CAN YOUR ATTRACTORS ADAPT? PERSISTENT READ AND WRITE IN ASSOCIATIVE MEMORY

Anonymous authors
Paper under double-blind review

ABSTRACT

Associative memory networks recover stored information by evolving incomplete cues toward learned attractors. Most existing formulations assume a fixed memory bank and therefore leave a central problem unresolved: how can memories be added without unpredictably damaging the landscape that supports future recall? Persistent ingestion changes the dynamics governing all future retrievals. We introduce **Attractor Lifecycle Control (ALC)**, a framework for persistent read and write operations over a versioned energy landscape. A read uses a transient query lens to reshape the geometry traversed by the state without altering the stored landscape. A write evaluates a candidate update for memory separation, attractor stability, and preservation of earlier content before modifying the system. A recurrent **Proposal Flow Model (PFM)**, trained by conditional flow matching, iteratively proposes bounded read schedules and candidate write repairs from the current landscape, operation constraints, and previous predictions. These outputs remain proposals: the dynamical system executes them and accepts only outcomes that satisfy the declared geometric and stability checks. ALC achieves **82.1% Recall@1** on LongMemEval-S and reaches a certified stored core on **91.8%** of queries. On controlled synthetic tasks derived from approximately 770,000 arXiv abstracts, ReadPFM improves certified retrieval success by **25%**, while WritePFM enables exact writes for **28.1%** of initially uncertifiable cases. Cost-fitted physical descent reduces NFE by **72.2%** at inference without changing terminal outcomes.

1 INTRODUCTION

Associative memories recover stored information through dynamics rather than explicit addresses. A partial or corrupted cue initializes a state, and an energy-decreasing evolution carries it toward a stored attractor. Hopfield networks established this view of content-addressable recall (Hopfield, 1982; 1984), building on the stability principle of Lyapunov's direct method (Lyapunov, 1992). Dense and modern variants have since increased storage capacity and connected energy-based retrieval to attention (Krotov & Hopfield, 2016; Demircigil et al., 2017; Ramsauer et al., 2021). Yet these systems are most naturally analysed after their memory landscape has been fixed.

Persistence makes the problem fundamentally different. Adding a memory changes the energy landscape that supports every future recall: it can create new critical points, deform existing basins, and reduce the separation between stored patterns. A write is therefore not simply an append to an index. It is a transformation of the dynamical system itself. Retraining or unconstrained parameter updates can absorb new data, but they do not directly establish which properties of the previous landscape remain intact. Existing work has explored fast test-time writes and online continual storage (Bartunov et al., 2020; Alonso & Krichmar, 2024); the broader lifecycle question remains: how can one repeatedly read from and write to an associative landscape while governing the consequences of each operation?

We introduce *Attractor Lifecycle Control* (ALC), a framework for persistent operations over a bounded, versioned energy landscape called *Remanence*. ALC treats reading as navigation through the current landscape and writing as a constrained transformation of that landscape. Its central separation is between proposing an operation and authorizing its outcome. A learned or optimized component may suggest how to retrieve or where to place a memory, but the underlying dynamics execute the proposal and operation-specific checks determine whether the resulting state is admissible.

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.

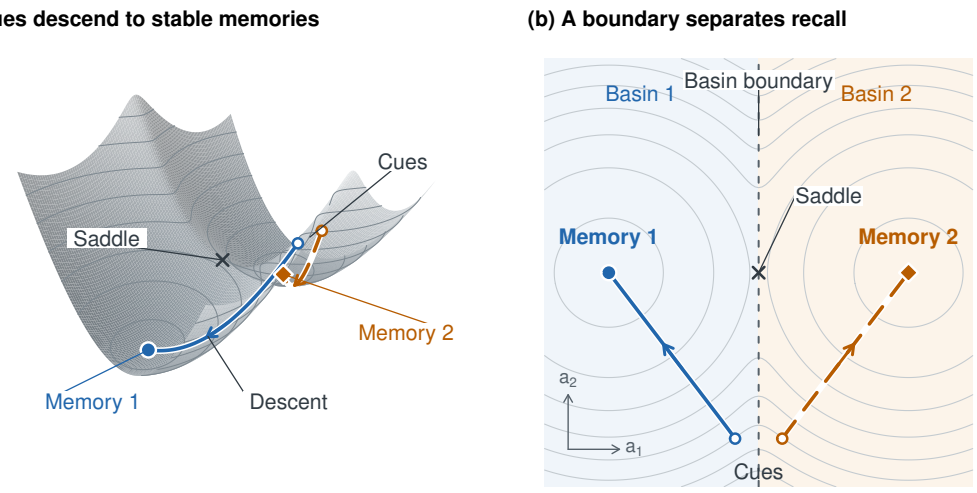

Figure 1: **Cues descend into stable memory attractors.** The energy surface and basin map show the same recall trajectories in an illustrative two-memory system. The dashed boundary separates their destinations. Construction details: Appendix E.1.

This separation addresses two distinct failures. During reading, ordinary energy descent may settle in an irrelevant basin. Precision-modulated retrieval can reshape motion away from equilibrium (Ghosh et al., 2026), but no mobility operator can release a state from an exact incorrect critical point: the energy gradient, and hence the controlled field, is already zero. During writing, an apparently useful candidate can damage the attractors that make earlier memories retrievable. ALC responds by combining transient query-conditioned guidance for reads with verified landscape transformations for writes. Temporary read guidance is removed before the terminal state is judged, whereas a write changes persistent memory only after its proposed landscape passes the declared checks.

ALC couples this physical substrate, the *Remanence Plant*, with a recurrent *Proposal Flow Model* (PFM). The plant retains authority over execution, versioning, and acceptance. PFM amortizes the construction of bounded read schedules and write repairs, refining its proposals in response to the landscape and the consequences of earlier attempts. We train these proposal fields by conditional flow matching (Lipman et al., 2023) and use recurrent self-conditioning inspired by Flow Reasoning Models (Helbling et al., 2026). Recurrence supplies additional proposal computation; it does not turn a neural prediction into a certificate.

Our contributions are : **1)** We formulate persistent associative memory as lifecycle control over a bounded, versioned energy landscape, with reads and writes represented as governed transformations of a shared dynamical state. **2)** We introduce ALC, combining a Remanence Plant with recurrent proposal flows while preserving a strict boundary between neural proposals, physical execution, and operation-specific acceptance. **3)** We characterize the equilibrium obstruction of mobility-only recall and develop complementary mechanisms for transient read guidance and verified persistent writes, evaluated separately at the proposal, dynamical, geometric, and semantic levels.

2 Related Work

Associative memory and energy-based retrieval. Hopfield networks cast content-addressable memory as convergence toward attractors of an energy-decreasing dynamical system (Hopfield, 1982; 1984). Dense and modern Hopfield networks increase capacity, connect retrieval to attention, and extend association across layers (Krotov & Hopfield, 2016; Demircigil et al., 2017; Ramsauer et al., 2021; Krotov, 2021; Hoover et al., 2022). These formulations primarily study representation and retrieval for a specified memory system; ALC studies how reads and writes share a landscape that changes over time.

Writable and continual memory. Adaptive bidirectional memories update associations online (Kosko, 1987); later systems learn fast update rules (Bartunov et al., 2020), combine quantized codes with network growth (Alonso & Krichmar, 2024), or use energy changes to study forgetting and replay (Takeda et al., 2026). They establish that online writes are possible. ALC adds a different operational requirement: every accepted write is a versioned landscape transformation checked against geometric and numerical conditions before commit.

Controlled and verifier-bounded dynamics. External inputs can alter Hopfield interactions and reshape retrieval geometry (Betteti et al., 2025), while dynamic-manifold models learn context-dependent deformations of the attractor landscape (Li et al., 2025). Precision-controlled associative dynamics instead modulate descent while preserving the underlying energy (Ghosh et al., 2026). Our numerical tools draw on state-dependent Riccati design, iterative local control, and model predictive control (Beeler & Cox, 2004; Çimen, 2008; Todorov & Li, 2005; Mayne et al., 2000). Related learning systems place an explicit stability or safety check around a neural proposal (Chang et al., 2019; Wabersich & Zeilinger, 2021). ALC applies this separation to associative memory: a transient query lens changes basin competition, mobility governs descent, and operation-specific checks retain authority over the executed read or proposed write.

Flow matching and recurrent self-conditioning. Neural ODEs and flow matching learn continuous vector fields and conditional transport paths (Chen et al., 2018; Lipman et al., 2023). Self-conditioning feeds earlier predictions into later refinement (Chen et al., 2023); Flow Reasoning Models extend this idea with detached recurrent rollouts (Helbling et al., 2026), related to on-policy aggregation (Ross et al., 2011). PFM combines Transformer attention, a Perceiver-style full-bank bottleneck, and DiT-style modulation (Vaswani et al., 2017; Jaegle et al., 2021; Peebles & Xie, 2023). Its learned flows produce bounded read schedules or write repairs; the Remanence dynamics and exact operation checks still determine the outcome.

3 Persistent Associative Memory

Persistent associative memory alternates two operations on a shared energy landscape: a **read** moves a query state with memory fixed, whereas an accepted **write** changes the landscape traversed by later queries (Figure 2).

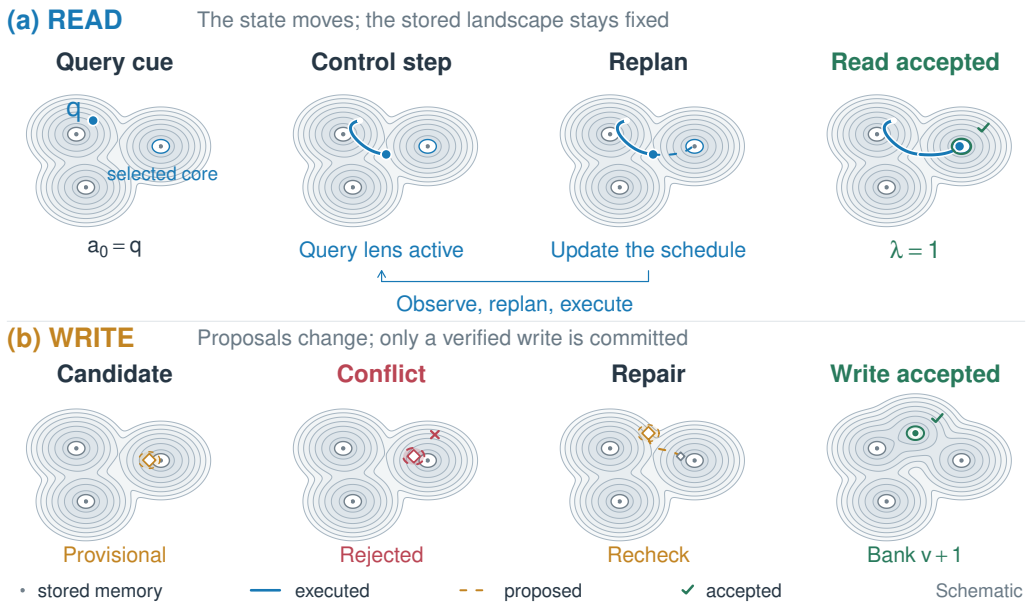

Figure 2: **Read control moves the state; verified writes reshape memory.** The stored landscape remains fixed during a read and changes only after an accepted write. Paths and proposals are schematic, not measured PFM rollouts.

3.1 A VERSIONED MEMORY LANDSCAPE

At version v, the bank contains K_v patterns $\boldsymbol{x}_i \in \mathbb{R}^n$ collected in $\boldsymbol{X}_v$. Let $\boldsymbol{Q}_S = \boldsymbol{R}_S^{-1}$, $\boldsymbol{S} = \boldsymbol{Q}_S^{1/2}$, and define

$$\boldsymbol{y}_i = \boldsymbol{S}\boldsymbol{x}_i, \qquad \boldsymbol{z} = \boldsymbol{S}^{-1}\boldsymbol{a}, \qquad \boldsymbol{Y}_v = [\boldsymbol{y}_1, \dots, \boldsymbol{y}_{K_v}]. \tag{1}$$

C0–C4 assigns label i the closed ball $\overline{B}(\eta\boldsymbol{y}_i, \rho_i)$ in whitened coordinates. Labels use stable identifiers rather than array positions. The committed version is

$$\mathcal{V}_v = (K_v, \boldsymbol{X}_v, \boldsymbol{Y}_v, \boldsymbol{\rho}_v, \mathcal{I}_v, \mathcal{P}_v; \mathcal{D}, \psi, h_v), \tag{2}$$

where $\boldsymbol{\rho}_v$, $\mathcal{I}_v$, $\mathcal{P}_v$, and h_v bind radii, ordered identifiers, payload lineage, and version content; $\mathcal{D}$ fixes the design and ψ the representation. Validity requires

$$\begin{aligned}\mathrm{Inv}(\mathcal{V}_v) \iff\ & \mathrm{Inv}_{\mathrm{format}}(\mathcal{V}_v) \wedge \mathrm{Inv}_{\mathrm{static}}(\mathcal{V}_v) \\ & \wedge \mathrm{Inv}_{\mathrm{core}}(\mathcal{V}_v) \wedge \mathrm{Inv}_{\mathrm{model}}(\mathcal{V}_v) \wedge \mathrm{Inv}_{\mathrm{lineage}}(\mathcal{V}_v),\end{aligned} \tag{3}$$

The static and core terms contain C0–C4; the remaining terms bind representation, identity, lineage, and content. Every operation pins one immutable version.

3.2 READING NAVIGATES THE CURRENT LANDSCAPE

An encoded query initializes $\boldsymbol{a}(0) = \boldsymbol{q}$. A full-bank scorer either selects a unique hypothesis i_q in stable-identifier order or refuses; no correct identifier enters inference. Conditioned on i_q, control evolves the actual state with $\mathcal{V}_v$ pinned, and all transient guidance vanishes before the exact local reader judges $\boldsymbol{z}_T = \boldsymbol{S}^{-1}\boldsymbol{a}_T$. Success requires

$$\mathrm{LocalRead}_{\mathcal{D}}(\mathcal{V}_v, \boldsymbol{z}_T) = (\mathrm{accepted}, \mathcal{I}_v[i_q], \varepsilon_{\mathrm{eq}}), \tag{4}$$

where $\varepsilon_{\mathrm{eq}}$ bounds the whitened equilibrium residual and all bindings match $\mathcal{V}_v$. Another certified core does not satisfy the selected hypothesis; payload resolution occurs only after Equation 4.

3.3 WRITING TRANSFORMS THE LANDSCAPE

For candidate $\boldsymbol{c}$ with $\boldsymbol{y}_c = \boldsymbol{S}\boldsymbol{c}$, the writer privately constructs $\widetilde{\mathcal{V}} \in \mathcal{A}_w(\mathcal{V}_v, \boldsymbol{y}_c)$, possibly using a repair $\|\widetilde{\boldsymbol{y}} - \boldsymbol{y}_c\|_2 \le \epsilon_w$ with an admissible radius and updated lineage. Publication obeys

$$\mathrm{Write}(\mathcal{V}_v, \boldsymbol{c}) = \begin{cases} \widetilde{\mathcal{V}}, & \begin{array}{l}\widetilde{\mathcal{V}} \in \mathcal{A}_w(\mathcal{V}_v, \boldsymbol{y}_c), \\ \mathrm{Verify}_{\mathrm{WRITE}}(\mathcal{V}_v, \widetilde{\mathcal{V}}) = \mathrm{ACCEPT}, \\ \mathrm{CAS}(\mathrm{head}(\mathcal{V}_v), \widetilde{\mathcal{V}}) = \mathrm{success}, \end{array} \\ \mathcal{V}_v, & \text{otherwise.} \end{cases} \tag{5}$$

The research artifact constructs, binds, and replays the complete verified successor, rejecting predecessor mismatch. Its guarantee concerns that successor; durable atomic publication remains a commit-layer assumption. A rejected or unresolved proposal leaves $\mathcal{V}_v$ unchanged.

Scope of success. Proposal quality, numerical execution, geometric acceptance, and semantic relevance are distinct. A local certificate is not semantic correctness; an accepted write is not globally optimal placement; budget exhaustion is not infeasibility.

4 ATTRACTOR LIFECYCLE CONTROL

The lifecycle in Figure 2 separates proposals from operations on a pinned memory version. We establish fixed-energy dissipation, certified local cores, and exact recovery of stored energy after lens removal; proofs are in Appendix A.

Theorems are symbolic consequences of stated assumptions; executable C0–C4 and local-read certificates bind one finite bank and tolerance, not an entire continuous trajectory.

4.1 THE REMANENCE PLANT

For a pinned bank $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K]$ and $\boldsymbol{R}_S \succ 0$, we use the log-sum-exp associative energy (Ramsauer et al., 2021; Ghosh et al., 2026)

$$E(\boldsymbol{a}; \boldsymbol{X}) = \frac{1}{2}\boldsymbol{a}^\top \boldsymbol{R}_S \boldsymbol{a} - \frac{\eta}{\beta} \log \sum_{j=1}^{K} \exp(\beta \boldsymbol{x}_j^\top \boldsymbol{a}). \tag{6}$$

With $\boldsymbol{\sigma}(\boldsymbol{a}) = \mathrm{softmax}(\beta \boldsymbol{X}^\top \boldsymbol{a})$, its gradient and Hessian are

$$\boldsymbol{g}(\boldsymbol{a}) = \boldsymbol{R}_S \boldsymbol{a} - \eta \boldsymbol{X} \boldsymbol{\sigma}(\boldsymbol{a}), \tag{7}$$

$$\boldsymbol{H}(\boldsymbol{a}) = \boldsymbol{R}_S - \eta\beta \boldsymbol{X} \big(\mathrm{Diag}\, \boldsymbol{\sigma} - \boldsymbol{\sigma}\boldsymbol{\sigma}^\top\big) \boldsymbol{X}^\top. \tag{8}$$

The physical read state follows

$$\dot{\boldsymbol{a}} = -\boldsymbol{\Pi}(\boldsymbol{a}, t) \boldsymbol{g}(\boldsymbol{a}), \qquad \pi_{\min} \boldsymbol{I} \preceq \boldsymbol{\Pi}(\boldsymbol{a}, t) \preceq \pi_{\max} \boldsymbol{I}. \tag{9}$$

Mobility changes motion where $\boldsymbol{g} \neq 0$, but not stored-energy critical points.

Theorem 4.1 (Dissipation and the equilibrium obstruction)**.** *Along every differentiable solution of Equation 9,*

$$\frac{\mathrm{d}}{\mathrm{d}t} E(\boldsymbol{a}(t); \boldsymbol{X}) = -\boldsymbol{g}(\boldsymbol{a})^\top \boldsymbol{\Pi}(\boldsymbol{a}, t) \boldsymbol{g}(\boldsymbol{a}) \leq -\pi_{\min} \|\boldsymbol{g}(\boldsymbol{a})\|_2^2 \leq 0. \tag{10}$$

If $\boldsymbol{g}(\boldsymbol{a}_c) = 0$, then $\dot{\boldsymbol{a}}_c = 0$ for every admissible mobility. Consequently, mobility-only SDRE, SLQ, or receding-horizon control cannot dislodge an exact wrong critical point or cross a higher stored-energy barrier enclosing the initial sublevel set.

