COMPETITOR INTERACTIONS AND RECALL LIFETIMES IN SPHERICAL LOG-SUM-EXP ASSOCIATIVE MEMORY

Anonymous authors
Paper under double-blind review

ABSTRACT

How long does a retrieved memory persist under noisy updates? We study this question in spherical log-sum-exp associative memory, defining recall loss as the target losing to competitors for a fixed, uninterrupted interval. We test two ways to simplify the dynamics: representing a bank by its target–competitor similarities, and tracking only the overlaps with the target and leading rival. Both omit information relevant to recall. In a controlled three-key example, banks with identical target similarities, initial logits, margin and entropy have mean recall times differing by a factor of approximately 1.74. A numerical solution of the joint dynamics closely reproduces this ratio without fitting a time scale. A separate control shows that the effect persists with distinct keys. In two pairs of 128-key banks, changing the collective drift changes the lifetime ratio even though the exit boundaries remain fixed. Within a bank, replicated short-time tests find different future gaps from states with the same target and leading-rival overlaps and leader. Conditional equilibrium averaging improves the held-out aggregate lifetime ratio over isolated-pair addition, although individual mean times remain systematically biased. These finite-dimensional examples show how interactions among competitors and unresolved state coordinates affect predictions of noisy recall lifetimes.

1 INTRODUCTION

A retrieved memory can be lost under repeated noisy updates, even in a system with ample storage capacity. To predict how long recall lasts, we need to know how competing memories affect the dynamics. One way to describe the competitors is to list their similarities to the target. This records how many competitors there are and how closely each resembles the target, but leaves their mutual similarities unspecified. For a fixed set of stored keys (a memory bank), we call this list the target spectrum.

We study whether this missing geometry matters in spherical log-sum-exp associative memory. The model has an explicit connection to tied key–value softmax attention (Ramsauer et al., 2021), and belongs to a class of dense associative memories with well-studied storage capacity (Krotov & Hopfield, 2016; Demircigil et al., 2017; Lucibello & Mézard, 2024). Starting from the target, we measure the mean time until recall is lost for a prescribed uninterrupted interval. The leading competitor may change before this interval is completed. This observable differs from a stationary escape rate.

The missing information matters in two ways. Across banks, similarities among competitors enter both the drift and the noise covariance of target–competitor gaps. Within a bank, the overlaps with the target and leading rival need not specify the short-time evolution, even when the leader's identity is known.

We test these two omissions directly. Planar banks with identical target similarities have different recall times; their joint operators predict the difference seen in simulation. In 128-key banks, removing the nonleading drift contributions and renormalizing the remaining weights changes the lifetime ratio while leaving the winner and exit rules unchanged. Finally, a local experiment and an independent replication find different future gaps from states with the same two resolved coordinates and leader under the tested discrete dynamics.

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.

We also evaluate a two-coordinate approximation that averages over the other coordinates using the full equilibrium distribution. On four banks held out when its predictions were frozen, it predicts a lifetime ratio of 1.289, compared with 1.349 observed and 2.11 from isolated-pair addition. It nevertheless underestimates the mean in one of the two conditions on all four banks, and only four of eight predefined 10% agreement checks pass. Thus, accounting for interactions improves the ratio without resolving the errors in individual lifetimes. The exact operators and controlled tests explain where the simpler descriptions lose information; the approximation measures how far equilibrium averaging recovers it.

Scope and related work. Capacity theory (Lucibello & Mézard, 2024), spherical thermal-retrieval maps (Petrova, 2026), and free-energy landscapes for binary-spin memories (Sumedha & Singh, 2026) address storage, alignment and equilibrium structure. Here we measure sustained loss after target initialization; equilibrium barriers alone need not determine transition rates (Sicard, 2018). Random-feature representations approximate energies and finite-step retrieval dynamics (Hoover et al., 2024); those bounds do not directly control our noisy confirmation time. Distribution-valued memories use Bures–Wasserstein geometry for Gaussians (Tankala & Balasubramanian, 2026) or Sinkhorn divergences with spherical Hellinger–Kantorovich flow (Mustafi & Mukherjee, 2026). Their retrieval and basin guarantees concern different state spaces and deterministic dynamics. Epanechnikov/LSR energies introduce compact support and additional minima (Hoover et al., 2025); our LSE lifetime results do not establish cross-kernel universality.

Convergence and transient retrieval (Mimura et al., 2026; Clark, 2026) and projection-memory methods (Thiede et al., 2019; Lorpaiboon et al., 2024) provide established tools for dynamical approximation. We use these tools to test which information is needed to predict spherical LSE recall lifetimes. The main controls use $\beta = 1$, $T = 0.2$, $N = 12, 18, 24$ and constructed or finite 128-key banks; they establish neither an exponential-load kinetic limit nor results for trained Transformers.

2 Full dynamics and confirmed recall loss

Let $a_0, \dots, a_{M-1} \in \mathbb{S}^{N-1}$ be unit memory keys and let $y \in \mathbb{S}^{N-1}$ denote the state. The target is a_0. Define

$$m_\mu = a_\mu^\top y, \qquad p_\mu(y) = \frac{e^{\beta N m_\mu}}{\sum_\nu e^{\beta N m_\nu}}, \qquad \bar{a}(y) = \sum_\mu p_\mu(y) a_\mu. \tag{1}$$

At initialization, the logits are $\beta N m_\mu(a_0)$, the overlap margin is $1 - \max_{i>0} m_i(a_0)$, and the entropy is $-\sum_\mu p_\mu(a_0) \log p_\mu(a_0)$. For the continuous dynamics, the generator $\mathcal{L}f(y)$ gives the instantaneous rate of change of the expected value of a smooth function f starting at state y:

$$\mathcal{L}f = \frac{T}{N} \Delta_{\mathbb{S}} f + [(I - yy^\top)\bar{a}(y)]^\top \nabla_{\mathbb{S}} f. \tag{2}$$

Here $\Delta_{\mathbb{S}}$ and $\nabla_{\mathbb{S}}$ are the Laplacian and gradient along the sphere. Its equilibrium surface density is proportional to $[\sum_\mu e^{\beta N m_\mu}]^{1/(\beta T)}$. The first term describes noise along the sphere; the second pulls the state toward the softmax-weighted key average. The sharpness parameter β and noise temperature T are distinct.

Direct simulations use the normalized finite-step recurrence

$$y_{n+1} = \frac{y_n + h\bar{a}(y_n) + \sqrt{2Th/N}\,(I - y_n y_n^\top) g_{n+1}}{\| y_n + h\bar{a}(y_n) + \sqrt{2Th/N}\,(I - y_n y_n^\top) g_{n+1} \|}, \qquad g_{n+1} \overset{\text{iid}}{\sim} \mathcal{N}(0, I). \tag{3}$$

The update applies the full drift $\bar{a}$ before normalization. In the small-step limit, normalization yields Equation (2); finite steps introduce discretization error. The continuum operator predictions below are compared with simulations at the reported steps, with step-size sensitivity reported separately.

