WHEN DOES GRADIENT FIDELITY SUPPORT LOW-RANK OPTIMIZATION DECISIONS?

Anonymous authors
Paper under double-blind review

ABSTRACT

Low-rank optimization restricts model updates to a limited set of directions. Gradient fidelity measures the fraction of the current gradient's energy captured by a candidate. We ask when higher fidelity justifies a candidate preference, accounting for update implementation and subsequent training. With candidate geometry fixed, projected descent preserves fidelity rankings, whereas a zero-moment LoRA+AdamW step can reverse them. In the studied regime, its local first-order score is approximately a weighted L_1 score of projected-gradient coordinates rather than the squared-energy score used by fidelity. Matched controls isolate adaptive coordinate normalization as a source of the zero-moment reversals. Using projected descent to remove implementation-induced reordering, we compute a first-order future-sensitivity score by differentiating a common reference continuation before evaluating candidate-specific terminal outcomes. The score identifies terminal winners in $4/5$ states in each of two language cohorts and $5/5$ vision states, versus at most $1/5$ per cohort for fidelity. Retrospectively, gains over lowest-fidelity and post hoc best-fixed rules are smaller than gains over fidelity. For a fixed finite candidate set, first-order score differences depend only on future sensitivity along candidate differences. A signed gain and residual bound give a sufficient condition for first-order preservation or reversal of local pairwise ordering. The condition certifies 26 of 60 evaluated pairs as first-order reversals; all 26 also reverse the local ordering in separate terminal evaluations. This geometry also specifies how to assess approximate future sensitivity by comparing induced pairwise-score errors with the reference first-order gaps they could overturn. Together, these results distinguish gradient capture from candidate value and show how update implementation and first-order future valuation can re-rank candidates.

1 INTRODUCTION

Low-rank methods can improve the efficiency of model training and adaptation by reducing trainable parameters or optimizer memory. The current full gradient is a natural signal for constructing low-rank update spaces: GaLore uses dominant gradient subspaces, and related initialization and rank-allocation methods likewise exploit gradient or spectral information (Zhao et al., 2024; Wang et al., 2024; Refael et al., 2025).

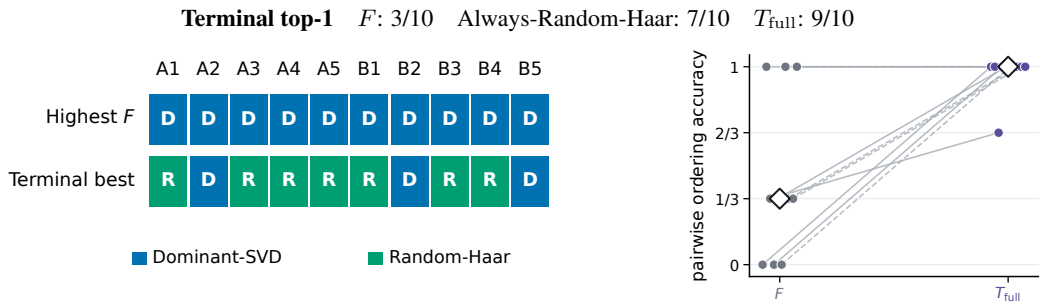
Gradient information can guide candidate construction. Fidelity measures the fraction of current gradient energy captured by a candidate, but using it to justify a preference requires a further link between gradient capture and the value of the implemented update or its terminal outcome. This link can fail at two points. First, a candidate must be turned into an update by a receiver, which specifies the parameterization and update rule, such as direct descent along the projected gradient (hereafter Direct) or LoRA+AdamW (Hu et al., 2022; Loshchilov & Hutter, 2019). The receiver may induce a different ranking from fidelity. Second, the update changes later parameters, gradients, and optimizer state, so its value for the current loss need not determine its value at the end of training.

This paper asks: when does higher current-gradient fidelity justify choosing one candidate over another? We analyze these two points separately. We begin by studying when the receiver preserves fidelity rankings, using Direct as a positive control that preserves them exactly. We then remove receiver-induced local reordering to examine whether the local ordering determines terminal rankings and to characterize the future information relevant to first-order comparisons within a fixed candidate

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.

set. A receiver-margin condition and a signed future-ordering condition identify sufficient regimes for the corresponding first-order ranking relations. We test whether the first-order future rankings agree with finite terminal outcomes using separately executed candidate continuations.

Our experiments exhibit ranking changes at both stages. At one Qwen3-8B state, a LoRA+AdamW step from zero-initialized optimizer moments reverses the fidelity ranking in all eight matched comparisons. Under Direct, the fidelity-maximizing Dominant-SVD candidate is the terminal winner in only $3/10$ language states. A future-sensitivity score, T_{full}, computed by differentiating a common reference continuation and fixed before candidate outcomes are evaluated, identifies the terminal winner in $9/10$ states; always-Random-Haar, the post hoc best fixed rule, gives $7/10$ (Figure 1).



(a) Current-fidelity and terminal winners.

(b) Pairwise ranking accuracy.

Figure 1: **Current-gradient fidelity can miss the terminal winner.** Dominant-SVD, SARA-style importance sampling, and Random-Haar are compared under the same Direct receiver on the ten Language A+B states at 75% of the scheduled training steps ($p = 0.75$). The top annotation reports terminal-winner selections: $3/10$ for fidelity, $7/10$ for always-Random-Haar, and $9/10$ for T_{full}. In this candidate pool, always-Random-Haar coincides with selecting the lowest-fidelity candidate and is also the best fixed rule identified post hoc. **(a)** Columns A1–A5 and B1–B5 denote states and compare the candidate selected by fidelity, which is Dominant-SVD in every state, with the terminal winner. **(b)** Points are states, lines connect scores within a state (solid for A, dashed for B), and diamonds are medians. Selection regret is in Appendix Figure 12.

Our contributions are as follows.

Receiver implementation can change candidate rankings. Matched controls link zero-moment LoRA+AdamW ranking reversals to coordinate normalization, and a closed-form score matches nearly all pairwise receiver rankings across five independently trained states. A controlled nonzero-moment comparison isolates second-moment normalization as a source of preference changes.

Local update rankings do not identify terminal rankings. A smooth convex construction permits opposite terminal rankings under identical local information. Across two language cohorts and one vision cohort, common-reference future sensitivity identifies more terminal winners than fidelity, with smaller gains over simple baselines in retrospective comparisons. The objective-matched switch-time control does not reproduce the future score's improvement over fidelity.

First-order comparisons depend only on future sensitivity along candidate differences. We characterize which components of future sensitivity affect first-order candidate comparisons and show how approximation errors enter pairwise score differences relative to the gaps they could overturn. A signed sufficient condition certifies first-order reversals for $26/60$ pairs on ten new language and vision states. In separate terminal evaluations, the local ordering is reversed in all 26 certified pairs and in 22 of the 34 uncertified pairs.

2 Scores and Controlled Comparisons

To analyze the two stages introduced above, we distinguish four quantities: F, S_R, T_{full}, and Q (Table 1). These quantities distinguish whether implementation preserves fidelity rankings ($F \to S_R$), how subsequent training revalues the implemented updates at first order ($S_R \to T_{\text{full}}$), and whether these future rankings agree with observed terminal rankings at finite intervention size ($T_{\text{full}} \to Q$).

Table 1: Definitions and information requirements for the four candidate-ranking quantities (larger is better); V denotes a low-rank candidate.

Score	What it measures	Information required
$F(V)$	Fraction of current gradient energy captured by the candidate	Current gradient and candidate geometry
$S_R(V)$	First-order decrease in the current loss from the implemented update	Current gradient and the receiver's actual update
$T_{\text{full}}(V)$	First-order decrease in the terminal objective from the candidate intervention	Candidate intervention, common reference continuation, and its sensitivity
$Q(V)$	Negative terminal objective after executing the candidate intervention	A separately executed candidate continuation

Current-gradient fidelity. Let $G = \nabla_W L$ be the gradient of the current training loss L with respect to a target weight matrix W, and let V denote a low-rank candidate with a fixed budget. Write Π_V for the Frobenius-orthogonal projection onto the candidate's allowed update space and $H_V = \Pi_V(G)$ for the captured gradient. We define

$$F(V) = \frac{\|H_V\|_F^2}{\|G\|_F^2} \in [0, 1]. \tag{1}$$

For nonzero G, fidelity depends only on the current gradient and candidate geometry, without specifying how an optimizer will use the candidate. For multiple target matrices, energies and inner products are summed over matrices. Our experiments use the right-acting projection $\Pi_V(G) = GP_V$, with $P_V = VV^\top$ and orthonormal columns in V (candidate constructions in Appendix C).

Implemented updates and local ordering. A receiver maps the current training state z_t, gradient, and candidate to an actual weight update, $\Delta W_R(V) = R(z_t, G, V)$, scored by its first-order decrease in the current loss:

$$S_R(V) = -\langle G, \Delta W_R(V)\rangle. \tag{2}$$

The actual finite-step loss change is measured separately (Section 3).

Direct as a positive control. Direct projected descent applies $\Delta W_{\text{Direct}}(V) = -\eta H_V$, using the same $\eta > 0$ for every candidate in a comparison. Orthogonal projection gives $S_{\text{Direct}}(V) = \eta\|H_V\|_F^2 = \eta\|G\|_F^2 F(V)$, so Direct preserves every strict fidelity ranking – a property we use to remove receiver-induced reordering when studying subsequent training. Because every candidate in a comparison shares η, fidelity also determines the update norm, $\|\Delta W_{\text{Direct}}(V)\|_F = \eta\|G\|_F\sqrt{F(V)}$. Selecting the lowest-fidelity candidate therefore also selects the smallest Direct update.

Future sensitivity and terminal quality. To evaluate how a candidate intervention affects terminal quality, we propagate an infinitesimal perturbation through the remaining training dynamics. Let the full state be $z_t = (\theta_t, \xi_t)$, where θ_t denotes the model parameters and ξ_t the optimizer/receiver state carried by the training update (e.g. Adam moments). With data order and randomness ω_t fixed, write $z_{t+1} = \Phi_t(z_t; \omega_t)$. For a candidate intervention direction e_V at switch time s, let $z_s^\tau = z_s + \tau e_V$, where τ scales the intervention. Define the tangent $\delta z_t(V) = \partial_\tau z_t^\tau|_{\tau=0}$. Then

$$\delta z_{t+1}(V) = D_z\Phi_t(z_t; \omega_t)\,\delta z_t(V), \qquad \delta z_s(V) = e_V. \tag{3}$$

Here $D_z\Phi_t(z_t; \omega_t)$ is the Jacobian of the update map with respect to the full state, evaluated along the reference continuation. We propagate tangents using Jacobian–vector products without forming the full Jacobian. For a terminal objective $J(z_T)$, define

$$T_{\text{full}}(V) = -\langle \nabla J(z_T), \delta z_T(V)\rangle = \left.\frac{d[-J(z_T^\tau)]}{d\tau}\right|_{\tau=0}, \tag{4}$$

the directional derivative of terminal quality along the candidate intervention, with positive values indicating a first-order decrease in terminal loss. In our trajectory experiments, the candidate intervention displaces otherwise frozen base-weight blocks at the switch. This displacement is retained while adapters and optimizer state continue to evolve. The terminal first-order effect therefore separates into a direct contribution from the retained displacement and a contribution transported through the evolving state:

$$T_{\text{full}}(V) = T_{\text{direct}}(V) + T_{\text{transported}}(V). \tag{5}$$

We compute T_{full} by forward JVPs along a common reference continuation and fix the resulting scores before evaluating candidate-specific terminal outcomes. An adjoint-based numerical cross-check is reported in Appendix D. To measure terminal quality, $Q(V) = -J(z_T^{(V)})$, we continue each finite-intervention branch from the same pre-intervention state under matched data order, randomness, and training protocol, while parameters and optimizer states may subsequently diverge. The terminal objective is mean cross-entropy on a fixed held-out subset of TRAIN; TEST is not used. We evaluate rankings using terminal-winner top-1 agreement, Spearman correlation, and pairwise accuracy. We measure selection quality by regret, defined as the selected candidate's terminal loss minus the lowest terminal loss among the evaluated candidates.

A sufficient condition for a general receiver. To expose how a receiver can reorder candidates, for $H_V \neq 0$, write $-\Delta W_R(V) = \alpha_V H_V + E_V$ with $\langle E_V, H_V \rangle = 0$, where $\alpha_V = \langle -\Delta W_R(V), H_V \rangle / \|H_V\|_F^2$ is the signed gain along the captured gradient and $\beta_V = \langle G - H_V, E_V \rangle / \|G\|_F^2$. For $H_V = 0$, we use the convention $\alpha_V = 0$ and $E_V = -\Delta W_R(V)$ (Appendix L). Then

$$\frac{S_R(V)}{\|G\|_F^2} = \alpha_V F(V) + \beta_V. \tag{6}$$

Thus candidate-dependent gains and residuals can alter the fidelity ranking.

Proposition 1 (Receiver ordering preservation) *Fix any reference gain $\eta > 0$. For a pair with $F_i := F(V_i) > F_j := F(V_j)$, define $D_{ij}(\eta) = |\alpha_i - \eta| F_i + |\alpha_j - \eta| F_j + |\beta_i| + |\beta_j|$. If*

$$\eta(F_i - F_j) > D_{ij}(\eta), \tag{7}$$

then $S_R(V_i) > S_R(V_j)$.

Here $\eta(F_i - F_j)$ is the normalized score gap under a common positive reference gain, and $D_{ij}(\eta)$ bounds the receiver-induced deviation from that gap.

Experimental units and state reuse. The receiver replication and nonzero-moment control share five Qwen states. Primary terminal-ranking experiments use three independent five-state cohorts, and the quotient test uses ten new states. The candidate-rule test reuses Language A+B states, while the $p = 0.50$ test revisits Language A seed identities (Appendix B).

3 Receiver Implementation Can Change Candidate Rankings

We first test the implementation stage identified in Section 1: can two candidates with a fixed fidelity ordering be ranked differently once they are turned into actual updates? We begin with Direct and LoRA+AdamW, then hold the LoRA installation fixed and vary only the update operation. We study a zero-moment step first and then a controlled nonzero-moment setting.

Candidates. CURRENT spans the installed LoRA-A row space. FORM uses the leading eigenvectors of a Gram matrix constructed from paired activations and backpropagated gradients. MARGINAL retains the singular values of both matrices but pairs their singular components by index instead of using measured cross-alignment. PERMUTED FORM reassigns FORM bases between target matrices of matching input dimension, changing the assigned subspaces rather than rotating a basis within one subspace. All constructions precede candidate outcomes; Appendix C gives the formulas.

Squared gradient energy and coordinate normalization. The receivers assign different local scores to the same projected gradient coordinates. Direct uses squared magnitudes, whereas a zero-moment AdamW step yields an approximately weighted L_1 score under the conditions below. Direct gives $S_{\text{Direct}}(V) = \eta \sum_m \|G_m V_m\|_F^2$, where m indexes target weight matrices. For LoRA, whose low-rank update is $(\alpha/r) B_m A_m$, installing V_m as $A_m = s_{A,m} V_m^\top$ with $B_m = 0$ gives factor gradient $g_{B,m} = c\, s_{A,m} G_m V_m$, where $s_{A,m}$ is the per-matrix installation scale for candidate V_m and $c = \alpha/r$ is the LoRA scaling factor ($c = 2$ in the Qwen3-8B experiments; installation protocol in Appendix C). With zero-initialized moments and zero weight decay, the first update is approximately sign-normalized when the clipped factor-gradient coordinates are large in magnitude relative to AdamW's numerical-stability constant ϵ. This yields $\widehat{S}_{L1}(V) = \text{lr}\, c \sum_m s_{A,m} \|G_m V_m\|_1$, where lr is the AdamW learning rate and $\|\cdot\|_1$ sums absolute matrix entries (exact score in Appendix C). Two candidates can therefore be ordered differently by fidelity and by the implemented-update score.

Matched update rules and measured utility. We test this at one Qwen3-8B-Base/GSM8K state using eight matched examples and FORM versus PERMUTED FORM. For LoRA+AdamW, we first merge the checkpoint adapter into the base weights, then install each candidate with $B = 0$ on an isolated model copy and take a single diagnostic *shadow step* with zero-initialized optimizer moments and zero weight decay. We record the effective dense update and same-example whole-model utility, defined as the loss decrease $U_R(V) = L_{\mathrm{pre}} - L_{\mathrm{post}}$. FORM has the larger squared-energy score in $8/8$ comparisons, while PERMUTED FORM has the larger approximate weighted-L_1 score in $8/8$. Direct preserves the fidelity ranking in all eight; LoRA+AdamW reverses it in all eight (Figure 2a). Under LoRA+AdamW, S_R ranks the two candidates consistently with realized same-example utility in $6/8$ comparisons, versus $2/8$ for fidelity (Appendix Figure 4). Holding state, candidates, target matrices, and installation fixed, an analytic counterfactual applying SGD to the LoRA factors (Factor-SGD) preserves the fidelity ordering in all eight comparisons with and without gradient clipping, whereas the matched zero-moment AdamW reconstruction reverses all eight under the same conditions. Together with the closed-form score, these controls attribute the reversals in this zero-moment comparison to adaptive coordinate normalization rather than LoRA factorization alone.

Replication across independently trained states. Across five independently trained Qwen3-8B-Base/GSM8K states, with FORM, PERMUTED FORM, and CURRENT rebuilt separately at each state, the same closed-form score reproduces $119/120$ pairwise receiver-score orderings and $39/40$ complete three-candidate rankings without refitting. LoRA+AdamW reverses the fidelity ordering of FORM and PERMUTED FORM in a majority of examples in every state, whereas the analytic Factor-SGD counterfactual preserves the fidelity ordering in all 80 comparisons with and without gradient clipping (Appendix C).