Proof sketch. The chain rule and spectral lower bound give the inequality; at a critical point every admissible mobility multiplies zero.

4.2 CERTIFIED WRITE LIFECYCLE

The verifier evaluates C0–C4 on the complete post-write bank. In the whitened coordinates of Equation 1, these conditions bound resident norms and radii, enforce positive competitor margins, make the fixed-point map $T(\boldsymbol{z}) = \eta \boldsymbol{Y} \boldsymbol{\sigma}(\boldsymbol{z})$ a strict self-map and contraction, and separate the certified balls. The exact inequalities and their derivation appear in Appendix A.14.

Theorem 4.2 (Certified local cores)**.** *Suppose a finite bank satisfies C0–C4 with $0 < \kappa_* < 1$ and $\delta_* > 0$. For every label i, the map T sends the closed ball $\overline{B}(\eta \boldsymbol{y}_i, \rho_i)$ into itself and is a contraction there with constant at most $1 - \kappa_*$. The ball contains exactly one fixed point $\boldsymbol{z}_i^*$, the whitened-energy Hessian is bounded below by $\kappa_* \boldsymbol{I}$ throughout the ball, and certified balls are disjoint with reserve δ_*. Thus every certified fixed point is a strict local minimum with an unambiguous resident label.*

Proof sketch. The directed margin bounds competing softmax mass; the remaining conditions establish self-mapping, contraction, curvature, and separation. See Appendix A.14.

The runtime verifier uses exact rationals and outward enclosures, so acceptance establishes C0–C4 for the named finite snapshot (not a Lean proof). The trusted write path reconstructs the full successor, checks it, and binds identity and lineage. Given sound verification and atomic compare-and-swap, acceptance preserves Equation 3; rejection or a stale predecessor keeps the old version. Bounded search implies neither feasibility nor optimality.

4.3 QUERY LENS AND CONTROLLED READ

Theorem 4.1 shows why target selection cannot be assigned to mobility alone. ALC instead uses a transient full-bank query lens. For frozen query log-priors $\boldsymbol{b}(\boldsymbol{q}) \in \mathbb{R}^K$ and $0 \leq \lambda \leq 1$, define

$$E_\lambda(\boldsymbol{a}; \boldsymbol{q}) = \frac{1}{2}\boldsymbol{a}^\top \boldsymbol{R}_S \boldsymbol{a} - \frac{\eta}{\beta} \log \sum_{j=1}^{K} \exp\big(\beta \boldsymbol{x}_j^\top \boldsymbol{a} + (1-\lambda)\alpha b_j(\boldsymbol{q})\big). \tag{11}$$

All residents participate; the bias is fully active at $\lambda = 0$ and absent at $\lambda = 1$. With

$$\boldsymbol{\sigma}^\lambda = \mathrm{softmax}\big(\beta \boldsymbol{X}^\top \boldsymbol{a} + (1-\lambda)\alpha \boldsymbol{b}(\boldsymbol{q})\big), \qquad \boldsymbol{g}_\lambda = \boldsymbol{R}_S \boldsymbol{a} - \eta \boldsymbol{X} \boldsymbol{\sigma}^\lambda, \tag{12}$$

the state follows $\dot{\boldsymbol{a}} = -\pi_s \boldsymbol{g}_\lambda$ while each stage holds λ fixed.

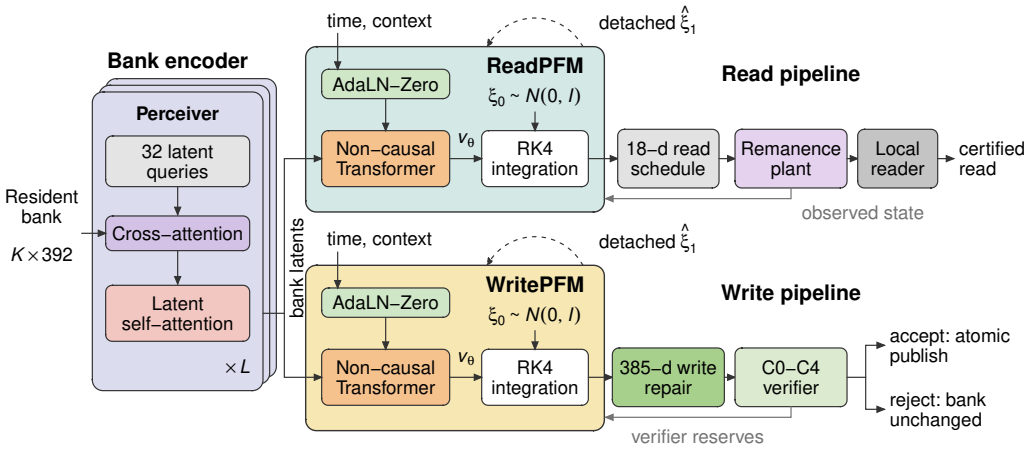

Figure 3: **Inside the recurrent Proposal Flow Models.** A Perceiver encoder compresses the complete resident bank into 32 latents that both models cross-attend to. The separately parameterized ReadPFM and WritePFM integrate a non-causal Transformer velocity field with RK4 from Gaussian noise and condition each refinement on the detached preceding endpoint $\hat{\xi}_1$. An 18-coordinate read schedule is executed by the Remanence plant; a 385-coordinate write repair must pass the exact C0–C4 verifier before it is published.

Proposition 4.3 (Fixed-stage dissipation and lens removal)**.** *For fixed* λ *and* $\pi_{\min} \le \pi_s \le \pi_{\max}$,

$$\frac{\mathrm{d}}{\mathrm{d}t} E_\lambda(\boldsymbol{a}(t); \boldsymbol{q}) = -\pi_s \|\boldsymbol{g}_\lambda(\boldsymbol{a}; \boldsymbol{q})\|_2^2 \le 0. \tag{13}$$

Moreover $E_1(\boldsymbol{a}; \boldsymbol{q}) = E(\boldsymbol{a}; \boldsymbol{X})$ *and* $\boldsymbol{g}_1(\boldsymbol{a}; \boldsymbol{q}) = \boldsymbol{g}(\boldsymbol{a})$ *exactly.*

Proof sketch. Differentiate the fixed-stage energy; at $\lambda = 1$ the query term is zero (Appendix A.19).

ALC carries the actual endpoint through $0 = \lambda_0 < \cdots < \lambda_L = 1$. Energy decreases within each fixed stage, but no monotonicity is claimed across lens changes. The final local reader accepts only the scorer-named certified core; other outcomes are typed refusal or `UNKNOWN`, not reachability claims. Current continuation is identity-whitened and numerically checked; terminal verification covers C0–C4, residual, and discrete roundoff, but not continuous-time error or semantics.

4.4 Optimization proposes; the plant decides

SDRE/DRE constructs local state-dependent mobility, while SLQ improves a finite open-loop schedule around a nonlinear rollout. Receding-horizon NMPC is planned but currently fails closed rather than relabeling SLQ as feedback. All remain subject to Theorem 4.1. Section 5 amortizes proposals; only the local reader or write verifier authorizes them.

5 Recurrent Proposal Flow Models

The mechanisms in Section 4 define which proposals can be executed or accepted, but repeatedly solving for those proposals is costly. We amortize this search with a recurrent *Proposal Flow Model* (PFM). Its two instances, ReadPFM and WritePFM, share an architectural pattern but have separate parameters and output spaces. Their predictions are never certificates: the Remanence plant executes read schedules, while the exact write verifier accepts or rejects candidate repairs.

5.1 Conditional flow and recurrent carry

Let $\boldsymbol{\xi}^\star$ denote a teacher proposal, ϵ a sample from a standard Gaussian in the corresponding proposal space, and $\mathcal{C}$ the pinned bank and operation context. We train the velocity field on the linear interpolant

$$\boldsymbol{\xi}_t = (1-t)\epsilon + t\boldsymbol{\xi}^\star, \qquad \boldsymbol{u}^\star = \boldsymbol{\xi}^\star - \epsilon, \qquad t \sim \mathcal{U}[0,1], \tag{14}$$

using the coordinate-weighted conditional flow-matching objective (Lipman et al., 2023)

$$\mathcal{L}_{\mathrm{PFM}} = \mathbb{E}\left[\left\|\boldsymbol{v}_\theta(\boldsymbol{\xi}_t, t, \mathcal{C}, \mathrm{sg}[\widehat{\boldsymbol{\xi}}^{r-1}]) - \boldsymbol{u}^\star\right\|_{\boldsymbol{M}}^2\right]. \tag{15}$$

Here r indexes recurrent attempts and sg stops gradients through the preceding clean estimate. A pooled condition token also predicts a typed outcome, giving

$$\mathcal{L}_{\mathrm{train}} = \mathcal{L}_{\mathrm{PFM}} + \alpha\ \mathrm{CE}(h_\theta(\mathcal{C}), y). \tag{16}$$

Training first fits the uncoupled flow field, then enables one detached self-conditioned rollout. At inference, fixed-grid RK4 integrates the same field from base noise to $t = 1$. The decoded proposal is executed or checked; the observed state or verifier reserves update $\mathcal{C}$ before the next attempt. This is a deterministic proposal-space ODE, rather than a stochastic diffusion process over memory contents.

5.2 Full-bank non-causal denoiser

Figure 3 shows the common computation. Each resident contributes its 384-dimensional vector and eight operation-specific scalars. A Perceiver-style bottleneck (Jaegle et al., 2021) uses thirty-two learned latent queries to cross-attend to all K residents, then passes them through latent self-attention blocks. Resident tokens have no positional encoding, so reordering stable bank identifiers does not alter the condition. This bottleneck costs $O(KL)$ attention for $L = 32$ latents and does not remove any memory from the bank.

Proposal tokens enter bidirectional Transformer (Vaswani et al., 2017) blocks with DiT-style AdaLN-Zero modulation (Peebles & Xie, 2023), cross-attention to the encoded condition, and a feed-forward layer. Non-causal attention is natural here because a proposal is optimized as a complete schedule or repair, not generated autoregressively. Detached recurrent carry follows the fixed-point forcing pattern of Helbling et al. (2026), but here the iterates are control proposals rather than task solutions.

5.3 Read and write proposal spaces

For six continuation stages, ReadPFM operates in $\mathbb{R}^{18}$: five logits allocate strictly increasing continuation levels from the currently accepted λ to 1, six control the query-bias scale, six allocate stage duration, and one controls scalar mobility. Floors and bounded transforms make every decoded schedule structurally executable. The plant applies the next segment, returns its actual state, and the model replans; only the terminal local reader can declare a certified core.

WritePFM instead operates in $\mathbb{R}^{385}$, representing a 384-dimensional center displacement and one radius coordinate. Its condition includes candidate-relative resident features and the latest C0–C4 reserves. The decoder enforces the movement and radius boxes, after which the exact verifier either commits the successor bank or returns rejection evidence for another attempt. We evaluate six paired capacity profiles spanning 5–68M parameters; capacity is an experimental variable, not part of the definition of ALC.

6 Experiments

We evaluate ALC as a lifecycle rather than as another nearest-neighbor retriever. The experiments separate four questions: can a long write stream be replayed and recertified; can the resulting bank support semantic recall; do learned proposal flows improve fixed read and write proposals under exact execution; and can this control be made cheaper without changing its terminal decision?

6.1 Evaluation design

Table 1 states the denominator and authority for each claim. The persistent experiment uses LongMemEval-S (Wu et al., 2025): 500 questions, of which 470 are answerable, 23,867 chronological sessions, and 124,334 round-level writes. A frozen MiniLM encoder maps content into the 384-dimensional Remanence state, but the reported ALC pipeline includes ingestion, exact write verification, versioned replay, scoring, physical descent, and terminal reading.

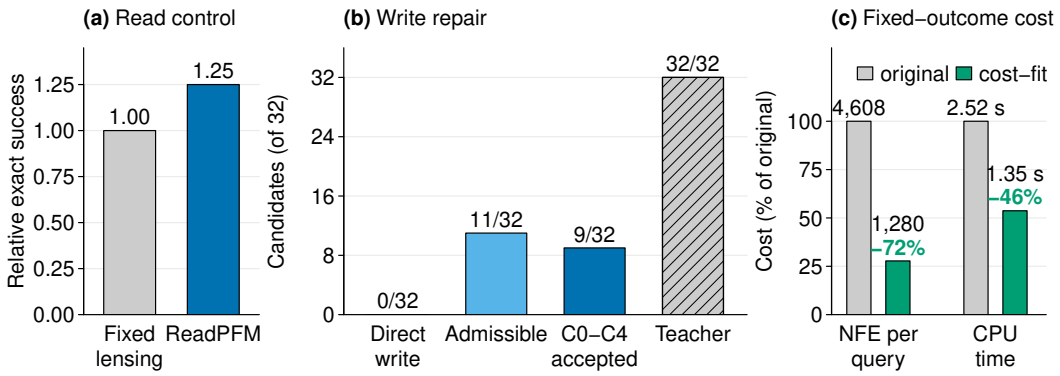

Figure 4: **Learned control outcomes and inference cost.** **(a)** ReadPFM improves exact terminal success by 25% over fixed query lensing on teacher-feasible controlled reads. **(b)** Of 32 rejected write candidates, 11 become geometrically admissible and nine pass exact C0–C4 authorization; the hatched bar is the finite teacher. **(c)** Cost fitting reduces NFE and CPU time while preserving all 8/8 terminal outcomes; bar labels give absolute values.

PFM control is evaluated separately on bank-disjoint synthetic records derived from approximately 770,000 arXiv abstracts. The read cohort has 56 trajectories over eight banks, including 24 teacher-feasible targets; the write cohort has 32 verifier-rejected candidates over eight banks. Frozen validation selects a 37.9M ReadPFM and a 5.2M WritePFM from six capacities per operation. These controlled tests isolate physical control and exact repair; they do not measure semantic relevance.

Table 1: Evaluation layers. The final column identifies the component that determines success rather than merely proposing it.

Experiment	Denominator	Primary metric	Authority
Persistent lifecycle	124,334 writes; 500 queries	replay, Recall@k, core	exact verifier/reader
Read control	24 feasible targets, 8 banks	exact terminal success	local reader
Write repair	32 rejected writes, 8 banks	exact C0–C4 acceptance	write verifier
Cost fit	8 queries + 40 starts	NFE, time, outcome match	replayed outcome

6.2 The complete persistent lifecycle

ALC accepts 117,476 of 124,334 chronological writes (94.5%): 117,431 are admitted and 45 are merged, while 6,858 fail the write policy. All 500 final journals replay and recertify, and required evidence survives ingestion for all 470 answerable queries. The archived tree retains six unreferenced intermediate snapshots; they do not enter any metric denominator and are disclosed in Appendix F.

The resulting system obtains **82.1% Recall@1**, 94.0% Recall@5, and 96.0% Recall@10 on LongMemEval-S. Physical descent reaches a certified stored core for 459/500 queries (91.8%), with no wrong-core or numerical outcomes; 361/470 answerable queries finish in a core containing the required evidence. These numbers measure different layers: Recall@k measures semantic selection, whereas core arrival measures geometric execution. The query lens acts on the selected hypothesis and cannot repair a scorer miss. For that reason, lexical and dense retrieval diagnostics are reported in Appendix F.2, rather than presented as the main ALC comparison.

6.3 Learned read and write control

Figure 4a compares the recurrent controller with the fixed query-lensing schedule while holding the bank, initial state, target, integrator, and terminal reader fixed. ReadPFM raises exact terminal success by a **25% relative improvement**. The small cohort makes this controlled development evidence, not

a broad reachability claim, but it establishes the intended causal result: a learned recurrent proposal changes physical retrieval beyond fixed lensing.

Every write-control example begins as an exact rejection. WritePFM makes 11/32 candidates geometrically admissible and its typed gate authorizes nine, so **28.1%** of initially uncertifiable cases become exact C0–C4 writes (Figure 4b). The finite teacher accepts 32/32, showing that the benchmark contains valid repairs and that substantial model headroom remains. Neither the proposal nor its outcome head can bypass the verifier.

6.4 Efficiency, scaling, and failure modes

On the frozen eight-query continuation cohort, a two-stage cost fit replaces the original six-stage schedule while preserving all eight terminal labels and certificates. It reduces NFE from 4,608 to 1,280 per query (-72.2%) and summed CPU time from 2.521 to 1.354 seconds (-46.3%). Outcomes also remain unchanged on 40 adversarial starts. Because the schedule was fitted on this cohort, this is an inference-cost ablation rather than an independently validated universal default.

All twelve PFM runs complete 50,000 updates, but scale alone does not improve the objective: frozen validation selects the 37.9M ReadPFM and 5.2M WritePFM, while larger read models become unstable late in recurrent training. Appendix E.3 reports the curves and full optimization protocol. The main unresolved failure is abstention: the LongMemEval-S gate falsely accepts 28/30 unanswerable queries. ALC therefore demonstrates persistent, verified operation and learned physical control, while calibrated answerability and semantic reranking remain open.

7 Limitations and Conclusion

ALC certifies geometric properties of a finite bank, not semantic correctness, global reachability, or optimal write placement. Its controlled PFM evidence is also narrow: read control uses 24 feasible targets, write repair uses 32 rejected candidates, and the cost fit is evaluated on eight queries. The present answerability gate remains poorly calibrated, and generalization beyond the fixed 384-dimensional representation, synthetic control records, and tested bank sizes is unresolved.

Within these boundaries, ALC demonstrates a persistent associative-memory lifecycle in which writes transform a versioned landscape, reads navigate the resulting dynamics, and learned proposal flows remain subordinate to exact execution and acceptance. The results establish that recurrent proposals can improve both controlled retrieval and certified write repair while preserving auditable outcomes. Broader semantic reranking, independently validated cost schedules, and full receding-horizon control are the next tests of this framework.

AI use statement

Generative-AI tools assisted with language editing, LaTeX drafting, and figure layout iteration. They did not generate experimental measurements or serve as an authority for mathematical claims or citations. The authors checked the final text, derivations, references, code, and reported results and take full responsibility for the submission.

Ethics statement

ALC is evaluated on public research text and synthetic control records and does not involve human-subject intervention. Persistent memory systems can retain sensitive or erroneous information and can make later retrieval behavior difficult to anticipate. The geometric certificates studied here do not establish factuality, fairness, privacy, or authorization. Deployment with personal or sensitive data therefore requires consent, access control, deletion and correction mechanisms, provenance auditing, and independent semantic safety checks.

Reproducibility statement

The model, control, and verification assumptions are stated in Section 3–5, with complete derivations and proofs in Appendices A–C. The evaluation design is summarized in Section 6; Appendix E

specifies dataset construction and preprocessing, split rules, training hyperparameters, model selection, and metric definitions, while Appendix F reports the extended ablations. Appendix G records deterministic numerical settings, artifact identifiers, and claim-to-evidence mappings. The anonymous supplementary material provides inference and evaluation code, frozen selected checkpoints, machine-readable configurations, sanitized result records, and commands for verifying the reported evaluations and regenerating their tables and figures. Architecture manifests, seeds, per-run summaries, and optimization curves document the training evidence; the complete private training and dataset-generation systems are outside the artifact.