We initialize at the target, $y_0 = a_0$. A winner maximizes m_μ with a fixed tie convention. Let $\Omega = \{y : m_0 > \max_{i>0} m_i\}$ be the target region. For the continuous process, confirmation occurs after an uninterrupted outside episode of duration $\Delta = 0.1$. Returning to Ω resets the clock; switching between non-target winners does not. For the sampled process, the corresponding endpoint is a prescribed number k of consecutive non-target-winning observations, with nominal $kh = \Delta$. The

primary observable is $U(a_0)$, where $U(y) = \mathbb{E}_y \tau_c$ for $y \in \Omega$ with the confirmation clock initially reset.

First crossing, confirmation, and the start of the final successful episode are three different observables. In particular, a sampled backdated start satisfies $\tau_r = \tau_c - (k-1)h$ on a completed path. A first-exit principal eigenvalue characterizes quasi-stationary survival, not automatically the inverse mean confirmation time from a_0. We use inverse means only to define the reported mean-time contrasts, without assuming exponentially distributed lifetimes.

3 What target similarities leave out

Write $G_{\mu\nu} = a_\mu^\top a_\nu$ for the full Gram matrix. Applying the spherical generator to functions of all overlaps gives

$$\mathcal{L}f(m) = \frac{T}{N} \sum_{\mu,\nu} (G_{\mu\nu} - m_\mu m_\nu)\, \partial_{\mu\nu} f + \sum_{\mu} \left[(Gp)_\mu - m_\mu (m \cdot p) - \frac{T(N-1)}{N} m_\mu \right] \partial_\mu f. \tag{4}$$

This is an exact representation on the Gram-constrained support. When $M > N$, the overlaps are not independent coordinates. An orthonormal basis of the key span gives a nonredundant projection. At rank N, the projected state lies on the sphere.

To see how competitor geometry enters, consider the gaps $\delta_i = m_0 - m_i$ and target similarities $\phi_i = G_{0i}$. Below, b_i is the mean instantaneous gap velocity; $\langle \delta_i, \delta_j \rangle_t$ denotes quadratic covariation, whose derivative gives the covariance of noise increments per unit time:

$$b_i(m) = \sum_j p_j (G_{0j} - G_{ij}) - \delta_i (m \cdot p) - \frac{T(N-1)}{N} \delta_i,$$
$$\frac{d\langle \delta_i, \delta_j \rangle_t}{dt} = \frac{2T}{N} \left[1 - \phi_i - \phi_j + G_{ij} - \delta_i \delta_j \right]. \tag{5}$$

Every key contributes to a shared drift. Rival–rival similarities also enter the noise covariance even when all target similarities are known.

Proposition 1 (Insufficiency of the target spectrum for gap dynamics)**.** *The target spectrum $(\phi_i)_{i>0}$ and number of keys do not determine the local generator of the target–competitor gaps, even at the target state with identical initial logits. This statement holds for spherical LSE diffusion with fixed $N \geq 3$, $\beta > 0$ and $T > 0$.*

Proof. Set $a_0 = e_1$ and choose $0 < \phi < 1$. Let $a_1 = \phi e_1 + \sqrt{1-\phi^2} e_2$ and $a_2 = \phi e_1 + \sqrt{1-\phi^2}(\cos\theta\, e_2 + \sin\theta\, e_3)$. For every θ, both target similarities are ϕ, and the logits and softmax weights at a_0 are unchanged. However, $G_{12} = \phi^2 + (1-\phi^2)\cos\theta$ varies with θ. Equation (5) therefore gives different off-diagonal gap covariance at a_0. Taking distinct angles strictly between zero and π avoids duplicate keys. Hence the local gap generator is not determined by the target spectrum. □

Different generators can still give the same mean lifetime. We therefore test the lifetime itself in the controlled banks below, using the full-system confirmation problem. Target spectra may remain useful predictors within a restricted distribution of banks.

4 Identical target spectra, different recall times

Consider two rivals at angles $+\vartheta, +\vartheta$ from the target (aligned) or at $+\vartheta, -\vartheta$ (separated), with $\cos\vartheta = 0.7$. Both banks have three keys, the same target spectrum, initial logits, top margin and entropy. Their rival Gram entries are 1 and $2(0.7)^2 - 1 = -0.02$, respectively. In the aligned bank the transverse force at the target is nonzero; in the separated bank it cancels. The aligned case uses duplicate rivals. We also rotate one rival away from the other while preserving both target similarities to test whether duplicates are essential to the effect.

At $N = 18$, direct simulation gives mean confirmation times of 24.74 (aligned) and 43.14 (separated): identical target similarities yield a lifetime ratio of 1.744. The joint operator below predicts 25.12 and 43.59, a ratio of 1.735, without fitting a time scale.

All keys lie in a plane, so the full diffusion closes exactly in the two coordinates of its projection onto that plane: rotations of the unresolved directions leave the projected dynamics unchanged. In polar coordinates (r, ϑ), its density and mobility are

$$w(r,\vartheta) \propto r(1-r^2)^{(N-4)/2}\left[\sum_{\mu} e^{\beta N r\cos(\vartheta-\vartheta_\mu)}\right]^{1/(\beta T)},$$
$$D(r,\vartheta) = \operatorname{diag}\left((1-r^2)/N,\, 1/(Nr^2)\right), \qquad \mathcal{L} = \frac{T}{w}\nabla\cdot(wD\nabla). \tag{6}$$

The spherical Jacobian and curvature are retained. Here the two projected coordinates suffice for exact dynamics. For high-rank banks, the conditional averaging used later is an approximation.

A shared confirmation problem. The calculation must count failed outside excursions followed by a return, as well as the final uninterrupted outside episode. We discretize the common operator with conservative positive face conductances. Let I and O denote cells inside and outside the target region, and let u_I be the mean confirmation time starting in I with the clock reset. Let A_I, A_O be the negatives of the within-region subgenerator blocks, including crossing losses, and R_{IO}, R_{OI} their nonnegative crossing blocks. For the resulting spatial chain with a deterministic outside clock,

$$[A_I - R_{IO}F_\Delta R_{OI}]u_I = \mathbf{1} + R_{IO}F_\Delta\mathbf{1}, \qquad F_\Delta = \int_0^\Delta e^{-tA_O}\,dt. \tag{7}$$

The term $R_{IO}F_\Delta R_{OI}$ accounts for returns before confirmation, while $R_{IO}F_\Delta\mathbf{1}$ accounts for time spent outside up to return or confirmation. Numerically, the fixed delay is approximated by k successive exponential waiting times of mean Δ/k (an Erlang clock). Their total duration concentrates around Δ as k increases. Elimination is exact for the corresponding augmented chain; spatial and clock-phase convergence must be checked separately. Implementation and independent two-key checks appear in Appendix B.

