SAME PREDICTOR, DIFFERENT LEARNERS: HOW LORA FACTORIZATION CHANGES ACTIVE LEARNING

Anonymous authors
Paper under double-blind review

ABSTRACT

Two low-rank adaptation (LoRA) learners can make identical predictions and equivalent updates on common data, yet choose different examples to label next. In rank-one Gaussian task families, exact refactorization under the same posterior-variance rule produces a label-cost ratio for the same prediction risk that grows without bound with pool size. The gap persists with arbitrarily accurate conditional covariance and nearly optimal information about the teacher: learning private coefficients improves few predictions, whereas learning a shared coefficient improves many. We formulate a factorization-independent acquisition Gram from effective updates and effective gradients, retaining the geometry of balanced factor gradients. With matched selection inputs and randomness, it preserves ordered purchases across equivalent factorizations; with corresponding common-data updates, it also preserves learning paths. Controlled multi-round experiments with Qwen3 classifiers isolate acquisition as the source of divergence and recover learner correspondence through common purchases or the invariant Gram. A retrospective analysis of repeated training continuations on MultiNLI finds repeatable per-example correctness changes despite similar mean accuracy. Equivalence on common training data therefore does not extend automatically to active learning: the acquisition geometry must also be specified.

1 INTRODUCTION

An active learner chooses which examples to label next. For a low-rank adaptation (LoRA) update $\Delta W = BA$ (Hu et al., 2022), doubling B and halving A leaves the predictor unchanged. Does it also leave acquisition unchanged? A selector built from Euclidean factor gradients can choose a different next batch, even after row normalization and conditioning on the same history (Figure 1).

PEFT adapter composition makes this choice concrete: concatenation and SVD can represent the same weighted update (Hugging Face, 2026). In six synthetic adapter-composition tests on CPU, without truncation or clamping, their predictors agree to numerical precision but raw factor-gradient acquisition chooses different examples (Appendix A.10).

Such gradients are used in instruction and preference data selection (Xia et al., 2024; Lin et al., 2026; Yuan et al., 2026). We isolate acquisition by matching update rules and optimizer states so that effective updates agree on common data, following factorization-aware optimization and coordinated Adam recipes (Yen et al., 2025; Schulman & Thinking Machines Lab, 2025). Each representation then selects its own labels from the same pool with matched randomness. Acquisition can break learner correspondence before either predictor changes. Transporting the reference gradient metric preserves the reference selector's purchases. Alternatively, acquisition can specify its geometry directly from the effective update.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

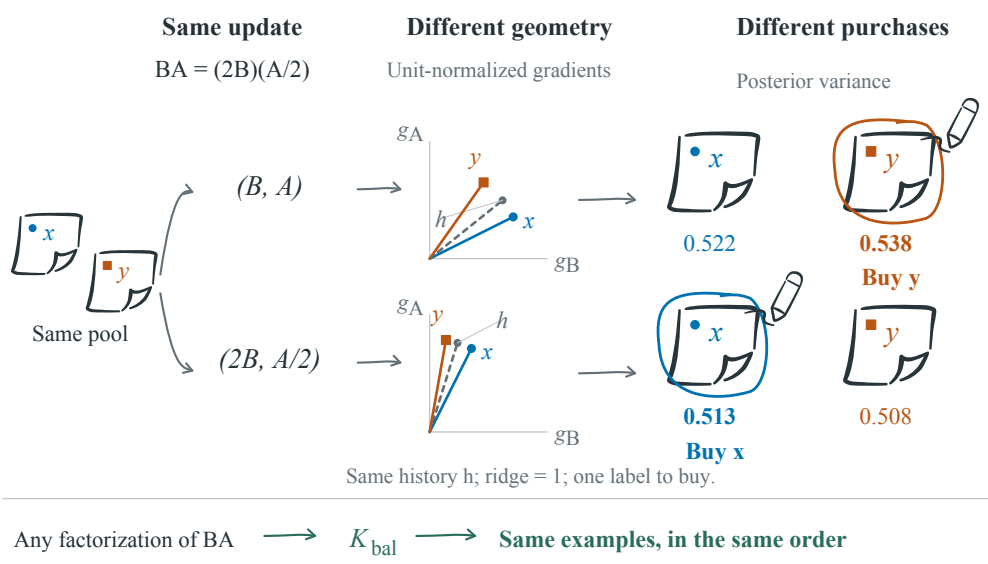


Figure 1: **The same predictor can choose different labels.** Exact rescaling changes normalized factor-gradient geometry and reverses the next choice from y to x, despite the same history h and ridge one (Section 2.2). Using K_{bal} instead preserves the full ordered selection sequence when the remaining selection inputs and randomness are matched.

Our Gaussian constructions show that accuracy on common data need not ensure comparable acquisition costs. In Theorem 2, two exact representations of one rank-one LoRA update use the same variance rule. Conditional covariance remains uniformly accurate after every common observation subset, and raw acquisition obtains nearly optimal true teacher information at every budget. Its kernel predictor also has nearly Bayes-optimal prior-averaged risk on any fixed purchase set (Appendix A.2). Yet its label-cost ratio relative to balanced acquisition grows without bound with pool size for a fixed prediction target.

For reference-free acquisition, we retain the sequential variance rule and use its balanced-factor gradient Gram, expressed through Balanced LoRA's inverse-Bures operator (Castin et al., 2026). Equal Grams preserve ordered purchases and, with equivalent common-data updates, the learning process. Four-round experiments on six Qwen3 classification fits identify acquisition as the break: raw purchases separate the learners, while forced-common purchases and independent inverse-Bures selection both restore their correspondence. Retrospective analysis on eight new MultiNLI fits finds repeatable per-example correctness changes despite similar mean accuracy.

2 Equivalence and Data Selection

We study sequential posterior-variance acquisition, its Gaussian label costs (Section 3) and its factorization-independent paths (Section 4.2).

2.1 From common data to independently selected data

Write $\theta \sim_{\Delta W} \theta'$ when all non-LoRA parameters agree and each LoRA block has the same effective update, $BA = B'A'$. These checkpoints give the same predictor. For training, we additionally require update rules and optimizer states that preserve the effective updates on common labeled minibatches. Appendix A.4 gives the corresponding scalar AdamW recipes.

Acquisition introduces a separate decision. Let U be a candidate pool, ω the selector randomness, b the budget, and z any declared auxiliary state held fixed by the intervention. The state z may contain pool statistics or task information, but not the LoRA factors or an equivalent encoding of those factors. Acquisition is *factorization independent* when, for every such (U, ω, b, z),

$$\theta \sim_{\Delta W} \theta' \quad \Longrightarrow \quad \mathrm{Select}(\theta, U, \omega, b; z) = \mathrm{Select}(\theta', U, \omega, b; z).$$

We call selecting a label a *purchase*, distinct from subsequently exposing the learner to it. Equality means the same ordered purchases. If it holds at each round, the corresponding learners receive common data and can continue to agree. If acquisition depends on the factors, that premise can fail before either predictor changes.

2.2 Factor gradients and sequential selection

For one LoRA block, every $Q \in GL(r)$ defines the exact refactorization

$$(B, A) \mapsto (BQ, Q^{-1}A), \qquad (BQ)(Q^{-1}A) = BA.$$

Let $G_i = \partial \ell_i / \partial (BA)$ be the effective gradient of example i. The factor gradients transform as covectors:

$$g_{B,i} = G_i A^\top, \quad g_{A,i} = B^\top G_i, \qquad g'_{B,i} = g_{B,i} Q^{-\top}, \quad g'_{A,i} = Q^\top g_{A,i}.$$

Increasing the scale of B shrinks g_B but enlarges g_A, changing their relative weights in Euclidean distance.

Proposition 1 (scalar rescaling changes gradient distances). For $Q = cI$, let d_B and d_A be the squared distances contributed by the two factor-gradient blocks. Up to a common positive factor,

$$d_c(i, j) = c^{-2} d_B(i, j) + c^2 d_A(i, j).$$

For two competing pairs, let Δ_B and Δ_A be their differences in d_B and d_A. Opposite signs give a crossing at $c_\star = (-\Delta_B / \Delta_A)^{1/4}$. For example, contributions $(4, 1)$ and $(1, 2)$ have total distances 5 and 3 at $c = 1$, but 5 and 8.25 at $c = 2$. Although $\kappa(cI) = 1$, the complete factor-coordinate Jacobian has $\kappa_2(J) = \max(c^2, c^{-2})$.

We study a unit-normalized, unsketched posterior-variance variant of ActiveDPO (Lin et al., 2026). Given its Gram K and previously selected examples S, the next unselected candidate maximizes

$$v_i(S) = \lambda^{-1} \left[K_{ii} - K_{iS} (\lambda I + K_{SS})^{-1} K_{Si} \right].$$

For nonzero concatenated factor gradients g_i, set $K_{ij} = \langle g_i, g_j \rangle / (\|g_i\| \|g_j\|)$. Here $\lambda > 0$ is the ridge parameter, and S includes the conditioning history and the current purchase prefix. The predictor is frozen while one batch is selected. Figure 1 uses $\lambda = 1$ and factor-block coordinates $h = (1, 4/5)$, $x = (2, 1)$, $y = (1, \sqrt{2})$. Normalizing each row before and after $(g_B, g_A) \mapsto (g_B/2, 2g_A)$ changes the next choice from y to x, despite conditioning on the same example h. For comparison, balanced acquisition uses factor gradients at $B = U\Sigma^{1/2}$, $A = \Sigma^{1/2} V^\top$, where $U\Sigma V^\top$ is the reduced SVD of the effective update. Write K_{bal} for their Gram and use its unit-normalized version in the same rule. Section 4.1 gives its factor-free expression and general path guarantee.

3 How Factorization Changes Label Cost

Different purchases can reach the same prediction target at different budgets. We construct Gaussian tasks in which refactorization changes whether acquisition favors shared or private information.

3.1 Shared and private information

A shared coefficient affects a group of inputs; a private coefficient affects one. Both are genuine predictive signal. Learning a private coefficient can be informative about the teacher yet improve few predictions. Scalar refactorization changes how the variance rule divides the budget between them. The frozen checkpoint defines acquisition features; a common, correctly specified Bayesian learner incorporates purchased labels and evaluates prediction risk.

Both constructions start with balanced factors, whose raw Gram equals K_{bal}. They compare the same raw selector in two exact representations of one update, keeping the Bayesian learner fixed.

Theorem 1 (label cost with informative observations). For every integer $m \geq 2$, there is a rank-one LoRA regression model, a realizable Gaussian teacher prior, a pool of $N = 2m^2$ distinct inputs, and four historical observations with the following properties. Use unit-normalized output gradients,

ridge $\lambda = 1$, and a frozen acquisition checkpoint. Evaluate prior-averaged squared error for the latent teacher mean at a uniform pool input. The balanced Gram equals the true prior covariance on the pool and history. After conditioning on history, both arms start at risk $2/3$. Raw acquisition from $(2B, A/2)$ needs more than m^2 new labels to reach risk $1/2$, even with the optimal Bayesian learner. The same sequential rule using K_{bal} reaches this target within $2m$ purchases, for every fixed tie order. Each of the first m^2 raw purchases supplies at least $\frac{1}{2}\log(3/2)$ nats of conditional information about the teacher. Including all four historical labels, the label-cost ratio is greater than $(m^2+4)/(2m+4) = \Omega(\sqrt{N})$.

Why the costs separate. The teacher has m groups of m inputs sharing a coefficient and m^2 singleton inputs, each with its own coefficient. Balanced unit gradients give the true covariance and cover every group within $2m$ purchases. Doubling B and halving A instead makes each fresh singleton outrank every group for the first m^2 purchases. A group observation improves prediction on m inputs; a singleton's private information improves one. With true latent posterior covariance C_S after the observations S, querying candidate i reduces expected risk by

$$\Delta R(i \mid S) = \frac{\|C_S[:, i]\|^2}{N(1 + (C_S)_{ii})}.$$

Pointwise variance measures uncertainty at the queried input; the numerator here measures how its observation affects predictions across the pool. Appendix A.1 gives the exact risks, rank-one realization and proof of the cost ratio.

3.2 The cost gap persists with accurate uncertainty

The next construction makes the uncertainty estimates arbitrarily accurate while retaining the cost gap. The refactorized rule also acquires nearly optimal information about the teacher at every budget. Thus neither poor covariance estimation nor uninformative observations explain the separation. Every input has a private signal of fixed strength, so within-group latent means are distinct.

Theorem 2 (accurate covariance, expensive prediction). For every $0 < \delta \leq 1/4$, there are arbitrarily large pools of size N in a rank-one Gaussian task using the selector and risk of Theorem 1. The construction has one common paid history label. After history and every further subset S of distinct candidate IDs, the true and refactorized acquisition covariances satisfy

$$C_S \succeq I/5, \qquad (1-\delta)C_S \preceq \widetilde{C}_S \preceq (1+\delta)C_S.$$

For every target in the fixed interval $[7/8, 29/32]$, balanced acquisition needs one new label, whereas raw needs more than $N/8$. Including history, their cost ratio exceeds $N/16$. At every budget, raw's information under the true teacher model is within $\delta^2/16$ nats of the globally optimal batch of the same size. At one new label, its prediction risk exceeds balanced acquisition's risk by more than $1/16$. Appendix A.2 gives the proof. Figure 2 uses $7/8$, the lower endpoint of this fixed target interval. The label-cost separation holds throughout the interval.

Why the same predictor buys private information first. Set $c = ((1+\delta)/(1-\delta))^{1/4}$. Row normalization changes the relative block weights from $1:1$ to $(1-\delta):(1+\delta)$. One history label then reverses the ordering between a shared group with private components and a family of mostly private observations. Balanced buys a shared label first; raw buys the entire private family before reaching the group. Both orderings are strict, despite uniformly accurate conditional covariance.

Separately, every raw prefix globally maximizes batch mutual information under the refactorized proxy; exact batch optimization retains the cost gap. With $N = 2m$ and $\delta = m^{-2}$, absolute covariance error is uniformly below $1/m$ and the largest teacher-information deficit is at most $1/(16m^4)$, while the cost ratio exceeds $m/8$. Each arm may also predict with its own kernel: on any fixed purchase set, raw's prior-averaged risk is within $1/(1-\delta^2)$ of Bayes risk. Appendix A.2 proves these extensions.

The scale of the amplification. Let Λ be the largest true covariance eigenvalue after paid history. For a fixed attainable risk improvement, the history-inclusive cost ratio of any complete, nonrepeating order to the best set is at most linear in Λ. Theorem 2 attains it as dimension and Λ grow.

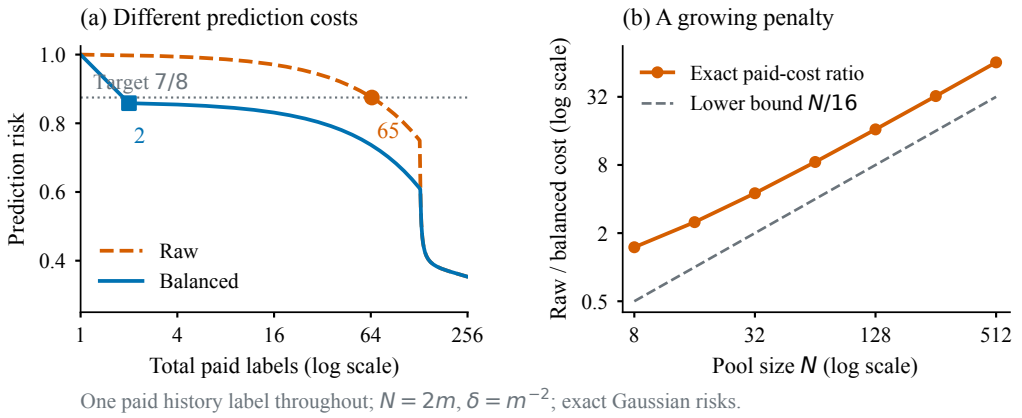


Figure 2: **Accurate covariance can still demand more labels for prediction** (Theorem 2). Left: in a 256-candidate pool, reaching the illustrative risk target $7/8$ requires 65 total raw labels versus two balanced labels, including one common history label. Right: the same target's cost ratio grows along $N = 2m$, $\delta = m^{-2}$, while covariance error and the largest additive teacher-information deficit over budgets vanish. Curves are exact latent-mean Bayes risks averaged over every pool ID; markers indicate first attainment.

At fixed δ, the separation tolerates score errors below $3\delta^2/64$, independently of pool size (Appendix A.2). The additional sequence $\delta = m^{-2}$ makes covariance and information errors vanish, but its strict ordering margin shrinks as $O(m^{-4})$. Comparable predictive reach and bounded variance instead give a fixed-task risk-slack guarantee with extra budget (Appendix A.2; Chamon & Ribeiro, 2017).

The failure depends on what acquisition rewards. A variance score examines one candidate; pool-risk reduction also credits improvements to other predictions. Exact LoRA refactorization realizes the shared-versus-private pivoting pattern (Chen et al., 2025; Epperly, 2025) while retaining accurate covariance and near-optimal teacher information. In Theorem 2's family, greedy reduction of the proxy's pool-wide risk reaches every stated true-risk target after one new label (Appendix A.2). Objective design chooses which predictive improvements to seek; factorization-independent acquisition makes that choice reproducible across representations. Its balanced Gram matches the true covariance here.

