiment asks whether that discrepancy is sufficient to learn the required coupling between beginnings and endings. On grid and modular maps, set $\gamma = 0.74$ and $M = \mathrm{diag}(\pi)(I - \gamma P_0)^{-1}$. The endpoint joint distribution is

$$J_{ij}(\theta, \eta) = \frac{e^{\theta_i} M_{ij} e^{\eta_j}}{\sum_{ab} e^{\theta_a} M_{ab} e^{\eta_b}}. \tag{119}$$

Target marginals μ_0, μ_T are normalized sums of two Gaussian components plus background 0.02, with width $0.55 \max\{\mathrm{std}(x), \mathrm{std}(y), 1\}$. The initial primary center minimizes $x + 0.1y$ and the terminal primary center is its farthest state. Secondary centers use indices $\lfloor N/4 \rfloor$ and $\lfloor 3N/4 \rfloor$. Sinkhorn scaling of the same M to the same marginals supplies the reference joint distribution. Its alternating updates are $a \leftarrow \mu_0/(Mb)$ and $b \leftarrow \mu_T/(M^\top a)$, with elementwise division. The default stopping tolerance on the larger marginal TV is 10^{-12}, within 20,000 iterations.

Frequency-difference updates. Starting at $\theta = \eta = 0$, update

$$\theta_i \leftarrow \theta_i + \alpha_t[\widehat{\mu}_0(i) - \widehat{q}_0(i)], \qquad \eta_i \leftarrow \eta_i + \alpha_t[\widehat{\mu}_T(i) - \widehat{q}_T(i)], \qquad \alpha_t = \frac{\alpha_0}{(1 + t/d)^p}. \tag{120}$$

The baseline is $(\alpha_0, d, p) = (1.2, 250, 0.72)$. Each vector is recentered to zero mean after the update and clipped to $[-18, 18]$. Exact-frequency conditions use the analytical marginals. Sampled-frequency conditions draw wake counts from the target marginals and replay endpoint pairs from the current joint table; joint draws preserve dependence between the two endpoints. The learning experiment evaluates these updates with the current ideal endpoint distribution. Recurrent settling and compressed-field errors are studied in the preceding experiments.

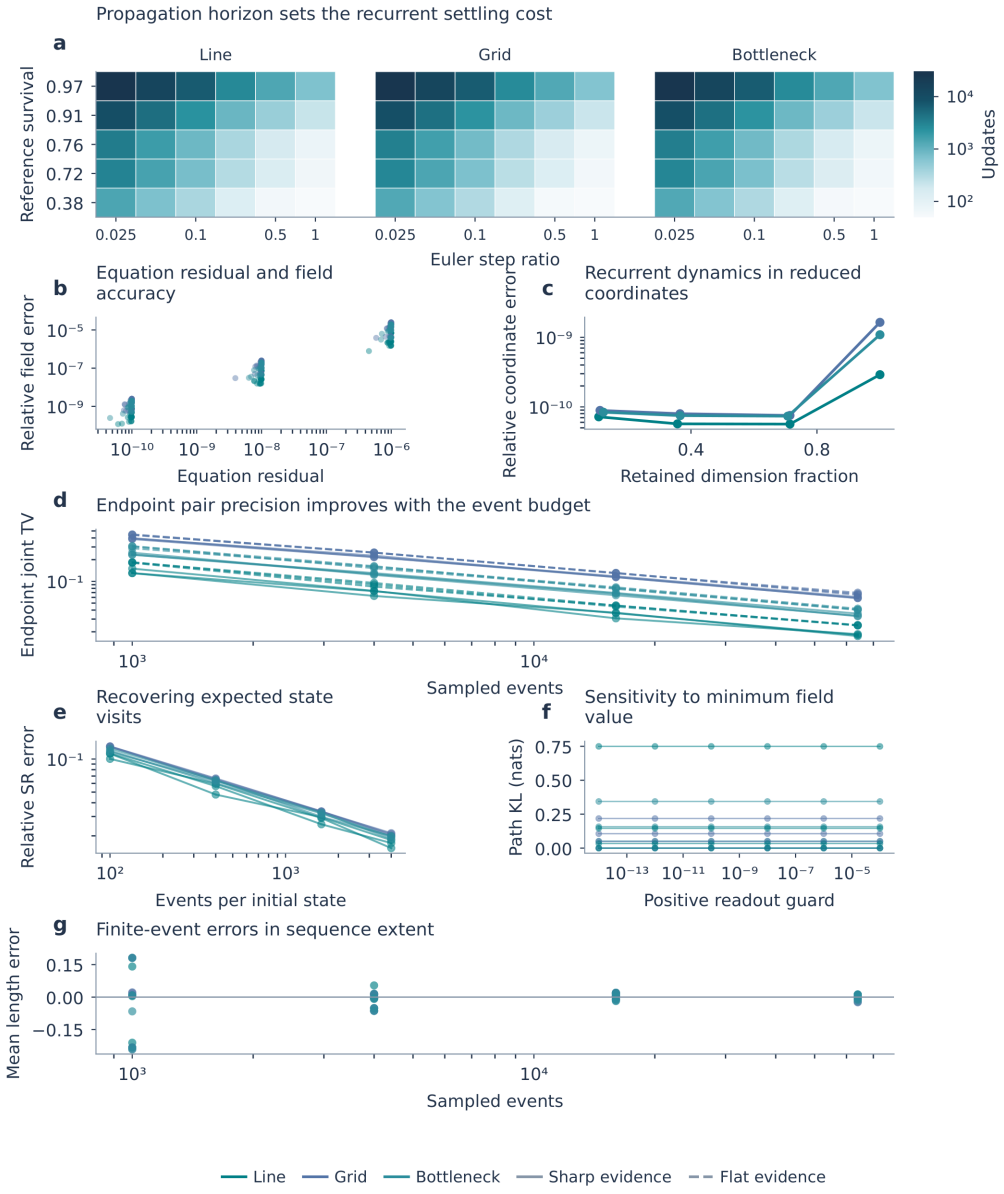

Figure 15: **Recurrent settling and finite sampling have separate precision budgets.** (a) Updates to residual 10^{-10} across three maps, all five survival values and six Euler ratios. Each cell is one integration; the common logarithmic color scale gives its settling cost. (b) Residual and field error for all 270 integrations. (c) Twelve compressed-loop equilibrium comparisons. (d) Sampled endpoint joint TV; solid curves use sharp evidence, dashed curves flat evidence. (e) Relative successor-representation error. Every sampling curve is one realization. (f) Positive-guard sensitivity, with one curve per graph–dimension pair. (g) Sample mean minus analytical mean event length for all 72 sampling conditions; the horizontal line marks agreement.

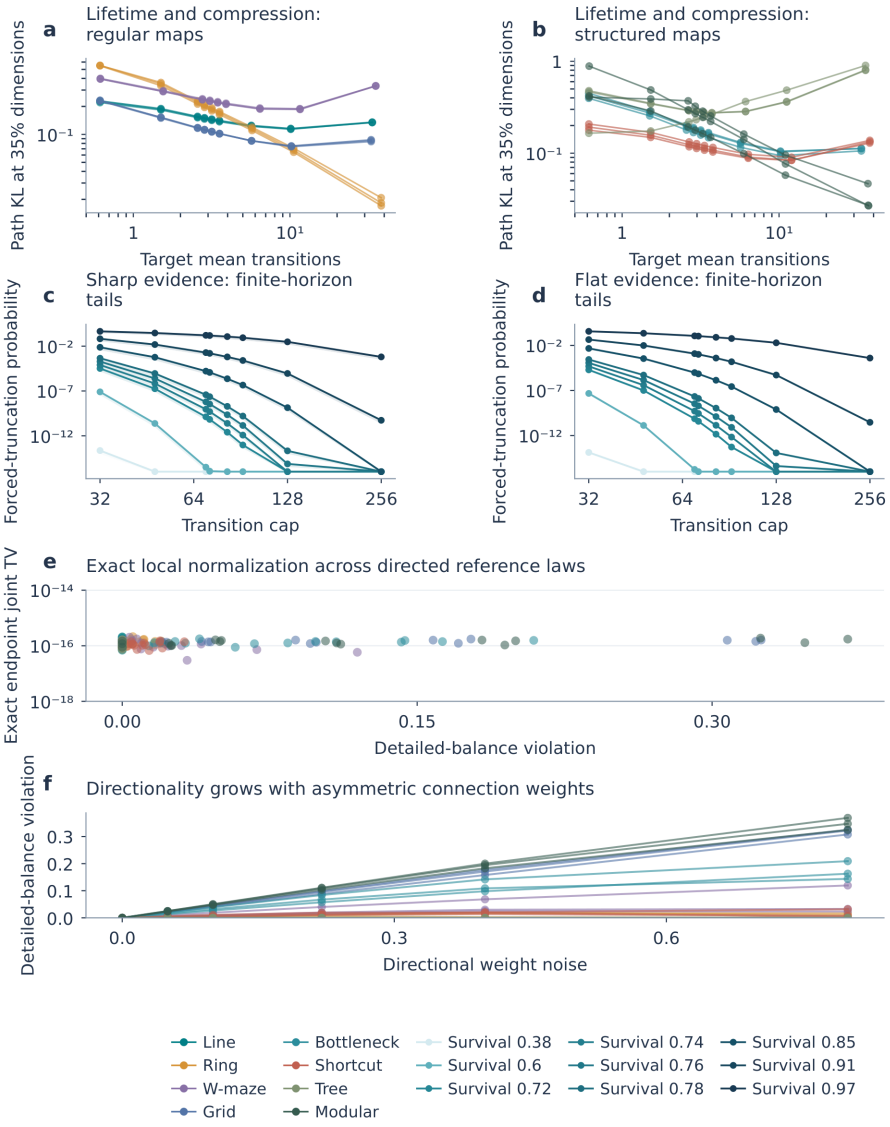

Figure 16: **Reference lifetime and directionality affect different computational properties.** (a,b) Target mean transitions and path KL at 35% dimension, with every graph realization followed through increasing survival. (c,d) Finite-horizon tails for sharp and flat evidence. Lines show the maximum forced-truncation probability over eight graphs; bands span their minimum–maximum range. (e) Analytical endpoint TV against detailed-balance violation for all 144 directed conditions. (f) Detailed-balance violation under increasing directional weight noise, retaining each realization. Graph colors agree throughout. Display floors are 10^{-8} for KL, 10^{-18} for endpoint TV and 10^{-16} for tails.

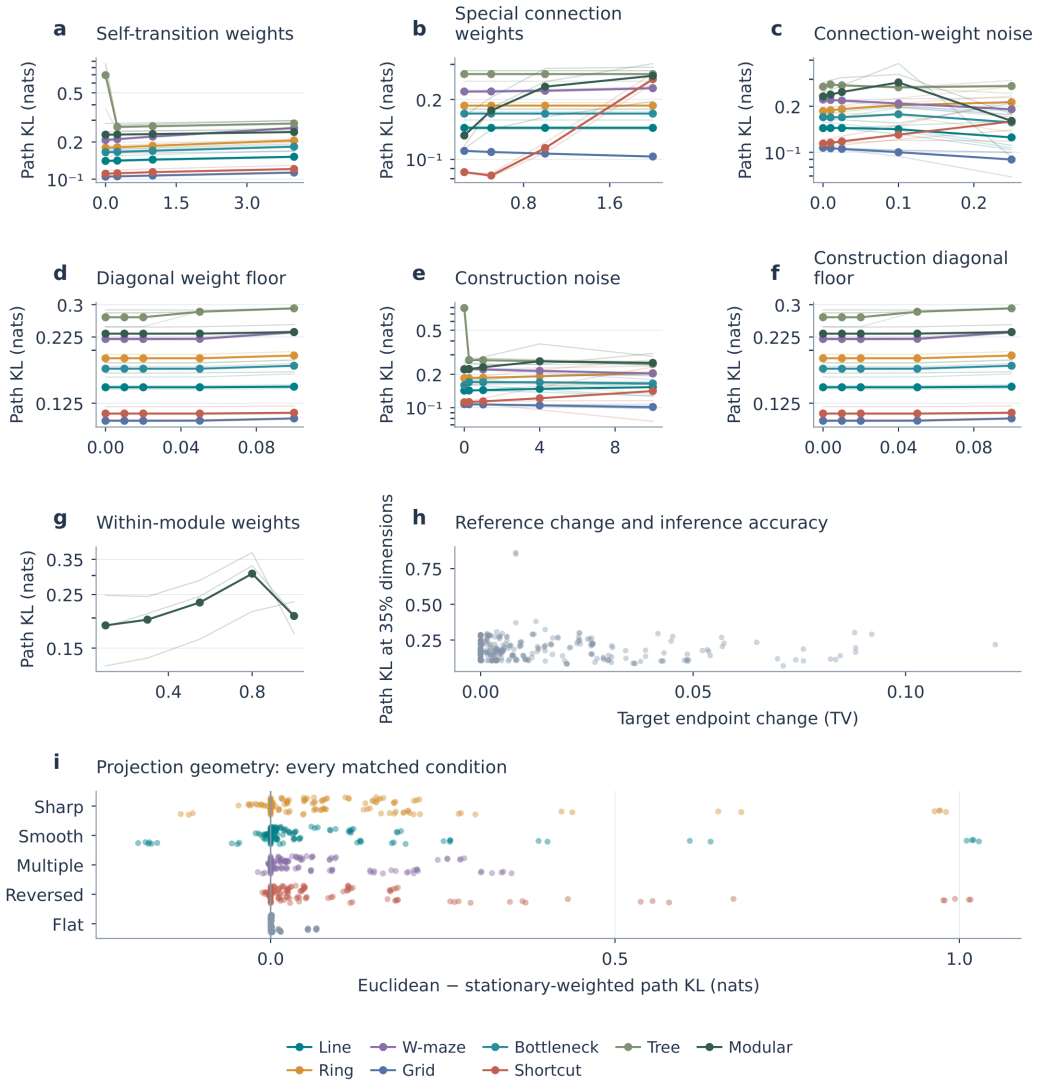

Figure 17: **Connection construction and projection geometry shape compressed inference.** (a–g) Parameter-response profiles at 35% dimension. Thin curves retain individual weight realizations and bold curves their means; every tested parameter value is shown. Within-module minimum weight applies to the modular graph. (h) Target endpoint change and compressed KL for the four direct weight manipulations. (i) All 720 matched Euclidean-minus-stationary-weighted KL differences, grouped by evidence family. Each point compares the same graph, realization and dimension; the zero line marks equal errors. Each modified reference is evaluated against its own analytical target.

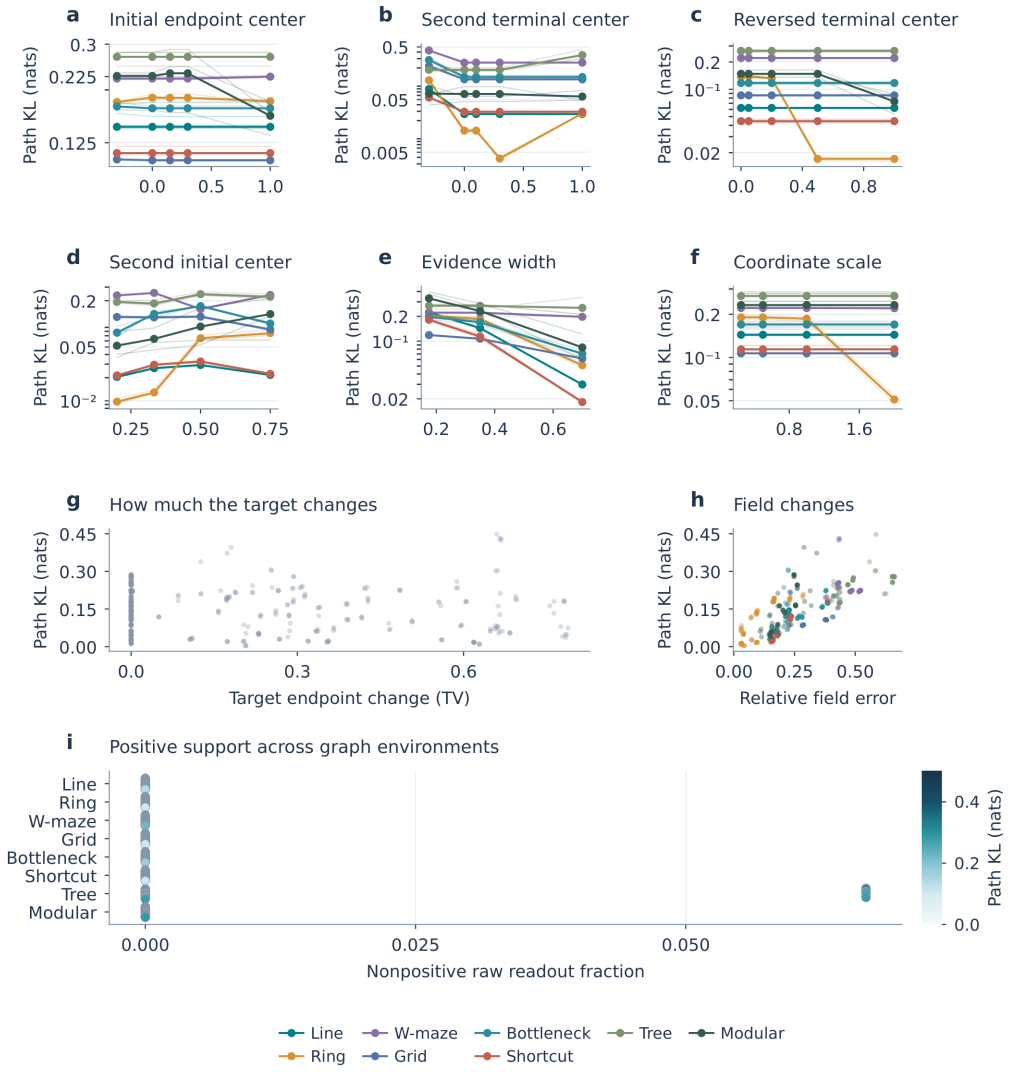

Figure 18: **Endpoint construction changes the target and its representational demands.** (a–f) Parameter-response profiles for centers, widths and coordinate scale at 35% dimension. Thin curves retain individual realizations; bold curves give means. Evidence is multiple-center in (b,d), reversed in (c), and sharp in (a,e,f). (g,h) Target endpoint change and field distortion versus path KL. (i) Nonpositive raw-field fraction by graph. Circle color encodes path KL on a scale from zero to 0.5 nats; small vertical offsets separate individual conditions within each graph. Each of (g–i) retains all 624 conditions; the complete values are provided in the source tables.

The four learning comparisons each use 9,000 updates. The sampled mode draws 16,384 wake samples per endpoint and 16,384 replay endpoint pairs per update. Final endpoint joint TVs are 0.002724 and 0.003246 on the grid for exact and sampled modes, and 0.000168 and 0.001113 on the modular map. Corresponding dual objective gaps are 9.84×10^{-5}, 1.07×10^{-4}, 4.71×10^{-7} and 4.12×10^{-6}. Endpoint-frequency differences guide the system toward a joint law that also preserves the connections supplied by the reference map. The learning rule need not count every complete trajectory: the dependence between beginnings and endings is carried by the common matrix M, while the two endpoint populations adjust their relative weights.

Learning schedules and sample budgets. The learning parameter scan contains 48 exact-frequency conditions and 24 sampled-frequency conditions. Exact conditions vary one parameter at a time on one realization per graph: $\alpha_0 \in \{0.3, 0.6, 1.2, 2.4, 4.8\}$, $d \in \{100, 250, 1000\}$, $p \in \{0, 0.5, 0.72, 1\}$, and clipping magnitude in $\{6, 12, 18, 30\}$. Target changes use width coefficients $\{0.275, 0.55, 1.1\}$, backgrounds $\{0, 0.005, 0.02, 0.1\}$, initial secondary-center fractions $\{0.1, 0.25, 0.5\}$, terminal fractions $\{0.5, 0.75, 0.9\}$, and localization coefficients $\{-0.2, 0.1, 0.3\}$. Shared baseline conditions are counted once. Sampled conditions cross two graphs, three independent weight realizations and batch sizes $\{256, 1024, 4096, 16384\}$.

Every 100 updates, the simulation evaluates the exact marginal TV and records the exact joint TV to the Sinkhorn reference. The stopping criterion is three consecutive marginal checks below 0.001, within 30,000 updates. Runs continue through at least update 9,000 to preserve a common comparison point. Forty-six of the 72 conditions meet this criterion; 26 reach the update limit. Joint TV and marginal TV quantify different errors, so their thresholds need not coincide.

At update 9,000, increasing the sampled batch from 256 to 16,384 lowers mean grid joint TV from 0.01211 to 0.003124, and modular TV from 0.009323 to 0.001282. All three realizations at the largest batch meet the stopping criterion on each graph. In exact-frequency learning, increasing α_0 from 0.3 to 4.8 changes grid joint TV at update 9,000 from 0.06916 to 1.88×10^{-6}, and modular joint TV from 0.01656 to 4.07×10^{-9}. These parameter comparisons describe convergence speed and frequency-estimation precision on the specified targets. Six Sinkhorn tolerance checks use 10^{-8}, 10^{-10} and 10^{-12} on the two graphs.

C.8 Analytical observables and illustrative conditions

The bottleneck illustration uses graph seed 20260524, smooth endpoint evidence, $\gamma = 0.76$, and 15 retained dimensions for its 41 states. These choices specify the exact and compressed fields used in Figures 2 and 10. The comparison holds the state space, reference transitions, and endpoint evidence fixed. All state and connection values are provided in the accompanying source tables.

Occupancy and connection use. For the local law (s, T, r), let $v = s^\top (I - T)^{-1}$ be the expected visits. The expected number of traversals of edge $i \to j$ is $v_i T_{ij}$. The difference $v_i T_{ij} - v_j T_{ji}$ gives net expected flow between two states. Self-transitions contribute to the calculation. Step-specific occupancy is $m_0 = s$, $m_{t+1} = T^\top m_t$. Its entry $m_t(i)$ is the probability that the event reaches state i after t transitions, including events that stop there. The reduction in total mass between successive steps is the probability of completion at that step. The displayed occupancy window extends through transition 24; all expected visits and endpoint quantities are calculated over the full infinite horizon.

The joint distribution of beginnings and endings. With $S = (I - \gamma P_0)^{-1}$, the exact endpoint coupling is

$$J^\star_{ij} = \frac{\pi_i \varepsilon_0(i) S_{ij} \varepsilon_T(j)}{\sum_\ell \pi_\ell \varepsilon_0(\ell) \phi_\ell}. \tag{121}$$

The compressed-field coupling is $J_{ij} = s_i [(I - T)^{-1}]_{ij} r_j$, with the locally normalized probabilities in Equation 116. Thus its endpoint error includes every change in initiation, transitions, and stopping induced by field approximation. In contrast, a direct substitution of the approximate field into the exact coupling would generally omit these normalization effects.

Summaries over computational conditions. Rank curves retain each integer k, and requested fractions use $k = \lceil fN \rceil$. The matched width comparison fixes graph, realization, and dimension

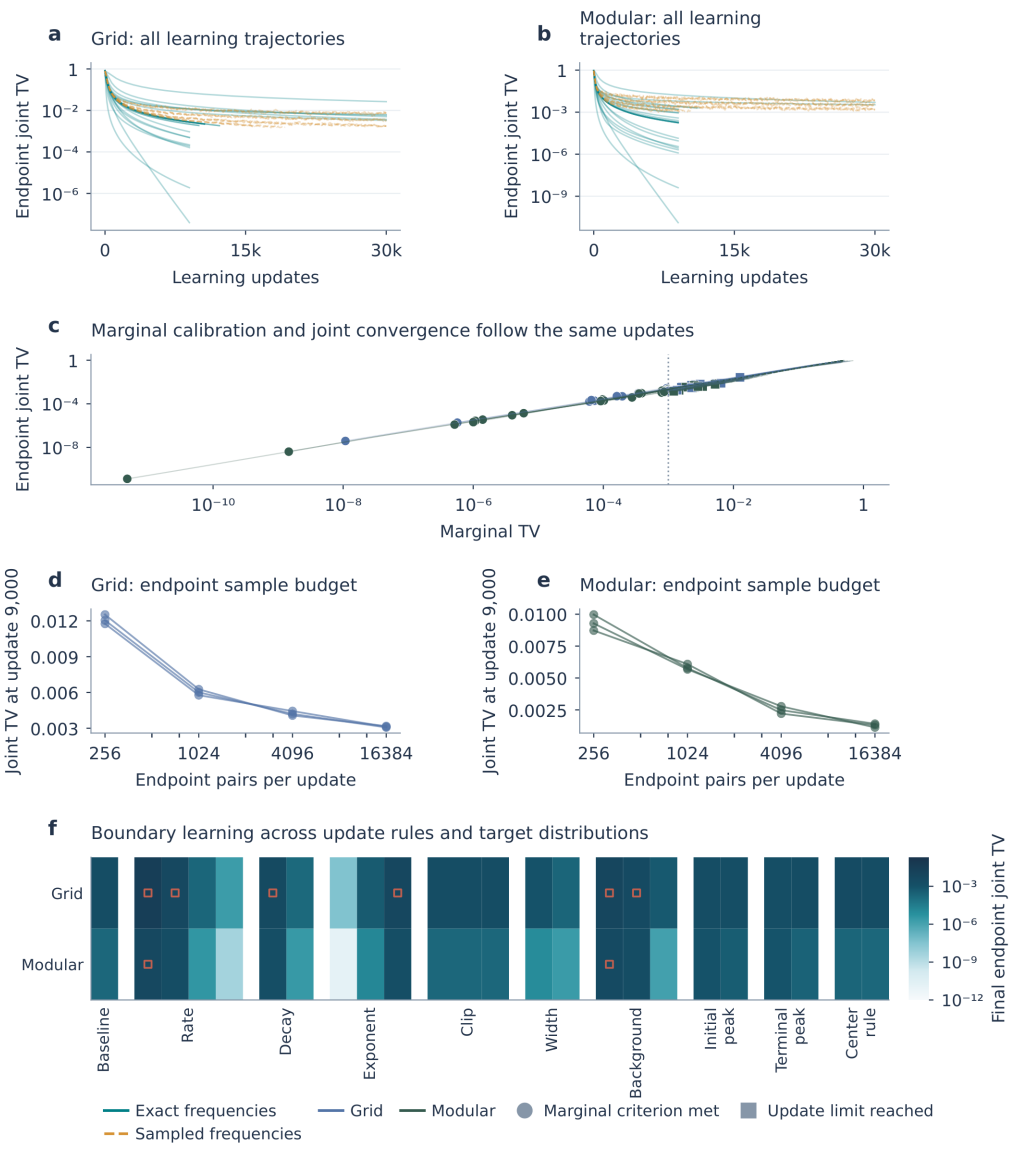

Figure 19: **Endpoint-frequency learning connects marginal calibration with joint convergence.** (a,b) Every checkpoint in all 72 learning conditions, grouped by map and exact or sampled frequencies. (c) Their joint–marginal error trajectories; circles mark runs meeting the marginal criterion, squares the update limit. The vertical line is marginal TV 0.001. (d,e) Joint TV at update 9,000 against sample batch size, with a curve per realization. (f) Every exact-frequency condition, grouped by schedule or target parameter and ordered by tested value. Each cell shows one final joint TV; an open terracotta square marks the update limit.

and measures $\mathrm{KL}_{\mathrm{sharp}} - \mathrm{KL}_{\mathrm{smooth}}$. The width–background summaries pool 18 combinations per graph, width, background, and dimension: three realizations, three initial locations, and two terminal locations. Medians use evaluable conditions; numerical failures remain counted separately. The reported fraction of nonpositive raw-field entries is $N^{-1} \sum_i \mathbf{1}\{(Cu)_i \leq 0\}$ before the positivity guard. Mean-length distortion is $|\mathbb{E}_Q K / \mathbb{E}_{Q^\star} K - 1|$ for the corresponding pair of generated and target laws. These summaries use the specified finite parameter grids. Their means and ranges describe those grids; deterministic matrix calculations have no sampling interval.