REFERENCES

Nicholas Alonso and Jeffrey L. Krichmar. A sparse quantized hopfield network for online-continual memory. *Nature Communications*, 15:3722, 2024. doi: 10.1038/s41467-024-46976-4. URL https://doi.org/10.1038/s41467-024-46976-4.

Sergey Bartunov, Jack W. Rae, Simon Osindero, and Timothy P. Lillicrap. Meta-learning deep energy-based memory models. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=SyljQyBFDH.

Scott C. Beeler and David E. Cox. State-dependent riccati equation regulation of systems with state and control nonlinearities. Technical Report NASA/CR-2004-213245, National Aeronautics and Space Administration, Langley Research Center, July 2004. URL https://ntrs.nasa.gov/citations/20040110953.

Richard Bellman. *Dynamic Programming*. Princeton University Press, Princeton, New Jersey, 1957.

Simone Betteti, Giacomo Baggio, Francesco Bullo, and Sandro Zampieri. Input-driven dynamics for robust memory retrieval in hopfield networks. *Science Advances*, 11(17):eadu6991, 2025. doi: 10.1126/sciadv.adu6991. URL https://doi.org/10.1126/sciadv.adu6991.

John C. Butcher. *Numerical Methods for Ordinary Differential Equations*. John Wiley & Sons, 2 edition, 2008. doi: 10.1002/9780470753767. URL https://doi.org/10.1002/9780470753767.

Jean Cerf. La stratification naturelle des espaces de fonctions différentiables réelles et le théorème de la pseudo-isotopie. *Publications Mathématiques de l'IHÉS*, 39:5–173, 1970. doi: 10.1007/BF02684687. URL https://doi.org/10.1007/BF02684687.

Ya-Chien Chang, Nima Roohi, and Sicun Gao. Neural lyapunov control. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://proceedings.neurips.cc/paper/2019/hash/2647c1dba23bc0e0f9cdf75339e120d2-Abstract.html.

Ricky T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://proceedings.neurips.cc/paper/2018/hash/69386f6bb1dfed68692a24c8686939b9-Abstract.html.

Ting Chen, Ruixiang Zhang, and Geoffrey Hinton. Analog bits: Generating discrete data using diffusion models with self-conditioning. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=3itjR9QxFw.

Tayfun Çimen. State-dependent riccati equation (SDRE) control: A survey. *IFAC Proceedings Volumes*, 41(2):3761–3775, 2008. doi: 10.3182/20080706-5-KR-1001.00635. URL https://doi.org/10.3182/20080706-5-KR-1001.00635.

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(2):288–299, 2017. doi: 10.1007/s10955-017-1806-y. URL https://doi.org/10.1007/s10955-017-1806-y.

Priyam Ghosh, Krish Jaiswal, and Sauhard Gupta. Precision-controlled dynamical systems for associative memory. Unpublished manuscript, 2026. URL https://doi.org/10.13140/RG.2.2.30155.14881.

Alec Helbling, Andrey Bryutkin, Mauro Martino, Duen Horng Chau, Nima Dehmamy, and Hendrik Strobelt. Flow reasoning models: Turning flows into efficient recurrent reasoners, 2026. URL https://arxiv.org/abs/2606.29150v3. arXiv preprint, version 3, 1 September 2026.

Benjamin Hoover, Duen Horng Chau, Hendrik Strobelt, and Dmitry Krotov. A universal abstraction for hierarchical hopfield networks. In *The Symbiosis of Deep Learning and Differential Equations II*, 2022. URL https://openreview.net/forum?id=SAv3nhzNWhw.

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. URL https://doi.org/10.1073/pnas.79.8.2554.

John J. Hopfield. Neurons with graded response have collective computational properties like those of two-state neurons. *Proceedings of the National Academy of Sciences*, 81(10):3088–3092, 1984. doi: 10.1073/pnas.81.10.3088. URL https://doi.org/10.1073/pnas.81.10.3088.

David H. Jacobson and David Q. Mayne. *Differential Dynamic Programming*. American Elsevier Publishing Company, New York, 1970.

Andrew Jaegle, Felix Gimeno, Andrew Brock, Andrew Zisserman, Oriol Vinyals, and João Carreira. Perceiver: General perception with iterative attention. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 4651–4664. PMLR, 2021. URL https://proceedings.mlr.press/v139/jaegle21a.html.

Bart Kosko. Adaptive bidirectional associative memories. *Applied Optics*, 26(23):4947–4960, 1987. doi: 10.1364/AO.26.004947. URL https://doi.org/10.1364/AO.26.004947.

Dmitry Krotov. Hierarchical associative memory, 2021. URL https://arxiv.org/abs/2107.06446.

Dmitry Krotov and John J. Hopfield. Dense associative memory for pattern recognition. In *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. URL https://proceedings.neurips.cc/paper_files/paper/2016/file/eaae339c4d89fc102edd9dbdb6a28915-Paper.pdf.

Joseph P. LaSalle. Some extensions of Liapunov's second method. *IRE Transactions on Circuit Theory*, 7(4):520–527, 1960. doi: 10.1109/TCT.1960.1086720. URL https://doi.org/10.1109/TCT.1960.1086720.

Chong Li, Taiping Zeng, Xiangyang Xue, and Jianfeng Feng. Dynamic manifold hopfield networks for context-dependent associative memory, 2025. URL https://arxiv.org/abs/2506.01303v3. arXiv preprint, version 3.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=PqvMRDCJT9t.

Aleksandr M. Lyapunov. The general problem of the stability of motion. *International Journal of Control*, 55(3):531–534, 1992. doi: 10.1080/00207179208934253. URL https://doi.org/10.1080/00207179208934253. English translation by A. T. Fuller of the work originally published in 1892.

David Q. Mayne, James B. Rawlings, Christopher V. Rao, and Pierre O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000. doi: 10.1016/S0005-1098(99)00214-9. URL https://doi.org/10.1016/S0005-1098(99)00214-9.

John Milnor. *Morse Theory*, volume 51 of *Annals of Mathematics Studies*. Princeton University Press, 1963. doi: 10.1515/9781400881802. URL https://doi.org/10.1515/9781400881802.

William Peebles and Saining Xie. Scalable diffusion models with transformers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4195–4205, 2023.

Lev S. Pontryagin, Vladimir G. Boltyanskii, Revaz V. Gamkrelidze, and Evgenii F. Mishchenko. *The Mathematical Theory of Optimal Processes*. Interscience Publishers, New York, 1962.

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. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=tL89RnzIiCd.

Stephane Ross, Geoffrey Gordon, and Drew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15 of *Proceedings of Machine Learning Research*, pp. 627–635. PMLR, 2011. URL https://proceedings.mlr.press/v15/ross11a.html.

Ken Takeda, Masafumi Oizumi, and Ryo Karakida. Continual learning in modern hopfield networks with an application to diffusion models, 2026. URL https://arxiv.org/abs/2605.27975.

Emanuel Todorov and Weiwei Li. A generalized iterative LQG method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proceedings of the 2005 American Control Conference*, pp. 300–306, 2005. doi: 10.1109/ACC.2005.1469949. URL https://doi.org/10.1109/ACC.2005.1469949.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, pp. 5998–6008, 2017. URL https://proceedings.neurips.cc/paper/7181-attention-is-all-you-need.

Kim Peter Wabersich and Melanie N. Zeilinger. A predictive safety filter for learning-based control of constrained nonlinear dynamical systems. *Automatica*, 129:109597, 2021. doi: 10.1016/j.automatica.2021.109597. URL https://doi.org/10.1016/j.automatica.2021.109597.

Di Wu, Hongwei Wang, Wenhao Yu, Yuwei Zhang, Kai-Wei Chang, and Dong Yu. LongMemEval: Benchmarking chat assistants on long-term interactive memory. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/d813d324dbf0598bbdc9c8e79740ed01-Paper-Conference.pdf.

A Mathematical Foundations and Certification Proofs

This appendix derives the results in Section 4 from the stored energy itself. The organization follows the proof style used for modern Hopfield networks (Ramsauer et al., 2021): definitions are fixed first, elementary identities are proved separately, and the main results are assembled only after every bound has been exposed. The runtime certificate implements the same inequalities with exact rationals and outward enclosures, but the proofs below are statements about real-valued functions.

A.1 Fixed-Bank PCAM Baseline and RW-PCAM Extension

The predecessor PCAM formulation (Ghosh et al., 2026) studies a fixed finite pattern matrix under a positive-definite precision and an admissible positive-definite mobility. In the notation of this paper, its continuous plant is

$$E_v(\boldsymbol{a}) = \frac{1}{2}\boldsymbol{a}^\top \boldsymbol{R}_S \boldsymbol{a} - \frac{\eta}{\beta}\log\sum_{j=1}^{K_v}\exp(\beta \boldsymbol{x}_{v,j}^\top \boldsymbol{a}), \qquad \dot{\boldsymbol{a}} = -\boldsymbol{\Pi}(\boldsymbol{a},t)\nabla E_v(\boldsymbol{a}). \tag{17}$$

PCAM established five fixed-bank conclusions: local well-posedness, Lyapunov dissipation with bounded trajectories, asymptotic invariance, formation of multiple wells under a sufficiently separated

Table 2: How the fixed-bank PCAM results enter ALC. "Break" names the first step in the inherited argument that no longer follows after a bank update.

Fixed-bank result	Required object	Break under a write	ALC disposition
Well-posed flow	One smooth finite-bank field	The field and even K_v may jump	Pin v during execution; publish a successor atomically
Dissipation and boundedness	One time-independent E_v	$\partial_t E$ has uncontrolled-sign terms	Prove dissipation per pinned segment; make no through-write claim
Asymptotic invariance	Compact invariant sublevel set	A write replaces that sublevel geometry	Apply only to an autonomous pinned segment
Multiple isolated wells	Fixed separation assumptions	Insertion can reduce old margins and curvature	Re-evaluate C0–C4 on the complete successor
Local equilibrium continuation	Nonsingular Hessian and fixed dimension	A K-changing update is not a smooth parameter path	Use IFT/Cerf analysis diagnostically; commit only after finite verification

Table 3: Notation used in the mathematical appendix.

Symbol	Definition	Role
$\boldsymbol{a}, \boldsymbol{x}_i, \boldsymbol{X}_v$	state, resident, version-v bank	physical state and persistent contents
$\boldsymbol{R}_S, \boldsymbol{Q}_S$	$\boldsymbol{R}_S \succ 0$, $\boldsymbol{Q}_S = \boldsymbol{R}_S^{-1}$	stored quadratic geometry and its inverse
$\boldsymbol{S}, \boldsymbol{z}, \boldsymbol{y}_i$	$\boldsymbol{S} = \boldsymbol{R}_S^{-1/2}$, $\boldsymbol{z} = \boldsymbol{S}^{-1}\boldsymbol{a}$, $\boldsymbol{y}_i = \boldsymbol{S}\boldsymbol{x}_i$	exact Euclidean whitening
$\eta, \beta, \boldsymbol{\sigma}$	attraction scale, inverse temperature, softmax weights	strength and concentration of associative attraction
$\boldsymbol{\Pi}, \pi_{\min}, \pi_{\max}$	positive-definite mobility and spectral slab	changes velocity, not fixed-energy equilibria
$\rho_i, d_{ij}, \gamma_{ij}$	core radius, whitened distance, directed margin	primitive C0–C4 geometry
$A_i, B_i, \kappa_*, \delta_*$	self-map bound, Jacobian bound, curvature and separation reserves	quantitative acceptance margins
$\lambda, \alpha, \boldsymbol{b}(\boldsymbol{q})$	continuation level, lens scale, query log-priors	transient read guidance removed at $\lambda = 1$

bank, and local continuation of nondegenerate equilibria under smooth energy perturbations. ALC does not discard this analysis. It exposes the assumptions on which those conclusions depend and places every read inside one interval on which they remain valid. The dissipation and invariance steps follow the classical Lyapunov–LaSalle lineage (Lyapunov, 1992; LaSalle, 1960); the new issue is that a persistent write changes the Lyapunov function itself.

Table 2 separates what is inherited from what must be added for persistent writes. This distinction matters because a write changes the function whose gradient defines the dynamics; it is not an ordinary control input to a fixed vector field.

For reference, the recurring symbols are collected below. This table also fixes a distinction that is easy to lose: $\boldsymbol{R}_S$ occurs inside the stored energy, whereas $\boldsymbol{\Pi}$ preconditions its gradient and does not change the critical-point equation while it remains invertible.

The boundedness part of the inherited argument can be seen directly from the present energy. It is included because compact sublevels are later needed by the invariance argument.

Proposition A.1 (Coercivity and continuation of pinned reads)**.** *Let* $r_{\min} = \lambda_{\min}(\boldsymbol{R}_S) > 0$ *and* $M_X = \max_j \|\boldsymbol{x}_j\|_2$*. For a nonempty finite bank,*

$$E(\boldsymbol{a}) \geq \frac{r_{\min}}{2}\|\boldsymbol{a}\|_2^2 - \eta M_X \|\boldsymbol{a}\|_2 - \frac{\eta}{\beta}\log K. \tag{18}$$

Hence E *is radially unbounded and every sublevel set is compact. If the mobility satisfies Equation 21 and is continuous in time and locally Lipschitz in state, every finite initial state generates a unique solution for the whole declared read horizon.*

Proof. For each resident, $\boldsymbol{x}_j^\top \boldsymbol{a} \leq M_X\|\boldsymbol{a}\|_2$. Therefore

$$\log \sum_{j=1}^{K} e^{\beta \boldsymbol{x}_j^\top \boldsymbol{a}} \leq \log K + \beta M_X \|\boldsymbol{a}\|_2. \tag{19}$$

The Rayleigh bound $\boldsymbol{a}^\top \boldsymbol{R}_S \boldsymbol{a} \geq r_{\min}\|\boldsymbol{a}\|_2^2$ proves Equation 18. Its positive quadratic term dominates the linear term as $\|\boldsymbol{a}\|_2 \to \infty$, so the sublevels are bounded; they are closed by continuity. Theorem 4.1 keeps a trajectory in its initial sublevel. Local Picard–Lindelöf existence plus boundedness rules out finite escape and extends the solution across the finite horizon. □

For an autonomous mobility $\boldsymbol{\Pi}(\boldsymbol{a})$, LaSalle's invariance principle then places every omega-limit set inside $\{\boldsymbol{a} : \nabla E(\boldsymbol{a}) = 0\}$. For explicitly time-varying mobility, the dissipation identity still gives

$$\int_0^T \|\nabla E(\boldsymbol{a}(t))\|_2^2 \,\mathrm{d}t \leq \frac{E(\boldsymbol{a}(0)) - \inf E}{\pi_{\min}}, \tag{20}$$

but convergence to one particular equilibrium requires additional regularity; ALC does not infer it from energy monotonicity alone.

A.2 Assumptions and Validity Domain

Fix a finite bank $\boldsymbol{X} = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_K] \in \mathbb{R}^{n\times K}$, $\boldsymbol{R}_S \in \mathbb{R}^{n\times n}$ symmetric positive definite, and scalars $\eta, \beta > 0$. Throughout a read, this bank, its ordered stable identifiers, its radii, and $\boldsymbol{R}_S$ are pinned to one immutable version. We assume that the mobility field is symmetric and satisfies

$$\pi_{\min}\boldsymbol{I} \preceq \boldsymbol{\Pi}(\boldsymbol{a}, t) \preceq \pi_{\max}\boldsymbol{I}, \qquad 0 < \pi_{\min} \leq \pi_{\max} < \infty. \tag{21}$$

The dissipation theorem is conditional on the existence of a differentiable solution. Standard local existence follows, for example, when $\boldsymbol{\Pi}(\boldsymbol{a}, t)$ is continuous in t and locally Lipschitz in $\boldsymbol{a}$, because the stored-energy gradient is smooth for a finite bank.

A broader formulation may also allow continuously moving patterns and a time-varying precision matrix. Those are not assumptions of the present theorems. Appendix A.7 derives the additional terms that appear in that setting and explains why the implemented lifecycle pins a version during execution and recertifies a complete successor at a write.

A.3 Exact Whitening Transformation

Let

$$\boldsymbol{Q}_S = \boldsymbol{R}_S^{-1}, \qquad \boldsymbol{S} = \boldsymbol{Q}_S^{1/2} = \boldsymbol{R}_S^{-1/2}, \qquad \boldsymbol{z} = \boldsymbol{S}^{-1}\boldsymbol{a}, \qquad \boldsymbol{y}_i = \boldsymbol{S}\boldsymbol{x}_i. \tag{22}$$

All square roots are the unique symmetric positive-definite square roots. Since $\boldsymbol{a} = \boldsymbol{S}\boldsymbol{z}$,

$$\boldsymbol{a}^\top \boldsymbol{R}_S \boldsymbol{a} = \boldsymbol{z}^\top \boldsymbol{S}^\top \boldsymbol{R}_S \boldsymbol{S}\boldsymbol{z} = \boldsymbol{z}^\top \boldsymbol{R}_S^{-1/2}\boldsymbol{R}_S\boldsymbol{R}_S^{-1/2}\boldsymbol{z} = \|\boldsymbol{z}\|_2^2, \tag{23}$$

$$\boldsymbol{x}_i^\top \boldsymbol{a} = \boldsymbol{x}_i^\top \boldsymbol{S}\boldsymbol{z} = (\boldsymbol{S}\boldsymbol{x}_i)^\top \boldsymbol{z} = \boldsymbol{y}_i^\top \boldsymbol{z}. \tag{24}$$

Substitution into Equation 6 gives the exact whitened energy

$$\widehat{E}(\boldsymbol{z}) = \frac{1}{2}\|\boldsymbol{z}\|_2^2 - \frac{\eta}{\beta}\log \sum_{j=1}^{K} \exp(\beta \boldsymbol{y}_j^\top \boldsymbol{z}). \tag{25}$$

No linearization or approximation has been made. In particular, $\boldsymbol{a} \leftrightarrow \boldsymbol{z}$ is bijective, so critical points and local minima correspond under the inverse transform.

The norm part of C0 is also preserved in the intended coordinates:

$$\boldsymbol{y}_i^\top \boldsymbol{Q}_S^{-1} \boldsymbol{y}_i = \boldsymbol{x}_i^\top \boldsymbol{S}\boldsymbol{R}_S\boldsymbol{S}\boldsymbol{x}_i = \|\boldsymbol{x}_i\|_2^2. \tag{26}$$

Thus C0's bound is a bound on the original stored representation even though the core geometry is evaluated after whitening.

