SELECTING SPECTRAL MEMORY FROM THE ORIGINAL GRADIENT

Anonymous authors
Paper under double-blind review

ABSTRACT

Selecting a spectral memory on the gradient used to fit it creates a dependence cost; independent evaluation avoids this dependence only by fitting candidates on noisier views. We keep every memory on the original observation and aggregate its polar action using Stein's unbiased risk estimate (SURE), conditional on the past. We establish the rank-boundary regularity and sequential risk comparison needed for this nonlinear setting. For strictly rectangular matrices, a predictable safety action makes the selected mixture continuously differentiable even at fitted rank loss, although individual corrected scores can have infinite variance. This justifies an exact, computable selection-response term and an expected-risk oracle against the unperturbed candidate rules under predictable feedback. The response is retained explicitly, not assumed small. Spectrum-saturated bounds connect the comparator risk to noise and memory bias without a lower bound on the true singular values. The leading-one oracle applies to the mixture and randomized isometric output; deterministic polar rounding has a factor-two guarantee. Across 272 fixed-plan synthetic streams, the rounded rule improves on one-way splitting in most cells, while symmetric splitting remains competitive and two zero/near-zero second-scale configurations incur substantial adaptation costs. The results identify when original-data fitting is useful and what dependence must still be charged.

1 INTRODUCTION

A long gradient history reduces observation noise but can lag behind a changing descent direction. For a spectral update, this tradeoff is geometric: the update is an isometry, and a weak singular direction can be poorly identified yet contribute little to current descent. We study *performance selection* among memory rules, rather than requiring an adaptive algorithm to reproduce a stationary baseline's actions exactly.

Let one noisy gradient be observed before the update,

$$G_t = S_t + \sigma Z_t, \qquad Z_t \sim N(0, I_{mr}), \qquad \mathcal{D}(S, P) = \|S\|_* - \langle S, P\rangle_F, \quad \|P\|_{\mathrm{op}} \le 1. \tag{1}$$

The entries of Z_t are fresh independent Gaussians with known $\sigma > 0$. The latent signal S_t may depend on past observations and actions. The loss measures the first-order decrease missed relative to the best operator-norm-constrained update, not a finite-step training objective.

Evaluating every memory on the observation that generated it is optimistic. If the candidates contain $\mathrm{polar}(G_t)$, then

$$\langle G_t, \mathrm{polar}(G_t)\rangle_F = \|G_t\|_* \ge \langle G_t, P\rangle_F \quad \text{for every contraction } P. \tag{2}$$

Uncorrected current gain therefore favors the noisiest, unaveraged fit.

Our approach is a safeguarded conditional SURE aggregate over spectral memories (Figure 1). Each candidate uses the original current sample and raw history. Its polar divergence corrects fitting optimism; differentiating the current-data weights identifies the remaining selection response. This response need not equal action disagreement, because the candidates themselves depend on the sample. A predictable safety action suppresses fitted rank-boundary singularities without changing any fitted candidate. The resulting comparison is to the actual original-data memory rules, rather than their higher-variance split-view surrogates.

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.

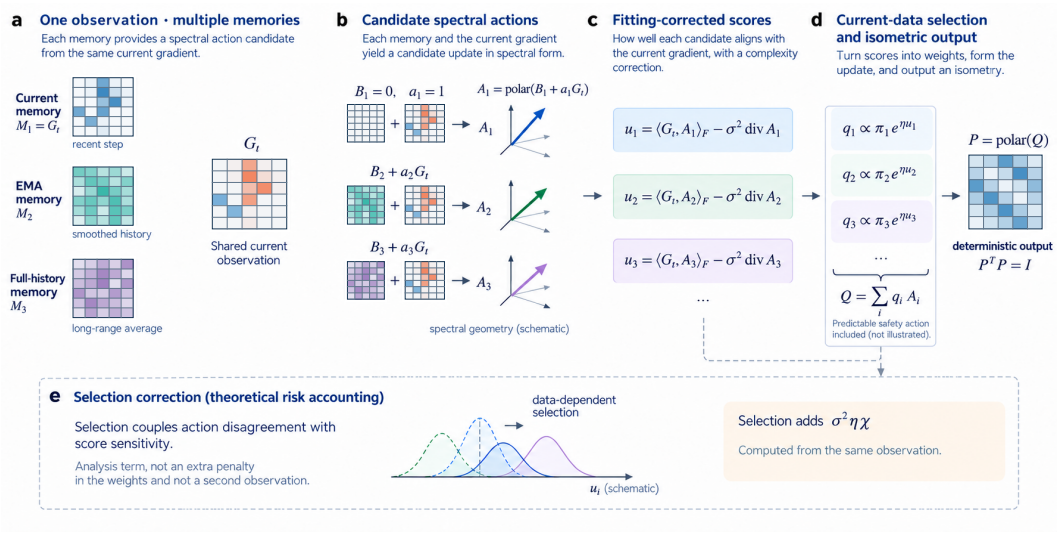

Figure 1: One observation, two dependence corrections. (a,b) Current, EMA and full-history memories use the same unmodified G_t. (c) The fitting correction enters each score. (d) Current-data weights form Q; the deterministic output is $P = \mathrm{polar}(Q)$. (e) $\sigma^2 \eta \chi$ accounts for selection in the analysis, not as an extra weight penalty. Matrix textures, geometry and density profiles are schematic; the predictable safety action is included but not illustrated.

Contributions. The central technical step is to establish inverse-spectrum integrability for arbitrary conditional means and a continuously differentiable selected map at empirical rank loss (Appendix B). This makes the aggregation identity valid for polar memories even when individual scores have infinite variance (Theorem 2). Combining it with a causal fixed-share potential gives an expected-risk oracle against original-data comparator paths, with an explicit signed response price (Theorem 3). Spectrum-saturated bounds then relate those comparators to noise, drift, and weak directions (Section 4). The experiments distinguish this risk accounting from deterministic rounding and compare both one-way and symmetric splitting (Section 5).

Relation to SURE and aggregation. Stein's identity (Stein, 1981), spectral divergence formulas (Candès et al., 2013), and polar differentiation (Gawlik & Leok, 2016) supply the classical starting point. Our weights are *exactly* conditional SURE exponential weights, and fixed-norm candidate rankings under descent loss coincide with squared-error rankings (Section 2; Appendix J). Dalalyan & Salmon (2012) prove sharp aggregation inequalities without splitting for affine estimators; their Appendix A.2, Lemma 2 also gives a general smooth-aggregate risk identity. The selection-sensitivity algebra is therefore not itself new. SURE tuning can leave excess optimism (Tibshirani & Rosset, 2019); corrected bounds for linear estimators are given by Cauchois et al. (2021). Here the additional work is to justify that accounting for rank-singular polar maps and carry it through predictable feedback and changing comparator paths, retaining the response rather than importing an affine-estimator bound.

Spectral updates and online selection. Li et al. (2026) study denoising before orthogonalization, and Jiang et al. (2026) develop efficient spectral matrix online learning. Gaussian coupled views provide independent fitting/evaluation observations with variances $(1 + \rho^2)\sigma^2$ and $(1 + \rho^{-2})\sigma^2$ (Oliveira et al., 2024); we retain the symmetric construction as a control. Our potentials use established exponential weighting, fixed share, and temperature adaptation (Herbster & Warmuth, 1998; de Rooij et al., 2014). Unlike observed-loss guarantees such as SMART's FTL comparison (Banerjee et al., 2024), the present protocol reveals a noisy gradient before acting but withholds latent loss. The conditional Stein step translates the observed-score potential into an expected-risk comparison on that same trajectory.

2 TWO KINDS OF OPTIMISM IN AN UNSPLIT OBSERVATION

We take $m > r \geq 1$; an unequal-sided matrix can be transposed to this convention. The strict inequality is substantive for *unregularized* polar scores, as shown below. Let $\mathcal{F}_{t-1}$ contain the past. Assume S_t is $\mathcal{F}_{t-1}$-measurable, Z_t is independent of it, and $\mathbb{E}\|S_t\|_F^2 < \infty$. A deterministic finite signal path satisfies these conditions. The candidate memories have the affine-current form

$$M_{t,i} = B_{t,i} + a_{t,i}G_t, \qquad A_{t,i} = \operatorname{polar}(M_{t,i}) \in \operatorname{St}(m,r), \quad i = 1, \ldots, N, \tag{3}$$

where $B_{t,i}$ and $a_{t,i} > 0$ are fixed given the past. We use fixed-kernel memories in the implementation: current gradient, full history, and three EMAs. No sample is reserved for validation. The coefficients need not be equal across candidates; full history has current gain $1/t$. A positive rescaling of M and a changes neither the action nor any score below. Predictable primary isometries may also be included: assign them zero divergence and score gradient equal to the action. They require no polar regularity argument; this special case is used in the fixed-library benchmark.

Fitting optimism. For a full-rank $m \times r$ matrix M with singular values s_j, define

$$\kappa(M) = (m-r)\sum_{j=1}^{r}\frac{1}{s_j} + 2\sum_{j<k}\frac{1}{s_j + s_k}. \tag{4}$$

This is the divergence of the polar map. It involves sums, not differences, of positive singular values, so repeated values cause no spectral-gap denominator. Write $d_{t,i} = a_{t,i}\kappa(M_{t,i})$. The raw-data score is

$$u_{t,i} = \langle G_t, A_{t,i}\rangle_F - \sigma^2 d_{t,i}. \tag{5}$$

Lemma 1 (Original-variance candidate score)**.** *For the candidates in* (3)*,* $\mathbb{E}[u_{t,i} \mid \mathcal{F}_{t-1}] = \mathbb{E}[\langle S_t, A_{t,i}\rangle_F \mid \mathcal{F}_{t-1}]$*. This holds without a positive lower bound on the true singular values and with arbitrary repeated true spectra.*

Stein's formula here needs integrability of the inverse singular values, which Appendix B proves for $m > r$ before justifying the identity by smooth approximation. In the minimal rectangular case the score need not have finite variance, so the analysis does not assume sub-Gaussianity as a substitute.

A predictable safety action. Include one additional action $A_{t,0} \in \operatorname{St}(m,r)$ that is fixed before the current observation, with $d_{t,0} = 0$ and $u_{t,0} = \langle G_t, A_{t,0}\rangle_F$. Its positive prior mass prevents every candidate's score from tending to minus infinity at the same fitted rank boundary. The implementation uses the previous full-history action, or a fixed isometry at initialization. It is not a true-signal oracle and does not replace the raw candidates. Its base mass is $(T+1)^{-2}$.

Selection optimism. For predictable positive prior π_t and temperature $\eta_t > 0$, set

$$q_{t,i} = \frac{\pi_{t,i}e^{\eta_t u_{t,i}}}{Z_t^{\mathrm{ew}}}, \qquad Q_t = \sum_{i=0}^{N} q_{t,i}A_{t,i}, \qquad \chi_t = \sum_{i=0}^{N} q_{t,i}\langle A_{t,i} - Q_t, \nabla_{G_t}u_{t,i}\rangle_F. \tag{6}$$

All terms defining χ_t are computable from the same observation and raw memory states, with the past, prior, and temperature held fixed. This is a risk-accounting term; it does not enter the weight update.

Exact SURE correspondence. Since $\|A_{t,i}\|_F^2 = r$, the conditional SURE for estimating S_t satisfies

$$\begin{aligned}\widehat{R}_{t,i}^{\mathrm{SURE}} &= \|G_t - A_{t,i}\|_F^2 + 2\sigma^2 d_{t,i} - mr\sigma^2 \\ &= \|G_t\|_F^2 + r - mr\sigma^2 - 2u_{t,i}, \qquad q_{t,i} \propto \pi_{t,i}e^{-\eta_t \widehat{R}_{t,i}^{\mathrm{SURE}}/2}.\end{aligned}$$

Thus the weight rule is a SURE aggregate, not a different unbiased-risk principle. The fixed candidate norms also make descent-loss and squared-error rankings equivalent. Appendix J derives the relation for the mixture, whose norm is not fixed.