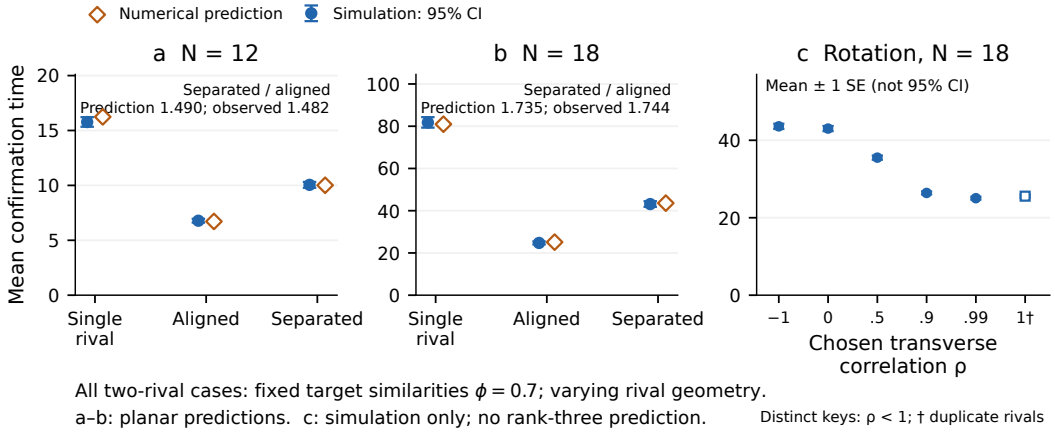


Figure 1: **Target similarities do not determine recall time.** (a,b) Planar banks: simulation means with pointwise 95% bootstrap intervals and shared-operator predictions without rate calibration. (c) A separate rotation control at $N = 18$: both target similarities remain 0.7 while the mutual geometry varies. Interior points have distinct rivals and generally rank three; no planar prediction is assigned to them. Error bars in (c) are one standard error, not 95% intervals; each point uses 4,096 paths. The $\rho = 1$ endpoint has duplicate rivals. Numerical settings and the complete control are in Appendix B.

The planar comparisons in Figure 1(a,b) use 4,096 paths per setting at $h = 0.0025$. At $N = 12$, the separated/aligned ratios are 1.482 observed and 1.490 predicted. This measured lifetime difference extends the local-generator result of Proposition 1. The distinct-rival rotation control in Figure 1c

shows that the effect is not confined to exactly repeated keys: means change from 43.02 to 25.05 between transverse correlations 0 and 0.99, with both rivals distinct. We analyze this simulation cohort separately from panels (a,b). Across the six archived planar settings, five predictions fall within the individual simulation 95% intervals; the additional $N = 12$ one-rival control slightly misses its interval. Independent checks at $h = 0.000625$ and 0.00015625 retain agreement of the ratios within sampling uncertainty: at $N = 18$ they are 1.757 and 1.751, against 1.735 predicted. Appendix B reports both absolute means and ratios at all four steps; these checks do not establish a continuum error bound.

5 Nonleading competitors change the common force

The isolated-pair baseline computes a mean confirmation time with each retained rival alone, sums the inverse means, and inverts that sum to predict a lifetime. Each pair is a different physical system. Adding these inverse means assumes that the separate solutions represent the channels of the original common force field. Exact addition of destination fluxes under a quasi-stationary law does not justify this step: those fluxes share one full-system density and eigenfunction. Target-start confirmation additionally involves returns and the common outside clock.

We now test a different omission: the force from nonleading competitors. Within each bank we change this force while preserving every recall-loss boundary. The test uses $N = 24$ dimensions. Each numbered bank realization has two conditions, A and B, each with the target and 127 rivals. In A, four rivals have target similarity 0.7 and the remaining 123 have similarity 0. In B, one has similarity 0.7 and the other 126 have similarity 0.52717. These conditions match the initial margin and softmax entropy. Their target spectra differ, whereas Section 4 holds the spectrum fixed. Appendix A specifies the random orientations and full precision. Using finite-N pair kinetics gives an approximate mean-time contrast near 2.11 for this 128-key construction. Throughout this comparison, the contrast is U_B/U_A: how much longer B retains the target than A. This approximate baseline neglects the 123 orthogonal rivals in A (Appendix C); its matched-cohort comparison appears in Section 7.

Our force intervention retains all 128 keys in the winner and confirmation rules, preserves the noise, and allows the leading rival to change. At each state, it recomputes the softmax drift using only the target and current leading rival. This changes both the vector contribution of the other keys and the normalization of the retained weights.

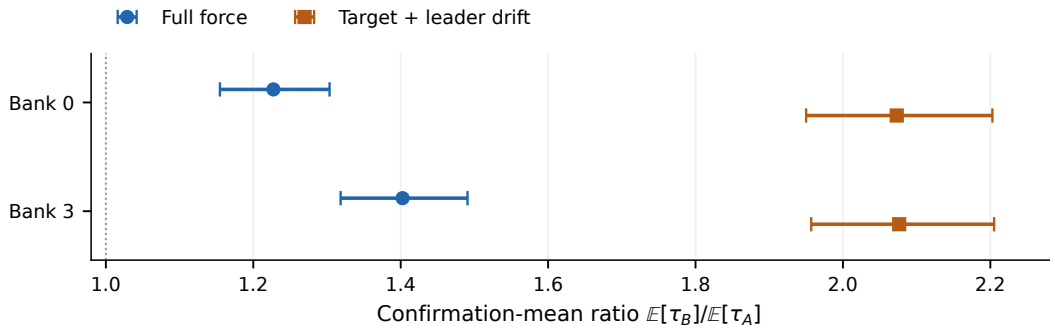


Figure 2: **Changing the drift with an unchanged recall-loss rule.** Full and target–leader-only drift are compared in two A/B bank pairs, labeled 0 and 3. A contains four strong rivals; B contains one strong rival and 126 moderate rivals. All keys remain in the winner and confirmation rules. Bars are pointwise 95% bootstrap intervals. Ratios are $\mathbb{E}\tau_B/\mathbb{E}\tau_A$. Removing nonleading force contributions while renormalizing the retained weights moves the contrast toward the isolated-pair baseline, predominantly by lengthening B. Force direction and weight normalization change together, so their effects remain confounded. The comparison covers only these two bank pairs.

The force comparison uses 2,048 paths per condition and bank at $h = 0.0025$. In bank 0, the full A/B means are 139.68/171.43, compared with 140.13/290.49 under target–leader drift. In bank 3, they are 124.11/174.07, compared with 138.12/286.79. Accordingly, the ratios $\mathbb{E}\tau_B/\mathbb{E}\tau_A$ change

from 1.227 and 1.403 to 2.073 and 2.076 (Figure 2). Nonleading competitors therefore have a large kinetic effect in these banks, especially in condition B. A common multiplier of the predicted times would cancel from these ratios, so it cannot account for the change caused by the force intervention.