A.4 Energy Gradient and Hessian Derivation

Write

$$s_j(\boldsymbol{z}) = \beta \boldsymbol{y}_j^\top \boldsymbol{z}, \qquad Z(\boldsymbol{z}) = \sum_k e^{s_k(\boldsymbol{z})}, \qquad \sigma_j(\boldsymbol{z}) = \frac{e^{s_j(\boldsymbol{z})}}{Z(\boldsymbol{z})}. \tag{27}$$

For an arbitrary perturbation $\mathrm{d}\boldsymbol{z}$,

$$\mathrm{d}\log Z = \frac{1}{Z}\sum_j e^{s_j}\,\mathrm{d}s_j = \sum_j \sigma_j \beta \boldsymbol{y}_j^\top \mathrm{d}\boldsymbol{z}. \tag{28}$$

Therefore

$$\nabla \widehat{E}(\boldsymbol{z}) = \boldsymbol{z} - \eta \sum_j \sigma_j \boldsymbol{y}_j. \tag{29}$$

Define the softmax-weighted mean and covariance

$$\overline{\boldsymbol{y}} = \sum_j \sigma_j \boldsymbol{y}_j, \qquad C_\sigma = \sum_j \sigma_j (\boldsymbol{y}_j - \overline{\boldsymbol{y}})(\boldsymbol{y}_j - \overline{\boldsymbol{y}})^\top. \tag{30}$$

Differentiating $\sigma_j = e^{s_j}/Z$ gives

$$\mathrm{d}\sigma_j = \sigma_j \left(\mathrm{d}s_j - \sum_k \sigma_k \mathrm{d}s_k \right) = \beta \sigma_j (\boldsymbol{y}_j - \overline{\boldsymbol{y}})^\top \mathrm{d}\boldsymbol{z}. \tag{31}$$

Consequently,

$$\mathrm{d}\overline{\boldsymbol{y}} = \sum_j \boldsymbol{y}_j\,\mathrm{d}\sigma_j = \beta C_\sigma\,\mathrm{d}\boldsymbol{z}, \tag{32}$$

$$\nabla^2 \widehat{E}(\boldsymbol{z}) = \boldsymbol{I} - \eta\beta C_\sigma. \tag{33}$$

Returning through the linear coordinate transform recovers Equation 7 and Equation 8.

It is useful to define the modern-Hopfield fixed-point map

$$T(\boldsymbol{z}) = \eta \sum_j \sigma_j(\boldsymbol{z}) \boldsymbol{y}_j = \eta \overline{\boldsymbol{y}}. \tag{34}$$

Equations 29–34 imply

$$\nabla \widehat{E}(\boldsymbol{z}) = \boldsymbol{z} - T(\boldsymbol{z}), \qquad DT(\boldsymbol{z}) = \eta\beta C_\sigma, \qquad \nabla^2 \widehat{E} = \boldsymbol{I} - DT. \tag{35}$$

Hence stationary points of the energy are exactly fixed points of T.

A.5 Local Equilibrium Continuation via the Implicit Function Theorem

The local continuation result inherited from PCAM is an application of the Implicit Function Theorem, and its scope is precise. Let $\theta \in \mathbb{R}^p$ parameterize any smooth quantity that actually enters the energy—for example a continuously perturbed resident, storage precision, or query-lens coefficient—and define

$$F(\boldsymbol{a}, \theta) = \nabla_{\boldsymbol{a}} E(\boldsymbol{a}; \theta). \tag{36}$$

Suppose $F(\boldsymbol{a}_0, \theta_0) = 0$ and $H_0 = D_{\boldsymbol{a}} F(\boldsymbol{a}_0, \theta_0) = \nabla^2_{\boldsymbol{a}} E(\boldsymbol{a}_0; \theta_0)$ is nonsingular. Then neighborhoods of θ_0 and $\boldsymbol{a}_0$ contain a unique differentiable critical-point branch $\boldsymbol{a}^*(\theta)$ satisfying

$$F(\boldsymbol{a}^*(\theta), \theta) = 0, \qquad D_\theta \boldsymbol{a}^*(\theta_0) = -H_0^{-1} D_\theta F(\boldsymbol{a}_0, \theta_0). \tag{37}$$

The formula follows by differentiating the first equality and solving the resulting linear system. If $H_0 \succ 0$, continuity of the smallest eigenvalue preserves positive definiteness in a sufficiently small neighborhood, so the continued critical point remains a strict local minimum.

Two boundaries are essential. First, an invertible mobility $\mathbf{\Pi}$ is absent from $F = \nabla E$: replacing one admissible mobility by another changes trajectory speed and direction away from equilibrium but does not move a fixed-energy critical point. Thus Equation 37 applies to parameters of E, not to mere positive-definite preconditioning. Second, appending a resident changes K_v and the defining log-sum-exp. Such a discrete update is not a smooth fixed-dimensional parameter path covered by the theorem. ALC therefore uses the IFT to explain local persistence, while the full successor verifier, not the IFT, authorizes a write.

For the current energy, the parameter derivative is explicit. If one resident follows a differentiable path $\boldsymbol{x}_r(\theta)$ while all other terms remain fixed, then

$$D_\theta F = -\eta \left[\sigma_r D_\theta \boldsymbol{x}_r + \beta \sum_j \sigma_j (\boldsymbol{x}_j - \overline{\boldsymbol{x}}) \boldsymbol{a}^\top D_\theta \boldsymbol{x}_j \right], \tag{38}$$

where only $j = r$ has a nonzero derivative and $\overline{\boldsymbol{x}} = \sum_j \sigma_j \boldsymbol{x}_j$. Substitution into Equation 37 shows why the inverse Hessian is the local sensitivity amplifier: as its smallest eigenvalue approaches zero, a small landscape perturbation can produce a large displacement of the associated equilibrium.

A.6 Cerf–Morse Theory: Fold Conditions and Local Normal Form

The IFT describes regular portions of a parameterized landscape. Cerf–Morse theory describes the generic singular events separating those portions (Cerf, 1970; Milnor, 1963). Consider a declared scalar path $s \mapsto \theta(s)$ and the family $E_s(\boldsymbol{a}) = E(\boldsymbol{a}; \theta(s))$. Write

$$F(\boldsymbol{a}, s) = \nabla_{\boldsymbol{a}} E_s(\boldsymbol{a}), \qquad H(\boldsymbol{a}, s) = D_{\boldsymbol{a}} F(\boldsymbol{a}, s). \tag{39}$$

Whenever $F = 0$ and H is nonsingular, the preceding IFT gives a unique local critical branch with constant Morse index. A possible birth–death event must therefore lie in

$$\Sigma = \{(\boldsymbol{a}, s) : F(\boldsymbol{a}, s) = 0,\ 0 \in \operatorname{spec} H(\boldsymbol{a}, s)\}. \tag{40}$$

Thus at least one Hessian eigenvalue is zero; this does not require the algebraically smallest eigenvalue to vanish.

Proposition A.2 (Generic scalar-path fold conditions)**.** *Let $(\boldsymbol{a}_0, s_0) \in \Sigma$. Suppose*

1. *the Hessian has a one-dimensional null space (corank one), $\ker H(\boldsymbol{a}_0, s_0) = \operatorname{span}\{\boldsymbol{v}\}$, with $\|\boldsymbol{v}\|_2 = 1$;*
2. *the parameter path is transverse to the singular direction, $c_s = \boldsymbol{v}^\top \partial_s F(\boldsymbol{a}_0, s_0) \neq 0$; and*
3. *the cubic coefficient $c_3 = D^3_{\boldsymbol{a}} E_{s_0}(\boldsymbol{a}_0)[\boldsymbol{v}, \boldsymbol{v}, \boldsymbol{v}]$ is nonzero.*

After eliminating the nonsingular directions, the stationary equation has the local reduced form

$$0 = c_s(s - s_0) + \frac{1}{2} c_3 \tau^2 + o(|s - s_0| + |\tau|^2), \tag{41}$$

where τ is displacement along $\boldsymbol{v}$. Thus one side of s_0 has two nearby critical points of adjacent Morse index and the other has none.

Proof. Decompose $\boldsymbol{a} - \boldsymbol{a}_0 = \tau \boldsymbol{v} + \boldsymbol{w}$ with $\boldsymbol{w} \perp \boldsymbol{v}$. The restriction of H to $\boldsymbol{v}^\perp$ is invertible because the Hessian has corank one. The IFT therefore solves the projected equation $P_{\boldsymbol{v}^\perp} F = 0$ uniquely as $\boldsymbol{w} = \boldsymbol{w}(\tau, s)$. Projecting the remaining equation onto $\boldsymbol{v}$ and Taylor expanding removes the constant and linear-in-τ terms because $F(\boldsymbol{a}_0, s_0) = 0$ and $H\boldsymbol{v} = 0$. The first surviving terms are the transverse parameter term $c_s(s - s_0)$ and the quadratic stationarity term $c_3\tau^2/2$, giving Equation 41. Their nonzero coefficients yield the standard saddle-node pair by solving for τ^2. □

This proposition does not establish that the structured family used by ALC is generic. It supplies checkable necessary local data for a fold: a critical point, Hessian corank exactly one, parameter transversality, cubic nondegeneracy, and full rank of the augmented system. Equal-critical-value crossings between distinct nondegenerate points are separate events. For a vector-valued parameter, the generic singular set is a codimension-one locus, not a finite set of isolated scalar values; no global catastrophe-theory claim is made here.

Most importantly, Cerf analysis is qualitative. It identifies how wells may appear or disappear along a smooth declared path, but it supplies none of the finite reserves m_*, κ_*, δ_* required by C0–C4. Conversely, C3 gives $\nabla^2 \widehat{E} \succeq \kappa_* \boldsymbol{I}$ throughout each accepted core, which excludes a fold inside that core for the verified snapshot. It says nothing about an arbitrary interpolation between two snapshots. The current writer therefore treats a write as a discrete transaction and certifies the complete endpoint; the current reader removes its lens before terminal acceptance. Cerf–Morse theory explains possible intervening topology but is not part of either runtime acceptance decision.

A.7 Version-Pinned Execution Semantics

The fixed-bank assumption is structural, rather than cosmetic. To see this, let both the precision and patterns vary smoothly while K remains fixed:

$$E(\boldsymbol{a}, t) = \frac{1}{2}\boldsymbol{a}^\top \boldsymbol{R}_S(t)\boldsymbol{a} - \frac{\eta}{\beta}\log \sum_j \exp\bigl(\beta \boldsymbol{x}_j(t)^\top \boldsymbol{a}\bigr). \tag{42}$$

At fixed $\boldsymbol{a}$,

$$\frac{\partial E}{\partial t} = \frac{1}{2}\boldsymbol{a}^\top \dot{\boldsymbol{R}}_S(t)\boldsymbol{a} - \eta \sum_j \sigma_j(\boldsymbol{a}, t)\, \dot{\boldsymbol{x}}_j(t)^\top \boldsymbol{a}. \tag{43}$$

Along $\dot{\boldsymbol{a}} = -\boldsymbol{\Pi}\nabla_{\boldsymbol{a}} E$, the full chain rule is

$$\frac{\mathrm{d}E}{\mathrm{d}t} = -\nabla E^\top \boldsymbol{\Pi} \nabla E + \frac{1}{2}\boldsymbol{a}^\top \dot{\boldsymbol{R}}_S \boldsymbol{a} - \eta \sum_j \sigma_j \dot{\boldsymbol{x}}_j^\top \boldsymbol{a}. \tag{44}$$

The last two terms have no fixed sign. A discrete insertion changes the dimension of the log-sum-exp and is not even represented by this derivative. The implementation therefore does not claim dissipation through a write. It executes reads against a pinned version and treats a write as a transactional transition to a separately certified successor.

A.8 Lyapunov Dissipation Proof

Proof. For a pinned bank, E has no explicit time dependence. Let $\boldsymbol{g}(\boldsymbol{a}) = \nabla_{\boldsymbol{a}} E(\boldsymbol{a}; \boldsymbol{X})$. Along any differentiable solution of Equation 9, the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}t} E(\boldsymbol{a}(t); \boldsymbol{X}) = \nabla E(\boldsymbol{a}(t))^\top \dot{\boldsymbol{a}}(t) \tag{45}$$

$$= -\boldsymbol{g}(\boldsymbol{a})^\top \boldsymbol{\Pi}(\boldsymbol{a}, t)\boldsymbol{g}(\boldsymbol{a}). \tag{46}$$

For any vector u, the Rayleigh-quotient form of Equation 21 gives $u^\top \boldsymbol{\Pi} u \geq \pi_{\min}\|u\|_2^2$. Taking $u = \boldsymbol{g}(\boldsymbol{a})$ proves

$$\frac{\mathrm{d}E}{\mathrm{d}t} \leq -\pi_{\min}\|\boldsymbol{g}(\boldsymbol{a})\|_2^2 \leq 0. \tag{47}$$

If $\boldsymbol{g}(\boldsymbol{a}_c) = 0$, then $-\boldsymbol{\Pi}(\boldsymbol{a}_c, t)\boldsymbol{g}(\boldsymbol{a}_c) = 0$ for every admissible mobility. Thus changing mobility can change the velocity away from critical points, but cannot move an exact critical point. □

Corollary A.3 (Sublevel and barrier obstruction)**.** *Let $\boldsymbol{a}(0) = \boldsymbol{a}_0$. Every differentiable trajectory remains in the stored-energy sublevel set $\{\boldsymbol{a} : E(\boldsymbol{a}) \leq E(\boldsymbol{a}_0)\}$. More generally, if a set Γ separates $\boldsymbol{a}_0$ from a target region and $\inf_{\boldsymbol{a}\in\Gamma} E(\boldsymbol{a}) > E(\boldsymbol{a}_0)$, no mobility satisfying Equation 21 can drive the trajectory into that region.*

Proof. Integrating Equation 47 gives $E(\boldsymbol{a}(t)) \leq E(\boldsymbol{a}_0)$. A continuous path from $\boldsymbol{a}_0$ to the target must intersect Γ, where its energy would be strictly greater than $E(\boldsymbol{a}_0)$, contradicting the first statement. □

This is the precise sense in which mobility-only SDRE, SLQ, or receding-horizon control cannot release an exact wrong equilibrium or cross a higher stored-energy barrier. It is not a claim that every noncritical initialization fails, nor that transient query-dependent energies cannot redirect a read.

A.9 C0–C4 Certification Conditions

For $i \neq j$, define

$$d_{ij} = \|\boldsymbol{y}_i - \boldsymbol{y}_j\|_2, \tag{48}$$

$$\gamma_{ij} = \eta(\boldsymbol{y}_i - \boldsymbol{y}_j)^\top \boldsymbol{y}_i, \tag{49}$$

$$\ell_{ij}(r) = \gamma_{ij} - d_{ij} r, \tag{50}$$

$$A_i(r) = \eta \sum_{j \neq i} d_{ij} e^{-\beta \ell_{ij}(r)}, \tag{51}$$

$$B_i(r) = \eta\beta \sum_{j \neq i} d_{ij}^2 e^{-\beta \ell_{ij}(r)}. \tag{52}$$

The directed quantity γ_{ij} is generally asymmetric: $\gamma_{ij} \neq \gamma_{ji}$. Therefore C1 is evaluated for every ordered pair, while C4 is geometrically symmetric.

For fixed constants $M, \rho_{\min}, \rho_{\max}, m_*, \epsilon_{\mathrm{SM}}, \kappa_*, \delta_*$ satisfying

$$0 < \rho_{\min} \leq \rho_{\max}, \qquad m_* > 0, \qquad 0 \leq \epsilon_{\mathrm{SM}} < \rho_{\min}, \qquad 0 < \kappa_* < 1, \qquad \delta_* > 0, \tag{53}$$

the certificate requires

$$\begin{aligned}
&\mathrm{C0}: \quad \boldsymbol{y}_i^\top \boldsymbol{Q}_S^{-1} \boldsymbol{y}_i \leq M^2, \qquad \rho_{\min} \leq \rho_i \leq \rho_{\max};\\
&\mathrm{C1}: \quad \ell_{ij}(\rho_i) \geq m_* \quad \text{for every ordered pair } i \neq j;\\
&\mathrm{C2}: \quad A_i(\rho_i) \leq \rho_i - \epsilon_{\mathrm{SM}};\\
&\mathrm{C3}: \quad B_i(\rho_i) \leq 1 - \kappa_*;\\
&\mathrm{C4}: \quad \eta d_{ij} \geq \rho_i + \rho_j + \delta_* \quad \text{for every } i \neq j.
\end{aligned} \tag{54}$$

C0 bounds the domain; C1 bounds competitor logits; C2 gives a declared self-map reserve; C3 gives contraction and curvature reserves; and C4 separates the certified cores. The following lemmas derive these roles rather than assuming them.

A.10 Soundness of Interval-Enclosed Verification

The mathematical certificate above uses real square roots and exponentials. The verifier instead keeps patterns, radii, and design constants as exact rationals. For each directed pair it computes the exact values

$$d_{ij}^2 = \|\boldsymbol{y}_i - \boldsymbol{y}_j\|_2^2, \qquad \gamma_{ij} = \eta(\boldsymbol{y}_i - \boldsymbol{y}_j)^\top \boldsymbol{y}_i, \tag{55}$$

and outward bounds

$$\underline{d}_{ij} \leq d_{ij} \leq \overline{d}_{ij}, \qquad \underline{\ell}_{ij} = \gamma_{ij} - \overline{d}_{ij}\rho_i \leq \ell_{ij}(\rho_i), \tag{56}$$

together with a rational exponential bound

$$e^{-\beta \underline{\ell}_{ij}} \leq \overline{e}_{ij}. \tag{57}$$

It then evaluates

$$\overline{A}_i = \eta \sum_{j \neq i} \overline{d}_{ij}\, \overline{e}_{ij}, \tag{58}$$

$$\overline{B}_i = \eta\beta \sum_{j \neq i} d_{ij}^2\, \overline{e}_{ij}. \tag{59}$$

Proposition A.4 (Conservative acceptance implies the real certificate)**.** *Suppose the exact-rational verifier establishes*

$$\begin{aligned}
&\underline{\ell}_{ij} \geq m_*, \qquad && \eta \underline{d}_{ij} \geq \rho_i + \rho_j + \delta_*,\\
&\overline{A}_i \leq \rho_i - \epsilon_{\mathrm{SM}}, \qquad && \overline{B}_i \leq 1 - \kappa_*,
\end{aligned} \tag{60}$$

along with C0. Then the corresponding real-valued bank satisfies C0–C4 in Equation 54.