Algorithm 1 Unsplit spectral-memory selection

Require: Known σ, dimensions $m > r$, horizon T; raw memory bank and base prior ω

1: **for** $t = 1, \dots, T$ **do**
2: Form predictable η_t, π_t using (10)–(11).
3: Observe the single current gradient G_t; update every raw memory.
4: Compute $A_{t,i}$ and scores (5); include the predictable safety action.
5: Set q_t, Q_t by (6); accumulate v_t from (9).
6: Output A_{t,I_t} with $I_t \sim q_t$, or deterministically output $\mathrm{polar}(Q_t)$.
7: **end for**

Theorem 2 (Safeguarded polar risk accounting)**.** *Under the preceding assumptions and the safeguarded priors of Algorithm 1,*

$$\mathbb{E}[\langle G_t - S_t, Q_t \rangle_F \mid \mathcal{F}_{t-1}] = \sigma^2 \mathbb{E}\left[\sum_i q_{t,i} d_{t,i} + \eta_t \chi_t \mid \mathcal{F}_{t-1} \right], \tag{7}$$

$$\mathbb{E}\left[\sum_i q_{t,i} u_{t,i} - \sigma^2 \eta_t \chi_t \mid \mathcal{F}_{t-1} \right] = \mathbb{E}[\langle S_t, Q_t \rangle_F \mid \mathcal{F}_{t-1}]. \tag{8}$$

The identities remain valid when the true signal is rank deficient. The selected map, unlike an individual inverse-spectrum score, has an integrable derivative at fitted rank boundaries.

The product rule gives $\operatorname{div} Q_t = \sum_i q_{t,i} d_{t,i} + \eta_t \chi_t$, the polar specialization of the smooth-aggregation identity discussed above. The substantive regularity step is that near a candidate's smallest singular value $s = 0$, its weight decays as $\exp(-c/s)$ against the predictable safety action, suppressing its derivative singularities. Appendix B gives the complete argument, including integrability under feedback.

Let

$$v_t = \sum_i q_{t,i} \|A_{t,i} - Q_t\|_F^2 = r - \|Q_t\|_F^2, \qquad \psi_t = \chi_t - v_t. \tag{9}$$

When actions are independent of the current sample, $\nabla u_i = A_i$ and $\chi_t = v_t$. For fitted polar actions, ψ_t measures the candidates' response to the sample, which Figure 3 resolves against independent quadrature.

3 An oracle for the original candidate rules

Let a base prior have mass $\omega_0 = \epsilon_0 \in (0, 1)$ on the safety action and equal masses $\omega_i = (1 - \epsilon_0)/N$ on the primary candidates. Put $L = \log[N/(1 - \epsilon_0)]$ and $\gamma = 1/(T + 1)$. We use

$$\eta_t = \frac{\sqrt{L}}{\sigma \sqrt{r + \sum_{s<t} v_s}}. \tag{10}$$

At $t = 1$, $\pi_1 = \omega$. At later times, with $k_t = \eta_t / \eta_{t-1} \le 1$,

$$\widetilde{\pi}_{t,i} = \frac{\omega_i^{1-k_t} q_{t-1,i}^{k_t}}{\sum_j \omega_j^{1-k_t} q_{t-1,j}^{k_t}}, \qquad \pi_{t,i} = (1 - \gamma)\widetilde{\pi}_{t,i} + \gamma \omega_i. \tag{11}$$

The power correction is part of the algorithm, not an optional implementation detail. All fitted actions still incorporate current G_t; predictability concerns the temperature and prior, not a one-step delay of the candidate gradients.

For a deterministic index path $a_{1:T} \in \{1, \dots, N\}^T$ with K constant runs, define

$$B_a^{\mathrm{raw}} = \mathbb{E} \sum_t \mathcal{D}(S_t, A_{t,a_t}), \qquad C_K = KL + (K - 1)\log(1/\gamma) + (T - K)\log \frac{1}{1 - \gamma}. \tag{12}$$

The index path is selected in the analysis, not by inspecting the current noise, and an infimum over such paths is outside the expectation. The actions in B_a^{raw} use exactly the original G stream, which is why the full-history bound (19) carries σ^2 rather than the $2\sigma^2$ of a split fitting view.

Theorem 3 (Original-data performance oracle). *Let $R_Q = \mathbb{E}\sum_t \mathcal{D}(S_t, Q_t)$. Algorithm 1 satisfies, for every fixed comparator path,*

$$\boxed{R_Q \le B_a^{\mathrm{raw}} + C_K \mathbb{E}[\eta_T^{-1}] + \mathbb{E}\sum_t \left[\sigma^2 \eta_t \chi_t - \frac{\mathrm{KL}(q_t\|\pi_t)}{\eta_t}\right].} \tag{13}$$

The integrand of the final expectation is observable without the true signal, and this signed selection price can be negative. This is an expected-risk comparison, not a realized-data risk certificate. The bound covers deterministic signal paths and predictable feedback with the stated moments.

The proof combines the pathwise exponential-weight potential with Theorem 2. Specifically,

$$\sum_t \langle q_t, u_t\rangle \ge \sum_t u_{t,a_t} - C_K/\eta_T + \sum_t \mathrm{KL}(q_t\|\pi_t)/\eta_t. \tag{14}$$

This algebra holds for unbounded realized scores. Individual comparator scores are unbiased by Lemma 1; the selected score incurs exactly $\sigma^2\eta_t\chi_t$. No concentration inequality is applied to the inverse-singular-value scores. Full proofs and all quantifiers are in Appendix C.

Corollary 4 (Disagreement plus response). *Write $\bar{v} = \mathbb{E}\sum_t v_t$ and $J_T = \sigma^2\mathbb{E}\sum_t \eta_t[\psi_t]_+$. Then*

$$R_Q \le B_a^{\mathrm{raw}} + c_K\sqrt{r+\bar{v}} + J_T, \qquad c_K = \sigma\left[\frac{C_K}{\sqrt{L}} + (1+\sqrt{2})\sqrt{L}\right]. \tag{15}$$

For $K = 1$, $C_1 \le L + 1$. No noise inflation multiplies the candidate risk or its variance.

The response budget J_T is not controlled by the temperature schedule alone. The exact oracle (13) retains signed response and the negative KL term; Table 2 compares its empirical accounting with mixture loss. Computability does not imply a small remainder, and (15) gives no universal $O(\sigma\sqrt{rT\log N})$ rate for arbitrary fitted memories.

The actual update norm. Sampling a candidate gives an isometry at every step and expected loss exactly R_Q. For deterministic output $P_t = \mathrm{polar}(Q_t)$,

$$\mathcal{D}(S_t, P_t) \le 2\mathcal{D}(S_t, Q_t). \tag{16}$$

To see this, write $Q = PK$, $0 \preceq K \preceq I$; $2Q - P = P(2K - I)$ is a contraction, so $\mathcal{D}(S,P) = 2\mathcal{D}(S,Q) - \mathcal{D}(S, 2Q - P)$. The factor two is sharp and it applies to our empirical main column, which uses deterministic rounding; mixture, sampled, and rounded losses are reported separately.

Delaying the selection. Using $Q_t^{\mathrm{pre}} = \sum_i \pi_{t,i}A_{t,i}$ removes selection optimism but pays $\mathbb{E}\sum_t \mathrm{KL}(\pi_t\|q_t)/\eta_t$ instead: current selection exchanges a mixability cost for a sensitivity cost. Appendix D interpolates the two.

4 Spectral efficiency and its boundaries

For $S_t = U_tH_t$, $U_t \in \mathrm{St}(m,r)$, $H_t \succeq 0$, the matrix disagreement is $\mathsf{V}_t = I - Q_t^\top Q_t \succeq 0$, with $\mathrm{tr}\,\mathsf{V}_t = v_t$. An exact decomposition gives

$$\mathcal{D}(S_t, Q_t) = \frac{1}{2}\mathrm{tr}(H_t\mathsf{V}_t) + \frac{1}{2}\|(Q_t - U_t)H_t^{1/2}\|_F^2. \tag{17}$$

It does not assume commuting matrices or distinct singular values. The task charges disagreement according to its current signal contribution, not according to the worst direction's inverse singular value.

For any analysis threshold $\mu > 0$, put $W_\mu = \mathbb{E}\sum_t \mathrm{tr}(\Pi_{H_t<\mu}\mathsf{V}_t)$. Then $\bar{v} \le 2R_Q/\mu + W_\mu$. Substituting in Corollary 4, with $B' = B_a^{\mathrm{raw}} + J_T$, yields

$$R_Q \le B' + \frac{c_K^2}{\mu} + c_K\sqrt{r + W_\mu + \frac{2B'}{\mu} + \frac{c_K^2}{\mu^2}}. \tag{18}$$

Fast comparison requires both controlled response J_T and little weak-direction disagreement W_μ; (18) charges them through B' and W_μ, respectively. The threshold is an analysis device, not an algorithm input. Rank deficiency alone determines neither budget.

The comparator uses the original information. For a fixed signal S with positive singular values λ_j, the unperturbed full-history candidate obeys

$$B_{\text{full}}^{\text{raw}} \le \sum_{t=1}^{T} \sum_{j:\lambda_j>0} \min\left\{2\lambda_j, \frac{(m+r)\sigma^2}{t\lambda_j}\right\}. \tag{19}$$

A symmetric-dilation proof appears in Appendix E. It charges unstable weak directions only by their own contribution. Crucially, the numerator is σ^2, not $2\sigma^2$ from a split fitting view. The oracle still pays adaptation cost; (19) alone is not the selector's risk bound.

For any deterministic normalized memory kernel $h_{t,i,s}$ on a deterministic signal path, let $B_{t,i}^{\text{bias}} = \sum_{s\le t} h_{t,i,s}S_s - S_t$ and $k_{t,i} = \sum_s h_{t,i,s}^2$. For a positive singular triple (u_j, v_j, λ_j) of S_t,

$$\mathbb{E}\mathcal{D}(S_t, A_{t,i}) \le \sum_{j:\lambda_j>0} \min\left\{2\lambda_j, \frac{\|B_{t,i}^{\text{bias}}v_j\|^2 + \|(B_{t,i}^{\text{bias}})^\top u_j\|^2 + (m+r)\sigma^2 k_{t,i}}{\lambda_j}\right\}. \tag{20}$$

Together with (13), this connects memory length, drift, observation noise, and spectral contribution without fitting on an inflated stream. It is not applied to data-selected kernels as if their coefficients were fixed. Matrix bias can overcharge harmless PSD amplitude changes; the exact comparator risk in (13) remains the sharper object.

A matching lower bound on a separated library. For a fixed library of $N \ge 4$ isometries separated by squared Frobenius distance at least $c_0 r$, some constant signal aA_i forces every one-observation procedure to incur $\Omega(\sigma\sqrt{rT\log N})$ loss, although the correct library action has zero loss. Here fitted responses vanish, $\chi_t = v_t$, and our oracle matches this order. Appendix F gives the entropy-testing proof. The result does not assert that every library is separated, that N can be arbitrarily large at fixed dimension, or that switching among data-fitted EMAs has this same minimax rate.

Why the regularity restrictions matter. Even with $G \sim N(0, I_2)$ and $A(G) = G/\|G\|$, the unbiased score $u = \|G\| - 1/\|G\|$ has infinite variance. The analysis therefore cannot assume bounded or sub-Gaussian score differences. With one such fitted action and safety action e_1, if q is the fitted weight and θ its angle to e_1, then

$$\chi = q(1-q)(1-\cos\theta)(2+\sigma^2/\|G\|^2), \qquad v = 2q(1-q)(1-\cos\theta). \tag{21}$$

The response correction cannot be bounded by a universal constant times disagreement pointwise. The safety weight suppresses the singular region and makes the selected correction integrable. For square matrices an additional issue arises: the scalar sign has derivative zero away from zero, yet nonzero Gaussian optimism. Its boundary contribution cannot be omitted. Our theorem is restricted to strictly rectangular matrices. Appendix B exhibits the missing contribution for scalar sign; it does not derive a general square-matrix boundary formula.

5 Validation: original-data fitting and the price of adaptation

The fixed final plan comprises 272 synthetic streams in 16 cells. Five core generators have 32 seeds each at $T = 512$: a strong constant signal, common-polar noncommuting PSD amplitudes, three polar segments, weak-direction reactivation, and continuous rotation. A longer stationary cell and two cells with zero/10^{-4} second singular values have 16 seeds each. Eight structural cells at 8×4 and 16×8 have eight seeds each, testing all-column PSD changes, unequal change times, and multiple rotating planes. Known latent signals permit direct evaluation of (1); these are controlled mechanism tests, not training benchmarks. Four core generators are retained byte-for-byte from the preceding package. Development used separate seeds 89010–89013; the implementation and final plan were then fixed without retuning or pooling development results (Appendix H).

