AMIYC ADD ME IF YOU CAN CONDITIONAL ADAPTATION HEADROOM AND RECOVERY IN EVOLVING MMOES

Anonymous authors
Paper under double-blind review

ABSTRACT

Multi-task mixture-of-experts (MMoE) models are commonly evaluated through routing behavior and end-to-end task quality, but neither establishes whether a trained state has exhausted the task-conditioned improvement available within its learned representation. We call this measurable opportunity conditional adaptation headroom. We introduce AMIYC, a framework that separates routing exposure from the task-visible expert state and distinguishes fresh adaptation opportunity, compatibility of earlier interventions, and recovery from a finite action bank. Across Amazon Reviews (Beauty, Grocery, Sports, and Toys), KuaiRand-Pure, and a production e-commerce ranking MMoE, bounded interventions reveal positive headroom on the tested cohorts, including near validation-selected checkpoints and when the same intervention coordinates participate in ordinary training. Alternative adapters show that the reference diagonal is not unique. As representations evolve, earlier operators may retain value, become harmful, or recover usefulness, while tested transport corrections do not consistently improve replay. Targeted Amazon studies further reveal reversals in checkpoint ordering after common-anchor fitting and associate a tested coupled adaptation schedule's deficit predominantly with its base trajectory. Realizing this opportunity remains constrained: most final raw gain in the tested continuation is inherited from the historical anchor, affine calibration substantially reduces the apparent loss advantage, and useful ranking improvements remain selective. Finite-bank availability, action selection, and secondary-quality constraints further limit recovery. Fresh-seed confirmation reuses previously examined data. Together, these findings distinguish bounded, state-relative opportunity from compatibility and operational recoverability, without implying adaptation saturation or a universal remedy.

1 INTRODUCTION

Does a well-trained multi-task model still admit useful task-specific adaptation, and does an adaptation fitted earlier remain useful as the model evolves? A shared model may retain accessible improvement near a validation-selected checkpoint, while an earlier operator may retain value, lose effectiveness, or become harmful. Repeating the fit requires additional supervision and computation; reuse avoids that work only when the operator remains compatible. The scientific question is therefore which opportunity exists at a state, which earlier adaptations remain useful, and which value a specified procedure can recover.

Post-training specialization is well established. MT-DNN and sparse multi-task experts refine individual tasks after joint training (Liu et al., 2019; Gupta et al., 2022); activation scaling and representation interventions adapt frozen models through restricted parameters (Liu et al., 2022; Wu et al., 2024). Trans-PEFT studies adaptation across evolving bases, while UpgradeBench distinguishes specialist quality, portability, and recovery resources (Gu et al., 2025; Chen & Zhang, 2026). Building on this work, AMIYC characterizes how model state, training history, and recovery resources determine accessible adaptation value. We test explanations for its divergence from reuse and recovery: checkpoint quality, already-trained coordinates, fitting variability, feedback into the base, inherited information, and calibration.

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.

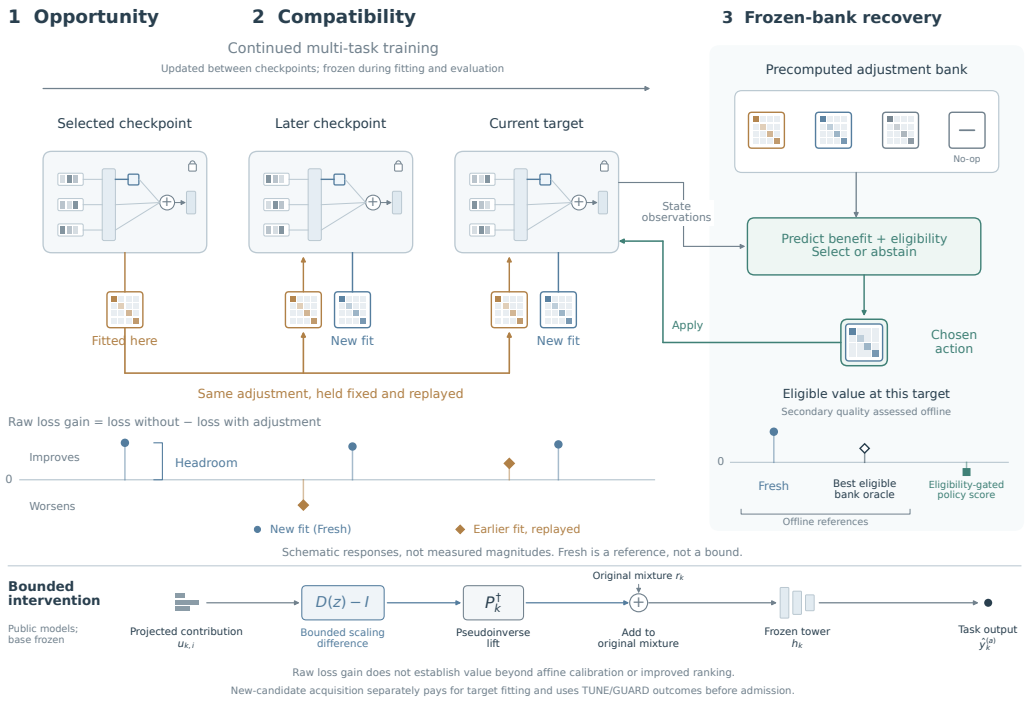

Figure 1: Current-state opportunity, replay compatibility, and frozen-bank recovery in AMIYC. Fresh fitting provides a local reference, while earlier operators are replayed unchanged. A witness-based selector chooses a bank action or abstains. Responses are schematic; Fresh is not an upper bound. The lower diagram shows the public-model intervention: a bounded projected scaling difference is lifted through $P_k^\dagger$ and added to the original mixture before the frozen task tower.

AMIYC (*Add Me If You Can*) studies these questions at the task-visible additive expert boundary of an MMoE (Ma et al., 2018). We define *conditional adaptation headroom* as held-out improvement under a prespecified bounded intervention at a frozen model state. With routing and network parameters fixed during each readout, we distinguish *opportunity*, measured by fresh current-state fitting; *compatibility*, measured by replay of an earlier operator; and *recovery*, measured by a declared selection or fitting procedure, including abstention (Figure 1). These quantities depend on the intervention family, budget, and information access. Fresh is a finite reference, not an upper bound.

Across Amazon Reviews 2023 (our four-category subset, Mixed-4Cat) (Hou et al., 2024), KuaiRand-Pure (Gao et al., 2022), and a production ranking MMoE, fresh fitting reveals positive headroom on the tested cohorts. It persists near high-quality checkpoints and on coordinates already trained with the base; alternative same-boundary adapters extend the response beyond the reference diagonal. Earlier operators can retain substantial value or become harmful, and tested transport corrections do not consistently improve primary gain. Frozen-bank experiments further separate available value from successful pre-intervention selection. New-candidate recovery additionally spends target fitting and validation resources before admission.

Targeted Amazon studies sharpen this framework. Native checkpoint loss ordering reverses after common-anchor fitting, and deferred, crossed, and shadow controls associate a tested coupled schedule's deficit predominantly with its base trajectory. Same-base accounting separates value retained from earlier operators from additional refitting gain. Rank-preserving affine calibration leaves a small positive fitted-endpoint advantage in the core cohort and an unresolved residual on fresh initializations. Selective admission retains useful ranking gains. These comparisons make the objective, historical information, and admission coverage explicit components of recovered value.

Our contributions form one empirical argument about *state-relative adaptation value*: (1) establish bounded opportunity across three MMoE environments and test its construct validity; (2) distinguish opportunity from compatibility through replay, transport, and controlled trajectory comparisons;

and (3) quantify recovery under finite-bank availability, selection, acquisition cost, calibration, and secondary-quality constraints. The environments provide complementary evidence, with the original cross-domain measurements and later controlled trajectories answering different parts of the argument. Together, the results support distinct measurements for retaining, refitting, or withholding an adaptation.

2 Measuring Adaptation at a Frozen State

A task-visible intervention. For task k, let the MMoE mixture and its projected expert contributions be

$$\boldsymbol{r}_k(\boldsymbol{x}) = \sum_i g_{k,i}(\boldsymbol{x})\boldsymbol{e}_i(\boldsymbol{x}), \qquad \boldsymbol{u}_{k,i} = \boldsymbol{P}_k g_{k,i}\boldsymbol{e}_i, \qquad \boldsymbol{u}_k = \sum_i \boldsymbol{u}_{k,i}, \tag{1}$$

where $g_{k,i} \geq 0$, $\sum_i g_{k,i} = 1$, and $\boldsymbol{P}_k$ is the exposed linear task projection. Additivity holds at this boundary even when the downstream tower h_k is nonlinear. Routing describes exposure; it does not determine behavioral value. The public models use four 128-dimensional experts and 64 projected coordinates. An action transforms one directed contribution through $\boldsymbol{D}(\boldsymbol{z}) = \mathrm{diag}(e^{z_1}, \ldots, e^{z_d})$, with identity $\boldsymbol{z} = 0$ and $|z_\ell| \leq 0.5$. The executed public-model intervention is

$$\begin{aligned} \boldsymbol{r}_k^{(a)} &= \boldsymbol{r}_k + \boldsymbol{P}_k^\dagger[\boldsymbol{D}(\boldsymbol{z}_a) - \boldsymbol{I}]\boldsymbol{u}_{k,i}, \qquad \hat{y}_k^{(a)} = h_k(\boldsymbol{r}_k^{(a)}), \\ \boldsymbol{u}_k^{(a)} &= \boldsymbol{u}_k + \boldsymbol{P}_k\boldsymbol{P}_k^\dagger[\boldsymbol{D}(\boldsymbol{z}_a) - \boldsymbol{I}]\boldsymbol{u}_{k,i}. \end{aligned} \tag{2}$$

The Moore–Penrose lift is defined without full row rank. Direct projected replacement additionally requires $\boldsymbol{P}_k\boldsymbol{P}_k^\dagger = \boldsymbol{I}$; this rank condition is not assumed for every historical state. Routing, experts, and tower stay fixed during fitting and readout. The source projection and lift also stay fixed across the later Amazon base trajectories. Pair support (i, j) specifies a directed expert and a companion used by the original pair-energy penalty; it is not a uniquely causal defect location.

Fitting and measurement contracts. An operator contains coefficients and source provenance; an action additionally specifies target task, support, direction, and dose. Original FIT optimizes task loss with identity and pair-energy penalties; SELECT chooses without EVAL outcomes. Kuai uses identity initialization and AdamW, whereas Amazon replay-cohort Fresh starts from the source operator. Later Amazon refits use the same coordinate box plus $\|\boldsymbol{z} - \boldsymbol{z}_0\|_2 \leq 1$, 20 FIT lists, 120 momentum-free SGD updates, and $0.01\,\mathrm{mean}[(\boldsymbol{z} - \boldsymbol{z}_0)^2]$ regularization. The historical anchor $\boldsymbol{z}_0$ supplies initialization, radius center, and penalty center. Box clipping followed by radial contraction enforces feasibility, not an exact Euclidean projection onto the intersection. These histories and feasible sets differ from identity-centered fitting (Appendices A.2.3, B).

For state s and action a, paired held-out effects are

$$\Delta L(s, a) = L_{\mathrm{ref}}(s) - L_a(s), \qquad \Delta Q(s, a) = Q_a(s) - Q_{\mathrm{ref}}(s). \tag{3}$$

Positive values improve primary loss or secondary quality. Each comparison names its reference: identity, a held operator, or a complete alternative system. *Conditional adaptation headroom* is the response of a prespecified finite current-state fitting procedure, called Fresh. It is not a global optimum, an equal-budget training advantage, or a guarantee that secondary quality improves. Separate FIT, selection, and EVAL roles prevent intervention-outcome leakage within a study; development-selected checkpoints and reuse across campaigns still limit external confirmation.

Pre-intervention state information. The witness $\boldsymbol{w}_{s,k,ij}$ combines exposure, spectral, geometric, interaction, and dimensional summaries, symmetric pair transforms, within-panel relative features, and past-only temporal context. It excludes seed identifiers, Fresh/replay outcomes, realized displacement, and future measurements. These are predictive descriptors, not sufficient or globally invariant coordinates. In particular, suppressing the tested interaction statistic can first improve and then worsen ranking, so that statistic alone does not prescribe a useful action (Appendix A.4).

3 Compatibility and Two Routes to Recovery

Replay and transport. Fresh fits at the current frozen target; Same replays an earlier operator from the same training lineage; Foreign replays one from an independent initialization. Replay

preserves source coefficients. A competing fixed-dose operator-space port of GradFix (Rinaldi et al., 2026) masks those coefficients using anti-gradient sign votes at target identity: 20 target-FIT units per action, no optimizer steps, and dose one. The published method's validation-scaled application and guarantees do not transfer automatically to this port. All comparisons preserve source support and direction and evaluate on matched target lists.

Frozen-bank recovery: predict before intervening. The original bank $\mathcal{B}$ is acquired on the training side of an outer seed, trajectory-role, or temporal split and frozen before the held-out target. A histogram gradient-boosted tree scorer predicts benefit and eligibility using only the witness and action provenance. It chooses the highest predicted benefit among candidates with predicted positive value and sufficient eligibility probability, otherwise no-op. For the original Amazon controller, eligibility requires NDCG gain ≥ -0.002. Three run-grouped inner out-of-fold folds support surface-aware isotonic calibration and cutoff selection under a 5% observed violation budget, using outer-training data only. Recovered temporal cutoffs and source qualifications appear in Appendix A.8.1; these are historical settings, not the later TUNE/GUARD rule. The observed score is

$$U_\epsilon(s,a) = \Delta L(s,a)\mathbf{1}[\Delta Q(s,a) \geq -\epsilon],$$
$$G^\star_{\mathcal{B}} = \mathbb{E}_s \max_{a \in \mathcal{B}\cup\{0\}} U_\epsilon(s,a), \quad G_\pi = \mathbb{E}_s U_\epsilon(s,\pi(s)). \tag{4}$$

Here $\mathbb{E}_s$ averages the evaluated panels. The maximum uses EVAL outcomes and is an offline oracle, not an executable selector. Comparing G_{Fresh}, $G^\star_{\mathcal{B}}$, and G_π on matched states separates availability from selection; Fresh can be exceeded by bank actions. Zeroing an ineligible outcome in U_ϵ does not undo execution, so G_π is not ungated deployed loss gain. Historical operator fitting and scorer training are paid before use; target witness extraction, scoring, and execution are newly incurred. No joint end-to-end controller timing is available.

Acquisition chronology: does fitting alter the base? Later Amazon branches share each seed's selected source weights and already fitted anchor $\boldsymbol{z}_0$. Over two additional native epochs, ordinary training applies identity, fixed continuation applies $\boldsymbol{z}_0$, and coupled periodic continuation refreshes the correction. Deferred fitting uses 120 updates on the final fixed-incumbent base. Shadow fitting instead distributes 120 updates across six saved states of that same fixed trajectory, without feeding its corrections into base training. Coupled coefficients are detached during base updates, but their forward effect remains. Crossed readouts exchange terminal coefficients between the fixed and coupled bases. Matching base exposure and new coefficient updates does not equalize historical acquisition cost or gradient-acquisition states.

Historical and objective controls. On the same frozen base and EVAL lists, define inherited gain $H = L(0) - L(\boldsymbol{z}_0)$, new gain $F = L(\boldsymbol{z}_0) - L(\boldsymbol{z}_1)$, and total gain $T = H + F$. This exact accounting is not a causal decomposition. Each fresh-weight trajectory has its own fitted anchor. A separate FIT-only positive affine logit map $\ell \mapsto a\ell + b$, $a > 0$, preserves ranking while changing BCE (Guo et al., 2017). Giving both identity and refit this two-parameter option measures a calibration-adjusted residual. It tests an alternative action family, not equal historical computation or a prospective calibration-aware selector.

New-candidate recovery: fit, then admit once. The final acquisition policy pays for target refitting and evaluates $\boldsymbol{z}_0 + \alpha(\boldsymbol{z}_1 - \boldsymbol{z}_0)$ for $\alpha \in \{0, 1/16, 1/8, 1/4, 1/2, 1\}$, plus identity. TUNE chooses minimum target BCE subject to point loss nonworsening and all three NDCG@10 changes at least -0.002 versus same-base identity; ties favor identity. One disjoint GUARD check requires point loss nonworsening and one-sided user-cluster t quality lower bounds at that floor, with per-head error level $0.05/3$. Failure returns identity without retry. The primary split has 4,915 TUNE, 3,277 GUARD, and 8,192 separate EVAL lists. Only the target head changes on this frozen base. Unlike frozen-bank prediction, this policy observes target candidate outcomes before application and incurs fitting and selection costs; its reported all-seed gains include actual identity fallbacks.

4 Experiments

Amazon Reviews 2023 (Mixed-4Cat) supplies intervention, trajectory, and selection studies; KuaiRand-Pure supplies eight original trajectories and a test of remedy transfer; Production supplies nine targets across three ranking-model lineages. Amazon's Rank, Next, and Future heads

Table 1: **Bounded response and its strongest construct controls.** Positive gain lowers loss; native scales and cohorts differ.

A. Original current-state response

Setting	Seeds / lineages	No-op loss	Fresh gain	Positive / tested
Amazon	2	1.07445918	+.028745	360/360
Kuai	8	.60602216	+.0006945	382/384
Production	3	.50758169	+.006864	9/9

B. Kuai checkpoint and trained-coordinate controls (8 seeds)

Quantity	Estimate	95% seed interval	Support
Pre/near-peak Fresh gain	+.000861	[.000777, .000944]	8/8 seeds
Trained-diagonal dev-AUC − ordinary	−.000114	[−.000451, .000254]	unresolved
Residual mean-pair gain	+.000704	—	95/96 positive

C. Matched-displacement alternatives (Amazon, 7 seeds)

Alternative	Parameters	Reference − alternative U_ϵ	95% seed interval
Rotated diagonal	64	+.00329	[−.00532, .01286]
Task-global diagonal	64	−.00203	[−.01043, .00770]
Rank-1 adapter	128	−.01608	[−.02710, −.00777]

A uses each setting's native loss and original cohort: Amazon averages Rank, Next, and Future losses; Kuai uses binary loss; Production uses pairwise ranking loss. Counts are dependent outcomes within the listed seeds/lineages. The original Amazon analysis and later 540-outcome snapshot are distinguished in Appendix A.5. B reports the descriptive all-pair residual mean; all 96 actions are eligible and the composite diagonal has a larger bound. C compares eligibility-gated value (Eq. 4); intervals spanning zero indicate unresolved contrasts.

model listwise immediate ranking, pointwise discrimination of that target, and short-horizon intent. Original opportunity pools their losses equally; later Next comparisons use list-averaged masked BCE and NDCG@10 on sampled lists. Different cohorts and native loss scales are never pooled. The final studies distinguish seven core seeds from five new training seeds.

Intervals aggregate over independent model seeds or lineages; heads, checkpoints, state–operator pairs, and resamples are repeated observations within them. Later-study intervals are exploratory and pointwise. The five fresh initializations use the stated evaluation population: 7,841 of 8,192 EVAL rows overlap earlier core EVAL or calibration data. Within-study FIT/SELECT/EVAL separation and cohort identities are detailed in Appendices A.2 and B.

4.1 Opportunity Survives Checkpoint and Coordinate Controls

Fresh fitting reveals positive response across all three settings (Table 1A). Kuai's pre/near-peak result rules out degraded late checkpoints as the sole explanation. Training the diagonal coordinates from initialization also leaves measurable residual response: 95 of 96 pair outcomes improve, with all 96 eligible. The residual multiplies the trained diagonal and therefore expands the composite bound. The result establishes response on trained coordinates; selected-model AUC and residual fitting use the distinct comparisons in Table 1B.

The same boundary supports useful response through several parameterizations. At matched displacement, the native diagonal's eligible-value contrasts with rotated and task-global diagonals span zero, while a rank-1 adapter with twice the parameters performs better (Table 1C). These controls separate support, parameter count, and coordinate basis. Output calibration addresses a further, distinct question: Section 4.4 measures its effect on the later Amazon endpoints. The original cohorts establish response under their bounded intervention contracts.

4.2 Compatibility Depends on State, Operator, and Objective

Replay and refitting answer different questions even when fitted coefficients move. Kuai's unchanged earlier operator retains 89.8% of the matched Fresh gain in Table 2A. Production supplies the opposite regime: at horizon $+5$, all three Same replays lose value while all three Fresh fits im-

Table 2: **Replay compatibility and fixed-dose target-gradient transport.** Matched primary and secondary outcomes can favor different actions.

A. Matched primary loss gain; GF is the tested GradFix port

Setting	Fresh	Raw replay	GF	GF − Raw [95% CI]
Kuai (8 seeds)	.000737	.000662	.000213	$-.000449$ $[-.000568, -.000341]$
Amazon Next (7)	.035538	.027652	.027230	$-.000422$ $[-.000854, -.000123]$

B. Amazon secondary outcomes on the same 378 actions

Quantity	Raw replay	GF	GF − Raw
Next ΔNDCG@10	$-.003112$	$+.002454$	$+.005566$
Observed eligible / actions	$232/378$	$278/378$	$+12.17$ pp
Eligible primary value U_ϵ	$+.014007$	$+.018156$	$+.004149$

C. Distinct replay regimes: local opportunity does not ensure compatibility

Target	Metric	Fresh	Same	Foreign	Same positive
Amazon evolved	U_ϵ	.017843	.014007	.001080	—
Production $+1$	ΔL	.005638	$-.000604$	$-.000105$	$1/3$
Production $+3$	ΔL	.008788	$+.002040$	$-.000077$	$2/3$
Production $+5$	ΔL	.006166	$-.003428$	$-.000099$	$0/3$

A uses matched no-op losses .606102 (Kuai e12–e14) and .309922 (Amazon Next), distinct from Table 1A. GF–Raw is negative in 8/8 and 6/7 seed means. B is exploratory, unadjusted secondary analysis at $\epsilon = .002$; counts are dependent actions. C uses 7 Amazon seeds and 3 Production lineages; Fresh is positive in all nine Production states. Alignment and donor controls appear in the text. The fixed-dose port does not reproduce every component of published GradFix.

prove. The horizon means summarize three independent lineages. Later Amazon resampling also finds greater within-state than between-state coefficient variation while 756/756 refits improve loss across seven seeds. Parameter movement alone therefore cannot identify representation drift or the need to refresh.