Proof. The first inequality in Equation 56 gives $\ell_{ij}(\rho_i) \geq \underline{\ell}_{ij} \geq m_*$, which proves C1. The exponential is decreasing in its margin argument, so

$$e^{-\beta \ell_{ij}(\rho_i)} \leq e^{-\beta \underline{\ell}_{ij}} \leq \overline{e}_{ij}. \tag{61}$$

Combining this with $d_{ij} \leq \overline{d}_{ij}$ bounds the exact A_i by $\overline{A}_i$. Since d_{ij}^2 is represented exactly, the same exponential bound gives $B_i \leq \overline{B}_i$. These are C2 and C3. Finally, $d_{ij} \geq \underline{d}_{ij}$ and the second gate in Equation 60 give C4. All inequalities are one-sided in the conservative direction, so an enclosure may refuse a truly feasible bank but cannot accept one on the wrong side of these bounds. □

A.11 Directed-Margin Softmax Concentration Bound

Lemma A.5 (Directed margin inside one certified ball)**.** *If $\boldsymbol{z} \in \overline{B}(\eta \boldsymbol{y}_i, \rho_i)$, then for every $j \neq i$,*

$$(\boldsymbol{y}_i - \boldsymbol{y}_j)^\top \boldsymbol{z} \geq \ell_{ij}(\rho_i). \tag{62}$$

Proof. Insert and subtract the center $\eta \boldsymbol{y}_i$:

$$(\boldsymbol{y}_i - \boldsymbol{y}_j)^\top \boldsymbol{z} = \eta (\boldsymbol{y}_i - \boldsymbol{y}_j)^\top \boldsymbol{y}_i \tag{63}$$

$$+ (\boldsymbol{y}_i - \boldsymbol{y}_j)^\top (\boldsymbol{z} - \eta \boldsymbol{y}_i). \tag{64}$$

Cauchy–Schwarz and membership in the ball give

$$(\boldsymbol{y}_i - \boldsymbol{y}_j)^\top (\boldsymbol{z} - \eta \boldsymbol{y}_i) \geq -d_{ij} \rho_i. \tag{65}$$

Combining the two expressions yields $\gamma_{ij} - d_{ij}\rho_i = \ell_{ij}(\rho_i)$. □

Lemma A.6 (Competitor-mass bound)**.** *On the same ball,*

$$\frac{\sigma_j(\boldsymbol{z})}{\sigma_i(\boldsymbol{z})} \leq e^{-\beta \ell_{ij}(\rho_i)}, \qquad \sigma_j(\boldsymbol{z}) \leq e^{-\beta \ell_{ij}(\rho_i)}, \tag{66}$$

and hence

$$1 - \sigma_i(\boldsymbol{z}) \leq \sum_{j \neq i} e^{-\beta \ell_{ij}(\rho_i)}. \tag{67}$$

Proof. The softmax normalizer cancels in the ratio:

$$\frac{\sigma_j}{\sigma_i} = \exp\big(\beta \boldsymbol{y}_j^\top \boldsymbol{z} - \beta \boldsymbol{y}_i^\top \boldsymbol{z}\big) \tag{68}$$

$$= \exp\big(-\beta (\boldsymbol{y}_i - \boldsymbol{y}_j)^\top \boldsymbol{z}\big). \tag{69}$$

Lemma A.5 gives the first inequality. Since $\sigma_i \leq 1$, multiplying the ratio by σ_i gives the per-competitor bound. Summing $1 - \sigma_i = \sum_{j \neq i} \sigma_j$ gives Equation 67. □

A.12 C2 Self-Mapping Guarantee

Lemma A.7 (Self-map with declared reserve)**.** *If C2 holds for resident i, then*

$$T\big(\overline{B}(\eta \boldsymbol{y}_i, \rho_i)\big) \subseteq \overline{B}(\eta \boldsymbol{y}_i, \rho_i - \epsilon_{\mathrm{SM}}) \subseteq \overline{B}(\eta \boldsymbol{y}_i, \rho_i). \tag{70}$$

Proof. For any point in the ball,

$$T(\boldsymbol{z}) - \eta \boldsymbol{y}_i = \eta \sum_j \sigma_j \boldsymbol{y}_j - \eta \sum_j \sigma_j \boldsymbol{y}_i \tag{71}$$

$$= \eta \sum_{j \neq i} \sigma_j (\boldsymbol{y}_j - \boldsymbol{y}_i). \tag{72}$$

The triangle inequality and Lemma A.6 yield

$$\|T(\boldsymbol{z}) - \eta \boldsymbol{y}_i\|_2 \leq \eta \sum_{j \neq i} \sigma_j d_{ij} \tag{73}$$

$$\leq \eta \sum_{j \neq i} d_{ij} e^{-\beta \ell_{ij}(\rho_i)} = A_i(\rho_i) \leq \rho_i - \epsilon_{\mathrm{SM}}. \tag{74}$$

□

A.13 C3 Contraction and Strong-Convexity Bound

Lemma A.8 (Anchored covariance inequality)**.** *For any anchor $c \in \mathbb{R}^n$,*

$$C_\sigma = \sum_j \sigma_j (\boldsymbol{y}_j - c)(\boldsymbol{y}_j - c)^\top - (\overline{\boldsymbol{y}} - c)(\overline{\boldsymbol{y}} - c)^\top. \tag{75}$$

In particular, with $c = \boldsymbol{y}_i$,

$$0 \preceq C_\sigma \preceq \sum_{j \neq i} \sigma_j (\boldsymbol{y}_j - \boldsymbol{y}_i)(\boldsymbol{y}_j - \boldsymbol{y}_i)^\top. \tag{76}$$

Proof. Let $u_j = \boldsymbol{y}_j - c$ and $\overline{u} = \sum_j \sigma_j u_j = \overline{\boldsymbol{y}} - c$. Expanding $\sum_j \sigma_j (u_j - \overline{u})(u_j - \overline{u})^\top$ gives

$$\sum_j \sigma_j u_j u_j^\top - \overline{u}\overline{u}^\top - \overline{u}\overline{u}^\top + \overline{u}\overline{u}^\top = \sum_j \sigma_j u_j u_j^\top - \overline{u}\overline{u}^\top, \tag{77}$$

which proves Equation 75. A covariance is positive semidefinite, and the subtracted outer product is positive semidefinite, so the Loewner-order upper bound follows. When $c = \boldsymbol{y}_i$, the $j = i$ summand is zero. □

Lemma A.9 (Uniform Jacobian bound)**.** *If C3 holds for resident i, then throughout its certified ball,*

$$\|DT(\boldsymbol{z})\|_2 \leq 1 - \kappa_*, \qquad \nabla^2 \widehat{E}(\boldsymbol{z}) \succeq \kappa_* \boldsymbol{I}. \tag{78}$$

Proof. For an arbitrary unit vector v, the anchored covariance bound gives

$$v^\top C_\sigma v \leq \sum_{j \neq i} \sigma_j \left[v^\top (\boldsymbol{y}_j - \boldsymbol{y}_i) \right]^2 \tag{79}$$

$$\leq \sum_{j \neq i} \sigma_j d_{ij}^2 \tag{80}$$

$$\leq \sum_{j \neq i} d_{ij}^2 e^{-\beta \ell_{ij}(\rho_i)}. \tag{81}$$

Taking the supremum over unit v and using $DT = \eta\beta C_\sigma$ gives

$$\|DT\|_2 \leq \eta\beta \sum_{j \neq i} d_{ij}^2 e^{-\beta \ell_{ij}(\rho_i)} = B_i(\rho_i) \leq 1 - \kappa_*. \tag{82}$$

Because DT is symmetric positive semidefinite, Equation 35 gives $\nabla^2 \widehat{E} = \boldsymbol{I} - DT \succeq \kappa_* \boldsymbol{I}$. □

Corollary A.10 (Lipschitz contraction on the ball)**.** *For any $\boldsymbol{z}, \boldsymbol{z}'$ in resident i's certified ball,*

$$\|T(\boldsymbol{z}) - T(\boldsymbol{z}')\|_2 \leq (1 - \kappa_*) \|\boldsymbol{z} - \boldsymbol{z}'\|_2. \tag{83}$$

Proof. The ball is convex, so the line segment $\boldsymbol{z}_\tau = \boldsymbol{z}' + \tau(\boldsymbol{z} - \boldsymbol{z}')$ remains inside it. The fundamental theorem of calculus gives

$$T(\boldsymbol{z}) - T(\boldsymbol{z}') = \int_0^1 DT(\boldsymbol{z}_\tau)(\boldsymbol{z} - \boldsymbol{z}')\, \mathrm{d}\tau. \tag{84}$$

Apply Lemma A.9 inside the integral. □

A.14 Certified-Core Existence and Uniqueness Proof

Proof. Fix resident i. C2 and Lemma A.7 show that T maps the complete metric space $\overline{B}(\eta \boldsymbol{y}_i, \rho_i)$ into itself. C3 and Corollary A.10 make T a contraction with constant at most $1 - \kappa_* < 1$. Banach's fixed-point theorem therefore gives a unique $\boldsymbol{z}_i^* \in \overline{B}(\eta \boldsymbol{y}_i, \rho_i)$ such that

$$T(\boldsymbol{z}_i^*) = \boldsymbol{z}_i^*. \tag{85}$$

By Equation 35, $\nabla \widehat{E}(\boldsymbol{z}_i^*) = 0$.

Lemma A.9 gives $\nabla^2 \widehat{E} \succeq \kappa_* \boldsymbol{I}$ throughout the ball. For any $\boldsymbol{z}$ in the ball, Taylor's theorem with integral remainder gives

$$\widehat{E}(\boldsymbol{z}) = \widehat{E}(\boldsymbol{z}_i^*) + \nabla \widehat{E}(\boldsymbol{z}_i^*)^\top (\boldsymbol{z} - \boldsymbol{z}_i^*) \tag{86}$$

$$+ \int_0^1 (1-\tau)(\boldsymbol{z} - \boldsymbol{z}_i^*)^\top \nabla^2 \widehat{E}(\boldsymbol{z}_i^* + \tau(\boldsymbol{z} - \boldsymbol{z}_i^*))(\boldsymbol{z} - \boldsymbol{z}_i^*)\, \mathrm{d}\tau \tag{87}$$

$$\geq \widehat{E}(\boldsymbol{z}_i^*) + \frac{\kappa_*}{2} \|\boldsymbol{z} - \boldsymbol{z}_i^*\|_2^2. \tag{88}$$

Thus the fixed point is a strict local minimum and is the only stationary point in its certified ball.

For $i \neq j$, the ball centers are $\eta \boldsymbol{y}_i$ and $\eta \boldsymbol{y}_j$, whose distance is ηd_{ij}. C4 gives

$$\|\eta \boldsymbol{y}_i - \eta \boldsymbol{y}_j\|_2 \geq \rho_i + \rho_j + \delta_*. \tag{89}$$

Therefore the two closed balls are disjoint and their boundary-to-boundary distance is at least δ_*. Since this holds for every pair, each certified core has an unambiguous resident label. □

A.15 Residual Localization and Discrete Convergence

Proposition A.11 (Residual-to-equilibrium distance)**.** *For any $\boldsymbol{z}$ in resident i's certified ball,*

$$\|\boldsymbol{z} - \boldsymbol{z}_i^*\|_2 \leq \frac{\|\nabla \widehat{E}(\boldsymbol{z})\|_2}{\kappa_*}. \tag{90}$$

Proof. The Hessian lower bound implies strong monotonicity of the gradient:

$$(\nabla \widehat{E}(\boldsymbol{z}) - \nabla \widehat{E}(\boldsymbol{z}_i^*))^\top (\boldsymbol{z} - \boldsymbol{z}_i^*) \tag{91}$$

$$= \int_0^1 (\boldsymbol{z} - \boldsymbol{z}_i^*)^\top \nabla^2 \widehat{E}(\boldsymbol{z}_i^* + \tau(\boldsymbol{z} - \boldsymbol{z}_i^*))(\boldsymbol{z} - \boldsymbol{z}_i^*)\, \mathrm{d}\tau \tag{92}$$

$$\geq \kappa_* \|\boldsymbol{z} - \boldsymbol{z}_i^*\|_2^2. \tag{93}$$

The fixed-point gradient is zero. Cauchy–Schwarz bounds the left side by $\|\nabla \widehat{E}(\boldsymbol{z})\|_2 \, \|\boldsymbol{z} - \boldsymbol{z}_i^*\|_2$. Cancel the distance when it is nonzero; the zero-distance case is immediate. Thus the residual supplies the computable bound $\|\boldsymbol{z} - \boldsymbol{z}_i^*\|_2 \leq \|\nabla \widehat{E}(\boldsymbol{z})\|_2 / \kappa_*$. □

Proposition A.12 (Contraction of the certified Euler update)**.** *For $0 < h \leq 1$, define*

$$G_h(\boldsymbol{z}) = \boldsymbol{z} - h \nabla \widehat{E}(\boldsymbol{z}) = (1-h)\boldsymbol{z} + hT(\boldsymbol{z}). \tag{94}$$

On a certified ball,

$$\|G_h(\boldsymbol{z}) - G_h(\boldsymbol{z}')\|_2 \leq (1 - h\kappa_*) \|\boldsymbol{z} - \boldsymbol{z}'\|_2. \tag{95}$$

If a floating implementation incurs one-step defects δ_k, its distance from the exact discrete trajectory obeys

$$e_{k+1} \leq (1 - h\kappa_*) e_k + \|\delta_k\|_2. \tag{96}$$

Proof. Using Corollary A.10,

$$\|G_h(z) - G_h(z')\|_2 \leq (1-h)\|z - z'\|_2 + h\|T(z) - T(z')\|_2 \tag{97}$$

$$\leq [(1-h) + h(1-\kappa_*)] \|z - z'\|_2 \tag{98}$$

$$= (1 - h\kappa_*) \|z - z'\|_2. \tag{99}$$

For the perturbed update $\widehat{z}_{k+1} = G_h(\widehat{z}_k) + \delta_k$, subtract $z_{k+1} = G_h(z_k)$ and apply the same Lipschitz bound plus the triangle inequality. The certified local reader computes an outward enclosure of this recursion and refuses if the iterate leaves its named ball or exceeds its roundoff budget. □

A.16 Finite Packing Bound

C4 implies a finite capacity for fixed design constants, although it does not say that every bank below the bound is feasible. From C0,

$$\|\boldsymbol{y}_i\|_2 = \|\boldsymbol{Q}_S^{1/2}\boldsymbol{x}_i\|_2 \le M\sqrt{\lambda_{\max}(\boldsymbol{Q}_S)} =: B. \tag{100}$$

Since $\rho_i \ge \rho_{\min}$, C4 implies

$$d_{ij} \ge \frac{2\rho_{\min} + \delta_*}{\eta} =: \Delta_{\mathrm{C4}}. \tag{101}$$

Open balls of radius $\Delta_{\mathrm{C4}}/2$ around the K whitened patterns are disjoint and all lie inside $B(0, B + \Delta_{\mathrm{C4}}/2)$. Comparing n-dimensional volumes yields

$$K \le \left(1 + \frac{2B}{\Delta_{\mathrm{C4}}}\right)^n. \tag{102}$$

This is a conservative geometric upper bound. It neither supplies an insertion algorithm nor proves feasibility below the bound. We do not use a stronger C1-derived packing expression because such a derivation requires additional norm-shell assumptions not present in C0–C4.

A.17 Full-Successor Verification and Transactional Invariance

Let $\mathsf{Cert}(\mathcal{V})$ denote the complete version predicate: C0–C4 together with format, identity, model, and lineage bindings. Suppose the published head is $\mathcal{V}_v$ and $\mathsf{Cert}(\mathcal{V}_v) = \mathrm{true}$. A write is performed in three logically distinct steps:

1. construct a private successor $\widetilde{\mathcal{V}}_{v+1}$ without mutating the head;
2. recompute Cert on the complete successor, including every old–old, old–new, and new–old contribution required by C0–C4;
3. publish only through compare-and-swap against the exact predecessor content hash.

Proposition A.13 (Inductive preservation of the committed invariant)**.** *Assume sound predicate evaluation, collision-resistant content bindings, and atomic compare-and-swap. After any finite sequence of completed write attempts,*

$$\mathsf{Cert}(\mathcal{V}_{\mathrm{head}}) = \mathrm{true}. \tag{103}$$

Proof. Proceed by induction over completed attempts. The base head is certified by assumption. For the inductive step, there are three cases.

1. If construction or verification fails, no publication is attempted and the certified predecessor remains the head.
2. If verification succeeds but compare-and-swap observes a stale predecessor, publication fails and the already published head is unchanged.
3. If verification and compare-and-swap both succeed, the new head is the exact successor for which Cert was recomputed as true.

Every completed attempt therefore leaves a certified head, proving the invariant. □

The proposition is a safety statement, not a completeness or optimality statement. It does not guarantee that a feasible repair exists, that a bounded teacher or WritePFM will find one, or that an accepted placement is globally optimal. It also explains why checking only the candidate's nearest resident is insufficient: inserting one column changes its directed C1 terms with every resident and adds nonnegative terms to every old resident's C2 and C3 sums.

A.18 Before/After Curvature Diagnostic

The runtime authority is full C0–C4 recertification. The following standard perturbation bound is nevertheless useful for interpreting how much a particular write changed local curvature.

Proposition A.14 (Symmetric Hessian perturbation)**.** *Let H and $H' = H + \Delta H$ be stored-energy Hessians evaluated at a common state before and after a write. Since both are symmetric,*

$$\lambda_{\min}(H) - \|\Delta H\|_2 \le \lambda_{\min}(H') \le \lambda_{\min}(H) + \|\Delta H\|_2. \tag{104}$$

Proof. By the Rayleigh-quotient characterization,

$$\lambda_{\min}(H') = \min_{\|v\|_2=1} \left(v^\top H v + v^\top \Delta H v \right). \tag{105}$$

For every unit v, $|v^\top \Delta H v| \le \|\Delta H\|_2$. Taking the minimum gives the lower bound; applying the same argument with $H = H' - \Delta H$ gives the upper bound. □

This proposition alone does not certify a write: it compares curvature at one state, while C0–C4 establishes a uniform post-write statement on every certified ball.

A.19 Query-Lens Dissipation Proof

For fixed query scores and fixed λ, define

$$Z_\lambda(\boldsymbol{a}; \boldsymbol{q}) = \sum_j \exp\!\big(\beta \boldsymbol{x}_j^\top \boldsymbol{a} + (1-\lambda)\alpha b_j(\boldsymbol{q})\big). \tag{106}$$

The query term is independent of $\boldsymbol{a}$. Repeating Equation 28 therefore gives

$$\nabla_{\boldsymbol{a}} E_\lambda = \boldsymbol{R}_S \boldsymbol{a} - \eta \sum_j \sigma_j^\lambda \boldsymbol{x}_j = \boldsymbol{g}_\lambda. \tag{107}$$

Proof. Within one stage, λ and π_s are constant and the state follows $\dot{\boldsymbol{a}} = -\pi_s \boldsymbol{g}_\lambda$. Hence

$$\frac{\mathrm{d}}{\mathrm{d}t} E_\lambda(\boldsymbol{a}(t); \boldsymbol{q}) = \nabla E_\lambda^\top \dot{\boldsymbol{a}} = -\pi_s \|\boldsymbol{g}_\lambda\|_2^2 \le 0. \tag{108}$$

At $\lambda = 1$, the multiplier $1 - \lambda$ is zero, so $Z_1 = \sum_j \exp(\beta \boldsymbol{x}_j^\top \boldsymbol{a})$. Therefore $E_1 = E$ and $\boldsymbol{g}_1 = \boldsymbol{g}$ exactly. □

There is intentionally no monotonicity claim across a stage boundary. At a fixed physical state, changing λ^- to λ^+ changes the energy by

$$E_{\lambda^+}(\boldsymbol{a}; \boldsymbol{q}) - E_{\lambda^-}(\boldsymbol{a}; \boldsymbol{q}) = -\frac{\eta}{\beta} \log \frac{Z_{\lambda^+}(\boldsymbol{a}; \boldsymbol{q})}{Z_{\lambda^-}(\boldsymbol{a}; \boldsymbol{q})}, \tag{109}$$

whose sign is not fixed for general score vectors. The guarantee is fixed-stage dissipation plus exact recovery of the stored energy at $\lambda = 1$, followed by the local certificate.