Computational materials. The accompanying files contain the 140 basic records, 21,678 extended condition records, 40 analytical checks, 726 basis-rank checks, numerical implementations, saved learning checkpoints, and figure source tables. The seeds, parameter grids, integration criteria, and sample budgets are specified above. Full dimension scans use the deterministic spectral construction in Section C.4, from which every tested basis can be reconstructed. The fixed-representation demonstrations retain outcomes and dimension settings; their archived inputs do not contain every fitted basis matrix or random execution state.

D EXPERIMENT 2: REUSABLE TEMPORAL STRUCTURE IN HIPPOCAMPAL POPULATION ACTIVITY

D.1 FROM A REFERENCE MAP TO A TEST OF SEQUENCE ORDER

Boundary conditioning gives a learned map a new use while preserving its relations among possible paths. This experiment asks whether those relations can be identified in hippocampal population activity. Consider several sequences that begin in the same state, end in the same state, and contain the same number of visits to each state. Endpoint weights multiply all these sequences equally. Their relative probabilities depend on the order in which the states occur. Comparing observed orders with rearrangements having these shared properties isolates the transition structure that the bridge preserves.

Experiment 2 develops this comparison in three steps. First, HC-3 establishes how the transition law is estimated from spikes and whether it predicts held-out events. Second, HC-11 tests reuse across PRE sleep, MAZE behavior, and POST sleep within fixed population coordinates. Third, HC-11 and DANDI:000978 test internal order after consecutive repeats are removed and possible next states are matched. The final analysis separates running from stationary task activity. Table 5 connects each comparison to its scientific purpose.

Table 5: The experimental logic. Every comparison uses held-out events and a fixed fitted reference.

Question	Comparison	What it establishes
Do transitions carry predictive information?	Ordered likelihood versus start-conditioned visitation readouts in HC-3.	The value of using the learned dynamics sequentially.
Can the reference be reused?	PRE and MAZE references each score PRE, MAZE, and POST events in HC-11.	Reusable relations together with epoch-specific prediction.
Does the population representation matter?	Independent 6-, 12-, and 28-state partitions and the most active cell.	Dependence on distributed activity and resolution.
Does internal order matter with boundaries fixed?	Observed jump sequences versus legal orders preserving endpoints and state counts.	The relative path preferences retained under endpoint weighting.
How does behavior affect the comparison?	Paired running and stationary estimates for each animal.	Condition-specific effects and their sampling uncertainty.

Two complementary measurements run through the experiment. Prediction gain compares the reference with a visitation readout on held-out events. An order effect compares an observed sequence with rearrangements under the same fixed reference. The first measures the accuracy of a predictive readout; the second identifies the temporal information expressed within a controlled set of candidate paths. A reference can prefer the observed order within that set even when another readout better predicts the broader collection of events. Experiment 3 then supplies explicit behavioral boundaries and asks how they reshape complete episodes.

A state here is a cluster of population spike-count vectors, identifying a recurring firing pattern across the recorded neurons. Events are detected from population activity, and position measurements identify running and stationary conditions. State labels consequently refer to neural patterns. Experiment 3 supplies a spatial interpretation through place-field decoding, and Experiments 3 and 4 examine correspondence with independently detected ripple-band activity.

D.2 DATA AND INDEPENDENT SAMPLING UNITS

The HC-3 analysis comprises 27 recording sessions, with 16 CA1 and 11 CA3 session assignments (Mizuseki et al., 2014; 2013). Its cohort summaries are session means. Multiple sessions originate from the same animal, so their intervals describe variation among recordings. The detailed recording-level analysis covers nine sessions from the ec013 recording series: ec013.374, ec013.387, ec013.665,

ec013.669, ec013.670, ec013.932, ec013.978, ec013.979, and ec013.982. Recording-level displays use these nine sessions; cohort summaries use all 27 recordings.

HC-11 supplies eight PRE–MAZE–POST recordings from four animals (Grosmark & Buzsáki, 2016; Grosmark et al., 2016). The sessions are Achilles_10252013, Achilles_11012013, Buddy_06272013, Cicero_09012014, Cicero_09102014, Cicero_09172014, Gatsby_08022013, and Gatsby_08282013. PRE and POST use annotated non-REM intervals. MAZE uses task intervals with finite estimated speed of at least 0.03 m/s. For the 12-state analysis, the data contain 42,119 PRE training events, 22,920 PRE test events, 7,344 MAZE test events, and 51,597 POST events. Retained CA1 pyramidal populations range from 41 to 96 units per session.

DANDI:000978, version 0.240511.0307 (Shin & Jadhav, 2024), supplies nine sessions from eight animals: ER1, JS14, JS15, JS17, JS21, JS34, KL8, and ZT2, with two recording days for ZT2. The analysis uses the designated PRE interval and the following task interval in each file. PRE in this dataset denotes the designated recording epoch, while HC-11 PRE specifically uses annotated non-REM sleep. Position is converted from centimetres to metres, and speed is recomputed from timestamped two-dimensional coordinates. Running task bins satisfy the same 0.03 m/s threshold. The task test set contains 3,504 original events and 3,222 retained jump events, with 16,937 jump transitions. HC-11 and DANDI group means first average sessions within each animal and then average animals equally.

D.3 From spikes to a state sequence

The representation is fitted once on its training support. Applying it to later spikes produces an event sequence without changing the meaning of a state. Table 6 summarizes which observations determine each fitted quantity and which observations are subsequently scored.

Table 6: Fitting support and temporal evaluation. All percentages refer to eligible bins within the stated support. Later epochs use the fixed PRE units, feature transform, and centroids.

Analysis	Fitted quantities and support	Held-out evaluation
HC-3	Threshold, standardization, centroids, and transitions: first 65% of recording bins.	Final 35% of recording bins.
HC-11 PRE reference	Units, threshold, transform, centroids, and transitions: first 65% of PRE NREM support.	Remaining PRE; final 35% of MAZE running support; POST NREM.
HC-11 MAZE reference	Same PRE state representation; transitions fitted on first 65% of MAZE running support.	Same PRE, MAZE, and POST test events as the PRE reference.
DANDI PRE reference	Units, threshold, transform, centroids, and transitions: designated PRE training support.	PRE test events and final 35% of eligible task support.
Behavioral comparison	The fixed PRE representation and jump reference.	Final 35% of running and stationary supports, each partitioned separately.

Temporal sampling and event construction. Spikes are counted in 20 ms bins. A bin is active when its total population count exceeds the fitted activity threshold, with a minimum threshold of two spikes. In HC-3 the threshold is the 65th percentile of positive population counts in the chronological training segment. In the PRE-defined HC-11 and DANDI analyses it is the 75th percentile of positive PRE training counts. The PRE threshold remains fixed when MAZE, task, or POST activity is scored.

Active bins are grouped into events when successive selected bins have at most three intervening inactive bins during task activity or HC-3 recording, and at most two during PRE or POST activity. An event requires at least three selected bins. Longer runs are divided into chunks of at most 50 selected bins; a final chunk shorter than three is discarded. Inactive gaps are omitted from the state string. Thus one step indexes the next selected population state, and the sequence likelihood is measured per state rather than per second of physical time.

Training support. HC-3 uses the first 65% of recording bins to fit the activity threshold, feature transform, centroids, and transition law; the remaining 35% is held out. Its pyramidal-unit metadata and spike-count eligibility are determined from the recording. HC-11 uses the first 65% of eligible PRE non-REM bins for every selection and fitted transform. Candidate CA1 pyramidal cells require at least 40 spikes on that PRE training support. Cells are ranked by their PRE training count, ties follow metadata order, and at most 96 are retained. The remaining PRE bins are held out. Eligible MAZE running bins are separately split chronologically at 65%; only the later part is used for MAZE test scoring. All eligible POST events are scored. DANDI similarly fixes the unit set from PRE training activity, requiring at least 40 spikes and a PRE training rate no greater than 5 Hz, with at most 96 retained units. Its PRE cutoff is fixed by the designated epoch, and its task running support is split chronologically at 65%.

Population states. For unit u, the feature is $\log(1 + c_u)$, where c_u is its bin spike count. Training means and standard deviations standardize each feature, with standard deviations bounded below by 10^{-3}. Euclidean k-means is fitted to active training vectors with 45 iterations and three random initializations. The initialization with the smallest mean squared distance is retained. At most 200,000 training vectors enter a fit, selected by the fixed random generator when this cap is needed. Test vectors are assigned to the closest fixed centroid. The primary alphabet has 12 states; independent 6-state and 28-state fits test resolution. A fourth representation records the identity of the unit with the largest raw count in the bin. Ties in this winner-cell representation follow the fixed unit order. It summarizes the most active unit; the other recorded neurons remain part of the underlying observation.

The base random seed is 7 for HC-3 and HC-11, with a $17k$ offset for a k-state clustering fit. DANDI uses base seed 20,260,710 and the same offset. Each state representation has its own fitted transition law. Raw likelihoods across different alphabets are therefore interpreted within their respective state spaces. Resolution comparisons use the likelihood difference against the corresponding visitation readout.

D.4 Reference and comparison readouts

Let n_{ij} count the training transitions from state i to state j, and let m be the alphabet size. The reference transition kernel is

$$P_0(i, j) = \frac{n_{ij} + \alpha}{\sum_{\ell=1}^{m} n_{i\ell} + m\alpha}, \qquad \alpha = 0.05. \tag{122}$$

The scored reference uses the stationary distribution π of this kernel for the initial state. The empirical event-start distribution is fitted separately for the start-frequency comparator. For an event $x_{0:K}$, the length-conditioned reference likelihood is

$$R(x_{0:K} \mid K) = \pi(x_0) \prod_{t=0}^{K-1} P_0(x_t, x_{t+1}). \tag{123}$$

The reported loss sums negative log likelihood over all test events and divides by their total number of states, including initial states. Conditioning on each observed length removes the shared stopping-length factor. This is appropriate because the event detector imposes minimum lengths and splits long runs.

Successor visitation. The successor matrix is $S = (I - \gamma P_0)^{-1}$. Here γ is the number of training transitions divided by that number plus the number of training events, clipped to $[0.02, 0.995]$. The SR-IID comparator fixes the observed initial state i_0, normalizes row $S(i_0, \cdot)$ to a probability vector q_{i_0}, and scores subsequent states as independent draws from q_{i_0}. It shares the initial probability $\pi(i_0)$ with the reference. The future-SR comparator instead normalizes row $(S - I)(i_0, \cdot)$. Removing the identity term excludes the visit at time zero while preserving later returns to the initial state. Both readouts retain the event's initial-state-dependent visitation tendency.

These two predictions use different operations on the same learned dynamics. A successor row answers how frequently future activity will visit each state, given a starting state. The path law updates the next-state prediction as the sequence develops. The full successor matrix retains the information needed for this operation: when γ is known, $P_0 = (I - S^{-1})/\gamma$. Numerical reconstruction recovers

both the reference kernel and its event likelihoods from that matrix. Thus the measured gain quantifies the predictive value of using the map sequentially. The matrix identity makes the same transition information available in both descriptions.

Exact event scores and units. For any start-conditioned visitation vector q_i, its event probability is

$$Q_q(x_{0:K} \mid K) = \pi(x_0) \prod_{t=1}^{K} q_{x_0}(x_t). \tag{124}$$

For a collection $\mathcal{E}$ of test events, the prediction gain is

$$G(Q_q, R) = \frac{\sum_{x \in \mathcal{E}} \log\{R(x \mid K_x)/Q_q(x \mid K_x)\}}{\sum_{x \in \mathcal{E}} (K_x + 1)}. \tag{125}$$

The initial-state terms cancel. A positive gain means that the reference assigns greater geometric-mean likelihood per scored state. Original-state and jump-state gains use the same expression on their respective event strings. Order-randomization effects below divide by the number of jump transitions, $\sum_x K_x$, because their comparison concerns transition products within paths whose initial state is held fixed.

Other baselines and model order. Three independent-draw baselines use the training state occupancy distribution, the training event-start distribution, or the uniform distribution. All use the same test events and conditioning on observed length. A second-order model estimates the next-state distribution from the preceding two states, with additive smoothing and a first-order initial transition. A reversible comparison symmetrizes stationary flux, $F_{ij} = \pi_i P_0(i, j)$, through $(F + F^{\mathsf{T}})/2$ and normalizes rows. It measures the predictive cost of detailed balance in the estimated dynamics. These comparisons refit the specified model. The order controls below instead keep the reference fixed while changing the tested sequence.

Visit-count prediction. For each observed start and length, the reference predicts visits by summing the corresponding rows of $I, P_0, \ldots, P_0^K$. Predictions and observed visit counts are averaged within each start state. A Pearson correlation uses the number of events at that start as its weight. This statistic measures finite-horizon visit prediction. It complements the ordered event likelihood.

D.5 A CONCRETE POPULATION SEQUENCE

The example recording ec013.374 contains 15 retained units and 270 held-out events. Its transition matrices are fitted on the chronological training interval and its event sequences come from the test interval. To show how spike activity becomes a state string, the example window is centered on the earliest held-out active run with at least three selected bins. The 2.4 s window begins at 700.33 s and contains 120 consecutive 20 ms bins. This selection uses event timing and length. The full collection of 270 events, including every selected state before termination, is included in Figure 20.

The fitted centroids give the standardized log-count pattern associated with each population state. They use all retained neurons. The state identifiers remain fixed when the trained map is applied to the test interval. Different alphabet sizes have independently fitted centroids and transition matrices, so a numerical label has meaning only within its own representation. Winner-cell labels instead identify the unit with the largest count in a bin. Figure 22 compares these representations on shared observations.

D.6 PREDICTION, ORDER, AND MODEL COMPLEXITY IN HC-3

The first comparison asks whether the reference's sequential readout provides useful predictive information. Relative to SR-IID, the 27-session mean gain is 0.0733 nats/state, with a session bootstrap interval of $[0.0550, 0.0912]$ and positive effects in 26 sessions. Gains against occupancy, start, and uniform independent draws are 0.2904, 0.3204, and 0.6021 nats/state, respectively (Table 7). The SR-IID comparison is informative because it already carries the visitation tendency associated with each starting state. The remaining gain measures predictive information in the succession of population patterns.

Table 7: HC-3 prediction comparisons. Gain is comparator loss minus reference loss, except for the explicitly labelled second-order comparison. Intervals summarize sessions; $n = 27$ recordings.

Comparison	Mean	Positive sessions	95% interval
R over SR-IID	0.0733	26/27	$[0.0550, 0.0912]$
R over occupancy IID	0.2904	27/27	$[0.2344, 0.3502]$
R over start IID	0.3204	27/27	$[0.2614, 0.3829]$
R over uniform IID	0.6021	27/27	$[0.5248, 0.6835]$
Shuffle loss minus observed loss	0.0812	27/27	$[0.0625, 0.1017]$
Second order over first order	-0.0799	1/27	$[-0.0999, -0.0602]$

The ordinary order control fixes the initial state, length, and state multiplicities, then permutes the remaining positions. The HC-3 reference score averages 199 randomizations per recording; SR-IID is evaluated on one paired rearrangement. The final state may change. Each draw starts from the original sequence, with up to 512 attempts to obtain a different order; sequences without a distinct permitted ordering retain their identity. The first-order model's loss increases by 0.0812 nats/state on average in the full cohort, whereas the SR-IID score is invariant to this rearrangement. In the nine sessions with individual diagnostics, the mean reference shuffle loss is 0.093700 nats/state, and the largest absolute SR-IID change is 2.22×10^{-16}. Their visit-count correlations average 0.920466, with range 0.763222–0.975175. The 27-session mean visit correlation is 0.9227. The reference therefore captures both the expected occurrence of population patterns and information about how those patterns follow one another. Rearranging an event preserves its constituent states while changing the transition relations that determine its path likelihood.

Adding a second state of memory reduces held-out performance by 0.0799 nats/state on average. At the present resolution and sample size, estimating a next-state distribution for every pair of preceding states costs more predictive accuracy than it gains. The first-order estimator is therefore the effective reference at this level of observation. Finer temporal or population measurements could resolve additional memory in the neural dynamics.

D.7 A PRE REFERENCE ACROSS BEHAVIORAL EPOCHS

HC-11 asks whether transitions learned from PRE can describe MAZE and POST activity. This is a prospective population-state comparison: units, threshold, standardization, centroids, and R_{PRE} are all learned from PRE training bins. A second reference, R_{MAZE}, is fitted to MAZE training events in the same PRE-defined alphabet. Both references score the same PRE test, MAZE test, and POST events. This gives six evaluations without fitting a POST model.

Table 8 shows complementary features of these evaluations. The PRE reference improves on future-SR in all three epochs, with positive animal means in all four animals. It also assigns higher likelihood to PRE and POST activity than the MAZE reference. The MAZE reference assigns higher likelihood to MAZE activity. The same population coordinates therefore support continuity across epochs and differences in their dynamics. For the theory, the useful observation is that a reference estimated from sleep before exploration already carries sequential information about subsequent activity. Fitting the reference during the task improves task prediction further. This pattern gives the reference a concrete computational role: it supplies reusable relations among population states, while the observed probability of a sequence can vary with the epoch. Experience, behavioral condition, sampling, and state occupancy all contribute to the activity from which each reference is estimated.

The original-state PRE-to-task comparison in DANDI yields a mean gain of -0.054533 nats/state against future-SR, with a descriptive animal bootstrap interval of $[-0.150860, 0.021231]$ and positive effects in two of eight animals. This dataset shows that cross-epoch sequence correspondence and prediction against an alternative readout are separate empirical questions.

A complementary HC-11 representation fit uses MAZE training vectors and a MAZE training activity threshold, while retaining PRE-based cell selection. Its PRE reference gains 0.0760 nats/state over SR-IID on MAZE test events, 0.0734 on PRE test events, and 0.0726 on POST events. The MAZE reference's MAZE gain is 0.097492. Since the alphabet itself uses task observations in this

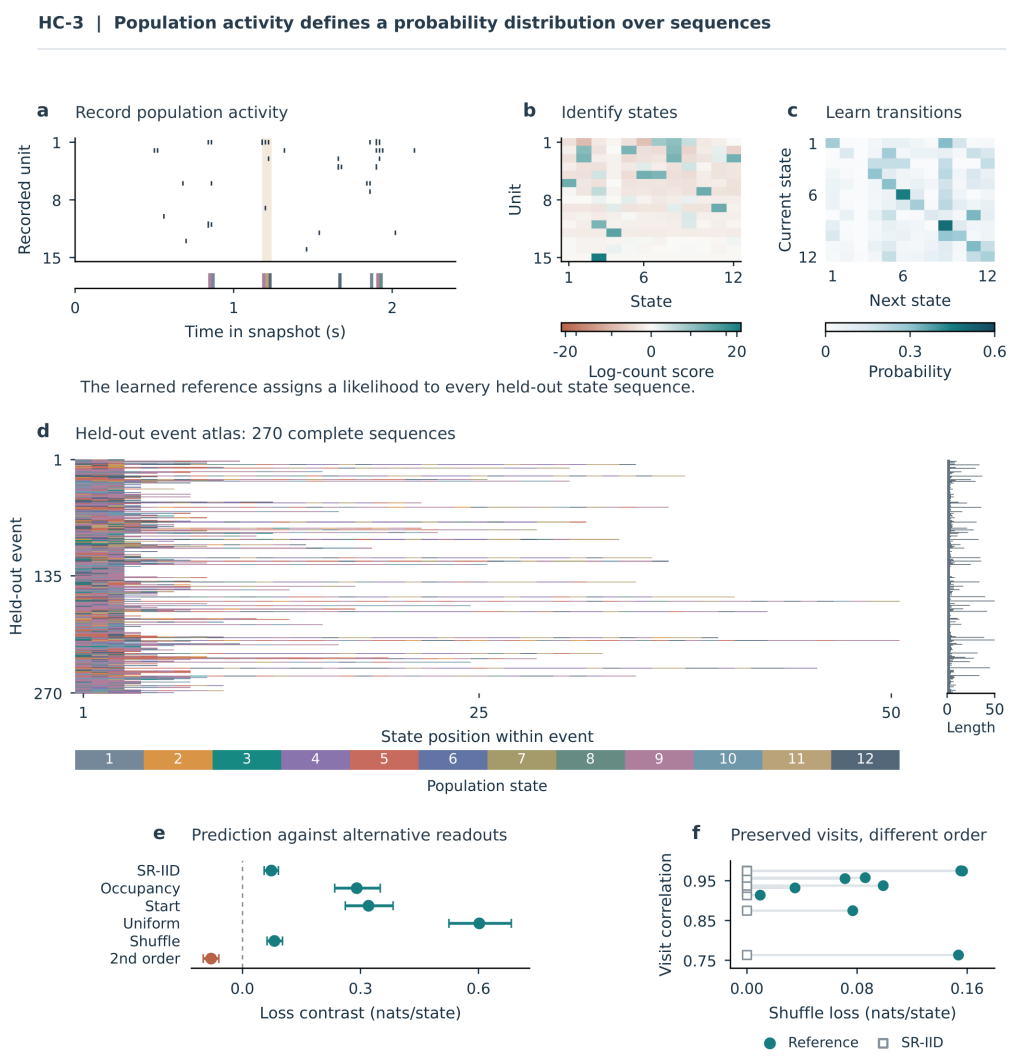

Figure 20: **From recorded firing to a reference over population sequences.** **a**, A 2.4-second snapshot of all 15 retained units in HC-3 session ec013.374. Marker area indicates spike count; shading locates the selected event and the strip shows its population states. **b**, Twelve training centroids in standardized log-count coordinates, with symmetric logarithmic colors. **c**, Learned transition probabilities from current (row) to next (column) states. **d**, All 270 held-out events in chronological order, with their lengths at right. Colors identify states; white positions follow event termination. **e**, Means and 95% session bootstrap intervals across 27 recordings. The first four rows give reference gains over visitation readouts; the fifth gives shuffle loss and the sixth gives second-order gain over first order. **f**, All nine available recording-level pairs of reference (filled circles) and SR-IID (open squares) shuffle losses, joined at the recording's visit-count correlation.

Table 8: Six HC-11 evaluations in a fixed PRE-defined 12-state alphabet. Each entry averages sessions within animal and then four animals equally. Losses and gains are in nats/state.

Reference	Test epoch	Reference loss	Gain over SR-IID	Gain over future-SR
PRE	PRE	1.852064	0.073533	0.055023
PRE	MAZE	1.693805	0.056127	0.093809
PRE	POST	1.851014	0.063928	0.054118
MAZE	PRE	2.076264	0.000559	0.028254
MAZE	MAZE	1.546673	0.055865	0.071534
MAZE	POST	2.080220	-0.000882	0.028398

analysis, these values describe correspondence within a task-defined representation. The PRE-defined comparisons in Table 8 establish transfer under a representation that was available before the task.

The crossed reference–epoch comparison. Each PRE or MAZE reference scores the same held-out events in all three epochs. Its comparison visitation readout is derived from that particular reference. Positive gains against visitation can therefore coexist with better task prediction by the MAZE-fitted reference. The shuffle comparison keeps the fitted model fixed and rearranges each event's states; it averages 49 reference-score randomizations per session and uses one paired rearrangement for SR-IID. Figure 21 retains the paired animal estimates.

D.8 State resolution and population information

The representation controls ask whether sequential information depends on a particular 12-state partition. In HC-3, the 6-state, 12-state, 28-state, and winner-cell representations produce mean gains over their respective SR-IID readouts of 0.0583, 0.0733, 0.0136, and 0.0578 nats/state. Their shuffle losses are 0.0544, 0.0812, 0.1115, and 0.0827 nats/state. Each representation retains a positive average order effect. The 28-state representation has the largest shuffle sensitivity and the smallest prediction gain. A finer partition exposes more distinctions between successive activity patterns and also creates more transition probabilities to estimate. The two effects separate the amount of detectable temporal structure from the accuracy with which a fitted reference predicts new sequences.