Isolating second-moment normalization with nonzero moments. To test nonzero memory, we build a fixed 16-step calibration history in each candidate's coordinates while keeping parameters fixed. For each candidate, let $\widehat{m}_V$ and $\widehat{v}_V$ denote the first- and second-moment quantities used by the AdamW update. Holding $\widehat{m}_V$ fixed, we compare $u_{\mathrm{M1}}(V) = \widehat{m}_V$ with $u_{\mathrm{Adam}}(V) = \widehat{m}_V/(\sqrt{\widehat{v}_V} + \epsilon)$, scored as S_{M1} and S_{Adam}. The moments may differ between FORM and PERMUTED FORM.

The first-moment-only rule prefers FORM in $40/40$ trials; adding the second-moment denominator shifts preference to PERMUTED FORM in $34/40$, with every state exceeding the preregistered threshold (Figure 2b). Thus second-moment normalization can change preference under controlled nonzero memory. The first-moment score is distinct from fidelity, and the weighted-L_1 approximation remains specific to the zero-moment case (Appendix P).

4 Local Rankings, Future Training, and Candidate Differences

4.1 Local Rankings and Terminal Outcomes

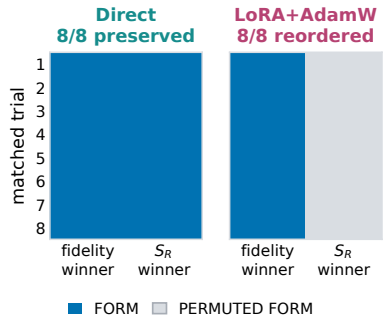
The receiver experiments identify one failure mode. We now remove it with Direct, which forces F and S_R to be co-ordered, and ask whether later training can still revalue the interventions.

Proposition 2 (Local updates do not identify terminal rankings) *Fix the switch state, current gradient, Direct receiver, two equal-budget rank-one candidate interventions, and terminal objective. Without restrictions on subsequent training dynamics, the same strict local ranking is compatible with opposite terminal rankings. Both futures can be realized by smooth future gradient-descent dynamics.*

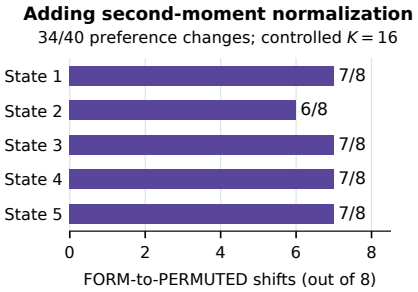
Appendix M gives a smooth convex construction in which the reversal persists for arbitrarily small interventions. We next test whether common-reference future sensitivity predicts terminal rankings at finite intervention size.

Prospective terminal prediction. At 75% of training, we compare four rank-8 Direct interventions, one for each of CURRENT, FORM, MARGINAL, and PERMUTED FORM. Each intervention applies a persistent displacement to the otherwise frozen base weights; adapters and optimizer states then evolve separately in the candidate branches. The branches share the pre-intervention state, future data sequence, randomness, and optimizer protocol (Appendix B). CURRENT also receives a projected-gradient intervention and is not the no-intervention reference.

What is gained beyond simple selection rules? Within each primary cohort, the observed complete fidelity ordering is unchanged across the five evaluated states, so F and $-F$ select fixed candidate



(a) Zero-initialized first step.



(b) Controlled nonzero moments.

Figure 2: **Receiver implementation can change candidate rankings. (a)** Eight matched examples at one Qwen3-8B state: Direct preserves the fidelity preference for FORM; LoRA+AdamW with zero-initialized optimizer moments reverses it. **(b)** After $K = 16$ calibration steps with model parameters held fixed, the first-moment-only receiver prefers FORM in all 40 trials. Adding the second-moment denominator while holding each candidate's first moment fixed changes the preference to PERMUTED FORM in 34/40 trials across five states.

Table 2: Primary terminal top-1 results at $p = 0.75$: correct selections out of five independently trained states per cohort. $-F$ selects the lowest-fidelity candidate, which is CURRENT in both language cohorts and PERMUTED FORM in Vision. The best fixed rule is always-CURRENT in each cohort, identified post hoc, and therefore coincides with $-F$ in the two language cohorts. Baseline comparisons are retrospective; T_{full} predictions were frozen before candidate outcomes. Spearman results are in Appendix J.

Cohort	Model / task	F	$-F$	Best fixed (post hoc)	T_{full}
Language A	RoBERTa / SST-2	1/5	4/5	4/5	4/5
Language B	RoBERTa / SST-2	0/5	3/5	3/5	4/5
Vision	DeiT-S / CIFAR-100	0/5	2/5	3/5	5/5

names in these samples, although the underlying bases and score values can vary. Fidelity has negative median within-state Spearman correlation with terminal quality in all three cohorts (-0.8, -0.4, and -0.8 for Language A, Language B, and Vision, respectively; Appendix J), consistent with the strong performance of reversed fidelity. Under Direct, reversed fidelity also selects the smallest update, so its performance cannot be attributed separately to candidate direction or intervention size. Applying $-F$ needs no terminal information, whereas identifying the best fixed rule uses the observed outcomes. Relative to this post hoc rule, T_{full} has the same top-1 count in Language A, one more correct selection in Language B, and two more in Vision (Table 2); these aggregate gains do not imply lower regret in every state (Appendix N.3).

Does objective matching explain the improvement? One alternative is that T_{full} helps simply because it uses the terminal objective. A retrospective switch-time control, $S_{\text{holdout},s}(V) = -\langle \nabla J(z_s), e_V \rangle$, uses that same objective without future differentiation. Its top-1 counts are $2/5$, $0/5$, and $2/5$ for Language A/B/Vision, versus $4/5$, $4/5$, and $5/5$ for T_{full}. Objective matching alone therefore does not explain the recovery (Appendix N.1).

Does either component of T_{full} suffice? In Language A, neither the retained direct component (T_{direct}: $3/5$) nor the transported component ($T_{\text{transported}}$: $2/5$) alone reaches the full score's $4/5$ terminal-winner count (Appendix H).

Does the pattern persist at an earlier switch? To assess robustness across switch times, we revisit the same five Language A seed identities at $p = 0.50$, with predictions fixed before executing all 20 candidate branches. Median pairwise accuracy is 0.167 for F and 0.833 for T_{full}; pairwise accuracy improves in three states and ties in two. Top-1 is $0/5$ for F and $3/5$ for T_{full}, matching both $-F$ and post hoc always-CURRENT; relative to always-CURRENT, T_{full} has lower terminal-loss regret in

two states and equal regret in three (Appendices N.3 and O). The earlier switch also changes the state, candidate geometry, and remaining training length, making this a robustness check across switch times.

Does the fidelity–terminal mismatch persist under other construction rules? We compare Dominant-SVD, SARA-style importance sampling, and Random-Haar candidates (Zhao et al., 2024; Zhang et al., 2025a; He et al., 2025) on the same ten Language A+B states under a common Direct receiver (Figure 1). Across these states, median pairwise accuracy is 0.333 for fidelity and 1.0 for T_{full}. Thus, the fidelity–terminal mismatch persists beyond the original candidate constructions.

4.2 First-Order Comparisons on a Finite Candidate Set

Having tested future sensitivity against terminal outcomes, we now ask which components of future sensitivity matter for first-order comparisons within a fixed candidate set. Pairwise score differences depend only on sensitivity along candidate differences: components orthogonal to the span of those differences cancel. With four candidates, this span has dimension at most three.

The span of candidate differences. For a fixed finite set of interventions $C = \{e_1, \dots, e_K\} \subset \mathcal{E}$, where $\mathcal{E}$ is the intervention space and $e_i = e_{V_i}$, define

$$\mathcal{U}_C = \operatorname{span}\{e_i - e_j : e_i, e_j \in C\}, \qquad \dim \mathcal{U}_C \le K - 1. \tag{8}$$

Let $\ell_{\text{fut}} : \mathcal{E} \to \mathbb{R}$ be the linear future-valuation functional satisfying $T_{\text{full}}(V) = -\ell_{\text{fut}}(e_V)$, so $\Delta T_{ij} := T_{\text{full}}(V_i) - T_{\text{full}}(V_j) = -\ell_{\text{fut}}(e_i - e_j)$.

Proposition 3 (Invariance on a finite candidate set) *For a fixed finite C, all first-order pairwise score differences depend only on the restriction of ℓ_{fut} to $\mathcal{U}_C$. Equivalently, adding any linear functional $\nu : \mathcal{E} \to \mathbb{R}$ that vanishes on $\mathcal{U}_C$ leaves every ΔT_{ij} unchanged.*

Two future functionals that agree on $\mathcal{U}_C$ therefore induce identical pairwise score differences on the candidate set (quotient formulation and proof in Appendix E.1). For a linear approximation $\widehat{\ell}_{\text{fut}}$, let $\widehat{\Delta T}_{ij} = -\widehat{\ell}_{\text{fut}}(e_i - e_j)$; then $\widehat{\Delta T}_{ij} - \Delta T_{ij} = -(\widehat{\ell}_{\text{fut}} - \ell_{\text{fut}})(e_i - e_j)$. Thus only approximation error on $\mathcal{U}_C$ affects pairwise differences; for $\Delta T_{ij} \neq 0$, an induced absolute pairwise-score error below $|\Delta T_{ij}|$ preserves the sign of ΔT_{ij}.

A signed sufficient condition. Within this candidate-visible space, we ask whether future valuation is close to a signed rescaling of the local functional. Using the Frobenius inner product, let g_{loc} and g_{fut} represent the Direct local and future functionals, so that $S_R(V) = -\langle g_{\text{loc}}, e_V \rangle$ and $T_{\text{full}}(V) = -\langle g_{\text{fut}}, e_V \rangle$. Project them to $a = P_{\mathcal{U}_C} g_{\text{loc}}$ and $b = P_{\mathcal{U}_C} g_{\text{fut}}$. When $a \neq 0$,

$$\lambda_U = \frac{\langle b, a \rangle}{\|a\|^2}, \qquad r_U = b - \lambda_U a, \qquad \langle r_U, a \rangle = 0. \tag{9}$$

With $\Delta S_{ij} = S_R(V_i) - S_R(V_j)$, the exact pairwise relation is

$$\Delta T_{ij} = \lambda_U \Delta S_{ij} - \langle r_U, e_i - e_j \rangle. \tag{10}$$

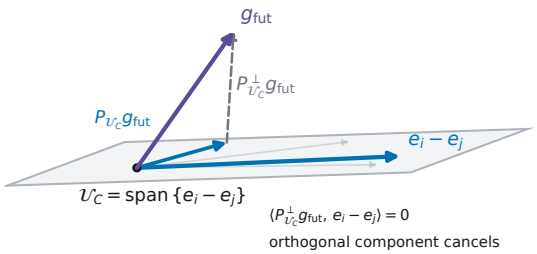
Whenever

$$|\lambda_U \Delta S_{ij}| > \|r_U\| \, \|e_i - e_j\|, \tag{11}$$

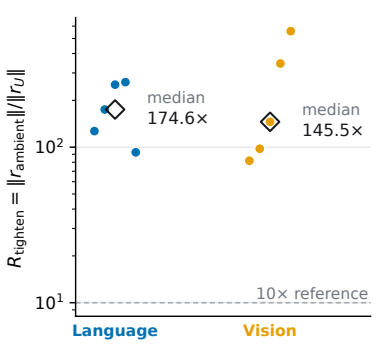
the first term determines the sign, giving $\operatorname{sign}(\Delta T_{ij}) = \operatorname{sign}(\lambda_U \Delta S_{ij})$. We call pairs satisfying Eq. 11 *certified*. For certified pairs, a positive λ_U certifies first-order preservation, whereas a negative λ_U certifies first-order reversal. Degenerate cases and tie conventions are in Appendix E.1.

Why the restriction matters empirically. On six archived discovery states used to motivate this restriction, an earlier full-space bound certified $0/36$ pairs. The candidate-difference span contained under 0.3% of future-sensitivity energy while retaining 60–84% of local-gradient energy, leaving the full-space residual dominated by components irrelevant to candidate comparisons.

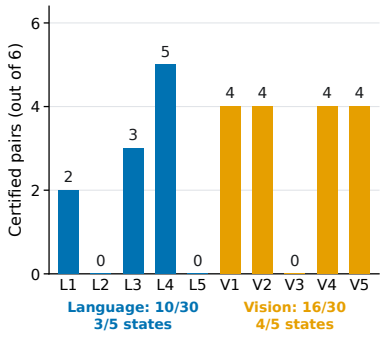
Evaluation on new states. With the candidate set, thresholds, and condition frozen in advance, we evaluate five new language and five new vision states, disjoint from all discovery states and other cohorts. The fitted λ_U is negative in $9/10$ states, including all seven states with at least one certified pair. The condition certifies $10/30$ language pairs across $3/5$ states and $16/30$ vision pairs across $4/5$, so all 26 certified pairs are reversals. We use the same λ_U for both residuals.



(a) Only sensitivity along candidate differences affects first-order comparisons.



(b) Reduction in residual norm.



(c) Certified first-order reversals.

Figure 3: **First-order comparisons depend only on future sensitivity along candidate differences.** **(a)** Schematic decomposition of future sensitivity into its candidate-visible component $P_{\mathcal{U}_C} g_{\text{fut}}$ and orthogonal component $P_{\mathcal{U}_C}^{\perp} g_{\text{fut}}$. Because $e_i - e_j \in \mathcal{U}_C$, the orthogonal component cancels from pairwise first-order differences. **(b–c)** Five new language and five new vision states, four candidates each. **(b)** Ratio of full-space to projected residual norms, both using the same fitted λ_U; each point is one state. **(c)** Pairs per state certified as first-order reversals by Eq. 11.

The median full-space-to-projected residual ratio is $174.6\times$ for language and $145.5\times$ for vision (Figure 3; Appendix E.2).

Do the certified rankings persist at the terminal outcome? After freezing the 26 certified pairs and their predicted signs, we execute all four candidate continuations on the same ten states. The terminal ordering reverses the local ordering in all 26 certified pairs and in 22 of the 34 uncertified pairs. T_{full} matches the terminal ordering in all 26 certified pairs and in 30 of the 34 uncertified pairs. Under Direct, F and S_R are co-ordered; because all 26 certified pairs are first-order reversals, T_{full} and $-F$ agree on them. Among the 34 uncertified pairs, the scores disagree on 14; T_{full} matches the terminal ordering in 11 and $-F$ in three (Appendix N.2). Eq. 11 is a sufficient condition for first-order sign transfer. The terminal evaluations support finite-intervention persistence of the certified predictions, while certification remains a first-order condition rather than a terminal-outcome classifier.

5 Related Work

Gradient-informed low-rank construction and allocation. GaLore uses dominant gradient subspaces (Zhao et al., 2024); LoRA-GA and LoRA-One use full gradients for initialization or alignment (Wang et al., 2024; Zhang et al., 2025b). LoRA-Pro adjusts factor gradients so the induced low-rank gradient better approximates the full-fine-tuning gradient (Wang et al., 2025). AdaLoRA, AdaRankGrad, RaLoRA, and IGU-LoRA adapt rank using gradient-derived signals (Zhang et al., 2023; Refael et al., 2025; Ye et al., 2026a; Cui et al., 2026). We study the decision validity of gradient-fidelity preferences after construction, under receiver implementation and subsequent training.

Changing subspaces and receiver semantics. ReLoRA and Flora merge, reset, or resample low-rank coordinates (Lialin et al., 2024; Hao et al., 2024); Online Subspace Descent, Breaking the Frozen Subspace, and GoLore update or sample subspaces during training (Liang et al., 2024; Zhang et al., 2025a; He et al., 2025). LoRA-RITE studies transformation invariance (Yen et al., 2025), while LDAdam, SubTrack++, LoRA-Pre, LoFT, and LORO/Riemannion adapt optimizer geometry to

low-rank coordinates (Robert et al., 2025; Rajabi et al., 2025; Wang et al., 2026; Tastan et al., 2026; Mo et al., 2025; Bogachev et al., 2026). We instead hold candidate geometry fixed and test whether the receiver preserves its fidelity ordering.

Gradient information and decision quality. Gur-Ari et al. and Song et al. study whether concentrated gradient directions are optimization-relevant (Gur-Ari et al., 2018; Song et al., 2025). We examine fidelity-based candidate preferences through implementation and subsequent training. For the latter, we use established finite-trajectory differentiation (Domke, 2012; Maclaurin et al., 2015; Franceschi et al., 2017). Data-driven linear optimization characterizes decision-relevant information (Bennouna et al., 2025; Ye et al., 2026b); our Proposition 3 gives a first-order characterization of the future information relevant to rankings within a fixed finite candidate set.

6 Discussion and Limitations

Evaluating candidates under the intended training procedure. Gradient capture and candidate preference answer different questions. Higher fidelity means that a candidate captures a larger fraction of the current gradient energy, but whether that candidate is preferable depends on how the receiver implements its update and how subsequent training values the intervention. For a fixed candidate set, first-order comparisons depend only on future sensitivity along candidate differences. These results motivate evaluating candidate preferences under the intended receiver and decision horizon, rather than treating fidelity as a stand-alone selection criterion. This two-stage view suggests two diagnostics. Isolated shadow updates can test whether the intended receiver preserves fidelity ordering in local first-order value. For approximate future sensitivity, assess its induced errors in candidate-pair score differences relative to the gaps they could overturn. Projecting onto the candidate-difference span focuses the error analysis on ranking-relevant components. The geometry identifies the decision-relevant information, while efficient approximation of that information is a separate computational problem.