The same limitation appears for a continuously varying lens. Since

$$\partial_\lambda E_\lambda = \frac{\eta\alpha}{\beta} \left\langle \boldsymbol{\sigma}^\lambda, \boldsymbol{b}(\boldsymbol{q}) \right\rangle, \tag{110}$$

the full derivative along $\lambda(t)$ is

$$\frac{\mathrm{d}}{\mathrm{d}t} E_{\lambda(t)}(\boldsymbol{a}(t); \boldsymbol{q}) = -\pi(t) \|\boldsymbol{g}_{\lambda(t)}\|_2^2 + \frac{\eta\alpha}{\beta} \left\langle \boldsymbol{\sigma}^{\lambda(t)}, \boldsymbol{b}(\boldsymbol{q}) \right\rangle \dot{\lambda}(t). \tag{111}$$

The second term has no fixed sign. Holding each stage fixed is therefore a mathematical condition for its Lyapunov claim, rather than an implementation convenience.

A.20 Numerical Trajectory Validation Scope

For a frozen read stage, fixed-grid classical RK4 (Butcher, 2008) approximates

$$\dot{\boldsymbol{a}} = f(\boldsymbol{a}) = -\pi_s \boldsymbol{g}_\lambda(\boldsymbol{a}) \tag{112}$$

using four vector-field evaluations per substep. NFE therefore counts evaluations of f; it is not a count of Transformer calls or a hardware-level FLOP measurement. The implementation checks finite values, the declared step grid, sampled stage-energy behavior, endpoint gradient and curvature tolerances, terminal residual, and agreement with a refined replay.

These checks provide reproducible numerical evidence and detect the failure classes represented in the test suite. They are not an interval enclosure of the complete RK4 trajectory. Exact authority is narrower: C0–C4 certifies a finite bank, and the terminal local reader uses the strong-convexity and residual bounds above inside a named certified ball. Semantic correctness is evaluated empirically and is not implied by any geometric theorem in this appendix.

B Lifecycle Protocols and Nonlinear Control Derivations

This appendix makes the operational boundary used throughout the system explicit. ALC is not a single neural network and a proposal is not a state transition. The persistent object at version v is

$$\mathcal{V}_v = (\boldsymbol{X}_v, \boldsymbol{\rho}_v, \mathcal{I}_v, \mathcal{P}_v, h_v^{\text{state}}, h_v^{\text{content}}, h_v^{\text{lineage}}), \tag{113}$$

where the hashes bind the numerical bank, payload identities, and predecessor chain. Every public operation names the version it observed. A write may publish $\mathcal{V}_{v+1}$ only after exact verification and an atomic compare-and-swap against v; a read never mutates Equation 113.

B.1 Recurrent Read Protocol

For a query $\boldsymbol{q}$, the frozen representation model produces a normalized vector in the same declared coordinate system as the residents. The scorer evaluates all residents in stable-identifier order and returns scores, answerability and margin diagnostics, and one proposed label i_q. These quantities form a *semantic hypothesis*; they are neither a certificate nor a restricted candidate set. The query-logit lens in Equation 11 still contains every resident.

The executable recurrent read is the following finite-state procedure.

1. Pin $\mathcal{V}_v$, the representation manifest, the scorer manifest, an initial physical state $\boldsymbol{a}_0$, and the numerical budget. Reject any identity, dimension, dtype, device, or finite-value mismatch.
2. Build the full-bank condition from $\boldsymbol{a}_r$, the query, every resident token, the accepted continuation level λ_r, the current bias, and the remaining horizon and step fractions.
3. Integrate the proposal-space ODE for ReadPFM from a reproducible base sample to $t = 1$. Decode a bounded continuation plan and select its next executable stage.
4. Hold the selected λ, bias, duration, and mobility fixed while advancing the Remanence state by a fixed RK4 grid. Record stage endpoints, numerical diagnostics, and the consumed budget.
5. If the stage passes its finite checks, update the observed state and accepted continuation level and return to step 2. A settling-only retry may be used when the gradient tolerance fails and budget remains. The model is never allowed to reset the physical state to a resident center.
6. At $\lambda = 1$, invoke the exact local reader on the actual numerical endpoint. Success requires that it names i_q; a different core, an outside-core endpoint, model failure, or exhausted budget yields a typed non-success outcome.

This is receding-horizon control in the operational sense that every new proposal conditions on the state produced by the previous segment. The implemented controller is not claimed to solve the global nonlinear model predictive-control problem. Algorithm 2 states the same transition system as executable pseudocode.

B.2 TRANSACTIONAL WRITE PROTOCOL

A write request contains the encoded candidate, payload identity and lineage, requested radius, movement budget, predecessor version, and the complete verification design. The trusted path evaluates inexpensive deterministic branches before learned repair:

1. **Replay or merge.** If identity and content already exist, the writer either returns an idempotent replay or verifies the declared merge.
2. **Direct admission.** Append the candidate and requested radius, reconstruct the complete successor bank, and evaluate C0–C4. If all predicates hold, construct a proposed commit.
3. **Repair.** When direct admission fails, build a full-bank WRITEPFM condition containing candidate-relative margins and verifier reserves. The model proposes a bounded center displacement and radius. Reconstruct and reverify the complete bank. Rejected proposals may be fed back as a detached recurrent carry together with new reserve diagnostics.
4. **Commit.** Bind the accepted proposal, verifier report, source manifests, payload lineage, and predecessor hash. Atomically publish only if the predecessor is still current; otherwise return an identity conflict without changing the bank.

The deterministic finite teacher searches a declared finite repair domain and can return `EXACT_FINITE_DOMAIN` or `NO_FEASIBLE_CANDIDATE_IN_DECLARED_DOMAIN`. The latter is not a proof that no continuous repair exists. Likewise, a neural rejection is not an infeasibility theorem. Algorithm 3 makes the verification and compare-and-swap boundaries explicit.

B.3 NONLINEAR CONTROL TEACHER FORMULATION

The nonlearned controller stack provides local baselines and proposal targets. This subsection derives the optimization problem that these routines approximate and marks the point at which the executable implementation stops.

Controlled mobility plant and finite-horizon objective. Let $D_1, \dots, D_m$ be fixed symmetric mobility directions and define

$$\boldsymbol{\Pi}(\boldsymbol{v}) = \boldsymbol{\Pi}_0 + \sum_{r=1}^{m} v_r D_r, \qquad f(\boldsymbol{a}, \boldsymbol{v}) = -\boldsymbol{\Pi}(\boldsymbol{v})\boldsymbol{g}(\boldsymbol{a}). \tag{114}$$

Admissible controls obey the inward spectral slab $\pi_{\min}\boldsymbol{I} \prec \boldsymbol{\Pi}(\boldsymbol{v}) \prec \pi_{\max}\boldsymbol{I}$ and remain inside a declared local state ball. For a certified stationary anchor $\boldsymbol{a}^\star$, put $\delta\boldsymbol{a} = \boldsymbol{a} - \boldsymbol{a}^\star$. The implemented continuous objective is

$$J(\boldsymbol{v}) = \frac{1}{2}\delta\boldsymbol{a}(T)^\top F \delta\boldsymbol{a}(T) + \frac{1}{2}\int_0^T \left(\delta\boldsymbol{a}^\top Q \delta\boldsymbol{a} + \boldsymbol{v}^\top C \boldsymbol{v}\right) \mathrm{d}t, \quad Q, F \succeq 0,\ C \succ 0. \tag{115}$$

The numerical implementation evaluates this cost on the original nonlinear plant, uses a fixed RK4 grid, and rejects any state, control, spectrum, or diagnostic outside its finite contract.

For orientation, Pontryagin's minimum principle gives first-order necessary conditions for an interior optimum (Pontryagin et al., 1962). With Hamiltonian

$$\mathcal{H}(\boldsymbol{a}, \boldsymbol{v}, \boldsymbol{p}) = \frac{1}{2}\delta\boldsymbol{a}^\top Q \delta\boldsymbol{a} + \frac{1}{2}\boldsymbol{v}^\top C \boldsymbol{v} + \boldsymbol{p}^\top f(\boldsymbol{a}, \boldsymbol{v}), \tag{116}$$

an extremal satisfies

$$\dot{\boldsymbol{a}} = \partial_{\boldsymbol{p}}\mathcal{H} = f(\boldsymbol{a}, \boldsymbol{v}), \qquad \boldsymbol{a}(0) = \boldsymbol{a}_0, \tag{117}$$

$$\dot{\boldsymbol{p}} = -\partial_{\boldsymbol{a}}\mathcal{H} = -Q\delta\boldsymbol{a} - A(\boldsymbol{a}, \boldsymbol{v})^\top \boldsymbol{p}, \qquad \boldsymbol{p}(T) = F\delta\boldsymbol{a}(T), \tag{118}$$

$$0 \in C\boldsymbol{v} + B(\boldsymbol{a})^\top \boldsymbol{p} + N_{\mathcal{U}}(\boldsymbol{v}), \tag{119}$$

where $\mathcal{U}$ is the spectral-feasible control set, $N_{\mathcal{U}}$ its normal cone, and the exact Jacobians are

$$A(\boldsymbol{a}, \boldsymbol{v}) = -\boldsymbol{\Pi}(\boldsymbol{v})\boldsymbol{H}(\boldsymbol{a}), \qquad B(\boldsymbol{a}) = \left[-D_1\boldsymbol{g}(\boldsymbol{a}) \quad \cdots \quad -D_m\boldsymbol{g}(\boldsymbol{a})\right]. \tag{120}$$

Equation 120 exposes the equilibrium obstruction algebraically: if $\boldsymbol{g}(\boldsymbol{a}_c) = 0$, then both $f(\boldsymbol{a}_c, \boldsymbol{v})$ and $B(\boldsymbol{a}_c)$ vanish for every admissible $\boldsymbol{v}$. No mobility optimizer, regardless of horizon, has a direction with which to leave that exact critical point.

State-dependent Riccati equation. The implementation uses an exact anchor-based state-dependent coefficient factorization before applying the Riccati approximation. By the fundamental theorem of calculus,

$$\boldsymbol{g}(\boldsymbol{a}) - \boldsymbol{g}(\boldsymbol{a}^\star) = \left[\int_0^1 \boldsymbol{H}(\boldsymbol{a}^\star + \tau\delta\boldsymbol{a})\,\mathrm{d}\tau\right]\delta\boldsymbol{a} = \overline{H}(\boldsymbol{a})\delta\boldsymbol{a}. \tag{121}$$

Since the anchor is stationary, the left side is simply $\boldsymbol{g}(\boldsymbol{a})$. Gaussian quadrature approximates the displayed integral, and the runtime rejects the factorization when its residual exceeds the declared tolerance. Freezing or predicting coefficients gives the local surrogate

$$\dot{\delta\boldsymbol{a}} \approx A(t)\delta\boldsymbol{a} + B(t)\delta\boldsymbol{v}. \tag{122}$$

For the quadratic approximation to Equation 115, the value ansatz $V(t,\delta\boldsymbol{a}) = \frac{1}{2}\delta\boldsymbol{a}^\top P(t)\delta\boldsymbol{a}$ yields the Hamilton–Jacobi–Bellman/Riccati reduction (Bellman, 1957)

$$-\dot{P} = Q + A^\top P + PA - PBC^{-1}B^\top P, \qquad P(T) = F, \tag{123}$$

$$K(t) = C^{-1}B(t)^\top P(t), \qquad \delta\boldsymbol{v}(t) = -K(t)\delta\boldsymbol{a}(t). \tag{124}$$

This is the finite-horizon differential Riccati equation used by the SDRE/DRE reference path (Beeler & Cox, 2004; Çimen, 2008). A nonsymmetric, non-positive-semidefinite, nonfinite, or high-defect Riccati path is a typed failure. The surrogate is then executed on the original nonlinear plant; surrogate agreement is reported rather than assumed. The complete numerical ordering appears in Algorithm 4.

Sequential linear-quadratic approximation. SLQ begins from a feasible nonlinear nominal rollout $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{v}})$ and linearizes the original plant along it:

$$\dot{\delta\boldsymbol{a}} = A(t)\delta\boldsymbol{a} + B(t)\delta\boldsymbol{v}, \quad A(t) = \partial_{\boldsymbol{a}} f(\bar{\boldsymbol{a}}, \bar{\boldsymbol{v}}), \quad B(t) = \partial_{\boldsymbol{v}} f(\bar{\boldsymbol{a}}, \bar{\boldsymbol{v}}). \tag{125}$$

Retaining the linear cost terms produces the value approximation $V = \frac{1}{2}\delta\boldsymbol{a}^\top P\delta\boldsymbol{a} + \boldsymbol{p}^\top\delta\boldsymbol{a} + c$. Along with Equation 123, its backward equations give

$$\boldsymbol{k} = C^{-1}(\ell_v + B^\top\boldsymbol{p}), \qquad K = C^{-1}B^\top P, \tag{126}$$

$$-\dot{\boldsymbol{p}} = \ell_x + A^\top\boldsymbol{p} - PB\boldsymbol{k}, \qquad \boldsymbol{p}(T) = F(\bar{\boldsymbol{a}}(T) - \boldsymbol{a}^\star), \tag{127}$$

$$\delta\boldsymbol{v} = -\boldsymbol{k} - K\delta\boldsymbol{a}, \qquad \ell_x = Q(\bar{\boldsymbol{a}} - \boldsymbol{a}^\star), \quad \ell_v = C\bar{\boldsymbol{v}}. \tag{128}$$

The candidate direction is evaluated by replaying the full nonlinear plant. Backtracking accepts only a spectrally feasible Armijo decrease,

$$\boldsymbol{v}^{(r+1)} = \boldsymbol{v}^{(r)} + \alpha_r\delta\boldsymbol{v}^{(r)}, \quad J(\boldsymbol{v}^{(r+1)}) \le J(\boldsymbol{v}^{(r)}) + c_1\alpha_r\langle\nabla J, \delta\boldsymbol{v}^{(r)}\rangle. \tag{129}$$

This establishes decrease for an accepted finite iteration, not global optimality (Jacobson & Mayne, 1970; Todorov & Li, 2005); see Algorithm 5 for the fail-closed replay procedure.

Receding-horizon NMPC requirements. At checkpoint r, an NMPC controller would solve a fresh constrained finite-horizon problem from the observed state $\boldsymbol{a}_r$, execute only its first segment, observe $\boldsymbol{a}_{r+1}$, reduce the remaining budget, and repeat (Mayne et al., 2000). If $\mathcal{S}$ denotes the finite-horizon solver and $\mathcal{P}_\Delta$ physical execution of one segment, the recurrence is

$$\boldsymbol{v}^\star_{r:r+N} = \mathcal{S}(\boldsymbol{a}_r, \mathcal{V}_v, \mathcal{B}_r), \qquad \boldsymbol{a}_{r+1} = \mathcal{P}_\Delta(\boldsymbol{a}_r, \boldsymbol{v}^\star_r), \qquad \mathcal{B}_{r+1} = \mathrm{Update}(\mathcal{B}_r). \tag{130}$$

Here $\mathcal{B}_r$ contains the remaining time, NFE, continuation, and retry budgets. Recursive feasibility or closed-loop stability would additionally require a terminal set, terminal controller, and a proof that the successor problem remains feasible. None of those follows from one successful SLQ rollout. The present code therefore exposes a fail-closed NMPC interface and does not relabel open-loop SLQ as feedback. The recurrent READPFM execution follows the observe–propose–execute structure of Equation 130, but its bounded proposal and terminal local-read checks are empirical and certificate-based authorities; they are not a proof that the nonlinear optimal-control problem was solved.

Table 4: Proposal, execution, and decision boundaries.

Claim	Proposer or executor	Authority
Semantic hypothesis	frozen scorer	benchmark relevance labels
Read schedule	READPFM, SDRE, or SLQ	bounded decoder
Physical endpoint	Remanence ODE + RK4	replayed numerical record
Stored-core arrival	local reader	residual, curvature, energy gates
Write repair	WRITEPFM or finite teacher	exact C0–C4 verifier
Persistent successor	atomic writer	predecessor and lineage bindings
Global optimality	none	not claimed
Semantic correctness	none	empirical evaluation only

B.4 Authority Boundaries and Typed Failures

Table 4 records which component is authoritative for each claim. All failures are retained in denominators; numerical exceptions are mapped to typed unknown or verification-failure outcomes rather than to success, infeasibility, or unreachability.

Theorem 4.1 concerns the continuous plant with a fixed landscape. RK4 traces are numerical evidence and are checked by finite step, finiteness, endpoint, and replay tests. Static C0–C4 certification and terminal local reading do not certify every point of an intervening numerical trajectory.

C Proposal Flow Architecture and Constrained Decoders

C.1 Operation-Specific Feature Tensors

The frozen representation width is $n = 384$. Each resident token concatenates its representation with eight scalars, giving an input width of 392. Padding is accompanied by a Boolean resident mask and never participates in attention or pooling. Stable identifiers determine record binding, while no positional encoding is added to resident rows.

Table 5: Per-resident scalar channels appended to the 384-dimensional vector.

ReadPFM	WritePFM
certified radius	certified radius
squared norm	squared norm
query cosine score	candidate distance
query log-prior	resident-to-candidate margin
current responsibility	candidate-to-resident margin
stored-energy logit	resident-to-candidate C1 reserve
distance from current state	candidate-to-resident C1 reserve
hypothesis indicator	C4 separation slack

ReadPFM additionally receives 15 global scalars: the accepted continuation level, current bias, remaining horizon and solver-step fractions, answerability and scorer-margin excesses, η, β, bias and mobility bounds, and gradient, curvature, and energy tolerances. WritePFM receives 12: requested radius, movement limit, minimum accepted radius, η, β, the norm bound, radius bounds, minimum directed margin, curvature and separation bounds, and the smoothing tolerance. Query or candidate vectors, current physical state, global scalars, flow time, and detached prior proposal are projected through separate channels before fusion.

C.2 Permutation-Invariant Full-Bank Conditioning

Let $R \in \mathbb{R}^{K\times 392}$ denote resident features and $M \in \{0,1\}^K$ their validity mask. A learned projection gives $H_R \in \mathbb{R}^{K\times d}$. Thirty-two learned latent vectors $L_0 \in \mathbb{R}^{32\times d}$ attend to every valid resident,

$$L_1 = L_0 + \mathrm{MHA}(\mathrm{LN}\, L_0, \mathrm{LN}\, H_R, \mathrm{LN}\, H_R; M), \tag{131}$$

followed by B_L pre-normalized latent self-attention/MLP blocks. Since the resident rows carry no positional code, a consistent permutation of resident features and mask leaves the encoded condition invariant. The attention cost is $O(KLd + L^2d)$ for $L = 32$, rather than $O(K^2d)$ resident self-attention. The exact verifier nevertheless evaluates all K residents; the bottleneck changes neural conditioning cost, not certification scope.