4 Preserving the Learning Process

We keep the sequential variance rule and replace its raw Gram with the balanced-factor gradient Gram, expressed through the effective update. Transporting a reference metric retains its purchases; this construction instead defines a reference-free balanced-gradient objective. The two need not select the same batch. It covers full-rank $GL(r)$ orbits and effective-update classes containing several orbits after rank loss.

4.1 An inverse-Bures acquisition Gram

For $\Delta W = U\Sigma V^\top$, the balanced factors $U\Sigma^{1/2}$ and $\Sigma^{1/2}V^\top$ split its scale symmetrically. Their gradient Gram uses Balanced LoRA's inverse-Bures geometry (Castin et al., 2026). Any reduced factors of the same update satisfying $\bar{B}^\top\bar{B} = \bar{A}\bar{A}^\top$ differ only by a common orthogonal transform and give the same Euclidean gradient Gram. The balanced split therefore specifies the geometry without a storage reference (Appendix A.3.1). For each LoRA block, set $P_L = (\Delta W \Delta W^\top)^{1/2}$ and $P_R = (\Delta W^\top \Delta W)^{1/2}$. These matrices weight the update's left and right singular directions by their singular values. Summing the resulting gradient inner products over blocks gives

$$K_{\mathrm{bal}}(i,j) = \sum_m \left\langle G_i^{(m)}, P_L^{(m)} G_j^{(m)} + G_j^{(m)} P_R^{(m)} \right\rangle_F .$$

Proposition 2 (inverse-Bures acquisition Gram). K_{bal} depends only on the effective updates and effective gradients, and is therefore identical throughout each $\sim_{\Delta W}$ class, including rank-deficient representations that need not share a single $GL(r)$ orbit. It equals the Gram of gradients in the reduced balanced factorization $\bar{B} = U\Sigma^{1/2}$, $\bar{A} = \Sigma^{1/2}V^\top$. At full inner rank it can be computed with thin QR factorizations and an $r \times r$ SVD, without materializing dense ΔW. The factor-free expression remains well-defined at rank deficiency. Appendix A.3.1 gives the proof, numerical-rank convention, cost analysis and CPU checks with general non-scalar transformations.

After rank loss, two small least-squares solves recover reduced balanced gradients. Numerical execution requires resolved ranks and solves; admission can differ across equivalent factorizations (Appendix A.3.1). In the audited AG News inverse-Bures trajectories, every pre-acquisition block retained rank eight under the original FP64 cutoff, with minimum $\sigma_8/\tau = 1.49 \times 10^9$ (Appendix Table 4). When all updates are zero, $K_{\mathrm{bal}} = 0$. We use warmed adapters; initialization requires a declared factor-independent rule such as candidate order or entropy.

4.2 From equal Grams to equal learning paths

The guarantees compose in three steps. First, equal effective updates and effective gradients give equal K_{bal} by Proposition 2; each loss must depend on the factors through BA. Second, a common pool, conditioning history and purchase prefix give equal conditioned scores. Matching auxiliary selection state, candidate order, tie rule and randomness then preserves the complete ordered batch. This step applies to selectors whose only factor input is K_{bal}; optimizer-derived features, clipping or coordinate sketches used in acquisition must also be transported or replaced.

Third, common replay randomness and corresponding updates on identical minibatches preserve learner correspondence into the next round. Our scalar AdamW recipes satisfy this additional condition (Appendix A.4); optimizer correspondence is not needed for the Gram identity itself.

Although the Qwen experiments store the full Gram, the selector needs only $\mathrm{diag}(K)$ and $K_{:,S}$. Exact column streaming preserves its path with $O(N|S| + |S|^2)$ kernel working memory, separate from model and gradient-feature storage (Appendix A.3.1).

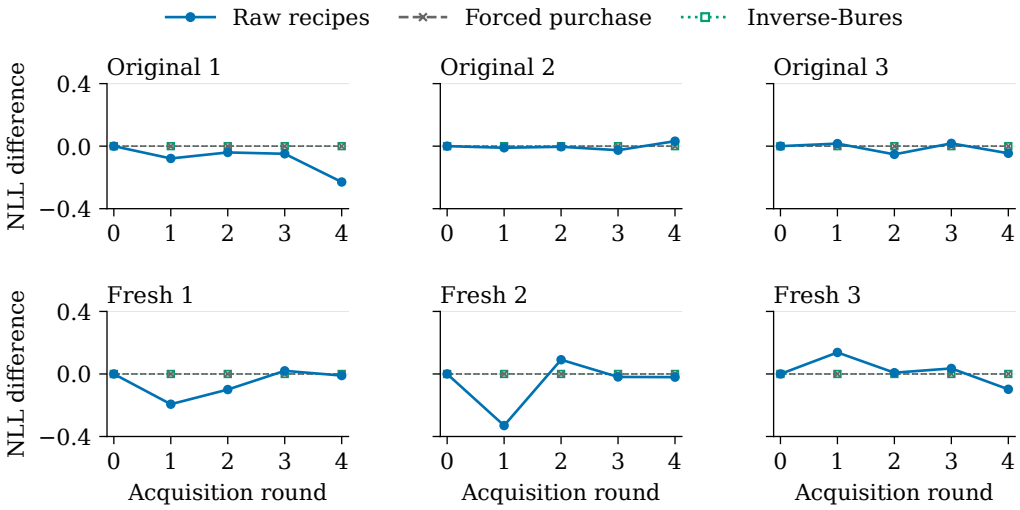
5 Experiments

Neural studies test learner correspondence on six AG News fits and repeatable prediction changes on eight MultiNLI fits (Appendix Table 3), separately from the Gaussian cost bounds. Frozen-gradient replays with D^2 selection (squared-distance sampling) also change raw purchases under refactorization, while inverse-Bures purchases remain matched (Appendix A.9).

5.1 Breaking and restoring a four-round learning process

We used six Qwen3-1.7B (Yang et al., 2025) fits on AG News (Zhang et al., 2015), with rank-8 LoRA on query and value projections in all 28 layers. Each fit started with 256 warm labels and 1,000 candidates, then acquired 128 labels per round. Three fits extended an initial two-round study; three new fits used disjoint warm and candidate data. The paired $c = 1, 2$ AdamW recipes coordinated factor scales and optimizer settings to preserve the effective-update trajectory on common data (Appendix A.4). They produced the same warm predictor. Both used normalized posterior-variance acquisition with independently extracted pseudo-label cross-entropy gradients. Conditioning started empty, then included refreshed past purchases; warm labels fitted the predictor but were excluded from this gradient-feature covariance proxy. Common candidate order fixed the first tied purchase.

Seven arms compared raw and inverse-Bures acquisition under both recipes, $c = 2$ forced to consume raw $c = 1$ purchases, and $c = 1$ random-purchase and old-label-only controls. Each retained optimizer state for 64 eight-example updates per round on ID-sorted cumulative data. From round three, this data exceeds the update budget: round four consumes 79–97 of 128 new labels, and purchases also change earlier-label replay. We compare the complete adaptive procedures at a fixed update budget, including this replay-mediated effect. Four-class negative log-likelihood (NLL) on 1,024 shared test examples measured learning. Fits are the six independent units; rounds and recipe pairs are repeated measurements (Appendix A.5).



Recipe pair	Prediction changes / 1,024 test examples	Absolute area M (secondary, NLL)	Equal endpoints / 24 fit-rounds
Raw recipes	34–514	0.014–0.113	0
Forced purchase	0	0	24
Inverse-Bures	0	0	24

Figure 3: **Raw acquisition separates the learners; both matched controls preserve their correspondence.** Per-fit held-out NLL contrasts ($c = 2$ minus $c = 1$) for raw, forced and inverse-Bures purchases. Forced-common purchases isolate acquisition from the update rule; independently computed inverse-Bures purchases preserve correspondence without sharing the selected IDs. Round j follows $128j$ new labels from 256 warm labels.

Before the predictors diverged, raw acquisition replaced 17–36 of 128 first-round labels, with set differences persisting across ridge and candidate-order checks (Appendix Table 5). Subsequent training separated the learners (Figure 3). Forcing $c = 2$ to consume raw $c = 1$ purchases recovered every held-out probability and the mapped parameter and optimizer states in all 24 fit–round endpoints. Independently computed inverse-Bures purchases preserved the same correspondence.

What changed in learning. Raw risk ordering reversed across rounds in five of six fits. The prespecified summary was $\mathcal{A}_{c=2} - \mathcal{A}_{c=1}$ for $\mathcal{A} = (\mathcal{L}_0 + 2\mathcal{L}_1 + 2\mathcal{L}_2 + 2\mathcal{L}_3 + \mathcal{L}_4)/8$, with $\mathcal{L}_j$ the round-j NLL. Mean contrasts were -0.032 NLL (six fits; descriptive 95% fit-bootstrap range $[-0.060, 0.002]$) and -0.035 (three fresh fits; $[-0.070, 0.033]$). The intervals span zero and do not establish a consistent directional advantage; signed mean risk does not test path correspondence.

The post hoc absolute area ranged from 0.014 to 0.113 NLL; both controls were zero.

All acquisition rules finished below warm and equal-update old-only risk, with fit-dependent risk and target-cost rankings (Appendix Figures 4–5).

Repeated continuations of the fixed first-round purchases left their conditional mean-NLL differences unresolved under the prespecified simultaneous intervals (Appendix A.8). Complete-consumption replay of the four-round curricula retained a 0.0044 NLL mean contrast after alignment to each fit–round's previously observed direction, mostly from round one; acquisition was not rerun (Appendix A.7).

5.2 Similar mean accuracy, different correct predictions

We retrospectively analyze repeated MultiNLI (Williams et al., 2018) continuations for per-example correctness changes despite similar mean accuracy. The runs originated in a prospective domain-direction study that did not meet its joint criterion (Appendix A.6).

Table 1: **Purchase changes repeatedly redistribute correct predictions across eight new fits.** Each fit has 16 paired continuations and the same 1,024-example MultiNLI panel. RMS estimates the magnitude of the per-example correctness-probability response; Δ accuracy is its signed mean (high minus low). Both are in percentage points (pp). r compares responses from the first and last eight streams. This endpoint analysis is retrospective.

Fit	Response RMS (pp)	Δ accuracy (pp)	Split-half r
1	7.6	-0.53	0.34
2	7.9	+0.96	0.41
3	6.5	+0.00	0.38
4	8.9	-0.73	0.49
5	9.1	+0.45	0.54
6	6.5	+0.04	0.45
7	9.8	+0.18	0.52
8	10.0	-0.59	0.64

We analyzed eight new Qwen3-1.7B fits. Two refactorizations giving high and low government purchase counts were fixed using public domain metadata on two development fits, then applied unchanged to all eight. Each new fit had 256 warm labels, 1,000 candidates and a 256-label budget. Sixteen paired continuations started from common warm parameters and AdamW state, matched shared-example positions and exposed every purchase four times. Unlike AG News, acquisition conditioned on warm-history gradients. The common test panel contained 512 previously unused government examples and 512 previously unused telephone examples (Appendix A.6).

Let $p_h(i)$ and $p_l(i)$ be the probabilities of answering example i correctly over training randomness under the two purchases. Cross-stream products estimate the mean squared response $N^{-1}\sum_i[p_h(i)-p_l(i)]^2$, avoiding the positive noise bias from squaring per-example sample means. We report its square root as RMS and contrast it with the signed mean accuracy change (Table 1).

All eight fits had positive response estimates and positive first/last-eight correlations. The equal-fit pooled RMS was 8.4 percentage points (approximate 95% fit-level interval 7.2–9.4, conditional on the test panel), while every fit's overall accuracy difference was below one point in magnitude. Removing domain–class mean responses retained 98.7% of the mean squared signal. The repeatability is within each fit and its fixed purchase pair across training streams; it does not imply that the same test examples benefit across different fits. Same-purchase controls provide a training-randomness reference, with different stream coupling and fewer independent differences; they are not a matched test that the purchase effect exceeds training noise (Appendix A.6.1).

6 Related Work

LoRA-gradient data acquisition. ActiveDPO uses projected, normalized implicit-reward gradient differences for sequential uncertainty sampling (Lin et al., 2026); we study its posterior-variance primitive with unsketched classification gradients. LESS targets influential examples through Adam-update features and checkpoint-weighted cosine scores (Xia et al., 2024); GradFiltering filters data through adapter gradients (Yuan et al., 2026). GIST explicitly studies coupled LoRA optimization geometry and builds a task-relevant subspace from validation gradients (Min et al., 2026). Our question is whether a specified decision survives exact refactorization. The cost bounds concern the variance rule; full LESS and GIST pipelines also require transformation rules for optimizer statistics, target features and projections.

STAR addresses uncertainty gaps and calibration in active LoRA fine-tuning (Zhang et al., 2024); C-PEAL calibrates uncertainty for parameter-efficient active learning in vision-language models (Lakshmi Narayanan et al., 2025). Calibration and factorization independence are distinct: a deterministic acquisition score based only on unchanged predictive probabilities, such as entropy, is already invariant.

Gauge-aware LoRA training and balancing. LoRA+ and LoRA-RITE address factor-specific learning rates and optimization state (Hayou et al., 2024; Yen et al., 2025); refactorization also appears in training and aggregation (Zhang et al., 2025; Lion et al., 2025; Chen et al., 2026). Balanced

LoRA supplies the inverse-Bures operator used here (Castin et al., 2026); PRISM uses tangent-Frobenius geometry for private per-example clipping and noise (Wang & Zhang, 2026). We use the former's per-example Gram to determine which labels are acquired. The new results concern acquisition paths and label costs, rather than the balancing operator itself. Appendix A.3.1 compares the invariant Grams.

Active-learning robustness and exact-coordinate audits. Gradient and Bayesian acquisition (Ash et al., 2020; 2021; Kirsch & Gal, 2022; Pinsler et al., 2019) have been studied under model and training variability (Lowell et al., 2019; Gallé & Mena Roldán, 2020; Munjal et al., 2022; Kirsch et al., 2023). Predictive multiplicity concerns comparably accurate models (Marx et al., 2020; Ganesh et al., 2025); our intervention holds the predictor fixed.

Prediction-oriented and loss-driven acquisition distinguish information from downstream prediction (Bickford Smith et al., 2023; Huang et al., 2026). Independent-versus-correlated pivoting examples (Chen et al., 2025, App. C.2) and experimental-design objectives (Chamon & Ribeiro, 2017; Epperly, 2025) provide related distinctions between shared and private information. The specific example in Chen et al. (2025, Theorem C.2) uses all-ones blocks and unfavorable tie-breaking. Our second construction retains fixed-strength private signals at every input and strictly reverses selection after one common paid observation. Exact refactorization of one LoRA update preserves uniformly accurate conditional covariance and near-optimal true teacher information, yet incurs a growing label-cost ratio. The spectral bound characterizes its scale. Model-mismatch approaches study how acquisition and prediction can differ (Rahmati et al., 2024). Jarl et al. (2026) correct boundary bias and introduce observation dependence in acquisition while leaving the prediction GP unchanged.

Implementation invariance and reparameterization-invariant inference motivate fixing the predictor (Sundararajan et al., 2017; Roy et al., 2024). Exact coordinate interventions also examine pruning decisions and representation measurements (Asiaee, 2026; Karasan, 2026). Acquisition instead changes future observations and training data; matched updates isolate its contribution to learner divergence.

7 Discussion

To retain a validated recipe's purchases under refactorization, transport the reference metric as $M' = JMJ^\top$ under Jacobian J; our controls recover the reference paths (Appendix A.5). To retain balanced-factor geometry without a storage reference, use K_{bal}. PRISM's tangent-Frobenius and ambient Grams specify other invariant objectives (Appendix A.3.1). Task validation chooses the objective; invariance preserves that choice across representations.

The Gaussian separation grows with dimension and shared-covariance scale. Neural risk and domain effects concern matched Qwen3-1.7B classification recipes and depend on training exposure and the evaluation panel. Numerical use of the factor-free Gram also requires resolved ranks and solves (Appendix A.3.1).

8 Conclusion

Equivalence on common training data does not automatically extend to active learning. Factor-gradient acquisition can choose different next labels before either predictor changes. The Gaussian tasks quantify a possible cost in labels; the neural controls isolate acquisition as the source of divergence and recover correspondence through common or factorization-independent purchases. A reusable learning rule must therefore specify not only how to update on data, but also how to choose that data.

AI USE STATEMENT

In this work, we used generative AI tools to assist with refining hypotheses, developing mathematical claims and proofs, designing and implementing experiments, synthetic numerical validation, and interpreting results. They also assisted with literature search, manuscript drafting and editing, translation, and figure preparation. Surveys, interviews, and qualitative coding are not applicable to this study. We reviewed the AI-assisted work, including the mathematical arguments, code, and interpretation; numerical reference checks and regression tests supported this review. We take responsibility for the manuscript and its accompanying artifacts.

ETHICS STATEMENT

This diagnostic study uses public benchmark data. It queries labels only in simulation and involves no human subjects, private data, or deployed annotation workers. The conclusions concern the audited architectures and factorization family; Section 7 describes the intended use.

REPRODUCIBILITY STATEMENT

Appendices A.1–A.4 give the assumptions, proofs, and matched update rules. The accompanying code includes the acquisition implementation, experimental configurations, data-split specifications, analysis scripts, and selected numerical records. Its `README.md` describes execution and the scope of the released materials.

REFERENCES

Jordan T. Ash, Chicheng Zhang, Akshay Krishnamurthy, John Langford, and Alekh Agarwal. Deep batch active learning by diverse, uncertain gradient lower bounds. In *International Conference on Learning Representations*, 2020.