Effect size and computational value. Across the 15 primary states, candidate terminal-loss spreads range from 0.05% to 2.7% of terminal loss. In the quotient cohort, absolute relative deviations from no intervention average 0.11% and reach 0.50% (Appendices J, N.2). Reversed fidelity and fixed-candidate rules already recover many terminal winners; T_{full} matches or exceeds the best fixed rule's aggregate top-1 accuracy in every evaluated setting, while its statewise regret gains vary (Appendix N.3). The current evidence supports ranking value, while the computational value of future information remains unresolved. A preregistered adaptive-lookahead rule reduces measured decision GPU time by 43.81% while matching median regret, at the cost of higher mean regret and lower top-1 accuracy, yielding a cost–quality tradeoff (Appendix I).

Scope of the evidence. Primary cohorts contain five states each; the reported effects characterize these evaluated states and settings. Receiver evidence covers a zero-moment step and a controlled $K = 16$ nonzero-memory protocol; optimizer memory acquired during ordinary training remains outside the current evidence. Primary terminal-ranking experiments use a switch at $p = 0.75$, with an earlier $p = 0.50$ check on the same seed identities. A separate fresh five-state cohort tests multiple lookahead fractions q for future scoring (Appendix I). Rank and candidate pools are fixed, with online adaptation left for future study. Both T_{full} and the signed condition use reference-future information.

7 Conclusion

When does gradient fidelity support a low-rank optimization decision? For local first-order comparisons, Proposition 1 gives a sufficient condition under which the intended receiver preserves the fidelity ordering. Under Direct, when Eq. 11 holds, the pair's fidelity ordering is preserved in first-order future value for $\lambda_U > 0$ and reversed for $\lambda_U < 0$. Both conditions are first-order; finite terminal rankings are evaluated through separately executed candidate continuations. In our experiments, zero-moment LoRA+AdamW reverses the fidelity ordering in all eight matched comparisons at one Qwen3-8B state. Under Direct, fidelity has negative median rank correlation with terminal quality in each primary cohort, whereas the first-order future score identifies more terminal winners than fidelity, while its gains over simple selection rules are smaller. Candidate-selection analysis should therefore evaluate candidates under the intended receiver and decision horizon, and assess future-score errors along the directions that distinguish the candidates.

AI USE STATEMENT

Generative AI tools were used to assist with code editing and implementation, as well as the drafting and editing of manuscripts and internal research texts. The authors verified all AI-generated suggestions before adopting them. The authors reviewed all AI-assisted work and bear full responsibility for the final content, arguments, and outputs of this submission.

REPRODUCIBILITY STATEMENT

The appendices document cohort provenance and independence, training and continuation protocols, candidate and receiver constructions, future-sensitivity computation and the timing and verification of prospective predictions, per-state results, numerical checks, and additional decision-validity controls. Complete proofs of Propositions 1, 2, and 3 are provided in Appendices L, M, and E.1, respectively. Appendix C describes how to apply the local comparison and interpret its limits. These materials specify the experimental and theoretical objects underlying the main-text claims.

REFERENCES

Omar Bennouna, Amine Bennouna, Saurabh Amin, and Asuman Ozdaglar. What data enables optimal decisions? an exact characterization for linear optimization. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-5534.

Vladimir Bogachev, Vladimir Aletov, Alexander Molozhavenko, Denis Bobkov, Vera Soboleva, Aibek Alanov, and Maxim Rakhuba. LoRA meets riemannion: Muon optimizer for parametrization-independent low-rank adapters. In *International Conference on Learning Representations*, 2026.

Xuan Cui, HuiYue Li, Run Zeng, Yunfei Zhao, Jinrui Qian, Wei Duan, Bo Liu, and Zhanpeng Zhou. IGU-LoRA: Adaptive rank allocation via integrated gradients and uncertainty-aware scoring. In *International Conference on Learning Representations*, 2026.

Justin Domke. Generic methods for optimization-based modeling. In Neil D. Lawrence and Mark Girolami (eds.), *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, volume 22 of *Proceedings of Machine Learning Research*, pp. 318–326, La Palma, Canary Islands, 21–23 Apr 2012. PMLR.

Luca Franceschi, Michele Donini, Paolo Frasconi, and Massimiliano Pontil. Forward and reverse gradient-based hyperparameter optimization. In Doina Precup and Yee Whye Teh (eds.), *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1165–1173. PMLR, 06–11 Aug 2017.

Guy Gur-Ari, Daniel A. Roberts, and Ethan Dyer. Gradient descent happens in a tiny subspace. *arXiv preprint arXiv:1812.04754*, 2018.

Yongchang Hao, Yanshuai Cao, and Lili Mou. Flora: Low-rank adapters are secretly gradient compressors. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 17554–17571. PMLR, 21–27 Jul 2024.

Yutong He, Pengrui Li, Yipeng Hu, Chuyan Chen, and Kun Yuan. Subspace optimization for large language models with convergence guarantees. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 22468–22522. PMLR, 13–19 Jul 2025.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022.

Vladislav Lialin, Sherin Muckatira, Namrata Shivagunde, and Anna Rumshisky. ReLoRA: High-rank training through low-rank updates. In *International Conference on Learning Representations*, 2024.

Kaizhao Liang, Bo Liu, Lizhang Chen, and Qiang Liu. Memory-efficient LLM training with online subspace descent. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-2054.

Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

Dougal Maclaurin, David Duvenaud, and Ryan Adams. Gradient-based hyperparameter optimization through reversible learning. In Francis Bach and David Blei (eds.), *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 2113–2122, Lille, France, 07–09 Jul 2015. PMLR.

Zhanfeng Mo, Long-Kai Huang, and Sinno Jialin Pan. Parameter and memory efficient pretraining via low-rank riemannian optimization. In *International Conference on Learning Representations*, 2025.

Sahar Rajabi, Nayeema Nonta, and Sirisha Rambhatla. SubTrack++: Gradient subspace tracking for scalable LLM training. In *Advances in Neural Information Processing Systems*, volume 38, 2025. doi: 10.52202/085713-1060.

Yehonathan Refael, Jonathan Svirsky, Boris Shustin, Wasim Huleihel, and Ofir Lindenbaum. AdaRankGrad: Adaptive gradient rank and moments for memory-efficient LLMs training and fine-tuning. In *International Conference on Learning Representations*, 2025.

Thomas Robert, Mher Safaryan, Ionut-Vlad Modoranu, and Dan Alistarh. LDAdam: Adaptive optimization from low-dimensional gradient statistics. In *International Conference on Learning Representations*, 2025.

Minhak Song, Kwangjun Ahn, and Chulhee Yun. Does SGD really happen in tiny subspaces? In *International Conference on Learning Representations*, 2025.

Nurbek Tastan, Stefanos Laskaridis, Martin Takáč, Karthik Nandakumar, and Samuel Horváth. LoFT: Low-rank adaptation that behaves like full fine-tuning. In *International Conference on Learning Representations*, 2026.

Shaowen Wang, Linxi Yu, and Jian Li. LoRA-GA: Low-rank adaptation with gradient approximation. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-1741.

Zhengbo Wang, Jian Liang, Ran He, Zilei Wang, and Tieniu Tan. LoRA-Pro: Are low-rank adapters properly optimized? In *International Conference on Learning Representations*, 2025.

Zhengbo Wang, Jian Liang, Ran He, Zilei Wang, and Tieniu Tan. Taming momentum: Rethinking optimizer states through low-rank approximation. In *International Conference on Learning Representations*, 2026.

Jingqi Ye, Haonan He, Minglei Li, Fujun Han, Tao Chen, and Peng Ye. Gradient intrinsic dimensionality alignment: Narrowing the gap between low-rank adaptation and full fine-tuning. In *International Conference on Learning Representations*, 2026a.

Yuhan Ye, Saurabh Amin, and Asuman Özdağlar. Learning decision-sufficient representations for linear optimization. In Steve Hanneke and Tor Lattimore (eds.), *Proceedings of Thirty Ninth Conference on Learning Theory*, volume 336 of *Proceedings of Machine Learning Research*, pp. 6924–6975. PMLR, 29 Jun–03 Jul 2026b.

Jui-Nan Yen, Si Si, Zhao Meng, Felix Yu, Venkata Sai Surya Subramanyam Duvvuri, Inderjit Dhillon, Cho-Jui Hsieh, and Sanjiv Kumar. LoRA done RITE: Robust invariant transformation equilibration for LoRA optimization. In *International Conference on Learning Representations*, 2025.

Haochen Zhang, Junze Yin, Guanchu Wang, Zirui Liu, Lin F. Yang, Tianyi Zhang, Anshumali Shrivastava, and Vladimir Braverman. Breaking the frozen subspace: Importance sampling for low-rank optimization in LLM pretraining. In *Advances in Neural Information Processing Systems*, volume 38, 2025a. doi: 10.52202/085713-0343.

Qingru Zhang, Minshuo Chen, Alexander Bukharin, Pengcheng He, Yu Cheng, Weizhu Chen, and Tuo Zhao. Adaptive budget allocation for parameter-efficient fine-tuning. In *International Conference on Learning Representations*, 2023.

Yuanhe Zhang, Fanghui Liu, and Yudong Chen. LoRA-one: One-step full gradient could suffice for fine-tuning large language models, provably and efficiently. In Aarti Singh, Maryam Fazel, Daniel Hsu, Simon Lacoste-Julien, Felix Berkenkamp, Tegan Maharaj, Kiri Wagstaff, and Jerry Zhu (eds.), *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 75513–75574. PMLR, 13–19 Jul 2025b.

Jiawei Zhao, Zhenyu Zhang, Beidi Chen, Zhangyang Wang, Anima Anandkumar, and Yuandong Tian. GaLore: Memory-efficient LLM training by gradient low-rank projection. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp (eds.), *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 61121–61143. PMLR, 21–27 Jul 2024.

A Experimental Evidence and Cohort Provenance

Cohort provenance. Language Cohort A uses seeds 20260880–20260884; the independently preregistered Language Cohort B uses 20260885–20260889. The vision confirmation uses seeds 20260909–20260913. The decision-challenge block explicitly reuses the ten Language A/B states and changes only the candidate-rule family; it is therefore rule-level breadth, not ten new independent states. The horizon audit uses a separate restart cohort, seeds 20260914–20260918, and the evidence-acquisition block uses a further fresh cohort, seeds 20260921–20260930. An earlier horizon-audit execution root was aborted before candidate construction because of a seed-registration collision and contributes no scientific result. For the horizon claim in Appendix I, the canonical table is the preregistered T_q/local-metrics table; a separate matched-short-rollout table in the audit record is treated as a secondary diagnostic and is not used for the headline horizon numbers.

Earlier-switch extension. The $p = 0.50$ test uses the five Language A seed identities, 20260880–20260884, at an earlier training checkpoint. Its 20 candidate continuations extend switch-time coverage; they do not constitute a new independent seed cohort. Results are given in Table 18, with execution details in Appendix O.

Controlled nonzero moments. The $K = 16$ receiver control reuses the five independently trained states from the zero-moment receiver replication, with eight FORM/PERMUTED FORM comparisons per state. Its receiver-state comparison is distinct from the original zero-initialized protocol; the state identities and reported results are listed in Appendix P. We report this protocol separately from the zero-initialized comparison, but it does not constitute an additional independent state cohort.

Statistical reporting. Given the small preregistered cohorts ($n = 5$ or 10), we report paired seed-level outcomes, discrete top-1/pairwise counts, medians, and regret comparisons rather than asymptotic significance tests. Examples and candidate pairs within a training state are matched comparisons, not independent state-level replications. Figure 9 and Section J report the paired outcomes rather than relying only on aggregate error bars.

B Experimental Protocols and Cohort Independence

Common language trajectory semantics. For the prospective RoBERTa trajectory cohorts, the frozen protocol uses rank $r = 8$ and a switch at normalized progress $p = 0.75$ (step 4700 of 6267). Direct interventions use

$$\Delta W_{m,V} = -\eta_s G_{m,s} V_m V_m^\top, \tag{12}$$

Table 3: Evidence supporting the main claims. "Prospective" means the score/rule was frozen before the relevant candidate outcome was opened.

Block	Role	Frozen result / boundary
Matched receiver	Fidelity → implemented-update score	One fixed Qwen3-8B state; 8 matched utility examples, not independent states/seeds. Direct: 0/8 reorderings; LoRA+AdamW: 8/8. Receiver-aware utility-order agreement 6/8 vs. fidelity 2/8.
Language cohort A	Prospective Local → Terminal	RoBERTa/SST-2, $n = 5$: top-1 1/5 local vs. 4/5 T_{full}; median Spearman -0.8 vs. $+1.0$. One top-1 prediction error; small terminal effect spread. Post hoc best-fixed top-1: 4/5.
Language cohort B	Independent prospective replication	Fresh RoBERTa/SST-2, $n = 5$: top-1 0/5 local vs. 4/5 T_{full}; paired association, ordering, and regret improve in the same direction. Post hoc best-fixed top-1: 3/5.
Vision trajectory	Independent cross-family confirmation	DeiT-S/CIFAR-100, $n = 5$: top-1 0/5 local vs. 5/5 T_{full}; median Spearman -0.8 vs. $+1.0$. Post hoc best-fixed top-1: 3/5.
Decision challenge	Prospective rule breadth	Reuses language cohorts A/B with Dominant-SVD, SARA-style, and random rule analogues: local top-1 3/10 vs. T_{full} 9/10; median per-state pairwise accuracy 0.333 vs. 1.0. Broadens candidate rules, not state count. Post hoc always-Random-Haar top-1: 7/10.
Horizon audit	Separate fresh prospective cohort	Fresh RoBERTa/SST-2, $n = 5$: $q \in \{0, .05, .10, .25, .50, 1.00\}$ gives top-1 $1, 3, 2, 3, 4, 3$ of 5; $q = 1.00$ is full horizon within this cohort, not Language A.
Evidence acquisition	Further separate fresh prospective cohort	Fresh RoBERTa/SST-2, $n = 10$: frozen short-horizon disagreement escalation matches median regret of always-$T_{.50}$ with 43.81% lower measured decision GPU time; mean regret worse, top-1 6/10 vs. 7/10.
Quotient confirmation	Prospective decision-relevant geometry	Fresh $5{+}5$ language/vision states, disjoint from all cohorts above and from the six discovery states: 10/30 language and 16/30 vision pairs certified (3/5, 4/5 state breadth); median tightening 174.6×/145.5× (Appendix E.2).
Objective-matched local control	Isolate objective matching from future transport	Retrospective; same held-out terminal TRAIN objective evaluated at the switch, no future differentiation. $S_{\mathrm{holdout},s}$ does not close the T_{full} gap in any of the three primary cohorts or the reused decision-challenge cohort (Appendix N.1).
Terminal evaluation	First-order certificate → finite terminal outcome	10 fresh quotient-confirmation states; 40 canonical single-scale candidate continuations. 26/26 pre-fixed certified reversal pairs preserve the T_{full} ordering at terminal Q; the corresponding agreement is 30/34 for uncertified pairs. This is an empirical completion, not a terminal-Q certificate (Appendix N.2).

with one candidate-independent positive scale η_s fixed from a canonical eager+fp32 shadow step. This is the right-action realization $\Pi_{V_m}(G_{m,s}) = G_{m,s}P_{V_m}$, with $P_{V_m} = V_m V_m^\top$. The continuation uses AdamW (learning rate 2×10^{-4}, weight decay 0.01, $\beta = (0.9, 0.999)$, $\epsilon = 10^{-8}$, gradient clipping 1.0) with a three-epoch schedule and 6% linear warmup. The terminal objective is mean cross-entropy on a fixed held-out TRAIN subset. Future data order and stochastic replay are fixed across candidates, and the derivative-consistent path uses eager attention and fp32 throughout.

Vision port. The DeiT-S/CIFAR-100 cohort preserves the same rank-8, single-switch, common-reference trajectory semantics while using the frozen vision recipe: learning rate 10^{-4}, batch size 64, 10 epochs, 5% warmup, and image size 224. The only deliberate implementation-path change relative to the historical vision recipe is eager+fp32 in place of the bf16/autocast path so that the primal and derivative maps are identical. Official GLUE validation and CIFAR-100 TEST were also evaluated as secondary diagnostics; the vision TEST evaluation followed the primary results. These secondary evaluators do not define J or Q for the main-text terminal-ranking comparisons.

Table 4: Frozen cohort identities and independence relationships recovered from preregistrations, executable configs, and result ledgers.

Block	Model / task	Scientific units	Independence / reuse
Qwen receiver	Qwen3-8B-Base / GSM8K	One step-64 state; 8 matched held-out example-level comparisons	Independent of all RoBERTa/DeiT cohorts; the 8 comparisons are not independent seeds.
Language A	RoBERTa-base / SST-2	seeds 20260880–20260884 ($n = 5$)	Fresh prospective cohort.
Language B	RoBERTa-base / SST-2	seeds 20260885–20260889 ($n = 5$)	Independent prospective replication of the frozen Language-A protocol.
Vision	DeiT-S / CIFAR-100	seeds 20260909–20260913 ($n = 5$)	Independent cross-family confirmation.
Decision challenge	RoBERTa-base / SST-2	reuses Language A+B ($n = 10$ states)	Same states/futures, new Dominant-SVD/SARA-IS/Random-Haar candidates.
Horizon audit	RoBERTa-base / SST-2	seeds 20260914–20260918 ($n = 5$)	Separate fresh restart cohort; earlier pre-candidate root aborted.
Evidence acquisition	RoBERTa-base / SST-2	seeds 20260921–20260930 ($n = 10$)	Further separate fresh actionability cohort.

Table 5: Held-fixed and varied factors for the main-text controlled comparisons. Each comparison addresses a specific alternative explanation.