Proposition C.1 (Permutation invariance of the bank condition)**.** *Let P be a permutation matrix acting jointly on resident rows and the resident mask. The valid bank latents produced by Equation 131, and therefore every downstream proposal, are unchanged when (R, M) is replaced by (PR, PM).*

Proof. The resident projection is row-wise, so it maps PR to PH_R. For one latent query, permuting keys and values permutes the masked attention logits and their softmax weights by the same P. Their weighted sum is therefore unchanged: $(P\boldsymbol{\omega})^\top(PH_R) = \boldsymbol{\omega}^\top H_R$. This holds for every latent and head. All subsequent latent self-attention, normalization, and MLP operations receive the same L_1, proving the claim. Stable identifiers remain outside this invariance: they bind records and outputs but do not impose a learned row order. □

The proposal is tokenized, combined with projected operation context, and processed by B_D bidirectional DiT blocks. In block b, flow time and the condition produce AdaLN-Zero scales, shifts, and residual gates. The block applies proposal self-attention, cross-attention to the bank latents, and a GELU feed-forward network. Zero initialization of the modulation and output projections makes the initial residual velocity near zero while Xavier initialization is used for ordinary linear and attention weights. LayerNorm starts with unit scale and zero bias, and learned bank latents use standard deviation 0.02.

C.3 Constrained Read-Schedule Decoder

For $S = 6$ stages, ReadPFM emits $3S = 18$ unconstrained coordinates: $S - 1$ continuation logits r^λ, S bias logits r^b, S duration logits r^τ, and one mobility logit r^π. For a vector $r \in \mathbb{R}^m$, define a floored simplex

$$\mathrm{FS}_\epsilon(r)_j = \epsilon + (1 - m\epsilon)\frac{e^{r_j}}{\sum_k e^{r_k}}. \tag{132}$$

If λ_c is the already accepted continuation level, the remaining interval is allocated by $s^\lambda = \mathrm{FS}_{\epsilon_\lambda}(r^\lambda)$ and

$$\lambda_0 = \lambda_c, \qquad \lambda_j = \lambda_c + (1 - \lambda_c)\sum_{k=1}^{j} s_k^\lambda, \qquad \lambda_{S-1} = 1. \tag{133}$$

The implementation uses float64 decoding, a minimum share of 10^{-6}, and representable-interior checks. Remaining horizon is allocated with another floored simplex whose minimum share is 10^{-3}. Each bias and scalar mobility is mapped strictly inside its declared interval with an affine sigmoid and a machine-representable margin. The default ranges are $b \in (0.25, 8)$, $\pi \in (0.25, 3)$, and total horizon 6. The execution adapter assigns at least one RK4 substep to every emitted stage before distributing the remaining steps by deterministic largest remainder.

Proposition C.2 (Read-decoder admissibility)**.** *If $0 < \epsilon < 1/m$, then $s = \mathrm{FS}_\epsilon(r)$ satisfies $s_j \geq \epsilon$ and $\sum_j s_j = 1$. Consequently, for $0 \leq \lambda_c < 1$ the mathematical decoder produces a strictly increasing continuation from λ_c to 1, and for remaining horizon $H > 0$ it assigns every retained stage positive duration.*

Proof. Softmax coordinates are positive and sum to one. Hence $s_j = \epsilon + (1 - m\epsilon)\mathrm{softmax}(r)_j > \epsilon$ and $\sum_j s_j = m\epsilon + (1 - m\epsilon) = 1$. Each continuation increment is $(1 - \lambda_c)s_j > 0$, while each simplex duration is $Hs_j > 0$. Float64 underflow or endpoint rounding is a separate representation issue; the implementation checks representable interiors and rejects a schedule whose realized differences are not strictly positive. □

C.4 Constrained Write-Repair Decoder

WritePFM emits $n + 1 = 385$ coordinates. Let $u \in \mathbb{R}^{384}$ be the raw displacement, D the movement limit, x_c the proposed candidate, and r_ρ the radius logit. The implemented radial–tanh decoder is

$$\begin{aligned} \Delta x &= \begin{cases} 0, & \|u\|_2 = 0, \\ D(1-\epsilon_{\text{int}})\tanh(\|u\|_2)\dfrac{u}{\|u\|_2}, & \|u\|_2 > 0, \end{cases} \\ \widetilde{x} &= x_c + \Delta x, \\ \widetilde{\rho} &= \rho_{\min}^{\text{acc}} + \sigma(r_\rho)(\rho_{\text{req}} - \rho_{\min}^{\text{acc}}). \end{aligned} \tag{134}$$

The request contract requires $\rho_{\text{req}} \geq \rho_{\min}^{\text{acc}}$ and compatible global radius bounds. The decoder establishes only the movement and radius box. C0–C4 are recomputed on the complete candidate successor; no neural score can authorize the write.

Proposition C.3 (Write-decoder admissibility)**.** *For a finite raw vector and a valid request with $D > 0$ and $0 < \rho_{\min}^{\text{acc}} < \rho_{\text{req}}$, the decoder in Equation 134 satisfies*

$$\|\widetilde{x} - x_c\|_2 < D, \qquad \rho_{\min}^{\text{acc}} < \widetilde{\rho} < \rho_{\text{req}}. \tag{135}$$

Proof. The zero branch has zero displacement. Otherwise normalization gives

$$\|\Delta x\|_2 = D(1-\epsilon_{\text{int}})\tanh(\|u\|_2) < D, \tag{136}$$

because $0 < 1 - \epsilon_{\text{int}} < 1$ and $0 < \tanh(\|u\|_2) < 1$. For finite r_ρ, $0 < \sigma(r_\rho) < 1$, so the affine radius map lies strictly between its ordered endpoints. Runtime next-representable-value checks preserve these open inequalities in float64 or reject the input. □

C.5 Conditional Flow-Matching Objective and Recurrent Training

For a teacher proposal $\boldsymbol{\xi}^\star$ and independently sampled $\boldsymbol{\xi}_0 \sim \mathcal{N}(0, \boldsymbol{I})$, training samples $t \sim \mathcal{U}[0,1]$ and uses

$$\boldsymbol{\xi}_t = (1-t)\boldsymbol{\xi}_0 + t\boldsymbol{\xi}^\star, \qquad \boldsymbol{u}^\star = \boldsymbol{\xi}^\star - \boldsymbol{\xi}_0. \tag{137}$$

The implementation validates finite rank-two tensors and supplies t as a batch column. The flow loss uses equal mass for physical coordinate blocks so that the single mobility or radius coordinate is not overwhelmed by the larger schedule or displacement block. If $\mathcal{G}_1, \ldots, \mathcal{G}_B$ partition the proposal coordinates, the declared diagonal metric is

$$\mathcal{L}_{\text{flow}} = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{B}\sum_{b=1}^{B}\frac{1}{|\mathcal{G}_b|}\sum_{d\in\mathcal{G}_b}\left(\boldsymbol{v}_{\theta,d}(\boldsymbol{\xi}_{t_i}, t_i, \mathcal{C}_i) - \boldsymbol{u}_{i,d}^\star\right)^2. \tag{138}$$

For ReadPFM the four blocks are continuation, bias, duration, and mobility; for WritePFM they are displacement and radius. A typed outcome cross-entropy term is added with unit weight. Stage A trains without a previous clean estimate. Stage B first produces one clean estimate, detaches it, and conditions the next prediction on that value. This prevents backpropagation through the recurrent rollout while training the model to correct its own proposals. Algorithm 1 gives the corresponding Stage-A, Stage-B, and proposal-sampling order.

At inference, fixed-grid RK4 solves $\mathrm{d}\boldsymbol{\xi}/\mathrm{d}t = \boldsymbol{v}_\theta(\boldsymbol{\xi}, t, \mathcal{C})$ for 16 proposal steps in frozen validation. Recurrent execution then updates the condition with the observed plant endpoint or verifier reserves; it does not feed a simulated memory state back as if it had been executed.

C.6 Model Capacity Configurations

All profiles use 32 bank latents, 384-dimensional external representations, dropout 0.1, and separate parameters for the two operations. Table 6 reports exact trainable counts from the frozen validation manifest.

The H8/H16 pair isolates head width at fixed parameter count; D12 isolates depth at fixed width. These are development profiles rather than named model families or claims of optimality.

Table 6: The six architecture profiles used in the paired sweep.

Profile	Width	FFN	Heads	Latent blocks	DiT blocks	Read/Write params
5M	192	768	4	2	4	5.044M / 5.190M
12M	256	1024	4	2	6	12.128M / 12.322M
38M-H8	384	1536	8	4	8	37.931M / 38.221M
38M-H16	384	1536	16	4	8	37.931M / 38.221M
53M-D12	384	1536	8	4	12	52.713M / 53.004M
67M	512	2048	8	4	8	67.155M / 67.542M

D Reference Algorithms and Typed Functional Semantics

This appendix connects the preceding equations to the training and inference cycles used by ALC. The algorithms are normative at the level of ordering, state transitions, and authority boundaries; tensor batching and kernel-level details are omitted. The Haskell fragments are typed reference specifications, chosen to make immutability and failure branches explicit. The production implementation is Python/PyTorch, so these listings are not claimed as a second implementation. Their names denote the abstract operators defined in this paper rather than implementation-specific APIs.

D.1 Conditional Flow-Matching Training and Sampling

Both proposal models use conditional flow matching on the straight path $\boldsymbol{\xi}_t = (1-t)\boldsymbol{\xi}_0 + t\boldsymbol{\xi}^\star$, following the continuous flow-matching construction of Lipman et al. (2023). Recurrent Stage B uses a detached clean prediction as self-conditioning, following the fixed-point training motivation of Helbling et al. (2026); it never backpropagates through physical execution.

Algorithm 1 PFM training and bounded proposal sampling

(a) Stage-A/Stage-B update

1. $(\mathcal{C}, \boldsymbol{\xi}^\star, y) \sim \mathcal{D}$; $\boldsymbol{\xi}_0 \sim \mathcal{N}(0, \boldsymbol{I})$; $t \sim \mathcal{U}[0,1]$.
2. $\boldsymbol{\xi}_t \leftarrow (1-t)\boldsymbol{\xi}_0 + t\boldsymbol{\xi}^\star$; $\boldsymbol{u}^\star \leftarrow \boldsymbol{\xi}^\star - \boldsymbol{\xi}_0$.
3. $c \leftarrow \varnothing$ (Stage A), or $\mathrm{sg}[\mathrm{FlowSolve}_\theta(\boldsymbol{\xi}_0', \mathcal{C})]$ (Stage B).
4. $(\widehat{\boldsymbol{u}}, \widehat{y}) \leftarrow f_\theta(\boldsymbol{\xi}_t, t, \mathcal{C}, c)$.
5. $\theta \leftarrow \mathrm{AdamW}[\mathrm{BlockMSE}(\widehat{\boldsymbol{u}}, \boldsymbol{u}^\star) + \mathrm{CE}(\widehat{y}, y)]$; reject nonfinite updates.

(b) Bounded proposal sampling

1. Validate $(\theta, \mathcal{C}, \mathrm{dtype}, \mathrm{device}, \mathrm{manifest})$.
2. Draw or replay $\boldsymbol{\xi}_0 \sim \mathcal{N}(0, \boldsymbol{I})$.
3. $\boldsymbol{\xi}_1 \leftarrow \mathrm{RK4}(\dot{\boldsymbol{\xi}} = f_\theta(\boldsymbol{\xi}, t, \mathcal{C}, c), 0, 1)$.
4. Require finite state and matching solver diagnostics.
5. Return $\mathrm{Decode}(\boldsymbol{\xi}_1)$ and outcome logits.

Algorithm 1 separates three spaces that must not be conflated. Flow integration moves a proposal variable; the bounded decoder maps that variable into a legal control box; only the Remanence plant or write verifier determines the resulting physical outcome. $\mathrm{BlockMSE}$ is defined in Equation 138, and ODE integration remains entirely in proposal space.

D.2 Recurrent Read Control with Query Lensing

The query lens changes the transient energy, while mobility changes only the gradient-flow metric. Equations 110 and 111 show why each lens segment is held fixed: the continuous $\dot{\lambda}$ term has no definite sign. The numerical stage uses classical RK4, whose four vector-field evaluations per substep define the reported NFE convention (Butcher, 2008).

Algorithm 2 Recurrent ReadPFM control with fixed-stage query lensing

1. Pin $(\mathcal{V}_v, i_q, \boldsymbol{a}_0, \text{manifest}, \text{seed}, \mathcal{B}_0)$.
2. $(r, \lambda_r, \boldsymbol{a}_r) \leftarrow (0, 0, \boldsymbol{a}_0)$.
3. **while** $\lambda_r < 1$ and $\mathcal{B}_r > 0$ **do**
   (a) $\mathcal{C}_r \leftarrow \mathrm{Condition}(\mathcal{V}_v, \boldsymbol{a}_r, \lambda_r, i_q, \mathcal{B}_r)$.
   (b) $(\lambda^+, b, \tau, \pi) \leftarrow \mathrm{DecodeRead}(\mathrm{ReadPFM}(\mathcal{C}_r))$.
   (c) $\boldsymbol{a}^+ \leftarrow \mathrm{RK4}(\dot{\boldsymbol{a}} = -\pi \boldsymbol{g}_\lambda(\boldsymbol{a}; \boldsymbol{q}), \tau)$.
   (d) Require stage checks; set $(\boldsymbol{a}_{r+1}, \lambda_{r+1}, \mathcal{B}_{r+1}) \leftarrow (\boldsymbol{a}^+, \lambda^+, \mathcal{B}_r - \tau)$.
4. Require $\lambda_r = 1$; return $\mathrm{LocalRead}(\mathcal{V}_v, \boldsymbol{a}_r, i_q)$.

The algorithm never replaces the plant endpoint with a resident center and never restricts the lens to a retrieved subset. Every recurrent proposal sees the complete pinned bank, and the final decision is made after the query bias has vanished exactly.

D.3 Certified Write Transaction

WritePFM is a repair proposer inside a transactional write. It cannot waive a failed geometric predicate, infer infeasibility from a search failure, or publish over a changed predecessor.

Algorithm 3 WritePFM repair, exact verification, and conditional publication

1. Validate request and pin predecessor $(\mathcal{V}_v, h_v)$.
2. If identity matches, return verified replay/merge.
3. $\mathcal{V}' \leftarrow \mathrm{DirectSuccessor}(\mathcal{V}_v, \text{request})$; return $\mathrm{CAS}(h_v, \mathcal{V}')$ if C0–C4 pass.
4. $\mathcal{C} \leftarrow \mathrm{Reserves}(\mathcal{V}')$; $(\Delta \boldsymbol{c}, \widetilde{\rho}) \leftarrow \mathrm{DecodeWrite}(\mathrm{WritePFM}(\mathcal{C}))$.
5. $\widetilde{\mathcal{V}} \leftarrow \mathrm{Successor}(\mathcal{V}_v, \Delta \boldsymbol{c}, \widetilde{\rho})$.
6. $\chi \leftarrow \mathrm{Verify}_{\mathrm{C0:C4}}(\widetilde{\mathcal{V}})$ using exact rationals and outward enclosures.
7. If χ rejects, return typed rejection and retain $\mathcal{V}_v$.
8. Bind $(\widetilde{\mathcal{V}}, \chi, h_v, \text{lineage})$; return $\mathrm{CAS}(h_v, \widetilde{\mathcal{V}})$.

D.4 SDRE/DRE and SLQ Numerical Procedures

The SDRE construction is a state-dependent local quadratic approximation (Beeler & Cox, 2004; Çimen, 2008); SLQ is a trajectory-local iteration in the differential dynamic-programming lineage (Jacobson & Mayne, 1970; Todorov & Li, 2005). Neither procedure changes the equilibrium obstruction because both actuate only the mobility channel.

Algorithm 4 Finite-horizon SDRE/DRE proposal and nonlinear replay

1. Validate anchor, Hessian, dimensions, costs, and mobility slab.
2. At each knot, compute $\overline{H}(\boldsymbol{a}) = \int_0^1 H(\boldsymbol{a}^\star + \tau(\boldsymbol{a} - \boldsymbol{a}^\star))\,\mathrm{d}\tau$ by declared Gaussian quadrature.
3. Require $\|\overline{H}(\boldsymbol{a})(\boldsymbol{a} - \boldsymbol{a}^\star) - \boldsymbol{g}(\boldsymbol{a})\|$ below the factorization tolerance.
4. $A \leftarrow -\boldsymbol{\Pi}_0 \overline{H}$; $B = [-D_1\boldsymbol{g}, \ldots, -D_m\boldsymbol{g}]$.
5. $P \leftarrow \mathrm{DRE}(A, B, C, F)$ backward from $P(T) = F$; require symmetry, PSD, finiteness, and defect tolerance.
6. $K \leftarrow C^{-1}B^\top P$; $\delta\boldsymbol{v} \leftarrow -K\delta\boldsymbol{a}$; enforce the slab.
7. Return $\mathrm{Replay}_{\mathrm{plant}}(\boldsymbol{v})$ and diagnostics.

Algorithm 5 SLQ improvement with nonlinear Armijo acceptance

1. $(\bar{\boldsymbol{a}}^{(0)}, \bar{\boldsymbol{v}}^{(0)}) \leftarrow \mathrm{Rollout}_{\mathrm{plant}}(\text{feasible schedule})$.
2. **for** $r = 0, \ldots, R-1$ **do**
   (a) $(A_r, B_r) \leftarrow (\partial_{\boldsymbol{a}} f, \partial_{\boldsymbol{v}} f)$ along the nominal trajectory.
   (b) $(P, \boldsymbol{p}, K, \boldsymbol{k}) \leftarrow \mathrm{BackwardRiccati}(A_r, B_r)$.
   (c) $\delta\boldsymbol{v} \leftarrow -\boldsymbol{k} - K\delta\boldsymbol{a}$; require a descent direction.
   (d) **for** $\alpha \in (1, 1/2, 1/4, \ldots)$, enforce the slab, replay the plant, and accept the first Armijo step.
   (e) If none is accepted, return line-search failure.
3. Return the final open-loop schedule and diagnostics (global certificate $=$ `false`).

An actual NMPC controller would invoke a finite-horizon solver again after every executed segment and would require recursive-feasibility and terminal-set arguments (Mayne et al., 2000). Algorithms 4 and 5 do not provide those guarantees; a dedicated receding-horizon control entry point therefore fails closed.

D.5 Haskell Reference Semantics

The first listing presents only the mathematical state transition. The bank and query are already closed over by the supplied evolution operator, so no storage, tensor, or framework interface appears. Only the transient state advances, and every failed precondition is represented by `Either`.

Listing 1: Schematic Haskell semantics for one recurrent read stage.

```
data ReadError = InvalidStage | NonFinite | NoBudget

data Control = Control
  { continuation :: Double
  , bias         :: Double
  , duration     :: Double
  , mobility     :: Double
  }

data ReadState a = ReadState
  { point     :: a
  , progress  :: Double
  , remaining :: Budget
  }
```

```
advance
  :: (Control -> a -> Either ReadError a)
  -> ReadState a -> Control
  -> Either ReadError (ReadState a)
advance evolve current control = do
  ensure (continuation control > progress current) InvalidStage
  budget' <- debit (remaining current) control
  endpoint <- evolve control (point current)
  ensure (finite endpoint) NonFinite
  pure ReadState
    { point = endpoint
    , progress = continuation control
    , remaining = budget'
    }
```

The write sketch exposes the abstract commit boundary. A certified successor can still fail publication when another writer has advanced the version; no mutation occurs before the atomic comparison succeeds.