The rule uses five raw memories, safety mass $(T+1)^{-2}$, and (10)–(11). Controls are the raw candidate bank, one-way and symmetric Gaussian splitting, predictable weights on current raw candidates, and uncorrected scores. All methods observe the same G_t; auxiliary noise appears only in split controls, with $\rho = 1$ and their original disagreement-temperature schedule. The main Raw column uses deterministic polar rounding; mixture and sampled outputs are reported separately. Appendix H specifies all controls, including the split rules' five-action libraries versus Raw's additional safety action.

Signal	n	**Raw**	Full (G)	Split	Symmetric	Best EMA
Static, strong	32	0.435	0.320	0.768	0.427	0.746
Common polar / PSD	32	0.057	0.027	0.092	0.049	0.070
Three polar segments	32	0.599	151.839	1.029	0.612	1.295
Weak reactivation	32	0.649	77.004	0.710	0.569	0.387
Continuous rotation	32	0.383	102.387	0.638	0.404	0.351
Static, $T = 2048$	16	0.397	0.307	0.811	0.429	2.301
Zero second scale	16	1.242	0.118	0.902	0.550	0.255
Second scale 10^{-4}	16	1.293	0.168	0.958	0.601	0.306

Table 1: Mean cumulative latent descent loss. Raw is the deterministic polar of the mixture, not the leading-one randomized rule. Split and Symmetric use one actual observation, perturbed only inside their fitting/evaluation views. Rows have $m = 3, r = 2, T = 512$, except the marked long row. Best EMA is the smallest *cell mean* of .5/.8/.95, not a per-time oracle.

What the controls establish. Across Tables 1 and 3, Raw has lower mean loss than one-way splitting in 14 of 16 cells, including all eight wider cells, but lower mean loss than symmetric splitting in only 7 of 16. These are descriptive counts, not significance tests or independent cross-cell trials. Symmetric splitting is therefore a substantive control, not a uniformly weaker method. Paired 95% Student intervals in Appendix Figure 5 show the uncertainty without multiplicity adjustment.

On the strong stationary cell, full-history loss is 0.320 using G versus 0.608 using the split fitting view. Raw reaches 0.435, compared with 0.768 for one-way and 0.427 for symmetric splitting. At $T = 2048$ the corresponding selector losses are 0.397, 0.811, and 0.429 (Appendix Figure 4). Keeping original-precision memories removes the forced inflation of each fitted view, but does not imply superiority to symmetric recombination. On the 8×4 and 16×8 static cells, Raw is near raw full history (0.606 vs 0.562; 2.042 vs 1.988), while one-way splitting gives 1.054 and 3.801.

Tracking and immediate selection. On three segments Raw has loss 0.599, compared with 1.029 for one-way splitting, 0.612 for symmetric splitting, and 1.295 for the best fixed EMA. Figure 2 shows weight moving toward short memories after changes and back toward longer memories within segments. Making only the weights predictable raises loss to 1.655, although every candidate still sees current G_t. Removing candidate correction instead gives losses 24.271 on the strong stationary cell and 2.134 on continuous rotation, consistent with (2). These ablations concern the scoring and selection policy; χ_t remains a diagnostic, not an additional algorithmic penalty.

Signal	Rule	$\widehat{B}$	$\widehat{A}_1$	$\widehat{B} + \widehat{A}_1$	Mixture	Rounded
Static, strong	Full	0.320	0.510	0.830	0.712	0.435
Three segments	EMA	1.295	-0.265	1.030	0.751	0.599
Weak reactivation	EMA	0.387	1.152	1.539	0.794	0.649
Continuous rotation	EMA	0.351	0.122	0.473	0.457	0.383
Zero second scale	Full	0.118	6.496	6.614	1.760	1.242

Table 2: Relating the oracle to the evaluated output. $\widehat{B}$ is a reference-rule mean; $\widehat{A}_1$ estimates the *complete* fixed-comparator remainder in (13), including complexity and signed KL/response accounting. Entries combine the reported means in Tables 1, 4, and 6; sums use their displayed precision. EMA denotes the cell-mean-selected Best EMA benchmark. These are descriptive estimates, not finite-sample risk certificates.

How informative is the oracle? Table 2 compares the full accounting with the mixture to which Theorem 3 applies. On rotation, the reported reference mean plus remainder is $0.351 + 0.122 = 0.473$, close to mixture loss 0.457. At zero second scale it is $0.118 + 6.496 = 6.614$, far above 1.760. Thus the accounting can be informative, but is not uniformly tight.

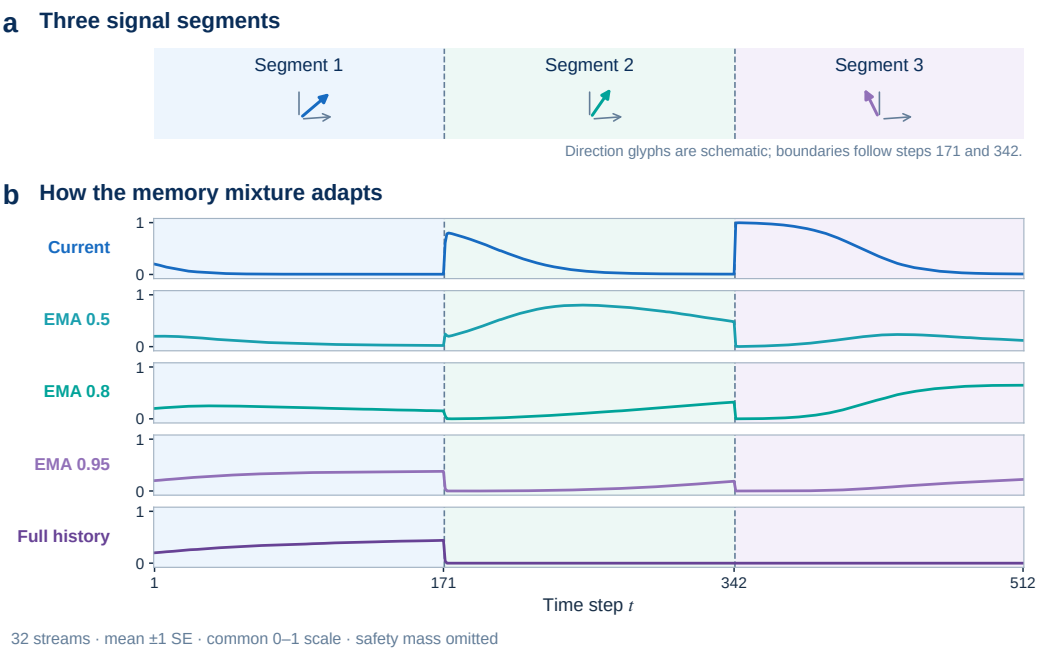

Figure 2: Adaptation is visible at each memory scale. Five primary weights on a common $[0, 1]$ scale, from current gradient to full history. Lines and bands are mean ± 1 standard error over 32 three-segment streams. Boundaries after steps 171 and 342 are not algorithm inputs; direction glyphs are schematic. The small safety mass is omitted; time points are not independent replicates.

Mixture versus rounded output. Rounding also matters: on strong static signals mixture loss 0.712 becomes 0.435 after rounding. The main table evaluates this rounded policy; its guarantee is the total-risk factor two in (16), not the mixture's leading-one oracle. Table 4 retains all three output conventions.

Weak directions and residual dependence. On the zero and near-zero second-scale cells, Raw has higher mean loss than both split controls. At exact zero its loss is 1.242, versus 0.118 for raw full history, 0.902 for one-way splitting, and 0.550 for symmetric splitting. Mean disagreement and positive response are 174.1 and 3.63, versus 2.03 and 0.032 in the strong stationary cell. Yet the wider common-polar cells also contain zero singular values and favor Raw over one-way splitting. The adverse cases therefore identify costly configurations, not a universal failure condition of rank deficiency. The relevant quantities are disagreement and response; their measured sizes do not prove the excess unavoidable. Figure 3 independently checks the selection identity and shows what replacing χ_t by v_t misses.

Numerical checks. Across the 272 streams, maximum isometry and direct loss-recomputation errors are 8.99×10^{-15} and 2.67×10^{-15}. Differential, quadrature, causal, candidate-parity, and pathwise-potential checks are detailed in Appendix I. Enabling response diagnostics leaves actions unchanged. These checks support the implementation but do not replace the proofs.

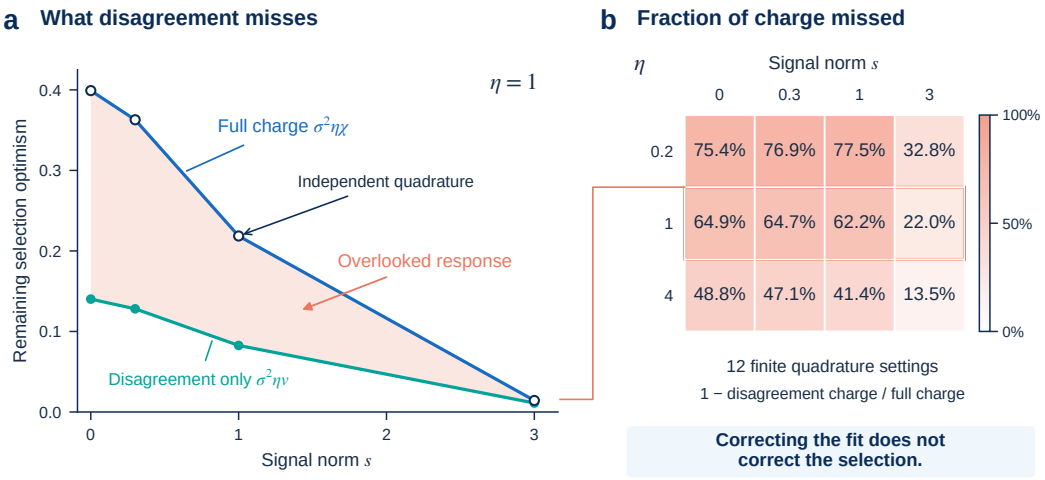

Figure 3: Fitting correction alone leaves a selection response. (a) At $\eta = 1$, the full charge matches independent quadrature; the shaded gap is the response missed by disagreement alone. Lines connect four evaluated signal norms, not additional measurements. (b) The percentage of full charge missed, computed as one minus the ratio of the two integrated charges, across all twelve fixed settings. Here $\sigma = 1$ and the fitted prior is .9. The maximum quadrature identity discrepancy is below 10^{-12}; the displayed finite grid is not a universal bound.

6 Discussion and Conclusion

Conditional SURE aggregation provides the starting point, while rank-boundary control makes it valid for the original-data polar memories studied here. A predictable safety action regularizes the selected map; the sequential oracle retains the exact response price and compares candidates on the unaltered gradient stream. Spectrum-saturated bounds explain why difficult weak directions should be charged by their signal contribution rather than a global inverse gap.

The scope is a finite affine-current memory bank under known Gaussian noise and strict rectangularity. The response can be large, and the separated-library lower bound does not establish minimax optimality for arbitrary fitted memories or switch budgets. Randomized isometric output inherits the mixture oracle; deterministic rounding has a factor-two guarantee. Computation requires multiple SVDs, and feedback comparisons share one realized trajectory rather than separately trained counterfactual optimizers.

The experiments support original-precision fitting as an alternative to one-way splitting, not dominance over symmetric splitting or the best fixed memory. Together with the regularity and risk results, they show how to reuse the observation without altering the candidate information budget, while keeping the remaining cost of adaptation explicit.

AI USE STATEMENT

Generative AI was used extensively for research formulation, mathematical derivation and proof writing, literature search, implementation, diagnostic experiments, interpretation, manuscript drafting, and figure/table generation. This assistance included implementation of synthetic-data generation; numerical samples were produced by fixed programs and random seeds, not directly supplied by a language model. Earlier remote development experiments used Codex; the new experiments reported in this manuscript were executed in the available research environment from fixed code and saved inputs. No aggregate result was invented, favorable trials substituted for adverse trials, or numerical test represented as a formal proof. The checks include executable algebraic tests, independent quadrature, code/data parity checks, and reference verification. They do not constitute an independent human proof audit or proof-assistant certification. The authors remain responsible for all submitted claims and disclosures.