Comparison	Held fixed	Varied or evaluated
Receiver mechanism	Training state, examples, candidate geometry, and matched installation for component controls	Update rule; fidelity ranking, implemented-update ranking, and actual one-step utility
Nonzero-moment control	Candidate geometry, parameters, and each candidate's first moment within the matched comparison	Presence of the coordinatewise second-moment denominator; receiver-score preference
Terminal prediction	Switch state, rank, receiver, continuation conditions, terminal objective	Candidate intervention; local and future scores against separately obtained terminal outcomes
Switch-time robustness	The experimental system and five Language A seed identities	An earlier switch checkpoint, which also changes the state, candidate geometry, and remaining training
Objective-matched control	Candidate interventions and the terminal evaluator	A switch-time derivative versus a derivative through future training
Candidate-rule comparison	The ten language states, Direct receiver, continuation, and evaluation	Candidate construction rules
Quotient condition	Within each new state, candidates and the fitted gain used in the residual comparison	Full-space versus projected residual; certification and subsequent terminal agreement

Prospective integrity. Across the prospective cohorts, freeze scripts refuse to seal predictions if a terminal-branch outcome already exists, hash the outcome-blind inputs used to construct candidates and scores, and require downstream analyses to re-verify those hashes.

C CANDIDATE AND RECEIVER DEFINITIONS

Table 6: Candidate constructions used by the current paper. All constructions are outcome-independent and frozen before terminal branches are opened.

Candidate	Frozen construction
CURRENT	Orthonormal basis of the installed LoRA-A row space at the switch checkpoint, obtained by deterministic QR; depends on the current training state, not on future outcomes.
FORM	Top-r eigenvectors of an accumulated formation-coupling Gram matrix built from activation/upstream-gradient pairs on the candidate-fit partition.
MARGINAL	Same estimator pipeline as FORM with the coupling core replaced by the marginal-only core.
PERMUTED FORM	A FORM basis computed for a different target matrix of matching input dimension and assigned through a deterministic, SHA-derived derangement; no reroll.
DOMINANT_SVD	Top-r right singular directions of the archived switch-state gradient; deterministic and fidelity-maximizing by construction.
SARA_IS	r right-singular directions importance-sampled without replacement with probability proportional to singular value; one SHA-derived draw per seed/matrix, no reroll.
RANDOM_HAAR	Haar-random rank-r orthonormal frame obtained by Gaussian QR with sign-fixed diagonal under the frozen SHA-derived seed.

Permutation terminology. Decision-experiment `PERMUTED_FORM` is a cross-target basis derangement, distinct from the sample-pairing (pairing-broken) control in Appendix F.

Method-analogue boundary. Dominant-SVD, SARA-IS, and Random-Haar reproduce matched *subspace-selection rules* on frozen states. They are not end-to-end benchmarks of GaLore, SARA, GoLore, Flora, or other named optimizers.

FORM/MARGINAL construction (executable definition). Let $\{(X_b, D_b)\}_{b=1}^{B}$ be the per-batch activation/upstream-gradient matrices on the frozen candidate-fit partition, $X_b \in \mathbb{R}^{n_b \times d_{\text{in}}}$, $D_b \in \mathbb{R}^{n_b \times d_{\text{out}}}$. Per batch, take truncated SVDs $X_b \approx U_{X,b} S_{X,b} V_{X,b}^\top$, $D_b \approx U_{D,b} S_{D,b} V_{D,b}^\top$ at internal rank $q = \min(32, n_b, d_{\text{in}}, d_{\text{out}})$, and set $\widehat{G}_b = (V_{X,b} K_b V_{D,b}^\top)^\top$, with $K_b = S_{X,b}(U_{X,b}^\top U_{D,b}) S_{D,b}$ for FORM and $K_b = \operatorname{diag}(\sigma_{X,b,1}\sigma_{D,b,1}, \ldots, \sigma_{X,b,q}\sigma_{D,b,q})$ for MARGINAL – i.e. MARGINAL replaces the joint core K_b with the index-matched diagonal of marginal singular-value products, equivalently replacing $U_{X,b}^\top U_{D,b}$ by I_q within the FORM core. The candidate Gram is the unweighted sum $C = \sum_b \widehat{G}_b^\top \widehat{G}_b$ over all candidate-fit batches, and the basis is C's top-r eigenvectors, $r = 8$: the exact frozen construction behind every FORM/MARGINAL basis in this paper.

Direct receiver. With $P_V = VV^\top$, Direct applies the optimizer-free update

$$\Delta W_{\text{Direct}}(V) = -\eta G P_V, \tag{13}$$

where the same $\eta > 0$ is shared by all candidates in a comparison; recall $\Pi_V(G) = GP_V$ (Section 2), so this instantiates $\Delta W_{\text{Direct}}(V) = -\eta \Pi_V(G)$. Therefore

$$S_{\text{Direct}}(V) = \eta \|GP_V\|_F^2 = \eta \|\Pi_V(G)\|_F^2 = \eta \|G\|_F^2 F(V) \tag{14}$$

and fidelity ordering is mathematically preserved. The archived receiver audit verifies this identity to maximum relative error 1.09×10^{-15}.

LoRA+AdamW receiver. Following LoRA (Hu et al., 2022), the effective weight is $W_{\text{eff}} = W_0 + cBA$, with $c = \alpha/r = 2$. The checkpoint adapter is merged into W_0 before candidate installation. The factor c is kept separate from A_{pre} and enters both the factor gradient and the reconstruction of the effective weight update. For $V \in \mathbb{R}^{d_{\text{in}} \times r}$ with orthonormal columns, the shadow receiver installs

$$A_{\text{pre}} = s_A V^\top, \qquad s_A = \|A_t\|_F / \|V^\top\|_F, \qquad B_{\text{pre}} = 0, \tag{15}$$

where A_t is the trained lora_A value of the currently-loaded step-64 checkpoint. Empirically $s_A \in [0.593, 0.613]$ across the eight primary matrices, identical across FORM/PERMUTED/CURRENT for a given matrix to float64 noise, and never 1, so $A_{\text{pre}} \neq V^\top$ elementwise (relative Frobenius difference 0.39–0.41). For the idealized fp64 installation $A_{\text{pre}} = s_A V^\top$, the row-space projector agrees with $VV^\top$ to within 1.7×10^{-15} relative Frobenius error. The executed receiver stores the factors in bf16, so this bound describes the idealized geometry. Candidate installation is function-preserving because $B_{\text{pre}} = 0$ exactly, verified for every arm; since $B_{\text{pre}}A_{\text{pre}} = 0$ for any A_{pre}, this fact alone does not identify the A_{pre} realization above. The realized dense action is reconstructed as

$$\Delta W_{\text{eff}} = 2\,(B_{\text{post}}A_{\text{post}} - B_{\text{pre}}A_{\text{pre}}) = 2\,B_{\text{post}}A_{\text{post}}. \tag{16}$$

The shadow AdamW step uses learning rate 10^{-4}, $\beta = (0.9, 0.999)$, $\epsilon = 10^{-8}$, zero weight decay, and gradient clipping 1.0; a freshly zero-initialized optimizer instance ($m = v = 0$, step counter 0) is constructed for every trial and every candidate, so no moment state is carried across candidates or trials. In the matched audit, S_R uses the archived canonical gradient directly, while the Direct receiver above uses a separately captured gradient from the same parent state; the two stored tensors agree exactly at the audited numerical precision, though they are not literally the same stored tensor object. S_R depends on candidate-specific receiver gain and residual terms, as defined in Eq. 6. This comparison is between composite receivers (Direct versus the full LoRA+AdamW realization); the matched first-step reconstruction below localizes the observed inversion at this Qwen state to the zero-initialized adaptive coordinate transformation specifically, without attributing it to LoRA factorization as such, and without claiming that AdamW generally causes fidelity reordering.

Same-batch utility. Realized same-batch utility is $U_R(V) = L_{\text{pre}} - L_{\text{post}}$ (larger is better), measured on the exact update example used to form the shadow update, with the model in train mode under `no_grad` for L_{post}; a paired disjoint held-out example under the same convention gives U_{heldout}. Both losses use the unmodified training objective under bf16 autocast. All eight primary target matrices are updated jointly in one AdamW step; utility is a single whole-model scalar per trial, not decomposed per matrix. A trial is a tie iff $|\Delta U_{\text{same}}| \leq 5\times$ the receiver's measured zero-delta numerical floor (exactly 0 for LoRA+AdamW); the $6/8$ vs. $2/8$ agreement counts reported in Section 3 are direct tallies, over the eight non-tied trials, against this realized-utility ordering. In three-candidate selection, LoRA+AdamW's top-1 hit rate is 0.75 versus 0.25 for raw fidelity, with median selection regret 0 versus 0.00207.

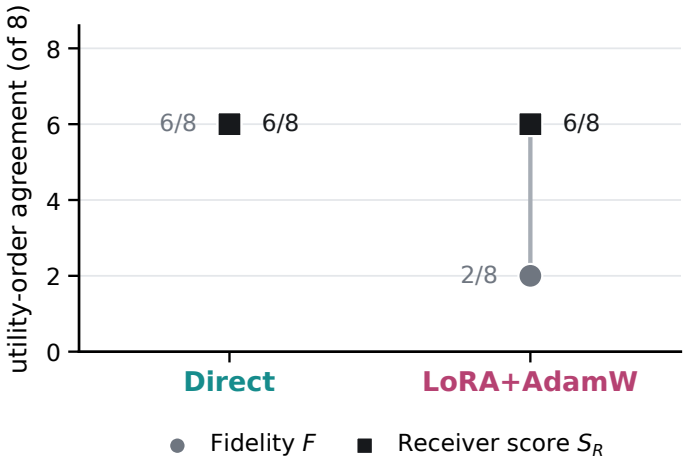


Figure 4: **Implemented-update scores and one-step utility.** Across eight matched examples at the original Qwen3-8B state, fidelity and S_R each agree with the measured utility ordering in 6/8 comparisons under Direct. Under zero-initialized LoRA+AdamW, agreement is 2/8 for fidelity and 6/8 for S_R. Agreement is evaluated against each receiver's own measured same-example whole-model loss decrease, with all target matrices updated jointly. The LoRA+AdamW results concern the zero-moment experiment in Section 3, not the $K = 16$ nonzero-moment control.

Method-native receiver breadth. Appendix Fig. 7 includes LDAdam and SubTrack++ only as descriptive method-native receiver audits. LDAdam's substitution/transport mechanics were traced

to source-line depth; SubTrack++'s stored receiver-ordering outputs were verified, but its internal transport path was not code-walked to the same depth. Neither is used for causal component attribution.

Using the local diagnostic. For a specified state and finite candidate set, retain a copy of the parameters, optimizer state, random state, and update example. Realize each candidate on an isolated copy under the receiver being evaluated, record its effective weight increment, and compute F and S_R from the same current gradient, summing over target matrices. With these updates available, directly comparing S_R answers whether the receiver preserves each local fidelity preference. The gain–residual decomposition explains the comparison; Proposition 1 supplies a sufficient margin condition, whose failure does not imply reversal. Candidate installation must preserve the current model function if it is intended to isolate update implementation; any optimizer-memory reset or transfer is part of the tested receiver and must be specified. Record unresolved comparisons separately using a tolerance fixed before candidate outcomes. Actual finite-step utility requires a separate loss evaluation, and terminal preferences require separate continuations. This diagnostic evaluates receiver-local rankings for the specified state and candidate set. Its cost includes computing the current gradient and realizing the candidate updates.

Matched first-step receiver mechanism. $B_{\mathrm{pre}} = 0 \Rightarrow g_A = 0 \Rightarrow A_{\mathrm{post}} = A_{\mathrm{pre}}$, so the ideal realized action stays in the candidate row space and the Eq. 6 residual β_V is structurally zero. With $s_V := S_R(V)/\|G\|_F^2$ (summing over target matrices before taking the ratio), the archived canonical-gradient CPU replay gives $\max|\beta_V/s_V| \approx 1.4097 \times 10^{-4}$ over all 24 trial-by-candidate cells (bf16 realization; no exclusions or denominator epsilon). This small numerical residual is a reconstruction diagnostic, not a measurement from stored executed in-branch gradients. A matched counterfactual isolates the operative operation, state/candidates/matrices/installation fixed, only the receiver operation varied (Table 7; within-state, not independent-state replication):

Table 7: Matched receiver counterfactual at the frozen Qwen state, with candidates and target matrices fixed as in Section 3. Factor-SGD rows are analytic counterfactuals; AdamW rows are reconstructions from archived canonical gradients. All comparisons use the same eight matched trials.

Matched receiver realization	FORM-vs-PERMUTED reorderings / 8
Direct projected descent	0/8
Factor-SGD (analytic; no clipping)	0/8
Factor-SGD (analytic; archived clipping)	0/8
Fresh-zero AdamW (replay; no clipping)	8/8
Fresh-zero AdamW (replay; archived clipping)	8/8

With $m{=}1, \ldots, 8$ indexing target matrices, Direct is $S_{\mathrm{Direct}}(V) = \eta \sum_m \|G_m V_m\|_F^2$; linear Factor-SGD is $S_{\mathrm{FSGD}}(V) = \gamma c^2 \sum_m s_{A,m}^2 \|G_m V_m\|_F^2$ up to the clipping factor, where $\gamma > 0$ is its common learning rate ($\gamma = 1$ in these analytic update-rule counterfactuals, not the executed AdamW learning rate). The clipping factor is one global candidate-dependent scalar across all 16 factor-gradient tensors. Direct and Factor-SGD preserve the ordering in the reported matched comparisons. Sign-saturated zero-initialized AdamW is numerically close to a weighted L_1-type functional, $S_R(V) \approx \mathrm{lr}\, c \sum_m s_{A,m} \|G_m V_m\|_1$: FORM wins squared-energy 8/8, PERMUTED wins weighted-L_1 8/8. Clipping was active in 16/24 trial-by-candidate cells (all eight FORM and eight PERMUTED FORM cells, $\kappa \in [0.185, 0.579]$); it was inactive in all eight CURRENT cells ($\kappa = 1$). Disabling clipping moves S_R only $\sim 10^{-4}$ relatively in this matched reconstruction and changes no ranking. *Provenance:* the in-branch gradients behind the archived factor gradients were not stored; this reconstruction substitutes the archived canonical gradient (pre-specified) and reproduces the archived FORM/PERM ordering 8/8 and S_R values at median relative error 0.88% (0.54%–4.18%) – a reconstruction with the numerical error reported here.

Evaluation on the full candidate set. A separate frozen audit tests whether this map explains the complete three-candidate set $\{\mathrm{FORM}, \mathrm{PERMUTED}, \mathrm{CURRENT}\}$ already in the archived audit (8 trials $\times$ 3 candidates), not only the headline pair. The verified closed-form score and its

parameter-free sign-saturated limit are

$$\widehat{S}_{\mathrm{cf}}(V) = \mathrm{lr} \sum_m \sum_{a,b} \frac{\kappa_V \, g_{B,m,ab}(V)^2}{\kappa_V \, |g_{B,m,ab}(V)| + \epsilon}, \qquad \widehat{S}_{L1}(V) = \mathrm{lr}\, c \sum_m s_{A,m}(V) \, \|G_m V_m\|_1, \tag{17}$$

with $g_{B,m}(V) = c\, s_{A,m}(V)\, G_m V_m$: $\widehat{S}_{\mathrm{cf}}$ is the exact zero-initialized AdamW first-step map on the same frozen canonical-gradient reconstruction as above (a reconstruction with the numerical error reported here), and $\widehat{S}_{L1}$ is its interpretive sign-saturated limit ($|\kappa_V g_{B,m,ab}| \gg \epsilon$), no coefficient fitted to the observed rankings. Against archived S_R over the same 8 trials as Table 7 (24 pairwise comparisons, 8 full three-candidate rankings), both $\widehat{S}_{\mathrm{cf}}$ and $\widehat{S}_{L1}$ match top-1 8/8, pairwise 24/24, and the exact full ranking 8/8 (median Spearman 1.0); raw fidelity F matches only 0/8, 16/24, 0/8 (median Spearman 0.5). Ordering agreement is exact; numerical agreement is not – median relative error 0.87% (0.52%–4.13%), consistent with the 0.88% figure above; the two independently coded evaluation routes agree to 3.4×10^{-16}. This uses the same idealized installation $A_{\mathrm{pre}} = s_A V^\top$ as above, consistent with the bf16-realized installation to below 0.3%; no S_R ties occurred. Retrospectively (not prospectively), both scores reproduce archived S_R's same-batch-utility agreement exactly (6/8 top-1, 22/24 pairwise). Once $G_m V_m$ and the receiver quantities above are available, the ordering can be evaluated in closed form without executing the shadow AdamW step ($\widehat{S}_{L1}$ needs only the weighted aggregation); agreement with receiver-local scores does not guarantee correct terminal rankings (Section 4). This is candidate-set breadth at one frozen state: all three candidates were already present in the frozen audit above, so this is a retrospective full-set evaluation, not a new cohort.

Replication across independently trained states. Five Qwen3-8B-Base LoRA runs used independent training seeds and fresh LoRA initializations under the same frozen GSM8K recipe; each was evaluated at the step-64 switch checkpoint. For each state, the FORM/PERMUTED/CURRENT candidate bases were rebuilt from that state's own step-64 activations under the same derangement rule, the same eight utility examples' canonical gradients were captured, and the LoRA+AdamW receiver was installed and stepped with read-only instrumentation archiving the true in-branch factor gradient for all $5 \times 8 \times 3 = 120$ trial/candidate cells. The original single-state reconstruction instead used archived canonical gradients. Neither $\widehat{S}_{\mathrm{cf}}$ nor $\widehat{S}_{L1}$ was refit: both use the same frozen constants ($c = 2.0$, $\mathrm{lr} = 10^{-4}$, $\epsilon = 10^{-8}$, clip$= 1.0$) applied to each fresh state's own quantities.