Target-gradient transport supplies a concrete competing procedure. The fixed-dose GradFix port loses primary gain relative to unchanged replay in both public settings, with negative paired seed intervals (Table 2). Yet its Amazon NDCG improvement is positive: $+0.005566$, interval $[0.001403, 0.011677]$, with seven positive seed means. Eligibility rises from 232/378 to 278/378 actions at the -0.002 tolerance. This exploratory secondary result was not optimized by the mask and shows that primary preservation and ranking quality can diverge. The comparison uses already-useful Same operators and the fixed target-gradient mask.

Cross-lineage reuse is harder on the evolved Amazon cohort. Same retains 78.5% of contemporaneous Fresh eligible value; raw Foreign is much weaker. FIT-only expert/coordinate alignment increases Foreign value to 0.002557, recovering 11.4% of the Same–Foreign gap. Even the nearest tested Foreign donor trails Same by 0.01373 ($[0.01078, 0.01729]$). The remaining gap motivates current-state fitting. Matched donor magnitudes also lose to constant magnitudes under stronger target-fitting controls, separating the value of a target direction from donor-magnitude specificity (Appendices A.6, B.3).

4.3 Native Quality and Fitted Value Can Reverse Their Ordering

The original tracking study motivates a stronger question than whether local fitting helps. On 63 evolved Amazon targets from seven seeds, holding the base fixed and fitting the same coefficients under ordinary MTL improves loss over the historical joint endpoint by 0.010477. Fresh also improves over the block-coordinate online endpoint by 0.01250. Updated coefficients lose to held coefficients at all 63 same-base readouts, while both retain positive mean identity-relative value. These paired loss comparisons establish useful fixed-state supervision and distinguish the tested online recipes. The original three-epoch continuation also retains selected-action loss gain $+0.011493$ and NDCG gain $+0.00392$ against matched evolved no-op, with positive seed intervals (Appendix A.7).

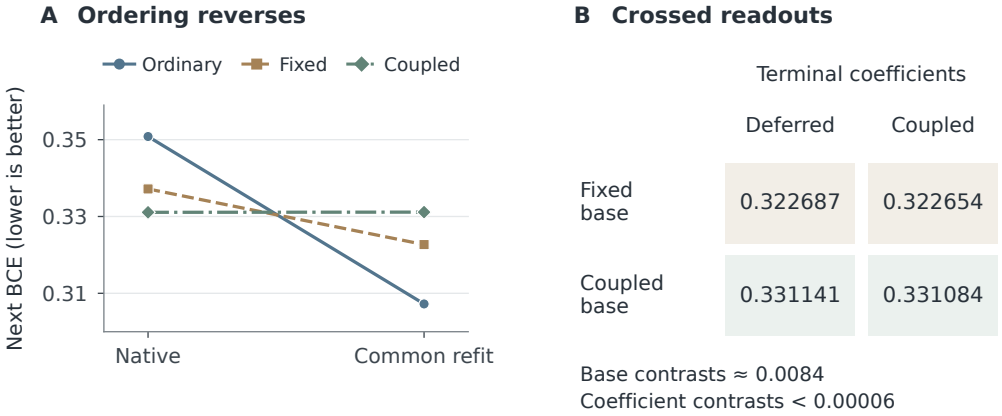

Figure 2: **Native versus recoverable endpoint ordering, and localization of the gap.** Left: mean Next BCE on 7 Amazon seeds. Each common refit uses the historical anchor as initialization, radius center, and penalty center. Right: crossing terminal coefficients without refitting changes loss far less than crossing bases. Deferred–coupled advantage is $+.008397$ $[.005983, .010873]$; shadow–deferred is $+.000036$ $[-.000007, .000101]$ with matched 120 new updates. Points/cells are means, not independent replications; intervals are paired seed summaries. Native operators differ by trajectory. Cross-base calibrated ordering and equal total historical cost are unestablished.

Stronger continuation controls preserve that distinction. Periodic SGD improves over fixed continuation by 0.006539 across seven positive seed means. Nevertheless, deferring the same number of coefficient updates improves its endpoint by another 0.008397 ($[0.005983, 0.010873]$). Ordinary continuation supplies a further planned secondary control: despite worse native loss, its common-anchor fitted endpoint beats fixed-common by 0.015457 ($[0.012996, 0.018812]$; 7/7 positive contrasts; Figure 2). Native and fitted endpoint quality therefore rank these bases differently. Historical initialization, radius center, and penalty center remain part of that contract; cross-base ordering after affine calibration is untested.

Crossed terminal readouts associate most of the deferred–coupled difference with the base path: mean base contrasts are approximately 0.0084, versus coefficient contrasts below 0.00006. Shadow fitting on saved fixed-trajectory states differs from deferred fitting by only $+0.000036$ ($[-0.000007, 0.000101]$), within the declared ± 0.0006 margin. Sequential fitting without feedback retains the deferred endpoint benefit. Together, the crossed and shadow controls associate the coupled deficit mainly with its base path. The optimizer result is conditional: SGD's Amazon advantage does not replicate consistently on Kuai, and a shared-feasible-set control identifies warm-start acceleration at early budgets. Continued base training retains 85.47% of the periodic system's mean advantage for another epoch with its operator applied (Appendices B.4–B.5.3).

4.4 Historical and New Gain Under Matched Output Calibration

Table 4 compares same-base identity, historical anchor, and refit. Core uses the ordinary-common endpoint above; fresh trajectories train ordinarily to epoch 12 with their own epoch-6 anchors. Both retained anchor value H and additional refit value F are positive in every reported seed. The anchor supplies 69.6% and 72.9% of total mean raw gain in core and fresh, respectively. This accounting credits the earlier fit and the newest 120 updates separately.

A two-parameter output control separates score correction from residual intervention value. FIT-only positive affine calibration of identity gains 0.166739 in core and 0.290854 in fresh, exceeding uncalibrated refit BCE improvement in every seed without altering ranking. After calibrating both endpoints, refit's residual is $+0.001598$ in core with a positive interval, but -0.003544 in fresh with an interval spanning zero. The matched control identifies a positive core residual and uncertainty about the fresh-cohort mean. Ranking changes supply a separate readout because positive affine correction preserves order. These affine comparisons use the later same-base cohorts; the original opportunity panels and cross-base ordering have different evaluation contracts.

Table 3: **Frozen-bank recovery and its distinct information contract.** No target candidate fitting or outcome readout is used by the selector.

A. Amazon temporal-forward selection (122 dependent panels)

Scorer	Within-state ρ	Coverage	Violations / acted	Bank capture
Action only	.431	$25/122$	$0/25$	4.9%
Witness + action	.919	$102/122$	$3/102$	75.3%

B. Matched eligible values; the oracle includes no-op

Setting	G_{Fresh}	$G^{\star}_{\mathcal{B}}$	G_{π}	Measured ratio
Amazon (122 panels)	not measured	.047710	.035940	Policy/bank 75.3%
Kuai (48 panels)	.001240	.000310	$-.000019$	Bank/Fresh 25.0%

C. Information and computation available before action

Contract	Frozen-bank selector	New acquisition (Table 4)
Inherited work	Operator bank; scorer training	Base training; historical anchor
New target work	Witness extraction; scoring	Refit; candidate TUNE/GUARD
Target outcomes	Not observed by selector	Observed before admission
Fallback	Predicted choice or no-op	One checked winner or identity
Reported value	Gated U_{ϵ}; raw gain separately	Executed gain including fallback

ρ is within-state Spearman correlation. Panel B reports eligibility-gated value; Amazon raw policy gain is $+.036825$. The original Amazon floor is $-.002$; fitted temporal cutoffs are $.85/.90$ for witness HistGB and $.95/.95$ for action-only HistGB. Coverage is 83.61%; violations/acted is 2.94%, with a one-sided 95% iid-panel upper bound of 7.43%. Amazon Fresh is unmeasured on these exact panels. The original configuration, predictions, decisions, and final-analysis inputs are bound by the experiment's checksum manifest. Kuai's task–pair prior and the measured cost components are detailed in Appendix A.8.

4.5 Recovery Separates Bank Availability, Selection, and Admission

Reusing a frozen bank. On the original 122 Amazon temporal-forward panels, state witnesses improve within-state action ranking from $\rho = 0.431$ to 0.919, and eligible bank-value capture from 4.9% to 75.3% (Table 3). Both scorers use the same bank and action descriptors, isolating the contribution of state witnesses. The policy acts on 102/122 panels, with 3/102 observed violations; the corresponding 95% iid-panel upper bound is 7.43%. Raw policy gain is separately recorded as $+0.036825$, versus $+0.035940$ after eligibility gating. The 75.3% ratio measures eligible bank capture on this population. The observed 2.94% violation rate meets the original 5% point-estimate budget; the uncertainty bound remains above that budget.

Kuai exposes the upstream availability bottleneck: an oracle with no-op can retain only 25.0% of contemporaneous Fresh eligible value from its frozen bank. The learned selector realizes slightly negative gated gain. Witnesses improve over weak context while a strong task–pair prior remains competitive. This distinguishes improving selection within an available bank from acquiring candidates that the bank lacks. Its cost matters: on three profiled Kuai states, searching 12 Fresh actions takes 28,896–53,520 backward operations and 15.9–16.6 minutes on a V100. The matched GradFix derivation uses 240 backward evaluations and no optimizer steps, excluding replay EVAL. These measured components quantify the fitting work that reuse can avoid; full controller cost additionally includes witness construction and selection.

Acquiring and admitting new candidates. The locked target-fitting policy uses a historical anchor, new fitting, and observed TUNE/GUARD candidate outcomes, unlike the bank scorer. Raw full refits breach ranking tolerances in 1/7 core and 3/5 fresh seeds despite positive loss gains. The policy admits 6/7 and 2/5, retaining positive all-seed mean loss and NDCG changes with zero observed active violations (Table 4). The fresh-cohort intervals reach zero. All admitted actions are full refits, inactive seeds choose identity at TUNE, and point and confidence-based checks select the same primary endpoints. Thus the primary result measures the complete candidate-selection and fallback procedure.

Identity fallback is substantive. Incremental attenuation moves toward the historical anchor and leaves the same three fresh failures at every tested strength; whole-vector shrinking toward identity

Table 4: **New acquisition: historical gain, affine challenge, and admitted recovery.** Same-base identity is the reference except for the calibrated residual.

Quantity	Core (7 seeds)	Fresh weights (5 seeds)
Inherited H / new F (loss gain)	$+.030342$ / $+.013248$	$+.051348$ / $+.019088$
Total raw refit gain T	$+.043591$	$+.070436$
95% seed interval	$[.034280, .054692]$	$[.056629, .083922]$
Affine identity gain	$+.166739$	$+.290854$
95% seed interval	$[.123486, .217607]$	$[.211221, .385937]$
Calibrated refit advantage	$+.001598$	$-.003544$
95% seed interval	$[.000590, .002522]$	$[-.008271, .001183]$
Raw refit Next ΔNDCG@10	$+.013338$	$-.062320$
95% seed interval	$[.003566, .022781]$	$[-.160004, .011670]$
Raw ranking failures / total	$1/7$	$3/5$
Policy loss gain, all seeds	$+.036746$	$+.022730$
95% seed interval	$[.021925, .051492]$	$[0, .050785]$
Policy Next ΔNDCG, all seeds	$+.014812$	$+.007952$
95% seed interval	$[.007393, .023069]$	$[0, .016984]$
Active / total; violations / active	$6/7$; $0/6$	$2/5$; $0/2$

Core uses ordinary-common; fresh weights use their own historical anchors and ordinary endpoint. $H+F=T$ up to displayed rounding; this is accounting, not causal attribution. The residual is $L(\text{calibrated identity}) - L(\text{calibrated refit})$. Affine fits preserve ranking and differ in action family and historical compute. Policy means include identity fallback; the primary 8,192-list TUNE/GUARD budget uses one choice, one check, no retry. Point and LCB versions choose identical full-refit/identity endpoints. Intervals are exploratory seed bootstraps on reused data; fresh EVAL substantially overlaps previously examined populations. Zero observed violations at the $-.002$ quality floor is not a population guarantee.

still leaves one failure at $1/16$. Moreover, ranking improvement of the final endpoint does not establish that the newest update created it: refit-minus-anchor NDCG intervals span zero in both cohorts. Prospective calibration-aware admission remains untested. The original bank and new-acquisition results quantify complementary routes to recovery, with explicit availability, information, cost, and quality requirements.

5 Discussion and Conclusion

AMIYC establishes an empirical separation between bounded adaptation opportunity, replay compatibility, and realized recovery in trained MMoEs. Across three environments, controlled readouts identify current-state response and contrasting replay regimes. Already-trained coordinates and alternative same-boundary adapters broaden the construct beyond one parameterization. State witnesses improve selection from the same Amazon bank, while Kuai distinguishes missing bank value from selection error. The continuation studies further show persistent behavioral value and a reversal between native checkpoint ordering and common-anchor fitted ordering.

The trajectory and objective controls explain why these distinctions matter. Crossed and shadow comparisons associate the tested coupled deficit mainly with its base path. Same-base accounting credits retained and newly fitted operator value separately, while affine and ranking readouts distinguish loss response from quality-eligible recovery. These results support evaluating an evolving base together with the adaptations it can retain or reacquire, specifying the objective, information, new computation, and fallback.

The findings apply to the evaluated MMoEs, intervention families, and continuation regimes. Matched affine comparisons across competing bases, evaluation on untouched populations, and prospective calibration-aware admission would extend the present evidence. Within the tested contracts, native checkpoint quality, compatibility of prior corrections, and recovered value are distinct measurable properties of an evolving multi-task model.

AI USE STATEMENT

In this work, we used generative AI tools (including OpenAI ChatGPT) to help refine research hypotheses and conceptual framing; provide feedback on research methodology and experiment design; assist with implementation, code review, and artifact construction; support interpretation and reconciliation of experimental results; and assist with drafting, editing, structuring, and formatting the manuscript and supplementary material. We also used generative AI for brainstorming, literature search and summarization, and improving the readability of research text.

Generative AI was not treated as an authoritative source of empirical measurements. We reviewed reported numerical results against the available experiment outputs, frozen manifests, and reconciled evidence record. The authors retain responsibility for the final content of this work, including text, claims, code, analyses, and artifacts produced with the aid of generative AI.

REPRODUCIBILITY STATEMENT

The supplement specifies datasets, model boundaries, fitting and evaluation roles, seed units, candidate access, and computational budgets. The accompanying evidence registry links each result family and main-table entry to its protocol and source records. The scientific archive contains experimental source snapshots, matched transport rows, campaign numerical panels, and verification entrypoints; the clean manuscript package contains the active TeX, bibliography, figures, registry, and compiled PDFs. Appendix D.1.4 provides the artifact index and checking entrypoints. Original experiment records include checkpoint and split manifests, fitted operators, and the hash-verified Amazon bank-analysis files. Production aggregate results and interface contracts accompany the public-data studies; Production raw data and checkpoints remain access-restricted. Reported independent units and within-study role separation are explicit, including reuse of evaluation populations across the later studies.

REFERENCES

Samuel K. Ainsworth, Jonathan Hayase, and Siddhartha Srinivasa. Git Re-Basin: Merging models modulo permutation symmetries. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=CQsmMYmlP5T.

Ye Chen and Weining Zhang. UpgradeBench: A decision-centric benchmark for upgrading fine-tuned llm specialists. *arXiv preprint arXiv:2608.20918*, 2026. URL https://arxiv.org/abs/2608.20918.

Zhao Chen, Vijay Badrinarayanan, Chen-Yu Lee, and Andrew Rabinovich. GradNorm: Gradient normalization for adaptive loss balancing in deep multitask networks. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 794–803. PMLR, 2018. URL https://proceedings.mlr.press/v80/chen18a.html.

Farzad Farhadzadeh, Debasmit Das, Shubhankar Borse, and Fatih Porikli. LoRA-X: Bridging foundation models with training-free cross-model adaptation. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/1add3bbdbc20c403a383482a665eb5a4-Abstract-Conference.html.

Chongming Gao, Shijun Li, Yuan Zhang, Jiawei Chen, Biao Li, Wenqiang Lei, Peng Jiang, and Xiangnan He. KuaiRand: An unbiased sequential recommendation dataset with randomly exposed videos. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management*, pp. 3953–3957. ACM, 2022. doi: 10.1145/3511808.3557624. URL https://kuairand.com/.

Naibin Gu, Peng Fu, Xiyu Liu, Ke Ma, Zheng Lin, and Weiping Wang. Adapt once, thrive with updates: Transferable parameter-efficient fine-tuning on evolving base models. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long*

Papers), pp. 14765–14783. Association for Computational Linguistics, 2025. doi: 10.18653/v1/2025.acl-long.719. URL https://aclanthology.org/2025.acl-long.719/.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q. Weinberger. On calibration of modern neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1321–1330. PMLR, 2017. URL https://proceedings.mlr.press/v70/guo17a.html.

Shashank Gupta, Subhabrata Mukherjee, Krishan Subudhi, Eduardo Gonzalez, Damien Jose, Ahmed H. Awadallah, and Jianfeng Gao. Sparsely activated mixture-of-experts are robust multi-task learners. *arXiv preprint arXiv:2204.07689*, 2022. URL https://arxiv.org/abs/2204.07689.

Hussein Hazimeh, Zhe Zhao, Aakanksha Chowdhery, Maheswaran Sathiamoorthy, Yihua Chen, Rahul Mazumder, Lichan Hong, and Ed H. Chi. DSelect-k: Differentiable selection in the mixture of experts with applications to multi-task learning. In *Advances in Neural Information Processing Systems*, volume 34, pp. 29335–29347, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/f5ac21cd0ef1b88e9848571aeb53551a-Abstract.html.

Yupeng Hou, Jiacheng Li, Zhankui He, An Yan, Xiusi Chen, and Julian McAuley. Bridging language and items for retrieval and recommendation. *arXiv preprint arXiv:2403.03952*, 2024. URL https://amazon-reviews-2023.github.io/.

Gabriel Ilharco, Marco Tulio Ribeiro, Mitchell Wortsman, Suchin Gururangan, Ludwig Schmidt, Hannaneh Hajishirzi, and Ali Farhadi. Editing models with task arithmetic. In *International Conference on Learning Representations*, 2023. URL https://openreview.net/forum?id=6t0Kwf8-jrj.

Gyunyeop Kim and Sangwoo Kang. Exploration-driven reinforcement learning for expert routing improvement in mixture-of-experts language models. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pp. 23592–23605, Suzhou, China, November 2025. Association for Computational Linguistics. doi: 10.18653/v1/2025.findings-emnlp.1282. URL https://aclanthology.org/2025.findings-emnlp.1282/.

Jaehan Kim, Minkyoo Song, Seungwon Shin, and Sooel Son. SafeMoE: Safe fine-tuning for MoE LLMs by aligning harmful input routing. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/34d3cf97696022b179171e5abda42c0b-Abstract-Conference.html.

Zhongyang Li, Ziyue Li, and Tianyi Zhou. Routing manifold alignment improves generalization of mixture-of-experts llms. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/249b8d8f41970822651435629e68a6e1-Abstract-Conference.html.

Bo Liu, Xingchao Liu, Xiaojie Jin, Peter Stone, and Qiang Liu. Conflict-averse gradient descent for multi-task learning. In *Advances in Neural Information Processing Systems*, volume 34, pp. 18878–18890, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/9d27fdf2477ffbff837d73ef7ae23db9-Abstract.html.

Haokun Liu, Derek Tam, Mohammed Muqeeth, Jay Mohta, Tenghao Huang, Mohit Bansal, and Colin A. Raffel. Few-shot parameter-efficient fine-tuning is better and cheaper than in-context learning. In *Advances in Neural Information Processing Systems*, volume 35, 2022. doi: 10.52202/068431-0142. URL https://proceedings.neurips.cc/paper_files/paper/2022/hash/0cde695b83bd186c1fd456302888454c-Abstract-Conference.html.

Xiaodong Liu, Pengcheng He, Weizhu Chen, and Jianfeng Gao. Multi-task deep neural networks for natural language understanding. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 4487–4496, Florence, Italy, July 2019. Association for Computational Linguistics. doi: 10.18653/v1/P19-1441. URL https://aclanthology.org/P19-1441/.

Jiaqi Ma, Zhe Zhao, Xinyang Yi, Jilin Chen, Lichan Hong, and Ed H. Chi. Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 1930–1939. ACM, 2018. doi: 10.1145/3219819.3220007. URL https://doi.org/10.1145/3219819.3220007.

Filippo Rinaldi, Giacomo Capitani, Lorenzo Bonicelli, Donato Crisostomi, Federico Bolelli, Elisa Ficarra, Emanuele Rodolà, Simone Calderara, and Angelo Porrello. Update your transformer to the latest release: Re-Basin of task vectors. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 51688–51702. PMLR, 2025. URL https://proceedings.mlr.press/v267/rinaldi25a.html.