REPRODUCIBILITY STATEMENT

The appendices contain complete proofs, algorithm constants, experiment generators, the frozen final plan, and per-cell summary tables. The full marginal and paired statistics accompany the code. The accompanying archive contains executable code, every original observation stream and latent test signal, all per-step comparison losses, weights, score and sensitivity traces, and independent seed summaries. To meet the upload-size limit, primary actions and mixtures are reconstructed from the saved observations and weights by the included preparation script; no saved observation or loss is replaced. Redundant candidate matrices and split views can be regenerated exactly up to numerical platform differences from these observations, frozen routines, and saved random seeds. The preparation script checks the reconstructed primary losses against the unchanged saved losses. The paper-only LaTeX package contains the manuscript, generated tables, vector figures, and bibliography. Development results and the preceding split-view pilot are kept separate from the 272 new streams. The stated arithmetic model and the limitations of floating-point rank handling are detailed in Appendix I.

ETHICS STATEMENT

The experiments use synthetic matrices, not human data or human participants. The main integrity risks are conflating a latent-risk oracle with observed-score regret, overlooking rank-boundary terms, overstating information efficiency, or treating small synthetic comparisons as optimizer superiority. The manuscript states these distinctions, includes unfavorable cases, and does not claim general training acceleration.

REFERENCES

Siddhartha Banerjee, Alankrita Bhatt, and Christina Lee Yu. The SMART approach to instance-optimal online learning. In *Proceedings of the Thirty-Seventh Conference on Learning Theory*, volume 247 of *Proceedings of Machine Learning Research*, pp. 426, 2024. URL https://proceedings.mlr.press/v247/banerjee24a.html. Extended abstract; full version arXiv:2402.17720.

Emmanuel J. Candès, Carlos A. Sing-Long, and Joshua D. Trzasko. Unbiased risk estimates for singular value thresholding and spectral estimators. *IEEE Transactions on Signal Processing*, 2013. doi: 10.1109/TSP.2013.2270464. URL https://arxiv.org/abs/1210.4139.

Maxime Cauchois, Alnur Ali, and John Duchi. A comment and erratum on "excess optimism: How biased is the apparent error of an estimator tuned by SURE?". arXiv:2112.14353, 2021. URL https://arxiv.org/abs/2112.14353.

Arnak S. Dalalyan and Joseph Salmon. Sharp oracle inequalities for aggregation of affine estimators. *The Annals of Statistics*, 40(4):2327–2355, 2012. doi: 10.1214/12-AOS1038. URL https://arxiv.org/abs/1104.3969.

Steven de Rooij, Tim van Erven, Peter D. Grünwald, and Wouter M. Koolen. Follow the leader if you can, hedge if you must. *Journal of Machine Learning Research*, 15(37):1281–1316, 2014. URL https://jmlr.org/papers/v15/rooij14a.html.

Evan S. Gawlik and Melvin Leok. Computing the Fréchet derivative of the polar decomposition. arXiv:1608.04491, 2016. URL https://arxiv.org/abs/1608.04491.

Mark Herbster and Manfred K. Warmuth. Tracking the best expert. *Machine Learning*, 32:151–178, 1998. doi: 10.1023/A:1007424614876.

Ruichen Jiang, Zakaria Mhammedi, Mehryar Mohri, and Aryan Mokhtari. Adaptive matrix online learning through smoothing with guarantees for nonsmooth nonconvex optimization. arXiv:2602.08232, 2026. URL https://arxiv.org/abs/2602.08232.

Xianliang Li, Zihan Zhang, Weiyang Liu, and Han Bao. Denoise first, orthogonalize later: Understanding momentum in Muon via spectral filtering. arXiv:2606.03899v2, 2026. URL https://arxiv.org/abs/2606.03899v2.

Natalia L. Oliveira, Jing Lei, and Ryan J. Tibshirani. Unbiased risk estimation in the normal means problem via coupled bootstrap techniques. arXiv:2111.09447v3, 2024. URL https://arxiv.org/abs/2111.09447v3. Version revised April 24, 2024.

Charles M. Stein. Estimation of the mean of a multivariate normal distribution. *The Annals of Statistics*, 9(6):1135–1151, 1981. doi: 10.1214/aos/1176345632.

Ryan J. Tibshirani and Saharon Rosset. Excess optimism: How biased is the apparent error of an estimator tuned by SURE? *Journal of the American Statistical Association*, 114(526):697–712, 2019. doi: 10.1080/01621459.2018.1429276. URL https://arxiv.org/abs/1612.09415.

A CONVENTIONS AND POLAR DIFFERENTIAL CALCULUS

Matrices are real. Vectorization identifies the Frobenius inner product with the Euclidean inner product. For $m \geq r$, $\mathrm{St}(m,r) = \{P : P^\top P = I_r\}$. A polar completion of a rank-deficient matrix is any measurable isometric maximizer of its inner product over this set. All results involving a target's zero singular values are independent of that completion on the target's null support. The algorithm and baseline use the same numerical SVD convention, but no theorem assumes that numerical conventions constitute exact arithmetic.

At full column rank, write $M = U\,\mathrm{diag}(s)V^\top$. The polar differential in direction E is

$$D\,\mathrm{polar}_M[E] = U\Omega V^\top + (I - UU^\top)EV\,\mathrm{diag}(s^{-1})V^\top, \quad \Omega_{jk} = \frac{F_{jk} - F_{kj}}{s_j + s_k}, \quad F = U^\top EV. \tag{A.1}$$

The diagonal of Ω is zero. This formula follows by differentiating $M = PH$, projecting onto the orthogonal complement of the columns of P, and solving the skew-symmetric Sylvester equation in an eigenbasis of H. It is also the Hessian of the nuclear norm on the full-column-rank set, hence self-adjoint as a map on the Frobenius space. Polar differentials and spectral divergences are established results (Gawlik & Leok, 2016; Candès et al., 2013); we supply their specific forms and regularity arguments because they determine the statistical correction here.

Taking the trace of this linear map gives $(m-r)\sum_j 1/s_j$ from the external-column-space directions and $2/(s_j + s_k)$ from every internal skew pair. Symmetric internal perturbations contribute zero. Therefore its divergence is exactly (4). Repeated positive singular values introduce no singular denominators. In a tied block, the scalar coefficients of the invariant spectral gradient agree, so the expressions do not depend on the basis chosen within that block.

The gradient of κ is

$$\nabla_M \kappa(M) = U\,\mathrm{diag}(\kappa_j')V^\top, \qquad \kappa_j' = -\frac{m-r}{s_j^2} - 2\sum_{k\neq j}\frac{1}{(s_j + s_k)^2}. \tag{A.2}$$

Holding the past fixed in $M_i = B_i + a_i G$, the complete score gradient used to evaluate the sensitivity is

$$\boxed{\nabla_G u_i = A_i + a_i D\,\mathrm{polar}_{M_i}[G] - \sigma^2 a_i^2 \nabla_{M_i}\kappa(M_i).} \tag{A.3}$$

For the predictable safety action, this gradient is simply A_0. This is an ordinary gradient with respect to *current* G, not a derivative through the entire online history. Formula (A.3) requires no differentiation of singular vectors across time or differentiation of the final rounded action.

Basic loss identities. If $S = UH$, $H \succeq 0$, and $P, U \in \mathrm{St}(m,r)$, expansion gives

$$\mathcal{D}(S,P) = \mathrm{tr}\,H - \mathrm{tr}[H\,\mathrm{sym}(U^\top P)] = \tfrac{1}{2}\,\mathrm{tr}[H(P-U)^\top(P-U)]. \tag{A.4}$$

For every contraction Q, the nuclear/operator duality ensures $\mathcal{D}(S,Q) \geq 0$. Writing $Q = PK$, $0 \preceq K \preceq I$, gives $\|2Q - P\|_{\mathrm{op}} \leq 1$, which proves (16) for every polar completion of Q. The scalar example $Q = 0$, $P = 1$, $S = -1$ attains the factor two. This algebra is a task-loss rounding guarantee, not a claim that rounding always improves a mixture.

B INTEGRABILITY, SAFETY, AND EMPIRICAL RANK BOUNDARIES

The following arguments establish the hypotheses of Gaussian integration by parts rather than infer them from full rank almost surely.

Lemma B.1 (A first inverse moment without a mean assumption)**.** *Let $M = C + \tau Z \in \mathbb{R}^{m\times r}$, $m > r$, with independent standard Gaussian entries in Z and any fixed C, $\tau > 0$. Then*

$$\mathbb{E}[s_{\min}(M)^{-1}] \leq \frac{r}{\tau}\,\mathbb{E}[\chi_{m-r+1}^{-1}] < \infty. \tag{B.1}$$

The bound is uniform over C.

Proof. Let d_j be the distance of column j to the span of the other columns. At full rank, $\|M^\dagger\|_F^2 = \mathrm{tr}[(M^\top M)^{-1}] = \sum_j d_j^{-2}$, so $s_{\min}^{-1} \le \sum_j d_j^{-1}$. Conditional on the other columns, their span has dimension $r-1$ almost surely. The projected jth column is a Gaussian vector in dimension $d = m - r + 1 \ge 2$ with covariance $\tau^2 I_d$ and an arbitrary mean. Its expected inverse radius is no larger than the centered value. One direct proof uses $x^{-1/2} = \pi^{-1/2}\int_0^\infty t^{-1/2}e^{-tx}\,dt$ and

$$\mathbb{E}e^{-t\|\tau z+c\|^2} = (1+2\tau^2 t)^{-d/2}e^{-t\|c\|^2/(1+2\tau^2 t)} \le (1+2\tau^2 t)^{-d/2}.$$

The centered inverse-radius moment is $\tau^{-1}2^{-1/2}\Gamma((d-1)/2)/\Gamma(d/2)$, finite for $d > 1$. Sum over columns. □

Proof of Lemma 1. Condition on $\mathcal{F}_{t-1}$, so B, a, S are fixed and $M = B + aS + a\sigma Z$. Use the smooth contraction $P_\epsilon(M) = M(M^\top M + \epsilon^2 I)^{-1/2}$. Its singular-value function is $f_\epsilon(s) = s/\sqrt{s^2+\epsilon^2}$. The Frobenius operator norm of its differential at a full-rank M is at most $1/s_{\min}(M)$: the internal symmetric, internal skew, and external coefficients are, respectively, divided differences of f_ϵ, $(f_\epsilon(s_j) + f_\epsilon(s_k))/(s_j + s_k)$, and $f_\epsilon(s_j)/s_j$. All are bounded by $1/s_{\min}$. The diagonal coefficient $f_\epsilon'(s)$ satisfies the same bound. Hence the current-data divergence is dominated in absolute value by $amr/s_{\min}$, integrable by Lemma B.1.

For each $\epsilon > 0$, ordinary Gaussian integration by parts applies to the smooth bounded map $P_\epsilon(B + aG)$: $\mathbb{E}\langle G - S, P_\epsilon\rangle_F = \sigma^2\mathbb{E}\,\mathrm{div}_G P_\epsilon$. The maps converge almost surely to $\mathrm{polar}(M)$, and their divergences converge to $a\kappa(M)$. Dominated convergence on both sides proves the assertion. This proof remains valid at rank-deficient true S because the moment argument permits every conditional mean. The candidate score is integrable, but a second moment was never required. □

Lemma B.2 (The positive safety component regularizes the selected map)**.** *Fix the past. Suppose $m > r$, all fitted current gains are positive, the safety prior is positive, and $0 < \eta < \infty$. Extend $e^{\eta u_i}A_i$ and $e^{\eta u_i}$ by zero on the fitted rank-deficient set of candidate i. The resulting soft mixture Q is continuously differentiable in G on the entire matrix space. Under the priors and temperature bounds of Algorithm 1, its derivative, selected score, and sensitivity are integrable when $\mathbb{E}\|S_t\|_F^2 < \infty$.*

Proof. Let $s_i = s_{\min}(M_i)$ and $d = m - r \ge 1$. Since $\kappa(M_i) \ge d/s_i$,

$$q_i \le \frac{\pi_i}{\pi_0}\exp\left\{\eta\left(2\sqrt{r}\,\|G\|_F - \frac{\sigma^2 a_i d}{s_i}\right)\right\}. \tag{B.2}$$

The first two differential formulas imply, with a dimension-dependent finite constant,

$$\|D_G A_i\| \le C_r a_i/s_i, \qquad \|\nabla_G u_i\|_F \le \sqrt{r} + C_r\left(a_i\|G\|_F/s_i + \sigma^2 a_i^2/s_i^2\right).$$

On a compact set of G, the exponential in (B.2) dominates every one of these inverse powers. Values and first derivatives of the unnormalized fitted component tend to zero at its rank-deficient set. The distance to that affine rank set is s_i/a_i; the same bound proves differentiability with zero derivative on the set itself. The safety component has strictly positive smooth denominator there. The sum and quotient therefore define a C^1 mixture, even when several fitted matrices lose rank simultaneously.