Table 8: Fresh-state replication of the closed-form receiver map, five independently trained Qwen3-8B-Base states, no refitting. Error is $\widehat{S}_{\mathrm{cf}}$'s relative error against realized S_R, median/max over the state's 24 trial-by-candidate score cells. Pass indicates that the preregistered acceptance criteria are met.

Fresh state	Map pairwise	Full rankings	Score check	L_1 check	Error (med./max)
Fresh-1	24/24	8/8	Pass	Pass	0.76% / 1.75%
Fresh-2	24/24	8/8	Pass	Pass	0.73% / 5.45%
Fresh-3	24/24	8/8	Pass	Pass	0.94% / 3.38%
Fresh-4	23/24	7/8	Pass	Pass	0.69% / 1.85%
Fresh-5	24/24	8/8	Pass	Pass	0.78% / 1.62%

Table 8 reports the archived scoring route. An independent float64 reconstruction gives $120/120$ pairwise agreements and $40/40$ complete-ranking agreements; the difference is confined to one nearly tied pairwise comparison. All five states clear the pre-frozen score agreement gate ($\geq 22/24$ pairwise, $\geq 6/8$ full-ranking, $\leq 5\%$ median error) and the L_1 ordering gate ($\geq 21/24$, $\geq 6/8$). Unlike the historical state, the fresh-state protocol archived the actual in-branch receiver gradients; deterministic replay passes all $120/120$ factor-gradient-replay-integrity checks (max relative error 0.054%). The parameter-free weighted-L_1 sign-limit interpretation recurs across the same five fresh states, tracking $\widehat{S}_{\mathrm{cf}}$ essentially identically within this regime. The analytic Factor-SGD counterfactual, computed from archived canonical gradients, preserves the fidelity ordering of FORM versus PERMUTED FORM in all 80 fresh-cohort comparisons (5 states $\times$ 8 trials $\times$ 2 clipping conditions), while the executed LoRA+AdamW receiver reverses that ordering in the large majority of trials in every state. The Factor-SGD clipping scalar is recomputed from the canonical gradient rather than copied from

the executed AdamW update. The two clipping conditions are deterministic evaluations of the same 40 matched trials nested within five states. This establishes within-regime state replication of the frozen receiver mechanism; cross-model, cross-task, and general AdamW behavior remain untested.

D Full-State Trajectory Sensitivity and Prediction Timing

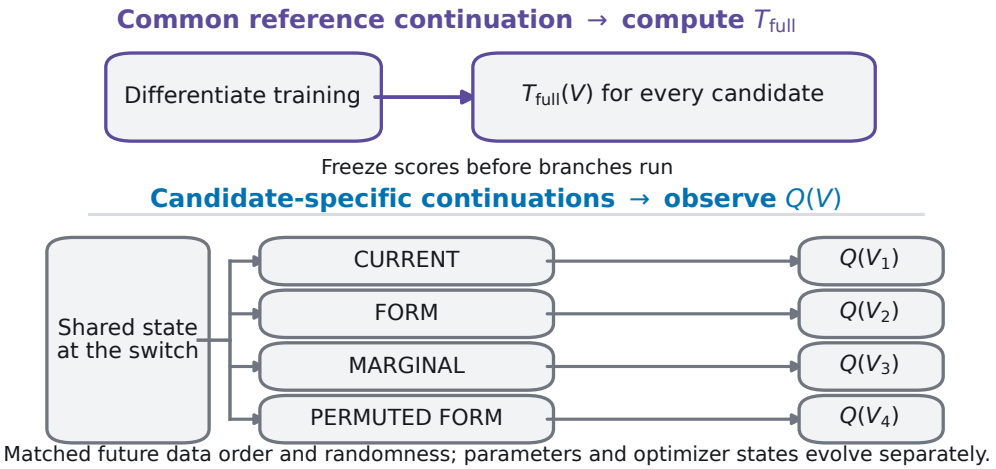


Figure 5: Controlled single-switch semantics. The upper path computes every prospective $T_{\text{full}} = T_{\text{direct}} + T_{\text{transported}}$ score along one reference continuation. After scores are fixed, the lower branches separately produce finite Q outcomes. Candidates start from the same state with matched data order and randomness; their parameters and optimizer states evolve separately. This protocol underlies Table 2.

Reported computation path. The derivative interpretation assumes that each update map is differentiable at its reference input and that the terminal objective is differentiable at the reference terminal state, with data order and randomness fixed. For every prospective headline result, the sealed score is produced by the frozen *forward-JVP tangent propagation* of the canonical training map. The tangent state follows the same optimizer/training dynamics as the reference trajectory while the future data order and stochastic sequence are treated as fixed exogenous inputs. The score contracts the terminal tangent with the gradient of the frozen terminal objective, using the definition in Eq. 4,

$$T_{\text{full}}(V) = -\langle \nabla J(z_T), \delta z_T(V) \rangle.$$

For a persistent switch of frozen base-weight blocks, the frozen implementation uses the exact accounting, with the decomposition in Eq. 5,

$$T_{\text{full}}(V) = T_{\text{transported}}(V) + T_{\text{direct}}(V),$$

where $T_{\text{transported}}$ is the effect propagated through trainable/optimizer state and T_{direct} is the terminal contribution of the persistent base-weight displacement. The correction coefficient is exactly one.

Forward- and reverse-mode agreement. The reverse-mode adjoint check is a later implementation verification, separate from the frozen forward-JVP predictions. The originally reported relative-error range of approximately 3.5×10^{-16}–4.1×10^{-15} for the score components concerns one language state. An independent contraction of archived adjoint tensors across the six discovery states reproduces the stored forward-JVP T_{full} scores with a maximum relative difference of 7.7×10^{-14}.

Common-reference future and outcome-blind scoring. All candidates at a state are scored on one common reference future: the same switch state, future data order, randomness replay, optimizer trajectory semantics, and terminal objective. Candidate-specific terminal branches are not used to construct T_{full}. The first-order score ordering is sealed before finite branches are opened. Freeze scripts refuse to write a prediction hash when an outcome artifact already exists and downstream analysis re-verifies the frozen hashes. Thus the terminal-ranking experiment tests transfer from a

common-reference linear sensitivity to separately executed finite branches rather than reading each candidate's own future outcome.

E Decision-Relevant Quotient Geometry

E.1 Theory: proof, annihilator formulation, and degeneracy conventions

Proof of Proposition 3. Fix the finite candidate set $C = \{e_1, \dots, e_K\} \subset \mathcal{E}$ and let $\mathcal{U}_C = \mathrm{span}\{e_i - e_j : e_i, e_j \in C\}$. Let $\ell, \ell' \in \mathcal{E}^*$ be two future-valuation functionals with $\ell' = \ell + \nu$ for some $\nu \in \mathcal{U}_C^\circ = \{\nu \in \mathcal{E}^* : \nu(u) = 0 \; \forall u \in \mathcal{U}_C\}$. For any $e_i, e_j \in C$, $e_i - e_j \in \mathcal{U}_C$ by construction, so $\nu(e_i - e_j) = 0$ and

$$\ell'(e_i - e_j) = \ell(e_i - e_j) + \nu(e_i - e_j) = \ell(e_i - e_j).$$

Hence $\Delta T_{ij} = -\ell(e_i - e_j) = -\ell'(e_i - e_j)$ for every pair in C: the two functionals induce identical pairwise future-order differences. Conversely, the map $\ell \mapsto (\ell(e_i - e_j))_{e_i, e_j \in C}$ factors through the quotient $\mathcal{E}^*/\mathcal{U}_C^\circ$ by definition of the annihilator, so no coarser object determines all pairwise differences in general. This proves the proposition; it is elementary linear algebra, not a claim about the training dynamics that produce ℓ_{fut}.

Riesz representation and the projector $P_{\mathcal{U}_C}$. The paper's experiments use the direct-sum Frobenius inner product over a state's full per-target-matrix action space, $\langle A, B\rangle = \sum_m \sum_{p,q} A_m[p,q] B_m[p,q]$, identifying $\mathcal{E}$ with $\mathcal{E}^*$. Under this identification $\mathcal{U}_C^\circ$ corresponds to $\mathcal{U}_C^\perp$ and the quotient map corresponds to the orthogonal projector $P_{\mathcal{U}_C}$. With reference action $e_{\mathrm{ref}} := e_{\mathrm{CURRENT}}$ (arbitrary but fixed before any alignment quantity is computed) and spanning matrix $D = [d_{\mathrm{FORM}}, d_{\mathrm{MARGINAL}}, d_{\mathrm{PERMUTED_FORM}}]$, $d_a = e_a - e_{\mathrm{CURRENT}}$, every pairwise difference in the four-candidate set lies in $\mathrm{span}(D)$ (with $d_{\mathrm{CURRENT}} := 0$), so D spans $\mathcal{U}_C$ without a separate six-column construction. Writing $K = D^\top D$ (a Gram matrix of Frobenius inner products) and K^+ its Moore–Penrose pseudoinverse, $P_{\mathcal{U}_C} g = D K^+ D^\top g$ for any g, and $\|P_{\mathcal{U}_C} g\|^2 = (D^\top g)^\top K^+ (D^\top g)$. Proposition 3 is an invariance statement about linear functionals, independent of this inner-product choice; the fitted gain, residual norm, and certificate in Eqs. 9–11 use the specified direct-sum Frobenius geometry. In the numerical implementation, the pseudoinverse is computed in float64 with `numpy.linalg.pinv(K, rcond=1e-15)`; the historical NumPy version is not independently recorded.

Proof of the signed transfer bound (Eq. 11). Use the conventions $S_R(V) = -\langle g_{\mathrm{loc}}, e_V\rangle$ and $T_{\mathrm{full}}(V) = -\langle g_{\mathrm{fut}}, e_V\rangle$. For $a \neq 0$, with $a = P_{\mathcal{U}_C} g_{\mathrm{loc}}$, $b = P_{\mathcal{U}_C} g_{\mathrm{fut}}$, $\lambda_U = \langle b, a\rangle / \|a\|^2$, and $r_U = b - \lambda_U a$, orthogonality of the least-squares residual gives $\langle r_U, a\rangle = 0$. For any pair $e_i, e_j \in C$, $e_i - e_j \in \mathcal{U}_C$, so $\langle g_{\mathrm{loc}}, e_i - e_j\rangle = \langle a, e_i - e_j\rangle$ and $\langle g_{\mathrm{fut}}, e_i - e_j\rangle = \langle b, e_i - e_j\rangle$ exactly (this is the same projection identity used in the proposition's proof, made concrete). Substituting $b = \lambda_U a + r_U$,

$$\Delta T_{ij} = -\langle b, e_i - e_j\rangle = -\lambda_U \langle a, e_i - e_j\rangle - \langle r_U, e_i - e_j\rangle = \lambda_U \, \Delta S_{ij} - \langle r_U, e_i - e_j\rangle ,$$

using $\Delta S_{ij} = -\langle a, e_i - e_j\rangle$. By Cauchy–Schwarz, $|\langle r_U, e_i - e_j\rangle| \leq \|r_U\| \, \|e_i - e_j\|$; whenever $|\lambda_U \Delta S_{ij}| > \|r_U\| \, \|e_i - e_j\|$, the residual term is too small in magnitude to change the sign of $\lambda_U \Delta S_{ij} - \langle r_U, e_i - e_j\rangle$, giving $\mathrm{sign}(\Delta T_{ij}) = \mathrm{sign}(\lambda_U \Delta S_{ij})$.

Degeneracy and tie conventions. If the computed norm of the materialized projection satisfies $\|a\| \leq 10^{-300}$ or its computed square satisfies $\|a\|^2 \leq 0$, the code flags the state as numerically degenerate and assigns zero certificate coverage instead of evaluating λ_U; it does not test the Gram quadratic form directly. Neither check triggered in the ten fresh states (minimum observed $\|a\| = 2.66$). If two candidates satisfy the executed pair-degeneracy check $\|e_i - e_j\| \leq 10^{-12}$, the pair is flagged and counted as uncertified in the denominator (not triggered; the minimum observed pair distance is 2.1×10^{-3}). The certification predicate $M_{ij}^U := |\lambda_U \Delta S_{ij}| - \|r_U\| \|e_i - e_j\| > 0$ uses strict inequality with no added or subtracted margin; an exactly zero computed margin is treated as not certified. All projection, gain, and residual arithmetic is carried out in float64.

E.2 Empirical evaluation and provenance

Why the ambient bound was loose. An ambient-space version of Eq. 11, using the same Cauchy–Schwarz argument without restriction to $\mathcal{U}_C$, certified none of 36 candidate pairs across six archived

states. A geometry audit of those states found that $\mathcal{U}_C$ contained at most 0.26% of future-sensitivity energy in each state, while retaining 60–84% of local-gradient energy. The resulting slack was concentrated largely outside the candidate-difference span. This motivated restricting the comparison to $\mathcal{U}_C$: components orthogonal to this span cannot affect pairwise scores but can enlarge an ambient residual bound.

Evaluation on new states. We evaluate the quotient-space condition on five new language and five new vision states, disjoint from the six discovery states and from all other cohorts in the paper. Table 9 reports the per-state results. The condition certifies 10/30 language pairs across 3/5 states and 16/30 vision pairs across 4/5 states. The median residual tightening ratios are $174.6\times$ and $145.5\times$, respectively. The fitted λ_U is negative in 9/10 fresh states; the only positive state certifies no pairs.

Table 9: Fresh quotient confirmation, per-state primary endpoints (all values recomputed from the stored per-state fresh-quotient summary). "Cert." is certified pairs out of 6; $R_{\text{tighten}} = \|r_{\text{ambient}}\|/\|r_U\|$ uses the same λ_U in both residuals.

Family	Seed	λ_U	sign	Cert. /6	Preserv. / Revers.	R_{tighten}
Language	2142314115	-0.00715	neg	2	0 / 2	126.9
Language	1528934958	-0.00287	neg	0	0 / 0	174.6
Language	129105507	-0.00318	neg	3	0 / 3	252.3
Language	1833964399	-0.01430	neg	5	0 / 5	262.0
Language	609928333	-0.00465	neg	0	0 / 0	92.7
Vision	1004992386	-0.01092	neg	4	0 / 4	81.8
Vision	1359762397	-0.03616	neg	4	0 / 4	97.8
Vision	722306861	$+0.00135$	pos	0	0 / 0	145.5
Vision	1342673783	-0.00988	neg	4	0 / 4	345.0
Vision	318185148	-0.00360	neg	4	0 / 4	557.0
Language total (breadth 3/5 states)				10/30	0 / 10	median 174.6
Vision total (breadth 4/5 states)				16/30	0 / 16	median 145.5

Preservation and reversal among certified pairs. For a pair satisfying Eq. 11, the gain's sign determines whether the local ordering is preserved or reversed. Its sign alone does not certify a pair: the residual must also satisfy the pairwise margin condition. All seven fresh states with nonzero certificate coverage have $\lambda_U < 0$, so all $10 + 16 = 26$ certified pairs are reversal certificates and none are preservation certificates. Direct recomputation of ΔT_{ij} confirms the predicted sign for all 26/26 certified pairs.

What the residual control isolates. The reported tightening ratio compares $\|r_{\text{ambient}}\| = \|g_{\text{fut}} - \lambda_U g_{\text{loc}}\|$ with $\|r_U\| = \|P_{\mathcal{U}_C}(g_{\text{fut}} - \lambda_U g_{\text{loc}})\|$ using the same quotient-fitted λ_U in both quantities. Thus the comparison isolates the effect of restricting the residual geometry to $\mathcal{U}_C$, rather than introducing a second fitted gain.

This fresh same-λ_U control is distinct from the historical ambient-space bound. The historical audit fit $\lambda_{\text{state}} = \langle g_{\text{fut}}, g_{\text{loc}}\rangle / \|g_{\text{loc}}\|^2$ directly in the full action space on the six archived discovery states. The fresh control instead holds the quotient-fitted λ_U fixed on ten disjoint confirmation states. The two quantities therefore use the same least-squares form but are fit in different spaces and on different states.

Discovery–confirmation separation. The original audit included the three language and three vision states for which both g_{loc} and full-state g_{fut} were available from frozen computations. This eligibility rule was fixed before pairwise evaluation and did not use candidate outcomes; the other nine archived states lacked frozen g_{fut}.

The ambient-bound failure and geometry audit motivated Proposition 3 and Eq. 11. The proposition, condition, and confirmation thresholds were fixed and hashed before applying the new condition to the six already-seen discovery states and before generating any fresh confirmation state. The non-confirmatory discovery check met all preregistered gates: 11/18 language and 12/18 vision pairs certified, state breadth 2/3 and 3/3, and median tightening $336.8\times$ and $166.9\times$, respectively. None of these six states contributes to the fresh 10/30 or 16/30 confirmation counts.

Implementation and execution checks. All 16 identity, orthogonality, and sign checks per fresh state pass: 160/160 in total. These checks verify implementation consistency rather than constitute the scientific result. The frozen lock specified a 10^{-9} floor convention for candidate-pair degeneracy, whereas the executed discovery and fresh-confirmation code used $\|e_i - e_j\| \leq 10^{-12}$; no applicable pre-execution amendment was found. This is an execution-protocol deviation. All 96 archived pairs (36 discovery and 60 fresh) have distances above both cutoffs, so the deviation does not change their degeneracy classifications or the reported certificate counts. Several infrastructure interruptions required retries with the identical preregistered seeds on re-selected GPUs; no seed, candidate, threshold, λ_U definition, norm, or cohort size was changed after results became visible. Full incident-level records are retained separately.