Jordan T. Ash, Surbhi Goel, Akshay Krishnamurthy, and Sham Kakade. Gone fishing: Neural active learning with fisher embeddings. In *Advances in Neural Information Processing Systems*, 2021.

Amir Asiaee. Causal mechanism reduction: Mechanism replacement for neural network pruning and abstraction. *arXiv preprint arXiv:2602.24266*, 2026.

Freddie Bickford Smith, Andreas Kirsch, Sebastian Farquhar, Yarin Gal, Adam Foster, and Tom Rainforth. Prediction-oriented bayesian active learning. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, pp. 7331–7348, 2023.

Valérie Castin, Kimia Nadjahi, Pierre Ablin, and Gabriel Peyré. Balanced LoRA: Removing parameter invariance to accelerate convergence. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://arxiv.org/abs/2605.31484.

Luiz F. O. Chamon and Alejandro Ribeiro. Approximate supermodularity bounds for experimental design. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/hash/0d9095b0d6bbe98ea0c9c02b11b59ee3-Abstract.html.

Jinqian Chen, Chang Liu, and Jihua Zhu. Beyond factor aggregation: Gauge-aware low-rank server representations for federated LoRA. *arXiv preprint arXiv:2605.06733*, 2026.

Yifan Chen, Ethan N. Epperly, Joel A. Tropp, and Robert J. Webber. Randomly pivoted Cholesky: Practical approximation of a kernel matrix with few entry evaluations. *Communications on Pure and Applied Mathematics*, 78(5):995–1041, 2025. doi: 10.1002/cpa.22234.

Ethan N. Epperly. *Make the Most of What You Have: Resource-Efficient Randomized Algorithms for Matrix Computations*. PhD thesis, California Institute of Technology, 2025. URL https://tropp.caltech.edu/dissertations/Epp25-Making-Most.pdf.

Matthias Gallé and José Mena Roldán. Robust active learning strategies for model variability. In *HAMLETS Workshop at NeurIPS*, 2020. URL https://openreview.net/pdf/c4134521809d14cef648d258f0a1de49e450f355.pdf.

Prakhar Ganesh, Hsiang Hsu, and Golnoosh Farnadi. Data as a lever: A neighbouring datasets perspective on predictive multiplicity. *arXiv preprint arXiv:2510.21303*, 2025. URL https://arxiv.org/abs/2510.21303v2.

Soufiane Hayou, Nikhil Ghosh, and Bin Yu. LoRA+: Efficient low rank adaptation of large models. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 17783–17806, 2024.

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.

Zhuoyue Huang, Freddie Bickford Smith, and Tom Rainforth. Loss-driven bayesian active learning. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 5140–5148, 2026. URL https://proceedings.mlr.press/v300/huang26b.html.

Hugging Face. PEFT v0.20.0: LoRA adapter composition source code. Versioned software source (add_weighted_adapter), 2026. URL https://github.com/huggingface/peft/blob/v0.20.0/src/peft/tuners/lora/model.py. Release v0.20.0; accessed September 9, 2026.

Hugging Face. PEFT checkpoint format. PEFT documentation, n.d. URL https://huggingface.co/docs/peft/developer_guides/checkpoint. Accessed September 9, 2026.

Sanna Jarl, Jens Sjölund, Jonathan J. S. Scragg, and Maria Bånkestad. Correcting boundary bias and observation independence in Bayesian experimental design, 2026. URL https://arxiv.org/abs/2602.01898v2.

Abdullah Karasan. Representation measurements under function-preserving reparameterizations. *arXiv preprint arXiv:2608.27020*, 2026.

Andreas Kirsch and Yarin Gal. Unifying approaches in active learning and active sampling via fisher information and information-theoretic quantities. *Transactions on Machine Learning Research*, 2022.

Andreas Kirsch, Sebastian Farquhar, Parmida Atighehchian, Andrew Jesson, Frédéric Branchaud-Charron, and Yarin Gal. Stochastic batch acquisition: A simple baseline for deep active learning. *Transactions on Machine Learning Research*, 2023.

Athmanarayanan Lakshmi Narayanan, Amrutha Machireddy, and Ranganath Krishnan. Optimizing active learning in vision-language models via parameter-efficient uncertainty calibration. *arXiv preprint arXiv:2507.21521*, 2025. URL https://arxiv.org/abs/2507.21521.

Xin Li and Dan Roth. Learning question classifiers. In *COLING 2002: The 19th International Conference on Computational Linguistics*, 2002. URL https://aclanthology.org/C02-1150/.

Xiaoqiang Lin, Arun Verma, Zhongxiang Dai, Daniela Rus, See-Kiong Ng, and Bryan Kian Hsiang Low. ActiveDPO: Active direct preference optimization for sample-efficient alignment. In *International Conference on Learning Representations*, 2026.

Kai Lion, Liang Zhang, Bingcong Li, and Niao He. PoLAR: Polar-decomposed low-rank adapter representation. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

David Lowell, Zachary C. Lipton, and Byron C. Wallace. Practical obstacles to deploying active learning. In *Proceedings of EMNLP-IJCNLP*, pp. 21–30, 2019.

Charles Marx, Flavio Calmon, and Berk Ustun. Predictive multiplicity in classification. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6765–6774, 2020. URL https://proceedings.mlr.press/v119/marx20a.html.

Guanghui Min, Tianhao Huang, Ke Wan, and Chen Chen. GIST: Targeted data selection for instruction tuning via coupled optimization geometry. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://arxiv.org/abs/2602.18584.

Prateek Munjal, Nasir Hayat, Munawar Hayat, Jamshid Sourati, and Shadab Khan. Towards robust and reproducible active learning using neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 223–232, 2022.

Robert Pinsler, Jonathan Gordon, Eric Nalisnick, and Jose Miguel Hernandez-Lobato. Bayesian batch active learning as sparse subset approximation. In *Advances in Neural Information Processing Systems*, 2019.

Amir Hossein Rahmati, Mingzhou Fan, Ruida Zhou, Nathan M. Urban, Byung-Jun Yoon, and Xiaoning Qian. Understanding uncertainty-based active learning under model mismatch, 2024. URL https://arxiv.org/abs/2408.13690.

Hrittik Roy, Marco Miani, Carl Henrik Ek, Philipp Hennig, Marvin Pförtner, Lukas Tatzel, and Søren Hauberg. Reparameterization invariance in approximate bayesian inference. In *Advances in Neural Information Processing Systems*, 2024.

John Schulman and Thinking Machines Lab. LoRA without regret. Thinking Machines Lab: Connectionism, 2025. URL https://thinkingmachines.ai/blog/lora/.

George P. H. Styan. On some inequalities associated with ordinary least squares and the Kantorovich inequality. In *Festschrift for Eino Haikala on his Seventieth Birthday*, volume 153 of *Acta Universitatis Tamperensis, Series A*, pp. 158–166. University of Tampere, 1983. URL https://webpages.tuni.fi/uta_statistics/misc/Styan-83-Tampere.pdf.

Mukund Sundararajan, Ankur Taly, and Qiqi Yan. Axiomatic attribution for deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 3319–3328, 2017.

Shihao Wang and Xueru Zhang. Prism: Gauge-invariant tangent-space differentially private lora. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. URL https://arxiv.org/abs/2606.00944.

Adina Williams, Nikita Nangia, and Samuel R. Bowman. A broad-coverage challenge corpus for sentence understanding through inference. In *Proceedings of NAACL-HLT*, pp. 1112–1122, 2018. doi: 10.18653/v1/N18-1101.

Mengzhou Xia, Sadhika Malladi, Suchin Gururangan, Sanjeev Arora, and Danqi Chen. LESS: Selecting influential data for targeted instruction tuning. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 54104–54132, 2024.

An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, et al. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025. URL https://arxiv.org/abs/2505.09388.

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

Zhihang Yuan, Chengyu Yue, Long Huang, Litu Ou, and Lei Shi. Uncertainty-aware gradient signal-to-noise data selection for instruction tuning. *arXiv preprint arXiv:2601.13697*, 2026.

Linhai Zhang, Jialong Wu, Deyu Zhou, and Guoqiang Xu. STAR: Constraint LoRA with dynamic active learning for data-efficient fine-tuning of large language models. In *Findings of the Association for Computational Linguistics: ACL 2024*, pp. 3519–3532, 2024. doi: 10.18653/v1/2024.findings-acl.209.

Xiang Zhang, Junbo Zhao, and Yann LeCun. Character-level convolutional networks for text classification. In *Advances in Neural Information Processing Systems*, volume 28, pp. 649–657, 2015.

Yilang Zhang, Bingcong Li, and Georgios B. Giannakis. RefLoRA: Refactored low-rank adaptation for efficient fine-tuning of large models. In *Advances in Neural Information Processing Systems*, volume 38, 2025.

A PROOFS AND EXPERIMENTAL DETAILS

Reading guide. Appendices A.1–A.3 give the cost proofs and geometry; Appendix A.4 specifies matched updates. The AG News and MultiNLI protocols begin in Appendices A.5 and A.6, followed by fixed-purchase and implementation checks. Reproduction entry points are in the accompanying code's `README.md`.

Table 2: Frequently used notation. Gaussian prediction risk and neural test NLL are distinct quantities.

Symbol	Meaning
$\Delta W = BA$	Effective update and its two LoRA factors
$G_i,\ g_i$	Effective-update gradient and concatenated factor gradient
S	Conditioning history and current purchase prefix
$K,\ K_{\mathrm{bal}}$	Acquisition Gram and balanced-factor gradient Gram
$C_S,\ \widetilde{C}_S$	True and refactorized conditional covariances
$R(S)$	Gaussian latent-mean prediction risk after observing S
$\mathcal{L}_j,\ \mathcal{A}$	Round-j neural test NLL and normalized curve area
Δ_{dom}	Difference of domain-specific NLL advantages
$n_{\mathrm{rep}},\ U_a$	Number of repeated streams and estimated mean squared correctness response
τ	Original full-rank FP64 rejection cutoff in the Qwen rank audit

Table 3: Experimental units and the question answered by each study. Repeated continuations reuse a fitted learner and its purchases; they are not additional model fits.

Study	Fits	Learning comparison and endpoint
AG News, adaptive	6	Four rounds; raw, forced-common and inverse-Bures paths locate the loss of learner equivalence. Risk curves compare the resulting learners (Appendix A.5).
AG News, replay	Same 6	Eight streams per fixed first-round purchase; 16 streams per complete four-round curriculum. These test conditional learning outcomes under specified continuation schedules (Appendices A.8, A.7).
MultiNLI, development	2	Fixed-purchase continuations develop a domain-composition contrast and examine per-example responses (Appendix A.8.1).
MultiNLI, new fits	8	Sixteen paired streams per fit test the frozen domain contrast. A subsequent response analysis measures repeatable per-example changes (Appendix A.6).

A.1 PROOF OF THEOREM 1: INFORMATIVE LABELS WITH DIFFERENT LEARNING VALUE

Rank-one model and true teacher. Let $m \geq 2$, $L = m^2$, and $N = 2L$. Use orthonormal left vectors $e_L, u_0, u_1, \ldots, u_m, u_*$ and right vectors $e_R, v_0, v_1, \ldots, v_L, v_*$. The scalar model is $f_{b,a}(X) = \langle ba^\top, X\rangle_F$, with acquisition checkpoint $(b, a) = (e_L, e_R)$. Draw all $\alpha_0, \ldots, \alpha_m$ and $\beta_0, \ldots, \beta_L$ independently from $\mathcal{N}(0, 1)$ and set

$$b_\theta = e_L + \sum_{i=0}^{m} \alpha_i u_i, \qquad a_\theta = e_R + \sum_{j=0}^{L} \beta_j v_j, \qquad Y(X) = \langle b_\theta a_\theta^\top, X\rangle_F + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, 1).$$

One teacher draw is shared by all observations; their noises are independent. Every teacher has rank one. Write $Z = u_* v_*^\top$ and define

$$X_{ik}^A = \frac{u_i e_R^\top + e_L v_0^\top}{\sqrt{2}} + \eta_{ik} Z, \quad i, k \in \{1, \ldots, m\},$$

$$X_j^B = \frac{u_0 e_R^\top + e_L v_j^\top}{\sqrt{2}} + \eta_j Z, \quad j \in \{1, \ldots, L\}.$$

Distinct bounded tags η, fixed before drawing the teacher, distinguish all inputs. The tag direction is annihilated by the teacher and checkpoint gradients. The complete candidate means nevertheless depend on the teacher: $\mu^A_{ik} = (\alpha_i + \beta_0)/\sqrt{2}$ and $\mu^B_j = (\alpha_0 + \beta_j)/\sqrt{2}$. Let history contain $h \geq 2$ distinctly tagged inputs of each type $u_0 e_R^\top$ and $e_L v_0^\top$, and put $D = h+1$, $H = 3D$. These $2h$ labels are independent-noise observations of α_0 and β_0, shared and charged in both arms. Their posterior variances are $1/D$. Theorem 1 uses $h = 2$; we retain h in the formulas below.

The prediction target is uniform over all N candidate IDs, including those already queried. Risk is expected squared error for their latent means, averaged over teacher and observed labels. The learner knows the prior and uses the posterior mean. Conditional squared-error decomposition makes this risk minimal among all estimators using the same observations. For a fresh noisy target label, risk and the target threshold are both exactly one larger.

Acquisition covariance. At $(b_c, a_c) = (ce_L, e_R/c)$, actual factor gradients are $(Xa_c, b_c^\top X)$. Put $a = 1/(1+c^4)$ and $b = c^4/(1+c^4)$ in the following feature expressions:

$$\phi_A = \sqrt{a}\, u_i \oplus \sqrt{b}\, v_0, \qquad \phi_B = \sqrt{a}\, u_0 \oplus \sqrt{b}\, v_j.$$

Unit-normalized history features are $u_0 \oplus 0$ and $0 \oplus v_0$. At $c = 1$, these rows are exactly the coefficients of the true Gaussian observation model on all candidate and historical inputs. Thus their Gram is the true prior covariance, and ridge-one conditioning gives the exact posterior latent variance. This is covariance calibration; the frozen checkpoint need not equal the posterior-mean predictor after history. At this unit-singular-value checkpoint, K_{bal} equals the balanced factor Gram. Both arms retain the same true prior and downstream learner; the raw selector retains the Euclidean normalized-gradient proxy after refactorization. Transporting the features instead would preserve the reference covariance and path.

The raw prefix. After q distinct B purchases and no A purchase, eliminate the observed private coordinates in the ridge precision matrix. For fresh candidates,

$$v_A = a + \frac{b}{D}, \qquad v_B = b + \frac{a}{D + qa/(1+b)} \geq b.$$

At $c = 2$ and $h = 2$, $a = 1/17$ and $b = 16/17$, so $v_B - v_A \geq 29/51 > 0$. Induction forces all L singleton labels before any group observation, independently of the fixed tie order.

The balanced prefix. At $c = 1$, after p distinct A groups and q singleton purchases, fresh variances are

$$v_A = \frac{1}{2} + \frac{1}{2(D + p/3)}, \qquad v_B = \frac{1}{2} + \frac{1}{2(D + q/3)}.$$

Both exceed $1/2$. A previously observed A group instead has variance at most $(2/3)/(1 + 2/3) = 2/5$: even history and its own first observation give that bound, and further conditioning only reduces variance. The rule avoids group repeats and selects the less-covered family, keeping $|p - q| \leq 1$. All m groups are covered by step $2m$. At that step $p = q = m$, including when group coverage first occurs at $2m - 1$: a fresh singleton then outranks every group repeat. These arguments apply to every candidate tie order.

Exact risks under the common learner. The two families involve disjoint latent coefficients and may be conditioned separately. Consider a common coefficient with posterior variance $1/D$ after history and r private coefficients observed once with coefficient $1/\sqrt{2}$ and unit noise. Each private-coordinate elimination adds $1/3$ to the common precision. An unobserved member's mean variance is therefore

$$U_h(r) = \frac{1}{2} + \frac{3}{2(r+H)}.$$

For an observed member, conditioning its private coefficient on the common one gives residual variance $1/3$ and common coefficient $\sqrt{2}/3$. Its mean variance is

$$S_h(r) = \frac{1}{3} + \frac{2}{3(r+H)}.$$

The raw prefix leaves all group means at variance $U_h(0)$. Averaging the singleton risks, with q observed and $L - q$ unobserved, gives

$$R_{\mathrm{raw}}(q) = \frac{U_h(0)}{2} + \frac{qS_h(q) + (L-q)U_h(q)}{2L} = \frac{1}{2} + \frac{1}{4D} - \frac{q}{12L} + \frac{9L - 5q}{12L(q+H)}, \qquad 0 \le q \le L.$$

For $h = 2$, posterior risk decreases under conditioning, while $R_{\mathrm{raw}}(L) = 1/2 + 1/[3(L+9)] > 1/2$. No learner can attain risk $1/2$ from any prefix of these L purchases. At balanced step $2m$, every group is observed once and m of the L singletons are observed, giving

$$R_{\mathrm{bal}}(2m) = \frac{S_h(m)}{2} + \frac{mS_h(m) + (L-m)U_h(m)}{2L} = \frac{5}{12} + \frac{12m - H - 5}{12m(m+H)}.$$

At $h = 2$, this is below $1/2$, since $m^2 - 3m + 14 > 0$. The target is attainable by either arm: after querying the entire pool, every candidate's variance is at most $2/5$ by conditioning on history and its own observation. Counting all four history labels gives the stated ratio greater than $(L+4)/(2m+4) = \Omega(\sqrt{N})$.