For the global moment assertion, put $z_i = \sigma^2 a_i/s_i$ and $H = 2\sqrt{r}\|G\|_F + \eta^{-1}\log(1/\pi_0)$. For $p = 1, 2$, the elementary bound

$$z^p \min\{1, e^{\eta(H - dz)}\} \le C_{p,d}(H + \eta^{-1})^p \qquad (z \ge 0) \tag{B.3}$$

is obtained by splitting at $z = 2(H + \eta^{-1})/d$ and maximizing the remaining exponential tail. Thus $q_i z_i$ and $q_i z_i^2$ are bounded by polynomials of degree at most two in $\|G\|_F$, with constants depending on η^{-1} and the safety-prior floor, not on B_i or a_i. The mixture derivative contains only weighted $D_G A_i$ and weighted score gradients, so the preceding bounds control it. They similarly control χ, $\sum q_i u_i$, and the selected correction.

In Algorithm 1, $\pi_{t,0} \ge \gamma\epsilon_0$, and $\sqrt{L}/(\sigma\sqrt{rT}) \le \eta_t \le \sqrt{L}/(\sigma\sqrt{r})$. These are deterministic finite bounds. Taking expectations uses only a finite second moment of G_t, hence of S_t. No second moment of every individual inverse-spectrum score is assumed. □

Proof of Theorem 2. Holding the past fixed, differentiate the finite softmax: $\nabla_G q_i = \eta q_i(\nabla_G u_i - \sum_j q_j \nabla_G u_j)$. The product rule yields

$$\operatorname{div}_G Q = \sum_i q_i \operatorname{div}_G A_i + \sum_i \langle A_i, \nabla_G q_i \rangle_F = \sum_i q_i d_i + \eta \sum_i q_i \langle A_i - Q, \nabla_G u_i \rangle_F.$$

Lemma B.2 justifies integration by parts for this selected map. Consequently $\mathbb{E}\langle G - S, Q\rangle_F = \sigma^2 \mathbb{E}[\sum_i q_i d_i + \eta\chi]$. Subtract the fitted divergence term from $\langle G, Q\rangle_F$ to obtain (8). All expressions are conditional on the same past, so the result applies equally to predictable feedback. □

A finite-variance shortcut would fail. For $m = 2, r = 1, S = 0, a = 1, B = 0, \sigma = 1$, let $R = \|G\|$. Its density is $re^{-r^2/2}$ on $(0, \infty)$, and the individual score is $u = R - R^{-1}$. It is integrable with $\mathbb{E}u = 0$, but

$$\mathbb{E}[u^2 \mathbf{1}_{R \geq \epsilon}] = \int_\epsilon^\infty (r - r^{-1})^2 r e^{-r^2/2} dr = \log(1/\epsilon) + O(1) \longrightarrow \infty. \tag{B.4}$$

Thus an ordinary variance-based expert bound cannot be imported without extra work. The pathwise log-sum-exp potential below does not make that assumption.

A square-matrix shortcut would also fail. For scalar $G \sim N(s, \sigma^2)$ and $A(G) = \operatorname{sign}(G)$, the ordinary derivative is zero almost everywhere, but

$$\mathbb{E}[G \operatorname{sign}(G)] - s\mathbb{E} \operatorname{sign}(G) = 2\sigma\varphi(s/\sigma) > 0, \tag{B.5}$$

where φ is the standard normal density. The missing derivative is a boundary term. In the smooth-approximation argument, domination by an integrable inverse singular value is precisely what fails for a scalar Gaussian. True full rank does not remove the issue: any $s \neq 0$ still has the positive term in (B.5). We do not apply the raw-polar theorem to square matrices by declaring that a zero-probability rank event is irrelevant.

C Complete pathwise potential and raw-data oracle

This proof separates deterministic exponential-weight algebra from its statistical interpretation. In particular, no comparator chosen from current noise is treated as a fixed unbiased estimator.

For any t, define the mix gain $m_t = \eta_t^{-1} \log \sum_i \pi_{t,i} e^{\eta_t u_{t,i}}$. The exact entropy identity is

$$\langle q_t, u_t \rangle = m_t + \mathrm{KL}(q_t \| \pi_t)/\eta_t. \tag{C.1}$$

Consider a fixed comparator index a_t among the equally weighted primary candidates. The update implies

$$m_t = u_{t,a_t} + \frac{\log \pi_{t,a_t} - \log q_{t,a_t}}{\eta_t}. \tag{C.2}$$

For $k_t \in (0, 1]$, concavity gives $Z_t^{\mathrm{temp}} = \sum_i \omega_i (q_{t-1,i}/\omega_i)^{k_t} \leq 1$. If the comparator stays at index i at t, then

$$\frac{\log \pi_{t,i}}{\eta_t} \geq \frac{\log q_{t-1,i}}{\eta_{t-1}} - L(\eta_t^{-1} - \eta_{t-1}^{-1}) + \eta_t^{-1} \log(1 - \gamma). \tag{C.3}$$

If a new run starts at $t > 1$, use instead $\pi_{t,i} \geq \gamma\omega_i$; at $t = 1$, use $\pi_{1,i} = \omega_i$.

Sum (C.2) along each constant run $[s, e]$. The $\log q$ terms telescope, and the terminal term $-\log q_{e,i}/\eta_e$ is nonnegative. The first run pays at most L/η_e. Each later run pays at most $L/\eta_e + \log(1/\gamma)/\eta_s$. Each stay pays $\log[1/(1-\gamma)]/\eta_t$. Because the temperature is nonincreasing, all inverse temperatures are at most $1/\eta_T$. Summing K runs and $T - K$ stays yields

$$\sum_t m_t \geq \sum_t u_{t,a_t} - C_K/\eta_T. \tag{C.4}$$

Combining with (C.1) proves (14). The proof is pathwise and permits any finite real scores. Exact singular fitted matrices have conditional probability zero under the model; the safety extension defines the algorithm even at those exceptional inputs.

Taking expectations in (14), use Lemma 1 for each fixed comparator index and Theorem 2 for the data-dependent mixture. Subtract both sides from $\mathbb{E}\sum_t \|S_t\|_*$ to obtain (13). The expected entropy term is finite because $\pi_{t,i} \geq \gamma\omega_i$ and the inverse temperatures have finite deterministic bounds. Lemma B.2 handles the remaining integrability. An infimum over deterministic comparator paths is therefore valid *outside* the expectation. A hindsight path minimizing true losses separately on each realized noise trace does not have this interpretation.

Proof of Corollary 4. Let $x_t = r + \sum_{s<t} v_s$. Since every candidate is an isometry, $0 \leq v_t \leq r \leq x_t$. Therefore

$$\frac{v_t}{\sqrt{x_t}} = \big(\sqrt{x_t + v_t} - \sqrt{x_t}\big)\big(1 + \sqrt{1 + v_t/x_t}\big) \leq (1+\sqrt{2})\big(\sqrt{x_{t+1}} - \sqrt{x_t}\big). \tag{C.5}$$

Discard the negative entropy term in (13) and write $\chi_t = v_t + \psi_t$. The contribution $\sigma^2 \sum_t \eta_t v_t$ is at most $(1+\sqrt{2})\sigma\sqrt{L}\sqrt{r + \sum_t v_t}$ by (C.5). The complexity term is at most $\sigma C_K \sqrt{r + \sum_t v_t}/\sqrt{L}$. Bound the remaining signed response by its positive part and use Jensen's inequality for the square root. Finally, with $\gamma = 1/(T+1)$, $(T-1)\log[1/(1-\gamma)] = (T-1)\log(1+1/T) \leq 1$. □

Scope of the derivation. Log-sum-exp, fixed share, temperature adaptation, and smooth SURE aggregation are established ingredients. Appendix J gives the exact correspondence. The task-specific steps are the safety regularity for polar memories and its use in a causal original-data comparison with explicit response and spectral costs. The oracle retains ψ_t; substituting $\chi_t = v_t$ for fitted candidates would remove a term that need not vanish.

D Predictable weights and an interpolation of the two prices

Use the same candidate bank, prior, posterior, and temperature states as the primary selector. For a predictable $\theta_t \in [0,1]$, let $w_t = (1-\theta_t)\pi_t + \theta_t q_t$ and $Q_t^{(\theta)} = \sum_i w_{t,i} A_{t,i}$. Only the posterior part of the weights differentiates with current data, so

$$\mathbb{E}\langle S_t, Q_t^{(\theta)}\rangle_F = \mathbb{E}\langle w_t, u_t\rangle - \sigma^2 \mathbb{E}[\theta_t \eta_t \chi_t]. \tag{D.1}$$

The entropy identities

$$\langle q_t, u_t\rangle = m_t + \mathrm{KL}(q_t\|\pi_t)/\eta_t, \qquad \langle \pi_t, u_t\rangle = m_t - \mathrm{KL}(\pi_t\|q_t)/\eta_t \tag{D.2}$$

give, for the same fixed comparator path,

$$R_{Q^{(\theta)}} \leq B_a^{\mathrm{raw}} + C_K \mathbb{E}\eta_T^{-1} + \mathbb{E}\sum_t \Bigg[(1-\theta_t)\frac{\mathrm{KL}(\pi_t\|q_t)}{\eta_t} + \theta_t\left(\sigma^2\eta_t\chi_t - \frac{\mathrm{KL}(q_t\|\pi_t)}{\eta_t}\right)\Bigg]. \tag{D.3}$$

The extremes are the predictable-weight control and current selector. The temperature states remain those of the same posterior process. Choosing θ_t after looking at current G_t would introduce an additional derivative; (D.3) must not be cited for an uncharged data-dependent switch.

In the final experiments, the predictable control takes $\theta_t = 0$ but its $A_{t,i}$ still use current G_t. Thus its three-segment deficit is not explained by candidates lacking the current sample. The difference is whether current evidence changes the weights immediately.

E Spectral geometry and original-noise memory bounds

Proof of (17) *and* (18). Expanding $\|(Q-U)H^{1/2}\|_F^2$ gives $\mathrm{tr}(HQ^\top Q) + \mathrm{tr}\,H - 2\langle UH, Q\rangle_F$. Adding $\mathrm{tr}[H(I - Q^\top Q)]$ and dividing by two yields the identity. The matrix $I - Q^\top Q$ is PSD because Q is a convex combination of contractions. In an eigenbasis of H, the contribution to its trace from eigenvalues at least μ is at most $\mu^{-1}\,\mathrm{tr}[H(I-Q^\top Q)]$. The remaining contribution is exactly the term defining W_μ. Sum and use the nonnegativity of the second term in (17) to obtain $\bar{v} \leq 2R_Q/\mu + W_\mu$.