6 Equal resolved coordinates do not fix the future

We now fix the bank and ask what information about its evolving state we must retain. Let ℓ be the leading rival and define

$$d = m_0 - m_\ell, \qquad s = (m_0 + m_\ell)/2, \qquad \omega = \sum_{j \notin \{0,\ell\}} p_j. \tag{8}$$

Here d is the target's lead over its strongest rival, s is their average overlap, and ω is the total softmax weight of all other keys. We ask whether states with the same (d, s) and leader identity ℓ have the same distribution of future target gaps. The test holds (d, s, ℓ) fixed and separates states by their background weight ω.

On bank 4B, we fix $d = 0.005$, $s = 0.8$ and the same strong leader. We sample full states with these exact overlaps, retain those with the specified leader, and weight them by the full Gibbs density. We then split the distribution at the weighted median of ω and simulate the lower- and higher-weight halves for time 0.1 at $h = 0.000625$. No lifetime residual enters this grouping. We fixed the primary settings after a separate exploratory bank-0 test. We already knew bank 4's lifetime outcomes, so we interpret the result as a controlled local test rather than held-out lifetime validation.

With 32,768 paths per half, the mean initial background weights are 0.101 and 0.199 in the low- and high-weight groups; their mean future gaps are 0.006292 and 0.004626, respectively. The future gap uses the strongest rival at the measurement time, which may differ from the initial leader. Their high-minus-low difference is -0.001666, with conditional 95% interval $[-0.002203, -0.001130]$. A replication using independent quadrature samples and independent noise, with 16,384 paths per half, gives -0.002163 $[-0.002916, -0.001410]$. For a rotational-symmetry null, a two-key system is instead split at the median of a remaining state coordinate; its background weight is zero. It gives -0.000401 $[-0.000947, 0.000145]$. Details of the conditional intervals and archived dependent replication appear in Appendix D.

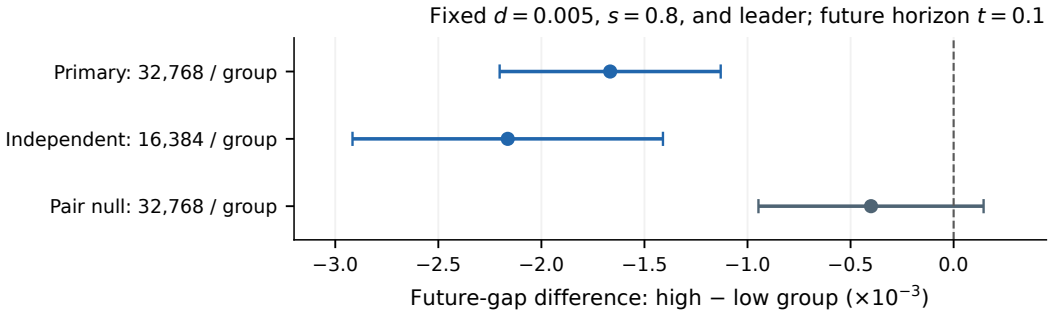


Figure 3: **Hidden-state dependence at identical coordinates and leader.** Differences in the mean future target gap at time 0.1 between high and low groups. The full-memory tests split by background weight; the two-key null splits by a remaining state coordinate, since it has zero background weight. The primary fixed-coordinate test and independent replication exclude zero; the two-key rotational null does not. Pointwise 95% normal-approximation intervals condition on finite initial quadrature. The experiment tests whether the matched coordinates suffice for local prediction; a lifetime correction and the sufficiency of background mass alone remain open.

The two groups start at identical (d, s, ℓ) but have different mean future gaps under the tested discrete update. These coordinates therefore fail to specify a common future-gap law for the tested initial distributions. Since the coordinates match exactly, finite bin widths cannot explain the difference. The split also selects other hidden features, including background-force direction, so it does not establish that adding ω would suffice or quantify lifetime bias. The intervals remain conditional on finite quadrature, whose error is not bounded by replication.

7 AN INTERACTION-AWARE APPROXIMATION AND ITS LIMITS

We next ask how useful a two-coordinate approximation can be despite this loss of state information. It accounts for all competitors by replacing the unresolved coordinates at each (d, s) with their conditional equilibrium distribution under the full model. This discards any dependence of their distribution on the path taken to reach (d, s). Let π be that equilibrium law, $\rho_G(d, s)$ its marginal density and $\phi_\ell = a_0^\top a_\ell$ the current leader's target similarity. The reduced state is (d, s); conditional averaging includes all leader identities. The approximate generator is

$$\overline{\mathcal{L}} f = \frac{T}{\rho_G} \nabla \cdot (\rho_G \overline{D} \nabla f), \quad \overline{D} = \frac{1}{N} \mathbb{E}_\pi \left[\begin{pmatrix} 2(1-\phi_\ell) - d^2 & -ds \\ -ds & (1+\phi_\ell)/2 - s^2 \end{pmatrix} \middle| d, s \right]. \tag{9}$$

We solve the return/confirmation problem for this operator. The weak formulation restricts the full Dirichlet form to functions of (d, s), combining all leader regions before discretization (Appendix C). Equilibrium quadrature determines the coefficients without fitting to lifetime observations. Unlike the exact planar projection, this is an approximate reduced model. Section 6 tests local predictive sufficiency for the discrete update at the stated step size; it does not prove a corresponding property of the continuous diffusion.

We froze predictions and the validation protocol before simulating full trajectories from four new banks. Each bank has two conditions with 2,048 paths each at $h = 0.0025$. The simulation horizon was 4,000 time units; all paths reached confirmation, so all eight samples have zero censoring. Figure 4 shows relative discrepancies for all eight A/B means. The aggregate contrast is the geometric mean of the four bankwise B/A ratios; its interval uses a hierarchical bank/trajectory bootstrap. The predicted aggregate contrast is 1.289, versus 1.349 $[1.276, 1.426]$ observed and 2.11 from isolated-pair addition. We use observed/predicted minus one throughout, so a positive discrepancy indicates underprediction. Across the eight means, this ranges from -2.22% to $+7.51\%$; for all four B conditions it is $+5.77$–$+7.51\%$. The individual means show the remaining error. Only three of eight predictions fall inside the simulation intervals, and four of eight pass the predefined 10% agreement check: the entire pointwise 95% interval must lie within $[0.9, 1.1]$ times the fixed prediction. We assess agreement separately for each condition, without simultaneous coverage. Containing the prediction and meeting the 10% agreement criterion are distinct tests.