F Supporting Representation Evidence

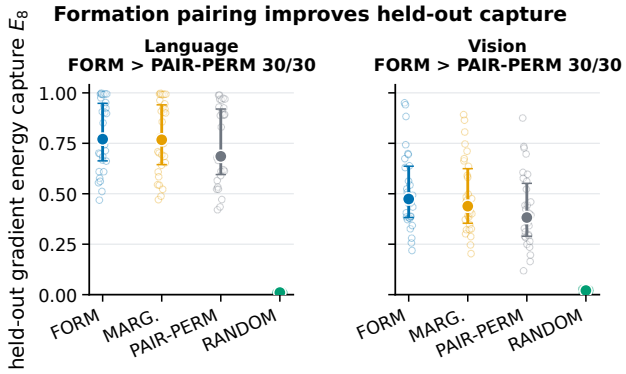
For a dense/token-linear operator $Y = XW^\top$, with $X \in \mathbb{R}^{n \times d_{\text{in}}}$ and $W \in \mathbb{R}^{d_{\text{out}} \times d_{\text{in}}}$, let $\Delta_Y := \nabla_Y L$ denote the upstream gradient (backpropagated output sensitivity) at the operator output. The corresponding weight gradient is

$$G := \nabla_W L = \Delta_Y^\top X. \tag{18}$$

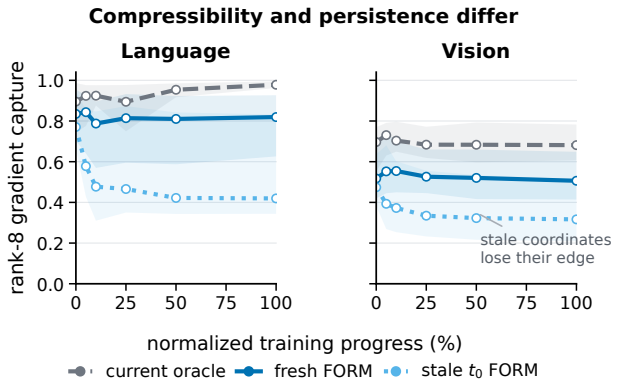
Matching $X^\top X$ and $\Delta_Y^\top \Delta_Y$ does not identify $\Delta_Y^\top X$: a sample permutation can preserve both marginal Gram matrices while changing the joint pairing and hence the formed gradient. To see this, let P be a row-permutation matrix applied to Δ_Y relative to fixed X, and set $\Delta'_Y = P\Delta_Y$. Since $P^\top P = I$,

$$(\Delta'_Y)^\top \Delta'_Y = \Delta_Y^\top P^\top P \Delta_Y = \Delta_Y^\top \Delta_Y,$$

while $X^\top X$ is unchanged, but in general $G' = (\Delta'_Y)^\top X = \Delta_Y^\top P^\top X$ need not equal $G = \Delta_Y^\top X$. Thus the two marginal Gram matrices do not determine the formed gradient. This motivates a pairing-broken control that changes the joint pairing while preserving both marginal Gram matrices. In matched real-model audits, coupling-aware FORM bases improve held-out gradient capture over pairing-broken controls across language and vision. Across training checkpoints, rank-8 compressibility can remain high while calibrated coordinates become stale and fresh bases repeatedly regain capture. These diagnostics characterize what the local fidelity surrogate captures and show that a fixed coordinate can be time-conditional; they are supporting representation evidence rather than additional headline decision transitions.



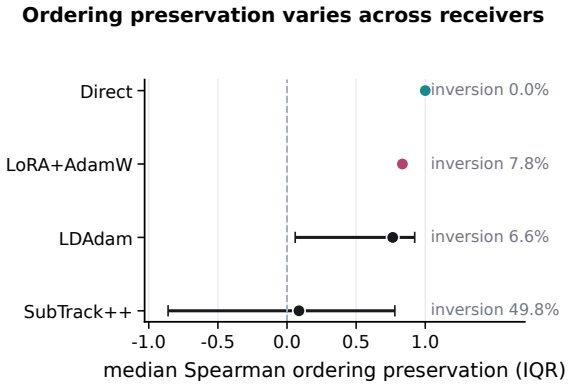
(a) Formation pairing.



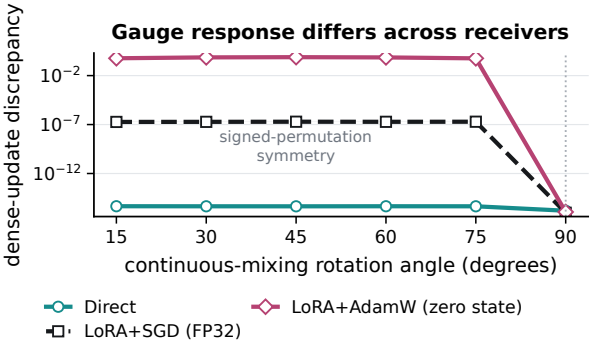
(b) Temporal persistence.

Figure 6: **Representation diagnostics for formation pairing and temporal persistence.** **(a)** Coupling-aware FORM bases capture more held-out gradient energy than the sample-pairing-broken PAIR-PERM control in all 30 target-matrix operators per family. PAIR-PERM row-permutes ΔY while holding X fixed; RANDOM is a random-orthogonal-subspace control. PAIR-PERM is distinct from the cross-target PERMUTED FORM candidate used in the decision experiments. Points are operators, with medians and IQRs summarized by the filled markers and bars. **(b)** Rank-8 gradient capture over training for the current oracle, freshly constructed FORM, and FORM calibrated once at t_0; the stale coordinates lose their initial advantage. These are operator-level diagnostics from one trained model per family rather than independent state-level replications.

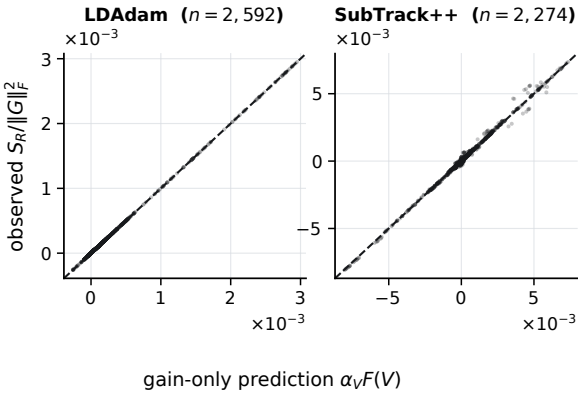
G Additional Receiver Results and Basis Conventions



(a) Receiver breadth.



(b) Gauge semantics.

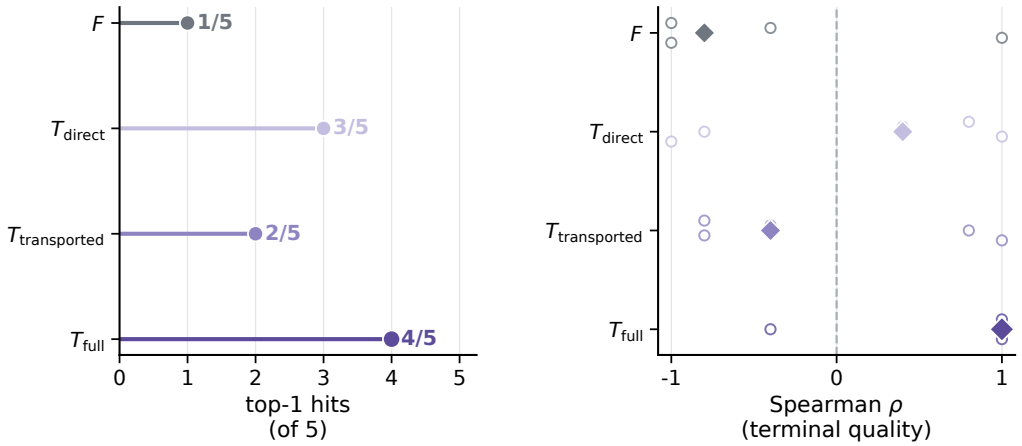


(c) Gain-only component vs. implemented-update score.

Figure 7: **Receiver-side diagnostics.** **(a)** Method-native audits show heterogeneous fidelity-to-receiver-score ordering preservation. Points show median Spearman correlation and bars show trial-level IQR; adjacent percentages give the median per-trial material-gap inversion rate. Direct and LoRA+AdamW use eight trials from the fixed Qwen3-8B dense-manifold audit, while LDAdam and SubTrack++ use 144 RoBERTa-base exact-PRE trials. **(b)** Dense-update discrepancy under continuous basis rotations for Direct, LoRA+SGD, and zero-state LoRA+AdamW, illustrating receiver-specific gauge behavior. The 90° point is a signed permutation for this rotation family. **(c)** The gain-only term $\alpha_V F(V)$ from Eq. 6 is compared with the observed normalized receiver score for LDAdam and SubTrack++, using 2592 and 2274 audited cells, respectively, across three seeds each.

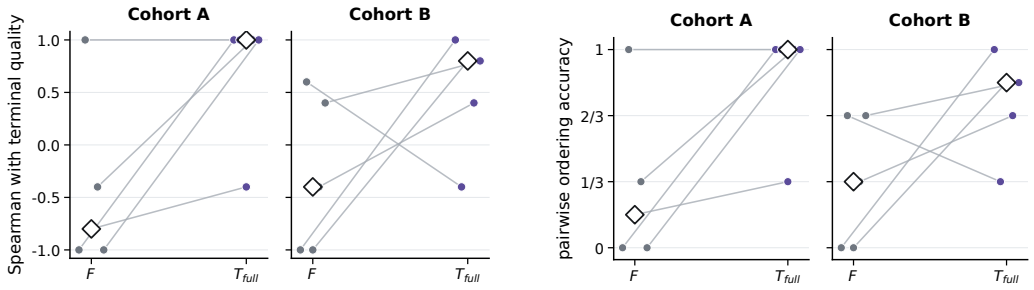
H PROSPECTIVE REPLICATION DIAGNOSTICS

These diagnostics concern the primary $p = 0.75$ experiments; the additional $p = 0.50$ results are reported separately in Table 18.



(a) Score-component selection. **(b)** Score-component association.

Figure 8: Component-level score ablation for the primary terminal-ranking results in Table 2, Language Cohort A ($n = 5$ states). T_{direct}/$T_{\text{transported}}$ are the persistent-direct/future-transported components of $T_{\text{full}} = T_{\text{direct}} + T_{\text{transported}}$. **(a)** Bars/points = top-1 hits out of 5. **(b)** Hollow points = each state's Spearman correlation with terminal quality (higher better); diamonds = cross-state median. F, T_{direct}, $T_{\text{transported}}$, T_{full} select the terminal winner in 1/5, 3/5, 2/5, 4/5 states; neither component alone ranks the winner first in all five states, and the complete score gives the strongest result.



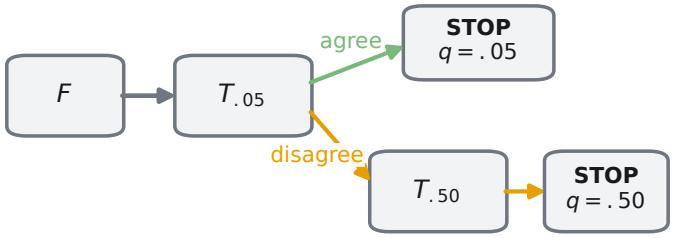
(a) Seed-wise association. **(b)** Pairwise accuracy.

Figure 9: Prospective language-cohort replication diagnostics, Language Cohort A and B ($n = 5$ states each). **(a)** Points = per-state Spearman correlation with terminal quality; lines join the same state's local-F and T_{full} values; diamonds = cohort median. **(b)** Points = per-state pairwise ordering accuracy; lines join the same state's local-F and T_{full} values; diamonds = cohort median. Across association and pairwise accuracy, T_{full} improves the local-fidelity ordering in both cohorts, retaining the preregistered cohort separation.

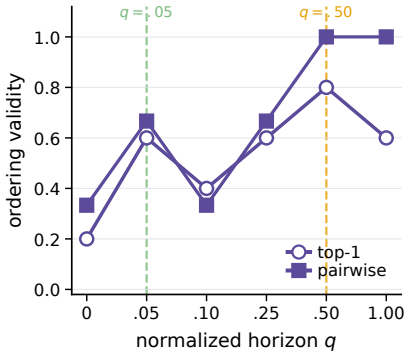
I COMPUTATIONAL COST OF FUTURE-CONDITIONED EVIDENCE

The quotient result in Section 4.2 concerns validity, while T_{full} and its quotient restriction require future-trajectory computation. We therefore study how performance changes with the amount of future computation available. We fix a disagreement rule on a separate fresh cohort to ask whether evidence

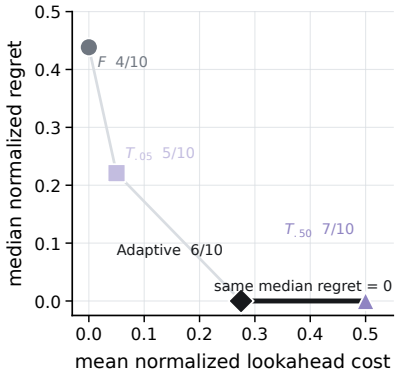
value is horizon-dependent: compare the local-fidelity winner to the winner from short-horizon evidence $T_{.05}$, and escalate to $T_{.50}$ only on disagreement. A preregistered horizon audit on a separate fresh five-state RoBERTa/SST-2 cohort shows validity is horizon-dependent and non-monotone – top-1 rates across evidence fractions $q \in \{0, .05, .10, .25, .50, 1.00\}$ are $1/5, 3/5, 2/5, 3/5, 4/5, 3/5$, so no tested short horizon is uniformly reliable (Figure 10b). On a further separate fresh ten-state cohort, the disagreement-triggered rule matches always-$T_{.50}$'s median regret (both 0) while using mean normalized evidence cost 0.275 versus 0.50 and reducing measured decision GPU time by 43.81%, though its mean regret and top-1 ($6/10$ versus $7/10$) are worse – a tradeoff between computational cost and prediction quality (Figure 10a,c). An exact cost-matched allocation control (Appendix Fig. 11) places the rule near the median of all 252 equal-cost allocations. These results characterize one fixed cost–regret tradeoff for future-conditioned evidence; optimizing the allocation policy remains open.



(a) Frozen disagreement-triggered evidence acquisition.

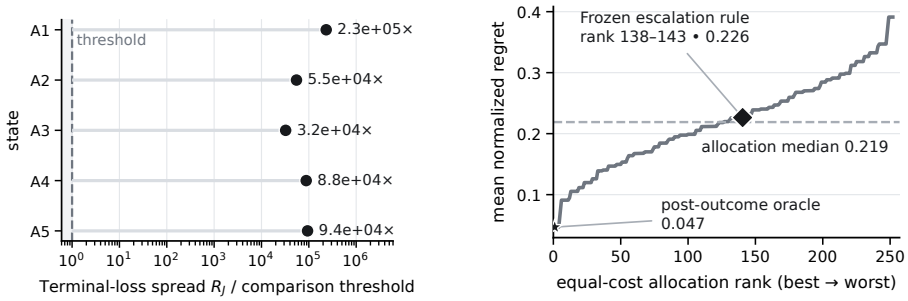


(b) Horizon-dependent validity.



(c) Cost–regret comparison.

Figure 10: Adaptive evidence acquisition trades computation against selection quality. **(a)** The preregistered disagreement rule compares the local-fidelity winner with the winner selected by $T_{.05}$ and escalates to $T_{.50}$ only on disagreement. **(b)** A separate five-state cohort is evaluated at $q \in \{0, .05, .10, .25, .50, 1.00\}$. Markers show evaluated settings; connecting lines are visual guides. "Pairwise" denotes median per-state pairwise accuracy. Agreement is non-monotone, and no short horizon selects the terminal winner in all five states. **(c)** Points represent separately evaluated policies on a ten-state cohort. Statewise normalized regret is the selected terminal loss minus the minimum candidate loss, divided by the terminal-loss spread across candidates; the plot reports the median across states. Compared with always-$T_{.50}$, the adaptive rule reduces mean normalized lookahead cost from 0.50 to 0.275 and decision GPU time by 43.81%, while retaining zero median regret. Its mean regret is higher and its top-1 count lower ($6/10$ versus $7/10$).



(a) Numerical threshold.

(b) Exact cost-matched allocation control.

Figure 11: Terminal-loss spread and cost-matched evidence allocation: Language Cohort A (seeds 20260880–20260884) for **(a)**, and the separate ten-state evidence-acquisition cohort for **(b)**. **(a)** R_J is the terminal-loss spread across candidates. Relative to the fixed 10^{-8} comparison threshold, $R_J/10^{-8}$ ranges from approximately 3.2×10^4 to 2.3×10^5. **(b)** The preregistered disagreement rule achieves mean normalized regret near the median of the 252 equal-cost allocations. Panel **(b)** uses the statewise normalized regret defined in Figure 10c, aggregated by the mean rather than the median across the ten states.

J Per-State Prospective Results

In Language Cohort A at $p = 0.75$, seed 20260882 is a top-1 miss for T_{full} (Table 10, row 3). Candidate terminal-loss spreads range from approximately 0.26% to 1.7% of terminal loss within this cohort, and from 0.05% to 2.7% across all 15 primary states.

Table 10: Language Cohort A at $p = 0.75$. Values are copied from stored prospective ranking tables; no value is reconstructed from a figure. ρ denotes within-state Spearman correlation with terminal quality, R denotes raw terminal-loss selector regret, and Pair denotes pairwise ordering accuracy; subscripts identify the scoring rule.