Information in every raw purchase. Before observing a fresh B_j, its private β_j retains variance one. Write $\mathcal{H}_q$ for history and the first q raw purchases. Its true conditional mean variance is at least $1/2$, so

$$I(\theta; Y_j^B \mid \mathcal{H}_q) = \tfrac{1}{2}\log\big(1 + \mathrm{Var}(\mu_j^B \mid \mathcal{H}_q)\big) \ge \tfrac{1}{2}\log(3/2) > 0.$$

This is true teacher information, not information measured by the altered acquisition proxy. The private coefficient also affects an input with positive target mass. A singleton label further informs the common α_0; the exact risk calculation above includes that benefit. The cost difference arises from how these genuine information gains are distributed across the target pool.

The construction uses Gaussian conjugacy, shared group means, and dimensions growing with m. It establishes a coordinate-induced cost separation, with unit-normalized inverse-Bures and tangent-Frobenius Grams coinciding here. Finite checks build the actual matrices and rank-one teachers, differentiate their outputs, and compare direct latent-precision and Gram conditioning. For $h = 2$, $m \in \{2, 3, 4, 8\}$ and three input orders, all 24 paths satisfy the formulas; teacher-realization and reference-covariance errors are below 5×10^{-16} in double precision. The finite-check record retains all paths and their exact posterior risks.

Relation to greedy pivoting. Chen et al. (2025, Theorem C.2) use a block-diagonal matrix with independent coordinates and repeated correlated blocks to show slow Nyström trace-error reduction under diagonal pivoting. The shared-versus-private pattern is the same. Our construction realizes the two acquisition covariances as unit-normalized derivatives of one rank-one LoRA update, with a strict post-history ordering under a fixed scalar refactorization. The endpoints are posterior prediction risk under a common noisy teacher and labels paid by both paths, including their shared history. The following extension additionally constrains proxy error after the shared history and every further common subset while retaining a target interval and a matched-budget risk gap independent of that error tolerance. These constraints distinguish the learning-cost statement from the matrix-approximation counterexample.

A.2 Proof of Theorem 2: accurate covariance and predictive label cost

One label can reveal a shared component that improves prediction on many inputs. The construction below makes that label almost indistinguishable from a private observation by information gain, yet very different by prediction risk. A single common history observation makes the two factorizations choose opposite sides of this distinction.

Construction. Fix $m \ge 4$, $0 < \delta \le 1/4$, and put

$$N = 2m, \qquad \rho = 3/4, \qquad \epsilon = 1/4, \qquad t = \delta\sqrt{\rho}/2, \qquad p = 1 - t^2.$$

Independent standard Gaussian coefficients give latent means

$$\mu_H = h, \qquad \mu_i^A = \sqrt{\rho}\,\alpha + \sqrt{\epsilon}\,\gamma_i, \qquad \mu_j^B = th + \sqrt{p}\,\beta_j, \quad 1 \le i, j \le m.$$

Every label adds independent unit Gaussian noise. The label of H is the one common paid history observation. The $2m$ distinct candidates form the uniform prediction target distribution. All risks average over the Gaussian teacher and all observed labels.

Use orthonormal left directions $e_L, u_0, u_1, \ldots, u_{2m}$ and right directions $e_R, v_0, v_1, \ldots, v_{2m}$. Define $X_i = \ell_i e_R^\top + e_L r_i^\top$, where

$$\begin{aligned} \ell_H &= u_0/\sqrt{2}, & r_H &= -v_0/\sqrt{2}, \\ \ell_i^A &= \sqrt{\rho/2}\, u_0 + \sqrt{\epsilon/2}\, u_i, & r_i^A &= \sqrt{\rho/2}\, v_0 + \sqrt{\epsilon/2}\, v_i, \\ \ell_j^B &= t u_0/\sqrt{2} + \sqrt{p/2}\, u_{m+j}, & r_j^B &= -t v_0/\sqrt{2} + \sqrt{p/2}\, v_{m+j}. \end{aligned}$$

At $(b, a) = (e_L, e_R)$, the predictor $b^\top X_i a$ is zero and its factor gradient is (ℓ_i, r_i), with each side's squared norm $1/2$. Add independent standard Gaussian coefficients to both factors along the displayed u and v directions. The tangent support of X_i eliminates products between these perturbations. The resulting rank-one teacher has exactly the latent means above: the normalized sum and difference of the two common coefficients give α and h; the private sums give γ_i and β_j.

Let L and R collect the left and right gradient rows, including H. For $c = ((1+\delta)/(1-\delta))^{1/4}$, replacing the factors with $(ce_L, e_R/c)$ preserves the predictor. Row normalization gives the full prior Grams

$$C^{\text{full}} = LL^\top + RR^\top, \qquad K^{\text{full}} = (1-\delta)LL^\top + (1+\delta)RR^\top.$$

The balanced Gram is the teacher covariance, and $(1-\delta)C^{\text{full}} \preceq K^{\text{full}} \preceq (1+\delta)C^{\text{full}}$. This comparison includes the history label.

Conditioning preserves accuracy. Write $J = J_m$ and $M = \epsilon I_m + \rho J$. Conditioning on the noisy history observation yields

$$C = \begin{pmatrix} M & 0 \\ 0 & pI_m + (t^2/2)J \end{pmatrix}, \qquad K = \begin{pmatrix} M - 2t^2 J & -t^2 J \\ -t^2 J & pI_m + (t^2/2)J \end{pmatrix}.$$

The true covariances of H with A and B are 0 and t; the raw ones are $-\delta\sqrt{\rho}$ and t. Both give H unit variance, so the conditioning subtraction is half the corresponding outer product.

For a coordinate-selection matrix E_S, noisy conditioning satisfies

$$z^\top F_S(C) z = \min_w \{(z - E_S w)^\top C (z - E_S w) + \|w\|^2\}.$$

Monotonicity and positive scaling of this variational form preserve the relative sandwich. Thus, after every common subset S of distinct candidate labels,

$$(1-\delta)C_S \preceq K_S \preceq (1+\delta)C_S.$$

Moreover $C \succeq I/4$ and $E_S E_S^\top \preceq I$, giving $C_S = (C^{-1} + E_S E_S^\top)^{-1} \succeq I/5$.

Strict paths and globally optimal information. For any conditioning set T, an unselected A has raw variance at most $1 - 2t^2$. An unselected B retains private variance $p = 1 - t^2$, because that private coefficient occurs in no other observation. Hence raw acquisition buys every B before any A, with a strict family margin at least t^2. The same inequality applies when exchanging an A in a batch for an unused B: its conditional log-determinant gain strictly increases. Every raw prefix therefore globally maximizes proxy batch information at its budget. Above budget m, every maximizer contains all B, and the remaining A choices have equal information by exchangeability.

Under C, the first A has variance one, exceeding B's variance $1 - t^2/2$. After one A, another A has variance at most $1 - \rho^2/2 = 23/32$, below each fresh B's private variance $p \geq 253/256$. Balanced consequently selects one A, all B, then the remaining A, regardless of within-family tie order. This path is globally true-information optimal at every budget. A set with no A improves by exchanging a B for the first A. A set with two A and an unused B improves by exchanging the second A for B, conditional on the rest of the set. These exchanges give exactly the balanced counts.

Fixed score tolerance. Suppose each arm selects the maximum reported score, with $|\widehat{v}_i(S) - v_i(S)| \leq \eta$ for every unselected candidate at each actual prefix. Raw's between-family comparisons then retain a gap of at least $t^2 - 2\eta$; balanced's first comparison retains at least $t^2/2 - 2\eta$. Thus $\eta < t^2/4 = 3\delta^2/64$ preserves raw's all-B-first behavior and balanced's first A purchase. Within-family order does not affect the risks, so the label-cost bounds below still hold. At fixed δ, this score tolerance is positive and independent of N.

Information accuracy at every budget. Write $I(S) = \frac{1}{2}\log\det(I + C_{SS})$ for teacher information conditional on the paid history. The information from q B labels is

$$I_B(q) = \frac{q}{2}\log(1+p) + \frac{1}{2}\log\left(1 + \frac{qt^2}{2(1+p)}\right).$$

For $1 \le k \le m$, the optimal-to-raw deficit is $\frac{1}{2}\log 2 - [I_B(k) - I_B(k-1)]$; it is zero for $k = 0$ and $k \ge m+1$. Consequently

$$0 \le \max_{0\le k\le N}[I_{\mathrm{opt}}(k) - I_{\mathrm{raw}}(k)] \le \frac{1}{2}\log\frac{2}{2-t^2}.$$

Each raw B label supplies at least $\frac{1}{2}\log(1+p) \ge \frac{1}{2}\log(509/256)$ nats. The private components also make every other purchase informative.

Prediction risk and paid-label cost. After one A purchase, A's average posterior variance is $23/32 - 7/(32m)$, while B's remains $1 - t^2/2$. Thus

$$R_{\mathrm{bal}}(1) = \frac{55}{64} - \frac{7}{64m} - \frac{t^2}{4} \le \frac{55}{64}.$$

After $q \le m$ raw purchases, A's variance remains one. Unpurchased B targets retain private variance p; purchased B targets retain at least $p/(1+p)$, even if the common coefficient were revealed. For $q \le m/4$,

$$\begin{aligned} R_{\mathrm{raw}}(q) &\ge \frac{1}{2} + \frac{1}{2}\left[p - \frac{q}{m}\frac{p^2}{1+p}\right] \\ &\ge \frac{477}{512} > \frac{29}{32}, \end{aligned}$$

using $p \ge 253/256$ and $p^2/(1+p) \le 1/2$. The initial risk is $1 - t^2/4 > 29/32$. Every target $\tau \in [7/8, 29/32]$ therefore requires exactly one new balanced label and more than $m/4$ new raw labels. Including the common history, the cost ratio exceeds $(1 + m/4)/2 > N/16$. At one new label the risk gap is at least $37/512 > 1/16$. All targets are attainable: observing each candidate's own noisy label reduces its latent variance to at most $1/2$.

A joint vanishing-error sequence. Take $\delta = m^{-2}$. The largest true covariance eigenvalue after history is $\Lambda = (3m+1)/4$. Since $C_S \preceq C$, the relative sandwich gives

$$\sup_S \|K_S - C_S\|_{\mathrm{op}} \le \delta\Lambda < \frac{1}{m}, \qquad \max_k [I_{\mathrm{opt}}(k) - I_{\mathrm{raw}}(k)] \le \frac{1}{16m^4}.$$

For the second bound, use $t^2 = 3/(16m^4)$ and $-\frac{1}{2}\log(1 - t^2/2) \le t^2/(4 - 2t^2) \le t^2/3$. The paid cost ratio exceeds $m/8 \to \infty$, while the one-label risk gap exceeds $1/16$. History remains one label throughout this sequence. The spectrum and dimension grow; the proved raw ordering margin $t^2 = 3/(16m^4)$ shrinks. These are exact-arithmetic path statements.

Learning with each acquisition kernel. For any fixed purchase set S, let $R_*(S)$ be true Bayes risk and $R_K(S)$ the true risk of the raw-kernel posterior-mean predictor, with both learners conditioning on all history and purchase labels. Then

$$R_*(S) \le R_K(S) \le \frac{R_*(S)}{1-\delta^2}.$$

The full-prior sandwich, including history, persists after adding the identical observation noise and taking joint targets and observations. The classical Kantorovich efficiency argument then applies (Styan, 1983, Theorem 1, Eq. (20)). Explicitly, whiten the true joint covariance. For a scalar target, prediction residuals form an affine space. Let r_0 minimize Euclidean squared norm and r_1 minimize the assumed squared norm with relative metric T, whose spectrum lies in $[1-\delta, 1+\delta]$. Stationarity gives $\langle r_1 - r_0, Tr_1\rangle = 0$, so

$$\frac{\|r_1\|^2}{\|r_0\|^2} \le \frac{\|r_1\|^2\|Tr_1\|^2}{\langle r_1, Tr_1\rangle^2} \le \frac{1}{1-\delta^2}.$$

Sum over targets. The bound averages over all labels, including history. Since variance-only acquisition is label-independent, it applies to each selected prefix. Bayes optimality preserves raw's cost lower bound and risk gap. Whole-pool raw risk is at most $1/[2(1-\delta^2)] \le 8/15 < 7/8$, preserving target attainability.

A predictive objective selects the shared label. Replace pointwise variance by the one-step decrease in pool-wide risk, $\|K[:,i]\|^2/[N(1+K_{ii})]$, or use C for the true-covariance rule. Both versions select A first and attain every stated target after one new label under the common Bayes evaluator. To verify the proxy case, put $u = \rho - 2t^2 \geq 93/128$. Its A score is at least $(m-1)u^2/(4m) \geq 3(93/128)^2/16 > 0.098$, whereas its B score is at most $1/(3m) + 5t^4/12 < 0.084$. Under C, the A lower bound improves to $3\rho^2/16$ and the B upper bound remains valid. Thus the objective determines which predictive improvements receive credit, even when either covariance is accurate.

Shared correlations set the scale of the separation. For a general Gaussian pool with unit observation noise, let C_0 be the latent covariance after n_0 paid history labels, $\Lambda = \lambda_{\max}(C_0)$, and $\Delta = R(\varnothing) - \tau > 0$. If the whole pool attains τ, each purchase reduces uniform-pool risk by at most $\Lambda^2/[N(1+\Lambda)]$, because $C_S^2 \preceq \Lambda C_S$. The best set therefore needs at least $N\Delta(1+\Lambda)/\Lambda^2$ new labels, while every complete nonrepeating order uses at most N. Its history-inclusive cost ratio to the best set is at most

$$\max\left\{1, \frac{\Lambda^2}{(1+\Lambda)\Delta}\right\}.$$

In the present construction, $\Lambda = (3m+1)/4 \leq 13m/16$ and $\Delta > 1/16$. The ratio is greater than $2\Lambda/13$ and less than 16Λ, matching the general linear spectral order. The positive private variance keeps every target nondegenerate, while the shared component lets one label improve many predictions.

A bounded-reach counterpart on a fixed task. After common paid history, let $C \succeq 0$ be a fixed Gaussian covariance, with unit observation noise and no repeated query IDs. More generally than uniform-pool risk, take $R(S) = \operatorname{tr}(WC_S)$ for fixed $W \succeq 0$. Let $d_i = (C_S)_{ii}$, $u_i = (C_S W C_S)_{ii}$ and $\rho_i = u_i/d_i$ for $d_i > 0$. Assume $(1-\delta)C_S \preceq K_S \preceq (1+\delta)C_S$, $0 \leq \delta < 1$, at each subset visited by proxy-maximum-variance acquisition. Suppose $d_i \leq v$ and $0 < \rho_{\min} \leq \rho_{\max} \leq H\rho_{\min}$ hold over unqueried IDs at every prefix before step t. Zero-variance columns have zero gain; positive-variance points with zero reach invalidate the positive minimum. Uniform-pool risk uses $W = I/N$; its common $1/N$ cancels in the reach ratio.

Fix $|O| \leq k$, $k \geq 1$, put $B = O \setminus S$, and write $D = \max_{i \notin S} d_i$. Calibration ensures the selected a has $d_a \geq \alpha D$, where $\alpha = (1-\delta)/(1+\delta)$. The true batch and one-point gains satisfy

$$J = \operatorname{tr}\big(WC_S[:,B](I + C_S[B,B])^{-1}C_S[B,:]\big) \leq \sum_{i \in B} u_i \leq k\rho_{\max} D,$$

$$g_a = \frac{u_a}{1+d_a} \geq \frac{\rho_{\min}\alpha D}{1+\alpha D} \geq \frac{\alpha J}{kH(1+\alpha v)}.$$

The first inequality uses $(I + C_S[B,B])^{-1} \preceq I$; it does not assume submodularity. Since $S \cup O$ contains O, $J \geq R(S) - R(O)$. Put $\xi = \alpha/[H(1+\alpha v)]$. Every positive gap contracts by $1 - \xi/k$, while a nonpositive gap stays nonpositive. Iteration gives

$$(R(S_t) - R(O))_+ \leq (1 - \xi/k)^t [R(\varnothing) - R(O)]$$

for arbitrary, possibly disjoint comparator sets. If all remaining variances vanish, the gap is already nonpositive. For $0 < \epsilon < R(\varnothing) - R(O)$, $\lceil (k/\xi)\log((R(\varnothing) - R(O))/\epsilon)\rceil$ steps suffice to reach $R(O) + \epsilon$, or the full pool does so if exhausted first. This is an instance of greedy gap contraction (Chamon & Ribeiro, 2017), with reach relating a variance selector to risk improvement.

For a concrete path-changing case, take two 20-ID blocks with covariances $\frac{1}{2}uu^\top$ and $\frac{9}{20}uu^\top$, where $u = (1,\ldots,1)^\top$, and $W = I/40$. Multiplying these blocks by 0.9 and 1.1 gives a $\delta = 0.1$ proxy whose first block choice strictly reverses the true variance rule. Its first eight steps have $H \leq 29/18$ and $v = 1/2$. The bound gives $R(S_8) \leq R(O_2) + 0.032$, before pool exhaustion, where O_2 buys one point per block and is optimal at budget two. Larger equal-size blocks leave this fixed-prefix bound unchanged. The condition limits risk slack, not exact-target label-cost ratios; applying it to neural learning would additionally require covariance calibration and a link to the learner's risk.