Listing 2: Schematic Haskell semantics for certified publication.

```
data WriteError e = Invalid e | Rejected e | VersionChanged
data WriteResult v = Replayed v | Published v

publish
  :: IO v
  -> (v -> r -> Either (WriteError e) (Either v p))
  -> (v -> p -> Either (WriteError e) v)
  -> (v -> v -> IO Bool)
  -> r
  -> IO (Either (WriteError e) (WriteResult v))
publish readVersion propose certify atomicSwap request = do
  base <- readVersion
  case propose base request of
    Left err             -> pure (Left err)
    Right (Left replay) -> pure (Right (Replayed replay))
    Right (Right proposal) ->
      case certify base proposal of
        Left err -> pure (Left err)
        Right successor -> do
          ok <- atomicSwap base successor
          pure $ if ok
            then Right (Published successor)
            else Left VersionChanged
```

Finally, the training sketch records the detached recurrence using only distribution sampling, interpolation, flow integration, and loss operators. The explicit `detach` boundary distinguishes recurrent conditioning from backpropagation through the entire refinement chain.

Listing 3: Schematic Haskell form of detached recurrent flow training.

```
stageB
  :: Flow -> Example -> Seed
  -> Either TrainError (Loss, Gradient)
stageB flow sample seed =
  let (zCarry, seed1) = gaussian seed  (target sample)
      (zBase,  seed2) = gaussian seed1 (target sample)
      carry = detach
            $ solve flow zCarry (context sample) Nothing
      time = uniform seed2
```

```
        zTime = interpolate zBase (target sample) time
        velocityTarget = target sample - zBase
        prediction = field flow zTime time (context sample) (Just
      carry)
        loss = blockMetric (velocity prediction) velocityTarget
             + classification (logits prediction) (label sample)
  in finiteGradient flow loss
```

These listings intentionally omit tensor layouts, batching, storage formats, random-generator serialization, framework calls, and implementation-specific identifiers. Their purpose is to expose immutability, gradient boundaries, and the failure branches that prevent a proposal from becoming a physical or persistent operation.

E Experimental Data and Evaluation Protocols

E.1 Attractor-Landscape Figure Construction

Figure 1 is an explanatory two-dimensional slice, not an empirical result. We set $\boldsymbol{R}_S = \boldsymbol{I}$, $\eta = 1$, $\beta = 6$, $x_1 = (-1.4, 0)$, and $x_2 = (1.4, 0)$. Writing $r = 1.4$ gives

$$E(a) = \tfrac{1}{2}\|a\|_2^2 - \frac{\eta}{\beta}\log\sum_{j=1}^{2} e^{\beta x_j^\top a}, \qquad \nabla E(a) = \begin{pmatrix} a_1 - \eta r \tanh(\beta r a_1) \\ a_2 \end{pmatrix}. \tag{139}$$

The origin is a saddle because the Hessian eigenvalues there are $1 - \eta\beta r^2 = -10.76$ and 1. The minima are $(\pm m, 0)$, where $m = \eta r \tanh(\beta r m)$, and $a_1 = 0$ is an invariant separatrix. We use a 121×121 grid on $[-2, 2]^2$ and fixed-grid RK4 with step 0.02 for 240 steps from $(-0.22, -1.55)$ and $(0.22, -1.55)$. No query lens or learned controller contributes to the figure.

E.2 Synthetic Proposal Dataset Construction

The source collection contains approximately 770,000 arXiv abstracts. A frozen encoder maps source text to normalized 384-dimensional vectors; the teacher pipeline then constructs bank-level control problems rather than treating abstracts as independent supervised examples. The resulting read and write training lanes each contain 20,590 records from 124 banks with $K = 64$. Bank identity, state, content, representation, design, source lineage, and proposal manifests are bound into every record.

Splits are assigned at the bank-lineage level before trajectories or repairs are generated. Consequently, residents, derived states, and write histories from one base lineage cannot be divided between training and evaluation merely by relabeling stable identifiers. Arrays are stored in a non-executable format with recorded shape, dtype, byte order, and partition coverage. Publication is atomic and never overwrites an existing dataset version.

Read targets contain the teacher schedule, initial state, scorer hypothesis, full bank, numerical diagnostics, and typed outcome. Write targets contain the rejected candidate, complete predecessor bank, finite-domain repair when one exists, exact verifier result, and the branch label. Accepted teacher examples are replayed before serialization. Malformed input, identity conflict, numerical failure, exhausted search, and absence of a witness in the declared finite domain remain distinct statuses.

E.3 PFM Training Configuration

We train six architectures independently for each operation. Every model receives 50,000 optimizer updates at global batch size 256: 25,000 Stage-A plain flow-matching updates and 25,000 Stage-B updates with one detached self-conditioned rollout. AdamW uses learning rate 3×10^{-4}, $(\beta_1, \beta_2) = (0.9, 0.999)$, zero weight decay, 500 warmup steps, gradient clipping at 1.0, and EMA decay 0.9999. Runs use BF16, fused AdamW, compiled models, seed zero, dropout 0.1, and numerical and metric checks every 100 steps. No validation result controls stopping.

The complete sweep therefore contains 600,000 optimizer updates and 153.6M example exposures. Six isolated A10 containers each execute one read and one write run with eight CPUs and 32 GiB

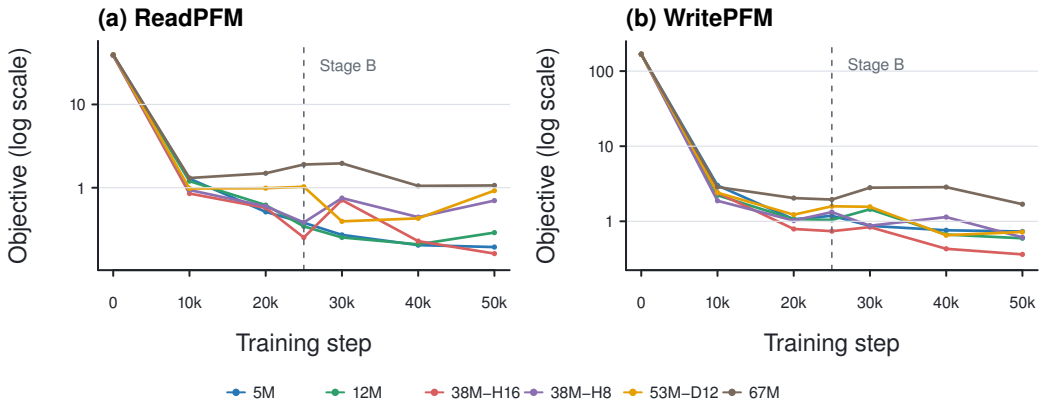

Figure 5: **Training dynamics across six capacities per operation.** Curves show authenticated 10k-step milestones; the dashed line marks the switch from plain flow matching to detached recurrent self-conditioning. Loss scales differ between read and write and are not compared across panels.

host memory. Observed runtime ranges from 1.68 to 3.15 hours per model, mean GPU utilization from 21.6% to 83.1%, and maximum device memory from 1.68 to 6.24 GiB of 23 GiB. The smaller models are launch/input limited, so these numbers do not support a hardware-efficiency claim.

E.4 Frozen Model-Selection Protocol

Validation is a separate, versioned run on the `val` partition. It loads EMA weights, uses stateless PCG64 Gaussian bases with seed 260915, deterministic CUDA, batch size 256, and a fixed-grid RK4 solver with 16 proposal steps. One recurrent self-conditioning pass is used. The six paired profiles are grouped into six A10 containers, with one read and one write job per container. The selection metrics are outcome accuracy and the coordinate-weighted proposal RMSE; they do not execute the proposals through the physical plant or exact write verifier.

This protocol selects the 37.9M-H16 ReadPFM (89.6% outcome accuracy; 4.10 weighted RMSE) and 5.2M WritePFM (96.7%; 3.82). Model selection and the subsequent closed-loop cohorts are deliberately separated: a lower vector error need not preserve the branch or terminal core that matters to ALC.

E.5 LongMemEval-S Lifecycle Evaluation

The persistent experiment uses all 500 questions of the cleaned LongMemEval-S split: 470 answerable and 30 abstention cases. Its histories contain 23,867 chronological sessions and 124,334 round-level write attempts. The source snapshot is immutable and validated before preprocessing. Text is encoded with the fixed `sentence-transformers/all-MiniLM-L6-v2` release, attention-mask mean pooling, 512-token truncation, L2 normalization, and quantization denominator 2^{20}. The observed embedding dimension is 384.

Every question produces a final record; a seed-committed, question-type stratified cohort of 100 questions additionally produces records at 25%, 50%, 75%, and 100% of its write stream. This yields 800 snapshot arrays and 4,000 lane runs. Evaluator-only evidence identifiers are excluded from the inference records and checked by a 4,000-record firewall audit. Recall metrics use only the 470 answerable questions; abstention uses its own 30-question denominator.

The semantic lane in this archived run uses the frozen scorer and native query-logit continuation. It does *not* execute a trained ReadPFM checkpoint; the archived aggregate explicitly marks the FRM lane as not produced. Accordingly, LongMemEval-S supports the persistent ALC lifecycle, semantic-selection, and physical-lensing claims. Learned ReadPFM improvement is reported only from the separate controlled synthetic cohort.

E.6 Controlled-System Read, Write, and Compute Cohorts

The read-control cohort contains 56 trajectories over eight held-out banks. Twenty-four cases are teacher-feasible and define the denominator for the fixed-lens versus recurrent-controller comparison. Bank, initial state, scorer target, numerical integrator, and terminal local-reader thresholds are held fixed; only the proposal schedule changes.

The write cohort contains 32 candidates rejected by direct exact admission over eight held-out banks. Each candidate is passed through WritePFM and then through the same complete-bank C0–C4 verifier. We report geometric admissibility separately from typed authorization. The finite-domain teacher is an upper-bound diagnostic within its declared search domain, not a global continuous optimum.

The cost study uses eight 384-dimensional queries against one replayed eight-resident bank. Its original schedule has six stages and 192 RK4 substeps per stage. The fitted schedule has two stages and 160 substeps per stage, preserving step size 0.125. It is selected on this same cohort, then checked at doubled resolution and on 40 deterministic adversarial starts. Consequently it is an engineering ablation, not held-out evidence for a universal inference schedule.

F Extended Experimental Results and Ablations

F.1 PFM Capacity Sweep

Table 7 reproduces the authenticated run-level summary. Accuracy is the mean outcome-head accuracy over the final ten logged checkpoints; gradient norm is measured before clipping over the same window. Read and write losses have different coordinate metrics and must not be compared across lanes.

Table 7: All continuous-PFM training runs. Each row uses one seed and 50,000 updates.

Model	Initial	Final	Minimum	Last-10	Acc.	Grad. norm
Read 5.0M	39.069	0.194	0.102	0.465	99.8	10.8
Read 12.1M	39.125	0.289	0.106	0.407	97.9	27.7
Read 37.9M-H8	38.697	0.701	0.256	0.739	80.4	450.1
Read 37.9M-H16	38.693	**0.162**	0.111	0.552	99.6	24.8
Read 52.7M-D12	39.631	0.924	0.243	0.807	77.8	342.2
Read 67.2M	39.472	1.066	0.646	1.712	70.4	12156.7
Write 5.2M	167.309	0.732	0.555	0.748	99.1	36.0
Write 12.3M	167.910	0.594	0.446	0.685	99.1	50.1
Write 38.2M-H8	168.285	0.611	0.355	0.753	98.6	42.4
Write 38.2M-H16	168.257	**0.362**	0.284	0.533	98.6	422.6
Write 53.0M-D12	168.097	0.721	0.383	0.596	98.2	55.2
Write 67.5M	168.187	1.689	0.997	1.978	95.0	72.4

The final minibatch favors the H16 models, while frozen validation selects the H16 read model and the much smaller 5.2M write model. The 38M head-count ablation is particularly informative: H8 and H16 have identical parameter counts, yet their late read behavior differs sharply. The 52.7M and 67.2M read models also show large pre-clipping gradients and weaker late accuracy. This evidence supports capacity-dependent optimization sensitivity rather than monotonic gains from scale.

F.2 LongMemEval-S Retrieval and Persistence Metrics

The write stream accepts 117,476/124,334 attempts (94.5%): 117,431 direct admissions, 45 merges, and 6,858 policy rejections. All required evidence survives ingestion for 470/470 answerable questions. The final scorer obtains 386/470 Recall@1, 442/470 Recall@5, and 451/470 Recall@10. Their Wilson 95% intervals are $[78.4, 85.3]$, $[91.5, 95.8]$, and $[93.8, 97.4]$ percentage points, respectively.

The native lens reaches a named stored core for 459/500 questions. On the answerable subset the rate is 431/470=91.7% with Wilson interval $[88.9, 93.9]$. It returns an evidence-bearing certified terminal

Table 8: Semantic retrieval on 470 answerable LongMemEval-S questions. These rows diagnose selection; they are not independent physical-control methods.

Selection lane	Recall@1	Recall@5	Recall@10
BM25 diagnostic	86.2	94.9	97.4
Raw MiniLM diagnostic	82.8	94.0	95.7
ALC frozen scorer	82.1	94.0	96.0

core for 361/470=76.8%, interval $[72.8, 80.4]$. The gap from scorer Recall@1 is a downstream refusal effect: in the paired records, 25 queries succeed at scoring but not at the certified terminal, and none show the reverse event (two-sided exact McNemar $p = 5.96 \times 10^{-8}$). Thus the present lens does not rerank a missed semantic hypothesis and can refuse a scorer hit.

Table 9: LongMemEval-S outcomes by answerable question type.

Type	n	Recall@1	Reached core	Evidence core
Knowledge update	72	90.3	95.8	87.5
Multi-session	121	79.3	91.7	74.4
Single-session assistant	56	100.0	96.4	96.4
Single-session preference	30	66.7	93.3	66.7
Single-session user	64	78.1	90.6	73.4
Temporal reasoning	127	78.0	87.4	68.5

The 100-question dynamic cohort contributes 400 checkpoint rows. Evidence first becomes retrievable at the 25%, 50%, 75%, and 100% checkpoints for 45, 25, 14, and 9 questions; seven are never retrievable. Mean rank drift is 0.357 over 210 comparable samples, and seven later-checkpoint records lose a previously available retrieval. This is direct evidence that writes change future access and that final-snapshot evaluation alone hides lifecycle behavior.

F.3 Closed-Loop PFM Control Outcomes

On 24 teacher-feasible read targets, recurrent ReadPFM achieves a 1.25-fold exact-terminal-success rate relative to the frozen query-lensing schedule when all other execution components are held fixed. On 32 direct-write rejects, WritePFM produces 11 geometrically admissible repairs and nine fully authorized successors, i.e., 28.1% exact write recovery. The finite teacher finds a witness for all 32 in its declared domain. These small, bank-disjoint cohorts establish that learned proposals can change exact physical outcomes; they do not estimate broad semantic generalization.

F.4 Inference Cost and NFE Ablation

Table 10: Same-cohort continuation cost fit. Both schedules preserve all eight terminal labels and certificates.

Measurement	Original	Cost-fit
Continuation stages	6	2
RK4 substeps per stage	192	160
NFE per query	4,608	1,280
Endpoint Hessians, total	48	16
CPU time, eight queries	2.521 s	1.354 s
Maximum energy increase	8.88×10^{-16}	8.88×10^{-16}
Maximum residual bound	1.02×10^{-15}	7.64×10^{-16}

The 72.2% NFE reduction becomes a 46.3% measured CPU reduction because exact certification, Hessian checks, and terminal reading remain fixed costs. On 40 deterministic adversarial starts, both schedules produce the same 23 target arrivals, 16 continuation refusals, and one settling-budget

outcome. Since the shorter schedule was chosen on the same eight-query cohort, the ablation supports a cheaper implementation point only.

G Reproducibility Specification and Artifact Provenance

G.1 Dataset, Protocol, and Checkpoint Identifiers

Table 11 records the human-readable dataset, cohort, and model identifiers needed to interpret the experiments. The anonymous supplement will provide the corresponding machine-readable configurations and provenance records.

Table 11: Principal frozen experimental artifacts.

Artifact	Frozen identifier
Read control cohort	56 trajectories; 8 bank-disjoint banks; 24 teacher-feasible targets
Write repair cohort	32 verifier-rejected candidates; 8 bank-disjoint banks
Capacity sweep	12 runs; 6 capacities per operation; 50,000 updates
Selected ReadPFM	37.9M-H16 EMA checkpoint selected by frozen validation
Selected WritePFM	5.2M EMA checkpoint selected by frozen validation
Persistent evaluation	Cleaned LongMemEval-S; 500 queries; 124,334 write attempts
Representation	`all-MiniLM-L6-v2`; 384 dimensions

The supplementary record binds every table row to the corresponding split, configuration, and checkpoint used for evaluation. This preserves traceability without turning the typeset appendix into an implementation inventory.

G.2 Deterministic Numerical Configuration

Training uses seed zero for model and data order. Frozen validation uses stateless PCG64 normal samples with seed 260915 and deterministic CUDA. The proposal ODE and physical plant use fixed-grid RK4; adaptive tolerances do not silently change the number of field evaluations. Plans record the stage grid, substep allocation, dtype, device, model configuration, and execution environment. Float decoding enforces finite values and representable open intervals; exact write verification converts accepted candidates to the rational domain before evaluating C0–C4.

Each stored bank version binds its predecessor, resident ordering, state, content, design, representation, payload lineage, and journal transition. Dataset splits bind base-snapshot and write-stream lineage in addition to the derived state hash. Recorded tensors and array artifacts are defensively copied or serialized so later in-place mutation cannot change authenticated evidence.

Verification targets observable invariants rather than successful execution alone: permutation and padding isolation, bounded decoded controls, monotone fixed-stage evolution, exact lineage binding, and preservation of the predecessor on every rejected write. All failure modes remain explicit in the recorded denominators.

G.3 Claim-to-Artifact Traceability

Table 12: Minimum evidence required to reproduce each reported claim.

Claim	Required artifacts
PFM learnability	training manifest, 12 metric logs, checkpoints, curve script
Model selection	validation protocol, base seed, EMA checkpoint identifiers, per-run metrics
Read improvement	bank-disjoint cohort, fixed baseline, recurrent traces, local reads
Write recovery	rejected inputs, PFM repairs, complete-bank verifier reports
LongMemEval recall	pinned qrels, records, scorer manifest, denominator audit
Persistent safety	journals, predecessor chain, replay log, fresh C0–C4 reports
NFE reduction	both schedules, RK4 counts, terminal labels, adversarial replay

No result in this paper should be reproduced from a plotted percentage alone. The typed rows and their manifests are the primary evidence; figures and tables are deterministic views of those records.