Filippo Rinaldi, Aniello Panariello, Giacomo Salici, Fengyuan Liu, Marco Ciccone, Angelo Porrello, and Simone Calderara. Gradient-sign masking for task vector transport across pre-trained models. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/59e4f5d009f95324a1a1aa9971b51e25-Abstract-Conference.html.

Hongyan Tang, Junning Liu, Ming Zhao, and Xudong Gong. Progressive layered extraction (PLE): A novel multi-task learning (MTL) model for personalized recommendations. In *Proceedings of the Fourteenth ACM Conference on Recommender Systems*, pp. 269–278. ACM, 2020. doi: 10.1145/3383313.3412236. URL https://doi.org/10.1145/3383313.3412236.

Runqian Wang, Soumya Ghosh, David Cox, Diego Antognini, Aude Oliva, Rogerio Feris, and Leonid Karlinsky. Trans-LoRA: Towards data-free transferable parameter efficient finetuning. In *Advances in Neural Information Processing Systems*, volume 37, 2024a. doi: 10.52202/079017-1957. URL https://papers.nips.cc/paper_files/paper/2024/hash/708fdc7911f11585ee7161518e509ae6-Abstract-Conference.html.

Zihan Wang, Deli Chen, Damai Dai, Runxin Xu, Zhuoshu Li, and Yu Wu. Let the expert stick to his last: Expert-specialized fine-tuning for sparse architectural large language models. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pp. 784–801, Miami, Florida, USA, November 2024b. Association for Computational Linguistics. doi: 10.18653/v1/2024.emnlp-main.46. URL https://aclanthology.org/2024.emnlp-main.46/.

Zhengxuan Wu, Aryaman Arora, Zheng Wang, Atticus Geiger, Dan Jurafsky, Christopher D. Manning, and Christopher Potts. ReFT: Representation finetuning for language models. In *Advances in Neural Information Processing Systems*, volume 37, 2024. doi: 10.52202/079017-2041. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/75008a0fba53bf13b0bb3b7bff986e0e-Abstract-Conference.html.

Prateek Yadav, Derek Tam, Leshem Choshen, Colin A. Raffel, and Mohit Bansal. TIES-Merging: Resolving interference when merging models. In *Advances in Neural Information Processing Systems*, volume 36, 2023. doi: 10.52202/075280-0310. URL https://papers.neurips.cc/paper_files/paper/2023/hash/1644c9af28ab7916874f6fd6228a9bcf-Abstract-Conference.html.

Youngsik Yoon, Siwei Wang, Wei Chen, and Jungseul Ok. When are experts misrouted? counterfactual routing analysis in mixture-of-experts language models. *arXiv preprint arXiv:2605.07260*, 2026. URL https://arxiv.org/abs/2605.07260.

Lijun Zhang, Tianbao Yang, Rong Jin, and Zhi-Hua Zhou. Dynamic regret of strongly adaptive methods. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 5882–5891. PMLR, 2018. URL https://proceedings.mlr.press/v80/zhang18o.html.

SUPPLEMENTARY MATERIAL

CONTENTS

A Foundational Evidence and Protocols **14**

A.1 Scientific roadmap and positioning 14

A.2 Experimental and Protocol Foundation 17

A.3 Observing the Task-Visible Additive Boundary 28

A.4 Structural Falsification and Measurement Boundaries 30

A.5 Frozen Intervention Response and Construct Validity 33

A.6 Lifecycle, Target Motion, and Replay Compatibility 37

A.7 Moving-Reference Tracking under Continued Optimization 46

A.8 Amortization, Bank Compatibility, and Deployment 51

A.9 Structured Action-Space Closure 57

B Recovery, Training Trajectories, and Objective-Dependent Value **60**

B.1 Experimental contracts and evidence locations 60

B.2 How much does local fitting recover, and does added complexity help? 61

B.3 Is fitted-operator motion evidence of representation motion? 61

B.4 Does the fitting remedy survive stronger and cross-domain controls? 62

B.5 Which training trajectory supports the stronger endpoint? 64

B.6 What limits admission, attenuation, and the value of more data? 66

B.7 Where does final gain come from, and what repeats on fresh weights? 68

C Interpretation and Scope **72**

C.1 Interpretation of State-Relative Adaptation 72

C.2 What the Controls Distinguish 76

C.3 Scope of the Current Conclusions 76

D Reproducibility and Evidence Navigation **77**

D.1 Reproducibility and Experimental Artifacts 77

D.2 Additional-study provenance and measured work 78

D.3 Cohort and Source Bindings 79

D.4 Scientific Evidence Map 79

A Foundational Evidence and Protocols

A.1 Scientific roadmap and positioning

AMIYC separates four questions that are easy to conflate but correspond to different estimands and experimental designs:

1. **State identification:** *What state is present* at the additive expert boundary?
2. **Current-state value:** *What bounded action carries value* at a frozen instance of that state?
3. **Replay compatibility:** *Does an action fitted elsewhere remain useful* after the underlying model state changes?
4. **Search amortization and deployment:** *Can repeated search be reduced* using only pre-intervention information and a finite available action bank, while respecting a separate secondary-quality constraint?

These questions map to the empirical levels that follow and should not be collapsed into a single notion of "repairability." The mature evidence hierarchy is

$$\begin{aligned} \text{observable state} &\longrightarrow \text{Fresh current-state opportunity} \\ &\longrightarrow \text{compatible available action} \\ &\longrightarrow \text{selection/abstention} \\ &\longrightarrow \text{secondary-quality eligibility.} \end{aligned} \tag{5}$$

The arrows mark successive evaluation stages, not logical implications. Structural observability does not prescribe an action; Fresh opportunity does not guarantee that a stale or foreign operator remains compatible; a compatible bank does not guarantee that a selector will recover its best action; and secondary-quality eligibility does not imply positive primary value.

The positive diagonal provides a bounded, identity-centered reference adapter for measuring current-state response and replay under a fixed contract. K1-D establishes residual response on coordinates trained with the base. A fixed, outcome-independent rotated basis and a prespecified rank-1 adapter extend the measurement across same-boundary families. These controls distinguish the measurable response from the choice of reference parameterization.

How to use this supplement. The foundational studies below establish the measurement surface, construct validity, replay behavior, and frozen-bank recovery. Appendix B then tests stronger fitting, training paths, admission, and objective-dependent value. Interpretation follows all experiments in Appendix C; the source contract and full evidence ledger are in Appendix D. Campaign identifiers locate source studies; the scientific questions determine the reading order.

Reader's question	Evidence location
What is measured, on which data and with which information?	A.2–A.4
Does bounded opportunity survive the construct controls?	A.5
When do replay and continued adaptation retain value?	A.6–A.7
What can a frozen bank recover, and what are its action limits?	A.8–A.9
What changes under stronger fitting and trajectory controls?	B.2–B.5
What survives admission, attribution, calibration and fresh weights?	B.6–B.7

A.1.1 Related work and positioning

AMIYC intersects several literatures that establish neighboring pieces of the problem but generally study different empirical objects. We group the closest work by the question it answers rather than by model family alone.

Multi-task optimization and task-specific residual value. Multi-task learning does not imply that the jointly selected model is individually saturated for every constituent task. MT-DNN, for

example, performs joint multi-task refinement and then further fine-tunes the resulting model on individual target tasks (Liu et al., 2019). Gradient-balancing and multi-objective methods instead modify the joint optimization trajectory to reduce task imbalance or conflicting progress (Chen et al., 2018; Liu et al., 2021). More directly for sparse expert models, Gupta et al. train task-aware mixture-of-experts models jointly across tasks and subsequently fine-tune the multi-task model on individual target datasets (Gupta et al., 2022). These studies establish post-training specialization. AMIYC investigates how its accessible value depends on the learned state and on the information available for retaining or refitting an adaptation.

Recommendation-oriented MTL work studies a related problem from the architecture side. MMoE uses task-specific gates over shared experts to model task relationships (Ma et al., 2018); PLE separates shared and task-specific extraction pathways to reduce negative transfer and the "seesaw" behavior of competing objectives (Tang et al., 2020); and DSelect-k replaces conventional top-k gating with differentiable sparse expert selection, including large-scale recommendation experiments (Hazimeh et al., 2021). These methods ask how experts should be shared, selected, or optimized during training. AMIYC measures opportunity and compatibility at frozen readouts of the task-visible expert state, holding routing fixed during each intervention.

Post-training adaptation and routing in trained MoEs. A second line of work shows that already-trained MoEs can benefit from restricted post-training adaptation. Expert-Specialized Fine-Tuning identifies experts most associated with a downstream task and updates those experts while freezing the remaining experts and modules (Wang et al., 2024b). Exploration-driven routing methods explicitly evaluate alternative expert paths and use their performance to improve the router (Kim & Kang, 2025). Recent analyses are even closer to the static AMIYC question: Yoon et al. show that a frozen trained MoE can contain lower-loss equal-compute routing alternatives and that a small router-only update can recover downstream value (Yoon et al., 2026); RoMA likewise improves downstream behavior through router-only post-training while leaving the remaining model fixed (Li et al., 2026). SafeMoE further demonstrates that routing decisions themselves can drift as an MoE is fine-tuned and regularizes the evolving router toward an earlier reference state (Kim et al., 2026).

Our question concerns a specified bounded response at a learned MMoE state, after task supervision and the relevant model coordinates have participated in ordinary training. Freezing routing during the readout makes that response comparable across current-state fitting, operator replay, and finite-bank selection.

Parameter-efficient adaptation as method versus state measurement. Adapters, low-rank updates, prompt tuning, bias-only tuning, and related parameter-efficient methods establish broadly that substantial task adaptation can be obtained while most pretrained parameters remain frozen. MoE-specific variants further exploit expert specialization or routed adaptation modules. AMIYC uses restricted adaptation to measure how the accessible response changes across states and recovery contracts. In our experiments, the positive-diagonal family is deliberately low-reach, identity-centered, and replayable because it serves as a *measurement instrument*. Its held-out response operationalizes a state-relative quantity under a declared intervention family and FIT–SELECT–EVAL protocol. K1-D makes this distinction especially important: the corresponding diagonal coordinates are already present and optimized from initialization, yet a newly fitted current-state response remains after the selected model is frozen. This establishes residual response on trained coordinates. Because the residual multiplies the trained diagonal, it does not establish an identical composite coefficient bound or an equal-budget comparison with further joint training. Activation rescaling and frozen-representation interventions themselves have prior art in $(IA)^3$ and ReFT (Liu et al., 2022; Wu et al., 2024).

Update portability and changing base models. A separate literature studies whether a task-specific update learned on one base model remains useful after the base changes. GradFix treats update compatibility as target-state dependent and uses the gradient geometry of the destination model to determine which components of a previously learned task vector remain appropriate (Rinaldi et al., 2026). UpgradeBench studies the same problem longitudinally across model releases and continued-pretraining checkpoints, showing that a specialization asset can lose portability as the underlying base model evolves, including within known training lineage (Chen & Zhang, 2026). Trans-PEFT instead trains update-resistant modules for evolving language models (Gu et al., 2025);

Trans-LoRA uses synthetic data to transfer an existing adapter (Wang et al., 2024a); and LoRA-X transfers across base models using a subspace-constrained adapter (Farhadzadeh et al., 2025). These are relevant portability methods with different representations, target information, and evaluation objectives, rather than matched controls for the fixed task-visible MMoE operator measured here. Related task-vector transport and model-rebasing work likewise treats a learned parameter delta as an object whose behavioral meaning depends on the representation frame in which it is applied. TIES-Merging resolves sign conflicts when combining task vectors from separate models (Yadav et al., 2023). Its merging objective differs from our single-source, target-state replay comparison; we do not report a matched TIES port in these experiments.

Target-fitted references in Trans-PEFT, UpgradeBench, and GradFix provide context for the state-controlled comparisons here. AMIYC adds controls during endogenous MMoE training: response on already-trained coordinates and high-quality checkpoints, operator motion without necessary behavioral decay, and fixed-state supervision versus online tracking. A successful transport rescue would change recovery from the tested bank without invalidating its raw replay measurements or the current-state response.

Task arithmetic treats learned updates as reusable vectors (Ilharco et al., 2023); Git Re-Basin and TransFusion address their coordinate dependence through permutation alignment (Ainsworth et al., 2023; Rinaldi et al., 2025). Our restricted expert and coordinate matching tests one such explanation, not all rebasing methods. TIES could combine aligned or transported donors to create new bank actions; our availability results concern the specified unmerged bank, not that enriched family. Trans-LoRA, LoRA-X, and Trans-PEFT additionally change data access, source fitting, or representation constraints, so they are related solutions rather than interchangeable zero-step controls.

State evolution, replay, and tracking. The longitudinal experiments also connect to the broader idea of tracking a changing reference. We use this language descriptively rather than claiming a new theory of nonstationary or online optimization. Dynamic-regret theory already formalizes comparison with moving references (Zhang et al., 2018); its guarantees do not directly apply to our endogenous nonconvex state changes. The source of motion here is endogenous: the task and evaluation protocol can remain fixed while the base representation produced by ordinary multi-task training changes, and the bounded high-value response changes with it. The relevant empirical comparison is therefore not simply whether an old update decays with age, but whether it remains behaviorally compatible with the *current* learned state.

This distinction is sharpened by the optimization controls. When the evolved representation is frozen, ordinary multi-task supervision moves the exposed coefficient coordinates toward the contemporaneous Fresh response. Under the two tested online conditions, however, the same coordinates do not track that reference while the representation co-evolves. We therefore describe the observation as a moving-reference/state-evolution tracking difficulty under the tested schedules, without identifying a unique optimizer, timescale, conditioning, or update-order mechanism.

From current opportunity to operational recovery. Finally, modular adaptation and expert-routing work often assumes that a useful module, route, or update is available to be selected. AMIYC separates that assumption from the current-state measurement itself. Fresh fitting estimates bounded opportunity after state-specific search; replay asks whether an operator fitted elsewhere remains compatible; a frozen action bank asks whether any compatible operator is actually available; and the learned policy asks whether pre-intervention information can identify or abstain from those actions. Independent secondary-quality eligibility is evaluated separately from primary value. Thus a state can exhibit Fresh opportunity even when no useful bank action remains, and an eligible action can still have nonpositive primary value.

Positioning of the present work. AMIYC characterizes state-relative adaptation value inside an evolving trained MMoE. A bounded current-state response is re-estimated at successive learned states; previously fitted operators are replayed against those contemporaneous references; and opportunity, compatibility, bank availability, selection, abstention, and secondary eligibility are measured separately. Construct, trajectory, and calibration controls test explanations for their divergence.

This positioning determines the role of the major experimental components. Structural diagnostics are observations of state rather than harm labels; bounded adapters are measurement instruments

rather than asserted mechanisms; Fresh is a finite protocol-defined current-state reference rather than a global optimum; Same and Foreign are provenance labels rather than compatibility classes; and deployment experiments ask whether measured opportunity remains operationally recoverable rather than assuming that successful current-state fitting directly supplies a reusable controller.

Terminology used throughout. We use four conventions consistently.

1. *Fresh* means the result of the fixed state-specific FIT–SELECT procedure at a frozen target checkpoint, with authoritative behavior measured only on EVAL. Fresh is a finite protocol-defined current-state reference, not a unique or global mathematical optimum.
2. *Same* and *Foreign* describe source provenance for replay, not compatibility classes. Same-lineage replay is often locally useful on Amazon and Kuai, but Production shows that it can become harmful, useful, and harmful again across descendant checkpoints. The operational object is *current-state compatibility*, not lineage membership or checkpoint age.
3. *Witness* denotes a pre-intervention state descriptor supplied to an outcome predictor. Witnesses are observational coordinates; they are not causal harm labels, sufficient state variables, or globally invariant coordinates across independently trained model realizations.
4. *Eligibility* is an empirical secondary-quality constraint. It is not a formal safety guarantee and is not a substitute for positive primary-objective value.

A.2 Experimental and Protocol Foundation

This section fixes the contracts under which all later results are interpreted. The project contains many fitted candidates and action evaluations but far fewer independent learned model states; those counts are not interchangeable. We therefore keep FIT, SELECT, EVAL, and ALIGN roles explicit, distinguish frozen-state effects from continued-training effects, and use seeds or lineages as the inferential units for cross-trajectory claims.

A.2.1 Datasets, tasks, models, and checkpoints

We use two public recommendation settings and one production ranking setting. Their roles are deliberately complementary rather than three repetitions of the same benchmark.

Amazon Reviews: four-category mixture. We construct Mixed-4Cat from Amazon Reviews 2023 (Hou et al., 2024), pooling Beauty and Personal Care, Sports and Outdoors, Toys and Games, and Grocery and Gourmet Food. Mixed-4Cat names this study's derived four-category mixture. The documented preprocessing selects up to 500,000 events per category from January 2022 through September 2023, with history length 20, minimum user sequence length 5, future horizon 3, and ranking lists containing one positive and 19 negatives. The resulting dataset contains approximately 2.0M recorded events yielding 946.9K ranking examples, 677.5K users, and 546.0K items. The public task names are *Rank*, *Next*, and *Future*: Rank is a listwise immediate-next ranking objective, Next is pointwise discrimination of the same immediate target, and Future is a short-horizon multi-target intent objective. The execution code keys these heads as `ranking`, `trx`, and `future_intent`, respectively; `trx` is the implementation name for the immediate-target Next head. The processed training and validation splits contain 767,453 and 89,747 ranking examples, respectively.

The Amazon MMoE has four experts, a 128-dimensional expert mixture, and a 64-dimensional exposed task projection in the recovered intervention backend. It supplies the largest controlled family of structural diagnostics, frozen actions, scope/capacity controls, continued-optimization experiments, and amortized action-selection studies. Seven independent seeds support the mature cross-seed analyses.

KuaiRand. KuaiRand-Pure contributes approximately 1.44M interaction examples from 27.3K users and roughly 7.5K items. Its MMoE has four experts, a 128-dimensional expert mixture, and a 64-dimensional exposed task projection. Eight independently initialized trajectories are densely observed and are used for lifecycle alignment, Fresh-operator motion, Same/Foreign replay, diagonal-from-initialization and normalization controls, and the transfer boundary of witness-conditioned selection.

The lifecycle coordinate is defined relative to each task trajectory's own validation-AUC peak. This is an empirical alignment device, not a causal phase boundary. The dense trajectory allows four questions to be separated: whether Fresh headroom exists, how large it is, how the underlying checkpoint quality changes, and whether a previously fitted operator remains compatible.

Production ranking system. The restricted production setting supplies an anonymized ranking-model replication. Operating volumes and unreleased architectural details are omitted; only the experimental intervention interface and aggregate responses are reported. Three independently initialized lineages are fixed in advance. Each contributes its own epoch-6 source checkpoint and descendant targets at $+1$, $+3$, and $+5$ epochs.

Production is a targeted replication rather than a third full benchmark of every AMIYC component. Its role is specific: Fresh adaptation remains positive at all nine predeclared target states, whereas Same replay changes sign across horizons and raw Foreign replay remains near zero. Thus the current-state headroom phenomenon survives the larger setting, while universal same-lineage persistence does not.

Table 5: Datasets, models, checkpoint structure, and scientific roles. Workload counts use each setting's native model-level unit; they should not be added across settings. FIT/SELECT/EVAL sizes are reported numerically only where a single frozen panel is authoritative; Amazon and Kuai contain multiple experiment-specific role populations under the common disjoint-role contract.

Setting	Data / task surface	Model / independent units	Checkpoints and role populations	Scientific role
Amazon Mixed-4Cat	946.9K ranking examples derived from about 2.0M recorded events; 677.5K users; 546.0K items. Heads: *Rank* (listwise ranking), *Next* (immediate-target discrimination, labeled `trx` in experiment logs), and *Future* (short-horizon intent).	35.28M parameters; 3 tasks $\times$ 4 experts $\times$ 128-D expert mixture / 64-D exposed projection. Seven independent seeds support mature cross-seed analyses.	Controlled frozen checkpoints plus matched evolved continuation states. FIT, SELECT, and EVAL are experiment-specific disjoint populations; EVAL is held out from intervention fitting/selection; parent development data can be reused for checkpoint selection. Amazon alignment uses a separate pre-intervention FIT/ALIGN population disjoint from EVAL. Historical Phase-II witness construction used a common $N_{\mathrm{probe}} = 20{,}000$ validation population per epoch, but this is *not* a universal intervention split size.	Primary controlled setting: structural diagnostics and falsification; intervention ladder; scope/capacity; same-boundary controls; continuation/tracking; composition/codec; witness-conditioned amortization and deployment decomposition.
KuaiRand-Pure	Approximately 1.44M interaction examples; 27.3K users; roughly 7.5K items. The two binary task heads in the recovered contract are *engagement* and *long_view*; their primary loss is binary cross-entropy and secondary metric is AUC.	28.6M parameters; 2 tasks $\times$ 4 experts $\times$ 128-D expert mixture / 64-D exposed projection. Eight independently initialized trajectories.	Dense lifecycle anchors include epochs $\{1, 11, 12, 13, 14, 21\}$, aligned descriptively to each task's own validation-AUC peak. Final amortization uses leave-one-seed-out outer evaluation with complete held-out checkpoints for threshold calibration; outer-test labels do not fit predictor or threshold. Frozen Fresh/replay analyses retain disjoint FIT/SELECT/EVAL roles under their experiment manifests.	Dense lifecycle and target-motion setting; pre/near-peak audit; Same/Foreign replay; K1-D and K1-N construct-validity controls; similarity/common-support boundary; negative-transfer boundary for fine-grained witness amortization.