HC-11 provides a further distinction. PRE references in the 6-, 12-, and 28-state population alphabets yield positive mean SR-IID gains in PRE, MAZE, and POST. Winner-cell gains for the same three tests are 0.025952, -0.089933, and -0.032724 nats/state. The population representation carries cross-epoch information that is lost when each time bin is reduced to the identity of its most active cell. This is consistent with a map expressed in relations among distributed activity patterns. More states also change epoch specificity: the 28-state MAZE reference gains 0.089023 on MAZE but -0.124160 on PRE and -0.102788 on POST. The complete 24-condition comparison is shown in Figure 22. It quantifies a tradeoff between resolution and reuse of the estimated dynamics.

D.9 Removing dwell and matching the legal transitions

An original event includes repeated labels, such as $(i_0, i_0, i_1, i_2, i_2, i_T)$. Its jump sequence is (i_0, i_1, i_2, i_T). Collapsing consecutive repeats within each detector event removes dwell counts while retaining event boundaries. Collapsed events with fewer than two states are discarded. The transition kernel, initial distribution, and successor matrix are refitted to the corresponding collapsed PRE training events.

Ordinary additive smoothing places positive mass on the diagonal even when all observed training jumps change state. The support-matched analysis explicitly sets each diagonal entry to zero and renormalizes the remaining row. It then recomputes the stationary distribution and successor matrix. At every scored transition, the visitation comparator excludes the immediately preceding observed state and renormalizes its future-SR probabilities. Both models consequently assign probability over the same legal next states. This support conditioning makes the comparison readout depend on the preceding state through the legality constraint.

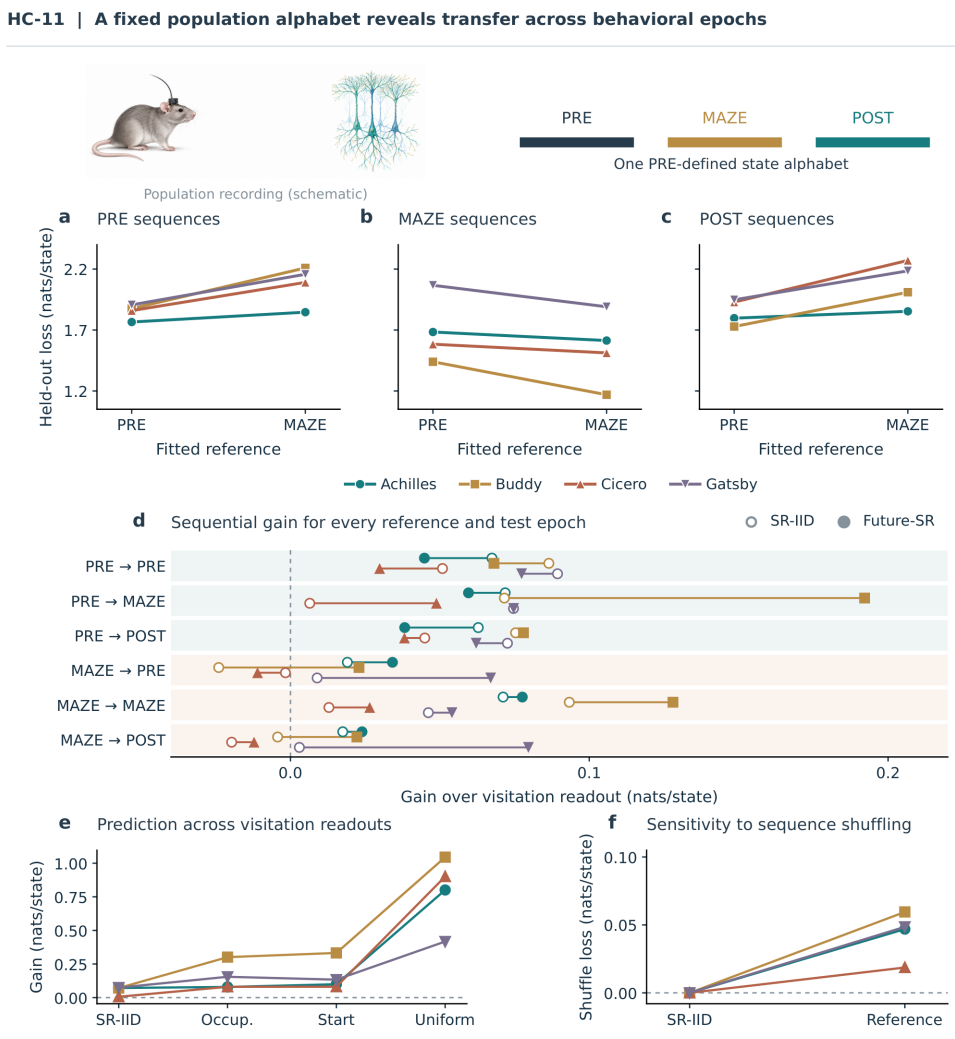

Figure 21: **Reference dynamics and visitation readouts across epochs. a–c**, Held-out losses under PRE and MAZE references for PRE, MAZE, and POST sequences. Each line joins the two reference scores for one animal; both references use the PRE-defined alphabet. Common vertical limits make the three tests comparable. **d**, All 24 animal comparisons across six reference–test routes. Segments join gains over SR-IID (open circles) and future-SR (filled animal symbols), using the reference named in each row. **e**, PRE-to-MAZE gains across four visitation comparators. **f**, Paired SR-IID and reference shuffle losses. Four animals and eight contributing sessions appear throughout, with sessions averaged within animal. The schematic depicts population recording and the PRE–MAZE–POST protocol.

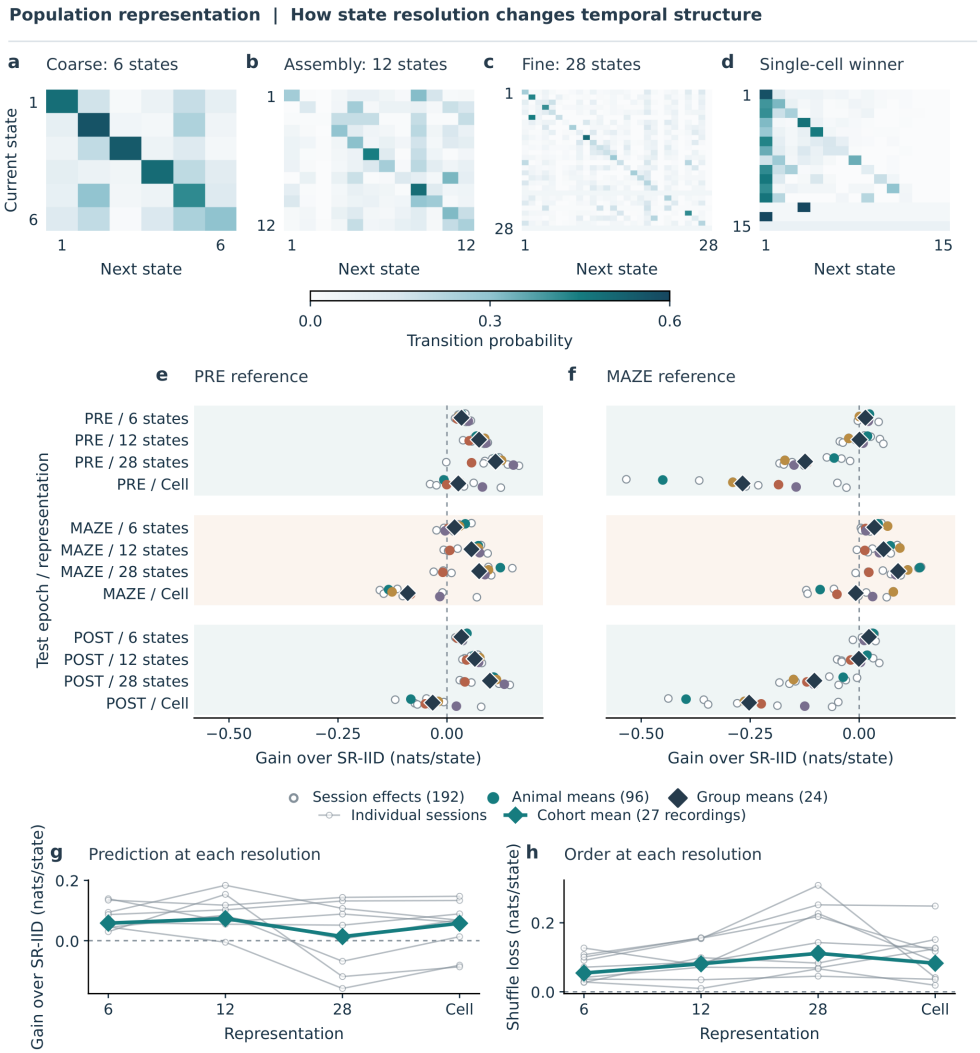

Figure 22: **Population resolution changes the structure and reuse of the reference. a–d**, Complete transition matrices from HC-3 session ec013.374 for 6, 12, and 28 population states and a 15-cell winner representation, sharing one probability scale. **e,f**, HC-11 gains for PRE and MAZE references across three test epochs and four representations. Open circles show all 192 session effects, filled circles show 96 animal means, and diamonds show 24 equal-animal means. Animal colors identify Achilles (teal), Buddy (ochre), Cicero (terracotta), and Gatsby (mauve). Shared limits include every observation. **g,h**, HC-3 prediction gain and shuffle loss. Thin profiles join each of nine available recordings across the four representations; teal diamonds give the separate 27-recording cohort means. Individual variation is shown directly. Effects are in nats/state.

The HC-11 future-SR gains for original states, ordinary collapsed states, and support-matched jump states are respectively 0.055023, 0.147361, and 0.022925 for PRE; the corresponding MAZE gains are 0.093809, 0.181977, and 0.041881. Each denominator refers to states in its own representation. The support-matched jump intervals are $[0.016324, 0.029561]$ for PRE and $[0.001813, 0.068723]$ for MAZE, with positive effects in four and three animals. DANDI task jump gain is -0.114172 nats/jump state, with interval $[-0.217011, -0.035197]$ and one positive animal out of eight. Dwell and support therefore explain part of the simpler contrasts, while the remaining predictive comparison depends on dataset and epoch.

D.10 Conditional tests of internal order

What is held fixed. Three related permutation analyses examine distinct levels of ordering. A legal block control divides a jump event into blocks of two states, fixes the first block, and uniformly permutes the labelled remaining blocks subject to legal, nonrepeating boundary transitions. A state-order control fixes the initial state, length, and exact state counts, and uniformly samples legal full orders. The endpoint control additionally fixes the final state. All score the same frozen PRE jump kernel. Each analysis uses 999 randomization draws, retains the identity in the legal null distribution, and retains events having only one legal ordering.

Uniform sampling with both endpoints fixed. For an event beginning at i_0 and ending at i_T, remove those endpoints from the multiset of state counts and call the remaining counts c. Let $C(c, i)$ be the number of legal completions when the most recently emitted state is i. Then

$$C(c, i) = \sum_{j \neq i:\, c_j > 0} C(c - e_j, j), \qquad C(0, i) = \mathbb{1}\{i \neq i_T\}. \tag{126}$$

The endpoint condition enters the boundary value. Counts are exact integers. An integer rank drawn uniformly from 0 to $C(c, i_0) - 1$ selects one unique legal sequence by successive unranking; equivalent count configurations are memoized to reduce storage. This construction samples state strings uniformly while respecting repeated state identities, fixed endpoints, and the exclusion of adjacent repeats. It needs no acceptance preference based on reference likelihood.

Why the sampling rule is uniform. At a partial sequence ending at i with remaining counts c, the legal completions split into disjoint sets according to the next state j. Their sizes are $C(c - e_j, j)$, and their sum is $C(c, i)$. Choosing j with probability $C(c - e_j, j)/C(c, i)$ and repeating this rule gives any complete legal order probability $1/C(c, i_0)$ by cancellation of successive completion counts. Integer unranking implements the same choice exactly. States with equal labels are counted through their remaining multiplicities, so exchanging two copies of the same state never creates a spurious additional string.

Conditional path probabilities. Let Ω be the legal jump strings sharing an event's length, endpoints, and state counts. Under an endpoint-weighted path law $P(x) \propto a(x_0)b(x_K)R(x)$,

$$P(x \mid \Omega) = \frac{\prod_{t=0}^{K-1} P_0(x_t, x_{t+1})}{\sum_{y \in \Omega} \prod_{t=0}^{K-1} P_0(y_t, y_{t+1})}, \qquad x \in \Omega. \tag{127}$$

All common factors cancel. The uniform null assigns each member of Ω equal probability; the fitted reference prefers some interior arrangements through its transition products. The empirical order effect asks whether observed arrangements tend to receive higher reference likelihood than those uniformly sampled alternatives. It tests this retained component of the bridge law without estimating new endpoint factors from the scored events.

Score and aggregation. For every session, observed and null transition losses are summed over all events and divided by the total number of jump transitions. The mean null loss minus the observed loss is the session effect. Events with a unique legal ordering contribute zero difference while remaining in the denominator. Session effects are averaged within animal, followed by equal animal weighting. The descriptive 95% intervals use 10,000 bootstrap samples, resampling animals and then sessions within the selected animal. Exact one-sided animal sign-flip tests enumerate all 2^A

Table 9: Internal order under a fixed PRE jump reference. Effects are mean null loss minus observed loss, in nats/jump transition. Endpoint-fixed intervals are descriptive hierarchical 95% bootstrap intervals.

Dataset	Test	Blocks	Start fixed	Both endpoints	Endpoint 95% interval
HC-11	PRE	-0.000196	-0.002323	0.001323	$[0.000358, 0.002278]$
HC-11	MAZE	0.000036	-0.000049	0.002669	$[0.000578, 0.004546]$
DANDI	PRE	0.001484	0.019221	0.010436	$[0.002096, 0.019582]$
DANDI	Task	0.003878	0.004491	0.006892	$[0.000242, 0.013438]$

assignments for A animals. These order-null analyses are exploratory and their displayed tests have no multiplicity adjustment.

With both endpoints fixed, the mean order effect is positive in all four dataset–epoch groups (Table 9). It is positive in four of four HC-11 animals in both epochs, and six of eight DANDI animals in both epochs. Session counts with positive effects are 7/8 and 5/8 for HC-11 PRE and MAZE, and 7/9 and 6/9 for DANDI PRE and task. The exact one-sided animal sign-flip values are 0.0625, 0.0625, 0.027344, and 0.058594, respectively. With four animals, 0.0625 is the smallest attainable value in this test.

Allowing the endpoint to vary changes the comparison. The initial-state-fixed effect is negative in HC-11 PRE and near zero in MAZE; legal block effects are also near zero in these epochs. With both endpoints fixed, the reference gives a small average preference to the observed interior arrangements. Each conditioning rule changes the alternatives against which the observation is compared. In DANDI, the endpoint-conditioned order effect is positive while the gain against the matched visitation readout is negative. The first result concerns relative probabilities within an event's constrained set of paths; the second concerns prediction across the broader event distribution. Their joint occurrence shows why the conditional test is useful: it asks specifically about the transition structure retained when a bridge changes endpoint probabilities.

An observed sequence and its legal alternatives. The illustrative held-out HC-3 event is selected by computing the median jump length across all 270 test events, then choosing the earliest event nearest that length among events with at least five jump states and three legal endpoint-preserving orders. This gives the fourth chronological event: original labels $(1, 9, 11, 1, 1, 1, 1, 9)$ and collapsed labels $(1, 9, 11, 1, 9)$. Exactly three legal orders share endpoints 1 and 9 and one interior occurrence each of states 1, 9, and 11. All three appear in Figure 23. Its selection depends on length and the number of legal alternatives. The cohort comparisons use every eligible event with the corresponding frozen PRE reference.

Joint evaluation within animals. Task prediction against future-SR and the endpoint-fixed order effect are measured on the same held-out jump events under the same PRE reference. The paired analysis retains animal identity and averages sessions within animals before equal animal averaging. Its 95% intervals are marginal intervals for the two statistics, obtained by 10,000 hierarchical bootstrap draws. The original-state task comparison uses animal-and-session resampling in HC-11 and animal resampling in DANDI. Each conditional-order comparison uses 999 uniform legal-order draws. These paired measurements establish how broad prediction accuracy and preference within constrained candidate sets vary together.

D.11 Running and stationary task activity

The task-state comparison asks whether the reference relationships occur during locomotion and during low-speed task periods. The same PRE units, state centroids, threshold, and kernel are applied to each condition. Finite-speed task bins are partitioned into running, at least 0.03 m/s, and stationary, below 0.03 m/s. Each support is split chronologically at 65%, and its final 35% is scored. Events remain within a single condition. Stationary sequences with fewer than 20 test events are retained where the score remains estimable, preserving paired animal coverage while exposing their greater uncertainty. These task pauses may include reward consumption, head movements, and other activity; they are defined by speed rather than sleep state.

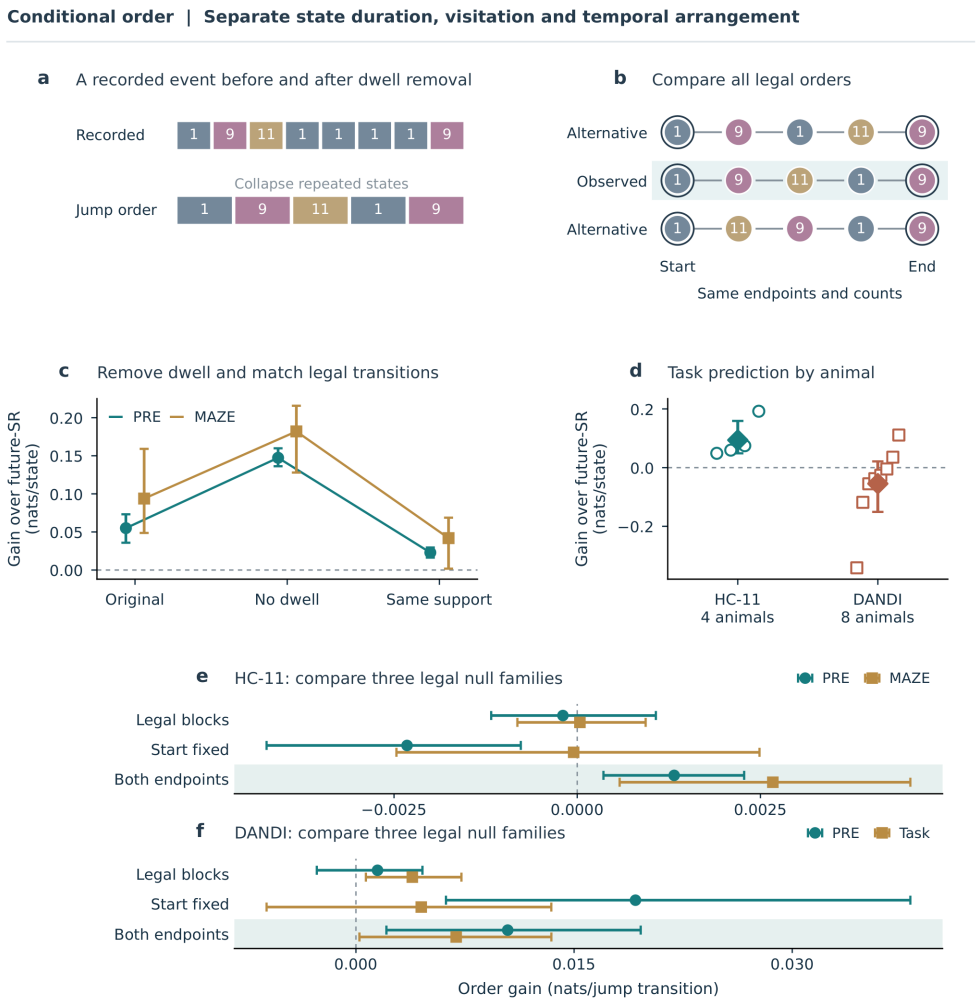

Figure 23: **Sequence constraints identify distinct contributions to temporal order.** **a**, The eight-bin labels of a measured HC-3 event and the five-state jump sequence obtained by collapsing consecutive repetitions. **b**, All three legal orders of that jump sequence with its endpoints and state counts fixed. Shading marks the observation; outlines mark the endpoints. **c**, HC-11 gains over future-SR before collapse, after collapse, and after matching legal jump support. **d**, Original-state task gains for all four HC-11 and eight DANDI animals. Open symbols show animal means; diamonds give group means. Descriptive 95% intervals resample animals and sessions in HC-11, and animals in DANDI. **e,f**, HC-11 and DANDI order gains for legal-block, initial-state-fixed, and endpoint-fixed nulls. Panels **c,e,f** show equal-animal means with descriptive hierarchical 95% intervals. Teal circles denote PRE; ochre squares denote the later task epoch.

HC-11 contains 7,344 running and 1,226 stationary jump events, with 27,592 and 2,660 transitions. Its endpoint-fixed order effects are 0.002669 and 0.003560 nats/jump, respectively. The stationary interval is $[-0.001762, 0.009669]$, and the paired stationary-minus-running effect is 0.000890, with exact one-sided sign-flip $p = 0.375$. Its matched-readout gains are 0.041881 and 0.024259 nats/jump state. The latter has interval $[-0.015356, 0.063029]$.

DANDI contains 3,222 running and 569 stationary jump events, with 16,937 and 1,582 transitions. Its order effects are 0.006892 and 0.003845 nats/jump; the stationary interval is $[-0.013483, 0.016297]$. Its paired order difference is -0.003047, with one-sided $p = 0.59375$. The matched-readout gains remain negative in both conditions: -0.114172 during running and -0.104968 during stationary activity. The stationary interval is $[-0.207129, -0.017693]$.

Figure 24 displays every animal pair and the intervals. Stationary events provide fewer transitions, and their broader intervals include effects on either side of zero for internal order. The paired estimates place the running-to-stationary difference within intervals that span zero. The negative DANDI readout comparison extends to stationary task activity.

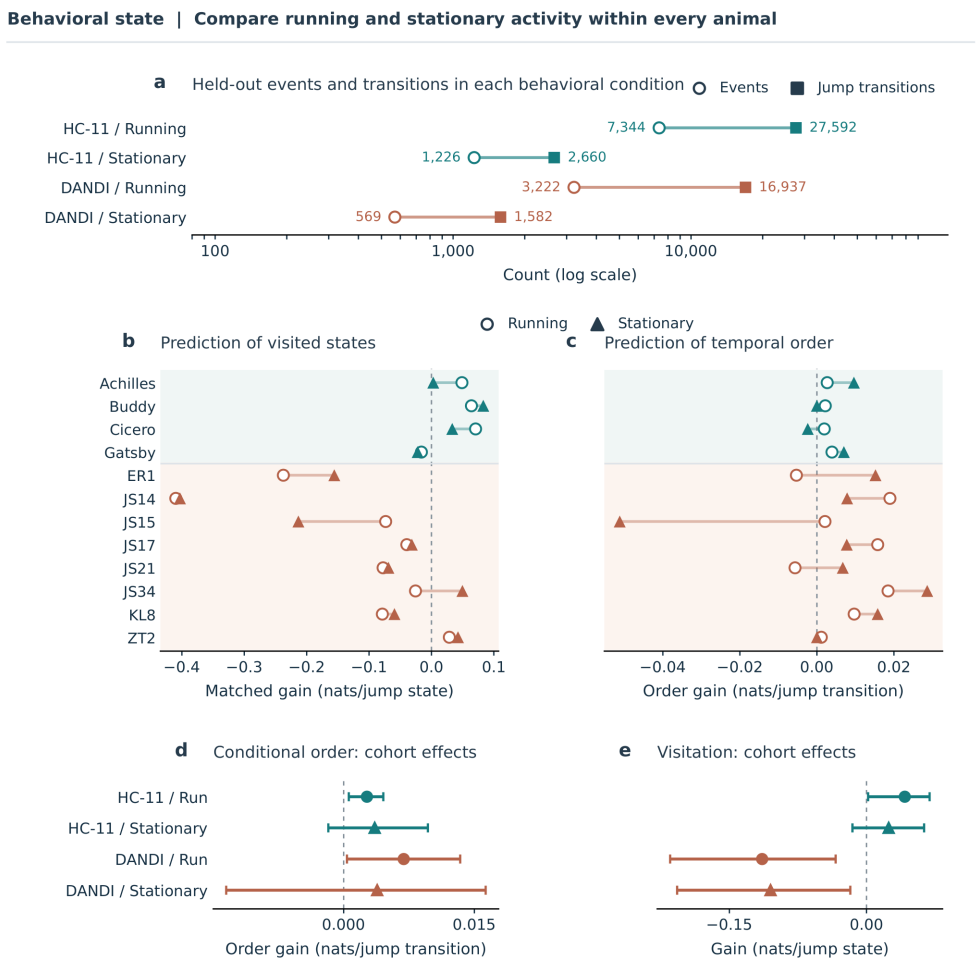

Figure 24: **Behavioral condition and reference structure. a**, Complete held-out event counts (open circles) and jump-transition counts (filled squares), joined within condition on a logarithmic axis. **b,c**, Matched visitation gains and endpoint-fixed order gains for every animal. Lines join running (open circles) and stationary activity (filled triangles); animal labels are shared across panels. Teal identifies HC-11 and terracotta identifies DANDI. **d,e**, Equal-animal means and descriptive hierarchical 95% intervals from 10,000 bootstrap draws. All four HC-11 and eight DANDI animals contribute. Each condition uses its final 35% temporal support; running requires speed at least 3 cm/s.

D.12 REPRODUCIBLE OUTPUTS

The package contains analysis programs, session and animal results, HC-3 state/event arrays, and figure source data. Dataset manifests identify source recordings and available inputs. Separate commands reconstruct the figures from stored results and estimate reference laws from recordings.

E EXPERIMENT 3: BEHAVIORAL BOUNDARIES AND COMPLETE HIPPOCAMPAL EVENTS

Hippocampal activity can represent a journey while the animal remains still. The theory connects such an event to two kinds of information: a learned map of possible transitions and the circumstances that favor a particular episode. Here, movement supplies place fields and transition probabilities. Nearby behavior supplies cues about beginnings and endings. After a traversal from A to B, an event at B can receive B as its initial cue and A as its terminal cue. Before a traversal from B to C, the analysis conditions on B and the subsequently observed destination C. The question is how these shared behavioral conditions change the probability assigned to recorded events.