A.3 Acquisition geometry and implementation

Lemma 1 (metric transport on admissible gradients). For a positive-definite metric $M \succ 0$ and quadratic gradient distance $d_M^2(i,j) = (g_i - g_j)^\top M(g_i - g_j)$, the transported metric $M' = JMJ^\top$

preserves all distances under the coordinate map. Let $\mathcal{S}$ be the linear span of admissible gradient displacements and set $K = J^{-1}MJ^{-\top}$. Keeping the same numerical M preserves every distance ordering on $\mathcal{S}$ if and only if, for some $\alpha > 0$,

$$x^\top Ky = \alpha x^\top My \qquad \text{for every } x, y \in \mathcal{S}.$$

Otherwise a two-branch reversal exists within $\mathcal{S}$. For the full covector space this reduces to $K = \alpha M$. For a LoRA block, admissible factor-gradient displacements lie in $\mathcal{S}_{B,A} = \{(GA^\top, B^\top G) : G \text{ has shape } d_{\text{out}} \times d_{\text{in}}\}$. This is a universal characterization over $\mathcal{S}$; a particular finite pool may preserve its orderings even when the two forms are not proportional on that space.

Consider a locally invertible parameter chart $\theta' = \psi(\theta)$ with $J = \partial\theta'/\partial\theta$. The chain rule gives $g' = J^{-\top}g$. Therefore, for $M' = JMJ^\top$,

$$(g'_i - g'_j)^\top M'(g'_i - g'_j) = (g_i - g_j)^\top J^{-1}JMJ^\top J^{-\top}(g_i - g_j) = (g_i - g_j)^\top M(g_i - g_j).$$

Assume $M \succ 0$ on the factor-gradient space and, for a fixed numerical M, write $K = J^{-1}MJ^{-\top}$. Let the columns of L be a basis for the admissible displacement subspace $\mathcal{S}$. The restricted quadratic forms are $M_{\mathcal{S}} = L^\top ML$ and $K_{\mathcal{S}} = L^\top KL$, both positive definite. If $K_{\mathcal{S}} = \alpha M_{\mathcal{S}}$, all admissible distances share one positive scale and every ordering is preserved. Conversely, whiten the restricted forms with $R = M_{\mathcal{S}}^{-1/2}K_{\mathcal{S}}M_{\mathcal{S}}^{-1/2}$. If R is not scalar, choose unit eigenvectors u, v with eigenvalues $\lambda_{\max} > \lambda_{\min}$ and any $1 < t < \lambda_{\max}/\lambda_{\min}$. The admissible displacements $x = LM_{\mathcal{S}}^{-1/2}u$ and $y = \sqrt{t}LM_{\mathcal{S}}^{-1/2}v$ satisfy $x^\top Mx < y^\top My$ but $x^\top Kx > y^\top Ky$. Thus every ordering on $\mathcal{S}$ is preserved exactly when the restrictions are proportional. Taking $L = I$ gives the full-space condition $K = \alpha M$ and proves Lemma 1.

For $\Delta W = BA$ and example-level effective gradient $G_i = \partial\ell_i/\partial(BA)$, matrix differentiation gives $g_{B,i} = G_iA^\top$ and $g_{A,i} = B^\top G_i$. Substituting $B' = BQ$ and $A' = Q^{-1}A$ gives $g'_{B,i} = g_{B,i}Q^{-\top}$ and $g'_{A,i} = Q^\top g_{A,i}$. Hence all realizable factor-gradient displacements belong to $\mathcal{S}_{B,A} = \{(GA^\top, B^\top G) : G \text{ has shape } d_{\text{out}} \times d_{\text{in}}\}$; applying the restricted statement above gives the exact condition for this LoRA block. When Q is orthogonal, both Frobenius inner products are preserved. For $Q = cI$, the squared distance is

$$d_c(i,j) = c^{-2}d_B(i,j) + c^2 d_A(i,j).$$

For the scalar counterexample in Section 2, consider two competing branches p and q and write $\Delta_B = d_B(p) - d_B(q)$ and $\Delta_A = d_A(p) - d_A(q)$. If $\Delta_B\Delta_A < 0$, the sign of $c^{-2}\Delta_B + c^2\Delta_A$ changes at $c^4 = -\Delta_B/\Delta_A$, although $\kappa(Q) = \kappa(cI) = 1$. On the full vectorized factor space, however, the Jacobian is $J = \text{diag}(cI_B, c^{-1}I_A)$ and $\kappa_2(J) = \max(c^2, c^{-2})$. This counterexample therefore excludes guarantees based only on $\kappa(Q)$; it does not call the full reparameterization unit-conditioned. Normalizing the concatenated vector divides the same chart-weighted inner products by chart-weighted norms, so it is not invariant in general.

A.3.1 Intrinsic balanced LoRA acquisition

We use the balanced factor map and inverse-Bures operator of Castin et al. (2026) but give them a different endpoint: a per-example acquisition kernel and its induced sequential decision path. Let $\Delta W = BA = U\Sigma V^\top$ be a reduced SVD over its nonzero singular directions and set $\bar{B} = U\Sigma^{1/2}$ and $\bar{A} = \Sigma^{1/2}V^\top$. When $\text{rank}(BA) = r$, the input factors have full inner rank and an invertible R satisfies $\bar{B} = BR$ and $\bar{A} = R^{-1}A$. It can be computed without materializing BA: thin factorizations $B = U_BR_B$ and $A^\top = V_AR_A$ reduce the SVD to the $r \times r$ core $R_BR_A^\top$, and one valid chart is $R = R_B^{-1}U_c\Sigma^{1/2}$.

To prove Proposition 2, consider a second input factorization $(BQ, Q^{-1}A)$ and independently balanced factors. Two reduced-SVD balanced factorizations of the same matrix differ by a common orthogonal O, including within repeated-singular-value subspaces. If R' maps the second input pair to its balanced pair, the factor identities give $QR' = RO$. Starting from the dual action yields

$$(g_BQ^{-\top})R'^{-\top} = g_BR^{-\top}O, \qquad R'^\top(Q^\top g_A) = O^\top R^\top g_A.$$

Right or left multiplication by O preserves Frobenius inner products. Summing the two factor blocks and all adapter modules therefore preserves the complete Gram matrix. Norms and squared distances

follow, as does every fixed-stream branch based on them. Unit normalization uses only the invariant norms. The QR and small-SVD computation costs $O((d_{\text{out}} + d_{\text{in}})r^2 + r^3)$ per full-rank block.

To prove the symmetric-split characterization, write any reduced factorization as $\widetilde{B} = UX$ and $\widetilde{A} = YV^\top$, where $XY = \Sigma$. Balance gives $X^\top X = YY^\top = X^{-1}\Sigma^2 X^{-\top}$. Multiplying on the left by X and on the right by $X^\top$ yields $(XX^\top)^2 = \Sigma^2$. The positive semidefinite square root is unique, so $XX^\top = \Sigma$ and $X = \Sigma^{1/2}O$ for an orthogonal O; consequently $Y = O^\top\Sigma^{1/2}$. The residual O preserves the Euclidean factor-gradient Gram, proving uniqueness of K_{bal} under the balanced left–right symmetry requirement.

For N candidates, forming the full Gram costs $O(N^2(d_{\text{out}} + d_{\text{in}})r)$ after transforming the per-example gradients. Sequential acquisition does not require that materialization. Let S be its current prefix and $A_S = \lambda I + K_{SS}$. The score vector is

$$\lambda^{-1}\big(\operatorname{diag}(K) - \operatorname{diag}\big(K_{:S}A_S^{-1}K_{S:}\big)\big)\,,$$

so only $\operatorname{diag}(K)$ and $K_{:S}$ are needed. After choosing j, the new column $K_{:,j}$ is evaluated from the factor-free expression in candidate blocks and appended to $K_{:S}$. This gives exactly the full-Gram scores and tie decisions. Initialize $S = \varnothing$ and $d = \operatorname{diag}(K)$, appending conditioning-history columns before new purchases. For the prefix before appending j, solve $A_S a = K_{Sj}$ once, form the residual column $r = K_{:,j} - K_{:S}a$, and update $d_i \leftarrow d_i - r_i^2/(\lambda + r_j)$. The next choice maximizes the updated d_i over remaining candidates. For final prefix length q, kernel working memory is $O(Nq + q^2)$; evaluating the q columns costs $O(Nq(d_{\text{out}} + d_{\text{in}})r)$. The released column-provider implementation solves each prefix system directly, costing $O(Nq^2 + q^4)$ beyond kernel evaluation, and is tested against full-Gram selection. Maintaining a bordered Cholesky factor would reduce that additional cost to $O(Nq^2 + q^3)$. A fixed coordinate sketch is cheaper but changes the kernel and does not inherit the exact guarantee.

The balanced kernel has a factor-free form that removes the mathematical full-rank condition. Define $P_L = (\Delta W \Delta W^\top)^{1/2} = U\Sigma U^\top$ and $P_R = (\Delta W^\top \Delta W)^{1/2} = V\Sigma V^\top$. Cyclicity of the trace gives

$$\begin{aligned} K_{\text{bal}}(i,j) &= \langle G_i\bar{A}^\top, G_j\bar{A}^\top\rangle_F + \langle \bar{B}^\top G_i, \bar{B}^\top G_j\rangle_F \\ &= \langle G_i, P_L G_j + G_j P_R\rangle_F\,. \end{aligned}$$

This expression depends only on ΔW and G_i, which proves gauge invariance directly. More strongly, it has the same value for any two factorizations of the same effective update, even when rank-deficient factorizations lie in different $GL(r)$ orbits. It remains defined when some singular values vanish and is the Gram matrix of the reduced balanced factors over the nonzero directions. Equivalently, with the operator of Castin et al. (2026), $K_{\text{bal}}(i,j) = \langle G_i, H_{\Delta W}[G_j]\rangle_F$; the new step here is its use as a per-example kernel and the path-level acquisition result. It is positive semidefinite because the first line is an explicit feature inner product. The original CPU tests compare the feature and principal-square-root forms, cover repeated singular values and dense gauges, and include exact rank-deficient and zero-update cases. The full-rank implementation used in the Qwen runs rejects a block when $\sigma_{\min} \le \epsilon_{64}\max(d_{\text{out}}, d_{\text{in}}, r)\max(\sigma_{\max}, 1)$, where ϵ_{64} is machine precision. It does not truncate the rank and continue. The rank-reduced implementation below recovers features on the effective update's retained support without forming a dense update or effective gradient.

Numerical ranks in the Qwen runs. Auditing all stored QR-core spectra across six fits, four rounds and both inverse-Bures trajectories (`balanced_c1`, `balanced_c2`) finds no rank rejection or admission disagreement (Table 4). Every block retains rank eight, at least 1.49×10^9 times above the original cutoff. These summaries use the original full-rank cutoff; the later rank-reduced policy below is separate. The states are taken immediately before acquisition: after warm training for round one, and after the preceding round's continuation thereafter. These two trajectories retain corresponding effective updates throughout; the diverging raw-acquisition arms are not the pair audited here.

Recovery after rank loss. Let $BA = U\Sigma V^\top$ have rank k, possibly smaller than the stored inner dimension r. Since $\operatorname{col}(BA) \subseteq \operatorname{col}(B)$ and $\operatorname{row}(BA) \subseteq \operatorname{row}(A)$, the systems

$$BL = U\Sigma^{1/2}, \qquad A^\top R = V\Sigma^{1/2}$$

Table 4: Production numerical-rank margins immediately before acquisition in the two inverse-Bures trajectories, `balanced_c1` and `balanced_c2`. Each AG News fit contributes four rounds $\times$ two corresponding representations $\times$ 56 blocks. These are repeated states, not independent fits. The cutoff τ is the original FP64 threshold above for the QR core $R_B R_A^\top$.

Fit	Rank-8 blocks	Min. σ_8/σ_1	Min. σ_8/τ	Max. $\kappa_2(R)$
Original 1	448/448	0.008	1.91×10^9	10.20
Original 2	448/448	0.005	1.76×10^9	12.07
Original 3	448/448	0.005	1.91×10^9	11.74
Fresh 1	448/448	0.008	2.75×10^9	9.26
Fresh 2	448/448	0.004	1.49×10^9	12.28
Fresh 3	448/448	0.008	1.86×10^9	10.19

have solutions. For compatible factor derivatives $g_{B,i} = G_i A^\top$ and $g_{A,i} = B^\top G_i$, any such solutions give

$$g_{B,i} R = G_i V \Sigma^{1/2}, \qquad L^\top g_{A,i} = \Sigma^{1/2} U^\top G_i.$$

Thus the recorded derivatives determine the reduced balanced features even when dead or cancelling columns prevent an invertible map between stored and balanced factors. The argument also applies to equal products in different $GL(r)$ orbits. It uses the two range inclusions, not covariance of the Moore–Penrose inverse under an arbitrary change of factors.

For a matrix-free implementation, take thin QR factorizations $B = Q_B R_B$, $A^\top = Q_A R_A$ and the reduced SVD $R_B R_A^\top = U_c \Sigma V_c^\top$. Solve $R_B L = U_c \Sigma^{1/2}$ and $R_A R = V_c \Sigma^{1/2}$ in least squares. These solves and the SVD involve matrices of size at most r; factor preprocessing still costs $O((d_{\text{out}} + d_{\text{in}})r^2 + r^3)$. Transforming each example costs $O((d_{\text{out}} + d_{\text{in}})rk)$. This computes the same inverse-Bures geometry of Castin et al. (2026); factor-based recovery of intrinsic derivatives also appears in PRISM (Wang & Zhang, 2026).

The new FP64 implementation uses the cutoff

$$t = 64\epsilon_{64} \max(d_{\text{out}}, d_{\text{in}}, r) \max(\sigma_{\max}, 1),$$

rejects values in $[t/8, 8t]$, and checks the two target-specific solve residuals. It reports the discarded spectrum and map norms: gradient error can be amplified by $\|L\|_2$ or $\|R\|_2$, and removing a positive singular direction changes the kernel. Small absolute truncation error alone does not guarantee a small change after row normalization. The algebraic identity holds on the nonzero support; this numerical policy makes unresolved cases explicit.

Local tests compare recovered Grams with an independent dense-update reference and autograd derivatives, including repeated singular values, rank loss, zero updates, cancelling columns, different factorization orbits, and the previous full-rank implementation. A 216-case sweep varies nine constructions, six seeds and four gauge condition numbers from 1 to 256. It resolves 210 cases, all with matching normalized acquisition paths; six are rejected as numerically unresolved. In those six, the reference factorization is resolved but its transformed version is rejected. Rejection raises an error before acquisition and returns no batch: numerical admission itself need not be factorization invariant. The path identity is algebraic; the numerical checks compare pairs for which recovery succeeds. Maximum-entry discrepancies, divided by the larger matrix entry magnitude with a denominator floor of one, are below 3×10^{-14} for both Grams and normalized Grams. The compact package includes 41 rank-recovery unit tests, including rejection and truncation counterexamples.

PRISM uses a different intrinsic per-example geometry for private optimization (Wang & Zhang, 2026). Writing $\Pi_U = UU^\top$ and $\Pi_V = VV^\top$, its orthogonal tangent projector induces

$$P_T(G) = \Pi_U G + G\Pi_V - \Pi_U G \Pi_V, \qquad K_{\text{PRISM}}(i,j) = \langle G_i, P_T(G_j) \rangle_F.$$

In the singular-vector basis, this tangent-Frobenius kernel assigns unit weight to components in the left or right tangent support and zero weight off both supports, counting their intersection once. By contrast, K_{bal} weights left-only component a by σ_a, right-only component b by σ_b, and an intersection component (a,b) by $\sigma_a + \sigma_b$. PRISM establishes gauge-invariant per-example clipping and noise on the rank-r tangent space; our construction selects the inverse-Bures weighting supplied by Balanced LoRA and propagates its Gram through a sequential acquisition program. Supplying

K_{PRISM} to the same Gram-driven selector likewise makes its fixed-stream path invariant; the two kernels encode different task geometries.

Normalization leaves zero-norm rows at zero; candidate order resolves ties. Both conventions preserve the path identity. At zero update, $P_L = P_R = K_{\mathrm{bal}} = 0$, giving an uninformative all-tie path. Switching to predictive entropy when $\|\Delta W\|_F = 0$ preserves invariance because the switch and both rules depend only on the predictor/effective update. Our experiments instead start with nonzero warm adapters.

A.4 Native training-recipe equivalence across acquisition

Adam admits equivalent LoRA initialization and learning-rate recipes (Schulman & Thinking Machines Lab, 2025). We use this known symmetry to test whether equivalent training runs remain equivalent after selecting their next data. For a scalar $c > 0$, initialize $A' = A/c$, $B' = cB = 0$. In an AdamW optimizer with denominator $\sqrt{\widehat{v}} + \epsilon$, transform its parameter groups as follows, keeping the moment coefficients fixed:

	learning rate	ϵ	weight decay
A'	η_A/c	$c\epsilon_A$	cw_A
B'	$c\eta_B$	ϵ_B/c	w_B/c

For A', the gradients and first moments scale by c, second moments by c^2; the transformed epsilon preserves the normalized Adam direction, and the learning rate gives an update divided by c. The B' argument is symmetric. Each learning-rate–decay product is unchanged. With corresponding minibatches, no coordinate clipping, and matching initial optimizer states, induction preserves $B'_t A'_t = B_t A_t$. This calculation also applies when continuation retains the corresponding moment states.