Table 5 continued.

Setting	Data / task surface	Model / independent units	Checkpoints and role populations	Scientific role
Production ranking system	Restricted ranked-list data; three anonymized task heads. Operating volumes and business labels are withheld.	Restricted multi-task ranking architecture; three independently initialized lineages.	Each lineage supplies epoch-6 source plus predeclared descendants e7/e9/e11 ($+1, +3, +5$), giving 12 frozen checkpoints total. One label-blind frozen panel is fixed as FIT 15K, SELECT 5K, and untouched EVAL 100K.	Targeted scale replication of Fresh/Same/Foreign and compatibility: Fresh positive in all nine target states; Same harmful/useful/harmful across horizons; Foreign near zero. Also supplies the eligibility-versus-primary-usefulness boundary.

A.2.2 State, pair, operator, and action semantics

For task k, let $e_i(x)$ denote expert i's output and let $g_{k,i}(x)$ denote the corresponding router weight. The routed expert contribution is

$$c_{k,i}(x) = g_{k,i}(x)e_i(x). \tag{6}$$

Let P_k denote the frozen linear projection exposed by the model diagnostics. The task-visible contribution is

$$u_{k,i}(x) = P_k c_{k,i}(x). \tag{7}$$

The additive state presented through this boundary is therefore

$$u_k(x) = \sum_i u_{k,i}(x). \tag{8}$$

We use the following vocabulary consistently.

- A *state* is a frozen model checkpoint or a matched evolved continuation checkpoint.
- A *pair* is an unordered expert support (i, j).
- A *direction* specifies which contribution is transformed in a directed pair-local action.
- An *operator* contains fitted numerical parameters together with the source state and provenance needed to interpret them.
- An *action* specifies how an operator is applied, including target task, pair/support, direction, scope, dose, composition, and source provenance.

This factorization is empirical, not bookkeeping. Pair identity is not the intervention. The same source coefficient vector can behave differently when its target state, scope, dose, task, or support changes. Conversely, different nominal actions can realize nearly identical held-out behavior. This distinction later motivates separate reporting of exact action agreement, operator replay, bank availability, and realized policy value.

A.2.3 Bounded positive-diagonal reference family

The primary reference family applies a bounded positive diagonal transform in the task-visible coordinate system:

$$D(z) = \mathrm{diag}(e^{z_1}, \ldots, e^{z_d}), \qquad D(0) = I. \tag{9}$$

For a directed pair-local action on expert i, the requested projected change is $[D(z_{\mathrm{eff}}) - I]u_{k,i}$, where dose is bounded as specified below. It is lifted into the raw mixture using P_k^+ and passed through the unchanged task computation. Equation 2 gives the exact forward; direct replacement of one projected contribution requires the stated range/rank condition. The operator dimension is set by the model's exposed projection.

The requested positive scaling preserves coordinate signs; its realized projected effect remains subject to the lift condition above. The family is useful because it is bounded, identity-centered, sparse in reach under pair-local support, easy to replay exactly, and comparatively conservative on the independent ranking metric in the Amazon intervention ladder. These properties motivate its use as a reference family; they do not make it the unique adaptation family or identify it as the mechanism behind residual headroom.

Fresh fitting. At target state s, SELECT retains an epoch and, where searched, a direction; EVAL measures that frozen candidate. Kuai fits both directions from identity; Amazon's matched target re-oracle instead retains the source direction and initializes from its coefficients. The executable contracts below distinguish these finite procedures. Fresh therefore returns one protocol-defined, state-specific representative within the tested family.

Replay. A Same replay applies an earlier source operator unchanged to a descendant state from the same training lineage. A Foreign replay applies an operator fitted on an independently initialized lineage. An aligned Foreign replay first applies only a prespecified, FIT-only family-preserving expert/coordinate permutation and then replays the transported diagonal unchanged. These are provenance labels; none is itself a claim of compatibility.

A.2.4 Executable operator and fitting contracts

The implementation distinguishes the raw routed mixture from its projected coordinates. Write $r_k = \sum_i c_{k,i}$, $u_{k,i} = P_k c_{k,i}$, and let P_k^+ be the Moore–Penrose pseudoinverse of the frozen projection exposed by the model diagnostics. In the recovered Amazon and Kuai implementations, $r_k, c_{k,i} \in \mathbb{R}^{128}$, $P_k \in \mathbb{R}^{64\times 128}$, and the positive diagonal has 64 coefficients. Thus 128 is the expert/mixture width; 64 is the projected intervention width. These dimensions should not be interchanged.

For the source expert selected by direction, the executed intervention is

$$r'_k = r_k + P_k^+ \big[D(z_{\text{eff}}) - I\big] P_k c_{k,i}, \qquad f'_k = T_k(r'_k), \tag{10}$$

where T_k is the original task computation for this backend. Consequently,

$$P_k r'_k = P_k r_k + P_k P_k^+ \big[D(z_{\text{eff}}) - I\big] u_{k,i}. \tag{11}$$

Direct diagonal scaling at the projected boundary is the full-row-rank special case, or holds when the requested projected delta lies in the range of P_k. The lifted expression is the unconditional execution definition. If P_k is the first tower weight, the original affine bias remains unchanged inside the tower. Identity gives exactly zero delta regardless of rank; baseline reproduction alone therefore does not certify the rank condition. The supplied intervention archives do not provide an all-checkpoint projection-rank audit. Architectural controls use their documented, model-specific boundary backends.

Direction, scope, and bounds. The code convention `i_to_j` transforms contribution i; `j_to_i` transforms contribution j. It does not transfer router mass between experts. For the task-global control, $c_{k,i}$ in Eq. 10 is replaced by r_k. The bounded source parameterization is $z = \gamma \tanh(\rho)$, $D(z) = \operatorname{diag}(\exp z)$, with $\gamma = 0.5$ in the recovered Amazon and Kuai contracts. The bound is on projected gains; raw displacement also depends on P_k^+. Amazon's fixed-dose replay wrapper reconstructs its raw parameter through

$$z_{\text{eff}} = \gamma \operatorname{clip}(\alpha z/\gamma, -0.999999, 0.999999)\,, \tag{12}$$

up to floating-point roundoff. It therefore equals αz only before saturation. In the matched GradFix experiment $\alpha = 1$; every archived source coordinate is inside this clipping threshold (maximum absolute log coordinate 0.3882531 for Amazon and 0.4858689 for Kuai).

Why the companion expert enters fitting. Although application modifies only the directed contribution, pair fitting uses both experts. With validity mask m_n, define $w_n = g_{k,i}(x_n) g_{k,j}(x_n) m_n$,

$v_n = u_{n,k,i} + u_{n,k,j}$, and $v'_n = u'_{n,k,i} + u'_{n,k,j}$. The recovered regularizers are

$$R_I = d^{-1} \sum_{q=1}^{d} (\exp z_q - 1)^2, \tag{13}$$

$$R_E = \frac{\sum_n w_n \left[\frac{\|v'_n\|^2 - \|v_n\|^2}{\max(\|v_n\|^2, \epsilon)} \right]^2}{\max(\sum_n w_n, \epsilon)}, \tag{14}$$

$$R_C = \frac{\sum_n w_n \max(0, -\langle u'_{n,k,i}, u'_{n,k,j} \rangle)}{\max(\sum_n w_n, \epsilon)}, \tag{15}$$

with $\epsilon = 10^{-12}$ and fitting objective $L_{\text{FIT}} + \lambda_I R_I + \lambda_E R_E + \lambda_C R_C$. The tensors denoted u' here are the requested transformed projected contributions used by the regularizer; their realized projected effect is given by Eq. 10. The companion expert therefore affects the fitted operator through pair energy and exposure, even when a fixed operator's applied delta depends only on the directed source expert.

Fresh is a source-specific finite protocol. The recovered Kuai primary oracle starts each direction at $\rho = 0$ and fits both legal directions using AdamW (learning rate 0.003, weight decay zero), 128-example minibatches, gradient clipping at 5, at most 100 epochs, patience 12, and minimum improvement 10^{-6}. It uses $(\lambda_I, \lambda_E, \lambda_C) = (0.01, 0.05, 0)$. Within a direction, the retained epoch minimizes the regularized SELECT objective. Direction selection then minimizes $L_{\text{SELECT}} + 0.05|R_{\text{pair}} - 1|$, where $R_{\text{pair}} = \sum_n w_n \|v'_n\|^2 / \max(\sum_n w_n \|v_n\|^2, \epsilon)$. The nominal split fractions are $0.30/0.10/0.60$, with split seed 20260715; the matched Kuai cohort has 6,140/2,025/11,835 actual examples. Identity initialization is not a separately retained candidate among the post-epoch SELECT candidates. Neither positive held-out gain nor dominance over Raw is guaranteed by this procedure.

Amazon's matched replay cohort uses an archived target re-oracle initialized from the original source operator, retaining its direction. Its target split seed is 20260723 and the native units are lists. This differs from the identity-started, two-direction source-oracle search and from Kuai's target Fresh fit. All 378 available historical target-profile records report 100 epochs, 300 optimizer steps, and retained epoch 99. The source runner supports the bounded parameterization and regularized objective above, but its CLI defaults are not evidence that every experiment used one universal recipe. Population sizes, initialization, selection, continuation, and scope must be read with the experiment-specific source contract.

For the 288 matched Kuai rows, archived health fields report zero nonfinite gradient steps, zero clipped-step fraction, and zero fraction at the bound; the largest recorded preclip norm is 0.082764. Amazon's compact historical profiles do not supply equivalent gradient-health fields. These records support the stated finite computations, not global optimality or a common convergence certificate. Historical Amazon worker time also includes loading, probe construction, and evaluation, and is not isolated optimization time.

Decision roles and the 20-unit transport budget. GradFix retains the source support and direction and uses unregularized native task-loss derivatives at the identity in effective log coordinates. Its fixed majority/normal mask takes no target optimizer steps and consumes no target SELECT outcomes. Amazon uses 20 complete lists, each with one positive and 19 negatives (400 item labels), with one mean masked-BCE gradient per list. Kuai uses 20 scalar examples, ten per class. Equal vote counts are not equal labeled-data or compute budgets. The deterministic sample reuse described in Appendix A.6.9 limits conclusions about alternative FIT resampling; it does not add independent replications. Kuai computes Raw EVAL parity before mask construction, but passes no such outcome to the mask function; Amazon freezes its masks before its new EVAL phase. The justified restriction is no EVAL information in the transport decision, rather than an identical execution ordering in both runners. These archived role separations do not establish that every choice in the wider research development process was preregistered.

Anonymous numerical release. The accompanying audit release supplies 288 Kuai and 378 Amazon curated matched rows, seed aggregates, anonymized sign-sufficient operator evidence,

source fingerprints, and a Python-standard-library verifier. It reconstructs all 666 masks, reaggregates the table means, and enumerates exact seed-level sign-flip tests. Source means and p-values are checked; archived bootstrap intervals are identified separately rather than described as newly generated. The numerical audit operates on anonymized records and sign-sufficient statistics. Original opportunity and transport cohorts retain separate baseline identities, as specified in Appendix A.5.

A.2.5 FIT–SELECT–EVAL

Individual frozen-state candidate fits use the following three-stage contract.

1. **FIT**: optimize candidate intervention parameters using only the fitting population.
2. **SELECT**: choose early stopping, direction, or candidate using a separate selection population.
3. **EVAL**: evaluate the frozen selected action on an intervention-held-out population.

EVAL is never used to fit an operator, choose its direction, align a Foreign operator, calibrate a learned-policy threshold, or construct target-state features. When FIT/SELECT searches over a finite candidate set, EVAL measures the behavior of the resulting search-and-selection procedure. Candidate rows are not independent replicates. Two summaries are explicitly retrospective: the bank maximum uses EVAL to describe availability, and the archived K1 best-eligible-pair summary maximizes EVAL over six fitted pairs. Neither is an executable selector or an unbiased estimate of population-optimal value. The main K1 result instead averages all six pairs. Base-checkpoint selection can reuse the parent development population later partitioned for intervention evaluation; disjoint intervention roles do not make these fresh-test estimates of the full training pipeline.

For the primary objective,

$$\Delta L(s, a) = L(s, \text{no-op}) - L(s, a), \tag{16}$$

so positive values denote improvement. For a prespecified secondary-quality tolerance ϵ,

$$U_\epsilon(s, a) = \Delta L(s, a)\, \mathbf{1}[\Delta Q(s, a) \geq -\epsilon]\,, \tag{17}$$

with $U_\epsilon(s, \varnothing) = 0$. For a finite available bank $\mathcal{B}$, define the best eligible bank value

$$G^\star_{\mathcal{B}}(s, k) = \max\left\{0,\ \max_{a \in \mathcal{A}_{\mathcal{B}}(s,k)} U_\epsilon(s, a)\right\}. \tag{18}$$

This is an eligibility-gated score: an ineligible executed action receives score zero even if its raw primary effect is negative. Gating does not undo execution or implement an outcome-informed no-op. Ungated primary effects, secondary changes, and violation rates must be evaluated separately. The tolerance is an empirical operating constraint, not a statistical or causal safety guarantee.

A.2.6 Inferential units and search-aware evaluation

Pair rows, directions, doses, and candidate actions drawn from the same checkpoint share the same learned representation state and are not independent model realizations. Seed or lineage is therefore the inferential unit for cross-trajectory claims, with checkpoint-level pairing preserved for within-trajectory contrasts. Seed-level bootstrap and paired intervals preserve that grouping. They are pointwise, conditional on the reported cohorts and analysis choices, and are not simultaneous confidence bounds across the research program. The Amazon selected-tail Clopper–Pearson bound is an iid-panel reference calculation, not cluster-aware coverage; its distinction is explicit in Appendix A.8.6.

The historical read-only search/selection audit classified its reviewed core contract as passed: FIT, SELECT, and EVAL are disjoint; preprocessing is nested inside the appropriate outer training split; and candidate rows are not used as inferential replicates. The expressive-family ladder reports each family's measured response. Matched displacement and capacity comparisons use the dedicated scope and same-boundary controls in Appendix A.5; the full ladder is not an all-family matched-budget comparison.

Table 6: Protocol, search variables, inferential units, and estimands for the principal AMIYC result families. "Search" describes legal FIT/SELECT choices; it does not turn candidate rows into independent statistical replicates.

Result family	FIT / SELECT / EVAL / ALIGN contract	Search / calibration	Inferential unit	Estimand and audit boundary
Structural AF falsification	Structural regularization is applied during training; the targeted interaction and held-out task metrics are measured on matched checkpoints/populations. This is not a Fresh adapter fit.	Prespecified AF strength, weighting, and detach variants; strongest claims use the matched stress sweep plus untreated-training controls.	Independent trained runs/ checkpoints; candidate structural measurements are descriptive rows.	Tests whether manipulating a plausible structural quantity produces monotonic behavior. It does not define an intervention-value objective.
Fresh frozen adapter	FIT optimizes bounded adapter parameters using only FIT. SELECT chooses early stopping/direction/candidate on a separate role. The selected frozen action is evaluated on intervention-held-out EVAL.	Legal direction and fitting trajectory/early stopping; action family and coefficient bounds are prespecified by experiment.	Seed/lineage for cross-trajectory claims; matched check-point/task/pair for within-state contrasts.	Frozen-state conditional adaptation response. Fresh is a finite protocol-defined current-state reference, not a mathematical ceiling or online oracle.
Intervention ladder / scope–dose	All compared frozen actions use matched base checkpoint and held-out EVAL; family-specific parameters are fit without EVAL. Matched-displacement analyses use realized displacement from the legal comparison set, not EVAL-driven family invention.	Family, pair/direction, scope, nominal dose; selected matched-displacement comparisons are prespecified/read-only from compatible outputs.	Seed-level paired contrasts; action rows remain clustered within state.	Compares response under action families/reach. Archived expressive families lack a complete all-family matched-displacement record; R0 explicitly retains that missing theorem as a boundary.
K1-D / K1-N construct validity	Base model is trained under the declared architectural control; checkpoint selection uses its development criterion. The selected model is then frozen and Fresh FIT–SELECT–EVAL is rerun.	Checkpoint selection is predeclared; no outcome-driven architecture sweep. K1-D exact diagonal degrees are present from initialization; K1-N affine LayerNorm is the fixed normalization alternative.	Eight Kuai seeds; task-seed blocks are descriptive subunits.	Tests parameter availability and one normalization alternative. K1-N is bounded by base-quality mismatch; K1-D measures residual response on trained coordinates; dev-AUC equivalence and the identical composite bound are not established. Mean-pair response is descriptive; historical best-pair summaries maximize EVAL.
A-Boundary rank-1 / rotated controls	Same Amazon boundary and data roles as the mature frozen adapter audit. Alternative adapters are fit/select/evaluated without target EVAL leakage.	One prespecified rank-1 family and one fixed outcome-independent rotated basis; realized displacement is matched to native diagonal where reported.	Seven-seed paired contrasts over 63 states.	Tests family/basis specificity. Rank-1 is not parameter-count matched; rotated result does not imply universal rotation invariance.
Kuai lifecycle / pre-peak audit	Fresh and replay outcomes are recomputed only on supported target states; lifecycle alignment uses each task's validation-AUC peak from development history.	No causal phase selection. Pre/near-peak subset is defined by $\ell \leq 0$; post-peak quality model uses degradation and epoch distance descriptively.	Eight seeds, with task/checkpoint pairing retained.	Separates existence of headroom from checkpoint quality. Does not establish adapted-late $>$ earlier-best on a common EVAL frame.

Table 6 continued.

Result family	FIT / SELECT / EVAL / ALIGN contract	Search / calibration	Inferential unit	Estimand and audit boundary
Same / Foreign replay	Source operator is already FIT/SELECT frozen. Target-state replay applies it unchanged; target EVAL is used only to measure realized replay outcome. Fresh target fit is a separate contemporaneous reference.	Source provenance and horizon are fixed by experiment; no target-state refitting for replay.	Seed/lineage; matched target-pair units inside lineage.	Measures current-state replay compatibility, not lineage causality or a monotonic staleness law.
Amazon permutation alignment	Donor/target expert and coordinate mapping is chosen using pre-intervention FIT/ALIGN data disjoint from EVAL. Intervention outcomes and EVAL labels are excluded from the mapping.	Hungarian expert matching from task-visible similarity, then coordinate matching from FIT-side correlation; donor diagonal is only permuted, not refit.	Seven seeds; matched target-pair relations.	Tests family-preserving permutation explanation only. Does not test unrestricted rotations/gauges.
Similarity / nearest-Foreign audit	Donor ranking uses only pre-intervention similarity descriptors; target replay EVAL remains untouched.	Amazon: FIT-only matched-expert CKA + matched-coordinate correlation. Kuai: 920-D pre-intervention representation descriptor distance.	Amazon seven seeds; Kuai eight seeds; matched relation panels.	Tests sufficiency of measured similarity. Kuai Same/Foreign supports do not overlap, so residual causal lineage after exact common-support matching is not identified.
Production targeted replay	Frozen predeclared panel: FIT 15K, SELECT 5K, EVAL 100K. Fresh is target-state fit; Same/Foreign are unchanged source replay. No learned selector or alignment is fitted.	Source e6 and targets e7/e9/e11 are fixed before outcomes; full 28-pair target census is aggregated within target state before lineage-level interpretation.	Three independent lineages; horizon pattern descriptive.	Targeted replication of current-state headroom and replay compatibility, not a third full benchmark of every AMIYC component.
Matched continuation	Starting state and continuation schedule are matched. No-op is an evolved ordinary-MTL branch, not the original source checkpoint. Selected/reference interventions are carried through continuation and evaluated at matched endpoints.	Condition assignment is prespecified; source selected/reference action already frozen.	Seven Amazon seeds with matched starting checkpoints.	Continued-trajectory effect. Raw continuation gain is not interchangeable with frozen Fresh gain.
Historical joint / online block-coordinate tracking	Base and coefficients evolve online under ordinary MTL; after the endpoint is reached, the endpoint is frozen and a contemporaneous Fresh reference is fit independently.	Historical joint versus one prespecified same-minibatch block-coordinate update schedule; no optimizer/schedule sweep.	Seven seeds; 63 matched evolved states supply geometry rows.	Endpoint tracking geometry relative to Fresh. Negative result rejects simple update-ordering rescue but does not identify the dynamical cause.
Frozen-base coefficient-only MTL	At each evolved state the base representation is frozen; only the same coefficient coordinates are optimized under ordinary MTL, then compared with independently fit Fresh.	Coefficient-only optimization budget fixed by protocol; no base updates.	Seven seeds / matched evolved states.	Fixed-state signal control. Positive alignment rejects intrinsic absence of local task signal; post-hoc recovery ratio is not a causal variance decomposition.

Table 6 continued.