The model specifies a distribution over complete paths, including the decision to finish. We therefore evaluate both the represented sequence and its extent. The complete decoded-path score assigns probability to one decoded sequence together with its number of bins. The conditional spike score holds that number fixed and sums over latent sequences to evaluate the observed spike counts. The first connects the model to the organization of an episode; the second asks how well it accounts for the neural observations when duration is already known. Table 10 states the quantities entering each comparison.

The analysis begins with a common movement-derived map and held-out population events. Computational deletions then locate the contributions of continuation, source normalization, and stopping. Varying spatial resolution distinguishes a more detailed representation from a more effective use of boundary information. Independent ripple detections finally connect the event comparisons to a physiological subset. Additional behavioral and sleep cohorts extend these predictive targets to well-to-well navigation, reward approach, and complete ripple events.

Table 10: Prediction targets in Experiment 3. A bin has nominal duration 20 ms in the main HC-3/HC-6 analysis. Physical event duration and retained bin count are recorded separately.

Score	Observation being evaluated	Information supplied to the model
Complete decoded path	A decoded state sequence and its bin count, jointly.	Behavioral endpoint cues; movement-derived reference.
Event length	The number of bins until completion, with paths summed out.	Same cues, reference, and stopping dynamics.
Path content at fixed length	The decoded sequence conditional on its bin count.	Same inputs plus the observed length.
Conditional spikes	The spike-count vectors, with all latent paths summed out.	Same inputs, place fields, and observed length.
Joint spikes and length in DANDI 001695	Spike-count vectors and termination, with paths summed out.	Movement-derived model and fixed geometric boundary evidence.

E.1 FROM MOVEMENT AND SPIKES TO REPRESENTED PATHS

The main cohort contains six HC3 sessions from three rats, ec013, ec016, and gor01, and one HC6 session from rat Bon (Mizuseki et al., 2014; 2013; Karlsson et al., 2015). HC3 provides approximately linear trajectories; HC6 provides a branching W-maze. The analysis includes 467 HC3 and 54 HC6 population events. Table 11 gives the biological sampling structure. CA1 and CA3 pyramidal cells enter the main analysis according to their available metadata and movement-related firing. The resulting predictions concern the recorded hippocampal ensembles, with their CA1 and CA3 composition specified below.

Movement and state space. The spatial representation determines which locations the model can distinguish. Movement samples exceed the session's 60th speed percentile. Relative immobility samples fall below its 32.5th percentile. These percentile criteria accommodate differing coordinate scales. The absolute immobility thresholds for the seven sessions in Table 11 are 8.2540, 10.3132, 10.0322, 5.7834, 1.6126, 8.0249, and 1.5246 in their recording coordinate units per second. HC3

Table 11: Main hippocampal cohort. Unit counts give CA1/CA3 place cells. Event classes are assigned from nearby behavior. Training and test counts sum to the complete event count in each session.

Dataset	Session	Units	Forward	Reverse	Free	Train/test
HC3	ec013.387	19/0	0	0	31	21/10
HC3	ec013.669	20/0	8	10	18	24/12
HC3	ec013.670	15/0	5	7	45	38/19
HC3	ec016.269	31/0	3	13	178	130/64
HC3	gor01 16:53:46	7/12	11	12	23	30/16
HC3	ec013.932	0/13	35	33	35	68/35
HC6	Bon day05 epoch02	11/10	10	17	27	36/18

uses 24 spatial states ordered along the principal trajectory. HC6 uses 27 states embedded in the two-dimensional maze and graph distances between state centers. A state is a spatial bin represented by the activity of the recorded ensemble.

Place fields and reference transitions. Movement observations supply both the neural templates for location and the transition probabilities between locations. A neuron is retained with at least 30 movement spikes and a peak smoothed firing rate of at least 0.5 Hz. Occupancy and spike counts are smoothed with a Gaussian of standard deviation one state, with 0.05 s occupancy regularization and a 0.01 spike pseudocount per state. Rates have a 0.02 Hz floor. HC3 behavioral transitions count successive movement samples whose state indices differ by at most three. A 0.25 pseudocount is added to self-transitions and immediate neighbors. Destination counts are smoothed with a Gaussian of standard deviation 0.7 states and normalized by row. HC6 adds 0.25 to self-transitions and 0.10 to each state's four nearest spatial neighbors before normalization. Its shortest-path distances are divided by the median finite nonzero graph distance. The stationary distribution of the resulting matrix P_0 provides π.

Population-event detection. Events are defined from population activity during relative immobility, before examining their relation to LFP ripples. Spikes are counted in 10 ms bins and population rate is smoothed with a Gaussian of standard deviation two bins. An event begins from a population-rate excursion above the 99.2nd percentile of the immobility background. Its boundaries extend through adjacent immobility bins above 0.45 times that threshold. Candidate durations are 40–350 ms; segments separated by less than 25 ms are merged and the duration criterion is reapplied. If fewer than 100 immobility bins are available, detection uses the 40th speed percentile. Independent LFP criteria enter the subsequent ripple analysis in Section E.6.

Neural observation model and decoding. Within an event, spikes are binned at 20 ms. Conditional on state i, the count of neuron c has a Poisson distribution with mean $\lambda_{ci}\Delta t$, where λ_{ci} is its movement place field. Writing the population observation in bin t as Y_t, the emission likelihood is

$$E_t(i) = \prod_c \frac{(\lambda_{ci}\Delta t)^{Y_{tc}}}{Y_{tc}!} \exp(-\lambda_{ci}\Delta t). \tag{128}$$

All state-independent Poisson terms are retained. The decoder forms binwise state evidence proportional to $\pi_i E_t(i)$. HC3 uses a uniform transition matrix for decoded path extraction, giving the binwise maximum of this evidence. HC6 uses P_0 in the maximum-probability sequence calculation to respect its graph representation. The spike-evidence panels of Figure 25 display normalized binwise state evidence. This difference between decoding rules enters the interpretation of intermediate-transition statistics below.

Population evidence and conditional route inference. The displayed neural evidence is $U_t(i) = \pi_i E_t(i) / \sum_j \pi_j E_t(j)$. A second image shows the latent-state posterior after combining all event spikes with the fitted reference transitions and both boundary functions. Its forward and backward

messages satisfy

$$F_0(i) = \pi_i e_0(i) E_0(i), \qquad F_t(j) = E_t(j) \sum_i F_{t-1}(i) P_{0,ij}, \tag{129}$$

$$B_{L-1}(i) = e_T(i), \qquad B_t(i) = \sum_j P_{0,ij} E_{t+1}(j) B_{t+1}(j). \tag{130}$$

The posterior is $Q_t(i) = F_t(i)B_t(i)/\sum_j F_t(j)B_t(j)$. Each recursion is rescaled after every step for numerical stability. The common fixed-length factor γ^{L-1} cancels. The two messages give the posterior a direct interpretation. The forward message measures how readily a path can arrive at a state while explaining the preceding spikes. The backward message measures how readily it can leave that state while explaining the remaining spikes and terminal evidence. Their product favors states compatible with the event on both sides. The decoded sequence is overlaid separately. Its differences from posterior concentration display the effect of including temporal and boundary information. Setting every E_t to one gives the corresponding path prior, included in the source tables. The difference between prior and posterior then identifies what the spikes contribute to the inference.

Examples are selected from three groups in the main test set: HC3 Forward, HC3 Reverse, and HC6 events with explicit boundaries. Each group contributes its event with the highest total spike count; session name and event identifier resolve ties. Selection depends on observed spike count independently of model score. The groups contain 22, 24, and nine events. The selected observations are ec013.670 event 13 (310 ms, 16 bins, 73 spikes, 14 active cells), ec013.669 event 14 (200 ms, ten bins, 54 spikes, 15 active cells), and Bon event 4 (160 ms, nine bins, 25 spikes, eight active cells). Figure 25 shows all three examples, and the main figure uses the HC3 Forward event. HC6 state indices identify graph nodes; their numerical separation has no metric meaning. The evidence, path prior, and posterior use each event's fitted boundary widths and the same reference matrix.

An event requires at least three decoding bins and seven spikes. Ensembles of fewer than 20 neurons require two active neurons and mean normalized posterior entropy at most 0.96. Larger ensembles require $\max\{3, \min[5, \lceil 0.18N \rceil]\}$ active neurons and entropy at most 0.92. The accepted HC3 paths can contain spatial jumps. Event duration in milliseconds is the difference between detected onset and offset. The likelihood uses the number of retained decoding bins, L, as its discrete duration variable. Bin edges are constructed at the nominal 20 ms spacing, and the last interval can be shorter; these conventions can yield, for example, five retained bins in an 80 ms detected event. The emission model uses the nominal 20 ms exposure for every bin. The millisecond duration and L are therefore reported separately, and all models share the same binning and exposure conventions.

Behavioral boundary evidence. Nearby traversals provide a reproducible correspondence between the animal's behavior and the conditions placed on a represented path. The closest traversal ending before, or beginning after, an event within five seconds determines its behavioral class. For a preceding traversal from A to B, the current location supplies the initial cue and A supplies the terminal cue, giving the Reverse class. For an upcoming traversal, current location supplies the initial cue and its destination supplies the terminal cue, giving the Forward class. Events without an eligible traversal receive the Free class and uniform evidence at both boundaries. Forward and Reverse label the ordering of the behavioral cues. The decoded neural sequence has its own represented direction. Terminal evidence for Forward events is conditioned on subsequently observed behavior.

Soft evidence around a cue state c is a Gaussian function of state distance with a 0.04 positive floor. Its width expresses uncertainty in the correspondence between behavior and represented location. The positive floor allows represented paths to depart from the cue centers. HC3 widths are measured in state-index distance, and HC6 widths in normalized graph distance. Reverse and Forward events use the same model in their own event-time order.



Figure 25: **From recorded activity and boundary cues to a represented episode.** Rows show HC3 Forward, HC3 Reverse, and HC6 events with explicit boundaries, selected by highest spike count within each held-out group. Detected milliseconds and retained bin counts are stated separately. Columns show normalized Poisson population evidence; initial and terminal evidence with normalized continuation field; the state prior conditioned on boundaries; and the posterior using all event spikes. Probability images share a square-root scale from zero to one. Terracotta marks identify the separately decoded states. HC3 states follow spatial order; HC6 states identify graph nodes. Header artwork illustrates the task concept; the panels contain measured or model-derived quantities.

E.2 What the path distribution predicts

For a path $x_{0:L-1}$ of L bins, the probability used in the main path comparison is

$$p(x_{0:L-1}, L \mid e_0, e_T) = \frac{\pi_{x_0} e_0(x_0)\, \gamma^{L-1} \prod_{t=0}^{L-2} P_0(x_t, x_{t+1})\, e_T(x_{L-1})}{\sum_i \pi_i e_0(i) \phi_i}, \qquad \phi = (I - \gamma P_0)^{-1} e_T. \tag{131}$$

Here $P_0 = P$, $e_0 = a$, $e_T = b$, and $\phi = h$ in the main-text notation, while $L = K + 1$ counts observed bins. The denominator sums the weights of all finite paths, including all lengths. It couples the choice of initial state to the paths and terminations available from that state. Complete-path prediction therefore evaluates represented sequence and event duration together. The corresponding initial distribution is proportional to $\pi e_0 \phi$.

For numerical evaluation of decoded paths, the implementation uses $\log(p + 10^{-12})$ for transition, initial, boundary, and normalization factors; the resolvent itself uses the fitted transition matrix. The same convention applies to every comparator. This finite-score convention permits evaluating independently decoded HC3 paths containing jumps with zero estimated transition weight.

The Reference model sets both evidence functions to one. Start and End retain the indicated evidence function. Swapped exchanges the cue centers while retaining the initial and terminal width assignments. The remaining controls retain start evidence and replace P_0 by a row-normalized biased kernel, with uniform terminal evidence in their complete path law. Table 12 states their factors explicitly.

The Pairwise affinity A_{ij} is the cosine similarity between population rate templates after a $\log(1 + r)$ transform and centering across neurons. Every biased kernel retains the support of P_0. At $\beta = 0$ its path law reduces to Start. Shortest and Reward give the same normalized kernel because their unnormalized factors differ by a quantity depending only on the source state. HC6 has no independently defined linear direction; its Direction entry uses $\beta = 0$ and equals Start. These relationships explain coincident numerical results.

SR-value receives the same event-specific terminal cue as Bridge through its resolvent value. At $\beta = 1$, its transition conditional on continuation equals the Bridge continuation kernel. Bridge also weights initiation by the continuation field and couples each continuation step to state-dependent stopping. The comparison with SR-value isolates the contribution of coordinating initiation, continuation, and termination around the same prediction of available future paths. Its scientific target is the probability of a complete represented event, including the decision to stop.

Conditional spike prediction. For the second predictive target, the observed event length is fixed and the hidden path is summed out:

$$p(Y_{0:L-1} \mid L, e_0, e_T) = \frac{\sum_{x_{0:L-1}} \pi_{x_0} e_0(x_0) e_T(x_{L-1}) \prod_{t=0}^{L-2} P_0(x_t, x_{t+1}) \prod_{t=0}^{L-1} E_t(x_t)}{\sum_{x_{0:L-1}} \pi_{x_0} e_0(x_0) e_T(x_{L-1}) \prod_{t=0}^{L-2} P_0(x_t, x_{t+1})}. \tag{132}$$

A forward recursion evaluates numerator and denominator in log space. This score accounts for uncertainty about represented state throughout the event. The factor γ^{L-1} cancels for Bridge at fixed length. Complete decoded-path likelihood measures probability assigned to a selected representation and its duration. Conditional spike likelihood measures probability assigned to the observations after summing over representations. Conditioning on length removes the common Bridge continuation factor from the second quantity.

E.3 Fitting and prediction on test events

Three sources of data have different roles. Movement observations establish the spatial representation, firing templates, and transition matrix. Training population events choose boundary widths and control-model biases. Held-out population events provide the paired prediction scores. The survival parameter requires a separate distinction because it affects the distribution over event lengths but cancels from the bridge likelihood once length is fixed.

Events are sorted by time within each behavioral class. Class index 1 modulo 3 enters the test set, and the remaining events enter training. There are 311/156 HC3 and 36/18 HC6 training/test events.

Table 12: Transition factors for the biased controls. Each kernel is the row normalization of $P_{0,ij}$ times the given factor. d_i denotes distance from state i to the event's target; s is the behavioral direction sign.

Control	Factor	Information used
Shortest	$\exp[\beta(d_i - d_j)]$	Target distance
Direction	$\exp[\beta s(j - i)]$	Linear direction
Reward	$\exp(-\beta d_j)$	Target proximity
SR-value	$\exp[\beta(v_j - v_i)]$	$v = \log[(I - \gamma P_0)^{-1} e_T]$
Pairwise	$\exp(\beta A_{ij})$	Population-template similarity

Table 13: Parameter calibration and evaluation in the main event analyses. Repeated analyses of an event use its original train/test assignment.

Analysis	Training choice	Scored observations
Main event comparison	Widths minimize class-balanced conditional-spike loss; controls select their bias. Bridge survival takes the first tied grid value, 0.58.	156 HC-3 and 18 HC-6 events; all models share movement fields and boundary calibration.
Operation deletions	Survival matches mean training-event bin count; other fitted parameters retained.	Same 174 test events, with path and length contributions paired.
Main resolution scan	Three folds; boundary widths refitted in each training fold and representation; survival 0.82.	Each of 460 HC-3 and 54 HC-6 events evaluated once on held-out spikes at fixed length.
Ripple restrictions	Main event fits retained.	Physiologically selected subsets of the 105 test events with matched LFP coverage.

Movement observations estimate place fields and reference transitions; training population events select the boundary widths and model biases. The training objective averages negative conditional spike likelihood per bin equally over the available behavioral classes.

The search grids are $\gamma \in \{0.58, 0.66, 0.74, 0.82, 0.88, 0.93\}$, $\sigma_0, \sigma_T \in \{1.4, 2.2, 3.4, 5.2, 7.0\}$, and $\beta \in \{0, 0.10, 0.25, 0.45, 0.75, 1.10\}$. Main comparisons use the Bridge-selected boundary widths and γ; biased controls select their own β on training events under this common cue calibration. The chosen width pairs in session order are $(1.4, 1.4)$, $(7.0, 3.4)$, $(7.0, 3.4)$, $(1.4, 1.4)$, $(1.4, 2.2)$, $(5.2, 1.4)$, and $(1.4, 2.2)$. All sessions use $\gamma = 0.58$. The Bridge conditional likelihood is identical across γ, so 0.58 is a grid choice. A continuation timescale is estimated from event lengths in the duration analysis. Section E.4 uses a separate duration calibration.

For comparator m, the session gain is the sum of paired event log-likelihood differences divided by the sum of their bin counts:

$$G_s(m) = \frac{\sum_{e \in s} [\log p_{\mathrm{Bridge}}(e) - \log p_m(e)]}{\sum_{e \in s} L_e}. \tag{133}$$

HC3 averages these gains equally across sessions within each animal and equally across animals. Its intervals use 20,000 hierarchical resamples of animals, sessions, and the stored paired within-session event-bootstrap draws. HC6 intervals resample events within its one session. Model parameters remain fixed during resampling. These intervals describe uncertainty conditional on the fitted models and the sampled animals. Sessions from ec013 share an animal and are nested accordingly.

Main result and endpoint specificity. Bridge improves the HC3 held-out path score over SR-value by 0.03193 nats/bin, with a 95% interval of $[0.00187, 0.08421]$. Gains over Start, End, and Pairwise are 0.03068 $[0.00050, 0.09725]$, 0.05293 $[0.00144, 0.12138]$, and 0.06089 $[0.02068, 0.11251]$. Against Reference, the gain is 0.03357 $[-0.01678, 0.12312]$. All nine named controls have positive

aggregate point estimates in both datasets; the complete comparisons are shown in the main figure and the accompanying source tables.

HC6 provides an application to branching space: its Bridge–SR-value gain is 0.00195 $[-0.00231, 0.00632]$ nats/bin and its Bridge–Reference gain is 0.00899 $[-0.00054, 0.01699]$. These intervals reflect variation within the recorded W-maze session. In HC3 events with explicit boundaries, the Bridge–SR-value gain rises to 0.13665 $[0.05159, 0.23708]$ nats/bin. This subset interval resamples animals and sessions, with their estimated event-level gains fixed. The 46 HC3 test events with explicit boundaries are the observations for which the model makes its strongest context-dependent prediction. Uniform boundaries in the remaining 110 events recover the unconditioned reference law; the biased controls also recover this law when their bias is constant across states. The full sample and boundary subset therefore answer different questions about the frequency and strength of contextual modulation. Their intervals use the resampling schemes specified above.

Forward and Reverse classes both show positive descriptive gains. Bridge–Start gains are 0.15618 and 0.13835 nats/bin in HC3, and 0.00351 and 0.00275 in HC6. Against Reference, the Forward and Reverse gains are 0.03265 and 0.14049 nats/bin in HC3, and 0.01271 and 0.01832 in HC6. Figure 26 places these aggregate effects alongside event composition and individual-session variation. The gain occurs in both temporal arrangements of behavior around an event. Together with the explicit-boundary comparison, this pattern links the event probability to the information supplied about its beginning and end.

The 174 held-out events comprise 119 Free, 25 Forward, and 30 Reverse observations. The 119 Free events have identical Bridge and SR-value scores because their uniform boundaries recover the reference law. Event-level differences are paired before averaging, and each event's complete-path and conditional-spike gains use its same observed bin count. The HC3 boundary subset contains 46 events; the corresponding HC6 subset contains nine. Figure 26 gives the complete event and recording distributions underlying the aggregate comparisons.

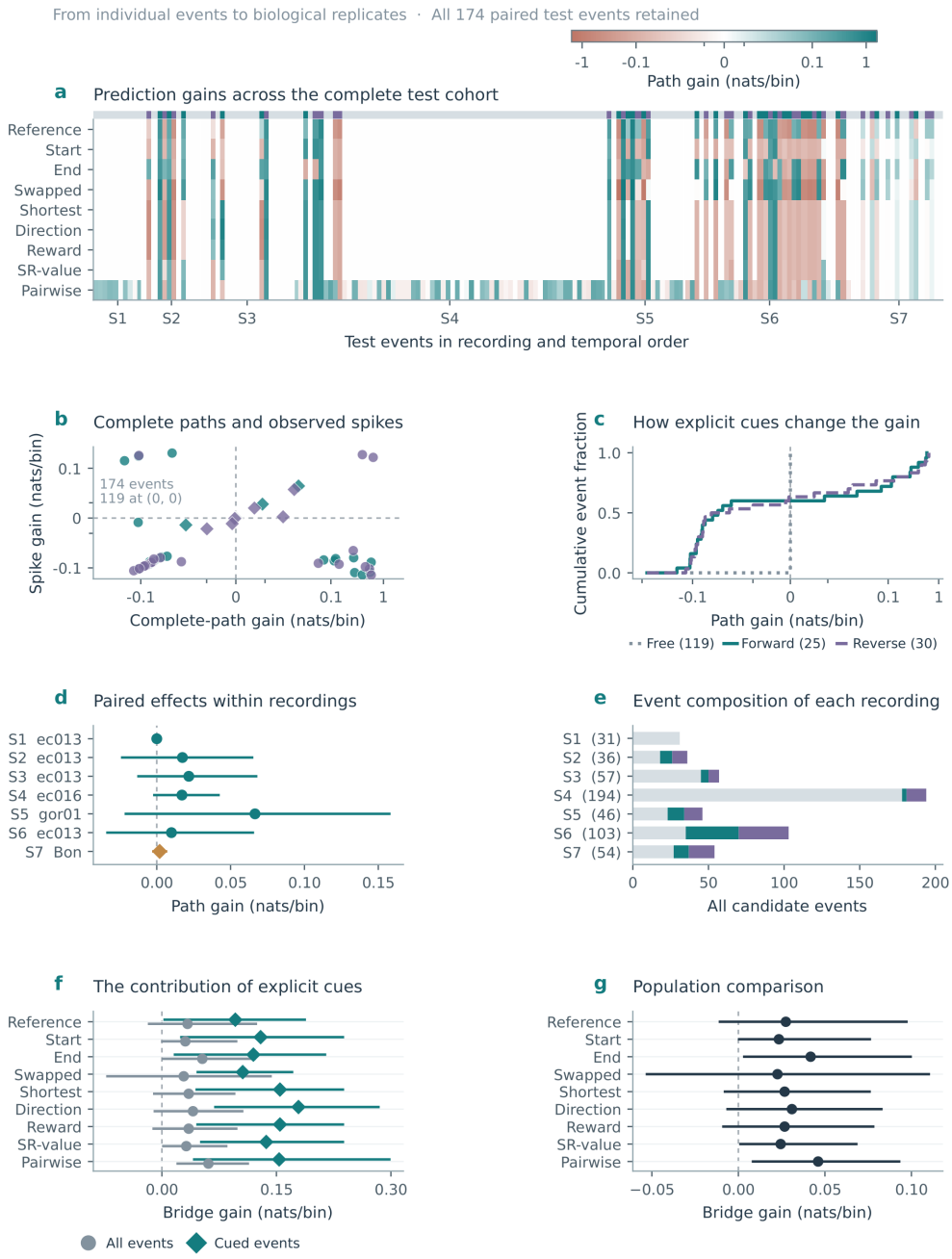

Figure 26: **Contextual path gains across events, recordings, and animals.** (a) The 174×9 array of paired path gains, ordered by recording and event time. The upper strip marks behavioral class; signed color uses symmetric logarithmic normalization. (b) Complete-path and conditional-spike gains for every test event, including 119 Free events at zero. (c) Empirical cumulative path-gain distributions. (d) Recording means and 95% intervals against SR-value. (e) Candidate-event composition by recording. (f) HC3 estimates for all events and events with explicit boundaries. (g) Animal-equal estimates across four animals and seven recordings. Intervals in f,g preserve the reported biological sampling hierarchy. Gains use nats per observed bin.

E.4 How continuation and stopping shape an event

A complete event has three coupled features: where it begins, how it progresses, and when it ends. The continuation field influences all three through the available future paths. Source normalization controls the strength of continuation drive relative to termination, and the stopping rule converts their competition into an event duration. To examine their computational contributions at the observed duration scale, each session uses $\gamma = 1 - 1/\overline{L}_{\text{train}}$, where $\overline{L}_{\text{train}}$ is its mean training-event length. The seven means are 7.28571, 6.58333, 7.34211, 5.51538, 6.06667, 5.14706, and 6.00000 bins. The resulting continuation factors range from approximately 0.8053 to 0.8627.

Write the complete model's initial distribution as $s_i \propto \pi_i e_0(i)\phi_i$, its transition conditional on continuation as $C_{ij} = P_{0,ij}\phi_j/(P_0\phi)_i$, and its stopping probability as $r_i = e_T(i)/[e_T(i) + \gamma(P_0\phi)_i]$. Three computational deletions retain the fitted parameters. Removing continuation-field input sets its drive to one, giving $s \propto \pi e_0$, $C = P_0$, and $r_i = e_T(i)/[e_T(i) + \gamma]$. Removing state-dependent stopping retains s, C and sets $r_i = 1 - \gamma$. Removing source normalization retains s, C and uses $r_i = e_T(i)/[e_T(i) + \gamma(d_i/\bar{d})(P_0\phi)_i]$, where d_i is the row sum of the unnormalized behavioral weight matrix and $\bar{d}$ its mean. These row sums are estimated behavioral weights. A shared 10^{-12} numerical smoothing term supports likelihood evaluation.

For any such model the complete path probability is $s_{x_0} \prod_{t=0}^{L-2}[(1 - r_{x_t})C_{x_t x_{t+1}}]\, r_{x_{L-1}}$. Letting $A = \operatorname{diag}(1 - r)C$, its length probability is $s^{\mathsf{T}} A^{L-1} r$. Subtracting its log from complete-path log probability isolates path content conditional on length. This identity separates an improved duration distribution from an improved distribution of represented paths at a given duration. All events are paired, and 5,000 hierarchical bootstrap draws produce the intervals. Deletions are evaluated without individual retraining.