Set $B' = B_a^{\mathrm{raw}} + J_T$. If $R_Q \le B'$, (18) is immediate. Otherwise $x = R_Q - B' > 0$ obeys $x^2 \le c_K^2[r + W_\mu + 2B'/\mu + 2x/\mu]$. The positive quadratic root is the displayed bound. No parameter in this argument is supplied to the algorithm. Small weak-spectrum disagreement and small response cost are sufficient conditions, not facts inferred merely from a positive signal scale. □

Lemma E.1 (Spectrum-saturated polar perturbation)**.** *Let $S = \sum_{j=1}^{q} \lambda_j u_j v_j^\top$, $\lambda_j > 0$, and $M = S + E$. For any isometric polar completion P of M,*

$$\mathcal{D}(S, P) \le \sum_{j=1}^{q} \min\left\{2\lambda_j, \frac{\|Ev_j\|^2 + \|E^\top u_j\|^2}{\lambda_j}\right\}. \tag{E.1}$$

Proof. For any matrix A, let $\mathcal{H}(A) = \left(\begin{smallmatrix} 0 & A \\ A^\top & 0 \end{smallmatrix}\right)$ and $\mathcal{T}(P) = \left(\begin{smallmatrix} 0 & P \\ P^\top & 0 \end{smallmatrix}\right)$. The matrix $I - \mathcal{T}(P)$ is PSD, has norm at most two, and annihilates the positive spectral space of $\mathcal{H}(M)$. Thus it is bounded by twice the nonpositive projection $\Pi_{M,\le 0}$. Put $w_j = (u_j; v_j)/\sqrt{2}$, a λ_j eigenvector of $\mathcal{H}(S)$. Projecting $(\lambda_j I - \mathcal{H}(M))w_j = -\mathcal{H}(E)w_j$ onto the nonpositive space gives $\|\Pi_{M,\le 0} w_j\| \le \|\mathcal{H}(E)w_j\|/\lambda_j$. It follows that $1 - u_j^\top P v_j \le \min\{2, 2\|\mathcal{H}(E)w_j\|^2/\lambda_j^2\}$. Use $2\|\mathcal{H}(E)w_j\|^2 = \|Ev_j\|^2 + \|E^\top u_j\|^2$, multiply by λ_j, and sum. The proof permits every completion on the null support and requires no gap between positive singular values. □

For a fixed deterministic normalized kernel, the error about S_t is $B_{t,i}^{\mathrm{bias}} + \sigma\sqrt{k_{t,i}}Z$ in distribution. Its expected squared action on v_j and adjoint action on u_j sum to $\|B_{t,i}^{\mathrm{bias}} v_j\|^2 + \|(B_{t,i}^{\mathrm{bias}})^\top u_j\|^2 + (m + r)\sigma^2 k_{t,i}$. Apply Lemma E.1 and the concavity of $x \mapsto \min(c, x)$ to obtain (20). The stationary full-history case has zero bias and $k = 1/t$, proving (19).

An unnormalized EMA is equivalent, for these actions and scores, to dividing its matrix and current gain by $1 - \beta^t$. Thus its normalized kernel has

$$h_{t,s} = \frac{(1-\beta)\beta^{t-s}}{1-\beta^t}, \qquad k_t = \frac{(1-\beta)^2(1-\beta^{2t})}{(1-\beta^2)(1-\beta^t)^2}. \tag{E.2}$$

The implementation can use unnormalized states without altering the theorem. The bias calculation is for deterministic signals and kernels. For feedback signals or data-selected kernels, the conditional score oracle still applies under its assumptions, but the above unconditional bias/noise split is not silently reused.

The information budget of splitting. The pair $X = G + \rho\sigma W$, $Y = G - \rho^{-1}\sigma W$ reconstructs G as $(X + \rho^2 Y)/(1 + \rho^2)$. Its two independent mean observations have variances $v_X = (1 + \rho^2)\sigma^2$ and $v_Y = (1 + \rho^{-2})\sigma^2$, with $v_X^{-1} + v_Y^{-1} = \sigma^{-2}$. There is no creation of a second original-precision sample. Our comparison eliminates the inflation of each fitted memory by keeping $M_i(G)$ unchanged; it does not claim a way to obtain free independent validation of the same precision. Dependence is charged through Theorem 2 instead.

F A FINITE-LIBRARY LOWER BOUND AND THE LIMITS OF A UNIVERSAL CLAIM

Proposition F.1 (A matched selection subproblem)**.** *Suppose $A_1, \dots, A_N \in \mathrm{St}(m, r)$, $N \ge 4$, satisfy $\|A_i - A_j\|_F^2 \ge c_0 r$ for $i \ne j$, where $c_0 > 0$ is fixed. There is a positive amplitude a such that, on the constant-signal family $S_t = aA_v$, every causal contraction-action procedure has worst-case cumulative loss at least $c\sigma\sqrt{rT\log N}$. The best fixed candidate has zero loss. For fixed current-data-independent candidates, Algorithm 1 has the matching order with a constant depending only on c_0, including after deterministic rounding.*

Proof. Choose $a = \sigma\sqrt{\log N/(16rT)}$ and put the uniform prior on v. The KL divergence between any pair of full T-sample Gaussian laws is at most $Ta^2\|A_i - A_j\|_F^2/(2\sigma^2) \le 2Ta^2 r/\sigma^2 = \log N/8$. The same bound holds for the mutual information under the uniform prior. The elementary entropy form of Fano's inequality gives error at least $1 - 1/8 - \log 2/\log N \ge 3/8$ for recovering v, even after observing all T samples.

At any time, decode v by the nearest library action to the algorithm's output P in Frobenius norm. A decoding error implies $\|P - A_v\|_F \geq \sqrt{c_0 r}/2$. Since $\|P\|_F^2 \leq r$, $\mathcal{D}(aA_v, P) = a(r - \langle A_v, P\rangle_F) \geq a\|P - A_v\|_F^2/2$. Each time has expected loss at least $3ac_0 r/64$, even with all observations granted in advance. Summing proves the order. The prior is used only to lower-bound worst-case risk.

For the upper bound, current-data-independent actions have $d_i = 0$, $\nabla u_i = A_i$, and $\chi_t = v_t$. The response budget vanishes. With the correct fixed index as comparator, Corollary 4 gives $O(\sigma\sqrt{rT \log N})$, and (16) adds at most two. This upper bound does not assume knowledge of a or v. □

The proposition is conditional on a separated library existing. It is not a packing theorem for arbitrary N at fixed dimensions. Nor does it establish minimax optimality over all fitted memory rules or over arbitrary switch counts. Those distinctions are essential to the paper's interpretation of an oracle comparison.

A direct check of the response term. For one fitted vector action $A_1 = G/R$, $R = \|G\|$, and safety $A_0 = e_1$, the scores are $u_1 = R - \sigma^2/R$, $u_0 = G_1$. Their gradients are $(1 + \sigma^2/R^2)A_1$ and A_0. Writing $c = A_1^\top e_1$ and fitted weight q, direct substitution gives

$$\chi = q(1-q)(2 + \sigma^2/R^2)(1-c), \quad v = 2q(1-q)(1-c), \quad \psi = q(1-q)(1-c)\sigma^2/R^2.$$

This proves (21). The ratio $\chi/v = 1 + \sigma^2/(2R^2)$ has no finite pointwise upper bound as $R \downarrow 0$, even though the safety-regularized expectation is finite. The main response price is therefore not replaceable by a universal multiple of disagreement through this argument.

G Predictable feedback, covariance, and output semantics

Every score calculation conditions on $\mathcal{F}_{t-1}$. Thus S_t, affine offsets and gains, the safety action, prior, and temperature may all depend on the previous observation/action path. The current noise must still be fresh Gaussian. The pathwise potential applies to the resulting realized score sequence, while conditional Stein identities supply the expectation step. This is why the main oracle permits parameter feedback without treating an adaptive signal as a deterministic Gaussian shift.

If f is lower bounded and L_f-smooth in Frobenius norm, $S_t = \nabla f(W_t)$, and $W_{t+1} = W_t - hP_t$ for isometric outputs, the descent lemma gives

$$\frac{1}{T}\sum_t \mathbb{E}\|\nabla f(W_t)\|_* \leq \frac{f(W_1) - f_*}{hT} + \frac{L_f h r}{2} + \frac{1}{T}\mathbb{E}\sum_t \mathcal{D}(S_t, P_t). \tag{G.1}$$

For randomized candidate output the last term is R_Q/T; for deterministic polar output it is at most $2R_Q/T$. Substituting the oracle gives a conditional performance consequence, not a universal convergence-to-zero or acceleration theorem: the raw comparator and response terms must themselves be useful. The comparison is on the same realized gradient trajectory, not to each candidate's independently trained parameter path.

Known nonisotropic Gaussian noise. The accounting identity has a direct covariance version. If $\mathrm{vec}(G - S)$ has known positive-definite covariance Σ, replace the candidate correction by $\mathrm{tr}(\Sigma J_{A_i})$ and define

$$u_i = \langle G, A_i\rangle_F - \mathrm{tr}(\Sigma J_{A_i}), \qquad \chi_\Sigma = \sum_i q_i \langle A_i - Q, \Sigma \nabla u_i\rangle_F,$$

using vectorization in the last expression. Then $\mathbb{E}\langle S, Q\rangle_F = \mathbb{E}\sum_i q_i u_i - \eta\mathbb{E}\chi_\Sigma$ whenever the same integrability conditions hold. The raw polar Jacobian is PSD, so a positive-definite Σ preserves the rank-boundary suppression. The pathwise potential is unchanged. We do not claim the isotropic complexity constants or experiments automatically extend to an estimated, misspecified, or non-Gaussian covariance. The implementation and new experiments use the declared scalar $\sigma^2 I$.

Actual actions, not rescaled comparisons. A sampled action uses fresh sampling randomness after q_t has been computed and has conditional expectation Q_t. Each realization has Frobenius norm $\sqrt{r}$. Deterministic polar rounding has the same actual norm and the pointwise factor-two guarantee;

a mixture alone generally has smaller norm. All three losses are saved separately. No theorem differentiates the final rank-deficient polar completion of Q_t, so its nonsmoothness does not invalidate the smooth selected-mixture argument.

H Reproducible experimental definitions and complete results

H.1 Fixed plan, provenance, and statistical units

The final experiment comprises 272 streams in 16 cells. The five core cells each have 32 seeds, 961100–961131. The longer static cell has 16 seeds, 961200–961215; the two weak-spectrum cells each have 16 seeds, 961300–961315. The eight wider cells each have eight seeds, 961400–961407. These are the actual seed ranges in `plan.json`. Reusing an integer seed in different cells does not create additional independent evidence about a common hypothesis; all summaries are computed within a cell. Noise generation uses NumPy `SeedSequence([seed,110])`; auxiliary control noise and optional action sampling use the distinct suffixes 230 and 430.

The current algorithm and these cells were fixed before the final runs. Development beforehand compared several raw-score/temperature choices on four distinct seeds, 89010–89013, using already familiar signal families. The resulting final algorithm retains the preceding disagreement schedule, with the base-prior correction needed for the added safety action. We did not alter the final settings after inspecting their outcomes, rerun unfavorable seeds, or pool the final trials with development. This is a research-driven controlled evaluation, not an untouched test of a hypothesis specified before any prior research. Earlier evidence-filter experiments and the 80-stream split pilot remain separate provenance, not observations contributing to the new means.

An independent stream is the statistical unit. We sum each loss over its complete horizon before reporting the mean, standard error $s/\sqrt{n}$, and descriptive interval $\bar{x} \pm t_{.975,n-1} s/\sqrt{n}$. Paired intervals are computed from within-seed differences, not from marginal standard errors. No correction for all comparisons is applied. "Best EMA" selects the smallest of three *cell means*; it is a descriptive benchmark, not an online oracle, a per-step hindsight selector, or an independent inferentially selected winner. The complete marginal and paired records, including all methods rather than only favorable comparisons, are in `results/seed_totals.csv` and `results/aggregate.json`.

H.2 Raw memories, actions, and information budget

The primary candidate bank contains current gradient, full history, and exponential averages with coefficients $\beta \in \{.5, .8, .95\}$. For the last three, initialize $M_{0,i} = 0$ and use

$$M_{t,i} = \beta_i M_{t-1,i} + (1 - \beta_i) G_t.$$

Their current coefficients a_i are $1 - \beta_i$. Full history uses $M_{t,\text{full}} = t^{-1} \sum_{s \le t} G_s$ and $a_{t,\text{full}} = 1/t$; current uses $M = G, a = 1$. Positive normalization of an EMA does not affect its polar or corrected score, provided a is rescaled together with M. We keep the unnormalized recurrences to avoid unnecessary operations. The sixth, safety action is the *previous* full-history polar, with $a_0 = 0$ and base mass $\epsilon_0 = (T+1)^{-2}$; at the first step it is the first r coordinate vectors. Primary masses are $(1 - \epsilon_0)/5$. The implementation stores the safety entry last, whereas the mathematical notation numbers it zero; this is only an index convention.