Result family	FIT / SELECT / EVAL / ALIGN contract	Search / calibration	Inferential unit	Estimand and audit boundary
Amazon amortized selection	Candidate bank is constructed only from outer-training sources and frozen. Target outcomes never create a new bank operator. Preprocessing/model fitting/calibration occur inside outer training; held-out target states are evaluated once.	Action-only versus witness+action; grouped, leave-role, temporal-forward splits. Threshold/calibration uses outer-training predictions only.	Held-out state panels, with seed/run grouping appropriate to split; 122 exact temporal-forward panels in Appendix A.8.	Finite-bank action-ranking/selection value. Temporal-forward is a compatible-regime result, not universal witness invariance.
Kuai LOSO amortization	Outer held-out seed is untouched by predictor and threshold fitting; complete training checkpoints are held out internally for threshold calibration.	Context-only, task–pair prior, full witness, within-panel permutation, lifecycle/staleness controls.	Eight leave-one-seed-out folds.	Tests transfer of fine-grained witness-to-action mapping. Negative boundary does not negate Fresh headroom or replay structure.
Fresh $\rightarrow$ bank $\rightarrow$ policy decomposition	G_{fresh} comes from state-specific Fresh protocol; $G^{\star}_{\mathcal{B}}$ is best realized eligible value actually present in the frozen bank with no-op allowed; G_{π} is learned-policy realized value.	Bank fixed before held-out target outcomes; selector sees only legal pre-intervention inputs.	Held-out state panels, then seed-level aggregation where applicable.	Separates opportunity loss before selection from selector ranking/abstention regret. Fresh is not assumed to be a global ceiling.
Selected-tail eligibility	Eligibility threshold/calibration is fixed without held-out target leakage; final acted violations and primary gains are reported separately.	Prespecified or post-hoc-described ϵ curves are labeled accordingly; no claim of formal safety certification.	Acted target panels; the reported Clopper–Pearson reference treats panels as iid, without a cluster correction.	Secondary-quality eligibility is a constraint, not primary usefulness. Amazon temporal-forward does not certify a 5% budget at one-sided 95% confidence.
Search/selection meta-audit (R0)	Read-only audit checks role disjointness, nested preprocessing, candidate clustering, and whether reported final outcomes are untouched EVAL results.	No new search experiment; audit of existing evidence/capabilities.	Independent states/seeds, not action rows.	Historical core contract classified `passed_core_contract`; v3.2 separately labels retrospective bank and K1 maxima and checkpoint-selection reuse. Missing archived matched-displacement comparability remains explicit.

A.2.7 TEMPORAL CONTINUATION ESTIMANDS

Frozen-checkpoint response and continued-training response are different estimands. A *no-op continuation endpoint* is not the original source checkpoint; it is the model obtained after continuing ordinary MTL under the same schedule without the tested intervention. Continuation effects must be compared with this evolved no-op endpoint.

A *Fresh evolved-state fit* is different again. The assigned branch first evolves; the resulting checkpoint is then frozen; only then is the reference adapter fit under FIT–SELECT–EVAL. This measures the contemporaneous frozen response landscape. It is not an online closed-loop controller that refreshes the action while the model evolves.

Consequently, a Fresh frozen-state gain and a deployed continuation gain are not two estimates of the same quantity. The former measures current-state response under a bounded frozen action family; the latter measures the realized consequence of carrying an action through continued optimization. Their difference is scientifically informative, but their raw gains are not numerically interchangeable.

A.2.8 ACTION PROVENANCE AND LEAKAGE AUDIT

Learned amortization models consume only information available before executing the candidate action. Target-state descriptors may include contextual, spectral, geometric, relational, dimensional, exposure, interaction, and past-only temporal summaries. Action descriptors may include task, family, scope, pair/support, direction, source provenance, and nominal dose.

Model/checkpoint identifiers used only as bookkeeping, seed identity, Fresh or reference outcomes, realized displacement, post-intervention measurements, and future observations are excluded from learned inputs. For temporal and role-shift evaluation, preprocessing and threshold calibration are fit inside the outer training split. Foreign alignment uses a pre-intervention ALIGN/FIT population disjoint from EVAL and never uses intervention outcomes to choose a mapping.

Table 7: Pre-intervention feature, provenance, and leakage audit. "Permitted" denotes information available before executing the candidate action and used only within the declared split or fitting population. Bookkeeping identifiers may remain in artifacts for joins, grouping, and reproducibility while being excluded from learned inputs.

Audit surface	Permitted information	Explicitly forbidden / isolated information	Nesting, separation, and current audit result
Target-state structural descriptor	Current pre-intervention spectral, geometric, relational, dimensional, exposure, and interaction summaries for the target task–pair state.	Post-action measurements; realized candidate response; Fresh/reference outcome; any feature computed from the candidate's held-out EVAL result.	Descriptors are computed before intervention and serve only as observational witnesses. No scalar descriptor, feature block, or bank representation is assumed to be a sufficient state variable.
Panel-relative context	Current-panel means/scales, percentiles/ranks, symmetric unordered-pair summaries, and related transforms computed from candidates visible in the same target panel.	Statistics using held-out outcomes or future panels; arbitrary ordered (i, j) identity where the declared representation is permutation-invariant.	Panel-relative transforms use pre-intervention quantities from the target panel only. The seven-seed Amazon invariant views explicitly remove ordered expert identity from the learned representation.
Past-only temporal descriptor	Deltas, recent slopes, persistence summaries, empirical lifecycle position, and checkpoint history observed at or before the target state.	Future checkpoints; future metric trajectories; future intervention outcomes; retrospectively known replay or Fresh labels at the target state.	Temporal features are past/current-only. Lifecycle peak may be used as a descriptive alignment coordinate where specified by protocol, but not as a future-outcome feature for a deployment model.
Action descriptor	Task, action family, pair/support, legal direction, scope, source provenance, nominal dose, and other pre-execution action semantics defined by the experiment.	Realized target displacement; post-intervention primary or secondary metrics; candidate outcome rank; Fresh value of the nominal target pair.	Supervision is the realized target outcome of the exact candidate action. Nominal action identity is never substituted for target-state behavioral value.
Source provenance	Source lineage/run/checkpoint role when provenance is part of the declared candidate semantics; operator parameters; source task/pair/direction; and legally available age or horizon.	Source or target intervention outcomes unavailable at decision time; target Fresh oracle value; hidden target seed identity used as a memorization feature.	Provenance contextualizes an action but does not establish compatibility. Same/Foreign designations are evaluation strata rather than causal predictors.
Bookkeeping identifiers	IDs retained in manifests solely for joins, grouping, reproducibility, or fold construction.	Model/checkpoint IDs, raw row IDs, seed identity, run identity, or analogous identifiers used directly as learned predictors when they permit memorization of held-out states.	Grouping variables remain available for splitting and inference while being excluded from the learned feature matrix. Amazon seven-seed evaluation and Kuai LOSO preserve the declared outer grouping.

Table 7 continued.

Audit surface	Permitted information	Explicitly forbidden / isolated information	Nesting, separation, and current audit result
Preprocessing	Transformation parameters estimated only from the corresponding training population for the declared evaluation stage.	Scaling, imputation, missingness filtering, constant filtering, correlation pruning, or feature selection fit using held-out outer-test states.	In Amazon LOSO evaluation, preprocessing is fit inside each outer fold. The final frozen Amazon controllers separately serialize their fitted preprocessing contracts for deployment-time transformation.
Benefit / eligibility model fitting	Outer-training candidate panels and their intervention labels, including legal cross-validation performed wholly inside the outer-training split.	Outer-test intervention labels, target Fresh outcomes, or future target states.	Held-out state panels remain untouched until evaluation. Candidate rows within a checkpoint panel are clustered observations and are not treated as independent model realizations.
Threshold / calibration fitting	Out-of-fold predictions generated from outer-training data; complete held-out training checkpoints where required by the Kuai protocol.	Outer-test labels, or thresholds/calibrators refit after inspecting held-out policy outcomes.	Calibration and action-threshold selection are nested inside outer training. The Kuai final MVP reserves complete checkpoints for threshold calibration rather than individual candidate rows.
Foreign permutation ALIGN	Pre-intervention task-visible contribution similarity, FIT-side matched-expert CKA, and matched-coordinate correlations on the dedicated alignment population.	EVAL labels; intervention outcomes; target replay value; assignments chosen to maximize held-out behavioral gain.	ALIGN/FIT is disjoint from EVAL. Expert and coordinate Hungarian assignments are determined before target replay evaluation.
Similarity / nearest-Foreign donor ranking	Pre-intervention Amazon functional similarity or Kuai 920-D representation-descriptor distance.	Replay outcomes used to choose the nearest donor; EVAL primary or secondary labels.	Similarity matching is outcome-blind. Kuai Same/Foreign distance ranges lack common support; this is reported as an identification boundary rather than repaired through outcome-informed matching.
EVAL population	Final measurement of an already frozen selected, fitted, aligned, or replayed action.	Operator fitting, early stopping, direction selection, alignment, feature construction, model calibration, or policy-threshold tuning.	EVAL is untouched by FIT and SELECT. Production uses a fixed label-blind panel at FIT 15K / SELECT 5K / EVAL 100K; the general core separation contract is recorded as passed in R0.
Inferential grouping	Seed/lineage and matched-checkpoint identifiers used for bootstrap, pairing, and fold construction.	Treating thousands of pair/action/dose/candidate rows as independent inferential replicates.	Cross-trajectory claims aggregate over independent seeds or lineages. Checkpoint-level pairing is retained for matched within-trajectory contrasts.
Within-panel descriptor-collision audit	The frozen pre-intervention learner-visible descriptor together with legal action identity, used to test whether distinct candidates become feature-identical within a selection panel.	Outcome fields introduced solely to disambiguate otherwise identical candidate descriptors.	Final Amazon audit covers 486 complete six-candidate panels (2,916 candidate rows). No distinct legal actions collide under the frozen absolute view (238 inputs), relative view (637 inputs), or their joint representation: 0 collision groups, 0 affected rows, and 0 affected panels. No outcome-derived field is used for disambiguation.

A.3 Observing the Task-Visible Additive Boundary

The structural program began before the final intervention question was known. Its purpose was to determine whether standard routing statistics adequately describe the representations that are ultimately combined for a task. The answer is negative but bounded:

$$\text{routing exposure} \not\Rightarrow \text{task-visible additive representation state.} \tag{19}$$

This does not imply that every geometric discrepancy is harmful. It establishes that router activity and downstream additive representation state are distinct measurement surfaces.

A.3.1 Three measurement surfaces

Raw expert space. Statistics of $e_i(x)$ describe the representation produced by an expert before task routing and task projection.

Routed exposure. Statistics of $g_{k,i}(x)e_i(x)$ describe which expert content is exposed to task k and with what gate mass.

Task-visible additive space. Statistics of $P_k(g_{k,i}e_i)$ describe the contribution in the coordinates actually combined at the exposed task boundary.

The distinction is empirical rather than merely notational. Raw-space variation can lie in a downstream null direction and never reach the task; conversely, a task projection can expose signed interaction that is absent in identity space. Balanced routing or high gate entropy therefore cannot certify the structure of the contribution eventually seen downstream.

Historical diagnostics illustrate this separation without serving as authority for the current intervention claims. They show that active, non-collapsed routing can coexist with low effective dimensionality, weak dominant-subspace agreement, and weak local neighborhood preservation. Table 8 retains those observations as provenance alongside the later measurement-surface controls.

A.3.2 Measurement families

The measurement bank intentionally contains complementary views rather than a single putative health coordinate:

- **Spectral**: effective rank, concentration, anisotropy, and spectrum summaries.
- **Geometric**: subspace overlap, orientation, projection relationships, and dominant-angle summaries.
- **Relational**: neighborhood preservation and local relational agreement over the same examples.
- **Dimensional**: same-index fidelity, support, covariance, correlation, coordinate energy, and population summaries.
- **Exposure**: gate mass, effective sample support, active coverage, and pair/task visibility.
- **Interaction**: signed relationships between task-visible contributions, including the opposing-addition probe used in the falsification program.
- **Dynamic/contextual**: past-only deltas, slopes, persistence, lifecycle context, and within-panel relative position.

The bank is designed for observability. Controlled perturbations can change rank, orientation, neighborhood structure, exposure, and signed interaction in different combinations; ordinary training can move those quantities while task quality improves. A descriptor becomes a *witness* only when it is later supplied to a predictor of intervention outcomes.

A.3.3 Task-visible interaction smoke test

An early falsification made the choice of measurement surface concrete. On a ReLU-ended Amazon expert stack, the raw/identity-space negative-interaction field was inactive ($L_{\mathrm{AF}} = 0$, destructive

interaction $= 0$), while the same gated contributions projected through the task-tower linear map produced a nonzero task-visible signal ($L_{\mathrm{AF}} \approx 4.86 \times 10^{-4}$ and destructive interaction $\approx 9.97 \times 10^{-4}$). Holding the router fixed while perturbing the additive contribution further separates exposure from the representation eventually added downstream.

The conclusion is not that the router is irrelevant. It is that routing alone does not identify the downstream additive state and that structural statements must specify the surface on which they are measured.

A.3.4 No single structural scalar is sufficient

The diagnostic dossier contains several independent counterexamples to a scalar "harm" interpretation. Untreated training can reduce opposing interaction while task quality improves. Geometric alignment can improve while neighborhood preservation does not. Dimensional fidelity can improve while another interaction statistic and task behavior move in the opposite direction.

The strongest dimensional counterexample is retained because it records a hypothesis being falsified rather than merely replaced. On one seed trajectory, average task quality falls from approximately 0.5099 to 0.5016 while same-index dimensional fidelity rises from 0.6010 to 0.6221 and an exposed dimensional-risk summary falls from 0.06624 to 0.06225; meanwhile the local destructive interaction grows from 6.62×10^{-4} to 8.27×10^{-2}. No one of these scalars supplies a monotonic ordering of task utility.

The resulting epistemic rule is simple: measurements describe state; paired behavioral interventions label the consequence of acting on that state.

Table 8: Structural observability at the additive expert boundary. The table consolidates current measurement-surface controls with historical diagnostic evidence that motivated the intervention program. Historical Amazon-category and LTS rows are retained as provenance only; they are not authority for the current Production Fresh/Same/Foreign results.

Evidence / setting	Manipulation or surface	Quantitative observation	Scientific use / boundary	Status
A. Measurement-surface controls				
Amazon task-visible smoke test	Same ReLU-ended gated expert contributions measured in raw/identity space versus after the exposed task-tower linear map	Identity/raw: $L_{\mathrm{AF}} = 0$, $E_{\mathrm{destr}} = 0$. Task-visible projection: $L_{\mathrm{AF}} \approx 4.86 \times 10^{-4}$, $E_{\mathrm{destr}} \approx 9.97 \times 10^{-4}$ (destructive fraction $\approx 1.81 \times 10^{-2}$).	A structural field can be exactly inactive on one representation surface and nonzero on the surface actually exposed downstream. Structural statements must therefore name their measurement surface.	Current support; FND-03 / E037
Fixed-router perturbation	Router exposure held fixed while the additive contribution is perturbed downstream	Gate assignment/exposure is unchanged while the task-visible additive state changes.	Directly separates *which expert content is exposed* from *the representation state ultimately added.* Does not imply that every representation difference is harmful.	Current support; FND-01
Raw structural null-space caveat	Compare raw expert disagreement with task-visible projected disagreement	A raw-space difference may be mapped into a downstream null direction; conversely, the task projection can expose interaction absent from an identity-space probe.	Raw expert geometry, routed exposure, and task-visible additive behavior are distinct observables. The projected surface is a local view, not the complete nonlinear tower.	Current boundary; FND-03 / E026,E051
B. Historical multi-lens diagnostic snapshot (supporting provenance)				
Beauty (Amazon category)	Single checkpoint; 4 experts / 6 pairs	$D_{\mathrm{nom}} = 0.91$; eRank $= 5.11/128$; anisotropy 0.644 ± 0.039; Overlap@20 3.65 ± 0.20; principal angle $67.96° \pm 0.83°$; NP@10 0.065 ± 0.001.	Active routing coexists with low effective rank, weak dominant-subspace overlap, and weak neighborhood agreement.	Historical only; E024

Continued on next page

Table 8 continued.

Evidence / setting	Manipulation or surface	Quantitative observation	Scientific use / boundary	Status
Grocery (Amazon category)	Single checkpoint; 4 experts / 6 pairs	$D_{\text{nom}} = 0.86$; eRank $= 5.97/128$; anisotropy 0.558 ± 0.040; Overlap@20 3.88 ± 0.23; principal angle $66.91° \pm 0.89°$; NP@10 0.087 ± 0.001.	Same observational separation; not a scalar task-harm claim.	Historical only; E024
Sports (Amazon category)	Single checkpoint; 4 experts / 6 pairs	$D_{\text{nom}} = 0.87$; eRank $= 7.02/128$; anisotropy 0.592 ± 0.044; Overlap@20 3.57 ± 0.18; principal angle $68.25° \pm 0.63°$; NP@10 0.070 ± 0.001.	Same observational separation; later falsification prevents interpreting these diagnostics as a monotonic intervention objective.	Historical only; E024
Toys (Amazon category)	Single checkpoint; 4 experts / 6 pairs	$D_{\text{nom}} = 0.88$; eRank $= 8.98/128$; anisotropy 0.485 ± 0.041; Overlap@20 3.62 ± 0.21; principal angle $67.85° \pm 0.72°$; NP@10 0.068 ± 0.001.	Informative counterexample to a universal router-health story: this setting showed healthier gate–spectral association than several others.	Historical only; E023,E024
Historical production LTS	Single checkpoint; 8 experts / 28 pairs	$D_{\text{nom}} \approx 0.91$ (gate entropy mean 1.8868); mean eRank $22.94/256$; anisotropy 0.314 ± 0.133; Overlap@20 2.03 ± 0.34; principal angle $74.33° \pm 1.44°$; NP@10 0.216 ± 0.045.	Direct early production-scale evidence that routing activity and representation compatibility are distinct observables. This is **not** the current Production transport experiment.	Historical only; E002,E086
C. Diagnostic-family counterexamples				
Router-awareness audit	Compare gate usage with expert spectral health; compare gate-weighted pair pool with unweighted pool	Examples: $\rho(g, \text{eRank}) = -0.80$ (Beauty), -0.80 (Sports), -0.69 (historical LTS). Gate-weighted ΔNP@10 is approximately 0 in the public Amazon categories and -0.004 in LTS.	Standard task-trained routing is not reliably compatibility-aware. The claim is not that routers always prefer fragile experts; Toys provides the opposite directional example.	Historical support; FND-01 / E023
Post-hoc alignment probes	Translation / rigid alignment / shared-subspace projection / dimensional removal, aggregated within each setting	Geometry improves strongly: ΔOverlap@20 ranges $+4.048$ to $+5.697$ and principal-angle change ranges $-14.48°$ to $-22.07°$. Cross-expert ΔNP@10 is non-positive in every setting (-0.014 to -0.004).	Improving coarse geometry does not imply recovery of relational organization. Geometry and relational compatibility are complementary measurements, not interchangeable definitions of "repair."	Historical support; FND-03 / E028
Synthetic lens calibration	Controlled perturbations: rank truncation, rotation/alignment, and shared, partial, or independent relational structure	The spectral, geometric, and relational bases respond differently to their targeted perturbations; e.g., rigid alignment changes overlap/angle while relational regimes separate NP@k from rank-dispersion behavior.	Justifies a multi-lens measurement bank but also shows why a compressed compatibility score can hide basis-specific disagreement.	Historical method provenance; FND-02 / E003
Seed-23 dimensional counterexample	Same trajectory, epoch $31 \to 61$	Avg task quality $0.5099 \to 0.5016$; same-index DFID $0.6010 \to 0.6221$; exposed dimensional risk $0.06624 \to 0.06225$; local destructive interaction $6.62 \times 10^{-4} \to 8.27 \times 10^{-2}$.	A dimensional witness can improve while another structural field deteriorates sharply and task quality falls. No individual measured scalar supplies a monotonic task-utility ordering.	Current support; FND-02 / E046,E068

A.4 Structural Falsification and Measurement Boundaries

The additive-fidelity program tested a scientifically plausible shortcut: if opposing task-visible expert contributions are harmful, then directly suppressing that interaction should monotonically im-

prove downstream behavior. The experiments falsify that stronger statement. This is the pivot between the structural-diagnostic and behavioral-intervention halves of AMIYC.

A.4.1 Probe definition and intervention variants

Let $\Psi_{k,ij}$ denote the tested local task-visible opposing-interaction quantity for task k and pair (i, j). Training variants changed regularizer strength, weighting, refresh behavior, and the gradient paths available through the gate and task projection. The principal stress tests detach the relevant branches so that the measured field cannot be reduced trivially through router avoidance or projection drift.

These variants are useful precisely because they manipulate the structural quantity without replacing the downstream task objectives. The probe is a falsification instrument, not the definition of the later adaptation objective.

A.4.2 Strong structural suppression is behaviorally non-monotonic

At the matched Amazon epoch-24 checkpoint, increasing AF pressure reduces the targeted destructive interaction monotonically: from approximately 5.08×10^{-4} untreated to 2.89×10^{-4} at a moderate setting and 8.55×10^{-5} at the strongest setting. Held-out NDCG@10 is not monotonic: $0.45836 \to 0.45916 \to 0.45791$. The strongest suppression therefore drives the measured interaction lower while degrading the ranking metric below the untreated checkpoint.

The failure is not rescued by attributing every observed structural change uniquely to AF. Ordinary untreated MTL itself can reduce destructive interaction while task quality improves. Aggressive AF can also compress an expert-diversity proxy without producing router collapse. The supported conclusion is therefore not that the opposite scalar should be optimized, but

$$\Psi \downarrow \not\Rightarrow \; U_\epsilon \uparrow . \tag{20}$$

A.4.3 Local saturation does not imply structural resolution

Later diagnostics exposed a second failure mode for direct structural penalties. A local opposing-interaction field can saturate near zero while broader structural-risk and dimensional descriptors remain active. A small scalar penalty does not imply a small gradient. Wherever the terms are differentiable,

$$\nabla(S_{ij}\Psi_{ij}) = S_{ij}\nabla\Psi_{ij} + \Psi_{ij}\nabla S_{ij}. \tag{21}$$

The gradient is zero in an inactive region where both $\Psi_{ij} = 0$ and $\nabla\Psi_{ij} = 0$, with finite differentiable S_{ij}. Near-zero field values alone establish neither condition. The historical scalar saturation observation therefore motivates behavioral intervention measurements, without proving that risk weighting cannot restore a gradient.