A.5 All-layer acquisition and continued learning

Model and paired training. Original 1–3 in Section 5.1 use seeds 1701, 2903, and 4307, respectively, and were extended from round two. Fresh 1–3 use seeds 2026090831–2026090833, respectively, and were trained from scratch. Warm/pool blocks are disjoint by ID and exact text; fresh blocks also exclude development/validation data. All share the 1,024-example test set. Query/value LoRA has rank eight, $\alpha = 16$ and 1,605,632 trainable parameters across 28 layers. FP32 eager evaluation uses the coordinated $c = 1, 2$ AdamW recipes in Appendix A.4: 96 warm updates, then 64 updates per 128-label acquisition, batch size eight and reference rate 10^{-4}, retaining optimizer moments. Corresponding states are checked through training; warm logits are bitwise equal and pseudo-label gradients covariant.

Class probabilities and losses. The prompt requests one topic digit: 0 (World), 1 (Sports), 2 (Business), or 3 (Sci/Tech). We use the tokenizer's chat template with thinking disabled and a generation suffix, truncating only the article to keep the prompt within 256 tokens. At the last prompt position, let z_k be the LM-head logit for digit k. All three stages use the four-class distribution $p(k \mid x) = \exp(z_k)/\sum_{j=0}^{3} \exp(z_j)$. Training minimizes $-\log p(y \mid x)$; acquisition differentiates $-\log p(\hat{y} \mid x)$ with the detached pseudo-label $\hat{y} = \arg\max_k p(k \mid x)$. Evaluation averages true-label NLL, computing the four-logit log-softmax in FP64. Training and gradient extraction use FP32.

Acquisition and replay. We apply the normalized posterior-variance rule of Section 2.2, with ridge one, to independently extracted pseudo-label cross-entropy gradients. Conditioning includes the branch's $128(j-1)$ past purchases in round j, but excludes warm data. Warm labels fit the predictor but do not enter the acquisition conditioning matrix. The resulting covariance is a gradient-feature proxy, not the Qwen posterior conditioned on all observed labels. MultiNLI conditions on warm-history gradients (Appendix A.6); it is a complementary protocol, not a same-task warm-history ablation. Gradients are refreshed under the current predictor; purchased points cannot be bought again. Feature, Gram and score arithmetic use FP64-promoted gradients. Selection takes the strict maximum, with frozen candidate order breaking exact ties and no near-tie tolerance. Normalization sets diagonals to one for rows with squared norm above 10^{-15}, zeroing others. All initial

Table 5: First-round purchased-set sensitivity on fixed warm checkpoints. Entries are ranges of replaced examples out of 128. Recipe contrasts cover three ridges and five paired orders per fit; order contrasts compare each of four permutations with the original order at each ridge. Inverse-Bures recipe contrasts are zero in all 15 configurations per fit.

Fit	Raw recipe change	Raw order change	Inverse-Bures order change
Original 1	16–21	3–8	4–10
Original 2	21–32	2–10	5–8
Original 3	20–26	1–9	4–9
Fresh 1	18–25	1–7	2–5
Fresh 2	35–41	3–11	4–9
Fresh 3	13–19	4–10	5–9

rows are valid, so empty history selects candidate zero. Pool labels remain inaccessible during extraction and selection. ID-sorted warm and purchased data feed the common minibatch procedure. Round one starts from equal predictors; later rounds use each branch's updated predictor and history.

Seven trajectories are run in every fit: raw acquisition under both native recipes, inverse-Bures acquisition under both recipes, $c = 2$ forced to consume raw $c = 1$ purchases, uniform random acquisition, and old-label-only continuation. The last two use $c = 1$. The forced arm uses the same IDs, not another independently sampled batch. After the second purchase, an additional raw $c = 1$ continuation in the original study changes only that round's minibatch stream. It calibrates training variation conditional on that purchase and starting state.

First-round configuration sensitivity. We replayed acquisition on the six saved warm checkpoints with ridges $0.1, 1, 10$ and five candidate orders: the original order and four fixed random permutations. Each configuration used the same order for both recipes, an empty history and a budget of 128. Raw recipes purchased different sets throughout this grid; inverse-Bures purchases agreed exactly. Table 5 reports each fit's range. Changing candidate order within either raw recipe replaced fewer examples than changing recipe in every tested fit. Inverse-Bures acquisition also responded to candidate order while remaining equal across the paired factors.

This retrospective check changes selection settings at a fixed predictor. The four-round learning study retains its original ridge and order. Replacement counts compare sets, consistent with sorting training examples by ID. The permutation seeds are 2026090911–2026090914 (PCG64), giving first candidates 108, 272, 149 and 437 instead of candidate zero. Direct selected-system solves verify the two choices at every raw pair's first divergence; the smallest relative winner–runner-up margin is 1.26×10^{-9}. This is the score gap divided by the largest absolute score among available candidates.

Consumption of purchased labels. The sampler traverses successive random permutations of the ID-sorted cumulative dataset and stops after 512 example exposures per round. That dataset contains 384, 512, 640, and 768 examples in rounds one through four. All 128 current purchases enter training in the first two rounds; 96–110 do so in round three and 79–97 in round four, across all six fits and acquisition arms (Table 6). Later purchases thus compete with earlier purchases and warm data for the fixed update budget. Old-only replays each of its 256 warm IDs twice per round.

Purchases also affect warm-example replay. The raw $c = 2$ minus $c = 1$ warm-exposure difference ranges over $[-7, 8]$, $\{0\}$, $[-7, 12]$, and $[-12, 11]$ across the four rounds; the corresponding ℓ_1 distances between per-warm-ID exposure counts are 80–133, 0, 70–90, and 88–111. Even in round two, when each available example is used once, ordering can differ. Both matched pairs retain exactly the same ordered purchases and minibatch streams in all 24 fit–rounds. The measured learning consequence therefore includes the purchase's effect on cumulative replay as well as the newly labeled examples: replay mediates the adaptive-procedure effect, rather than isolating batch value at equal exposure.

Four-round risks and observed label costs. The complete risk grid uses zero to 512 acquired labels in steps of 128, with 256 common warm labels. All seven arms run 64 updates per round; old-only therefore remains a zero-new-label, equal-update reference. We evaluate the process at the five trained checkpoints. The primary normalized area and raw signed contrast are defined in

Table 6: Consumed training data in the four-round experiment. Each entry is the minimum–maximum across all six fits; every row has 512 example exposures. Current and earlier columns give distinct purchased IDs consumed (U) and total exposures (E). The two matched pairs share exact streams and are grouped here.

Arm	Round	Warm E	Current U	Current E	Earlier U	Earlier E
Raw $c = 1$ / forced $c = 2$	1	334–352	128	160–178	0	0
	2	256	128	128	128	128
	3	196–205	100–107	100–107	200–216	200–216
	4	161–180	79–90	79–90	244–265	244–265
Raw $c = 2$	1	340–346	128	166–172	0	0
	2	256	128	128	128	128
	3	198–208	97–110	97–110	197–209	197–209
	4	165–174	80–92	80–92	248–263	248–263
Inverse-Bures $c = 1, 2$	1	335–345	128	167–177	0	0
	2	256	128	128	128	128
	3	200–214	96–109	96–109	201–205	201–205
	4	158–182	80–97	80–97	240–267	240–267
Random	1	332–348	128	164–180	0	0
	2	256	128	128	128	128
	3	200–208	102–107	102–107	200–206	200–206
	4	168–175	81–87	81–87	252–260	252–260
Old only	1	512	0	0	0	0
	2	512	0	0	0	0
	3	512	0	0	0	0
	4	512	0	0	0	0

Section 5.1. The original-three, fresh-three, and all-six means use equal fit weights. Their descriptive 95% percentile intervals resample fits 10,000 times, using seeds 2026090840, 2026090841, and 2026090842 respectively, with linear quantile interpolation. These are central 95% resampling ranges conditional on the shared test rows; with three or six fits, their population coverage is not established. The original R1/R2 outcomes informed the extension. The four-round schedule, six-fit split, signed-area primary contrast, and complete target-attainment analysis were fixed in the execution packet before any extension training began. The fresh cohort is reported separately. Tables 7 and 8 give the final risks and primary area summaries. Both matched controls preserve probabilities and mapped parameter/optimizer states at all 24 fit–round endpoints, with corresponding-state checks through every training update.

As a secondary magnitude summary added after inspecting the four-round curves, we use $M = \sum_{j=0}^{4} w_j |\mathcal{L}_{c=2,j} - \mathcal{L}_{c=1,j}|$, with the same trapezoidal weights $w = (1, 2, 2, 2, 1)/8$. Unlike the primary signed contrast, M measures trajectory separation without cancellation across rounds. Table 7 retains both summaries for every fit. Fresh 3 illustrates why the summaries differ: raw $c = 1$ has lower risk through round three, but $c = 2$ has lower final risk; its signed area favors $c = 1$, while its absolute area records separation in both directions.

The secondary endpoint is the entire first-observed attainment function

$$H_{\text{obs}}(\tau) = \min\{128j : \mathcal{L}_j \le \tau,\ j = 0, \ldots, 4\},$$

with status *not observed* when the set is empty. Partitioning nonnegative targets at every observed risk yields 22 half-open intervals per fit, including the unbounded final interval. Figure 5 shows acquired-label costs (add 256 for total labels) and flags targets old-only reaches at the same or an earlier checkpoint. Above warm risk, all costs are zero. These are observed grid costs under the stated schedule; targets reached by old-only do not establish a need for new labels.

Accuracy and prediction decisions. Original 1's final inverse-Bures and old-only predictors both have 82.5% accuracy, but NLL 0.595 and 1.277: NLL also measures confidence and remains the primary risk.

Table 7: Four-round outcomes. Final risks are mean held-out NLL. The primary signed area Δ measures direction; the secondary absolute area M measures separation without cancellation across rounds. The forced and inverse-Bures recipe pairs coincide exactly.

Fit	Raw $c = 1$	Raw $c = 2$	Inv.-Bures	Random	Old only	Δ	M
Original 1	0.687	0.457	0.595	0.542	1.277	-0.071	0.071
Original 2	0.410	0.441	0.473	0.410	0.641	-0.006	0.014
Original 3	0.449	0.403	0.421	0.431	0.740	-0.010	0.027
Fresh 1	0.402	0.391	0.421	0.430	0.705	-0.070	0.080
Fresh 2	0.448	0.428	0.421	0.446	0.608	-0.067	0.113
Fresh 3	0.517	0.419	0.441	0.483	1.308	0.033	0.058

Table 8: Primary paired area contrast by fit cohort. Ranges are central 95% fit-bootstrap percentiles conditional on the shared test rows, not calibrated population-coverage intervals at these small fit counts.

Cohort	Fits	Mean Δ	95% resampling range
Original	3	-0.029	[-0.071, -0.006]
Fresh	3	-0.035	[-0.070, 0.033]
All	6	-0.032	[-0.060, 0.002]

Controls and training-stream reference. The original three fits supply the same-purchase training-stream reference (Table 9). Their raw arms replace 20/21/23 of 128 purchases in round one, 47/59/65 in round two, and 47/64/65 of the 256 cumulative labels. Round two propagates different learning histories, rather than adding a same-function intervention. All six independently constructed inverse-Bures Gram pairs and paths agree exactly; all 24 column-streamed paths match full-Gram selection. Three first-round transport controls recover the raw reference. Weighting features by relative learning rates also recovers raw $c = 1$ geometry and paths in all 12 replays. This recipe-aware control retains a reference recipe; K_{bal} does not.

Specifically, for reference learning rates η_A, η_B and paired rates η'_A, η'_B, the transported features are

$$\widetilde{g}'_A = (\eta'_A/\eta_A)g'_A, \qquad \widetilde{g}'_B = (\eta'_B/\eta_B)g'_B.$$

The native recipes have $\eta'_A = \eta_A/c$ and $\eta'_B = c\eta_B$, which cancel $g'_A = cg_A$ and $g'_B = g_B/c$. These are feature weights; the corresponding quadratic metric weights are their squares.

Reproduction. The compact code package checks the saved four-round curves and paired controls with `analysis/ag_news.py`, without inference. The package's `README.md` describes the included training components and numerical inputs.

Table 9: Same-purchase training-stream reference. These rows compare an alternate final-round minibatch stream with the original raw $c = 1$ stream, holding its cumulative labels and starting state fixed. Intervals resample test rows conditional on the one available stream pair per fit.

Fit	Δ NLL [95%]	Δ acc. (pp) [95%]	Disagree / 1,024
1	-0.065 $[-0.104, -0.027]$	$+2.441$ $[+0.391, +4.492]$	130
2	-0.023 $[-0.079, +0.030]$	$+1.465$ $[-1.562, +4.492]$	289
3	-0.046 $[-0.071, -0.022]$	$+4.688$ $[+2.734, +6.738]$	127

A.6 MULTINLI PROTOCOL AND ANALYSES

Prospective comparison: development and new instances. The two development fits in Appendix A.8.1 supplied public government-purchase counts for all 81 groupwise scalings. We chose one ordered pair maximizing its minimum signed quota difference across these fits, breaking ties by the original grid order. In group order (early query, early value, late query, late value), the fixed low and high scales are $(1/2, 1/2, 2, 2)$ and $(2, 2, 1, 1/2)$, respectively. The first 14 of 28 layers form the

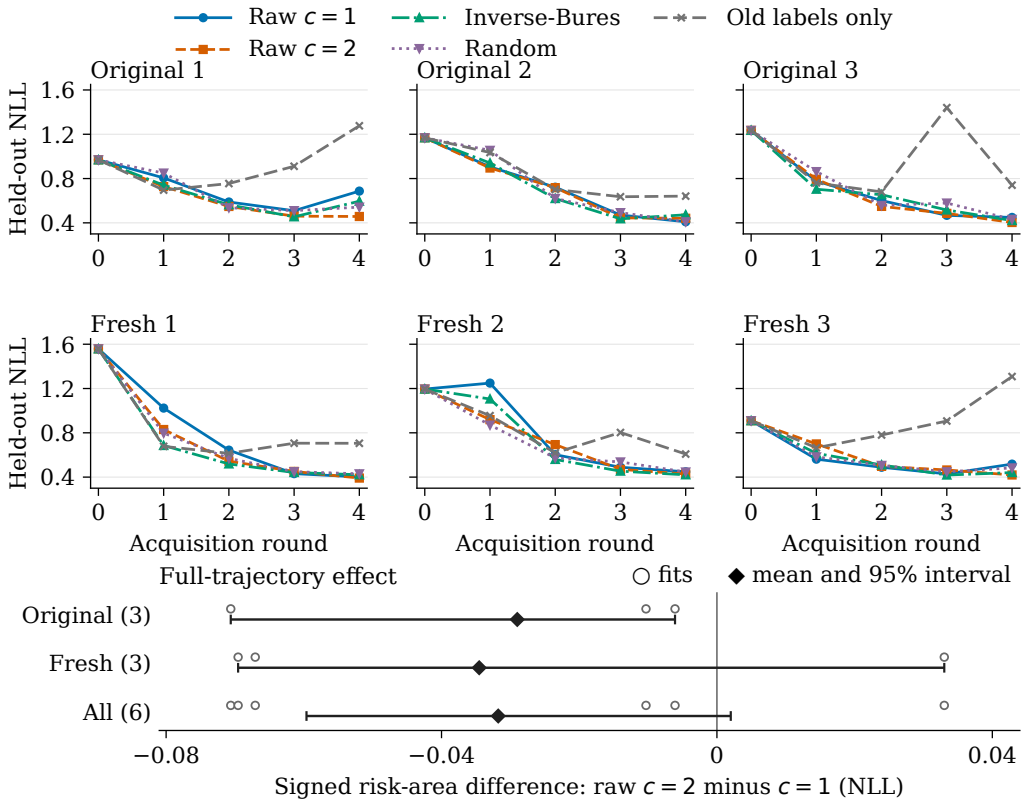


Figure 4: **All acquisition rules improve on warm and old-only risk by round four, with different intermediate trajectories.** Each panel shows one fit's held-out NLL at five checkpoints. Forced-purchase $c = 2$ coincides exactly with raw $c = 1$, and both inverse-Bures recipes coincide exactly. Old-only uses the same update budget without acquiring labels; the other rules add 128 labels each round. The lower panel repeats the primary signed contrasts and descriptive fit-bootstrap intervals for comparison with the paired-difference view in Figure 3.

early half. Every block uses $(B, A) \mapsto (cB, A/c)$. No new fit's labels or learning outcomes enter this choice.

Eight new Qwen3-1.7B fits use rank-8 query/value LoRA, scale $\alpha = 16$, 96 warm updates of batch size eight and AdamW learning rate 10^{-4}. Each has 128 warm, 500 pool and 128 development examples per domain. The common final panel contains 512 examples per domain from previously unused MultiNLI training-source premise groups. Input-hash ordering, premise-group separation and encoded-input exclusion remove overlap between these partitions, fits and the earlier 4,048 inputs. The panel is held out from fitting and acquisition; it is not the official matched validation set. Labels are not balanced by class.