Removing continuation-field input lowers complete-path scores by 0.25717 $[0.20049, 0.30146]$ nats/bin in HC3 and 0.25721 $[0.22585, 0.28789]$ in HC6. The length terms are 0.23449 and 0.25199, accounting for 91.2% and 98.0% of those differences. At the training-calibrated duration scale, continuation-field input contributes chiefly through the distribution of event length. A future-compatible continuation signal changes how long the recurrent process is likely to persist; the paired decomposition quantifies that consequence separately from its effect on represented content. Removing source normalization lowers scores by 0.04979 $[0.00606, 0.11094]$ and 0.01127 $[0.00473, 0.01853]$. Removing state-dependent stopping gives small differences, 0.00138 $[-0.00058, 0.00378]$ and 0.00006 $[-0.00001, 0.00019]$.

The explicit-boundary subset contains 46 HC3 and nine HC6 test events. Its continuation-field gains remain positive, 0.14768 $[0.08746, 0.25242]$ and 0.26521 $[0.23601, 0.29318]$. Source-normalization gains are 0.07228 $[0.00402, 0.19428]$ and 0.00549 $[-0.00160, 0.01444]$. Figure 27 shows the duration contribution alongside the other terms. The duration variable is the number of bins in the detected population event. The deletion results identify which features of that event distribution depend on each circuit term. They supply a computational link from the proposed neural operations to measurable initiation, progression, and termination statistics.

For every test event, the path-content contribution equals complete-path gain minus length gain. Both contributions are divided by that event's observed bin count. The same 174 paired observations enter each deletion comparison, with the explicit-boundary subset evaluated separately. This decomposition assigns a change in score to event extent or to represented content at fixed extent without changing the underlying observations (Figure 27).

Internal sequence diagnostics. Additional analyses examine how the complete path account distributes probability within an event. A logistic regression predicts terminal versus continuing bins from $\log e_T(x_t) - \log \phi(x_t)$, controlling for current bin index, target distance, occupancy, and binwise decoding entropy. Predictors are standardized, and a session requires at least 40 eligible bins and five terminal bins. These fits use eligible non-Free events from the full candidate pool. Their median controlled coefficient is 0.054 across six estimable sessions, with four positive coefficients. In the higher-span subset the median is 0.038 across five sessions, with three positive coefficients.

For progression, events are grouped by behavioral class and boundary side. Groups require at least five events; a step is evaluated only when at least five events reach it, up to eight steps. Empirical state histograms are compared with normalized forward propagation, backward compatibility, and their

product. The median cross-entropy gain of the product over forward propagation is 0.138 nats/step, positive in four of six sessions. Its gain over backward compatibility is 0.722 nats/step, positive in all six. The product combines how probability arrives from the initial boundary with how it remains compatible with the terminal boundary. Its progression gain gives a within-event counterpart to the complete-path comparison.

A separate test scores decoded internal transitions under P_0 against uniform and distance-based alternatives. HC3 session medians are -8.59 and -3.92 nats/transition, with six negative sessions for each contrast; HC6 values are 3.066 and 0.061. The contrast depends on the decoded-path representation: HC3 allows independent-bin jumps, whereas the HC6 sequence decoder uses P_0. These diagnostics measure agreement between the decoded sequence and behavioral transition probabilities. The contrast between HC3 and HC6 follows the different temporal assumptions used in their decoders and therefore belongs to the interpretation of the neural readout.

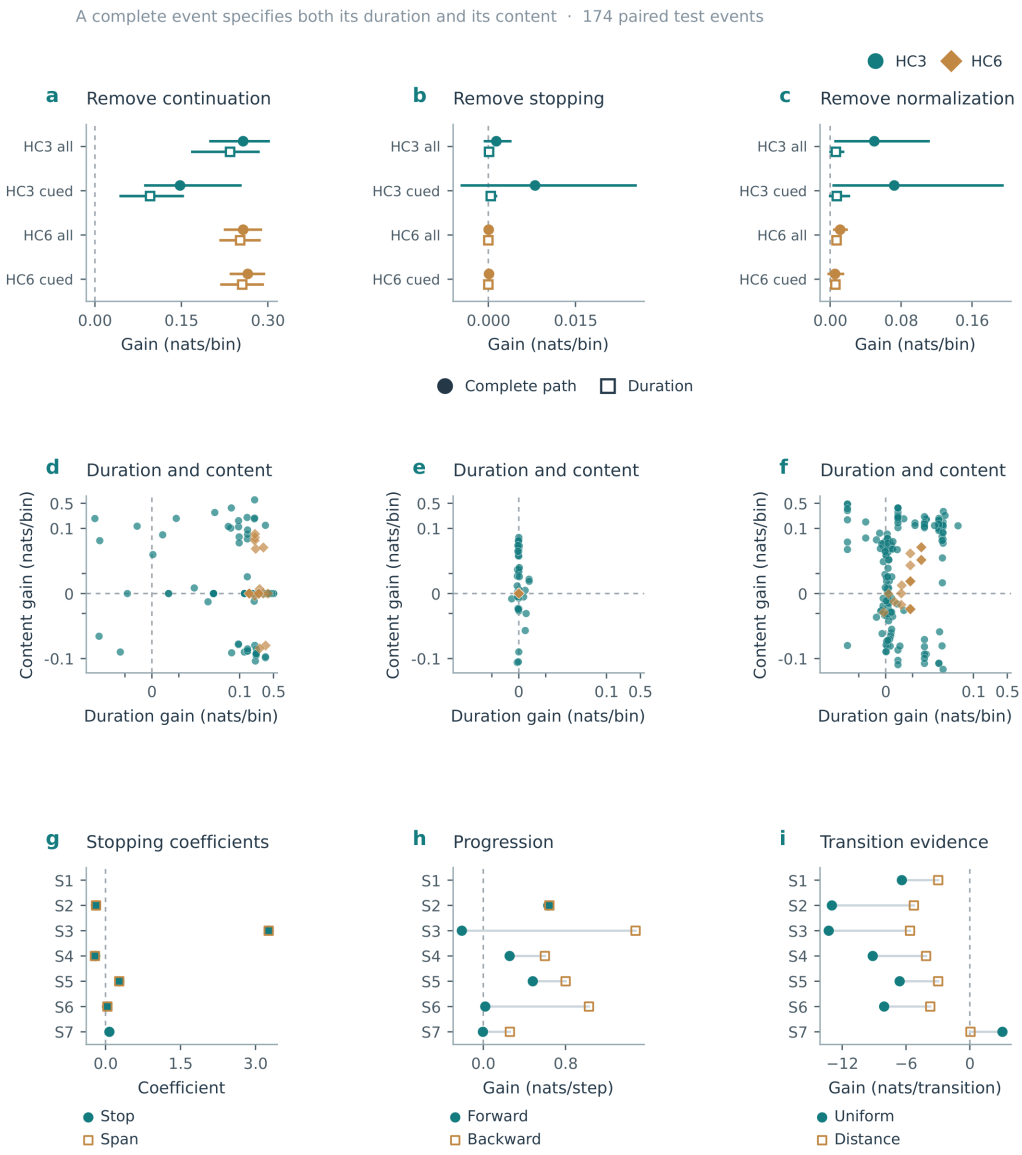

Figure 27: **How continuation, stopping, and source normalization shape complete events.** The three columns remove continuation-field input, state-dependent stopping, or source normalization. Panels a–c show complete-path and length gains with 95% paired-bootstrap intervals for full cohorts and explicit-boundary subsets. Panels d–f decompose gains for all 174 test events into length and path content conditional on length, using symmetric logarithmic axes. Panels g–i show stopping coefficients, progression gains, and internal-transition gains in their native units. Segments pair available session estimates. Deletions share the fitted parameters and the reference survival calibrated from training-event lengths.

E.5 SPATIAL RESOLUTION AND PREDICTION OF HELD-OUT SPIKES

A finer map can improve spike prediction by distinguishing nearby firing fields, even when the use of boundary information contributes little. Changing the number of states changes both place-field templates and reference transitions. To assess this choice using identical observations, the state-count analysis fixes neurons, counts, event windows, and fold membership while fitting each state representation anew. Three-fold evaluation holds every event out once, giving 460 HC3 and 54 HC6 events. This sample is distinct from the main fixed test sets. Boundary widths are selected within each training fold. The score is Equation (132), and γ is fixed at 0.82. Paired differences across representations use 2,000 animal/session/event bootstrap draws in HC3 and event draws in HC6.

HC3 scans $N \in \{18, 20, 22, 24, 25, 27, 30, 32, 35\}$. Relative to $N = 24$, the three coarser representations lose 0.05249, 0.03494, and 0.01306 nats/bin, with their pointwise 95% intervals below zero. Finer representations show gains from 0.00044 to 0.01053, with intervals spanning zero. HC6 scans $N \in \{23, 25, 27, 29, 31, 33, 35\}$. At $N = 27$, Bridge exceeds SR-value by 0.02729 $[0.01454, 0.04233]$ nats/bin. Increasing to 29 or 35 lowers prediction, while the apparent gain at 33 has an interval spanning zero. Figure 28 shows every scanned representation and the same-state model comparisons. The curves describe sensitivity within the tested ranges.

Behavioral and offline extensions. Five additional analyses assess the same-state comparison across tasks. The three behavioral datasets score held-out spike likelihood conditional on their observed lengths and behavioral cues, in nats/s. DANDI000447 uses familiar-environment segments. The HC3 extension uses a strict CA1 core with two wake cross-fitting roles weighted equally within each session. It contains 166 test events from two animals and four sessions, including 153 events with explicit endpoint cues. DANDI001695 evaluates joint spike and event-length probability for 2,065 held-out events from four animals, in nats/event. These tasks retain their own state counts and scoring definitions.

Cohort identity determines the task being scored. In the behavioral extensions, the observation is the spike sequence during a traversal whose boundaries and duration are supplied. In the sleep extension, the observation is an entire ripple event and its termination. All comparisons hold the observed counts and sampling windows fixed across representations. The following protocols specify how each cohort supplies its map, boundary evidence, and training data.

Well-to-well behavior in DANDI 000978 and 000447. DANDI 000978 supplies W-track learning trials (Shin & Jadhav, 2024). DANDI 000447 records CA1 and prefrontal activity as rats encounter novel, familiar, and revisited W-track environments; the analysis uses the familiar epoch from version 0.250724.2224 (Shin & Jadhav, 2025; Tang et al., 2023). It retains all nine ordered combinations of its three well labels, including trials returning to the same well. Trials are sorted by start time within the evaluated behavioral epoch and split into the first 65% for training and the remaining 35% for testing. Every state partition, firing-rate estimate, transition matrix, and endpoint likelihood is fitted on the training trials of that fold.

Spike counts use 20 ms bins with the actual exposure of a partial final bin. CA1 units require at least ten training spikes and a training rate no greater than 10 Hz; at most 96 units are retained, with at least eight required. Training positions are partitioned by Euclidean k-means with three initializations and 50 iterations, using seed $20{,}260{,}712 + 1009f$ for fold ordinal f. For each state, firing rates are $(n+0.5)/(o+0.5)$, where n is the spike count and o is occupancy in seconds. Adjacent valid position bins supply transitions; every transition count receives pseudocount 0.05 before row normalization. Initial-state counts receive pseudocount 0.5.

Start and end well labels are converted into evidence through their training associations with the first and final spatial states. State-by-well counts receive pseudocount 0.5 and are normalized across well labels at each state. Multiplying the initial-state prior by the observed start-well likelihood gives the initial distribution; the observed end-well likelihood supplies terminal evidence. Held-out intermediate position samples supply no factor in the spike likelihood. The score sums over latent spatial paths and divides the paired total log-likelihood difference by total observed trial exposure. Its unit is nats/s.

Table 14: State-resolution extension cohorts at their adopted N. Parentheses give 95% paired hierarchical bootstrap intervals. Animals are counted within each cohort; the HC3 core overlaps the main cohort.

Cohort	Animals	Sessions	Observations	N	Bridge–SR-value
000978	8	9	843	25	$0.01099\ (0.00511, 0.01644)$ nats/s
000447	5	5	96	25	$-0.00242\ (-0.01671, 0.01025)$ nats/s
001539	7	7	200	16	$0.04518\ (-0.13097, 0.17016)$ nats/s
HC3 CA1	2	4	166	25	$0.00182\ (-0.04264, 0.06922)$ nats/bin
001695	4	4	2065	25	$1.29538\ (0.48318, 2.31064)$ nats/event

Reward navigation in DANDI 001539. This dataset records hippocampal, prefrontal, and olfactory activity during an odor and place association task (Shin, 2025; Symanski et al., 2022). We use version 0.250815.1203 and evaluate rewarded navigation from the end of a holding period to reward entry. Trials are divided chronologically at 65% before selecting rewarded trials. Eligible observations require a holding period of at least 0.5 s, navigation lasting 0.2–15 s, and at least four 50 ms navigation bins. Actual bin exposures enter the Poisson likelihood. Position is assigned causally from the preceding sample within 100 ms. At least 80% of training navigation bins require valid position; the held-out likelihood receives the initial location and reward-goal identity, with intermediate location left latent.

Training reward-entry positions define two goal clusters. Each goal requires at least eight training and four test trials. The state representation is a separate partition of training navigation positions, with three k-means initializations, 50 iterations, and seed 20,260,716. CA1 cells require at least 20 training spikes and a rate no greater than 20 Hz. Between four and 32 cells are retained. Firing rates use 0.5 spike and 0.5 s regularization, transition counts receive 0.05, and initial-state counts receive 0.5. Initial evidence is a spatial Gaussian around the observed start; terminal evidence is a Gaussian around the assigned training goal center. Their common width is the median positive nearest-neighbor distance between state centers. Held-out goals are assigned through the training goal geometry. The seven estimable animals contribute 200 held-out navigation observations; the eighth animal lacks the required number of training-active CA1 units.

Spatial cross-fitting in the HC-3 CA1 extension. The four recordings are ec013.387, ec013.669, ec013.670, and ec016.269. Alternating 60 s waking blocks provide two complementary estimation roles. One block set estimates reference transitions and the other estimates CA1 firing templates; a second fit exchanges the roles. Their session contributions receive equal weight. Event windows, spike counts, CA1 unit identities, and train/test membership remain fixed within each spatial-resolution scan. Actual partial-bin exposures are retained. Initial width is first selected on the training Start-model likelihood. Bridge then selects terminal width, and SR-value selects terminal width and its gain parameters on the same training observations. Widths scale by $N/25$, preserving their approximate spatial meaning when state count changes. The resulting 166 held-out event evaluations include 153 with explicit boundary evidence. Conditional-spike gains are divided by bin count and combined over estimation roles, sessions, and animals.

Complete ripple events in DANDI 001695. This dataset records hippocampal and retrosplenial activity in mice during rest, spatial behavior, and sensory stimulation (Gonzalez & Voroslakos, 2026). The four mice contribute their selected earliest recordings from version 0.260319.2023. Movement defines a linear spatial coordinate and its state bins. Reference and firing templates use valid running samples above 5 coordinate units/s; official ripple intervals are excluded from this fitting support. CA1 cells require at least 20 running spikes and mean running rate between 0.05 and 15 Hz, with eight to 96 cells retained. Rates use 0.5 spike and 0.5 s regularization. Transition pseudocounts are 0.5 on self and neighboring states and 0.01 elsewhere; occupancy plus 0.01 s per state supplies the initial prior.

Official ripple intervals qualify when they last 20–500 ms, overlap annotated NREM by at least 95%, have complete LFP coverage, and occur before or after the movement interval. Within PRE and POST separately, chronological event indices 2 modulo 3 enter the test set. Spikes are counted in half-open 20 ms intervals with the actual exposure of the final interval. The fixed initial evidence is

centered on the track midpoint with width $2N/25$; terminal evidence is a symmetric mixture at the two track ends with width $1.5N/25$. Both use a 10^{-4} floor. These are geometric boundary conditions shared by all events.

Reference survival is selected from $\{0.50, 0.60, 0.70, 0.76, 0.82, 0.88, 0.92, 0.96, 0.98\}$ to minimize the absolute log difference between model mean length and training mean length. The constant stopping comparator uses the reciprocal training mean length. Its continuation choices and initial distribution are inherited from the bridge. SR-value uses the shared reference and endpoint evidence, with a constant stopping probability and training-selected continuation gain. Forward summation includes survival at each intermediate bin and completion at the final bin, giving a joint likelihood for spikes and event length. The 2,065 held-out events are scored in nats/event. This comparison directly evaluates the event-level stopping distribution alongside neural observations.

Cohorts adopting 25 states scan 15, 20, 25, 30, and 35 states. DANDI001539 adopts 16 and scans 8, 12, 16, 20, and 24. One DANDI001539 animal, CS42, has two training-active CA1 units, below the four-unit criterion, and is unestimable throughout the scan. The other seven animals supply its estimates. Across N, the observations, unit sets, windows, and splits remain identical. Bridge and SR share the spatial representation, reference kernel, and neural templates. SR selects $\gamma \in \{0.50, 0.82, 0.95\}$ and $\beta \in \{0, 0.25, 0.75, 1.10\}$ using training observations, without an extra terminal factor in its path law.

The 75 comparisons comprise Bridge versus SR, Bridge versus the adopted representation, and Bridge versus Start or the constant stopping-rate model, each across five N values and five cohorts. Figures 29 and 30 organize these comparisons by cohort. Intervals use 2,000 paired hierarchical resamples conditional on fitted models and are pointwise across N. DANDI000978 and DANDI001695 favor Bridge over SR at the adopted representation. The other cohorts have intervals spanning zero. Finer states improve mean prediction in all three behavioral cohorts, whereas additional offline resolution gives less consistent benefit. The same-state contrasts identify the contribution of boundary-conditioned path structure after the spatial representation is shared. The resolution contrasts identify the contribution of changing that representation. Together they distinguish two ways in which a more detailed predictive model can improve.

Cohort composition and score aggregation. The main HC3 animals are ec013, ec016, and gor01; HC6 uses Bon. DANDI 000978 uses ER1, JS14, JS15, JS17, JS21, JS34, KL8, and ZT2. DANDI 000447 uses AM2, JS17, JS21, JS34, and ZT2. DANDI 001539 uses CS31, CS33, CS34, CS35, CS39, CS41, and CS44. The HC3 CA1 extension uses ec013 and ec016; DANDI 001695 uses M01, M02, M03, and M05. The full session and animal identifiers are retained in the source tables. Within a cohort, every state-count comparison uses the same observation support and aggregation hierarchy. Resolution gains compare Bridge at two representations; same-resolution model gains compare Bridge with a control fitted on that same representation. Figures 28–30 report both quantities.

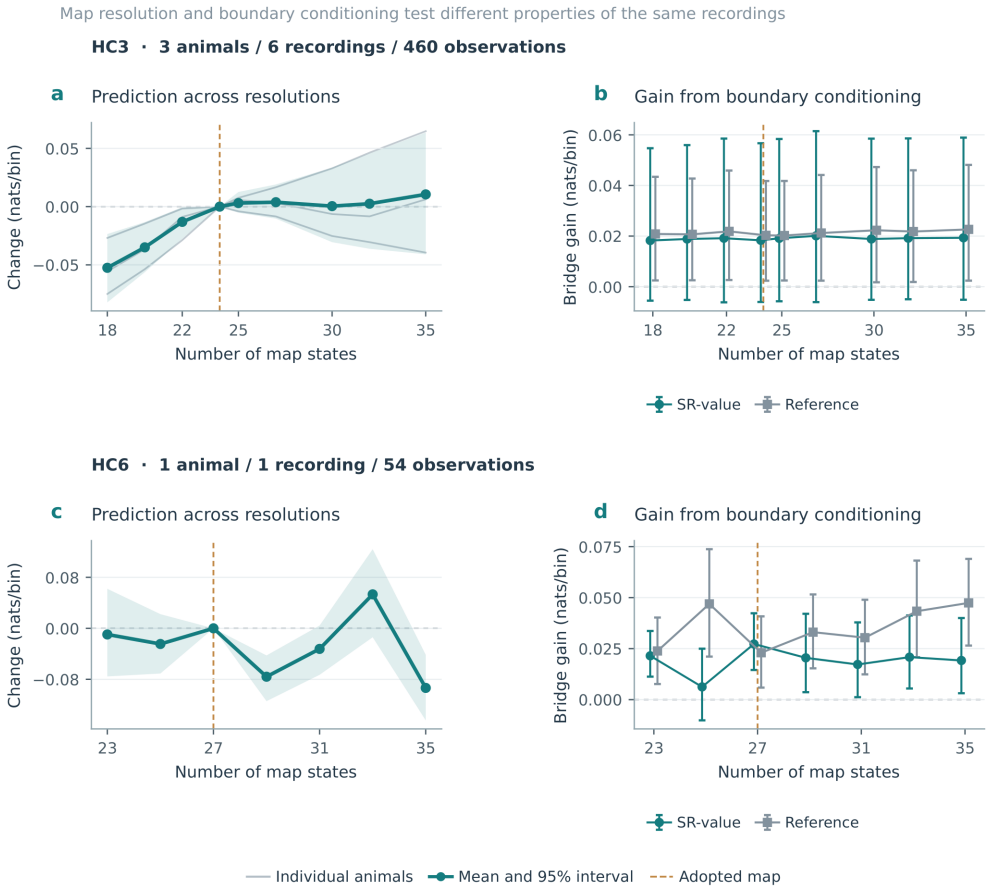

Figure 28: **Spatial resolution and conditional spike prediction.** HC3 and HC6 each show changes from the adopted representation (a,c) and same-state Bridge comparisons with SR-value and Reference (b,d). Resolution panels combine all 34 animal values with cohort means and pointwise 95% intervals; all 48 aggregate comparisons are displayed. Amber lines mark 24 or 27 states. Three-fold evaluation holds out each of 460 HC3 and 54 HC6 events once. Scores condition on event length and use nats/bin; intervals use 2,000 paired-bootstrap draws.

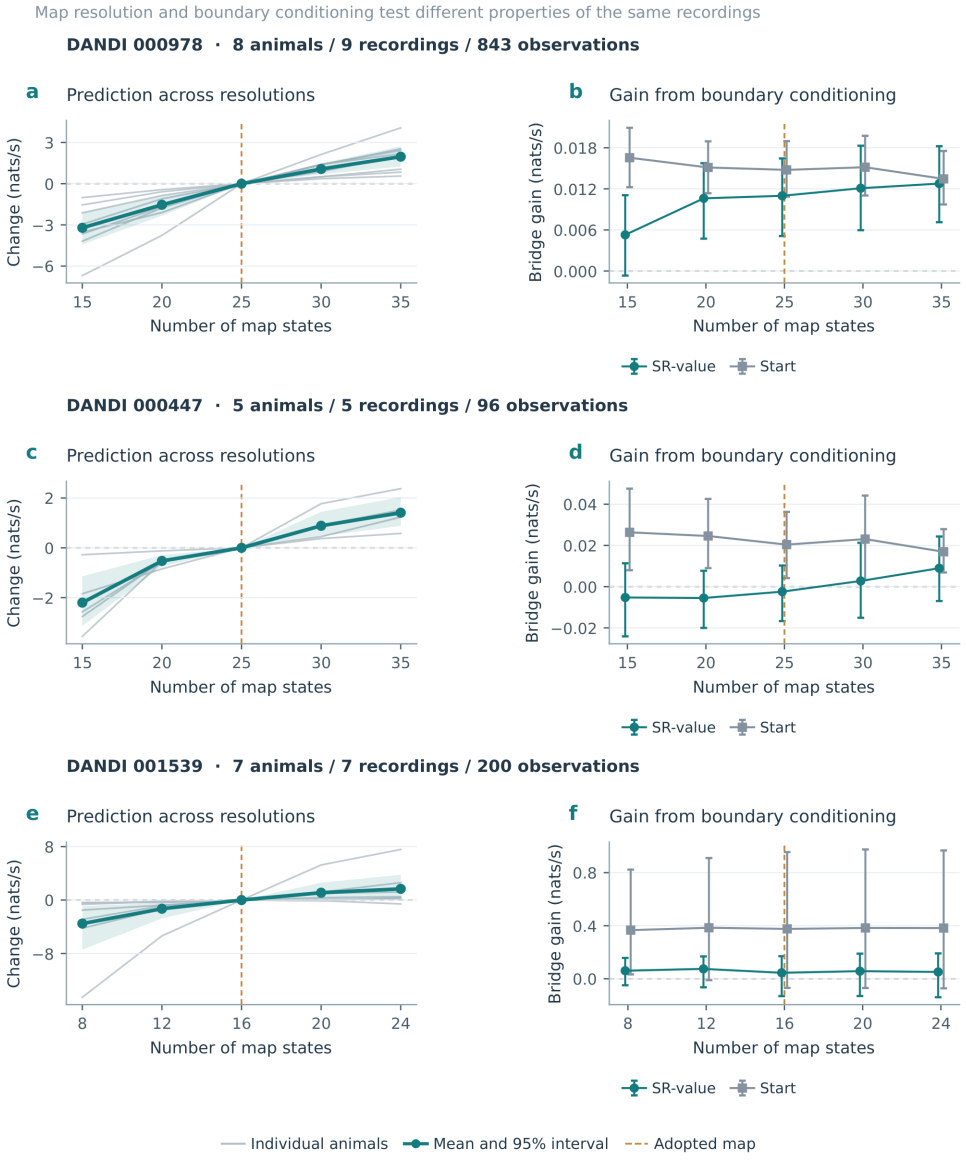

Figure 29: **Spatial detail and boundary information in behavioral prediction.** Rows show DANDI000978, 000447, and 001539. Left panels (a,c,e) combine individual animal changes from the adopted representation with cohort means and pointwise 95% intervals. Right panels (b,d,f) compare Bridge with SR-value and Start at the same resolution. The cohorts contain eight, five, and seven estimable animals with 843, 96, and 200 held-out observations. All 45 aggregate comparisons appear in nats/s. Amber lines mark adopted state counts. Observations, neurons, and test membership are fixed within each scan.