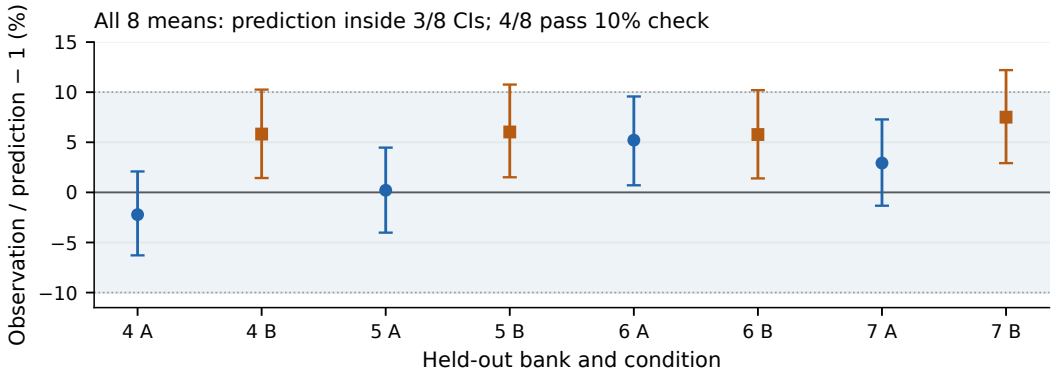


Figure 4: **Quantitative improvement with a persistent residual.** The frozen interaction-aware closure is evaluated on four banks held out at the time of prediction. All eight means are retained. The displayed relative discrepancy is observed/predicted minus one; positive values indicate underprediction. The aggregate contrast is much closer to observation than isolated-pair addition, while every predicted B mean remains below its observed mean and only four of eight 10% agreement checks pass: the entire pointwise 95% interval lies inside $[0.9, 1.1]$ times the fixed prediction. Subsequent diagnostics do not replace these primary predictions.

Separating numerical sensitivity from missing state. The lifetime discrepancies cannot yet be separated into numerical error and error from discarding state information. On bank 4B, clock refinement and tested starting-point evaluations have small effects, but spatial and quadrature sensitivity persists. A trajectory-based check of the predicted remaining lifetime also changes under

mesh refinement; its residual is no longer statistically resolved in the independent finer-mesh repeat. Coupled simulations detect a change at the first time-step reduction but do not resolve one at the next. Appendix D reports these diagnostics, including unstable refinements. They do not revise the frozen predictions or bound error across all eight cases.

Memory from unresolved coordinates. Eliminating unresolved coordinates can make future evolution depend on the past, so its effect cannot generally be captured by multiplying all predicted times by one constant. For confirmation times, elimination can also change the accumulated-time term and the reconstruction from the initial state. The standard operator identity in Appendix D expresses this dependence. Identifying the relevant missing coordinates and the sign of the confirmation-time bias requires further analysis.

8 DISCUSSION

The planar examples answer the question posed at the start: target similarities alone do not determine recall lifetimes. Similarities among rivals affect the joint dynamics, and solving those dynamics reproduces the measured lifetime difference. The force intervention shows that nonleading competitors also matter in the tested 128-key banks. These results support treating loss and return as one interacting problem rather than adding solutions for isolated pairs. Competitors can switch along a path, and no quasi-stationary initialization is required.

Choosing a reduced state poses a further problem. In the local test, identical target and leading-rival overlaps, with the same leader, admit different mean future gaps. Equilibrium averaging accounts for all keys and improves the aggregate lifetime ratio, but it discards how the unresolved coordinates depend on the preceding path. Such dependence can produce memory: a Markov description requires the future, conditional on the current resolved state, to be independent of the past. The local test identifies lost predictive information. Its contribution to the lifetime bias remains unresolved alongside numerical error, and the later reconstruction tests have yet to yield a reliable correction (Appendix D).

These conclusions concern finite, controlled systems. The force test covers two A/B bank pairs, and the exact-coordinate test covers one local fiber with replication. The four banks used for quantitative validation were held out when predictions were frozen, then used for later development. The experiments do not establish exponential-load asymptotics, uniform accuracy across parameters, or computational superiority over direct simulation. They provide concrete tests for a proposed reduction: whether it distinguishes banks with different competitor geometry, and whether states mapped to the same coordinates have the future dynamics needed to predict confirmation times.

AI USE STATEMENT

Generative AI tools assisted the refinement of theoretical models, and mathematical proofs, implementation code, literature review, and manuscript proofreading. Reported numerical observations come from recorded simulations. Mathematical derivations, numerical outputs, and references were checked against the calculations, archived records, and cited sources described in this work. The authors are responsible for the final scientific claims, code, references, and manuscript.

ETHICS STATEMENT

The experiments use synthetic stored patterns. There are no human-subject experiments or personal data. Conclusions concern the specified associative-memory dynamics and do not establish performance of trained Transformer systems.

REPRODUCIBILITY STATEMENT

The appendices specify the dynamics, confirmation rule, numerical approximations, experimental protocols, and statistical procedures. The supplementary package provides simulation and analysis code, per-trajectory endpoint records, local-test observations, saved continuation fields, frozen predictions, and figure-generation scripts. A version-pinned environment and documented commands reproduce the main empirical summaries, figures, and residual analyses from these records; separate launchers support fresh experiments. A result-to-data map and file checksums document the provenance of the supplied materials. The package records the checks performed, numerical reproduction differences, and the additional quadrature inputs required for re-solving operators on the original draws. Simulation uncertainty and numerical approximation error are reported separately.

REFERENCES

David G. Clark. Transient dynamics of associative memory models. *Physical Review E*, 113(5): 054301, 2026. doi: 10.1103/42y2-bsh1. URL https://doi.org/10.1103/42y2-bsh1.

Mete Demircigil, Judith Heusel, Matthias Löwe, Sven Upgang, and Franck Vermet. On a model of associative memory with huge storage capacity. *Journal of Statistical Physics*, 168:288–299, 2017. doi: 10.1007/s10955-017-1806-y. URL https://link.springer.com/article/10.1007/s10955-017-1806-y.

Benjamin Hoover, Duen Horng Chau, Hendrik Strobelt, Parikshit Ram, and Dmitry Krotov. Dense associative memory through the lens of random features. In *Advances in Neural Information Processing Systems*, volume 37, pp. 23549–23576, 2024. doi: 10.52202/079017-0742. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/29ff36c8fbed10819b2e50267862a52a-Abstract-Conference.html.

Benjamin Hoover, Zhaoyang Shi, Krishnakumar Balasubramanian, Dmitry Krotov, and Parikshit Ram. Dense associative memory with Epanechnikov energy. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-3947. URL https://proceedings.neurips.cc/paper_files/paper/2025/hash/ab989891ac2a4b35ee76d80139cec5fb-Abstract-Conference.html.

Dmitry Krotov and John J. Hopfield. Dense associative memory for pattern recognition. In *Advances in Neural Information Processing Systems*, volume 29, 2016. URL https://papers.nips.cc/paper_files/paper/2016/hash/eaae339c4d89fc102edd9dbdb6a28915-Abstract.html.

Chatipat Lorpaiboon, Spencer C. Guo, John Strahan, Jonathan Weare, and Aaron R. Dinner. Accurate estimates of dynamical statistics using memory. *The Journal of Chemical Physics*, 160(8): 084108, 2024. doi: 10.1063/5.0187145. URL https://doi.org/10.1063/5.0187145.