Acquisition uses the full row-normalized factor-gradient Gram, ridge one, warm-history conditioning and a common candidate order. Both fixed charts pass the frozen bitwise task-logit, hidden-state and transported-gradient checks in every fit. Independently balanced and transported controls recover their respective ordered paths. Learning starts only after all eight acquisitions are sealed. Both arms restore the same warm parameters and complete AdamW state, then run 256 updates with four warm and four purchased examples per batch. Each label appears four times; shared IDs occupy matched positions. All 16 paired streams per fit completed, and eight same-purchase duplicates match full final states and logits bitwise.

Estimand and inference. For each fit we average the 16 paired stream contrasts $\Delta_{\mathrm{dom}} = (\mathcal{L}_{\mathrm{low,G}} - \mathcal{L}_{\mathrm{high,G}}) - (\mathcal{L}_{\mathrm{low,T}} - \mathcal{L}_{\mathrm{high,T}})$. The reported mean and two-sided Student-t interval use eight equally weighted fit means and seven degrees of freedom. This model-approximate in-

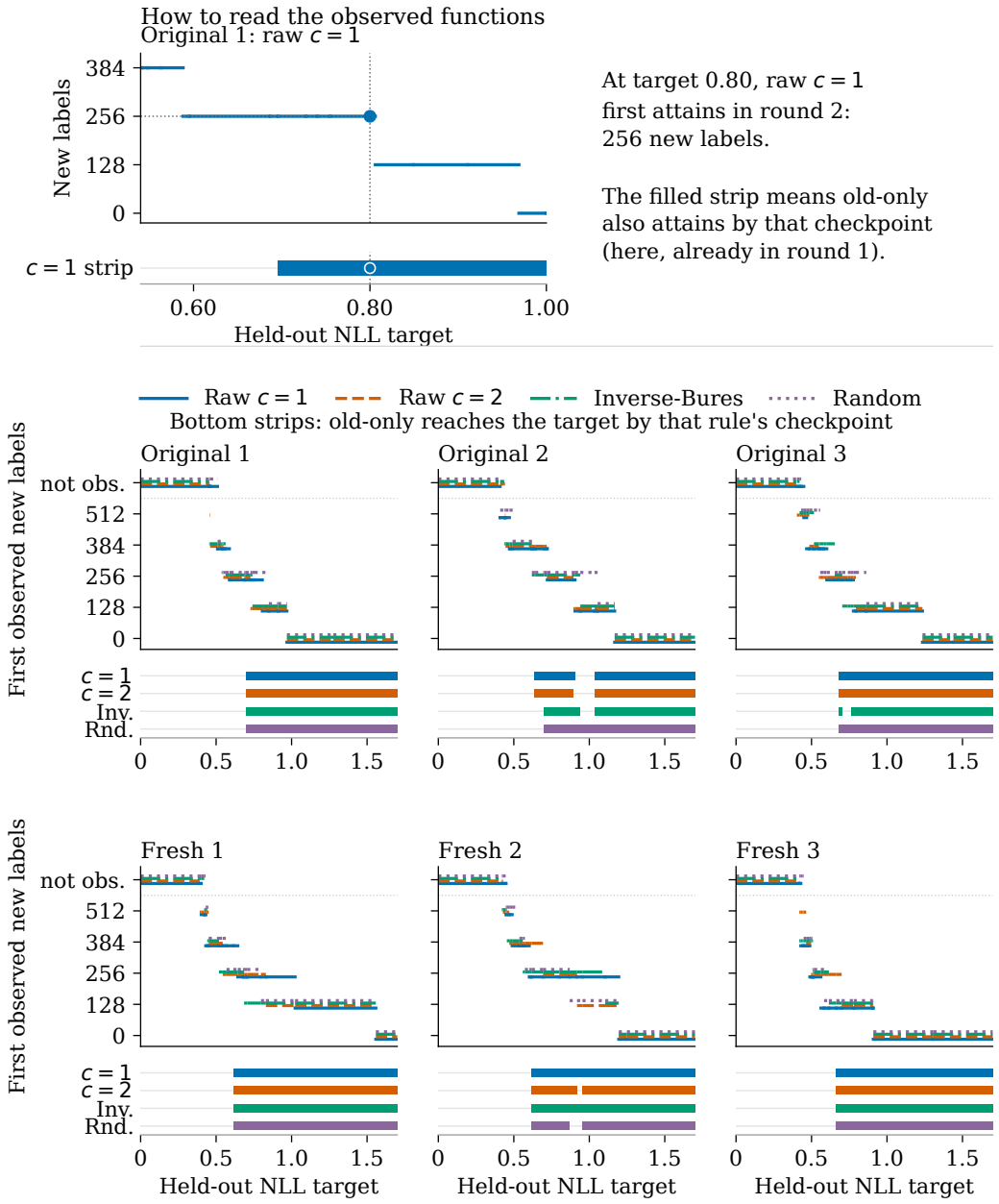


Figure 5: **The observed budget needed to reach a target depends on the factorization and the target.** The upper guide illustrates how to read one observed curve and its old-only strip. The six lower panels show all four first-observed attainment functions as horizontal segments, with slight vertical offsets to separate coincident values. The "not observed" level is categorical, not a numeric cost; above the warm risk all costs are zero. The four strips below each small panel mark targets reached by old-only at the same or an earlier checkpoint than each rule (Inv.: inverse-Bures; Rnd.: random). Add 256 warm labels to finite costs. These are observed grid costs on the executed update schedule.

terval conditions on the common fixed panel; it does not include resampling that panel. The new estimates do not pool development outcomes. For calibration, alternate rows within each domain–class stratum form two folds. For each arm and stream, a positive global temperature in $[0.05, 20]$ is fitted on one fold and scored on the other, then reversed. Both domains share the temperature;

Table 10: Relative-domain NLL contrasts for the fixed refactorization pair on eight new fits. Positive Δ_{dom} means that the high arm's NLL advantage is greater on government than telephone. Two-sided 95% model-approximate intervals use eight fit means (16 paired streams each) and condition on the common test panel.

Measure	Mean Δ_{dom}	95% interval	Positive fits
Raw NLL	0.0175	[0.0012, 0.0339]	6/8
Calibrated NLL	0.0094	[0.0010, 0.0178]	7/8

Table 11: All eight prospective fits under the same fixed rule. Quota is the high-minus-low government count within a 256-label purchase; replaced is 256 minus the purchase-set overlap. Contrasts average the 16 paired streams within each fit.

Fit	Quota shift	Replaced	Raw Δ_{dom}	Calibrated Δ_{dom}
1	13	77	-0.0046	0.0061
2	29	88	0.0047	-0.0010
3	27	89	0.0170	0.0009
4	12	79	0.0244	0.0086
5	26	89	-0.0006	0.0011
6	19	77	0.0350	0.0179
7	14	68	0.0536	0.0284
8	19	91	0.0108	0.0133

argmax predictions are unchanged. All 512 fitted temperatures lie between 1.20 and 3.89, away from the bounds.

The frozen joint criterion required positive raw and calibrated lower bounds, at least seven positive raw fits, and raw mean at least $\log(1.02) = 0.0198$ (a resource screen, not an established minimum important difference). Both lower bounds are positive, but only six raw fits are positive and the mean is 0.0175. The joint criterion is therefore unmet.

Table 12 gives absolute outcomes. Overall high-minus-low NLL is -0.0122 ($[-0.0405, 0.0161]$), or 0.0005 ($[-0.0089, 0.0099]$) after calibration. These intervals and the separate domain intervals include zero; the interaction does not establish an overall advantage. The compact code package retains the domain-comparison statistics alongside the eight-fit response records.

Evaluation-panel sensitivity. A post hoc analysis jointly resamples fits and panel premise representatives, preserving domain–calibration-fold strata and paired streams, and refits temperatures in each replicate. Across 1,999 replicates, raw and calibrated Δ_{dom} have central 95% ranges $[-0.0094, 0.0450]$ and $[-0.0029, 0.0222]$. Both include zero. These sensitivity ranges hold domains, source selection, calibration folds and observed streams fixed; they neither propagate corpus uncertainty nor guarantee population coverage. They do not replace the frozen criterion.

A.6.1 Per-example responses across the eight new fits

We applied the response estimator in Appendix A.8.1 to every endpoint of the eight-fit study: 16 paired streams, both purchase arms and all 1,024 test rows. The response analysis was specified after the domain results were observed; it uses the existing data and does not replace the prospective domain criterion. For correctness, $d_s(i) = a_{\mathrm{high,s}}(i) - a_{\mathrm{low,s}}(i)$, where a is the binary indicator of a correct prediction. Under independent continuation streams conditional on the fitted learner, purchases and panel, U_a estimates $N^{-1}\sum_i(\mathbb{E}_s d_s(i))^2$. Negative estimates are retained; the displayed RMS is $\sqrt{U_a}$ when nonnegative.

Table 13 gives delete-one-stream jackknife standard errors. The eight fits, not their 128 paired streams or the within-fit stream pairs, are the independent fitted units. The mean $U_a = 0.007026$ has an approximate Student-t interval with seven degrees of freedom of $[0.005144, 0.008908]$, conditional on the common panel. Its square root gives the main-text pooled RMS and interval. The

Table 12: Absolute outcomes, averaged equally over all eight fits and 16 streams. G: government; T: telephone. Overall averages weight the two domains equally. NLL is in nats; accuracy is in percent.

Metric	Low G	High G	Low T	High T	Low all	High all
Raw NLL	0.6238	0.6028	0.6930	0.6896	0.6584	0.6462
Calibrated NLL	0.4409	0.4367	0.5230	0.5282	0.4820	0.4824
Accuracy (%)	84.4101	84.3903	79.6936	79.6585	82.0518	82.0244

Table 13: Correctness-response estimates for all eight new fits. Each cell gives $10^3 U_a$ with one delete-one-stream jackknife SE in parentheses, also scaled by 10^3. Centered responses subtract domain–class means within each stream. Same-purchase low/high controls use eight disjoint stream differences; their noise coupling differs from the 16-stream paired purchase response. Negative estimates are retained.

Fit	Purchase	Centered	Same low	Same high
1	5.82 (1.33)	5.66 (1.16)	0.73 (2.02)	3.45 (6.22)
2	6.23 (1.26)	6.07 (1.19)	0.38 (6.52)	3.84 (6.58)
3	4.20 (1.47)	4.35 (1.38)	−2.83 (4.04)	0.73 (7.18)
4	7.98 (1.89)	7.97 (1.58)	4.39 (5.95)	−1.50 (7.01)
5	8.28 (1.92)	7.88 (1.74)	8.61 (8.08)	9.00 (10.81)
6	4.21 (1.52)	4.26 (1.41)	3.14 (4.03)	3.80 (2.66)
7	9.51 (1.38)	9.45 (1.30)	−0.66 (3.45)	2.09 (7.81)
8	9.99 (2.49)	9.82 (1.96)	−1.46 (3.77)	−4.39 (3.17)

signed mean accuracy difference is -0.027 points, with interval $[-0.508, 0.453]$; it measures a different aspect of learning.

We used two fixed splits of the streams. First/last-eight response correlations were 0.34–0.64; even/odd correlations were 0.33–0.66, with positive cross-products in all eight fits under both splits. Groups of examples with positive or negative first-half mean responses retained their respective group-mean signs in the other half, also after reversing the split. These are descriptive checks on the selected groups, not tests that every example retains its sign. Centering each stream within the six domain–class groups left mean $U_a = 0.006933$, or 98.7% of the uncentered estimate. Centered uncertainty uses fixed-panel per-fit jackknife estimates.

The same-purchase controls subtract the first eight streams from the last eight within each arm, giving eight disjoint stream differences. Their low/high mean U_a values are 0.001539/0.002128, with approximate fit-level intervals $[-0.001552, 0.004630]$/$[-0.001221, 0.005477]$. These controls have different noise coupling and fewer difference units than the paired purchase contrast. Fit 5 has control estimates as large as its purchase response, with larger standard errors. We use them as randomness references rather than a matched test of purchase-effect superiority.

NLL and true-class probability responses gave positive U and positive split-half correlations in all eight fits. Their mean squared responses were 0.13554 and 0.004637, with approximate fixed-panel fit intervals $[0.09359, 0.17750]$ and $[0.003341, 0.005934]$. These outcomes share the same predictions and do not count as independent replications.

For panel sensitivity, we drew eight fits with replacement and 512 rows within each domain, sharing each row draw across fits and holding the observed streams fixed. Across 1,999 draws (seed 20260914161), the 95% percentile range for mean U_a was $[0.005238, 0.009014]$; the corresponding NLL and probability ranges were $[0.09693, 0.18345]$ and $[0.003537, 0.005922]$. These are crossed fit/panel sensitivity ranges, not coverage-guaranteed intervals. Unlike the preceding domain-contrast analysis, no calibration is fitted here. All estimates reconstruct from the frozen logits and weights with `analysis/multinli_response.py`.

A.7 Complete consumption of four-round purchase curricula

This control uses every saved raw purchase from all four rounds of the six AG News fits. Both arms start from the same canonical warm parameters and complete AdamW state. Each round consumes

Table 14: Complete-consumption continuations over 16 paired streams per fit. Rounds report mean raw $c = 2$ minus $c = 1$ NLL. Aligned means sign each contrast by its frozen old direction before averaging over rounds. The final two columns give post hoc signed and absolute trajectory areas averaged over streams, using the original five-point trapezoidal weights. Absolute areas include training-stream variability.

Fit	R1	R2	R3	R4	Aligned	Area	Abs. area
Original 1	-0.0099	-0.0075	-0.0120	0.0029	0.0067	-0.0070	0.0268
Original 2	-0.0242	-0.0224	0.0004	0.0014	0.0119	-0.0114	0.0274
Original 3	0.0097	0.0283	-0.0110	-0.0069	-0.0057	0.0059	0.0413
Fresh 1	-0.0228	-0.0000	0.0080	-0.0024	0.0083	-0.0040	0.0263
Fresh 2	-0.0140	0.0041	-0.0137	0.0153	0.0041	-0.0040	0.0398
Fresh 3	0.0002	-0.0008	0.0009	-0.0044	0.0011	-0.0005	0.0292

every available example once, in minibatches of eight. Cumulative sizes 384, 512, 640 and 768 therefore give 48, 64, 80 and 96 updates. Parameters and optimizer state continue across rounds. Six fits and 16 paired streams give 192 four-round paths. This replays the acquired curricula; it does not rerun adaptive selection.

Each paired round uses uniform marginal permutations with common IDs in identical positions. Independently permuted exclusive IDs fill the remaining slots. A label acquired at different rounds can have different earlier exposure, so purchase timing remains part of the contrast. The model, rank, precision, warm states and 1,024-example test panel are those of Section 5.1.

Before running these continuations, we fixed all 24 old risk directions $h_{fr} = \operatorname{sign}(\mathcal{L}^{\mathrm{old}}_{f2r} - \mathcal{L}^{\mathrm{old}}_{f1r})$. For fresh paired stream k, the single primary observation is

$$Z_k = \frac{1}{24} \sum_{f=1}^{6} \sum_{r=1}^{4} h_{fr}(\mathcal{L}_{f2kr} - \mathcal{L}_{f1kr}).$$

Positive expectation means that the previously lower-risk arm retains an average advantage under the complete-consumption learner. Stream k groups independently assigned seeds across all six fits. All 16 groups, old directions and the mean endpoint were fixed before execution. The 95% interval is $\bar{Z} \pm t_{15,.975} s_Z/4$, under an IID Gaussian model for Z_k, conditional on these fits, curricula, directions and panel.

The mean is 0.0044 NLL with interval $[0.0003, 0.0085]$. Five of six fit contributions are positive. Six of the 24 cell means oppose their frozen old directions (Table 14). Round-wise signed means are 0.0135, 0.0008, 0.0039 and -0.0005 NLL; the first round contributes about 76% of the joint point estimate. A post hoc summary restricted to rounds 2–4 gives 0.0014 NLL, with a marginal interval $[-0.0032, 0.0060]$ under the same Gaussian model.

For comparison with Figure 3, we also summarize each replay by its signed area $\sum_j w_j D_j$ and absolute area $\sum_j w_j |D_j|$, where $D_j = \mathcal{L}_{c=2,j} - \mathcal{L}_{c=1,j}$, $D_0 = 0$ and $w = (1, 2, 2, 2, 1)/8$. These post hoc areas average -0.0035 and 0.0318 NLL, respectively, across all fits and streams (Table 14). The absolute area of each fit's mean path instead averages 0.0086: averaging absolute paths includes stream variability, whereas taking the absolute value after averaging is a plug-in summary of the mean trajectory. Neither summary decomposes the old risk gap, because complete consumption changes the learning schedule.

The analysis uses all endpoints, joint observations and 96 trajectory summaries for the descriptive intervals. Local verification recomputed 768 endpoints from saved logits and checked all 198 commits.

A.8 Repeated learning from fixed purchases

We measured the two first-round purchases under a common continuation rule. Each of the six fits reused its saved raw $c = 1$, raw $c = 2$, inverse-Bures and random sets, each containing 128 labels. Every arm restored the same $c = 1$ warm parameters and AdamW moments, configuration and update counter. No acquisition was repeated. Eight independent continuation streams per fit were

Table 15: Held-out NLL averaged over eight continuation streams per fit. All arms start from the same warm parameters and AdamW state. Δ is raw $c = 2$ minus raw $c = 1$; its intervals jointly cover the six conditional means at 97.5% under the paired normal model. Inv.: inverse-Bures; Rnd.: fixed random purchase; Old: old-only.