All primary candidates see precisely the original G_t, without auxiliary perturbation or held-out coordinates. The score and its derivative hold the memory before this observation fixed. The selection response χ_t is calculated for validation and reporting; it is not an additional, unproved penalty inside the exponential-weight update. The algorithm's weights use (5), the temperature (10), and the base-prior-preserving recursion (11).

The main reported action is $\mathrm{polar}(Q_t)$. We separately save Q_t's loss and the actual loss of sampling a candidate from q_t. Only the sampled rule has expected loss exactly R_Q; deterministic rounding has the factor-two guarantee. The predictable-weight ablation uses $\mathrm{polar}(\sum_i \pi_{t,i} A_{t,i})$ from the *same* current candidate bank and weight process. It is the mathematically specified alternative output, not an independently retuned learner. The uncorrected control reruns the same selector architecture with $u_i = \langle G, A_i \rangle$ and its own weights.

For the prior one-way split control, independent W_t gives $X_t = G_t + \sigma W_t$ and $Y_t = G_t - \sigma W_t$. Five candidates are updated on X_t and evaluated on Y_t, each of conditional variance $2\sigma^2$. The prior selector uses its original uniform five-candidate prior, fixed share $1/(T+1)$, and disagreement schedule with evaluation noise standard deviation $\sqrt{2}\sigma$. Its code is retained unchanged. The symmetric control runs the opposite direction as well and outputs $\mathrm{polar}((Q_{X\to Y} + Q_{Y\to X})/2)$. It uses no second real gradient, but approximately doubles control-side candidate fitting and selection. The split controls have five fitted candidates, while the raw rule additionally has the small-mass predictable safety action required by its proof. We do not hide this library difference or attribute every empirical difference solely to variance.

All actual matrix outputs have operator norm one and Frobenius norm $\sqrt{r}$. Equal norm does not imply equal finite-step curvature cost in neural-network training. All experiments here evaluate the specified first-order task on shared signal/observation trajectories.

H.3 Five core generators and weak-spectrum cells

Use zero-based time $u = 0, \dots, T-1$, $U_0 = [e_1, e_2] \in \mathbb{R}^{m\times 2}$, planar rotation $R(\theta)$, and row rotation $R_{23}(\theta)$ of the second and third output coordinates.

Strong static. Set $a = 4\sigma\sqrt{m}$ and $S_u = U_0 \operatorname{diag}(a, .3a)$. The core cell has $T = 512$, $m = 3$, $\sigma = .1$; the long cell has $T = 2048$ with the same amplitude and noise. Thus the long cell does not quietly reduce the signal with the horizon.

Common polar and noncommuting amplitudes. Set $V_u = R(.17u)$, $\lambda_{1,u} = .9 + .3\sin(.071u)$, and $\lambda_{2,u} = .08 + .175[1 + \cos(.13u)]$. Use

$$S_u = U_0 V_u \operatorname{diag}(\lambda_{1,u}, \lambda_{2,u}) V_u^\top.$$

The polar is fixed while the PSD amplitude matrices generally do not commute.

Three polar segments. Select the frames U_0, $R_{23}(1.25)U_0$, and $R_{23}(-1.1)U_0$ according to $\min(2, \lfloor 3u/T \rfloor)$. Multiply by $R(.09u)\operatorname{diag}(1, .45)R(.09u)^\top$. Exact cuts are $\lceil T/3 \rceil$ and $\lceil 2T/3 \rceil$. No rule receives these change times.

Reactivation. Let $z = u/T$, $\theta_u = \pi \min\{1, \max\{0, (z - .25)/.35\}\}$. The second amplitude is .3 before $z = .25$, .003 on $[.25, .6)$, and .45 afterward. Use $S_u = R_{23}(\theta_u)U_0 \operatorname{diag}(.9, a_u)$.

Continuous coupled rotation. Use $S_u = R_{23}(3u/T)U_0 R(.032u)\operatorname{diag}(.9, .35)R(.032u)^\top$. The latter four core cells have $T = 512$, $m = 3$, and $\sigma = .04$. Their generator module is copied byte-for-byte from the preceding package. They are not newly invented favorable signal functions.

Adverse weak spectra. Set $S = U_0 \operatorname{diag}(1, \lambda_2)$ with $\lambda_2 \in \{0, 10^{-4}\}$, $T = 512$, $m = 3$, and $\sigma = .1$. True rank loss is permitted by the theorem. These cases nevertheless expose a considerable empirical selection cost. We keep them in the main table, not just in this appendix.

H.4 Wider structural cells

The eight wider cells cross $(m, r) = (8, 4), (16, 8)$ with static, common-polar PSD motion, a three-stage frame change, and multiple-plane continuous rotation. All use $T = 512$, $\sigma = .04$, and the same fixed selector parameters.

For static, switching, and continuous cells, the base spectrum is $\mathrm{geomspace}(1, .08, r)$. For the common-polar cell it is the concatenation of $r/4$ ones, $r/4$ equally spaced values from .6 to .2, and $r/2$ zeros. Multiply its entries by $1 + .15\sin(.017u + i)$ for coordinate $i = 0, \dots, r-1$. This tests a changing joint PSD support as well as repeated or zero instantaneous scales.

Let each column i of the left isometry occupy its own plane (e_i, e_{r+i}). For switching, use angle $c_j\theta_i$ with θ_i equally spaced from .6 to 1.3, cuts $\lfloor .35T \rfloor$, $\lfloor .71T \rfloor$, and $c_j \in \{0, 1, -.8\}$. For continuous

Size	Signal	**Raw**	Full (G)	Split	Symmetric	Best EMA	SE (Raw)
8×4	static	0.606	0.562	1.054	0.625	1.397	0.048
8×4	common	1.022	0.696	1.475	0.876	1.203	0.099
8×4	switch	4.394	154.780	7.128	4.199	8.694	0.117
8×4	rotating	5.113	277.874	7.795	6.471	5.207	0.109
16×8	static	2.042	1.988	3.801	2.069	5.196	0.036
16×8	common	2.110	1.667	3.336	1.950	3.966	0.117
16×8	switch	11.503	272.718	18.916	11.154	25.728	0.135
16×8	rotating	16.297	491.747	21.358	16.708	16.555	0.140

Table 3: All wider-cell mean losses, eight seeds per row. Every cell uses one raw G stream shared by methods. These are tests at widths four and eight, not large-layer timing benchmarks or unrestricted-width minimax experiments.

Signal	Mixture	Sampled	**Rounded**	Predictable	Naive	SE (Rounded)
Static, strong	0.712	0.727	0.435	0.431	24.271	0.026
Common polar / PSD	0.124	0.141	0.057	0.056	3.491	0.003
Three polar segments	0.751	0.724	0.599	1.655	1.714	0.011
Weak reactivation	0.794	0.789	0.649	0.776	2.219	0.030
Continuous rotation	0.457	0.459	0.383	0.384	2.134	0.006
Static, $T = 2048$	0.703	0.768	0.397	0.393	100.468	0.023
Zero second scale	1.760	1.755	1.242	1.230	6.972	0.156
Second scale 10^{-4}	1.811	1.801	1.293	1.282	7.024	0.156

Table 4: Output conventions and score ablations. Mixture is the conditional expected sampled loss, not a full-norm action. Sampled and Rounded contain actual isometric outputs. Predictable uses the same fitted actions with predictable weights; Naive omits candidate fitting correction. The adverse near-null output costs are retained.

motion use

$$\theta_{u,i} = 2\left(1 + \frac{i\sqrt{2}}{r-1}\right)\frac{u}{T} + .12\sin(.037(i+1)u).$$

Static and common-polar cases have zero left angles. In each nonstatic case, form a right rotation V_u by successively rotating adjacent column pairs $(i, i+1)$ by $.013\sqrt{i+2}\,u$. The signal is $U_u V_u \operatorname{diag}(\lambda_{u,i}) V_u^\top$. These definitions are explicit programmed generators; right and left motions affect more than two columns.

H.5 Output and paired-comparison accounting

Across all 16 cells, Raw has lower reported mean loss than one-way splitting in 14 and than symmetric splitting in 7. These counts are descriptive; reused seed integers across cells do not supply independent cross-cell trials. The two adverse second-scale configurations show that preserving original noise variance alone does not ensure a small selection cost. Rank deficiency is not itself a failure condition: both wider common-polar cells include zero singular values and improve over one-way splitting. The response in (13) can remain appreciable even when a true weak direction contributes little direct loss. Table 2 places the full fixed-comparator accounting alongside mixture and rounded losses; the complete output and cost records remain in Tables 4 and 6.

Signal	Raw − Split	95% interval	Raw − Best EMA	95% interval
Static, strong	-0.333	[-0.418, -0.248]	-0.311	[-0.352, -0.271]
Common polar / PSD	-0.035	[-0.042, -0.027]	-0.013	[-0.018, -0.007]
Three polar segments	-0.431	[-0.457, -0.404]	-0.696	[-0.719, -0.674]
Weak reactivation	-0.061	[-0.122, -0.000]	0.262	[0.196, 0.329]
Continuous rotation	-0.255	[-0.270, -0.241]	0.032	[0.030, 0.035]
Static, $T = 2048$	-0.414	[-0.569, -0.258]	-1.905	[-1.998, -1.811]
Zero second scale	0.339	[-0.023, 0.701]	0.987	[0.657, 1.317]
Second scale 10^{-4}	0.335	[-0.024, 0.693]	0.987	[0.657, 1.317]

Table 5: Raw deterministic loss minus control loss, with descriptive paired 95% Student intervals. Negative favors Raw. These comparisons use independent seed totals, not time points; no simultaneous adjustment is claimed. Best EMA is chosen by the smallest cell mean as described above.

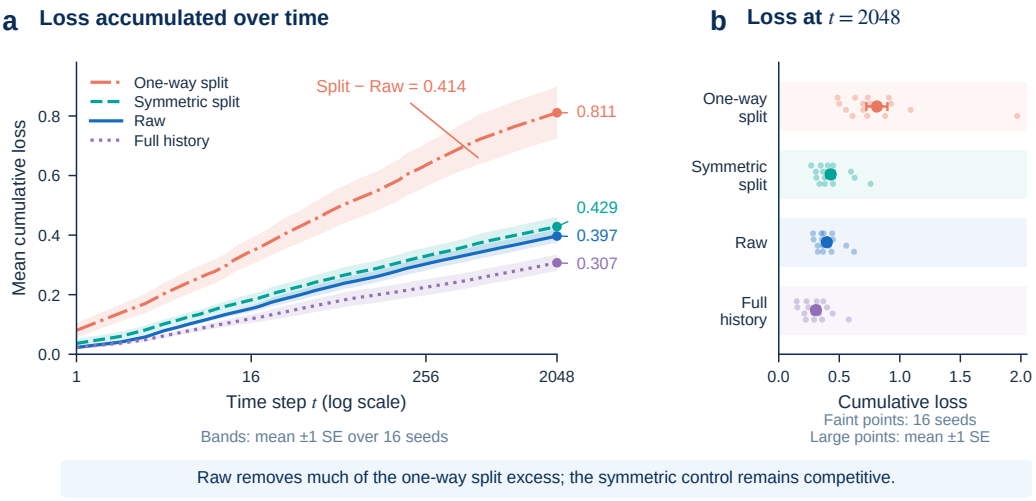

Figure 4: Stationary precision without a claim of dominance. (a) Cumulative losses on the long stationary cell ($T = 2048$); the horizontal axis is logarithmic. (b) Every faint point is one of the 16 independent seed totals. Curves, large points and error bands/bars show means and ± 1 standard error, not 95% intervals. All methods observe the same original stream; Full history uses G. Raw closes much of the one-way split excess, while symmetric splitting remains competitive.

I Numerical checks and implementation boundaries

The numerical programs are falsification and reproducibility checks; no proof-assistant, independent human certification, interval-arithmetic certification, or global floating-point stability theorem is claimed.