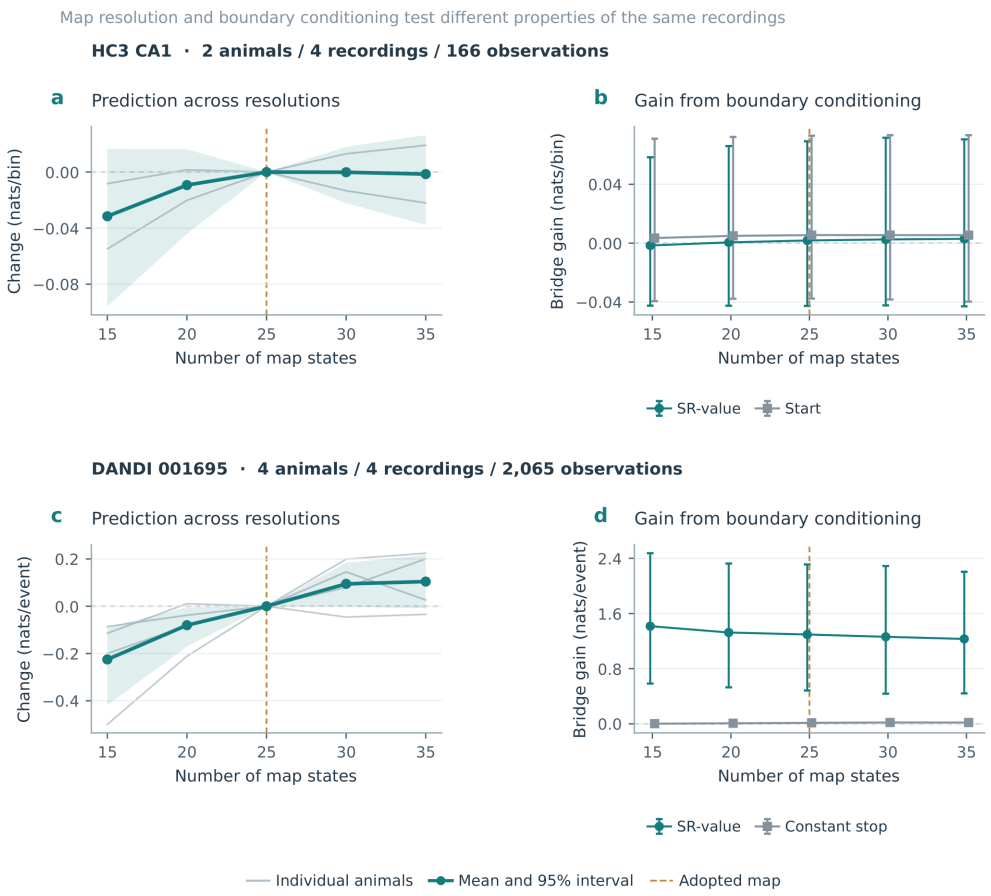

Figure 30: **Spatial detail and boundary information in offline events.** The HC3 CA1 core contains two animals, four recordings, and 166 test events; DANDI001695 contains four animals, four recordings, and 2,065 test events. Resolution panels (a,c) overlay individual animal changes, cohort means, and pointwise 95% intervals. Same-state model comparisons appear in (b,d). HC3 scores spikes conditional on length in nats/bin; DANDI001695 scores joint spike and length probability in nats/event. All 30 aggregate comparisons are shown. Amber lines mark adopted state counts; intervals use paired hierarchical resampling.

E.6 INDEPENDENT RIPPLE EVIDENCE

Independent ripple detections provide a physiological restriction of the main population-event sample. Available lists cover ec013.387, ec013.669, ec013.670, and ec016.269: two animals, four sessions, 318 candidate events, and 105 main test events. The remaining sessions lack matched usable lists. The unit of observation remains the population event, even when several ripple intervals overlap it.

Detections use a CA1-associated channel, 150–250 Hz filtering, a Hilbert envelope, and 4 ms smoothing. Robust standardized boundary and peak thresholds are 2 and 5. Accepted intervals last 20–200 ms, contain at least 80% immobility, and have an immobile peak. Raw-signal and first-difference artifact criteria include a 10 ms exclusion margin; short gaps are merged according to the detector rules. These lists establish correspondence with ripple-band activity. They contain no additional low-frequency sharp-wave morphology criterion or NREM staging.

The primary selection requires overlap between the closed population-event and ripple intervals. Sensitivities require a ripple peak inside the population event, interval gap at most 25 or 50 ms, ripple coverage of at least 80%, or complete-window coverage. Coverage is measured from the union of the intersecting ripple intervals. All boundary widths, biases, γ, and main train/test identifiers are fixed before filtering. Retained test counts are 55, 52, 58, 58, 31, and 11, respectively. The two gap criteria select the same test events.

Each subset reports complete-path gain and fixed-length spike gain separately. The paired change from all covered events to the overlap subset is computed within shared animal/session resamples, resampling matched and unmatched events separately and reconstructing the complete covered group. This preserves the dependence of the subset and its parent. All intervals use 20,000 hierarchical draws and condition on the fitted models and observed group counts.

In the covered sessions, Bridge–SR-value path gain is 0.01503 $[0.00037, 0.03292]$ nats/bin. The overlap subset gives 0.02390 $[-0.00762, 0.06730]$; its paired change is 0.00887 $[-0.01361, 0.03735]$. Individual-session overlap gains are 0, -0.00306, 0.02869, and 0.03925. All nine named controls have positive aggregate path-gain point estimates in this subset, with all intervals spanning zero. Complete-window coverage gives 0.04845 $[0, 0.17266]$, with the positive contribution concentrated in ec016.

For conditional spike prediction, the overlap Bridge–SR-value gain is -0.01678 $[-0.03668, 0.00070]$ nats/bin. At 80% coverage it is -0.02732 $[-0.05995, -0.00586]$. Ripple selection retains the positive path-score direction, while its conditional-spike comparison favors SR-value in the point estimate. The joint path law and the spike observation distribution therefore respond differently to physiological selection. Event composition also changes: explicit-boundary events rise from 15/105 to 12/55. Within that class, the covered versus overlap path gains are 0.14856 and 0.13919. The complete selection comparisons and recording coverage appear in Figure 31. The analysis connects the event-based model comparison to independently measured ripple-band activity and makes the dependence on event composition visible.

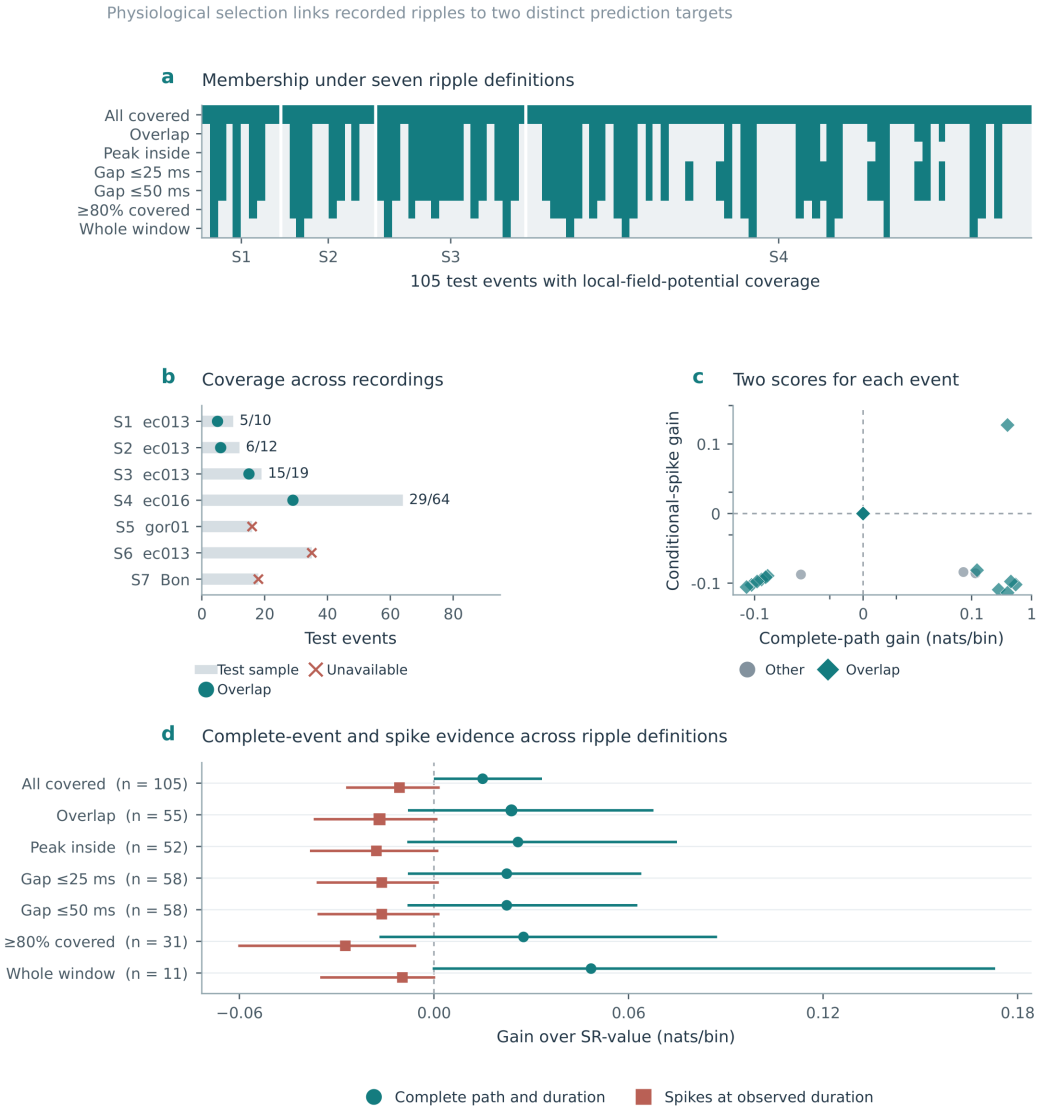

Figure 31: **Ripple correspondence and the organization of a complete event.** (a) Membership of all 105 covered test events under seven physiological selection rules, ordered by recording and time. (b) Test and overlap counts; crosses identify recordings without usable LFP lists. (c) Paired complete-path and conditional-spike gains, with 90 observations at zero on both axes. Color and shape identify ripple overlap. (d) Both predictive targets under each selection rule, on a shared axis with event counts and 95% intervals from 20,000 paired hierarchical draws. Gains are relative to SR-value. Events are nested within two animals and four recordings.

F EXPERIMENT 4: EXPERIENCE AND THE REDISTRIBUTION OF SLEEP PATHS

F.1 HOW EXPERIENCE CHANGES THE USE OF A LEARNED MAP

A learned map makes many episodes possible. Experience can change which of those episodes are frequently expressed. The endpoint rule achieves this by comparing the frequencies of beginnings and endings in waking experience with their frequencies in generated sequences. An experienced endpoint receiving too little probability gains weight; an endpoint already frequent in generation supplies a larger negative update. The reference retains the connections through which these changes affect whole paths.

The experiment follows sleep before experience (PRE), waking behavior, and sleep after experience (POST). A PRE fit supplies the model's endpoint expectation. Waking data supply the target frequencies. Their difference defines an update before POST events are scored. We ask whether the change in sleep sequences favors that direction and whether a finite step in the same direction improves prediction of later events. The step size is selected using other animals. Following subsequent sleep blocks and repeated learning cycles then relates this predictive change to the observed evolution of endpoint distributions.

HC-11 expresses the calculation in decoded position during novel-maze experience. DANDI expresses it in recurring population firing patterns during W-track learning. These complementary representations connect the proposed computation to spatial content and to distributed neural activity. Initial and terminal contributions, alternative directions, state resolution, and transfer to HC-31 identify which parts of the endpoint calculation carry the prediction. Figures 32–40 connect these measurements to individual animals and learning cycles.

HC-11 contains eight complete PRE-sleep, novel-maze, POST-sleep sessions from four rats: Achilles (two sessions), Buddy (one), Cicero (three), and Gatsby (two). Each session includes approximately 45 minutes of exploration of a novel maze. Investigator-supplied NREM intervals allow sleep events to be restricted to that state. The main event set contains 16,035 PRE and 12,246 POST population bursts. Their complete time intervals lie within a single NREM interval.

DANDI:000978 contains nine single-day W-track learning sessions from eight rats. ER1, JS14, JS15, JS17, JS21, JS34, and KL8 each contribute one session with eight sleep, behavior, sleep cycles. ZT2 contributes two sessions with four and three cycles. We define a cycle from three immediately consecutive epochs, with behavior between two sleep/rest epochs. This gives 63 cycles, of which 57 satisfy the behavioral and neural sample requirements. All nine sessions and eight animals contribute to the resulting estimates. Epochs are classified using the trial table; the population-burst analysis uses the supplied sleep/rest epochs without additional NREM staging. Both cohorts are public recordings; representations and reference transitions remain fixed within each evaluated session.

Table 15: The two primary recording designs. Event counts describe the local cross-fitted analysis. Unique sleep epochs are counted once in DANDI even when an epoch belongs to adjacent cycles.

Property	HC-11	DANDI 000978
Biological sample	Four rats; eight sessions.	Eight rats; nine sessions.
Experience	Approximately 45 min of novel-maze exploration per session.	Repeated W-track learning epochs; 63 candidate cycles.
Main state representation	12 spatial states, with latent paths decoded from CA1 spikes.	16 population prototypes assigned from CA1 spike counts.
Reference fitting	Spatial transitions during maze running.	Prototype jumps during the first behavioral epoch.
Sleep support	Complete population windows within annotated NREM.	Supplied sleep/rest epochs.
Population events	16,035 PRE and 12,246 POST.	113,359 events; 57 estimable cycles.
Ripple-containing events	6,482 PRE and 6,619 POST.	13,081 events; 53 estimable cycles.

F.2 State representations, reference transitions, and behavioral endpoints

Spatial states in HC-11. For a circular maze, position is represented by angle about the maze center. For a linear maze, the leading principal component of two-dimensional position is rescaled using its first and 99th percentiles. The main representation divides the resulting coordinate into 12 states. CA1 pyramidal cells are identified from the recording metadata and retained if they emit at least 20 spikes during maze running. Running requires speed at least 0.05 normalized track lengths per second. Cell selection and spatial tuning use waking activity.

For cell c and position state s, let N_{cs} be the running spike count and O_s the running occupancy. With mean running rate $\bar{\lambda}_c$, the initial rate estimate is

$$\lambda_{cs}^{(0)} = \frac{N_{cs} + \tau\bar{\lambda}_c}{O_s + \tau}, \qquad \tau = 0.25\,\mathrm{s}. \tag{134}$$

A Gaussian kernel with standard deviation one spatial state smooths the rates, and a floor of $0.02\,\mathrm{Hz}$ maintains positive emissions. A 20 ms decoding bin with counts n_{ct} has conditionally independent Poisson emissions

$$E_t(s) = \prod_c \frac{(\lambda_{cs}\Delta t)^{n_{ct}} e^{-\lambda_{cs}\Delta t}}{n_{ct}!}, \qquad \Delta t = 0.020\,\mathrm{s}. \tag{135}$$

Forward and backward recursions integrate the likelihood over latent position paths. The endpoint statistics therefore preserve decoding uncertainty.

Reference transitions are counted during maze running at approximately 128 ms intervals. A count of 0.05 is added to self-transitions and neighboring track states before row normalization. Circular tracks also connect the first and last states. The reference initial distribution π is the stationary distribution of this matrix P_0. Behavioral paths are nonoverlapping running segments that accumulate approximately half a track length of travel. Their initial and final states supply the two waking endpoint distributions. We require at least ten such segments and add 0.5 to every endpoint-state count before normalization.

Population states in DANDI. The first 65% of the earliest sleep/rest epoch defines a fixed population representation. Units with region metadata equal to CA1 are eligible when they emit at least 40 spikes in this segment and have mean rate below 5 Hz. We retain at most 96 eligible units, ranked by spike count, and require at least eight. Population activity is binned at 20 ms. If X_t is the population count, active bins satisfy

$$X_t \geq c_{\mathrm{act}}, \qquad c_{\mathrm{act}} = \max\{2, Q_{0.75}(\{X_t : X_t > 0\})\}. \tag{136}$$

Each active-bin vector is transformed with $\log(1 + n_{ct})$ and standardized using training means and standard deviations. Clustering into 16 groups uses three fixed initializations, up to 45 Lloyd iterations, and the solution with lowest squared reconstruction error. Empty clusters are initialized again at the observation with greatest current error. The resulting centers define neural population prototypes. Subsequent active bins are assigned to the nearest center using the fixed cell set, threshold, and standardization. Prototype identity describes a firing pattern; it has no imposed spatial coordinate.

The reference matrix uses valid trials in the first behavioral epoch. A valid trial has finite ordered boundaries, lies entirely in one behavioral epoch, and joins different nonmissing start and end wells. Gaps longer than two inactive bins break a state sequence. Consecutive repeated states are collapsed to form a jump chain. Transition counts receive a pseudocount of 0.05 in every matrix entry and are row-normalized; the representation must yield at least 20 jumps and eight represented states. This reference matrix is fixed throughout the session. For ZT2 on 2024-05-03, the first behavioral epoch precedes the earliest sleep epoch used to construct prototypes; subsequent evaluated sleep, behavior, sleep cycles use the same fixed representation.

For each later behavioral trial, its earliest and latest active prototypes contribute to the waking endpoint targets. At least 20 endpoint pairs are required per behavioral epoch, with count smoothing of 0.5 per state. A separate correct-trial target requires at least ten pairs. Within each dataset, waking targets, sleep statistics, and reference transitions share one state space.

F.3 Population events and ripple screening

HC-11 detection uses 10 ms population spike bins smoothed with a 20 ms Gaussian kernel. Bins above the 99th percentile of NREM activity are connected with a three-bin closing operation and

padded by 20 ms on each side. Eligible windows last 40–350 ms, contain at least seven spikes from three active cells, and provide at least two nonempty decoding bins, with at most 18 bins. Complete containment in one NREM interval is required.

DANDI detection joins active bins across gaps of at most two inactive bins. Inactive bins link a segment without receiving a state identity. Segments are split chronologically into blocks of at most 50 active states; final blocks with fewer than three raw states are discarded. Consecutive repetitions are then collapsed. All paths originating from one detection segment remain grouped during cross-fitting. Counting each sleep epoch once gives 113,359 population events before ripple screening.

The physiological sensitivity analysis retains an eligible population window when it contains a qualifying ripple peak on a selected CA1 channel. Channel choice and normalization use preceding sleep. A fourth-order Butterworth filter, applied forward and backward, isolates 150–250 Hz activity; the Hilbert amplitude is smoothed at 4 ms. Peaks must reach five standard deviations above the calibration mean, with boundaries at two standard deviations and ripple duration 40–250 ms. Candidates with raw deviations exceeding 20 calibration standard deviations are rejected. The original population-event boundaries remain fixed. This produces ripple-containing population bursts; sharp-wave morphology across hippocampal layers is outside this detection criterion.

The ripple-containing HC-11 sample has 6,482 PRE and 6,619 POST events. DANDI retains 13,081 events across sleep/rest epochs and 53 estimable cycles. The 63 candidate cycles yield 107 estimable fold slots; a cycle contributes to the primary mean when both folds are estimable. Six cycles lack sufficient behavioral endpoint pairs under both event definitions; four additional cycles lack sufficient neural events after ripple screening. Eighty-nine population windows lack full LFP coverage. All eight animals remain represented.

F.4 Path likelihood and the endpoint score

The reference transition matrix expresses the relationships learned between successive states. Initial and terminal weights determine how often paths with different endpoints are expressed. Updating these weights redistributes probability among complete paths. Paths that share both endpoints retain their relative probabilities under the reference, so the learning rule changes the use of the map through a small set of boundary variables. Write $z = (\theta, \eta)$ for the initial and terminal log weights. Conditional on a path having L transitions, its probability is

$$p_z(x_{0:L} \mid L) = \frac{\pi(x_0) e^{\theta(x_0)} \prod_{t=0}^{L-1} P_0(x_t, x_{t+1}) e^{\eta(x_L)}}{Z_L(z)}, \qquad Z_L(z) = \sum_{i,j} \pi_i e^{\theta_i} (P_0^L)_{ij} e^{\eta_j}. \quad (137)$$

The normalized joint endpoint matrix is proportional to $\pi_i e^{\theta_i} (P_0^L)_{ij} e^{\eta_j}$. Its row and column sums give $q_{0,L}$ and $q_{L,L}$. The model marginals q_0, q_T average these quantities over the preceding-sleep training length distribution, weighting events equally. Conditioning on length concentrates the comparison on the distribution of paths at the observed duration.

For an observed prototype sequence, differentiation of Equation (137) gives

$$\nabla_\theta \log p_z = \mathbf{e}_{x_0} - q_{0,L}, \qquad \nabla_\eta \log p_z = \mathbf{e}_{x_L} - q_{L,L}. \quad (138)$$

Here $\mathbf{e}_i$ denotes a point mass at state i. In HC-11 the likelihood sums over paths with emissions from Equation (135). Differentiating that sum replaces each point mass by its posterior endpoint distribution given the event. Thus both analyses measure data-supported endpoints relative to model endpoints, with the spatial analysis averaging over its latent-state uncertainty.

Potentials are centered to remove additive constants that leave normalized probabilities unchanged. They are fitted by penalized mean log likelihood using a smoothness coefficient of 0.02 and mean-squared L_2 coefficient of 10^{-4}. Smoothness uses neighboring spatial states in HC-11 and the three-nearest-neighbor prototype graph in DANDI. Optimization starts at zero potentials with L-BFGS-B. Cross-fitting allows up to 600 iterations and checks the penalized gradient norm against 5×10^{-4}. All evaluation scores use data likelihood, without the fitting penalty.

F.5 Temporal cross-fitting and the local prediction

The local and finite-update analyses use different fits for different questions. The local analysis fits one temporal half of PRE and compares independent PRE and POST groups at that common fit. The finite-update analysis fits the PRE event set and asks how a prescribed update predicts POST in an animal excluded from step-size selection. Keeping these roles explicit allows the direction of change and the usefulness of a finite learning step to be read separately.

Within each sleep epoch, events are sorted chronologically and divided as evenly as possible into early and late groups. HC-11 cuts occur between connected components of overlapping windows, including transitive overlaps. DANDI cuts occur between complete detection segments. Ties use the earliest allowable cut. The split depends on event times and grouping.

There are two folds. The first fits late PRE and evaluates early PRE against early POST. The second fits early PRE and evaluates late PRE against late POST. At least ten events are required in each DANDI training and testing group. Neighboring cycles sharing a sleep epoch retain the same temporal split.

At the PRE fit, a positive waking residual calls for greater weight on an endpoint state, and a negative residual calls for less. A likelihood gradient describes which of these changes would increase support for a set of observed sleep events. Comparing the POST and PRE gradients therefore measures how this preference changes across the intervening waking episode. For fold f, the PRE training fit z_f defines the residual

$$r_f = (\mu_0 - q_0(z_f),\ \mu_T - q_T(z_f)). \tag{139}$$

Let $g(D; z_f)$ be the mean endpoint likelihood gradient of event set D at this fixed fit. The local statistic is

$$\mathcal{T}_f = \langle g(D_{\mathrm{POST},f}; z_f) - g(D_{\mathrm{PRE},f}; z_f),\ r_f \rangle . \tag{140}$$

Equivalently, it is the derivative at $\alpha = 0$ of the POST-minus-PRE mean likelihood difference along $z_f + \alpha r_f$. A positive value means that the observed change across behavior favors the update predicted by the waking endpoint residual. Evaluation of PRE and POST at the same fitted potentials makes the comparison local to one common model.

The two folds are averaged equally. HC-11 then averages sessions within animals and animals within the cohort. DANDI first averages cycles within sessions, then sessions within animals, then animals within the cohort. Each level receives equal weight. ZT2 therefore contributes one animal to the DANDI mean. Units are nats/event per unit step for HC-11 and nats/burst per unit step for DANDI. Their numerical effect sizes are reported separately.

The initial and terminal contributions sum to $\mathcal{T}_f$. A supplementary directional cosine divides their joint inner product by the product of the concatenated residual and score-change norms. Blocks are centered before this calculation. Cosine is undefined if either norm is at most 10^{-12}; otherwise it describes orientation independently of magnitude. Cosines are computed within folds and averaged through the same hierarchy.

Length matching. To examine a possible change in event-length composition, we recompute score differences separately for lengths shared by PRE and POST, with common weights

$$w_L = \frac{n_{\mathrm{PRE},L} + n_{\mathrm{POST},L}}{\sum_{J \in \mathcal{L}_{\mathrm{common}}} (n_{\mathrm{PRE},J} + n_{\mathrm{POST},J})}. \tag{141}$$

The resulting mean is a comparison at a shared length distribution. Coverage is reported relative to the eligible PRE and POST samples.

Structured directions. Let u be uniform over states. We evaluate the full residual $\mu - q$, waking bias $\mu - u$, and current-sleep deficit $u - q$ at the same score difference. The latter two correspond to the Wake-only and Replay-only controls. Their scores add exactly to the full-residual score. HC-11 additionally shifts both targets by half the spatial state count. DANDI substitutes the next behavioral epoch's targets within the same session, where available. Negative residuals provide the algebraic sign reversal of the local statistic.

Generic tangent directions assess geometric specificity. Their starting and terminal blocks are separately centered and rescaled to the corresponding residual norms. HC-11 uses 128 directions

Table 16: Local endpoint calibration. Means weight animals equally. Ripple denotes population windows containing a qualifying ripple peak. The initial and terminal entries are the two additive contributions to $\mathcal{T}$.