Fit	Raw 1	Raw 2	Inv.	Rnd.	Old	Δ [simultaneous interval]
Original 1	0.692	0.692	0.604	0.571	0.732	0.001 $[-0.044, 0.045]$
Original 2	0.625	0.612	0.700	0.647	0.671	-0.013 $[-0.106, 0.079]$
Original 3	0.596	0.622	0.558	0.584	0.670	0.026 $[-0.012, 0.064]$
Fresh 1	0.634	0.598	0.571	0.586	0.634	-0.036 $[-0.095, 0.024]$
Fresh 2	0.622	0.625	0.623	0.631	0.679	0.003 $[-0.065, 0.071]$
Fresh 3	0.553	0.550	0.582	0.542	0.697	-0.004 $[-0.091, 0.083]$

shared across the five arms, including old-only, for 240 trajectories. The actual seeds differed across fits.

Each purchase arm ran 128 eight-example updates: four warm slots and four new-label slots per minibatch. Each new label appeared exactly four times. Warm examples occupied identical slots in all five arms; common raw purchases occupied identical new-label slots in the two raw arms. Old-only filled the other four slots with warm-data replay. This controls update count and warm-data ordering while measuring the benefit of replacing half of the replay slots with purchased examples. It differs from the cumulative, 64-update rule in the four-round experiment.

For fit f and continuation stream s, the primary contrast was $d_{fs} = \mathcal{L}(S_{c=2}, s) - \mathcal{L}(S_{c=1}, s)$ on the same 1,024 test examples. We estimated its conditional mean over continuation streams. Six paired Student-t intervals used seven degrees of freedom and a Bonferroni family error budget of 0.025. A separate 0.025 budget covered the 36 contrasts between each raw purchase and the inverse-Bures, fixed random and old-only arms. A comparator reversal required both relevant simultaneous intervals to exclude zero on opposite sides. The design, all eight streams, primary NLL endpoint and error allocation were fixed before training.

All 48 raw pairs changed 23–163 test predictions. Mean differences ranged from -0.036 to 0.026 NLL; all six primary intervals included zero. None of the 18 comparator pairs met the opposed-interval criterion. The simultaneous RMS interval was $[0, 0.081]$ NLL: conditional mean-value differences remain unresolved, rather than establishing equal value.

For conditional means μ_f and their six primary intervals $[l_f, u_f]$, the RMS summary is $\sqrt{\sum_f \mu_f^2/6}$. Mapping the simultaneous rectangle through this function gives

$$\left[\sqrt{\frac{1}{6}\sum_f \max\{l_f, -u_f, 0\}^2}, \quad \sqrt{\frac{1}{6}\sum_f \max\{|l_f|, |u_f|\}^2}\right].$$

It inherits the primary family's nominal 97.5% coverage under the same paired normal model, without a separate error allocation.

Table 16 reports compute-matched label benefits. Deleting any one stream retained the raw mean's sign only in Original 3 and Fresh 1. The paired-normal intervals condition on the six fits, fixed purchases and shared panel; eight streams cannot rule out rare tails. Repeating the saved random set does not average over random acquisition. The analysis used all arm risks and paired observations for the secondary intervals and leave-one-out checks.

A.8.1 Domain-directed purchases

Two Qwen3-1.7B/MultiNLI development fits used government and telephone data, 256 warm labels and 1,000 candidates. Independently scaling early/late query/value LoRA groups by $c \in \{1/2, 1, 2\}$ gives 81 factorizations $(cB, A/c)$ with factor-Jacobian condition number at most four. Acquisition retained the normalized factor-gradient Gram, ridge one, warm history and candidate order. Public genre metadata, without answer labels, identified the first minimum- and maximum-government paths in the fixed enumeration.

Table 16: Compute-matched label benefit: old-only minus purchase-arm NLL. Each entry gives a paired mean and a pointwise descriptive 95% normal-model interval. Positive values favor buying labels over warm-data replay. These pointwise intervals are not the simultaneous intervals used to judge comparator reversals.

Fit	Raw 1	Raw 2	Inverse-Bures	Random
Original 1	0.040 $[-0.086, 0.167]$	0.040 $[-0.080, 0.160]$	0.128 $[0.051, 0.206]$	0.161 $[0.069, 0.253]$
Original 2	0.046 $[-0.050, 0.141]$	0.059 $[-0.009, 0.127]$	-0.029 $[-0.193, 0.135]$	0.023 $[-0.087, 0.134]$
Original 3	0.075 $[0.017, 0.132]$	0.048 $[-0.026, 0.123]$	0.112 $[0.065, 0.160]$	0.086 $[0.008, 0.165]$
Fresh 1	-0.000 $[-0.095, 0.094]$	0.036 $[-0.042, 0.113]$	0.062 $[0.023, 0.102]$	0.047 $[-0.022, 0.117]$
Fresh 2	0.057 $[-0.045, 0.159]$	0.054 $[-0.054, 0.161]$	0.055 $[-0.045, 0.155]$	0.048 $[-0.034, 0.130]$
Fresh 3	0.143 $[0.044, 0.243]$	0.147 $[0.030, 0.264]$	0.115 $[-0.003, 0.234]$	0.155 $[0.057, 0.253]$

Table 17: **Domain-directed purchases and conditional learning.** d_G and d_T are low-minus-high government-purchase NLL on government and telephone, respectively. The primary interaction is Δ_{dom}. Intervals use eight paired continuation streams per fixed fit; the lower panel reports both arms and domains.

Fit	Contrast	Mean NLL difference	Conditional 95% interval
1	d_G	-0.0097	[-0.0574, 0.0380]
1	d_T	-0.0355	[-0.0961, 0.0252]
1	$\Delta_{\mathrm{dom}} = d_G - d_T$	0.0258	[-0.0213, 0.0729]
2	d_G	-0.0160	[-0.0836, 0.0516]
2	d_T	-0.0562	[-0.1225, 0.0100]
2	$\Delta_{\mathrm{dom}} = d_G - d_T$	0.0402	[-0.0049, 0.0854]
Fit	Purchase	NLL (G / T)	Accuracy (G / T, %)
1	Low government	0.497 / 0.578	84.96 / 80.52
1	High government	0.506 / 0.613	84.47 / 79.98
2	Low government	0.520 / 0.550	83.57 / 81.32
2	High government	0.536 / 0.607	84.74 / 80.69

The endpoints bought 141 versus 161 government labels in fit 1 and 126 versus 146 in fit 2, within budgets of 256, replacing 96 and 78 examples. All four preserved task logits and hidden states bitwise on 1,256 inputs, with matching transported gradients. Their 20-label ranges fell below the original 32-label resource screen. A separate learning follow-up retained these endpoints and the reserved evaluation panel; it was fixed after acquisition results but before learning outcomes.

Both arms restored the same 96-update warm parameters and complete Adam state. Eight paired streams per fit ran 256 updates of four warm and four new examples, exposing each purchase four times and matching shared-ID slots. All 32 trajectories completed before scoring on 512 reserved examples per genre. Two same-purchase duplicates reproduced final states and logits bitwise.

For domain $D \in \{G, T\}$, let $d_D = \mathcal{L}_D(S_{\mathrm{low}}) - \mathcal{L}_D(S_{\mathrm{high}})$, where $\mathcal{L}_D$ is held-out NLL and the subscripts denote government quota. The primary contrast is $\Delta_{\mathrm{dom}} = d_G - d_T$. Positive Δ_{dom} favors government relatively; absolute government improvement additionally requires $d_G > 0$. Table 17 uses paired Student-t intervals with seven degrees of freedom, conditional on each fit, purchases and panel. Neither positive mean has a lower bound above zero, so the prespecified two-fit direction criterion is unmet; both d_G means are negative. The contrast includes within-domain purchase changes, not just quota.

Fixed-precision follow-up. After the pilot, a separate cohort used 64 fresh paired streams per fit, holding purchases, learner, exposure law and the previously observed panel fixed. All 256 trajectories completed; two same-purchase duplicates reproduced final states and logits. The eight pilot streams are not pooled with these.

The primary contrast for this cohort is $C = \frac{1}{2}[(\mathcal{L}_G(S_{\mathrm{high}}) - \mathcal{L}_G(S_{\mathrm{low}})) + (\mathcal{L}_T(S_{\mathrm{high}}) - \mathcal{L}_T(S_{\mathrm{low}}))]$. The fixed-size criterion required positive lower bounds for both two-sided 95% Student-t intervals (63 degrees of freedom). Fit 2 has a positive mean cost, but fit 1 remains unresolved: the joint

Table 18: **Mean learning effects under 64 fresh paired streams.** C is high-minus-low government-purchase NLL, averaged equally over the government (G) and telephone (T) panels; Δ_{dom} is the secondary relative-domain contrast. Intervals condition on each fixed fit, purchase pair and evaluation panel. The lower panel reports both arms.

Fit	Contrast	Mean NLL difference	Conditional 95% interval
1	Overall cost C	0.0048	[-0.0117, 0.0213]
1	Domain interaction Δ_{dom}	0.0159	[0.0019, 0.0299]
2	Overall cost C	0.0291	[0.0165, 0.0416]
2	Domain interaction Δ_{dom}	0.0377	[0.0262, 0.0491]
Fit	Purchase	NLL (G / T)	Accuracy (G / T, %)
1	Low government	0.502 / 0.584	85.26 / 80.49
1	High government	0.499 / 0.597	84.93 / 79.96
2	Low government	0.525 / 0.530	83.45 / 81.63
2	High government	0.535 / 0.578	84.35 / 80.86

Table 19: **Repeatable response profiles differ across fixed purchases.** U_ℓ and U_a estimate mean squared per-example differences in expected NLL and correctness, respectively; $\pm$ denotes delete-one-stream jackknife SE, not a confidence interval. r_ℓ, r_a compare the fixed first and last halves. Negative U estimates are retained. All eight fits are shown; these retrospective diagnostics condition on the saved purchases and test rows.

Fit	Streams	U_ℓ	r_ℓ	$10^3 U_a$	r_a
AG Original 1	8	0.026 ± 0.011	0.44	1.95 ± 0.84	0.16
AG Original 2	8	0.006 ± 0.005	0.13	0.10 ± 1.29	-0.04
AG Original 3	8	0.005 ± 0.002	0.63	0.49 ± 2.94	-0.03
AG Fresh 1	8	0.031 ± 0.007	0.76	2.83 ± 1.09	0.23
AG Fresh 2	8	-0.001 ± 0.017	-0.11	-0.73 ± 3.64	-0.19
AG Fresh 3	8	0.013 ± 0.013	0.29	-0.03 ± 3.29	-0.03
NLI 1	64	0.125 ± 0.007	0.96	9.57 ± 0.86	0.82
NLI 2	64	0.095 ± 0.007	0.95	8.58 ± 0.59	0.85

criterion is unmet (Table 18). Neither lower bound exceeds 0.02 NLL. Both secondary Δ_{dom} means are positive, while absolute government advantages remain unresolved. Table 18 gives the contrasts and arm outcomes.

Repeatable per-example responses. We reanalyzed the saved endpoints of all six eight-stream AG News purchase pairs and both 64-stream MultiNLI pairs. No new purchases or training runs were selected. For a fixed fit with N test rows and n_{rep} training streams, let $d_s(i)$ be the second purchase's NLL minus the first's under continuation stream s, and let $\mu_i = \mathbb{E}_s d_s(i)$. Cross-stream products estimate the squared repeatable response without adding the variance of a single training run:

$$U_\ell = \frac{\sum_i [(\sum_{s=1}^{n_{\mathrm{rep}}} d_s(i))^2 - \sum_{s=1}^{n_{\mathrm{rep}}} d_s(i)^2]}{N n_{\mathrm{rep}}(n_{\mathrm{rep}} - 1)}, \qquad \mathbb{E}U_\ell = \frac{1}{N}\sum_i \mu_i^2.$$

Replacing NLL by the binary correctness indicator gives U_a. Table 19 reports both estimates and delete-one-stream jackknife standard errors. These are descriptive conditional estimates, not new confirmatory tests; cross-stream pairs are not independent observations. The fixed first/last split measures response agreement over continuation randomness, with the purchases, warm state and test rows held fixed.

In the two MultiNLI fits, U_ℓ is 0.125 and 0.095 NLL squared, while U_a corresponds to descriptive root-mean-square correctness-frequency changes of 9.8 and 9.3 percentage points. The overall accuracy changes are -0.43 and $+0.06$ points. Defining positive and negative NLL-response groups on the first 32 streams gives last-half means of $+0.109/-0.231$ and $+0.285/-0.056$, respectively. Both reverse-half checks retain this opposition. The same-purchase negative controls instead compare disjoint stream pairs; their four U_ℓ estimates range from -0.0027 to 0.0034, with standard errors 0.0087–0.0124. Their noise coupling differs from that of the paired purchase contrast.

Table 20: Complete nested pool–budget scan. Values average four independently fitted adapters and ten paired selector streams per recipe; J denotes selected-set Jaccard similarity. "Natural J" compares the training-emitted chart with the intrinsic balanced kernel; $Q = 2$ columns compare exact-equivalent factor charts.

Recipe	b/n	n	b	$Q = 2\,J$	**Unit** J	**Replaced /** b	**Natural** J
AdamW	1/12	120	10	.345	.279	5.42	.448
AdamW	1/6	120	20	.288	.249	11.43	.354
AdamW	1/3	120	40	.371	.341	18.70	.424
AdamW	1/12	240	20	.174	.179	14.25	.209
AdamW	1/6	240	40	.195	.201	27.05	.224
AdamW	1/3	240	80	.305	.298	42.75	.313
AdamW	1/12	300	25	.166	.160	18.10	.170
AdamW	1/6	300	50	.193	.192	34.00	.200
AdamW	1/3	300	100	.299	.299	54.20	.303
LoRA+	1/12	120	10	.198	.259	6.88	.502
LoRA+	1/6	120	20	.187	.249	13.85	.371
LoRA+	1/3	120	40	.316	.327	20.90	.419
LoRA+	1/12	240	20	.133	.157	15.43	.209
LoRA+	1/6	240	40	.180	.192	27.93	.216
LoRA+	1/3	240	80	.297	.304	43.45	.315
LoRA+	1/12	300	25	.114	.138	20.00	.184
LoRA+	1/6	300	50	.175	.176	35.23	.210
LoRA+	1/3	300	100	.284	.286	55.85	.315

The response estimators, centering operations and ensemble-probability diagnostics are implemented in `analysis/profile_metrics.py`.

A.9 LoRA TRAINING-RECIPE, POOL AND BUDGET STUDY

Eight Qwen3-1.7B adapters form four paired task–seed blocks, each with AdamW and LoRA+ fits sharing initialization, examples, minibatch order and 60 updates. AdamW uses 2×10^{-4} for both factors; LoRA+ uses that rate for A and 8×10^{-4} for B. Rank-8 LoRA ($\alpha = 16$) acts on query/value projections in the final two layers. Each fit uses 120 class-balanced examples, batches of four, at most 256 tokens and no weight decay. Prompts request one class digit: TREC (Li & Roth, 2002) uses ABBR/DESC/ENTY/HUM/LOC/NUM as 0–5, and AG News uses World/Sports/Business/Sci-Tech as 0–3. Frozen, nested class-balanced pools have 120, 240 or 300 candidates; budget fractions are $1/12$, $1/6$ and $1/3$. This auxiliary scan is described here and summarized in Table 20; its GPU launcher is not included in the compact release.

Per-example gradients are computed once per fit. The $Q = 2I$ comparison and smaller pool–budget cells are analytical replays of those gradients. The raw and unit-normalized D^2 selectors share the max-norm first-center rule and ten inverse-CDF streams. All eight fits pass the update, logit and pseudo-label equality checks. The inverse-Bures matched control changes none of the 720 raw or 720 unit-normalized ordered batches. Table 20 reports all nine cells per recipe.

At the $120/20$ condition, all 40 streams per recipe split. LoRA+ reduces the median balance condition from 5.128 to 4.579, but its paired $Q = 2I$ Jaccard change is positive on both TREC fits and negative on both AG News fits: recipe ordering is task dependent.

A.10 A PUBLIC ADAPTER-COMPOSITION EXAMPLE

PEFT's `add_weighted_adapter` exposes concatenation and SVD merge representations (Hugging Face, 2026). We compare these public calls on six fixed synthetic CPU tests. Each combines two $8 \to 6$ linear blocks as $\tanh(W_1 x) + W_2\,\mathrm{reverse}(x)$, with two distinct rank-two adapters per block, unit adapter scales and equal merge weights. Concatenation and unclamped SVD at output rank four represent the same weighted update; default lower-rank SVD would truncate it. With PEFT 0.20.0 and float64, maximum update and logit differences are below 10^{-15}, and pseudo-labels agree. From 64 candidates after four conditioning examples, the ridge-one, unit-normalized

maximum-variance rule replaces 6–9 of 12 purchases. Independently computed K_{bal} preserves the ordered path; transporting SVD gradients back recovers the concatenation raw path. Their maximum normalized-Gram errors are 1.09×10^{-15} and 1.17×10^{-15}, versus a minimum winner margin of 3.14×10^{-5}.

Ordinary save/load retains factors and purchases bitwise, as expected from the checkpoint format (Hugging Face, n.d.). A single-adapter SVD request also retains the path because the API dispatches it to linear merging. This example concerns an explicit composition choice, not automatic serialization drift.

Figure assets. The icons in Figure 1 are from Streamline Freehand (CC BY 4.0), scaled and recolored.