Carlo Lucibello and Marc Mézard. Exponential capacity of dense associative memories. *Physical Review Letters*, 132(7):077301, 2024. doi: 10.1103/PhysRevLett.132.077301. URL https://doi.org/10.1103/PhysRevLett.132.077301.

Kazushi Mimura, Junichi Takeuchi, Yuto Sumikawa, Yoshiyuki Kabashima, and Anthony C. C. Coolen. Dynamical properties of dense associative memory. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/21a57aa30ccfec4e2b12ac122b1da7b3-Abstract-Conference.html.

Aratrika Mustafi and Soumya Mukherjee. Sinkhorn based associative memory retrieval using spherical Hellinger Kantorovich dynamics. *arXiv preprint arXiv:2603.20656*, 2026. doi: 10.48550/arXiv.2603.20656. URL https://arxiv.org/abs/2603.20656.

Frank Noé and Feliks Nüske. A variational approach to modeling slow processes in stochastic dynamical systems. *Multiscale Modeling & Simulation*, 11(2):635–655, 2013. doi: 10.1137/110858616. URL https://doi.org/10.1137/110858616.

Tatiana Petrova. Thermal robustness of retrieval in dense associative memories: LSE vs LSR kernels. *arXiv preprint arXiv:2603.13350*, 2026. doi: 10.48550/arXiv.2603.13350. URL https://arxiv.org/abs/2603.13350.

Hubert Ramsauer, Bernhard Schäfl, Johannes Lehner, Philipp Seidl, Michael Widrich, Thomas Adler, Lukas Gruber, Markus Holzleitner, Milena Pavlović, Geir Kjetil Sandve, Victor Greiff, David Kreil, Michael Kopp, Günter Klambauer, Johannes Brandstetter, and Sepp Hochreiter. Hopfield networks is all you need. *arXiv preprint arXiv:2008.02217v3*, 2021. URL https://arxiv.org/abs/2008.02217v3.

François Sicard. Computing transition rates for rare events: When Kramers theory meets the free-energy landscape. *Physical Review E*, 98(5):052408, 2018. doi: 10.1103/PhysRevE.98.052408. URL https://doi.org/10.1103/PhysRevE.98.052408.

Sumedha and Abhishek Singh. Free energy landscape of dense associative memory. *arXiv preprint arXiv:2607.19195*, 2026. doi: 10.48550/arXiv.2607.19195. URL https://arxiv.org/abs/2607.19195.

Chandan Tankala and Krishnakumar Balasubramanian. Dense associative memory for Gaussian distributions. *arXiv preprint arXiv:2509.23162v2*, 2026. doi: 10.48550/arXiv.2509.23162. URL https://arxiv.org/abs/2509.23162v2. Version 2, revised 31 January 2026.

Erik H. Thiede, Dimitrios Giannakis, Aaron R. Dinner, and Jonathan Weare. Galerkin approximation of dynamical quantities using trajectory data. *The Journal of Chemical Physics*, 150(24):244111, 2019. doi: 10.1063/1.5063730. URL https://doi.org/10.1063/1.5063730.

A EXPERIMENTAL DESIGN AND STATISTICAL CONVENTIONS

Lifetime controls use $\beta = 1$, $T = 0.2$, unit keys, target initialization and Equation (3), projecting noise but not drift before normalization. Ties retain the target; rival switches do not reset the clock. Confirmation requires 40 consecutive outside observations at $h = 0.0025$ or 160 at $h = 0.000625$. We report lifetimes at confirmation, after the required run of outside observations; the background-fiber test measures a future gap.

128-key banks. At $N = 24$, take $a_0 = e_1$. A key of prescribed target similarity ϕ is $a = \phi e_1 + \sqrt{1-\phi^2}\, v$, where v is a normalized independent Gaussian vector in $e_1^{\perp}$. Condition A contains four rivals at $\phi = 0.7$ and 123 at $\phi = 0$; B contains one of the same strong rivals and 126 at $\phi = 0.5271703466250011$. These choices match the initial top margin and softmax entropy. Thus, A and B differ in their competitor populations. Within each condition, the force ablation retains all keys and replaces only the drift with the renormalized target–current-leader softmax mean.

Each numbered realization specifies an A/B pair of banks. Development realizations 0–3 precede test realizations 4–7, which use NumPy bank seeds 931005–931008. We froze predictions and the protocol before simulating validation trajectories; later diagnostics are post-validation. Validation uses 2,048 paths per condition, $h = 0.0025$ and horizon 4,000; all 16,384 paths reach confirmation. Simulation seeds are $2100001 + 20000\, b + 10000\, j$, for realization b and condition $j = 0, 1$. The force comparison uses realizations 0 and 3, with 2,048 paths per condition and force law. These constructed banks are not an iid exponential-load ensemble.

Intervals and estimands. Main lifetime intervals are pointwise percentile intervals from 10,000 within-setting trajectory-bootstrap draws. Force-ratio intervals resample the A and B trajectories. For the four-realization test, the contrast is $\exp[\frac{1}{4}\sum_{b=4}^{7} \log(\overline{\tau}_{bB}/\overline{\tau}_{bA})]$; the hierarchical bootstrap resamples realizations and their trajectory means. This weighting gives each realization equal weight. Four realizations give limited information about between-realization variability. The predefined 10% agreement check requires a complete pointwise 95% interval inside $[0.9\widehat{U}, 1.1\widehat{U}]$ for fixed prediction $\widehat{U}$. It differs from interval coverage, 90%-interval TOST, and simultaneous equivalence. Numerical predictions carry no certified error intervals.

B LOW-RANK OPERATOR, CLOCK AND GEOMETRY CONTROLS

For keys in a rank-r subspace, write their coordinates as b_μ and let z be the orthogonal projection of y into that subspace. Rotational invariance gives the exact projected generator

$$\mathcal{L}f = \frac{T}{N}(I - zz^{\top}) : \nabla^2 f + \left[\sum_{\mu} p_\mu b_\mu - z\sum_{\mu} p_\mu b_\mu^{\top} z - \frac{T(N-1)}{N} z\right] \cdot \nabla f.$$

For $r < N$ its unit-ball density is proportional to $(1 - \|z\|^2)^{(N-r-2)/2}[\sum_\mu e^{\beta N b_\mu^{\top} z}]^{1/(\beta T)}$. At full rank, the support is the sphere. In the planar formula (6), the divergence is with respect to coordinates $dr\, d\vartheta$; the Jacobian r is already in w.

Common clock. Conservative finite-volume discretization uses shared positive face conductances. Let A_I, A_O be the negatives of the within-region subgenerator blocks, including crossing losses, and R_{IO}, R_{OI} contain crossing rates. A k-phase Erlang clock of rate $r_k = k/\Delta$ gives

$$\begin{aligned} A_I u_I &= \mathbf{1} + R_{IO} u_1, \\ (A_O + r_k I) u_j &= \mathbf{1} + R_{OI} u_I + r_k u_{j+1}, \qquad u_{k+1} = 0. \end{aligned}$$

Eliminating phases yields Equation (7) with $F_k = \sum_{j=0}^{k-1} r_k^j (A_O + r_k I)^{-j-1}$ in place of F_Δ. For fixed spatial discretization $F_k \to \int_0^\Delta e^{-tA_O}\, dt$. Elimination exactly matches augmented-chain assembly; finite grids and clock phases still approximate diffusion with deterministic confirmation.