Dataset	Events	$\mathcal{T}$	Initial	Terminal	Matched length	Positive animals
HC-11	MUA	0.065127	0.034795	0.030333	0.063212	3/4
HC-11	Ripple	0.079737	0.038485	0.041251	0.078986	3/4
DANDI	MUA	0.009086	0.004126	0.004961	0.009076	7/8
DANDI	Ripple	0.005257	0.001704	0.003553	0.004938	6/8

and DANDI uses 256, generated deterministically before evaluating POST scores. DANDI maps SHA-256 keys containing cycle, fold, direction, endpoint block, and state coordinate to entries in $(-1, 1)$; centering gives zero-sum directions. Direction indices are preserved through aggregation. The full residual's percentile among these scores describes its relative alignment in this reference set.

F.6 Experience predicts the direction of endpoint reallocation

The change across experience favors the waking residual in both representations and under both event definitions (Table 16). The positive cohort means arise from identifiable contributions at initial and terminal states, whose spatial organization is visible in Figure 32. HC-11 animal values for MUA and ripple-containing events are respectively Achilles, 0.122127 and 0.170405; Buddy, -0.003508 and -0.013963; Cicero, 0.045610 and 0.036255; and Gatsby, 0.096280 and 0.126249. DANDI has seven positive animals in the MUA analysis and six after ripple screening. Its matched set of 53 cycles gives means 0.007658 and 0.005257, respectively, showing that the smaller ripple estimate also occurs within the shared cycle sample. Deleting any single animal preserves a positive mean in each dataset and event definition.

Initial and terminal states each contribute positively at the cohort level. Matching event lengths leaves the mean scores close to their original values (Figure 34), linking the effect to endpoint allocation across paths of comparable duration. In HC-11, common lengths cover 99.75% of PRE and 99.96% of POST MUA events, and 99.35% and 99.83% after ripple screening. Corresponding DANDI coverage is 99.42% and 99.45% for MUA, and 94.79% and 94.97% for ripple-containing events. Mean directional cosines are 0.362 and 0.316 in HC-11, and 0.0895 and 0.0598 in DANDI.

The decomposition into waking bias and current-sleep deficit asks what the positive score means biologically. Waking bias favors endpoints prevalent in behavior; the sleep deficit reduces the influence of endpoints already prevalent in preceding sleep. Their sum is the full residual. The animal-level control matrices in Figure 33 show how these contributions vary across recordings. HC-11 full-residual means exceed the Wake-only, Replay-only, and shifted-target means: 0.065127 compared with 0.003702, 0.061425, and 0.061695 for MUA; 0.079737 compared with 0.003633, 0.076103, and 0.076837 for ripple-containing events. DANDI gives 0.009086 compared with -0.000616, 0.009702, and 0.013112 for MUA, and 0.005257 compared with -0.001852, 0.007109, and 0.008338 after ripple screening. In DANDI, the current-sleep deficit carries the dominant positive component, and the next-target direction has a larger mean than the immediately adjacent target. Much of the DANDI effect therefore reflects redistribution away from endpoints already overrepresented in preceding sleep. The similar scores for successive waking targets indicate a substantial shared component in the fixed population representation. The recordings associate sleep reorganization with both experienced endpoints and the distribution already expressed by the network. The full residual exceeds all generic tangent scores in both MUA analyses and in ripple-containing HC-11; its percentile is 94.92 in ripple-containing DANDI.

Resolution changes preserve positive cohort means. HC-11 estimates at 12, 16, and 24 states are 0.065127, 0.054342, and 0.053252 for MUA, and 0.079737, 0.078323, and 0.059888 after ripple screening. DANDI estimates at 12 and 16 states are 0.006570 and 0.009086 for MUA, and 0.004120 and 0.005257 after screening. These settings retain all animals. With eight DANDI prototypes, only four animals and 29 MUA or 25 ripple-containing cycles satisfy the representation and sample criteria; their positive means describe that reduced sample.

Animal-level one-sided sign flips give descriptive tail proportions of 0.125 for both HC-11 event definitions, and 0.01172 and 0.04297 for DANDI. The calculation enumerates all 2^m sign assignments to m animal effects and compares their equal-weight means to the observed mean. Its reference assumptions are independent animals and sign-symmetric effects under the null. State-resolution and analysis choices were examined on these recordings, so these values characterize the fitted analysis. Figures display all animal values. Behavioral time-course intervals use 10,000 animal bootstrap resamples after aggregation within animals.

F.7 From a learning direction to prediction of subsequent sleep

Predictive effect of a finite update. A positive local score specifies a direction of change. Its predictive use also requires an update size: a small displacement can improve fit while a larger displacement in the same direction can overshoot. Figure 35 follows this step from the training likelihood landscape to prediction in a held-out animal. Normalization makes the likelihood a curved function of endpoint weights, so the update size is selected using the other animals. We fit a baseline to PRE, take the residual step $z' = z + \alpha r$, and evaluate the POST mean likelihood gain

$$\Delta\ell(\alpha) = \ell_{\mathrm{POST}}(z + \alpha r) - \ell_{\mathrm{POST}}(z). \tag{142}$$

For each animal, the remaining animals select α from zero and 61 logarithmically spaced values between 0.01 and 1000. The held-out animal receives that selected step. The same selected step is used for component, shifted-target, and reversed-direction controls. Equation (140) measures how directional preference changes across behavior; the finite gain measures prediction after an actual update. The paired measurements ask whether an experience-related local direction also yields useful prediction at a finite step.

HC-11 mean finite gains are 0.035904 nats/event for population bursts and 0.044951 for ripple-containing bursts. Four and three animals, respectively, show positive gains. The population-burst finite-step sample includes 12,255 POST events; its eligibility permits nine windows that cross an NREM boundary. The ripple-containing finite-step sample uses the strict NREM definition and 6,619 POST events. Descriptive hierarchical 95% intervals are $[-0.002243, 0.078448]$ and $[-0.019446, 0.110966]$. DANDI gains are 0.002739 and 0.013723 nats/burst, with seven and six positive animals. Reversing the update gives negative mean gains, -0.089016 and -0.120636 in HC-11 and -0.008598 and -0.042017 in DANDI. DANDI Replay-only and shifted-target finite gains remain close to or above full-residual gains. A model fitted to POST provides a descriptive reference for the improvement available from those events. The finite updates support predictive use of the endpoint residual, with the control comparisons locating the contribution of its two terms.

Relation to proportional scaling. Proportional scaling offers a second expression of endpoint correction: increasing an endpoint weight in proportion to the ratio between its target and current frequency. For HC-11 we compare the centered $\log(\mu/q)$ direction to $\mu - q$, matching endpoint-block norms (Figure 39). Their mean directional cosine is 0.839 for MUA and 0.831 for ripple-containing events. The proportional direction gives positive local scores, 0.032402 and 0.042041, with 11 of 16 folds and three of four animals positive in each analysis.

A numerical initial-then-terminal proportional sweep decreases mean fitted endpoint TV from 0.546849 to 0.150909 for MUA and from 0.573596 to 0.148103 for ripple-containing events. The fixed-length mixture has dual $\mu_0 \cdot \theta + \mu_T \cdot \eta - \sum_L w_L \log Z_L$; the log-ratio sweep supplies a proportional correction, while an exact block fit solves the corresponding mixture-marginal constraint. Solving each coordinate for this mixture gives endpoint TV values 0.150798 and 0.148247. Every fold contracts its fitted residual. The proportional sweep changes POST likelihood more favorably than held-out PRE likelihood by 0.221897 and 0.295873 nats/event. Both absolute likelihood changes are negative: POST changes are -0.508526 and -0.532886, and PRE changes are -0.730423 and -0.828759. The relative preference for POST and the large-step loss in absolute fit describe distinct properties of this update.

Evolution within POST sleep. We divide HC-11 POST events into three consecutive groups of approximately equal count and fit each group's endpoint distribution. Under a common length distribution, the sum of initial and terminal TV distances to the waking targets is 1.00599, 1.03266, and 1.04504 across the three MUA groups, and 1.03286, 1.09105, and 1.09604 after ripple screening.

No session shows strict monotonic contraction; one of four animals shows a net decrease under either definition. Across the observed POST groups, the endpoint distribution continues to vary and its distance to the waking target increases on average. The experience-related direction and the subsequent temporal evolution therefore describe two stages of sleep reorganization. Across experience, the likelihood change favors the waking residual. During the recorded POST period, endpoint allocation continues to move, with average distance to that target increasing.

F.8 Repeated experience stabilizes sleep as behavior becomes more efficient

Repeated cycles allow the endpoint analysis to be placed within learning itself. DANDI supplies up to eight successive cycles within a day, during which performance improves and exploration becomes more concentrated. We summarize early cycles 1–2, middle cycles 3–5, and late cycles 6–8 (Figures 36 and 37). Each stage average is first calculated within sessions and animals. Neural stabilization compares endpoint marginals between adjacent sleep epochs:

$$D_{\text{sleep}} = \tfrac{1}{2}\big[\mathrm{TV}(q_0^{\mathrm{PRE}}, q_0^{\mathrm{POST}}) + \mathrm{TV}(q_T^{\mathrm{PRE}}, q_T^{\mathrm{POST}})\big], \qquad \mathrm{TV}(p,q) = \tfrac{1}{2}\sum_i |p_i - q_i|. \quad (143)$$

This distance asks how much the sleep distribution continues to change as experience accumulates. Distance to the waking target asks how closely its endpoint frequencies reproduce the experienced endpoints. The two distances distinguish stability across sleep epochs from agreement with behavior.

Behavioral accuracy increases from 0.5663 in the early stage to 0.8542 in the middle and 0.9001 in the late stage. Adjacent-sleep distance declines from 0.1611 to 0.1406 to 0.1264 for MUA events and from 0.2464 to 0.2086 to 0.1710 for ripple-containing events. These estimates pool the specified cycle ranges within each stage. Local calibration scores vary across cycles: their stage means are 0.001871, 0.010696, and 0.008747 for MUA, and -0.005375, 0.009411, and 0.004840 after ripple screening. Sleep distributions become more stable as behavioral accuracy rises, while the local score varies across the learning period.

Position measurements characterize the behavioral change. Time fractions use position-sample exposure weights. Well zones have radii 10, 15, or 20 cm; stationary samples have speed below 5 cm/s, and running samples exceed 10 cm/s. Spatial coverage is the fraction of session-wide occupied 5 cm grid cells visited during a behavioral epoch. Trial duration and valid trials per minute describe task efficiency.

Within 15 cm of a well, stationary time rises from 0.2006 to 0.2520 to 0.2734. Grid coverage falls from 0.8214 to 0.6859 to 0.6514. Median trial duration falls from 29.50 s to 22.41 s to 22.36 s, and valid trials per minute rise from 1.72 to 2.54 before settling at 2.41. Running outside the 15 cm well zone remains close to 0.26 throughout. Together these observations describe more efficient, spatially concentrated behavior with greater endpoint dwell. The behavioral measurements are shared by the two neural event analyses because ripple screening changes sleep-event selection.

F.9 Transfer across cohorts and structure within sleep paths

HC-31 transfer. To test the portability of the update magnitude, we evaluate a fixed value of $\alpha = 1.7782794100389228$, selected in HC-11, on the HC-31 recording cohort from a 48 m linear track (Rich et al., 2014; 2022). Author-linearized run1 data define 16 spatial states, decoding, and behavioral endpoints. Running speed is at least 0.03 m/s; units must emit at least 20 run1 running spikes. Reference transitions use a 128 ms lag and count smoothing of 0.05. Behavioral endpoint fragments span half the track. High-MUA bursts in the complete declared sleep1 and sleep3 rest epochs provide PRE and POST samples, with the HC-11 duration, spike-count, and active-cell requirements. Four of five attempted animals have at least 40 bursts in both periods. The excluded animal has 20 PRE and 59 POST bursts. The four estimable animals contribute 46, 30, 50, and 38 decoder cells; PRE counts are 570, 220, 313, and 329, and POST counts are 523, 182, 632, and 566. Mean POST gain over persistence is 0.062632 nats/event. Animal gains are 0.071937, 0.308592, 0.041060, and -0.171060, with descriptive hierarchical 95% interval $[-0.103700, 0.234610]$ from 5,000 resamples. The one-sided animal sign-flip tail proportion is 0.3125. This transfer tests a fixed update in another cohort and shows substantial between-animal variation.

HC-3 temporal characterization. Endpoint reallocation also raises a question about the paths connecting those endpoints. We examine both the coupling of initial and terminal states and the temporal organization of intervening states in one rat, ec016, with one waking session and three sleep sessions (Figure 40). Thirty CA1 pyramidal cells, 126 traversals, and 43 position states define the representation. Candidate detection uses 10 ms timing, channelwise robust normalization of 120–250 Hz LFP features, the maximum across three selected CA1 channels, high and low thresholds of 1.5 and 0.5, merging gaps up to 20 ms, and 30 ms padding. Overlapping or touching padded windows are merged. Eligible windows last 40–350 ms, contain at least seven selected-cell spikes in two nonempty decoding bins, and provide at least two bits of endpoint information. A local HMM uses spatial width 1.4 and the stationary reference prior. Sleep-state interpretation follows the recorded sleep-session labels.

The three sessions retain 656, 540, and 188 events. Mean endpoint TV values are 0.7114, 0.7903, and 0.7780; event-bootstrap 95% intervals using 2,400 resamples are $[0.694, 0.730]$, $[0.773, 0.809]$, and $[0.751, 0.809]$. Initial-to-final contraction, defined as first minus last distance, is -0.0666 with interval $[-0.1012, -0.0323]$. The endpoint model improves held-out negative log likelihood relative to the reference by 0.1845, 0.3724, and 0.1318 nats/event. These single-animal measurements show predictive endpoint structure alongside a nonmonotonic distance trajectory. Their event-level uncertainty describes variation within this recording.

The HC-3 distance target is the normalized sum of waking initial and terminal marginals, each smoothed over one spatial state. Its reference endpoint coupling is proportional to $\mathrm{diag}(\pi)(I - 0.82P_0)^{-1}$. Five-fold event cross-validation compares this reference to an initial-only fit, a terminal-only fit, a two-potential fit, and products of independently fitted endpoint marginals. The independent fits use either posterior endpoint mass or the maximum-posterior endpoint pair. Training endpoint counts receive a pseudocount of 0.001 per state pair. Two-potential fitting uses proportional scaling with marginal tolerance 10^{-12} and at most 10,000 iterations. Evaluation is posterior-weighted endpoint cross-entropy. The two-potential model's S3 gain relative to initial-only and terminal-only fits is -0.02998 and -0.01881 nats/event, respectively.

Intermediate-order comparisons preserve each decoded path's endpoints and number of states. For each eligible event, 64 interior permutations or independent uniform interior draws supply the controls; a third control interpolates a shortest direct path between the endpoints. Mean transition-score advantages over permutation are 39.19, 28.38, and 23.83 nats/event across S1–S3, and advantages over uniform interiors are 325.64, 308.51, and 297.73. The direct-path comparison gives -1.102, -0.760, and -0.665 nats/event. Posterior-averaged transition scores exceed an independent-state readout by 4.958, 5.092, and 5.205 nats/step. These comparisons characterize paths under the specified spatial decoder and reference model. Events with fewer than three states have no interior, giving 656, 539, and 187 eligible intermediate-path comparisons.

F.10 Endpoint summaries and illustrative conditions

A spatial example of the learning direction. The example is the Achilles session on 25 October 2013, selected by alphanumeric session order. Its endpoint distributions are complete-session PRE and POST fits in the same 12-state partition, together with the waking target frequencies. The local-score example uses ripple-containing events and temporal fold 0. At each state, its contribution is the product of the waking residual and the PRE-to-POST gradient change evaluated at the same PRE training fit. Summing initial-state contributions gives 0.141620 nats/event per unit step; summing terminal-state contributions gives 0.148037. Their total is 0.289658, with joint cosine 0.832664. The same calculation is performed for all 16 folds across the eight sessions. Complete-session endpoint distributions provide descriptive comparisons; the cross-fitted score uses its specified independent PRE and POST groups (Figures 32 and 33).

Contributions across learning cycles. All 63 cycles from nine sessions and eight rats enter the behavioral summaries. At successive cycle positions 1–8, MUA contributes 7, 8, 8, 8, 6, 6, 7, and 7 estimable session cycles, totaling 57. Ripple-containing events contribute 6, 8, 8, 8, 6, 5, 6, and 6, totaling 53. The behavioral observations are shared by the two event definitions. Stage summaries pool cycles 1–2, 3–5, and 6–8 within sessions and animals. Behavioral time courses and stage intervals use 10,000 animal bootstrap resamples. These counts specify the sample behind the temporal estimates in Figures 36 and 37.

When behavioral measures are compared as percentage changes, the transformation is $100(m_t/m_1 - 1)$, where m_t is the cohort mean at cycle t and m_1 its mean at the first cycle. The underlying trial durations, rates, speeds, coverage, and time fractions remain available in their native units. Percentage change summarizes variation in measures with different physical units; all inferential intervals use the native measures and their stated animal aggregation.

Numerical correction and temporal change. The fitted endpoint error after a proportional sweep averages initial and terminal TV. The temporal HC-11 POST analysis reports their sum. Both are functions of the same two endpoint marginals, with a factor of two between the conventions. A local likelihood derivative, a finite-step predictive gain, and a change in endpoint distance quantify different properties of an update. The separate estimates in Figure 39 retain those definitions throughout.

Endpoint coupling and interior paths in HC-3. The illustrative empirical S1 coupling sums posterior endpoint-pair mass across all 656 events, adds 0.001 to each of the 43×43 pairs, and normalizes to unit mass. The corresponding model illustration is the two-potential fit from S1 training fold 0. Its endpoint constraints come from that training fold and its pairwise structure from the reference kernel. Held-out coupling comparisons use all five folds in each of S1–S3 (Figure 40). The displayed empirical and fitted matrices are descriptive examples with these different fitting supports.

For temporal endpoint distance, first-period minus second-period changes are -0.078934 for S1 to S2, 0.012301 for S2 to S3, and -0.066633 for S1 to S3. Their 95% intervals from 2,400 event bootstrap resamples are $[-0.103336, -0.054120]$, $[-0.023606, 0.046279]$, and $[-0.101165, -0.032316]$. Interior-path comparisons contain 656, 539, and 187 eligible events. Each event keeps its beginning, ending, and number of states. The intermediate-path comparison uses a shortest direct path between the same endpoints; the interior-order comparison uses posterior-averaged transitions against an independent-state readout. All these estimates describe the single HC-3 rat and preserve event-level variation.

F.11 Computational materials

The experiment directory provides analysis code, state and model caches, event and fold tables, complete cycle records, and machine-readable statistics. Source tables contain every plotted value, and the plotting program rebuilds all figures. Dataset manifests specify access to the raw recordings; the accompanying caches and sufficient statistics reproduce the reported comparisons.



Figure 32: **How spatial endpoint changes produce the local learning score.** (a,b) Waking residual (ochre) and PRE-to-POST likelihood-gradient change (teal), in the same spatial order for initial and terminal endpoints. The example uses ripple-containing events from Achilles on 25 October 2013, temporal fold 0, selected by session and fold order. (c,d) Their coordinate products; teal and terracotta identify positive and negative contributions. (e) All coordinate products across 16 folds from eight sessions and four rats. Initial and terminal states share one signed scale. A, B, C, and G denote Achilles, Buddy, Cicero, and Gatsby; session numbers follow alphanumeric order within rat. Each row sums to its local learning score.

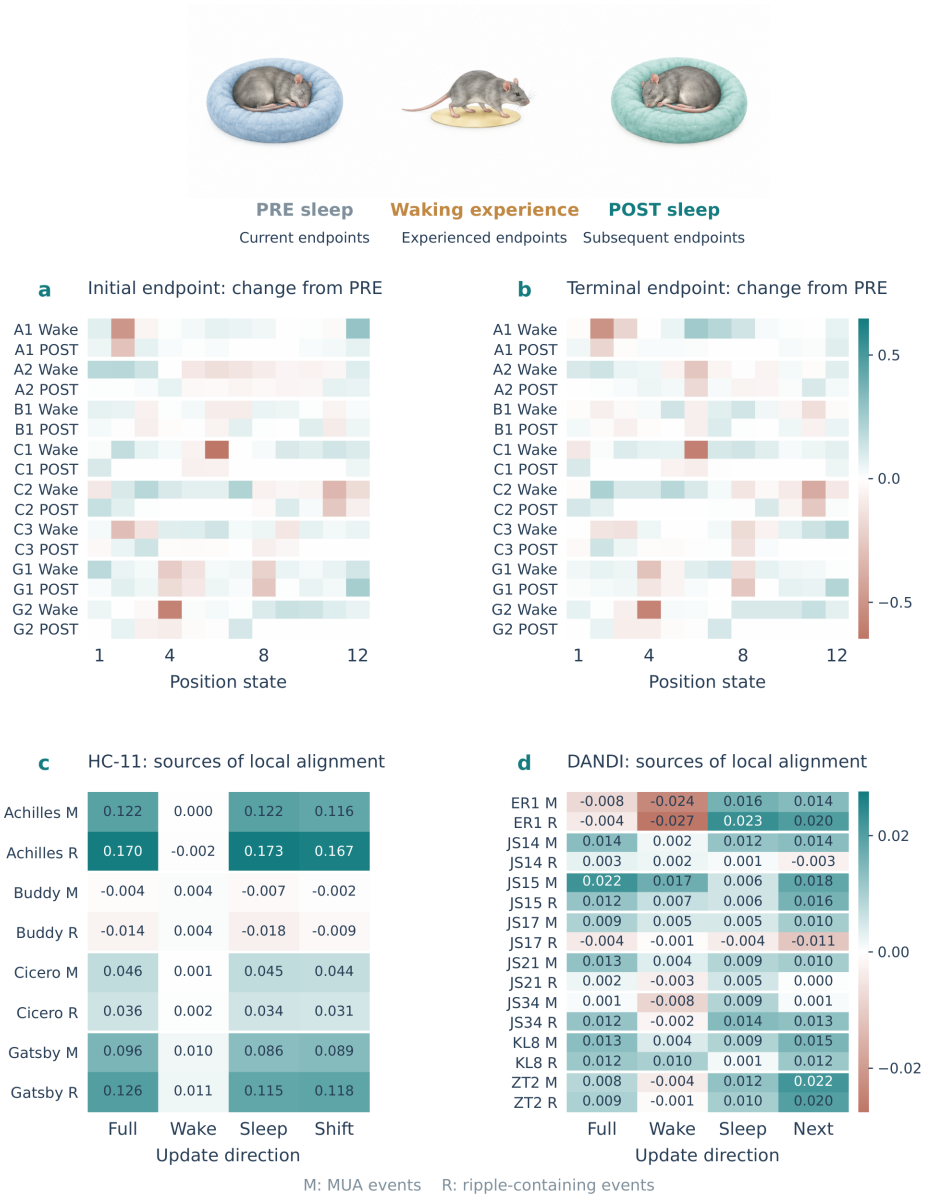

Figure 33: **Endpoint distributions and the animal structure of local learning.** (a,b) Initial and terminal probability changes in all eight HC-11 sessions. Each session has consecutive waking-minus-PRE and POST-minus-PRE rows; both panels share a signed probability scale. Session identifiers follow Figure 32. (c,d) Every animal's local scores under full-residual, waking-bias, current-sleep-deficit, and comparison-target directions. The comparison target is spatially shifted in HC-11 and drawn from the next waking epoch in DANDI. M and R identify MUA and ripple-containing estimates. Printed values use cohort likelihood units, with separate signed color ranges. Spatial distributions use complete-session fits; local scores use temporal cross-fitting. Header artwork illustrates the sleep–experience–sleep task.

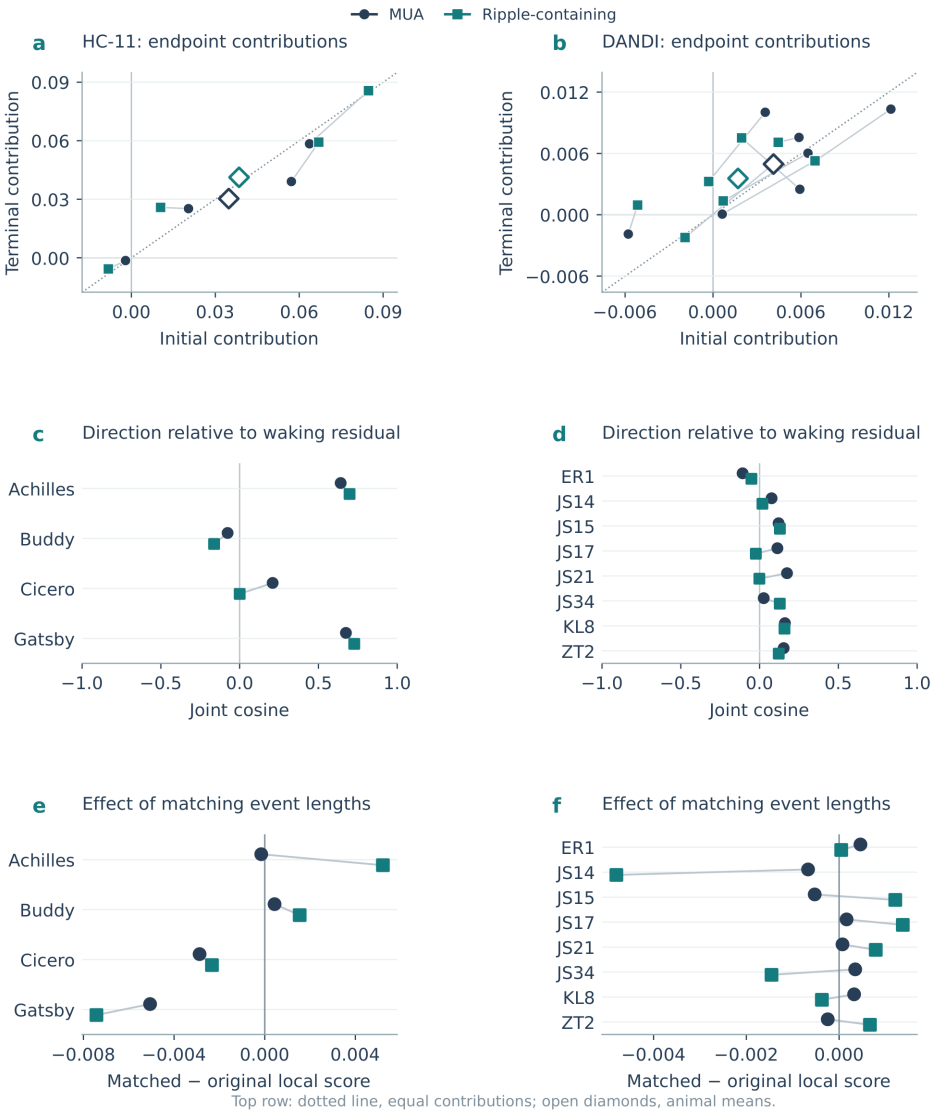

Figure 34: **Endpoint contributions, direction, and event length.** (a,b) Initial versus terminal score contribution for every animal. The two coordinates sum to its local score; the dotted line marks equal contributions and open diamonds give cohort means. (c,d) Joint cosines with animal identifiers on a common dimensionless scale. (e,f) Changes in local score after matching event lengths; zero marks an unchanged estimate. Lines pair event definitions within animals. All four HC-11 and eight DANDI animals contribute. Source tables give the scores before and after matching.