A.4.4 Dimensional-fidelity invariance under positive diagonal scaling

For the correlation-based same-index dimensional-fidelity witness, let u_q, v_q have finite, nonzero variance on the same fixed population. A constant positive scale d_q leaves Pearson correlation unchanged:

$$\rho(d_q u_q, v_q) = \rho(u_q, v_q), \qquad d_q > 0. \tag{22}$$

This identity concerns actual coordinate scaling with unchanged population weights. For the lifted operator in Eq. 2, it applies when the requested projected change is realized exactly (the stated rank/range condition). Otherwise $P_k P_k^+$ can mix coordinates, and invariance of the realized witness is not guaranteed. Under direct scaling, behavioral gain need not change this particular descriptor.

A.4.5 Interpretation boundary

The falsification is local to the tested structural quantities. It does not establish that structural measurements are useless. In the compatible Amazon regime, multivariate pre-intervention descriptors later improve action-value prediction. The narrower and stronger conclusion is that no tested single scalar is validated as a sufficient calibrated harm score or monotonic intervention objective.

The project therefore retains the structural dossier not because its earliest mechanistic interpretation survived, but because the failures determine what the subsequent behavioral intervention program must measure.

Table 9: Structural-objective falsification. Panel A is the matched Amazon epoch-24 AF stress sweep: stronger pressure monotonically suppresses the targeted opposing-interaction statistic while held-out ranking quality is non-monotonic. Panel B records independent controls showing why structural measurements are retained as descriptors rather than promoted to monotonic intervention objectives.

Condition / control	AF strength	Structural observation	Behavioral / safeguard observation	Interpretation
A. Matched epoch-24 AF stress sweep				
Untreated	$\lambda_{\mathrm{AF}} = 0$	$E_{\mathrm{destr}} = 5.0793 \times 10^{-4}$ (relative 1.000)	NDCG@10 $= 0.45836$	Reference state.
Weak AF	$\lambda_{\mathrm{AF}} = 10$	$E_{\mathrm{destr}} = 4.7153 \times 10^{-4}$ (relative 0.928)	NDCG@10 $= 0.45855$	Small structural suppression; no monotonic behavioral law inferred.
Moderate AF	$\lambda_{\mathrm{AF}} = 100$	$E_{\mathrm{destr}} = 2.8943 \times 10^{-4}$ (relative 0.570)	NDCG@10 $= 0.45916$	Structural suppression and ranking improvement happen to coincide at this setting.
Aggressive AF	$\lambda_{\mathrm{AF}} = 1000$	$E_{\mathrm{destr}} = 8.5510 \times 10^{-5}$ (relative 0.168)	NDCG@10 $= 0.45791$; expert-diversity proxy 0.6557 versus 0.6760 untreated; no router collapse observed in the safeguard audit.	Strongest structural suppression produces worse held-out ranking than untreated. Therefore $\Psi\downarrow \not\Rightarrow U_\epsilon\uparrow$.
B. Independent falsification and boundary controls				
Ordinary-training trajectory	No AF intervention	Untreated E_{destr} falls approximately $7.27 \times 10^{-4} \rightarrow 2.78 \times 10^{-4}$ over training.	NDCG rises approximately $0.1272 \rightarrow 0.4807$ over the same broad trajectory.	Structural motion is not uniquely caused by AF. Ordinary task training can reduce the same observable while quality improves; this is a trajectory control, not a matched causal estimate.
Static pair-risk shortcut	Static positive pair-risk penalty	On the gate simplex, the penalty can be driven toward zero by concentrating routing on a vertex / avoiding co-routing.	Auxiliary loss can improve without changing the expert representation basis.	A structural prior can create a router-avoidance shortcut. Lower structural exposure is not sufficient evidence of representation improvement.
Weighting / gradient-path variants	Uniform, destructive, hybrid; attached versus detached routing paths	Variants occupy a narrow task-performance region; changing the weighting or detachment rule can move the measured field without uniquely improving downstream metrics.	Mature audits did not show a unique behavioral advantage for one structural weighting or simple detached-gate variant.	Weakens the claim that a particular structural weighting rule or gradient path is itself the source of downstream value.
Local saturation	Local opposing-interaction field driven near zero	Refreshed broader structural-risk / dimensional descriptors remain active, while the weighted scalar tends to zero if S_{ij} stays bounded and $\Psi_{ij} \rightarrow 0$.	Scalar saturation does not establish gradient saturation. The product rule in Eq. 21 requires separate derivative conditions.	Local collision suppression and remaining intervention-recoverable opportunity are different objects.
Dimensional-fidelity invariance	Constant $d_q > 0$; finite nonzero variances	Exact identity: $\rho(d_q u_q, v_q) = \rho(u_q, v_q)$.	With fixed population weights and directly realized coordinate scaling, behavior can change while this correlation is unchanged.	For the lifted action, exact scaling needs the rank/range condition. Otherwise PP^+ may mix coordinates and this invariance is not guaranteed.

Continued on next page

Table 9 continued.

Condition / control	AF strength	Structural observation	Behavioral / safeguard observation	Interpretation
Dimensional scalar counterexample	See structural-observability table, seed 23	DFID increases and exposed dimensional risk falls while destructive interaction increases by over two orders of magnitude.	Task quality falls ($0.5099 \rightarrow 0.5016$).	Retained here only as a cross-reference to avoid duplicating the full numeric row: no tested dimensional scalar is a sufficient calibrated harm score.

A.5 Frozen Intervention Response and Construct Validity

The structural program in Appendices A.3–A.4 establishes what can be observed and which scalar shortcuts fail. A.5 asks the next question directly: given a frozen state, what bounded changes actually alter held-out behavior, and which interpretations of that response survive matched controls? The section therefore separates three objects: (i) the empirical response surface over intervention families, (ii) scope/capacity and realized dose, and (iii) construct validity of the positive-diagonal reference family.

The original balanced Amazon intervention campaign underlying the ladder contains 30 frozen checkpoints, 540 selected task–pair states, and 2,730 intervention-action rows. Those rows describe a response surface; they are not 2,730 independent model replicates. Later seed-level controls use their declared seed or lineage units for inference, consistent with Appendix A.2.6.

Original and expanded Amazon snapshots. The original Phase-II analysis underlying Table 1A reports 360 positive-diagonal outcomes, all loss-positive, with mean gain 0.02874498 and mean NDCG change 0.001763. A later action-result snapshot contains 540 diagonal outcomes, also all loss-positive, with mean gain 0.02854423 and 540 stored operators. These are separate populations: the original 360-row membership is not reconstructed from the expanded snapshot. The headline retains the original analysis; the larger snapshot corroborates positive response without replacing that cohort.

A.5.1 Full frozen intervention ladder

The Amazon ladder evaluates pair removal, direct interaction suppression, positive coordinate-wise scaling, scale/sign freedom, near-identity rewriting, and a smaller orthogonal Procrustes-style panel under paired frozen-state evaluation.

Removal and direct suppression. Pair removal is usually harmful: across 540 selected task–pair states, mean primary change is -0.064244 (median -0.041692), and only 18.5% of rows improve primary loss; mean NDCG change is -0.089303, with 13.3% positive. In contrast, directly suppressing the original local opposing-interaction field yields mean primary change of only approximately 1.19×10^{-7}, with essentially zero positive-loss rate. The pair contribution is therefore often useful even when a local structural interaction is measurable, and suppressing that interaction does not recover the remaining frozen-state value.

Positive diagonal. The bounded positive diagonal improves primary held-out loss in all 360/360 Amazon pair-states in the mature ladder. Mean $\Delta L = +0.028745$, median $+0.025777$, mean $\Delta\text{NDCG@10} = +0.001763$, and the median NDCG change is $+0.000721$. The ranking response is not uniformly positive: 57.2% of states improve NDCG, and 40.56% improve primary loss while worsening NDCG. This is exactly why primary headroom and secondary-quality preservation are tracked separately.

More expressive transforms. Scale/sign and near-identity transforms increase average primary recovery to $+0.051110$ and $+0.080206$, respectively, with positive primary loss gain in every one of their 360 evaluated states. Their mean NDCG changes are instead negative (-0.001796 and -0.004830). Among primary-improving states, NDCG worsens in 39.17% of scale/sign cases and

49.72% of near-identity cases. More capacity can therefore expose additional primary value without preserving the independent ranking metric.

Orthogonal alignment. The 30-state Procrustes panel has positive mean primary gain ($+0.007683$) and mean NDCG change ($+0.000608$), but it is not competitive with the best restricted action on the same states: Procrustes minus best-restricted mean primary gain is -0.061547, and Procrustes wins on 0/30 states. This is evidence against the simple historical hypothesis that an unrestricted orthogonal registration is the missing solution. It is not a general statement that rotations cannot help, and it does not conflict with the later rotated-basis diagonal control, which tests a different family under matched constraints.

The reference-family response also replicates beyond this Amazon ladder. Fresh positive-diagonal fitting is primary-positive in 382/384 completed Kuai pair outcomes and in all 9/9 predeclared Production target states. These support counts remain in their native within-setting units; loss magnitudes and row counts are not pooled across datasets.

A.5.2 Action-conditioned response

The ladder retires the idea that one scalar pair score can stand in for intervention value. Pair usefulness under removal, primary headroom under rescaling, incremental cross-coordinate capacity, and secondary-quality eligibility are related but non-interchangeable outcomes. For pre-intervention state descriptor $w(s, k, p)$ and action a,

$$Y = F(w(s, k, p), a)\,. \tag{23}$$

The intervention bank supplies behavioral labels for F; structural severity does not supply them a priori. This action-conditioned view is retained even when two actions share the same nominal pair, because scope, dose, direction, parameterization, and source provenance can change the realized behavior.

A.5.3 Scope, dose, and matched displacement

Pair-local support was initially attractive because it supplies sparse, interpretable containment. The later controls show that pair locality should not be interpreted as a uniquely causal defect locus. At the same nominal dose, a full task-global action moves the representation about $3.30\times$ farther in the appendix summary, so the raw scope comparison confounds locus with realized reach. Historical record E058 instead reports $0.2562/0.0694 \approx 3.69$; the supplied baseline lacks the rows needed to reconcile their populations or weighting. Both remain descriptive historical summaries; the matched contrasts below support the scope conclusion.

The decisive comparison uses an equal-capacity 64-dimensional task-global diagonal and matches realized displacement. Pair-local minus global64 eligible utility is -0.002025 with 95% CI $[-0.010429, 0.007701]$: the difference is unresolved, and equivalence is not established. By contrast, pair-local exceeds a one-parameter scalar by $+0.003621$ with 95% CI $[0.001113, 0.006620]$. The supported conclusion is therefore about controllable reach and capacity, not a uniquely bad pair: pair locality provides sparse support, containment, and a convenient action bank, while coordinate capacity and realized dose determine how much the action can change.

Table 10 consolidates the intervention response, cross-setting Fresh support, and matched scope/capacity evidence.

A.5.4 Expressive-family capacity boundary

The frozen ladder measures response across a range of intervention families. The global64/scalar and same-boundary experiments separately control realized displacement and expose the role of parameterization. These comparisons retain their stated capacity and quality contracts; the full ladder describes a response frontier rather than an all-family matched-budget ordering.

A.5.5 Normalization validity control

A natural construct-validity alternative is that post-hoc coordinate-wise gain simply compensates for missing normalization at the additive boundary. K1-N adds a prespecified affine LayerNorm-

Table 10: Frozen intervention response, replicated Fresh support, and matched scope/capacity controls. Panel A reports the Amazon frozen-response distribution; Panel B reports the reference adapter across settings in each setting's native evaluation unit; Panel C gives the final matched scope/capacity contrasts. Row-level positive rates in Panel A describe the response surface and are not treated as independent statistical replicates.

A. Amazon frozen intervention ladder

Action	N	mean ΔL	med. ΔL	$P(\Delta L > 0)$	mean ΔNDCG@10	$P(\Delta N > 0)$
Pair removal	540	-0.064244	-0.041692	18.5%	-0.089303	13.3%
Interaction suppression	540	$\approx 1.19\times10^{-7}$	–	≈ 0	≈ 0	–
Positive diagonal	360	$+0.028745$	$+0.025777$	100%	$+0.001763$	57.2%
Scale/sign	360	$+0.051110$	$+0.048530$	100%	-0.001796	59.7%
Near identity	360	$+0.080206$	$+0.066315$	100%	-0.004830	49.4%
Procrustes	30	$+0.007683$	$+0.004432$	96.7%	$+0.000608$	53.3%

Among states with positive primary-loss gain, NDCG worsens for 40.56% of positive-diagonal states, 39.17% of scale/sign states, and 49.72% of near-identity states. On the 30-state Procrustes subset, Procrustes minus the best restricted action has mean primary-loss-gain difference -0.061547, and Procrustes wins on 0/30 states. These quantities establish a primary/secondary tradeoff; they do not identify a universal ordering over adapter families.

B. Fresh positive-diagonal response across settings

Setting	Evaluation unit	Fresh result	Positive support
Amazon Mixed-4Cat	pair-state	$\mathrm{E}[\Delta L] = +0.02875$	360/360
KuaiRand	pair-state	positive held-out loss gain	382/384
Production	target-state	$\mathrm{E}[\Delta L] = +0.006864$	9/9

C. Scope, capacity, and realized reach (Amazon)

Comparison	Params	Matching / reach	Result
Task-global full scope vs. pair-local	–	same nominal dose; global displacement $\approx 3.30\times$ local	unmatched; scope confounded with realized reach
Pair-local − global64	64 vs. 64	matched realized displacement	$\Delta U_\epsilon = -0.002025$; 95% CI $[-0.010429, 0.007701]$
Pair-local − scalar	64 vs. 1	matched realized displacement	$\Delta U_\epsilon = +0.003621$; 95% CI $[0.001113, 0.006620]$

Positive ΔL denotes primary-objective improvement. In Panel C, positive pair-local-minus-comparator ΔU_ϵ favors pair-local application. The older full-scope comparison is retained only as a reach diagnostic; the intrinsic scope claim uses the matched global64/scalar controls. Cross-setting loss magnitudes in Panel B are not compared numerically.

style boundary normalization during ordinary training (512 added affine parameters), selects checkpoints by its declared development criterion, freezes the selected model, and reruns the mature Fresh positive-diagonal audit.

The retrospective EVAL-best-eligible-pair summary is large and positive: mean gain is 0.008506 with 95% seed interval $[0.006634, 0.010356]$, and all 8/8 seed means are positive. However, the selected base model is materially worse than the ordinary architecture: selected dev-AUC minus ordinary is -0.009634 with 95% CI $[-0.010132, -0.009011]$, and selected dev loss is higher by approximately 0.028248. The tested normalization therefore does not eliminate the phenomenon, but the base-quality mismatch prevents the stronger conclusion that every quality-matched normalization explanation has been falsified. K1-N is a bounded control, not an unfinished one.

A.5.6 Parameter availability from initialization

K1-D places bounded positive-diagonal coordinates in the model from initialization (512 coefficients: two tasks, four experts, 64 coordinates) and trains them jointly with ordinary MTL. All 16 task–seed blocks depart from identity, with mean log-coefficient ℓ_2 norm 0.1137. Selected dev-AUC minus ordinary is -1.144×10^{-4} (95% CI $[-4.512, 2.542] \times 10^{-4}$); the interval crossing zero is not an equivalence test.

After freezing the selected checkpoint, the residual action multiplies the trained projected diagonal in this control's direct projected-boundary backend. Its log-coordinates and the training-time coordinates are each bounded by 0.5, but their sum is not audited under the original single-diagonal bound. The result therefore tests residual response on already-trained coordinates, not saturation within an identical composite feasible set or an equal training budget.

The descriptive mean over all six fitted pair candidates in each of 16 blocks is 7.0426×10^{-4}; 95/96 pair gains are positive and all 96 satisfy the secondary tolerance. This is the main-paper result. The recovered aggregate summaries do not retain the seed-by-pair values needed for its seed-level CI; no independent-replication claim is made from the pair count.

For provenance, the archived headline 9.5028×10^{-4} instead first filters pairs by EVAL eligibility and then maximizes EVAL gain within each block. Its eight-seed bootstrap interval is $[8.165 \times 10^{-4}, 1.0785 \times 10^{-3}]$, with eight positive seed means. This is an optimistic retrospective maximum, not a SELECT-chosen action or selection-adjusted estimate. The K1-N 0.008506 summary uses the same EVAL-best-pair aggregation. These maxima are retained explicitly as historical summaries rather than silently relabeled.

A.5.7 Same-boundary low-rank adapter baseline

A prespecified rank-1 family tests whether the effect is specific to the reference diagonal while holding the intervention surface fixed:

$$u' = (I + UV^\top)u. \tag{24}$$

The comparison uses the same task-visible boundary, data roles, FIT–SELECT–EVAL logic, and matched realized displacement. At that displacement, rank-1 is stronger than the native positive diagonal: native-minus-rank1 eligible utility is -0.016083 (95% CI $[-0.027095, -0.007768]$), and native-minus-rank1 primary-loss gain is -0.019335 (95% CI $[-0.030158, -0.011962]$). Rank-1 uses 128 parameters rather than 64, so the result establishes adapter-family non-uniqueness, not equal-capacity rank-1 dominance.

A.5.8 Rotated-basis diagonal ablation

To test whether the response requires the learned native task-visible axes, one fixed, outcome-independent orthogonal basis R is chosen and a diagonal is fit in that basis:

$$u' = R^\top D(z) R\, u. \tag{25}$$

The native and rotated diagonals both use 64 fitted parameters, inherit the native-selected support, and are compared at matched realized displacement. Native minus rotated eligible utility is $+0.003292$ with 95% CI $[-0.005321, 0.012860]$, so the eligible-utility difference is not clearly separated from zero. On the primary objective, the rotated family is better on average: native-minus-rotated gain is -0.005742 with 95% CI $[-0.009629, -0.002301]$. Native axes are therefore not necessary in this tested comparison. One fixed basis does not establish universal rotational invariance.

Table 11 collects the normalization, from-initialization, rank-1, and rotated-basis controls. Together they broaden the empirical object while narrowing its interpretation: the positive diagonal remains a bounded, replayable reference adapter, not the phenomenon, a unique optimum, or an identified mechanism.

Table 11: Construct-validity controls on the bounded positive-diagonal reference adapter. Panel A tests training-time alternatives to post-hoc headroom; Panel B compares alternative same-boundary adapter families at matched realized displacement. The table reports the measured contrasts; the surrounding text states the surviving interpretation and boundary.

A. Training-time alternatives

Control	Added params	Selected-base audit	Residual Fresh audit	Status
K1-N affine normalization	512	Dev-AUC -0.009634 vs. ordinary; 95% CI $[-0.010132, -0.009011]$; dev-loss $+0.028248$	EVAL-best-pair gain $+0.008506$; 95% CI $[0.006634, 0.010356]$; 8/8 seed means positive	Bounded
K1-D diagonal from initialization	512	Dev-AUC -1.144×10^{-4}; 95% CI $[-4.512, 2.542]\times10^{-4}$; 16/16 task–seed blocks nonidentity; mean $\|z\|_2 = 0.1137$	EVAL-best-pair gain $+9.503\times10^{-4}$; 95% CI $[8.165\times10^{-4}, 1.0785\times10^{-3}]$; 8/8 seed means positive	EST.

B. Alternative adapters at the same task-visible boundary (Amazon, 7 seeds)

Alternative	Params	Native−Alt. ΔL	Native−Alt. U_ϵ	Identified boundary
Rotated diagonal	64	-0.005742; 95% CI $[-0.009629, -0.002301]$	$+0.003292$; 95% CI $[-0.005321, 0.012860]$	native axes not necessary in this one-basis test
Rank-1 adapter	128	-0.019335; 95% CI $[-0.030158, -0.011962]$	-0.016083; 95% CI $[-0.027095, -0.007768]$	family non-unique; not parameter matched

K1-N leaves substantial residual headroom but materially worsens selected base quality, so it does not cleanly falsify every quality-matched normalization account. K1-D shows response on already-trained coordinates; dev-AUC equivalence and an identical composite bound are not established. Its unmaximized all-pair mean is 7.0426×10^{-4} (95/96 positive, 96/96 eligible), with no recovered seed CI. Panel A maxima use EVAL for pair selection and their intervals are conditional, not selection-adjusted. In Panel B, positive native-minus-alternative values favor the native positive diagonal. The rotated control uses one fixed outcome-independent basis and native-selected support; the rank-1 comparison matches realized displacement but uses 128 rather than 64 parameters.

A.6 Lifecycle, Target Motion, and Replay Compatibility

Fresh fitting measures the bounded adaptation opportunity available at a frozen current state; replay asks a different question: whether an operator fitted at another state remains behaviorally compatible when applied unchanged. The three settings separate these objects cleanly. Amazon and Kuai contain regimes of substantial local Same replay, whereas the three Production lineage means change sign across the measured horizons. The operational quantity is therefore *current-state replay compatibility*, not lineage identity, elapsed checkpoint age, or nominal action identity.

This section keeps both the regularity and its boundary. Local continuity is a real empirical regime and is what makes finite operator banks plausible at all. It is not a universal persistence law, a monotonic staleness curve, or evidence that lineage membership causally determines compatibility.