Planar grids reach 120×960 radial/angular cells with 128 phases; 256-phase checks use 90×720. The largest final spatial-grid change across six settings is 0.461%; the remaining error has no

certified bound. An independent orthonormal two-key chart checks the planar special case. With drift removed, $v(y) = (a_0 - a_1)^\top y/\|a_0 - a_1\|$ on the hemisphere $v > 0$, killed at $v = 0$, satisfies $-\mathcal{L}v = T(N-1)v/N$. Direct planar simulation samples the unresolved Gaussian norm using a χ^2_{N-3} variable, giving the same normalized discrete kernel as the full-coordinate recurrence. Each setting in Figure 1(a,b) uses 4,096 paths at $h = 0.0025$. We report all six settings, including the $N = 12$ one-rival prediction that falls just outside its pointwise interval.

Step-size checks. Independent, uncensored cohorts of 4,096 paths per condition, with nominal confirmation duration $\Delta = 0.1$, extend the planar audit to $h = 0.000625$ and 0.00015625. The complete comparison is:

N	h	Aligned mean	Separated mean	Separated/aligned [95% CI]
12	0.01	6.763	9.823	1.452 [1.399,1.506]
12	0.0025	6.781	10.049	1.482 [1.427,1.539]
12	0.000625	6.804	9.923	1.458 [1.405,1.513]
12	0.00015625	6.605	10.080	1.526 [1.469,1.584]
18	0.01	25.563	42.792	1.674 [1.607,1.744]
18	0.0025	24.738	43.136	1.744 [1.673,1.818]
18	0.000625	25.328	44.503	1.757 [1.687,1.832]
18	0.00015625	25.532	44.708	1.751 [1.681,1.825]

Intervals use 10,000 within-condition trajectory-bootstrap draws. The operator ratios 1.490 and 1.735 are compatible with these intervals; this is not a certified continuum limit or an equivalence result.

Distinct-rival control. At $N = 18$, fix $a_0 = e_1$, $a_1 = 0.7e_1 + \sqrt{0.51}e_2$ and $a_2 = 0.7e_1 + \sqrt{0.51}(\rho e_2 + \sqrt{1-\rho^2}e_3)$. Then $G_{12} = 0.49 + 0.51\rho$ while both target similarities remain fixed. Intermediate configurations have distinct rivals and generally rank three; no planar prediction is assigned. Each setting uses 4,096 full-coordinate paths at $h = 0.0025$ with horizon 1,200 and no censoring. Seeds are $1500001 + 10000j$ in the listed order. Table 1 reports one standard error for each mean. We analyze this cohort separately from the planar cohorts in Figure 1(a,b).

Table 1: Rotation at fixed target spectrum. The interior points use distinct keys.

ρ	G_{12}	Mean confirmation time	Standard error
-1	-0.0200	43.599	0.678
0	0.4900	43.023	0.669
0.5	0.7450	35.515	0.534
0.9	0.9490	26.398	0.384
0.99	0.9949	25.048	0.367
1	1.0000	25.571	0.378

C HIGH-RANK APPROXIMATION AND ALL HELD-OUT OUTCOMES

Let $\rho_G(d, s)$ denote the full Gibbs marginal of the coordinates in Equation (8). Equation (9) gives the conditional-equilibrium operator. Here $\phi_\ell = a_0^\top a_\ell$ and ℓ is the current leading rival. Writing $r(y) = \max_{j>0} m_j(y)$, the projection is $\xi(y) = (m_0 - r, (m_0 + r)/2)$. Leader identity is not retained. If ρ_ℓ is the pushforward density of the restriction of π to leader region ℓ (without renormalizing its mass), then

$$\rho_G = \sum_\ell \rho_\ell, \qquad \rho_G \overline{D} = \sum_\ell \rho_\ell \mathbb{E}_\pi[D_\ell \mid \xi, \ell], \qquad \mathcal{E}(f, g) = T \int \rho_G \nabla f^\top \overline{D} \nabla g \, dd \, ds,$$

where D_ℓ is the instantaneous tensor inside the expectation in Equation (9), including $1/N$. At leader ties, ξ is continuous but its derivative changes. Restricting the full Dirichlet form to $f \circ \xi$ combines the leader regions in a single weak formulation. Numerically, all leader regions contribute to each coordinate bin before assembly. Symmetric conductances $c_{ij} = c_{ji} \geq 0$ give probability

current $c_{ij}(P_i/M_i - P_j/M_j)$ for bin probabilities P_i and Gibbs masses M_i. They conserve chain probability without recovering individual full-system inter-leader fluxes. The continuum closure restricts the full first-exit Dirichlet form to functions of (d, s), yielding a Ritz upper bound on the principal killed eigenvalue (Noé & Nüske, 2013). Neither its numerical discretization nor target-start confirmation means inherit that bound.

Importance quadrature mixes normalized von Mises–Fisher proposals centered on the target, all keys and all target–rival midpoints. The frozen prediction uses 2,097,152 samples per bank/condition, mesh 0.005, a nonnegative wide-stencil mobility decomposition (maximum direction 3), 128 Erlang phases and a reflecting cutoff outside the target region at $d = -0.25$. Mesh 0.01 and cutoff -0.15 are separate checks. No rate multiplier is fitted to lifetime observations. At $N = 18$, the two-key quadrature control gives 73.66, 79.29, 81.36, 84.50 at meshes 0.02, 0.01, 0.005, 0.0025 with unchanged sample count, against independent reference 80.97. Finer bins at a fixed sample budget do not ensure convergence.

Table 2: All frozen predictions and observed confirmation means. Each row has 2,048 paths, zero censoring and a pointwise 95% simulation interval. Discrepancy is observation/prediction minus one, as in Figure 4; positive values indicate underprediction.

Realization	Condition	Prediction	Observation [95% CI]	Discrepancy
4	A	134.65	131.66 [126.18,137.46]	-2.22%
4	B	159.82	169.12 [162.11,176.21]	$+5.82\%$
5	A	131.36	131.64 [126.08,137.23]	$+0.21\%$
5	B	164.33	174.24 [166.81,182.01]	$+6.03\%$
6	A	121.57	127.90 [122.43,133.20]	$+5.21\%$
6	B	166.16	175.75 [168.48,183.11]	$+5.77\%$
7	A	118.09	121.55 [116.52,126.69]	$+2.93\%$
7	B	160.42	172.46 [165.10,179.99]	$+7.51\%$

The geometric-mean B/A contrast is 1.288571 predicted, 1.349337 observed, with hierarchical interval $[1.275848, 1.426240]$. Three predictions lie in their simulation intervals; four of eight pass the predefined 10% agreement check. The isolated-pair baseline $4/(1 + 126/R_{\mathrm{pair}}) \simeq 2.11$ uses the finite-N, confirmation-matched weak/strong pair mean-time ratio $R_{\mathrm{pair}} = 140.591982\ldots$. The four-channel numerator neglects the 123 orthogonal rivals in A, so this is an approximate isolated-pair baseline. We compare it with the 128-key confirmation cohort defined above.