I.1 Differential, geometric, and causal checks

The final validation ran 180 random differential instances. Maximum relative finite-difference errors were 2.24×10^{-8} for candidate-score gradients, 3.73×10^{-9} for polar divergence, and 8.52×10^{-10} for the selected-map Jacobian trace. One thousand geometry/rank tests checked the variance identity and deterministic rounding: maximum variance-identity error was 6.22×10^{-15}, maximum isometry error 5.76×10^{-15}, and no positive rounding-inequality violation occurred.

Forty-eight short random score trajectories were compared with a dynamic program over comparator paths with one, two, or three runs. The pathwise potential had no positive violation. Candidate parity with the unchanged raw baseline bank was within 2.09×10^{-15}. Enabling diagnostics and

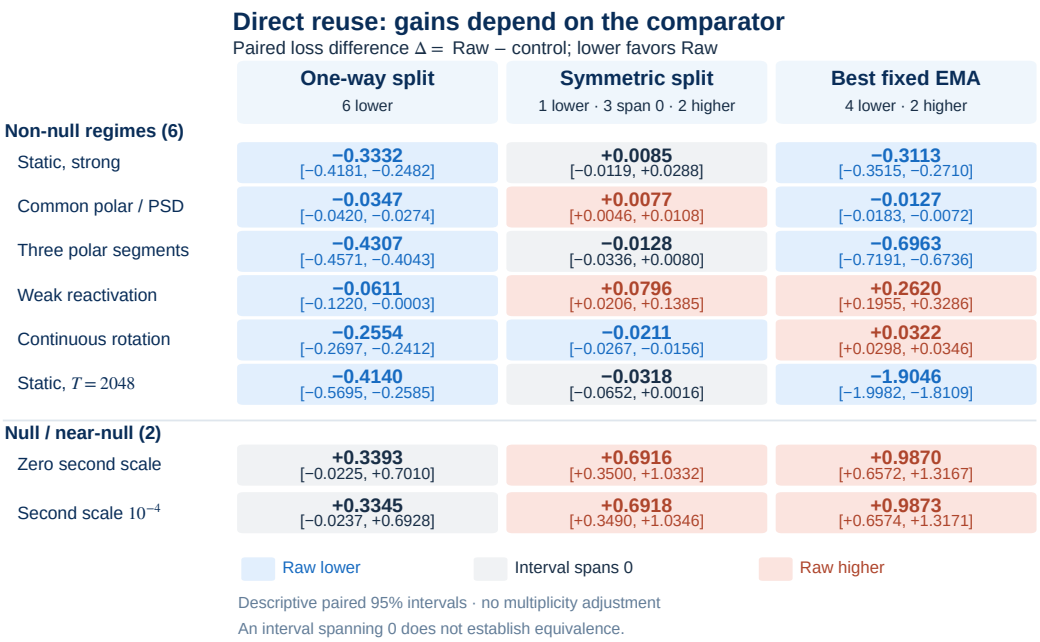

Figure 5: Direct reuse does not dominate every comparator. Each cell shows mean Raw-minus-control loss and a descriptive paired 95% Student interval (32 seeds per core cell; 16 per long or adverse cell). Blue and coral indicate intervals wholly below and above zero; gray spans zero and does not establish equivalence. Column counts refer only to the six non-null rows. Best EMA is selected by cell mean, not by seed or time. Intervals have no multiplicity adjustment; color does not encode effect size.

Signal	Disagreement $\widehat{\mathbb{E}\sum v_t}$	Response $\widehat{J}_T$	Oracle accounting $C_1\mathbb{E}\widehat{\eta_T^{-1}} + \mathcal{O}_T$	Mean safety weight
Static, strong	2.03	0.032	0.510	1.14e-04
Common polar / PSD	0.55	0.004	0.144	9.36e-06
Three polar segments	0.55	0.001	-0.265	2.26e-06
Weak reactivation	52.76	0.508	1.152	3.70e-06
Continuous rotation	0.29	0.001	0.122	3.99e-07
Static, $T = 2048$	2.28	0.030	0.521	2.03e-06
Zero second scale	174.11	3.634	6.496	7.16e-06
Second scale 10^{-4}	174.64	3.635	6.506	7.19e-06

Table 6: Observed disagreement and response accounting, summed within each seed and then averaged. Here $\mathcal{O}_T = \mathbb{E}\sum_t[\sigma^2\eta_t\chi_t - \mathrm{KL}(q_t\|\pi_t)/\eta_t]$ is the signed price in the oracle, and $J_T = \sigma^2\mathbb{E}\sum_t \eta_t[\psi_t]_+$. The displayed estimates include the fixed-comparator complexity where indicated. They do not certify a uniformly small remainder.

changing future observations left the corresponding earlier actions exactly unchanged. Scaling the entire gradient/noise problem and using the prescribed scaled temperature changed actions by at most 3.53×10^{-15}. Tests approaching fitted rank loss verified the intended vanishing weight of the singular fitted action relative to the safety action. A separate truncated second-moment calculation retained the logarithmic divergence of the Rayleigh example rather than assuming finite score variance.

Across all 272 completed streams, recomputing the loss from the saved true signal and actual output gave maximum error 2.67×10^{-15}; maximum output orthogonality error was 8.99×10^{-15}. The saved deterministic outputs never violated the pointwise factor-two inequality beyond the tested

tolerance. No exact empirical rank deficiency was encountered in these Gaussian trials and all reported metrics were finite.

I.2 An independent integral check of the selected-score identity

For $m = 2, r = 1$, use the current normalized vector and safety action e_1, prior masses $.9, .1$, noise standard deviation one, and signal se_1. The twelve fixed settings are $s \in \{0, .3, 1, 3\}$ and $\eta \in \{.2, 1, 4\}$. Polar-coordinate quadrature uses a periodic angular sum with 2,048 nodes and adaptive radial integration. The upper radial limit is $\max(14, s + 12)$; the resulting Gaussian-tail truncation is negligible at the reported numerical precision, but we do not present it as a rigorous interval bound. Both the total-optimism identity and the score-after-fitting-correction identity agree to below 10^{-12}; quadrature-reported error estimates are approximately 10^{-11}.

For example, at $s = 0, \eta = .2$, the remaining selection optimism is about $.135526$, while charging only the independent-action disagreement gives about $.033314$. The response term accounts for the difference. Figure 3 shows the $\eta = 1$ charge decomposition and the missed-charge fractions for all twelve settings. This is a direct independent numerical integration of an explicit conditional formula, not a large-sample claim that a selected score is pointwise unbiased. The score may be negative, and finite-seed averages of its correction need not equal their expectation.

I.3 Exact definitions versus numerical safeguards

The theorem defines full-rank polar actions in exact real arithmetic, and the differentiable selected-map extension at the measure-zero rank boundary. The code uses float64 SVD. It does not add an artificial singular-value floor, silently change σ, clip a large inverse-spectrum score, or tune a noise radius from the true signal. At an exact numerical zero singular value, the fitted candidate receives the limiting score $-\infty$ and weight zero; the predictable safety action ensures a finite normalizer. This is the extension used in the proof, not a square-matrix sign-boundary claim. Overflow or a nonfinite derivative in an otherwise allegedly full-rank input triggers an error instead of being silently accepted. The final trials encountered no such error.

The primary weight update requires singular values for d_i, but not the more expensive score derivative. Computing χ_t adds linear algebra for reporting, not a different policy. For N affine memories the persistent storage is $O(Nmr)$; fitted SVDs cost $O(Nmr^2)$ by a standard thin factorization, plus an additional SVD for deterministic rounding. Exact Jacobian-vector products can reuse these factors, without materializing an $mr \times mr$ Jacobian. This is not a one-EMA-cost claim. No GPU, large-layer, complete-construction, or matched-risk runtime speedup is inferred from our small CPU study.

I.4 Released records and rerunning

The package contains source, the fixed plan, generator and previous-control modules, all 272 original S/G arrays, weights, scores, divergence, response accounting, and each method's per-time loss. The submission stores these records in a compressed data archive. Before analysis, `code/prepare_submission_data.py` unpacks them and reconstructs primary mixtures and polar outputs from the saved observations and weights using the original candidate implementation. No observation, saved loss, seed, or diagnostic trace is regenerated or replaced by this preparation step. Redundant fitted-action tensors and artificial control views can be regenerated from the saved raw observations, fixed seeds, and released code; the compact release does not pretend to be a byte-for-byte copy of every working-array file. The complete code path regenerates the controls as well as the raw rule, and a clean-directory replay checks selected saved results.

Validation scripts reproduce the differential, potential, geometry, integrability-example, quadrature, and dataset checks. Summarization scripts generate every table from per-seed records; figures are generated from these summaries and the separate quadrature JSON. Environment information is included. Historical aggregate files and the preceding split-pilot seed totals are excluded from the submission supplement and do not contribute to the new table counts. The generator and split-control modules needed by the current experiments are retained in `provenance/`.

J EXACT SURE CORRESPONDENCE AND CONTRIBUTION SCOPE

Fix t and condition on $\mathcal{F}_{t-1}$; suppress time indices and put $n = mr$. The signal, prior, temperature, and pre-observation memory states are fixed. The following algebra makes the SURE correspondence explicit.

Candidate loss and weights. For every isometry A, $\|A\|_F^2 = r$, so

$$\|S - A\|_F^2 = 2\mathcal{D}(S, A) + c(S), \qquad c(S) = \|S\|_F^2 + r - 2\|S\|_*. \tag{J.1}$$

Thus squared-error and descent-loss comparisons between candidates differ only by a common additive quantity and a factor two. Using Lemma 1, the conditional SURE of each candidate is

$$\widehat{R}_i = \|G - A_i\|_F^2 + 2\sigma^2 d_i - n\sigma^2 = K(G) - 2u_i, \qquad K(G) = \|G\|_F^2 + r - n\sigma^2, \tag{J.2}$$

$$q_i = \frac{\pi_i e^{-\eta \widehat{R}_i/2}}{\sum_j \pi_j e^{-\eta \widehat{R}_j/2}}. \tag{J.3}$$

The common $K(G)$ factor cancels; the safety action is included with $d_0 = 0$.

Selected-mixture risk. Let $v = r - \|Q\|_F^2 = \sum_i q_i \|A_i - Q\|_F^2$. Expanding the squared norm and using Theorem 2 gives

$$\begin{aligned} \widehat{R}_Q &:= \|G - Q\|_F^2 + 2\sigma^2 \operatorname{div}_G Q - n\sigma^2 \\ &= \sum_i q_i \widehat{R}_i - v + 2\sigma^2 \eta \chi \\ &= \|G\|_F^2 + \|Q\|_F^2 - n\sigma^2 - 2\left(\sum_i q_i u_i - \sigma^2 \eta \chi\right). \end{aligned} \tag{J.4}$$

Lemma B.2 supplies the regularity and integrability needed for $\mathbb{E}[\widehat{R}_Q \mid \mathcal{F}_{t-1}] = \mathbb{E}[\|S-Q\|_F^2 \mid \mathcal{F}_{t-1}]$. Unlike a candidate, Q need not have squared norm r:

$$\|S - Q\|_F^2 = 2\mathcal{D}(S, Q) + c(S) - v. \tag{J.5}$$

The variance term $-v$ is therefore essential when translating between the mixture's squared-error and descent risks. A sampled isometry has conditional expected descent loss $\mathcal{D}(S, Q)$; deterministic polar rounding instead uses (16).

What requires additional proof. Equation (J.4) is the finite-library specialization of the general smooth-aggregate identity in Dalalyan & Salmon (2012, Appendix A.2, Lemma 2), with their temperature $\beta = 2/\eta$ and per-coordinate risk normalization. The difficulty here is justifying this identity for inverse-spectrum scores and rank-singular polar maps: almost-sure full rank is insufficient, individual scores may have infinite variance, and square matrices can carry a boundary contribution. Appendix B proves the needed moments and safety regularity directly. Appendix C then combines the conditional identity with the changing-temperature, fixed-share potential to compare original-data candidate paths under predictable feedback.

SURE tuning likewise distinguishes candidate-wise unbiasedness from selected-rule risk (Tibshirani & Rosset, 2019; Cauchois et al., 2021). Here (21) prevents replacing polar response pointwise by a universal multiple of disagreement. The spectral comparator bounds and localized budgets describe when the original-data comparison can be useful.