A.6.1 Kuai Fresh opportunity exists before and near peak

A first lifecycle alternative was that Fresh headroom merely repairs late checkpoint degradation. Restricting the Kuai audit to states at or before each task trajectory's own validation-AUC peak rejects that strong explanation. Mean Fresh gain is 8.61×10^{-4} (95% bootstrap interval $[7.77, 9.44]\times10^{-4}$), with positive seed means in all eight training seeds. The early anchor means are similar: 8.45×10^{-4} at epoch 1 and 8.01×10^{-4} at epoch 11.

Replay continuity is already present in this same high-quality regime. Earlier Same actions retain 91.2% of contemporaneous Fresh value (95% CI $[83.6\%, 99.2\%]$), and Same exceeds Foreign by 7.45×10^{-4} (95% CI $[6.62, 8.34] \times 10^{-4}$), with positive seed means in $8/8$ seeds.

The lifecycle audit also corrected a reporting denominator rather than silently carrying an earlier rate forward: missing Fresh-reference rows had previously been counted as negatives. On authoritative

non-null anchors the Fresh-positive counts are $49/49$ pre/near peak and $47/47$ post peak, $96/96$ overall. The correction does not change seed means, replay contrasts, or the quality regressions.

A.6.2 THE LARGER LATE GAP IS CHECKPOINT-QUALITY COUPLED

Post-peak Kuai states have a larger raw Fresh gap, with mean gain about 1.424×10^{-3}. Its magnitude is strongly associated with degradation of the underlying checkpoint: the correlation with dev-loss degradation is 0.838. Let $Y = \Delta L_{\mathrm{Fresh}}$ be raw primary-loss gain at a task checkpoint, d_{loss} validation-loss degradation, and h_{peak} epochs since the task's selected validation peak. The descriptive model $Y = \beta_0 + \beta_d d_{\mathrm{loss}} + \beta_h h_{\mathrm{peak}} + \alpha_{\mathrm{task}} + \gamma_{\mathrm{seed}} + \varepsilon$ adjusts for task and seed. Its degradation coefficient remains positive ($\beta = 0.2044$, 95% CI $[0.1460, 0.2602]$), whereas the epoch-from-peak coefficient is -1.26×10^{-5} with interval $[-3.03 \times 10^{-5}, 3.57 \times 10^{-6}]$ crossing zero.

The supported lifecycle statement is therefore deliberately asymmetric: Fresh headroom is already present around high-quality checkpoints, while the larger raw late gap is strongly quality-coupled. The available intervention-EVAL frames do not place an adapted late checkpoint and the earlier validation-selected checkpoint on one authoritative comparison population. Accordingly, "adapted late $>$ earlier best" is unidentified rather than a negative result.

A.6.3 FRESH TARGET MOTION

Persistent Fresh value does not imply a stationary operator. Relative to the epoch-11 Kuai Fresh fit, mean log-coefficient displacement grows from 0.346 at epoch 12 to 0.595 at epoch 14 while coefficient cosine falls from 0.936 to 0.866. Preferred direction changes much less often, from 7.3% to 13.5%. The fitted coefficients can therefore differ while nominal action semantics remain comparatively stable. These comparisons lack same-state repeated-fit controls and do not partition state evolution from fitting variability.

Exact winning-pair identity overstates this motion because the frozen response landscape contains near ties. The median best-versus-second gain gap is only about 5×10^{-6}, approximately 0.43% of the winning gain, and an average of 2.15 of the six pairs lie within 99% of the winner. Only one of sixteen seed–task trajectories keeps the same argmax pair at every audited anchor. We therefore treat Fresh as a state-specific high-value reference within the bounded search protocol, not as evidence for a unique current-state optimum vector.

Production supplies an independent motion audit at a larger boundary. Its preferred direction changes in only 15–23% of pair states, yet source-to-target coefficient cosine is low and non-monotonic across $+1, +3, +5$. Thus discrete pair or direction identity is not sufficient to describe motion of the high-value response.

A.6.4 AMAZON FRESH, SAME, AND FOREIGN

Amazon provides a second replay regime over 63 evolved target states and 378 target-pair states. These Amazon summaries use eligibility-gated U_ϵ, not raw primary loss gain. Mean Fresh value is 0.017843 and Same replay retains 0.014007, or 78.5% of Fresh, whereas raw Foreign replay is only 0.001080. This is substantial local compatibility, but it is not a stationary-target result: Fresh still exceeds the stale Same operator at the descendant state.

Nor is Amazon retention a simple function of elapsed horizon. Descriptive retention over its three horizon blocks is approximately 79.1%, 83.2%, and 73.5%. Together with Kuai and the Production result below, this argues against fitting a universal monotonic stale-decay law.

A.6.5 FAMILY-PRESERVING PERMUTATION ALIGNMENT

A raw cross-lineage comparison can be confounded by exchangeable experts and coordinate permutations. We therefore construct an explicit FIT-only gauge correction before replay. Experts are matched from task-visible contribution similarity with a Hungarian assignment; coordinates within matched experts are matched from absolute correlation using a second Hungarian assignment. The donor diagonal is transported only by this permutation,

$$D_{\mathrm{aligned}} = P^\top D_{\mathrm{source}} P, \tag{26}$$

Table 12: Lifecycle, checkpoint quality, and motion of the Fresh reference. Panel A separates the existence of Kuai Fresh headroom from the larger raw post-peak gap; Panel B measures short-horizon motion of the fitted Kuai operator; Panel C gives the independent Production operator-motion audit. The quality model and the nine-state Production correlations are descriptive rather than causal.

A. Kuai lifecycle and checkpoint-quality audit

Quantity	Estimate	95% interval / support	Role
Pre/near-peak Fresh gain	8.61×10^{-4}	$[7.77, 9.44]\times10^{-4}$; 8/8 seeds >0	headroom exists around high-quality checkpoints
Same/Fresh, pre/near peak	0.912	$[0.836, 0.992]$	substantial local replay already present
Same−Foreign, pre/near peak	7.45×10^{-4}	$[6.62, 8.34]\times10^{-4}$; 8/8 seeds >0	replay gap precedes late degradation
Fresh at epoch 1 / epoch 11	$8.45/8.01\times10^{-4}$	descriptive anchors	early-headroom stability
Authoritative non-null Fresh anchors	49/49 pre/near; 47/47 post	96/96 overall positive	denominator-corrected audit
Post-peak Fresh gain	1.424×10^{-3}	descriptive mean	larger raw late gap
$\mathrm{corr}(G_{\mathrm{fresh}}$, dev-loss degradation)	0.838	–	strong quality coupling
Mean dev-loss degradation	≈ 0.00314	–	descriptive scale
$\beta_{\mathrm{degradation}}$, joint descriptive model	0.2044	$[0.1460, 0.2602]$	positive after adjustment
$\beta_{\mathrm{epoch-from-peak}}$, same model	-1.26×10^{-5}	$[-3.03\times10^{-5}, 3.57\times10^{-6}]$	interval crosses zero
Simple $\beta_{\mathrm{dev\ loss}}$	0.1614	$[0.1258, 0.2018]$	descriptive quality association

B. Kuai Fresh-operator motion from the epoch-11 reference

Target	$\|\Delta z\|_2$	$\cos(z_{11}, z_t)$	Direction switch	Behavioral/action context
e12	0.346	0.936	7.3%	near-tie audit summarized below
e13	0.453	0.898	11.5%	
e14	0.595	0.866	13.5%	

Only 1/16 seed–task trajectories keeps the same argmax pair across all audited anchors. Exact winner identity therefore moves more often than the high-value region itself; the audit does not identify a unique current-state optimum vector.

C. Production source-to-target operator motion

Horizon	Direction change	$\|\Delta z\|_2$	Mean cosine	Same-direction cosine
$+1$	20.2%	4.787	0.117	0.148
$+3$	15.5%	4.756	0.318	0.347
$+5$	22.6%	5.501	-0.059	-0.037

Across the nine Production target states, post-hoc descriptive correlations are approximately $r = 0.92$ between source–target coefficient cosine and Same gain, and $r = 0.91$ between cosine and the Same–Foreign gap. These nine-state correlations are exploratory clues only, not inferential or causal results. The lifecycle data do not place adapted late checkpoints and the earlier validation-selected checkpoint on one authoritative intervention-EVAL frame; "adapted late $>$ earlier best" therefore remains unidentified, not negative.

with no intervention outcome or target EVAL label used to choose the mapping.

The alignment is nontrivial rather than cosmetic: mean expert CKA increases from 0.498 under nominal indexing to 0.759 after matching, expert mappings are identity in only 5.3% of relations, matched coordinate absolute correlation averages 0.663, and coordinate mappings are identity in only 1.35%. Alignment raises Foreign value from 0.001080 to 0.002557, but closes only 11.4% of the raw Same–Foreign gap. The residual Same-minus-aligned-Foreign gap is 0.011451 and remains positive in all seven seed blocks. Simple family-preserving permutation ambiguity therefore explains a real but minority component of the observed portability loss; unrestricted rotations, gauges, and broader functional equivalences remain outside this control.

A.6.6 LINEAGE VERSUS MEASURABLE FUNCTIONAL SIMILARITY

The residual Same advantage after permutation alignment does not make lineage a causal variable. We next ask a narrower read-only question: among the actually available independent-run donors, does selecting the donor closest to the target under pre-intervention measurements rescue replay?

For Amazon, donor ranking uses only FIT-side matched-expert CKA and matched-coordinate absolute correlation. The nearest observed Foreign donor still trails Same by 0.0137349 on average (95% CI $[0.0107844, 0.0172852]$), with a positive contrast in all seven seeds over 378 target-pair relations.

Kuai supplies a richer representation-level audit using 920 pre-intervention features. Same exceeds the nearest Foreign selected by this representation metric by 7.4977×10^{-4} on average (95% CI $[5.9662, 9.1623] \times 10^{-4}$), positive in $8/8$ seeds. The decisive identification boundary is the support geometry: Same distance has mean 0.1564 and maximum 0.2734, whereas Foreign distance has mean 1.4191 and minimum 0.7278. The ranges do not overlap. The tested measurements therefore do not make observed Foreign operators portable, but they cannot identify a residual causal effect of lineage after exact representation-distance matching because that matched population does not exist in the observed data.

A.6.7 PRODUCTION-SCALE TARGETED REPLICATION

Production is deliberately a targeted replication rather than another full AMIYC benchmark. Three independent lineages are fixed in advance, each with an epoch-6 source and targets at $+1$, $+3$, and $+5$; held-out outcomes are measured on an untouched 100k-list EVAL population. The resulting transport corpus contains nine target states and the full 28-pair census at each target.

Fresh opportunity survives this scale change cleanly. Mean Fresh gain is 0.005638, 0.008788, and 0.006166 at $+1, +3, +5$, respectively, and all nine predeclared lineage-by-horizon target states are positive; overall mean Fresh gain is 0.006864.

Same replay, however, is non-monotonic: mean gain is -0.000604 at $+1$, $+0.002040$ at $+3$, and -0.003428 at $+5$. Raw Foreign remains close to zero throughout. The state-level pair census shows that these sign changes are coherent within target states rather than being driven by a small number of pairs: for example, seed 17 has 28/28 Same-positive pairs at e7 and e9 but 0/28 at e11; seed 23 moves from 0/28 at e7 to 27/28 at e9 and back to 0/28 at e11.

This result changes the mature transport claim. Optimization history can induce a regime of local replay compatibility, but Same provenance and elapsed horizon are context rather than certificates. With only three independent Production lineages, the horizon and state patterns are treated descriptively; pair rows are not promoted to independent population replicates.

Table 14: Production targeted replication of current-state Fresh opportunity and replay compatibility. Three independent lineages use epoch-6 sources and predeclared targets at $+1, +3, +5$; target pair rows are repeated observations within lineage, not independent population replicates. Panel A reports horizon means, Panel B lineage aggregates, and Panel C the nine target states to show pair-level sign coherence.

Row	Fresh	Same	Foreign	Same/Fresh	Same−Foreign	Additional audit
A. Horizon means over three independent lineages						
$+1$ / e7	0.005638	-0.000604	-0.000105	-10.7%	-0.000499	ΔNDCG F/S/X: $+.005303/-.000222/-.000042$
$+3$ / e9	0.008788	0.002040	-0.000077	23.2%	0.002117	ΔNDCG F/S/X: $+.008768/+.002196/-.000013$

Continued on next page

Table 14 continued.

Row	Fresh	Same	Foreign	Same/Fresh	Same−Foreign	Additional audit
+5 / e11	0.006166	−0.003428	−0.000099	−55.6%	−0.003330	ΔNDCG F/S/X: +.006343/ − .003425/ + .000023
Overall	0.006864	−0.000664	−0.000094	−9.68%	−0.000571	Fresh positive at 9/9 target states
B. Aggregate by training lineage						
seed 17	0.011015	0.002371	−0.000221	21.5%	0.002591	Same positive overall
seed 23	0.004470	−0.001887	−0.000238	−42.2%	−0.001648	Same negative overall
seed 29	0.005106	−0.002476	0.000179	−48.5%	−0.002655	Same negative overall
C. Target-state coherence over the full 28-pair census						
17/e7	0.010752	0.002591	−0.000159	–	0.002749	Same-positive pairs 28/28
17/e9	0.016917	0.007411	−0.000191	–	0.007602	Same-positive pairs 28/28
17/e11	0.005378	−0.002890	−0.000313	–	−0.002577	Same-positive pairs 0/28
23/e7	0.005068	−0.003660	−0.000137	–	−0.003523	Same-positive pairs 0/28
23/e9	0.004440	0.001103	−0.000246	–	0.001349	Same-positive pairs 27/28
23/e11	0.003902	−0.003104	−0.000332	–	−0.002772	Same-positive pairs 0/28
29/e7	0.001095	−0.000741	−0.000018	–	−0.000723	Same-positive pairs 0/28
29/e9	0.005006	−0.002396	0.000205	–	−0.002600	Same-positive pairs 0/28
29/e11	0.009218	−0.004291	0.000349	–	−0.004640	Same-positive pairs 0/28

Positive values denote primary-loss improvement. The frozen transport corpus contains 9 targets and 1,008 evaluated outcome rows on untouched 100k-list EVAL. Pair-level coherence shows that the sign reversal of Same is state-wide rather than driven by one exceptional pair. With only three independent lineages, horizon and state patterns are treated descriptively; row counts are not promoted to population-level significance. Production operator-motion statistics are reported with the lifecycle/motion table, while secondary-quality tolerance curves are deferred to the deployment section because eligibility and usefulness are distinct estimands.

A.6.8 TARGET-AWARE RESCUE OF FOREIGN REPLAY AND ITS CONTROLS

Campaign 1 also applies target-gradient masking and SELECT dose choice to weak Foreign replay under the contract in Appendix A.7.5. Mean raw primary gain rises from 0.001522305 for unchanged replay to 0.007836957 for the majority mask: a paired $+0.006314652$ (95% interval $[0.0040202, 0.0092232]$), positive in all seven target-seed means. After FIT-only alignment, the corresponding values are 0.003474861 and 0.008524369, a $+0.005049508$ contrast ($[0.0029448, 0.0075870]$), again positive in all seeds. The NDCG contrasts are respectively $+0.002056400$ and $+0.001640580$. This target-aware, dose-selected procedure answers a different question from the fixed-dose Same-replay port in Appendix A.6.9.

The controls narrow the explanation. Shuffling the donor coefficient coordinates before masking gives mean gains 0.008460430 raw and 0.008486958 aligned. Majority-minus-shuffled differences are -0.000623473 and $+0.000037411$; both seed intervals include zero. Advantage over unchanged Foreign replay therefore does not establish distinctive information in the original donor coordinate placement. Matched-support/norm masking does beat its random-support control by $+0.006120377$ raw and $+0.006384790$ aligned, but this supports the tested target-informed procedure, not a unique donor mechanism. The matching applies before independent dose selection.

Table 13: Replay compatibility, family-preserving alignment, and measured-similarity audits in the public settings. Fresh is fit at the target state; Same replays a frozen operator from the same training run; Foreign replays an independent-run operator without target refitting. Amazon alignment and donor ranking use only pre-intervention FIT/ALIGN information disjoint from EVAL.

A. Frozen-operator replay

Target / regime	Fresh	Same	Foreign	Same/Fresh	Same−Foreign
Amazon, overall	0.017843	0.014007	0.002557†	78.5%	0.011451
Kuai e12	0.000648	0.000617	−0.000130	95.2%	0.000747
Kuai e13	0.000723	0.000660	−0.000125	91.3%	0.000785
Kuai e14	0.000840	0.000708	−0.000122	84.3%	0.000830
Kuai mean	0.000737	0.000662	−0.000126	89.8%	0.000787

Amazon reports eligibility-gated U_ϵ; Kuai reports raw ΔL. †The Amazon row uses FIT-only family-preserving aligned Foreign for the displayed comparison; raw Foreign is 0.001080. Amazon spans 63 evolved target states / 378 target-pair states. Descriptive Same/Fresh retention across its three horizon blocks is approximately 79.1%, 83.2%, and 73.5%; no smooth stale-decay law is inferred. Kuai Foreign is raw and unmapped. Absolute gain magnitudes are not compared across datasets.

B. Amazon family-preserving permutation alignment

Quantity	Result	Audit / support
Raw → aligned Foreign	$0.001080 \to 0.002557$; $\Delta = +0.001477$	95% CI $[0.000690, 0.002264]$; positive 6/7 seeds
Fraction of raw Same–Foreign gap closed	11.4%	about 88.6% of the raw gap remains
Same − aligned Foreign	0.011451	95% CI $[0.00875, 0.01452]$; positive 7/7 seeds
Expert correspondence	nominal CKA 0.498 → matched CKA 0.759	expert mapping identity 5.3%
Coordinate correspondence	matched absolute correlation 0.663	coordinate identity fraction 1.35%
ALIGN population	2,048 common FIT rows per relation	all 378 relations; no labels or intervention outcomes

C. Nearest-observed-Foreign similarity audits

Setting	Matching geometry	Same − nearest Foreign	Support boundary
Amazon	FIT-only matched-expert CKA + matched-coordinate $\lvert\rho\rvert$	mean 0.0137349; median 0.0128327; 95% CI $[0.0107844, 0.0172852]$; 7/7 seeds positive	378 matched target-pair relations from 2,268 Foreign replay rows
Kuai	920-D pre-intervention representation descriptor; standardized RMS distance	mean 7.4977×10^{-4}; median 6.5459×10^{-4}; 95% CI $[5.9662, 9.1623] \times 10^{-4}$; 8/8 positive	288 target relations, 2,016 Foreign rows; Same closer than nearest Foreign in 100% of relations

Kuai's measured representation supports do not overlap: Same distance mean 0.1564, max 0.2734; Foreign distance mean 1.4191, min 0.7278. The nearest-Foreign/Same distance ratio has median about 6.75 (range about 3.07–20.45). Across Foreign rows, representation similarity itself is only weakly associated with gain (Spearman about 0.02). Thus the tested similarity variables do not rescue observed Foreign replay, but the absence of common support prevents identifying a residual causal effect of lineage after exact representation matching. Amazon alignment tests only expert/coordinate permutations; unrestricted rotations or broader gauges are not ruled out.

Quality remains conditional. Identity-relative eligibility increases from 69.89% to 86.38% in the raw comparison and from 68.69% to 85.67% after alignment. These rates average dependent action outcomes; they are not prospective bank-policy coverage or population safety guarantees. The full method table includes mean-sign, few-shot, scalar, random and historical Same/Fresh references. In particular, the latter retain substantial primary gain with low quality eligibility.

The repeated Same/Fresh cells do not provide thousands of independent target refits. Deduplicating donors and modes leaves 378 target–pair references. Their primary values agree with the baseline transport release to at most 1.2×10^{-7}. Two Fresh NDCG values differ between exports, by up to 0.000474, with no eligibility change. We retain the original values for each protocol and record the unresolved source discrepancy; neither export silently replaces the other.

Table 15: Campaign 1 Foreign procedures: full ladder. Equal target-seed means after replicate/action/state aggregation.

Family	Raw G_L	Raw G_Q	Pass %	Aligned G_L	Aligned G_Q	Pass %
Identity	+0.000000	-0.000000	100.00	+0.000000	-0.000000	100.00
Short local refit	+0.004071	+0.000400	91.89	+0.004098	+0.000336	92.15
Archived Fresh	+0.035538	-0.011068	58.99	+0.035538	-0.011068	58.99
Majority mask	+0.007837	+0.001759	86.38	+0.008524	+0.001568	85.67
Matched majority (avg.)	+0.007676	+0.001728	86.35	+0.008354	+0.001519	85.39
Mean-sign mask	+0.008237	+0.001119	82.85	+0.008932	+0.001243	83.95
Random support (avg.)	+0.001556	+0.000556	91.18	+0.001969	+0.000739	89.90
Historical Same	+0.027652	-0.003112	61.38	+0.027652	-0.003112	61.38
Scalar attenuation	+0.002301	+0.000640	88.05	+0.002969	+0.001022	89.02
Shuffled majority (avg.)	+0.008460	+0.001823	85.63	+0.008487	+0.001620	85.98
Unchanged Foreign	+0.001522	-0.000298	69.89	+0.003475	-0.000073	68.69

Each ordinary family has 2,268 actions per mode; random/shuffled families have three raw replicas per action. Pass is EVAL identity-relative quality eligibility, not policy coverage. Fixed Same/Fresh references are repeated across donor/mode slots, not new fits. The pilot is contained in the full ladder.

Table 16: Campaign 1 Foreign procedures: nested pilot. Equal target-seed means after replicate/action/state aggregation.