D HIDDEN-STATE TEST AND UNRESOLVED ERROR

The exact-fiber test fixes $d = 0.005$, $s = 0.8$ and a strong leader on bank 4B. Residual directions are sampled on the corresponding sphere, configurations with another leader are rejected, and the full Gibbs weight is applied. The coarea factor is constant on this fiber. The weighted median of background softmax mass splits the initial distribution into halves. The primary test uses 32,768 futures per half; independent-quadrature, disjoint-noise replication uses 16,384 per half. At horizon 0.1 and $h = 0.000625$, the high-minus-low future-gap differences are respectively -0.001666 $[-0.002203, -0.001130]$ and -0.002163 $[-0.002916, -0.001410]$. The two-key rotational null splits by the weighted median of the third ambient coordinate, since background weight is zero. It yields -0.000401 $[-0.000947, 0.000145]$. Intervals use a normal approximation with independent-group standard errors, conditional on finite initial quadrature. A partially shared-noise repeat is not independent evidence. Primary settings follow an exploratory bank-0 test; bank-4 lifetimes were already known. The tests assess local prediction; lifetime accuracy and the sufficiency of background mass remain open.

Numerical contributions to the lifetime residual. The following post-validation checks assess numerical and trajectory consistency.

Operator discretization. On bank 4B, increasing Erlang phases from 32 to 128 to 512 gives 159.396, 159.819, 159.925. Reflecting boundaries at $d = -0.15$ to -0.35 do not materially change the mean at the tested resolution. Spatial and quadrature errors remain coupled: mesh 0.0025 gives 168.102 with two million importance samples but 161.047 with eight million, versus observed 169.119. The

closer value is not a corrected prediction. A singular bank-0B refinement selected an isolated starting component. Selecting the component with the greatest inside Gibbs mass among those reaching both regions repairs the solve (164.344 at mesh 0.0025) but does not certify starting-node or continuum error; bank-4 frozen results are unchanged.

The table reports target-start mean predictions for bank 4B. Each row uses an independent quadrature, extended with independent samples. Unstable results are retained.

	2,097,152 points		8,388,608 points	
Repeat	Mesh 0.005	Mesh 0.0025	Mesh 0.005	Mesh 0.0025
1	160.929	194.196	158.619	161.438
2	158.935	166.388	158.097	160.944
3	160.645	168.667	158.628	161.283

At mesh 0.01, the two-million-point means are 155.395, 153.574, 155.344. The sample standard deviation at mesh 0.005 is 1.078; it does not bound numerical bias. The nearest occupied starting bin also varies. Budgets within each replicate share samples. On the eight-million-point operators, geometry-filtered local affine evaluations at the boundary start, using all three tested radii 0.025, 0.04 and 0.06, change the nearest-bin values by less than 0.4. Unweighted affine fits use at least six centers, a rank-three radius-scaled design and a condition number below 10^6. Centers satisfy the necessary pair-ellipse union and $s + d/2 \le 1$; this filter is not a sufficient description of high-rank support. All radii are retained; sensitivity persists at $(0.3, 0.8)$. Two independent 8,388,608-point two-key controls give 81.661 and 81.931 at mesh 0.0025, versus independent confirmation references 80.965730 (orthonormal chart) and 80.962443 (planar chart), both with 128 phases and $\Delta = 0.1$. These diagnostics neither bound high-rank error nor replace frozen predictions.

Simulation timestep. On bank 4B, 1,024 independent Brownian-coupled path pairs sum four fine ambient Gaussian increments per coarse step, retaining separate states and clocks. Reducing h from 0.0025 to 0.000625 changes the mean by -2.343 $[-4.252, -0.755]$; 512 new pairs at 0.000625 versus 0.00015625 give $+0.093$ $[-0.863, 1.144]$. Intervals bootstrap pairs; the nested timing pilot is not independent. These measure sensitivity, not a certified continuum limit or transferable correction.

Additional state coordinate. A third background coordinate was also unstable under refinement (bank-0A: 131.19 to 212.55), so the favorable bank-0B result does not establish a reliable repair.

Trajectory-weighted residual sensitivity. Two independent bank-4B cohorts of 128 paths yield 8,486 reset-state visits, each tested with 16 independent blocks ($h = 0.0025$). Blocks stop at confirmation or the first target return after five time units. We sum estimates of $\mathbb{E}_y[D + AV(Y')] - V(y)$ along each path, where D is block duration, A is the survival indicator and V is a fixed continuation prediction. The diagnostic evaluates a separate post-validation numerical candidate. For this candidate ($V(a_0) = 158.619$), mesh-0.005 sums are 13.78 $[6.61, 21.19]$ and 6.53 $[0.49, 12.63]$; mesh 0.0025 gives 11.01 $[3.81, 18.35]$ and 3.91 $[-2.22, 10.05]$. Pointwise intervals bootstrap whole source paths. All three 8M-point finer-mesh repeat intervals include zero. The 200-block cap censors one path; its unknown tail precludes a full-lifetime error bound. These checks neither refit predictions nor isolate missing-state error. The supplement includes the protocol and results.

Scope of the memory identity and later reconstruction. For reversible $A = -\mathcal{L}$ killed at $\partial\Omega$, let P be orthogonal conditional averaging in $L^2(\pi|_\Omega)$, $Q = I - P$ and $A_{XY} = XAY$. For real $z > 0$ and compatible block domains and inverses, on $\operatorname{Ran} P$ the standard resolvent identity is

$$P(z + A)^{-1}P = [z + A_{PP} - A_{PQ}(z + A_{QQ})^{-1}A_{QP}]^{-1}. \tag{10}$$

The subtracted operator is z-dependent and positive semidefinite (also at $z = 0$ if inverses exist), generally not a scalar correction. Confirmation augmentation need not be self-adjoint, and elimination can change the reward and starting-state reconstruction. Neither the first-exit Ritz bound nor the local fiber test attributes the B-condition confirmation-time residual to memory.

Galerkin tests (Thiede et al., 2019; Lorpaiboon et al., 2024) compared 16-block histories with one-block restarts: 26,214,400 steps per arm per bank, two development banks, and 2,048 fresh validation paths at each of four sources. Bias reduction was 1.729 percentage points (nominal simultaneous interval $[-3.576, 6.236]$, conditional on frozen pools), below the required three-point improvement; neither superiority nor equivalence was established.