Seed	F sel.	T_{full} sel.	Terminal	ρ_F	ρ_T	R_F	R_T	Pair$_F$	Pair$_T$
20260880	FORM	CURRENT	CURRENT	-1.0	1.0	2.327e-3	0	0	1
20260881	FORM	FORM	FORM	1.0	1.0	0	0	1	1
20260882	FORM	FORM	CURRENT	-0.8	-0.4	3.108e-4	3.108e-4	0.167	0.333
20260883	FORM	CURRENT	CURRENT	-0.4	1.0	6.996e-4	0	0.333	1
20260884	FORM	CURRENT	CURRENT	-1.0	1.0	9.397e-4	0	0	1

Table 11: Independent Language Cohort B aggregate diagnostics and independent DeiT-S/CIFAR-100 per-state results. ρ denotes within-state Spearman correlation with terminal quality, R denotes raw terminal-loss selector regret, and Pair denotes pairwise ordering accuracy; subscripts identify the scoring rule. Top-1 hit entries are binary indicators for (F, T_{full}).

Language-B metric	Local F	T_{full}
Top-1 hits (of 5)	0	4
Median Spearman	-0.4	0.8
Median pairwise accuracy	0.333	0.833
Median regret	1.156e-3	0
Regret no-worse (of 5)	–	4
Regret strict improvement (of 5)	–	4

Vision seed	Terminal	Top-1 hit (F, T_{full})	ρ_F	ρ_T	R_F	R_T	Pair$_F$	Pair$_T$
20260909	PERMUTED FORM	$(0, 1)$	-1.0	1.0	3.054e-4	0	0	1
20260910	CURRENT	$(0, 1)$	-0.8	1.0	6.829e-4	0	0.167	1
20260911	CURRENT	$(0, 1)$	-0.8	1.0	3.566e-4	0	0.167	1
20260912	CURRENT	$(0, 1)$	-0.8	1.0	5.863e-4	0	0.167	1
20260913	PERMUTED FORM	$(0, 1)$	-1.0	1.0	4.550e-4	0	0	1

Table 12: Decision-challenge, horizon-audit, and evidence-acquisition policy summaries from frozen result tables. The decision-challenge block reuses Language A/B states; the horizon-audit and evidence-acquisition blocks are separate fresh cohorts. For the decision-challenge block, D/S/R denote Dominant-SVD/SARA-IS/Random-Haar. Regret entries are in raw terminal-loss units; Figures 10c and 11b instead report statewise spread-normalized regret.

Decision-challenge block	F top-1	T_{full} top-1	Terminal winner D/S/R
Cohort A (5 states)	1/5	4/5	1/0/4
Cohort B (5 states)	2/5	5/5	2/0/3
Pooled (10 states)	3/10	9/10	3/0/7

Horizon q	Top-1 (of 5)	Median pairwise	Median regret
0	1	0.333	1.862e-4
.05	3	0.667	0
.10	2	0.333	9.635e-6
.25	3	0.667	0
.50	4	1.000	0
1.00	3	1.000	0

Acquisition policy	Top-1 /10	Mean regret	Median regret	Norm. cost
F	4	1.824e-4	9.652e-5	~ 0
$T_{.05}$	5	1.132e-4	5.230e-5	.05
Adaptive	6	6.685e-5	0	.275
$T_{.50}$	7	4.217e-5	0	.50
$T_{1.00}$	7	4.217e-5	0	1.00

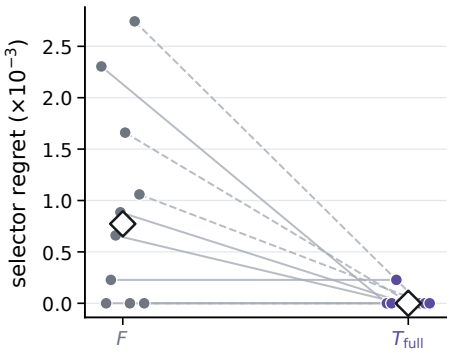


Figure 12: Selector regret for the low-rank selection-rule comparison (Figure 1), same ten reused Language Cohort A+B states, Dominant-SVD/SARA-IS/Random-Haar pool. Selector regret = selected candidate's terminal loss minus the pool minimum (lower is better; not a no-intervention comparison); points = per-state values, lines join the same state's F/T_{full} regret (solid: A, dashed: B), diamond = cross-state median. T_{full} reduces selector regret relative to current fidelity across the pooled ten states; not every state strictly improves.

K Scope of the Claims

We distinguish gradient capture, implemented local first-order value, first-order future value, finite-step utility, terminal candidate quality, and performance relative to no intervention.

Receiver-local validity. Direct preserves fidelity ordering exactly, while Proposition 1 gives a sufficient, not necessary, condition for receiver-order preservation. Neither result guarantees finite-step utility or terminal performance. The zero-moment matched receiver analysis localizes the observed Qwen inversion to the tested adaptive coordinate transformation, without implying that AdamW generally reorders fidelity. The controlled $K = 16$ experiment isolates the coordinatewise second-moment denominator at fixed candidate-specific first moments; it does not characterize optimizer memory acquired during ordinary training.

Future valuation and selection. The prospective experiments evaluate the complete score $T_{\text{full}} = T_{\text{direct}} + T_{\text{transported}}$ against separately executed terminal outcomes. Objective-matched and simple-rule controls assess what the future information adds beyond the terminal objective and fixed selection rules. T_{full} is neither Bellman-optimal Q^* nor an online-available optimal selector, and the reported ranking improvements do not establish that its computational cost is justified (Appendices D, I, and N).

First-order geometry and finite outcomes. Proposition 3 shows that, for a fixed finite candidate set, first-order pairwise score differences depend only on future sensitivity restricted to the candidate-difference span. The associated signed condition is sufficient rather than necessary and does not certify finite terminal outcomes or coverage for untested candidate families. The separate finite-outcome evaluation found that all 26 pre-fixed certified reversal pairs retained their predicted T_{full} ordering at terminal Q, as did 30 of the 34 uncertified pairs; this empirical completion does not turn the first-order condition into a terminal-Q certificate. The six discovery states in Appendix E.2 do not contribute to the fresh 10/30 or 16/30 confirmation counts.

Evaluation units and supporting evidence. Language Cohort A is kept separate from retrospective correction runs; the $p = 0.50$ extension reuses its five seed identities rather than adding an independent cohort. The vision cohort is independent, whereas the candidate-rule comparison reuses Language A/B and therefore adds rule breadth rather than new states. Formation and persistence diagnostics support the interpretation of gradient representation but are not additional headline decision transitions.

L Receiver Ordering Guarantee and Direct Positive Control

The main text gives the pairwise interpretation; here we record the explicit bound.

Decomposition behind Eq. 6. For $H_V = \Pi_V(G) \neq 0$, α_V is the orthogonal-projection coefficient of $-\Delta W_R(V)$ onto H_V, making E_V orthogonal to H_V. Hence $\langle G, E_V\rangle = \langle G - H_V, E_V\rangle$, which yields Eq. 6. If $H_V = 0$, α_V is not unique; we use the convention $\alpha_V = 0$ and $E_V = -\Delta W_R(V)$.

Proof of Proposition 1. Write $S_i = S_R(V_i)$ and $S_j = S_R(V_j)$. For any fixed $\eta > 0$, Eq. 6 gives

$$\begin{aligned}\frac{S_i - S_j}{\|G\|_F^2} &= \eta(F_i - F_j) + (\alpha_i - \eta)F_i - (\alpha_j - \eta)F_j + \beta_i - \beta_j \\ &\geq \eta(F_i - F_j) - |\alpha_i - \eta|F_i - |\alpha_j - \eta|F_j - |\beta_i| - |\beta_j|.\end{aligned}$$

Thus Eq. 7 makes the right-hand side strictly positive and implies $S_i > S_j$. The bound is sufficient, not necessary.

Direct corollary. For Direct projected descent, choose η in Proposition 1 to be the common positive gain multiplying the projected current gradient. For each candidate with $H_V \neq 0$, $\alpha_V = \eta$; for $H_V = 0$ the stated convention sets $\alpha_V = 0$ but $F(V) = 0$. In either case $|\alpha_V - \eta|F(V) = 0$ and $\beta_V = 0$. Hence $D_{ij}(\eta) = 0$, including a zero-projection candidate, and, restating the identity in Eq. 14,

$$S_R(V) \propto \|\Pi_V(G)\|_F^2 \propto F(V).$$

The fidelity ordering is therefore exactly the receiver-local action ordering. This control makes the terminal experiment stricter than an unconstrained "local score versus future outcome" correlation: if terminal ordering changes, the failure occurs after the receiver-local implication has been deliberately satisfied.

M Non-identifiability of Terminal Ordering from Receiver-Local Action

This appendix constructs two admissible future training systems that share all specified local information but induce opposite terminal rankings, proving Proposition 2.

What is held fixed. Across the two future systems constructed below, the switch state z_s, the current gradient G, the Direct receiver R_{Direct}, the candidate subspaces V_i, V_j, the candidate interventions e_i, e_j, and the terminal objective Q are all identical. Only the post-switch future training dynamics differ.

Construction. Work in vectorized candidate-action coordinates $x = (x_1, x_2) \in \mathbb{R}^2$ with switch-time reference $x_s = 0$, and let $\varepsilon > 0$ be arbitrary. Take the smooth, strongly convex current loss

$$\ell_\varepsilon(x) = \varepsilon(2x_1 + x_2) + \tfrac{1}{2}\|x\|^2,$$

so the formed current gradient at the switch is $G := \nabla\ell_\varepsilon(0) = \varepsilon(2, 1)$. Take two equal-budget rank-1 candidates $V_i = \mathrm{span}\{(1, 0)\}$ and $V_j = \mathrm{span}\{(0, 1)\}$, with $\Pi_{V_i}(G) = P_i G$, $\Pi_{V_j}(G) = P_j G$ for the orthogonal projectors $P_i = \mathrm{diag}(1, 0)$, $P_j = \mathrm{diag}(0, 1)$ – the rank-1 specialization of the $P_V = VV^\top$ convention of Appendix C. Under Direct projected descent with common gain $\eta = 1$, the applied switch interventions are

$$e_i = -\Pi_{V_i}(G) = (-2\varepsilon, 0), \qquad e_j = -\Pi_{V_j}(G) = (0, -\varepsilon).$$

Local ordering. From Eq. 2 and the Direct identity $S_{\mathrm{Direct}}(V) = \eta\|\Pi_V(G)\|_F^2$ (Appendix C), with $\eta = 1$,

$$S_{\mathrm{Direct}}(V_i) = 4\varepsilon^2 \;>\; \varepsilon^2 = S_{\mathrm{Direct}}(V_j) \qquad \text{for every } \varepsilon > 0 :$$

candidate i is strictly, ε-independently preferred by the local action score.

Terminal objective. Fix the same smooth terminal objective in both future systems, $J(x) = \exp(x_1 + x_2)$, $Q(x) := -J(x)$, matching the terminal-quality convention $Q = -J$ of Section 2.

Future system A. One gradient-descent step of size 1 on the smooth convex future loss $f_A(x) = 0$ leaves the state unchanged: $\Phi_A(x) = x$. Hence

$$Q_A(V_i) = -e^{-2\varepsilon} \;>\; -e^{-\varepsilon} = Q_A(V_j) \qquad \text{for every } \varepsilon > 0,$$

so under Future A the terminal ordering agrees with the local ordering.

Future system B. One gradient-descent step of size 1 on the smooth convex future loss $f_B(x) = \frac{9}{20}x_1^2$ – whose Hessian $\mathrm{diag}(9/10, 0)$ is positive semidefinite, so f_B is convex – gives $\Phi_B(x) = x - \nabla f_B(x) = (x_1/10,\, x_2)$. Applying Φ_B to the same two interventions,

$$\Phi_B(e_i) = (-\varepsilon/5, 0), \qquad \Phi_B(e_j) = (0, -\varepsilon),$$

so

$$Q_B(V_i) = -e^{-\varepsilon/5} \;<\; -e^{-\varepsilon} = Q_B(V_j) \qquad \text{for every } \varepsilon > 0,$$

since $e^{-\varepsilon/5} > e^{-\varepsilon}$: under Future B the terminal ordering is reversed.

Same reference trajectory. Without intervention, $x = 0$ is a fixed point of both Φ_A and Φ_B: the no-intervention reference trajectory is identical across the two future systems, so the ambiguity below is not an artifact of comparing different observed states.

Non-implication. Futures A and B share the switch state, current gradient, Direct receiver, candidate subspaces, interventions, local scores, and terminal objective, and both are ordinary one-step gradient descent on smooth convex losses realizing $\Phi = x - \eta_f \nabla f$ rather than an arbitrary hand-designed map. Yet Future A gives $Q(V_i) > Q(V_j)$ and Future B gives $Q(V_i) < Q(V_j)$. Hence

$$\boxed{S_{\mathrm{Direct}}(V_i) > S_{\mathrm{Direct}}(V_j) \quad \not\Longrightarrow \quad Q(V_i) > Q(V_j)}$$

without additional restrictions on the future training dynamics.

Arbitrarily small interventions. Every inequality above is exact for every $\varepsilon > 0$, not merely asymptotic: $4\varepsilon^2 > \varepsilon^2$, $e^{-2\varepsilon} < e^{-\varepsilon}$, and $e^{-\varepsilon/5} > e^{-\varepsilon}$ all hold strictly for all $\varepsilon > 0$, so the construction scales with $\|e_i\|, \|e_j\| \to 0$ while preserving the same local ordering and the opposite terminal orderings. The failure is therefore not caused only by a large finite intervention leaving the local linear regime: local-action information remains insufficient arbitrarily close to the reference state whenever the future transport is left unconstrained.

Relation to T_{full}. The receiver-local score sees only the switch-time action e_V; Futures A and B apply exactly the same action and differ only in how it is transported afterward. At the shared reference terminal state $x_T = 0$, $\nabla J(0) = (1, 1)$, and with $D\Phi$ the Jacobian of the future map, $T_{\mathrm{full}}(V) = -\langle \nabla J(0), D\Phi\, e_V \rangle$ (Eq. 4). For Future A, $D\Phi_A = I$, giving $T_{\mathrm{full}}(V_i) = 2\varepsilon > \varepsilon = T_{\mathrm{full}}(V_j)$; for Future B, $D\Phi_B = \mathrm{diag}(1/10, 1)$, giving $T_{\mathrm{full}}(V_i) = \varepsilon/5 < \varepsilon = T_{\mathrm{full}}(V_j)$. The future-conditioned score therefore tracks exactly the piece of information a receiver-local score omits: the transport Jacobian of the remaining training dynamics. This does not make T_{full} a finite-Q certificate. Whether its first-order ordering survives finite outcomes is a separate empirical question addressed in Section 4 and Appendix N; T_{full} remains a reference for terminal-ranking prediction for the transition this appendix shows is otherwise unidentified.

N Additional Decision-Validity Controls

N.1 Objective-matched switch-time control

Section 4.1 asks whether T_{full}'s recovery reflects future-conditioned information or can instead be explained by the mismatch between the switch-time local objective underlying F/S_R and the held-out objective defining terminal Q. To separate these two explanations, we retrospectively evaluate an

objective-matched switch-time score that uses the same held-out TRAIN objective as terminal Q, but is evaluated at the switch state with no differentiation through the future trajectory:

$$S_{\mathrm{holdout},s}(V) = -\langle \nabla J(z_s), e_V \rangle ,$$

where J is the identical frozen held-out TRAIN evaluator that already produces every stored terminal objective value, and e_V is the same frozen candidate intervention direction used elsewhere in the paper. Because $S_{\mathrm{holdout},s}$ uses the terminal-matched objective at the switch but omits future differentiation, it isolates objective matching from future-conditioned revaluation. A persistent gap to T_{full} therefore shows that objective matching alone does not account for the observed ordering recovery. This control is retrospective: it is computed after terminal outcomes are already on file, and it is not used to select or predict any new outcome. An independent finite-difference check validated the implementation before any full-cohort computation, to relative error below 10^{-8} at fp64 (achieved errors 3.9×10^{-10}–7.6×10^{-10} against a pre-specified $\leq 10^{-6}$ implementation tolerance) across all three implementation paths.

Table 13: Objective-matched switch-time control versus local fidelity and T_{full}. Top-1 is hits out of the family's state count; Pairwise and Spearman are medians; Regret is the median selection regret. Decision-challenge A+B reuses the Language A+B states with the Dominant-SVD/SARA-IS/Random-Haar candidate family (Section 4).

Family	Metric	F	$S_{\mathrm{holdout},s}$	T_{full}
Language A ($n = 5$)	Top-1	1/5	2/5	4/5
	Pairwise	0.167	0.833	1.000
	Spearman	-0.80	$+0.80$	$+1.00$
	Regret	6.996e-4	3.108e-4	0
Language B ($n = 5$)	Top-1	0/5	0/5	4/5
	Pairwise	0.333	0.333	0.833
	Spearman	-0.40	-0.60	$+0.80$
	Regret	1.156e-3	1.156e-3	0
Vision ($n = 5$)	Top-1	0/5	2/5	5/5
	Pairwise	0.167	0.167	1.000
	Spearman	-0.80	-0.80	$+1.00$
	Regret	4.550e-4	3.054e-4	0
Decision-challenge A+B ($n = 10$)	Top-1	3/10	3/10	9/10
	Pairwise	0.333	0.333	1.000
	Spearman	-0.50	-0.50	$+1.00$
	Regret	7.729e-4	4.440e-4	0

Comparing T_{full} with $S_{\mathrm{holdout},s}$ by per-state selection regret, T_{full} has lower regret in nine of the 15 primary states, equal regret in five, and higher regret in one. The exception is Language B seed 20260886, with regrets 3.2747×10^{-4} for T_{full} and 2.3556×10^{-4} for $S_{\mathrm{holdout},s}$. Across the ten decision-challenge states, the corresponding counts are six lower, four equal, and zero higher. The objective-matched switch-time control does not reproduce the cohort-level top-1 gains of T_{full}.