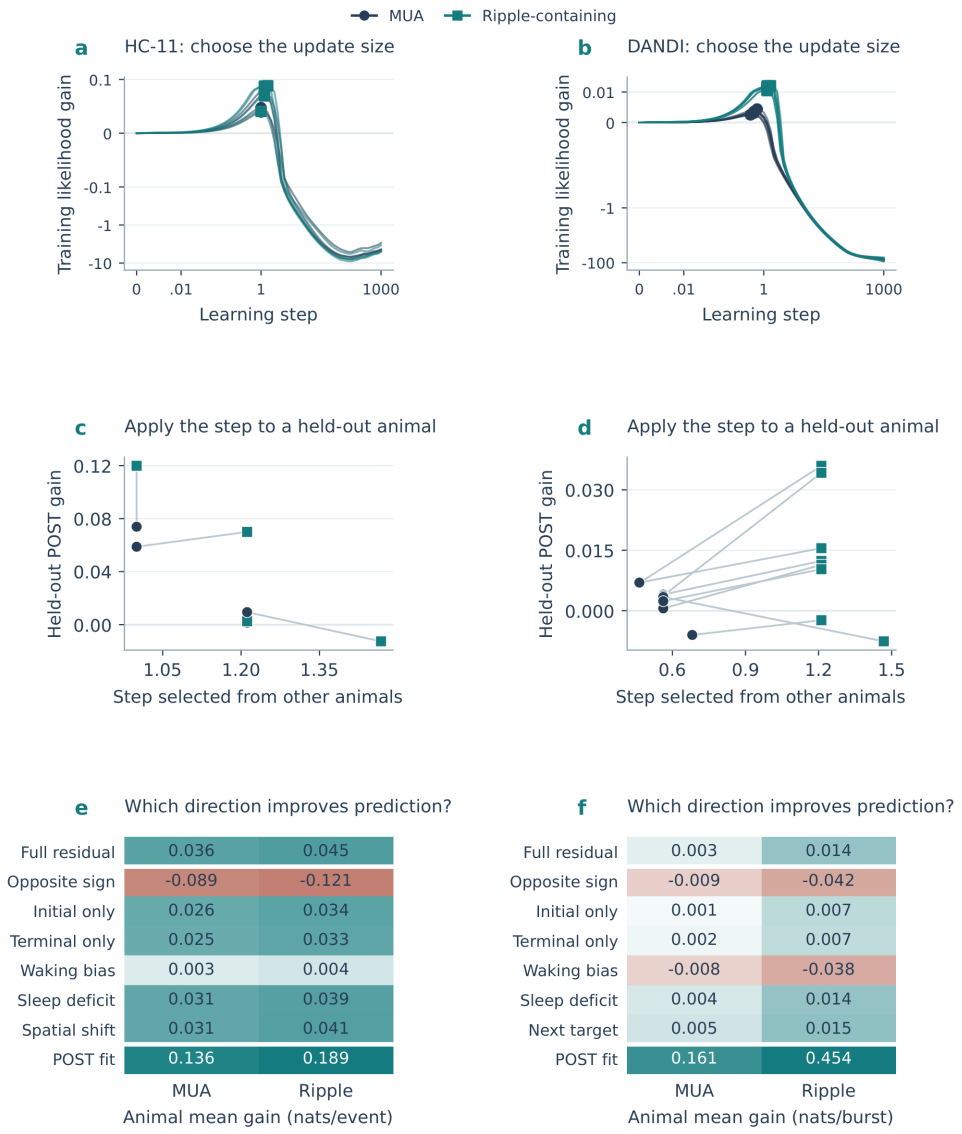

Figure 35: **From the learning-step landscape to held-out prediction.** (a,b) Training gains across all 62 candidate steps for every leave-one-animal-out split, relative to each zero-step objective. Markers locate selected maxima; axes are symmetric logarithmic. (c,d) Selected steps and their held-out POST gains, with event definitions paired within animal. (e,f) Animal-equal POST gains for eight update directions and two event definitions. Signed symmetric-logarithmic colors use cohort-specific ranges and a 0.01-unit linear region. HC-11 uses nats/event and DANDI nats/burst. The POST-fit comparison provides a descriptive reference. Step selection is separate for each event definition.

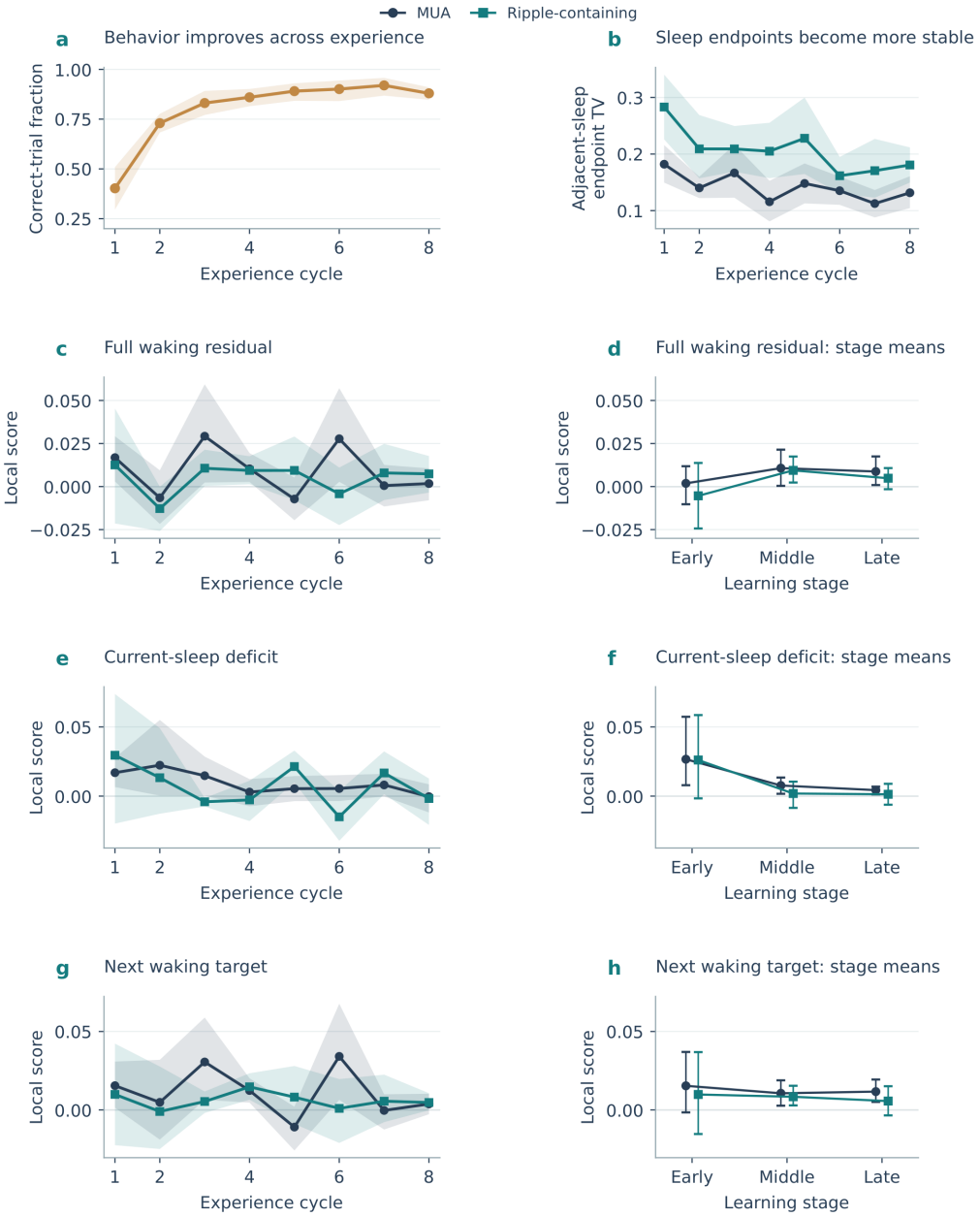

Figure 36: **Behavioral learning and sleep endpoint evolution across experience.** (a,b) Behavioral accuracy and adjacent-sleep endpoint distance across successive learning cycles. (c–h) Local scores for the full waking residual, current-sleep deficit, and next waking target, with cycle trajectories and corresponding stage estimates on common row scales. Early, middle, and late stages comprise cycles 1–2, 3–5, and 6–8. Lines give animal-equal means; bands and error bars give 95% intervals from 10,000 animal bootstrap resamples. The behavioral dataset includes 63 cycles from nine sessions and eight rats. Figure 37c gives contributing neural cycles.

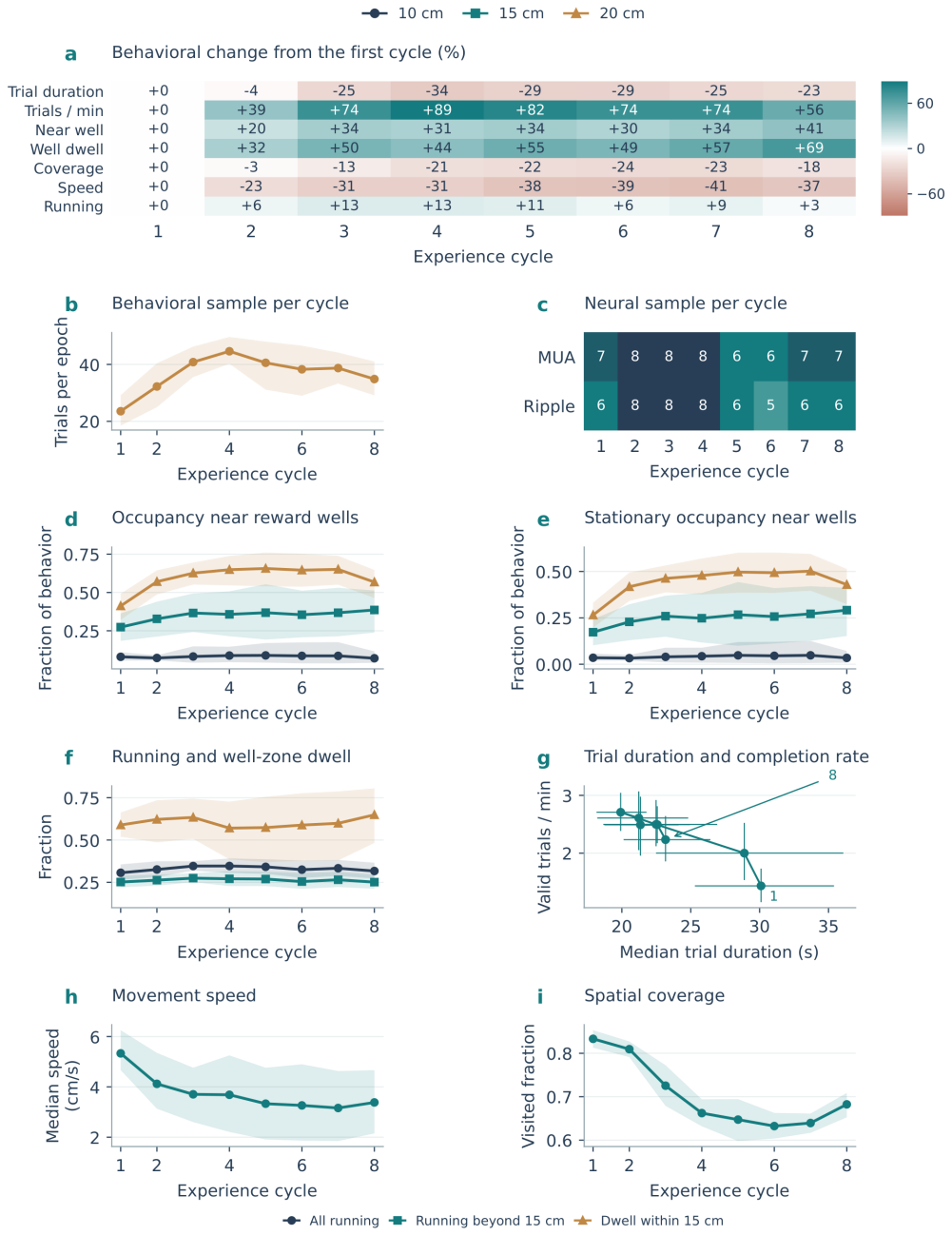

Figure 37: **The behavioral context of endpoint learning.** (a) Percentage change from each measure's first-cycle mean, shown as a seven-measure, eight-cycle matrix with annotated values and signed colors. (b) Waking endpoint trials. (c) Neural session cycles: 57 MUA and 53 ripple-containing. (d,e) Well-zone occupancy and stationary occupancy at three radii. (f) Running and conditional well-zone dwell. (g) Trial efficiency in cycle order, with first and last cycles labeled. (h,i) Speed and spatial coverage. Native-scale curves use animal-equal means and 95% animal-bootstrap intervals; (g) gives separate horizontal and vertical intervals. All 63 behavioral cycles enter the summaries.

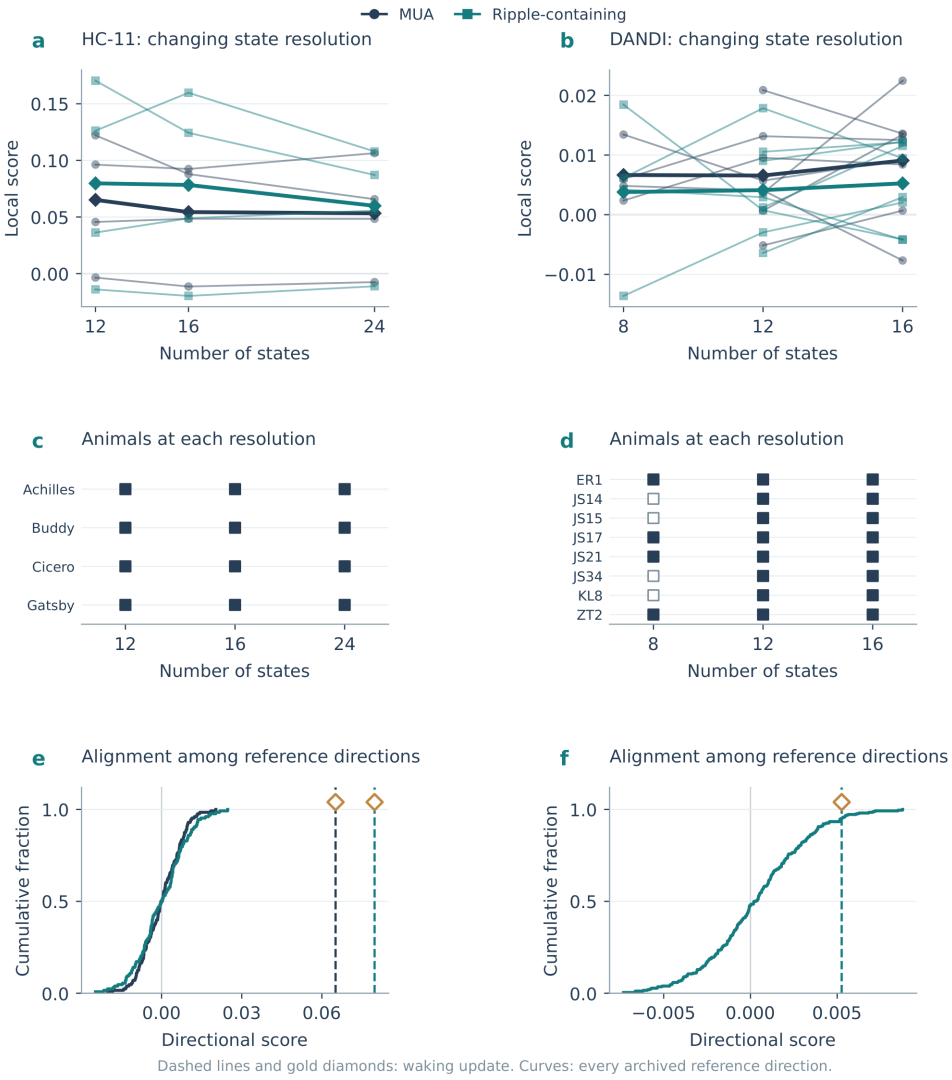

Figure 38: **State resolution and the specificity of the learning direction.** (a,b) Each animal across available state counts, with equal-animal means shown by thick traces and diamonds. (c,d) Estimate availability: filled squares indicate an estimate and open squares its absence. DANDI has four animals at eight prototypes and eight at 12 and 16; HC-11 has four throughout. (e,f) Empirical cumulative distributions of 128 reference-direction scores per HC-11 event definition and 256 ripple-containing DANDI scores. Dashed colored lines and gold diamonds locate the full-residual score. Initial and terminal direction norms are matched separately.

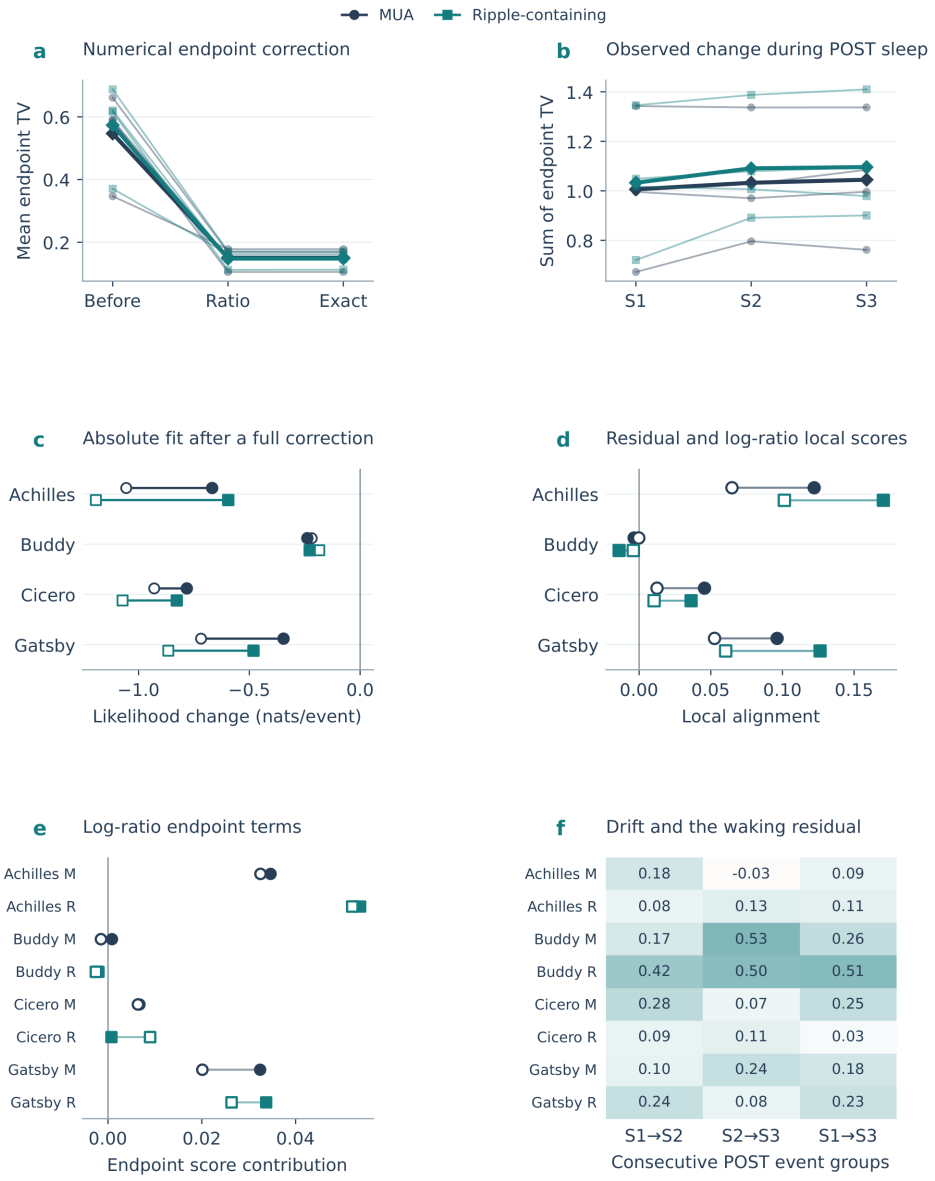

Figure 39: **Numerical endpoint correction and observed sleep evolution.** (a) Endpoint TV before fitting, after proportional correction, and after exact fitting. (b) Distance to waking endpoints across consecutive POST groups. Thin traces show animals; thick traces and diamonds give equal-animal means. (c) Held-out PRE (open) and POST (filled) likelihood changes after one proportional sweep, paired within animal and event definition. (d) Paired residual (filled) and log-ratio (open) local scores on one axis. (e) Initial (filled) and terminal (open) log-ratio contributions. (f) Cosines between observed drift and the waking residual. All four HC-11 animals contribute. Panel (a) averages endpoint TV; (b) sums it.

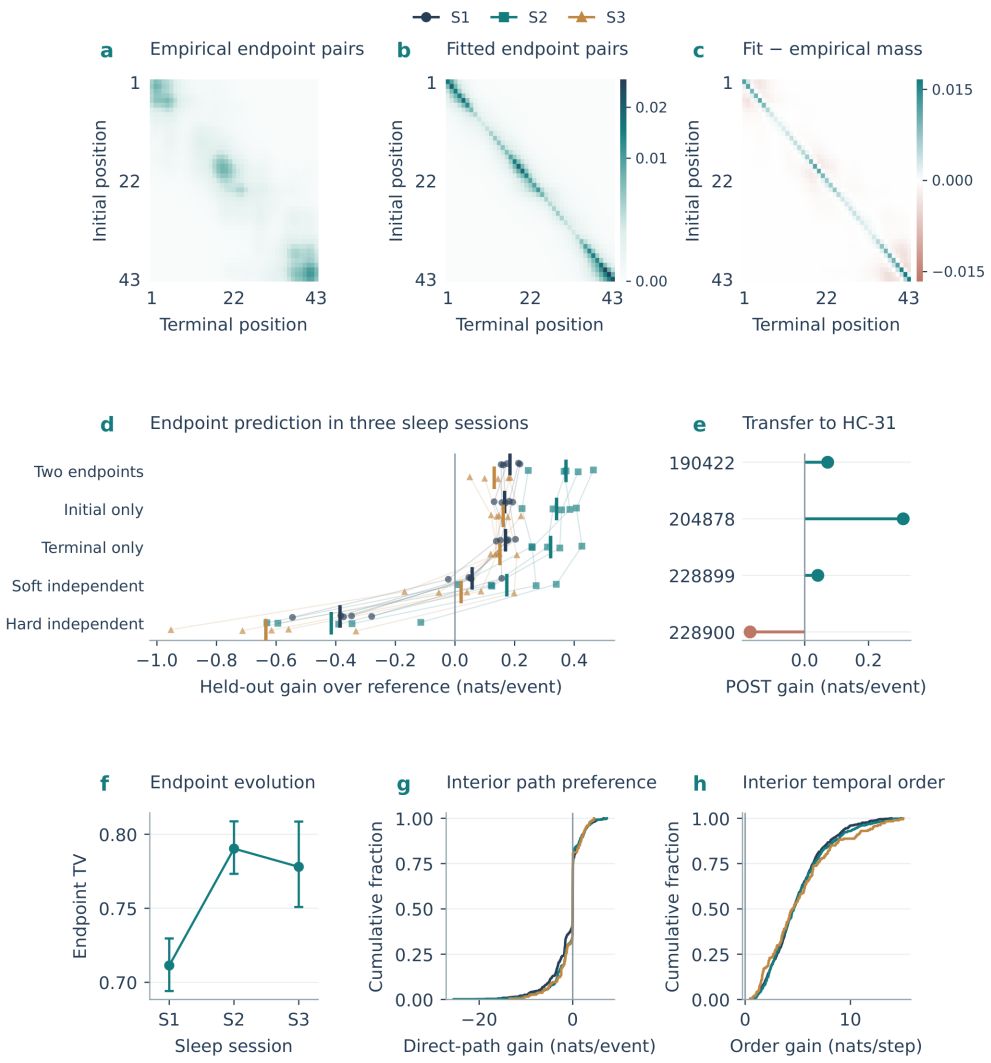

Figure 40: **Endpoint structure, predictive comparisons, and transfer.** (a,b) Empirical and fitted HC-3 S1 endpoint coupling on one square-root probability scale. (c) Fitted minus empirical mass. (d) Gains over reference for all five held-out folds per session; lines follow the same fold across models and vertical ticks give fold means. (e) Four estimable HC-31 animals under the fixed transferred step. (f) HC-3 endpoint TV with 2,400-resample event-bootstrap intervals. (g,h) Empirical cumulative distributions of intermediate-path and interior-order contrasts for 656, 539, and 187 eligible events in S1–S3. HC-3 comprises one rat. Appendix F.10 gives definitions and temporal contrasts.