Cheap constant-selector context. As a further descriptive control, fixed candidate-agnostic selectors (always-CURRENT, always-FORM, always-MARGINAL, always-PERMUTED_FORM, and an analytic uniform-random selector) were also evaluated; because these selectors do not rank candidates, only top-1 rate and regret are defined for them, not pairwise accuracy or Spearman correlation. In all three primary cohorts, always-CURRENT attains a higher top-1 rate (0.6–0.8) and lower median regret than either F or $S_{\mathrm{holdout},s}$ achieve as ranking rules; the best constant selector in each primary cohort (always-CURRENT in all three) is identified post hoc and is reported as a descriptive fact about this candidate pool at this switch point, not as a rule that was prospectively selected or as evidence that CURRENT is intrinsically superior in general. T_{full} matches or exceeds every constant selector's top-1 rate and has no worse median regret in each primary cohort.

N.2 Terminal Evaluation of the Certified Pairs

Section 4.2 certified 26 of the 60 fresh pairs (10/30 language, 16/30 vision) as first-order $S_R \to T_{\text{full}}$ reversal certificates. We complete this analysis with an empirical, finite-outcome test: for the same 10 fresh quotient-confirmation states and the same 4 already-frozen candidates per state, we execute the canonical Direct continuation at unit intervention amplitude ($\tau = 1$ in Section 2, with the switch-time learning rate already included in e_V) to the state's own terminal horizon and evaluate the identical frozen held-out TRAIN objective used throughout the paper – 40 candidate continuations in total, with no new parent-state training. These comparisons use the same frozen TRAIN-subset evaluator as the first-order scores. The 26 certified pair identities and their first-order predictions were fixed before any of these candidate-specific terminal outcomes was computed.

Table 14: Finite-outcome completion of the fresh quotient certificates. "T→Q preserved" counts pairs where the T_{full}-predicted sign matches the sign of ΔQ_{ij}; "S→Q reversal" counts pairs that remain a reversal of the local S_R preference at Q; "$|R| < |\Delta T|$" uses $R_{ij} = \Delta Q_{ij} - \Delta T_{ij}$ at unit intervention amplitude $\tau = 1$.

Population	Pairs	T→Q preserved	S→Q reversal	$\lvert R\rvert < \lvert\Delta T\rvert$	median $\lvert R\rvert/\lvert\Delta T\rvert$
Certified language	10	10/10	10/10	10/10	0.230
Certified vision	16	16/16	16/16	12/16	0.111
Certified pooled	26	26/26	26/26	22/26	0.143
Uncertified pooled	34	30/34	22/34	27/34	0.178

At the state level, T_{full} gives top-1 hits in $9/10$ states pooled (Language $5/5$, Vision $4/5$) versus local S_R's $0/10$ (Language $0/5$, Vision $0/5$). Uncertified means only that Eq. 11 did not certify the pair's $S_R \to T_{\text{full}}$ sign; it is not a prediction that finite-Q ordering should fail, and the 30/34 preservation rate above is reported as a secondary descriptive control, not as evidence that the certificate discriminates finite-Q outcomes.

Comparison with reversed fidelity. Because F and S_R have identical strict orderings in this Direct-receiver evaluation, reversing the local ordering is equivalent to ranking candidates by $-F$. Across all 60 pairs, T_{full} and $-F$ agree with the terminal ordering on 56 and 48 pairs, respectively. Both are correct on 45 pairs; only T_{full} is correct on 11, only $-F$ on 3, and neither on 1.

All 26 certified pairs are first-order reversals, so the certificate implies agreement between T_{full} and $-F$ on their first-order ordering. Their agreement with the finite terminal ordering on all 26 pairs is a separate empirical finding. Among the 34 uncertified pairs, both scores are correct on 19 pairs, only T_{full} on 11, only $-F$ on 3, and neither on 1. Thus, all differences in terminal pairwise accuracy between these scores occur outside the certified set in this evaluation. The 60 pairs are nested within ten states, not independent samples.

Matched no-intervention context. For the same 10 states and 40 candidate continuations, we evaluate a no-intervention reference with identical switch state, data order, randomness, optimizer continuation, horizon, and terminal evaluator. CURRENT is numerically close to this reference: its terminal loss is lower in 8/10 states, with an absolute difference of at most 3.8×10^{-5}. FORM, MARGINAL, and PERMUTED FORM improve on the reference in 0/10, 1/10, and 3/10 states, respectively. Overall, 12/40 candidate continuations improve on the reference; the absolute relative deviation $|Q - Q_{\text{ref}}|/|Q_{\text{ref}}|$ has mean 0.11% and maximum 0.50%. These small deviations contextualize the ranking results; they do not establish a performance advantage over no intervention.

N.3 Retrospective Reversed-Fidelity and Fixed-Candidate Baselines

We compare reversed fidelity, fixed-candidate rules, and T_{full} using the archived candidate outcomes. Reversed fidelity selects $\arg\min_V F(V)$ and needs no terminal information at selection time. A fixed-candidate rule selects the same named candidate construction in every state of an evaluated setting; the resulting subspace may still depend on the state. The best fixed rule is identified separately within each setting after observing its outcomes. The rules maximizing top-1 accuracy and minimizing mean terminal-loss regret coincide in all seven settings, with a unique optimum under each criterion

and no tie-breaking. These baseline analyses are retrospective; the original T_{full} predictions were fixed before the corresponding candidate-specific terminal outcomes were evaluated.

Table 15: Terminal-winner selection across evaluated settings. Top-1 denominators count evaluated states within each setting. Best fixed is identified post hoc from terminal outcomes, under either criterion described above. Decision challenge reuses the Language A+B checkpoints with a different candidate pool; $p = 0.50$ reuses Language A seed identities at an earlier switch. These rows do not represent independent replications.

Setting	$-F$ top-1	Best fixed top-1	T_{full} top-1	Best fixed rule
Language A	$4/5$	$4/5$	$4/5$	CURRENT
Language B	$3/5$	$3/5$	$4/5$	CURRENT
Vision	$2/5$	$3/5$	$5/5$	CURRENT
Decision challenge	$7/10$	$7/10$	$9/10$	Random-Haar
Quotient language	$4/5$	$4/5$	$5/5$	CURRENT
Quotient vision	$1/5$	$4/5$	$4/5$	CURRENT
Language A, $p = 0.50$	$3/5$	$3/5$	$3/5$	CURRENT

Observed stability of fidelity rankings. The complete F ordering is unchanged across the evaluated states in six of the seven settings. Quotient vision has two distinct orderings: FORM and MARGINAL exchange the highest-fidelity position, while PERMUTED FORM remains the lowest-fidelity candidate. Consequently, reversed fidelity makes the same candidate choice within every evaluated setting: CURRENT in the language settings, PERMUTED FORM in the vision settings, and Random-Haar in the decision challenge. These are observations about the sampled candidate pools, not a claim that fidelity values, score gaps, or the candidate subspaces are independent of the state.

Selection gains and remaining errors. T_{full} matches or exceeds the best fixed rule's aggregate top-1 accuracy in all seven settings, with a strict improvement in four. This does not imply statewise dominance. For example, in quotient vision, T_{full} and always-CURRENT both identify the winner in $4/5$ states, but always-CURRENT has lower regret in one state. Relative to reversed fidelity, T_{full} has lower, equal, and higher regret in 15, 22, and 3 of the 40 state–candidate-pool evaluations, respectively. These counts are descriptive and are not treated as 40 independent observations.

Table 16: Paired statewise regret comparisons against the post hoc best fixed rule. W/T/L means T_{full} has lower/equal/higher raw regret, respectively, using exact equality; no cross-setting independence is assumed.

Setting	W	T	L
Language A	1	3	1
Language B	1	3	1
Vision	2	3	0
Decision challenge	3	6	1
Quotient language	1	4	0
Quotient vision	1	3	1
Language A, $p = 0.50$	2	3	0

Regret definition and aggregation. For each state, let $J(V) = -Q(V)$ be its archived terminal objective loss, with lower J (higher Q) preferred. Raw terminal-loss regret is $R(r) = J(V_r) - \min_{V \in \mathcal{C}} J(V)$ for rule r and that state's evaluated pool $\mathcal{C}$. Table 17 reports the arithmetic mean and median of these statewise regrets separately within each setting, in raw terminal-loss units, without normalization or cross-task averaging. This differs from the dimensionless spread-normalized regret in Figure 10, which divides by the state's terminal-loss spread. The pool optimum here is only the best of the three or four executed candidates; it is not a global optimum and does not bound the gains available from untested alternatives.

The archived analysis uses exact floating-point equality for score, terminal, and regret ties. No score-selection or terminal-optimum ties occur in these evaluations; regret ties are retained. Table 16 compares regrets within each state before counting lower/equal/higher outcomes. It does not subtract medians or aggregate paired differences. Repeated values when rules make identical selections are the same outcomes, not independent evidence.

Table 17: Raw terminal-loss regret $R(r) = J(V_r) - \min_{V \in \mathcal{C}} J(V)$. Mean and median are over states within each evaluated setting, with the denominators in Table 15; units are those of the archived terminal objective loss, not percentages or normalized regret. Best fixed names are in that table. Scientific notation retains small nonzero regrets.

Setting	Rule	Mean regret	Median regret
Language A	$-F$	1.09704×10^{-4}	0
	Best fixed	1.09704×10^{-4}	0
	T_{full}	6.21523×10^{-5}	0
Language B	$-F$	8.32259×10^{-5}	0
	Best fixed	8.32259×10^{-5}	0
	T_{full}	6.54938×10^{-5}	0
Vision	$-F$	1.84818×10^{-5}	8.68060×10^{-6}
	Best fixed	3.75184×10^{-6}	0
	T_{full}	0	0
Decision challenge	$-F$	1.00441×10^{-4}	0
	Best fixed	1.00441×10^{-4}	0
	T_{full}	2.27716×10^{-5}	0
Quotient language	$-F$	2.64476×10^{-5}	0
	Best fixed	2.64476×10^{-5}	0
	T_{full}	0	0
Quotient vision	$-F$	1.49998×10^{-5}	9.40670×10^{-6}
	Best fixed	3.11678×10^{-6}	0
	T_{full}	1.65809×10^{-6}	0
Language A, $p = 0.50$	$-F$	3.59904×10^{-4}	0
	Best fixed	3.59904×10^{-4}	0
	T_{full}	1.15984×10^{-6}	0

Cases where reversed fidelity has lower regret. The following list reports all three cases, with selected candidates and terminal losses. Here $\Delta R = R(-F) - R(T_{\text{full}}) = J(V_{-F}) - J(V_{T_{\text{full}}})$ is a within-state difference in raw terminal-loss units; a negative value favors reversed fidelity.

- Language A, seed 20260882: $-F$ selects CURRENT ($J = 0.1220906662$); T_{full} selects FORM ($J = 0.1224014278$); $\Delta R = -3.10762 \times 10^{-4}$.
- Language B, seed 20260886: $-F$ selects CURRENT ($J = 0.1097420955$); T_{full} selects PERMUTED FORM ($J = 0.1098340069$); $\Delta R = -9.19114 \times 10^{-5}$.
- Decision challenge, seed 20260882: $-F$ selects Random-Haar ($J = 0.1221056666$); T_{full} selects Dominant-SVD ($J = 0.1223333824$); $\Delta R = -2.27716 \times 10^{-4}$.

The two occurrences of seed 20260882 use the same checkpoint with different candidate pools; they are not independent checkpoint replications. In Language B, neither selection is terminal-best (MARGINAL is). These cases identify selection errors, without attributing them to a checkpoint-level cause or a nonlinear remainder.

In quotient vision, the loss against always-CURRENT occurs at seed 722306861: T_{full} selects MARGINAL ($J = 0.5980151133$), while CURRENT has $J = 0.5980068228$ and zero regret. The excess raw regret of T_{full} is 8.29046×10^{-6}.

O EARLIER-SWITCH ROBUSTNESS TEST

Question and evaluation units. The additional test asks whether the qualitative local-versus-future ranking pattern observed at $p = 0.75$ persists at $p = 0.50$ under the frozen experimental system. It uses seed identities 20260880–20260884, already used by Language A, with four candidate continuations per earlier checkpoint. Its primary measure is within-state pairwise ranking accuracy, with terminal top-1 as a secondary measure. The five earlier checkpoints are not counted as additional independent seed replications. Cross-switch differences in state, gradient, candidate geometry, and remaining training preclude attributing the results to horizon length alone.

Prediction timing and completion. The execution summary confirms that predictions were frozen with 124 input hashes before any candidate-specific terminal branches were present. All 20 candidate branches and three no-intervention repetitions completed, with no final branch failures. All five states have six resolved candidate-pair comparisons and a resolved terminal top-1. The minimum reported gap is approximately 177 times the fixed 10^{-8} comparison threshold. This establishes resolution under the stated comparison rule; the threshold is not a measured noise floor.

Results and predefined interpretation. Table 18 reports all five states. The cohort medians are $P_F = 0.167$, $P_{T_{\text{full}}} = 0.833$, and $\Delta P = +0.667$, where the last quantity is the median of the within-state differences. Three states improve and two tie; none worsens. Top-1 counts are $0/5$ for F and $3/5$ for T_{full}. These results meet the predefined criterion for poor local ordering with recovery by full-state future sensitivity, with no unresolved comparisons. These results extend the qualitative ordering pattern to the earlier switch in the tested system.

Table 18: Earlier-switch results at $p = 0.50$ on the five Language A seeds. Pairwise accuracy is the primary measure; ΔP is computed within each state. All six pairwise comparisons and the terminal top-1 are resolved in every state under the fixed comparison rule. Values are reported to three decimal places.

Seed	P_F	$P_{T_{\text{full}}}$	ΔP
20260880	0.000	1.000	+1.000
20260881	0.167	0.833	+0.667
20260882	0.000	1.000	+1.000
20260883	0.667	0.667	0.000
20260884	0.833	0.833	0.000
Median	0.167	0.833	+0.667

Comparison with simple baselines. The retrospective analysis in Appendix N shows that $-F$ (reversed fidelity) and the post hoc best fixed rule, always-CURRENT, also achieve $3/5$ top-1 accuracy. Relative to always-CURRENT, T_{full} has lower terminal-loss regret in two states and equal regret in three. Thus its top-1 accuracy matches these simple baselines, despite the improvement over fidelity.

Execution deviations and provenance. The archived execution records document repaired implementation omissions, use of a GPU outside the frozen environment inventory, and omission of a redundant determinism retest specific to $p = 0.75$. Affected work was rerun after the implementation fixes, with no change to frozen scientific decisions. The additional GPU use is recorded as authorized, although artifact verification could not independently recover that authorization record after the documented stop. This provenance gap is separate from the numerical results, which were cross-checked against the archived outputs. No rescue seeds, threshold changes, or additional switch points were introduced.

The archive contains 124 prediction-freeze input hashes and a checksum manifest with 98 records. The two manifests cover different artifact sets.

P CONTROLLED NONZERO-MOMENT RECEIVER COMPARISON

Intervention and score. The controlled protocol constructs nonzero optimizer memory using a fixed $K = 16$ history in each candidate's own factor coordinates, with model parameters held fixed during calibration. The comparison then holds each candidate-specific first moment fixed and contrasts a first-moment-only action with the action obtained by adding the coordinatewise second-moment denominator. The scores compare the implemented dense weight updates through Eq. 2. Equality of the first moment is imposed within a candidate across update rules, not across FORM and PERMUTED FORM. The contrast measures a change from S_{M1} to S_{Adam}; it does not assume either score equals fidelity.

Per-state results and predefined criterion. All five states and all 40 primary FORM/PERMUTED trials completed. The first-moment-only score prefers FORM in every trial. Adam's normalized score prefers PERMUTED FORM in 34 trials, distributed as shown in Table 19. The predefined rule requires at least $5/8$ directional preference shifts within a state and at least $4/5$ states meeting that threshold; all five states satisfy it.

Table 19: Controlled nonzero moments at $K = 16$. Rows correspond to States 1–5 in Figure 2b, in order. Each state contains eight primary trials. The final column counts changes from an M1 preference for FORM to an Adam preference for PERMUTED FORM.

Seed	M1 prefers FORM	Adam prefers PERMUTED	Preference shifts
322587038	8/8	7/8	7/8
559422056	8/8	6/8	6/8
1039487108	8/8	7/8	7/8
1499157004	8/8	7/8	7/8
339743707	8/8	7/8	7/8
Total	40/40	34/40	34/40

No reverse-direction transitions are reported, but M1 prefers FORM in every trial, leaving no M1-PERMUTED preferences to reverse. This zero count adds no symmetric-comparison evidence. Inspection of the archived trial results confirms that the six remaining trials preserve the FORM preference; there are no ties.

Mechanistic interpretation. Adding the denominator induces these preference shifts conditional on the calibrated moments and matched implementation. First-moment directions still influence the normalized action. The result neither extends the weighted-L_1 approximation to nonzero moments nor establishes the same behavior in normally trained AdamW states. The measured outcome is the receiver score, rather than one-step utility or terminal performance.

Implementation verification. Independent recomputation from the raw per-seed accounting tables verifies all 80 candidate-level update reconstructions within the predefined 1% relative-error tolerance; observed relative error is approximately 0.29% (0.285%–0.289%). The validation record also includes 40 CURRENT-arm sanity cells. The reconstruction error is the relative L_2 difference between $-\mathrm{lr}\,\widehat{m}/(\sqrt{\widehat{v}}+\epsilon)$ and the realized parameter increment. These checks validate update implementation; the mechanism evidence comes from the matched denominator intervention. No integrity failure or scientifically meaningful execution deviation is recorded. The archived verification record reports successful checksum verification for all 283 files.

Checkpoint identity and provenance. Artifact verification confirms the moment-update, action-scaling, and accounting-error conventions against the executed code and raw per-seed tables. Checkpoint SHA256 comparison, rather than seed-label matching, establishes that these five states are the same physical checkpoints used by the zero-moment receiver replication, not an additional independent state cohort.