Family	Raw G_L	Raw G_Q	Pass %	Aligned G_L	Aligned G_Q	Pass %
Identity	+0.000000	-0.000000	100.00	+0.000000	-0.000000	100.00
Short local refit	+0.004824	+0.000426	90.62	+0.005658	+0.000400	89.58
Archived Fresh	+0.040880	-0.003928	45.83	+0.040880	-0.003928	45.83
Majority mask	+0.007414	+0.002074	90.62	+0.008887	+0.001564	76.04
Matched majority (avg.)	+0.007285	+0.001956	89.58	+0.008670	+0.001554	77.43
Mean-sign mask	+0.007424	+0.000745	84.38	+0.009497	+0.000239	71.88
Random support (avg.)	+0.001099	+0.000165	94.10	+0.001440	-0.000048	87.85
Historical Same	+0.031280	+0.001151	45.83	+0.031280	+0.001151	45.83
Scalar attenuation	+0.001189	+0.000480	89.58	+0.001458	+0.000271	90.62
Shuffled majority (avg.)	+0.008748	+0.002122	82.29	+0.009067	+0.002732	88.54
Unchanged Foreign	-0.001191	-0.000181	62.50	+0.001770	-0.001573	59.38

Each ordinary family has 96 actions per mode; random/shuffled families have three raw replicas per action. Pass is EVAL identity-relative quality eligibility, not policy coverage. Fixed Same/Fresh references are repeated across donor/mode slots, not new fits. The pilot is contained in the full ladder.

Table 17: Campaign 1 prespecified Foreign contrasts. Positive differences favor the left procedure.

Mode	Left	Right	Metric	Mean	95% interval	+ seeds
aligned	Maj.	Unchanged	G_L	+0.005050	[+0.0029448, +0.0075870]	7/7
aligned	Maj.	Unchanged	G_Q	+0.001641	[+0.0009133, +0.0023411]	6/7
aligned	Maj.	Local refit	G_L	+0.004426	[+0.0034741, +0.0057552]	7/7
aligned	Maj.	Local refit	G_Q	+0.001232	[+0.0006937, +0.0018469]	7/7
aligned	Maj.	Shuffled	G_L	+0.000037	[-0.0002983, +0.0003632]	4/7
aligned	Maj.	Shuffled	G_Q	-0.000053	[-0.0003146, +0.0001643]	3/7
aligned	Matched	Random	G_L	+0.006385	[+0.0046654, +0.0086688]	7/7
aligned	Matched	Random	G_Q	+0.000780	[+0.0003573, +0.0012418]	6/7
raw	Maj.	Unchanged	G_L	+0.006315	[+0.0040202, +0.0092232]	7/7
raw	Maj.	Unchanged	G_Q	+0.002056	[+0.0005719, +0.0039422]	6/7
raw	Maj.	Local refit	G_L	+0.003765	[+0.0029635, +0.0046525]	7/7
raw	Maj.	Local refit	G_Q	+0.001359	[+0.0004066, +0.0024672]	7/7
raw	Maj.	Shuffled	G_L	-0.000623	[-0.0013998, +0.0000503]	2/7
raw	Maj.	Shuffled	G_Q	-0.000065	[-0.0005572, +0.0005129]	2/7
raw	Matched	Random	G_L	+0.006120	[+0.0046816, +0.0079599]	7/7
raw	Matched	Random	G_Q	+0.001172	[+0.0005517, +0.0018529]	7/7

Intervals are original 20,000-draw target-seed bootstrap intervals, conditional on the fixed donor bank and shared data. Both majority–shuffled intervals include zero. All prespecified contrast vectors and means are separately retained and reconstructed; no multiplicity correction is implied.

A.6.9 TARGET-GRADIENT TRANSPORT ON ALREADY-USEFUL REPLAY

The GradFix-style experiment tests a specific alternative to unchanged Same replay, not another Fresh fit. For each previously fitted source operator, Raw replays its log-diagonal coefficients unchanged at the target state. The port uses coordinate-wise anti-gradient sign votes at the identity from 20 target-side FIT units to apply the frozen `max`-majority, `normal` hard mask to those source coefficients. It retains the source-selected expert support and direction. It takes no optimization steps at the target, does not choose an action from target EVAL, and evaluates the masked and unchanged operators on the same untouched target population. Kuai's FIT units are balanced scalar examples; Amazon's are complete lists with one positive and 19 negatives. Kuai reuses two deterministic, task-specific 20-example sets across its target states; Amazon reuses the same 20 positional list identifiers across 63 target states. Repeated votes and action rows are not independent inferential units; some probe populations differ, so identical positional identifiers need not denote identical list content. This is an operator-space adaptation of GradFix (Rinaldi et al., 2026), not a claim that the original full-model method has been reproduced unchanged. In particular, the published method tunes an application scale using validation data; our frozen probe fixes dose at one for both Raw and masked replay. It does not test the full validation-scaled method or establish a general failure of GradFix.

Specifically, for source log-diagonal $\boldsymbol{z}^{\mathrm{src}}$ and native loss ℓ_n of target FIT unit n, the frozen implementation computes

$$v_j = \operatorname{sgn}\left(\sum_{n=1}^{20} \operatorname{sgn}\left(-\left.\frac{\partial \ell_n}{\partial z_j}\right|_{\boldsymbol{z}=\mathbf{0}}\right)\right), \qquad z_j^{\mathrm{GF}} = z_j^{\mathrm{src}} \mathbf{1}\left[\operatorname{sgn}(z_j^{\mathrm{src}}) = v_j\right]. \tag{27}$$

Thus `max` means a simple majority of per-unit signs, with no vote-margin threshold: a zero sum returns $v_j = 0$, and a tied coordinate with a nonzero source update is zeroed. Source coordinates already at identity remain at identity. Amazon uses the gradient of its masked item-wise binary-cross-entropy *averaged within each list*; Kuai uses one scalar loss per example. The archived matched rows give mean zero-vote fractions of 15.23% and 3.63% of coordinates for Kuai and Amazon, respectively. The original execution code and archived rows were checked against the addendum's source hashes; the raw execution archives are not distributed, but anonymized matched rows and sign-sufficient mask records are included in the v3.2 numerical release (Appendix D.1.4).

On Kuai epochs 12–14, eight independent seed means all favor unchanged replay on primary loss: paired GradFix-minus-Raw gain is -4.49×10^{-4} (seed-bootstrap 95% CI $[-5.68, -3.41] \times 10^{-4}$). The mask retains 39.9% of source coordinates. On Amazon's 378 matched actions across seven seeds, the mean paired difference is -4.22×10^{-4} (95% CI $[-8.54, -1.23] \times 10^{-4}$; six of seven seed means negative), while the mask retains 84.2% of coordinates. Exact two-sided seed sign-flip p values are 0.0078125 (Kuai) and 0.03125 (Amazon). These are native-loss comparisons within each target cohort. Their no-op losses are 0.606102 and 0.309922, respectively, and are not Table 1A's original Fresh-population baselines.

Amazon's secondary response has the other sign: on the same 378 actions, the paired GradFix-minus-Raw Next NDCG@10 difference is $+0.005566$ (seed-bootstrap 95% CI $[+0.001403, +0.011677]$; seven of seven seed means positive; exact seed sign-flip $p = 0.015625$). Mean eligible primary value rises from 0.014007 to 0.018156 at the prespecified NDCG tolerance $\epsilon = 0.002$. Eligibility is observed on EVAL; it was not enforced before masking or execution. The NDCG comparison is exploratory and unadjusted for multiplicity; the eligibility and gated-value changes are descriptive (their exact two-sided seed sign-flip p values are 0.0625 and 0.078125). Recomputing eligibility at NDCG floors 0 and -0.005 also favors the mask (Raw 53.97% vs. GradFix 63.49%, and Raw 68.25% vs. GradFix 79.63%); these are alternative summaries of the same 378 actions. The experiment modifies already-useful Same replay, so it does not establish rescue of harmful Same or Foreign operators, nor a general failure of target-aware transport. The master report records mean retained source log-norm of 60.91% (Kuai) and 95.42% (Amazon). Their distinct gradient units and reused FIT sets preclude a causal explanation for this cross-domain difference.

Cross-setting synthesis. Across Amazon, Kuai, and Production, the robust invariant is narrower than universal trajectory-local persistence. Current-state conditional adaptation headroom replicates; fitted references differ across learned states, without isolating fitting variability; replay can retain substantial local value, lose it, and later realign; and raw cross-lineage portability is weak in

the tested family. A tested target-gradient mask does not improve mean primary reuse in the two local Same-replay cohorts, although Amazon's secondary ranking response improves. The available functional and representation similarities do not rescue observed Foreign operators, but neither do they establish lineage as an independent causal variable. The object carried forward into the tracking and deployment analyses is therefore *current-state compatibility*.

Table 18: Target-gradient masking of unchanged Same replay on matched target states. Primary comparisons use native losses; the Amazon secondary endpoint is exploratory.

A. Primary loss improvement ΔL

Setting	Fresh	Raw	GradFix	GradFix − Raw	Negative seeds
Kuai e12–e14	$+0.000737$	$+0.000662$	$+0.000213$	-0.000449	$8/8$
Amazon Next	$+0.035538$	$+0.027652$	$+0.027230$	-0.000422	$6/7$

B. Amazon exploratory secondary response, 378 paired actions

Quantity	Raw	GradFix	Difference
Next ΔNDCG@10	-0.003112	$+0.002454$	$+0.005566$
Eligible (ΔNDCG@10 ≥ -0.002)	$232/378$	$278/378$	$+12.17$ pp
U_ϵ, observed eligible primary value	$+0.014007$	$+0.018156$	$+0.004149$

Scope. Twenty target-FIT gradient units are used per action; no target optimization or EVAL-based action choice. A: seed-bootstrap 95% intervals for the paired differences, in 10^{-4} loss units, are $[-5.68, -3.41]$ (Kuai) and $[-8.54, -1.23]$ (Amazon). B: paired NDCG interval $[+0.001403, +0.011677]$, unadjusted for multiplicity. The two domains and the original Table 1A have distinct cohorts.

A.7 Moving-Reference Tracking under Continued Optimization

The frozen-state results establish that a bounded current-state adaptation can carry held-out value. They do not establish that ordinary continued training will preserve, improve, or track that adaptation while the underlying model state changes. The Amazon continuation program therefore separates three estimands that are easy to conflate: persistence of a previously selected intervention relative to its matched evolved no-op branch; endpoint geometry of online coefficient adaptation relative to a contemporaneous Fresh refit; and fixed-state coefficient optimization after the evolved representation is frozen.

The historical conclusion is narrower than an intrinsic "objective mismatch." Ordinary task supervision contains useful local coefficient information when the representation is fixed, while neither of the two original online schedules tracks its contemporaneous Fresh reference. These outcomes establish failure of those recipes; they do not identify representation motion as its cause. Later repeat-fitting controls, actual live adaptation, and crossed/shadow comparisons further restrict that interpretation (Appendices B.3, B.5.1, B.5.2, and B.5.4).

A.7.1 Matched continuation: value can persist without action identity

The Phase-A continuation panel begins from 21 Amazon checkpoints: seven seeds crossed with the three prespecified roles *best-primary*, *best-multitask*, and *last*. Each checkpoint is continued for three epochs under matched arms, and every intervention effect is defined relative to the evolved no-op branch from the same starting checkpoint. This control is essential because ordinary continuation itself improves the model; comparing an intervened endpoint to its pre-continuation source would overstate persistence.

The selected pair-local intervention remains useful after continuation. At the three-epoch endpoint, selected-minus-no-op Next loss gain is 0.0114932 (95% seed-cluster CI $[0.0087051, 0.0151824]$), positive in all seven seeds. Next NDCG@10 improves by 0.00392 (95% CI $[0.00196, 0.00563]$), and HR@10 by approximately 0.01007 (95% CI $[0.00603, 0.01397]$).

The behavioral endpoint is more stable than exact action identity. Selected and the frozen loss-oracle arm have almost identical aggregate primary continuation value descriptively: selected-minus-oracle loss gain is $+0.0001029$ with 95% CI $[-0.0004340, 0.0009415]$, even though they select the same pair and direction at only $7/21$ starting checkpoints. This unresolved contrast does not establish equivalence. Pair-local choices also have similar final Next-loss effects but materially different listwise behavior: selected beats random by $+0.00231$ NDCG@10 (95% CI $[0.00064, 0.00459]$) and bottom by $+0.00409$ (95% CI $[0.00122, 0.00720]$).

The immediate-to-final attenuation is largely shared across intervention arms. Selected retains about 40.4% of its initial matched advantage after three epochs; the oracle arm retains 41.0%. In the five-seed observer subset, the selected operator itself changes very little over continuation ($\|\Delta_3\|_2/\|\Delta_0\|_2 \approx 1.0036$ and cosine ≈ 0.99819), while approximately 96% of its instantaneous loss effect remains at epoch three. These observer statistics are useful persistence measurements, but their support is 15 checkpoints rather than the full 21.

A.7.2 Historical online coefficient adaptation does not track Fresh

A second continuation condition makes the selected diagonal coefficients trainable while the base MMoE continues to evolve under the ordinary MTL objective. This Phase-B intervention remains beneficial relative to matched no-op continuation: Next loss gain is 0.0103416 with seed-cluster 95% CI $[0.00745, 0.01389]$. It is nevertheless consistently weaker than the frozen Phase-A selected intervention. Phase-B minus Phase-A Next loss gain is -0.001152 (95% CI $[-0.001378, -0.000931]$), with negative differences in all seven seeds and all three checkpoint roles. Next NDCG@10 and HR@10 show the same direction of change.

The later contemporaneous-direction audit asks the more precise question. At the evolved endpoint, the net coefficient displacement produced by historical joint base/coefficient co-adaptation is compared with an independently re-fit Fresh direction at that same frozen endpoint. Across 63 matched evolved states, mean cosine is -0.316871, distance progress -0.153148, and target projection -0.105758; the seed-level means are negative in all seven seeds. These are endpoint

Table 19: Matched Amazon continuation separates behavioral persistence from exact action identity. All Phase-A effects in Panels A–B are measured against the matched evolved no-op endpoint from the same starting checkpoint. The full panel contains 21 checkpoints (7 seeds $\times$ 3 prespecified roles); observer-vector persistence in Panel C is available on 15 checkpoints from five seeds. Task-global is included as a scope control and used substantially larger normalized dose, so its absolute gain is not an intrinsic scope comparison.

A. Three-epoch continuation effects relative to matched no-op

Arm	ΔL_{Next}	$\Delta\mathrm{NDCG@10}_{\mathrm{Next}}$	$\Delta\mathrm{HR@10}_{\mathrm{Next}}$	Role in audit
Bottom pair	$+0.01171$	-0.00018	$+0.00367$	uninformed low-rank pair
Random pair	$+0.01244$	$+0.00160$	$+0.00612$	uninformed pair control
Selected	$+0.01149$	$+0.00392$	$+0.01007$	calibrated pair-local action
Frozen loss oracle	$+0.01139$	$+0.00399$	$+0.01027$	reference-derived pair-local action
Task-global	$+0.03219$	$+0.00535$	$+0.01865$	broader / higher-dose scope control

B. Paired inference and action-identity boundary

Contrast / quantity	Estimate	95% CI / support	Reading
Selected $-$ no-op, Next loss	$+0.0114932$	$[0.0087051, 0.0151824]$; 7/7 seeds > 0	persists
Selected $-$ no-op, Next NDCG@10	$+0.00392$	$[0.00196, 0.00563]$	listwise gain
Selected $-$ no-op, Next HR@10	$+0.01007$	$\approx [0.00603, 0.01397]$	retrieval gain
Selected $-$ oracle, Next loss	$+0.0001029$	$[-0.0004340, 0.0009415]$	unresolved
Selected $-$ random, Next NDCG@10	$+0.00231$	$[0.00064, 0.00459]$	higher NDCG
Selected $-$ bottom, Next NDCG@10	$+0.00409$	$[0.00122, 0.00720]$	higher NDCG
Selected/oracle exact pair+direction agreement	$7/21$	starting checkpoints	identity $\neq$ value

C. Immediate-to-final persistence

Quantity	Selected	Oracle	Support / boundary
Immediate matched advantage	0.02842	0.02776	full Phase-A effect summary
Final matched advantage	0.01149	0.01139	full 21-checkpoint panel
Fraction retained	0.404	0.410	no-op closes much of initial gap
Operator norm ratio $\|\Delta_3\|_2/\|\Delta_0\|_2$	1.0036	–	15-checkpoint observer subset
Operator cosine $\cos(\Delta_3, \Delta_0)$	0.99819	–	15-checkpoint observer subset
Instantaneous loss effect retained at epoch 3	$\approx 96\%$	–	observer subset; descriptive

measurements of net adaptation. They are not claims that every minibatch gradient is anti-aligned with Fresh. The compact release retains reported geometry summaries but not their complete vector origins and normalization implementation; exact geometry recomputation remains unavailable.

A.7.3 Fixed-state ordinary MTL contains the useful coefficient signal

The strongest static-objective control freezes each evolved representation and optimizes only the same coefficient coordinates using ordinary MTL task supervision. At fixed state the geometry reverses: mean cosine to Fresh is $+0.602$, distance progress $+0.213$, and target projection $+0.375$, with all seven seed blocks switching from negative online-joint geometry to positive fixed-state geometry.

The behavioral result is equally important. Frozen coefficient-only MTL improves primary loss by 0.010477 relative to the historical joint endpoint (95% CI $[0.008430, 0.012833]$). Fresh retains a smaller primary-loss advantage of 0.002003 (95% CI $[0.001497, 0.002826]$), while Fresh-minus-

frozen-coefficient eligible utility is -0.001958 with interval $[-0.006379, 0.001190]$, not clearly separated from zero. The frequently quoted 83.9% gap-recovery ratio is therefore descriptive: this is a post-hoc fixed-state control, not a causal decomposition of the historical online trajectory.

This control falsifies the strong claim that the ordinary task objective categorically lacks local information about the useful bounded coefficient direction. It does not establish that the particular fixed-state optimizer or convergence budget is unique.

A.7.4 PRESPECIFIED ONLINE BLOCK-COORDINATE DECOUPLING DOES NOT RESCUE TRACKING

A prespecified R1 control tests a narrower online explanation: perhaps the historical anti-tracking is caused primarily by simultaneous base/coefficient update ordering. The block-coordinate schedule separates the two online updates within the same minibatch while preserving ordinary MTL supervision and the predeclared continuation budget. No learning-rate, schedule, or optimizer sweep is performed after observing the result.

The schedule changes the evolved base state but leaves tracking geometry almost unchanged. Joint versus block-coordinate cosine is -0.316871 versus -0.317440; distance progress is -0.153148 versus -0.152646; target projection is -0.105758 versus -0.105528. The block-minus-joint differences all have seed-bootstrap intervals crossing zero. Block-coordinate cosine remains negative in $7/7$ seeds, and Fresh still exceeds the block-coordinate endpoint on primary loss by 0.012498 (95% CI $[0.010031, 0.015585]$), again positive in all seven seeds.

This is not a null perturbation of the base trajectory. Relative to the historical joint endpoints, the decoupled schedule changes mean ranking loss by about $+0.01226$, Next loss by $+0.00254$, Next NDCG by -0.00735, and future-intent loss by about -0.00924. Despite this altered task tradeoff, coefficient tracking remains essentially the same. The Fresh gap also grows modestly over the three audited continuation horizons while distance progress and target projection become more negative. We treat that temporal pattern descriptively; it is not a fitted timescale law.

A.7.5 CAMPAIGN 1: FROZEN-STATE RESCUE AND REPLAY CONTRACT

The first post-baseline campaign rereads archived Amazon continuation states; it performs no new continuation training. Its full ladder uses seven target seeds $\{13, 17, 23, 29, 31, 37, 41\}$, the roles best-primary, best-multitask and last, and horizons $+1, +2, +3$. There are 63 rescue states, 63 Foreign-replay states, and 21 crossed base/coefficient jobs. The two-seed pilot has eight rescue, eight Foreign and four swap jobs. All 3,376 pilot non-swap EVAL rows repeat exactly in the full ladder, including selected-coefficient hashes. The pilot is retained for provenance and is not additional independent replication.

Let B denote the fixed readout base, z_a its held historical coefficients, and z the candidate. Here

$$G_L(z) = L_B(0) - L_B(z), \quad dL(z) = L_B(z_a) - L_B(z),$$
$$G_Q(z) = Q_B(z) - Q_B(0), \quad dQ(z) = Q_B(z) - Q_B(z_a).$$

The primary public task is Next (called `trx` by the implementation), with native masked item BCE averaged within complete lists. Quality is Next NDCG@10, with absolute tolerance -0.002; HR@10 is retained separately. The exported Campaign 1 `utility` is $G_L\,\mathbf{1}[G_Q \geq -0.002]$, not the later campaign's incumbent-relative dual-floor utility. Every table below reports raw gains and eligibility separately.

The rescue anchor is the held coefficient vector evaluated on the same evolved base as the archived online vector. Their difference does not compare the full historical online training path with the separate frozen-coefficient continuation path. The Foreign anchor is identity. For each of six donors, six expert pairs are examined in raw and FIT-aligned coordinates. This yields 2,268 state–donor–pair actions per mode and ordinary family. Each action has its own SELECT choice; the reported Foreign result averages these procedures and is not a selector choosing the best donor or action bank.

The inherited implementation defines a 64-coordinate effective log diagonal bounded by $|z_j| \leq 0.5$. It estimates target anti-gradient signs from 20 complete FIT lists, each with one positive and nineteen negatives. The majority mask keeps coordinates of the proposed change $z - z_a$ whose nonzero signs agree with the target votes at z_a; ties are dropped. The mean-gradient mask is a separate

Table 20: Amazon moving-reference tracking audit. Panel A records the historical behavioral precursor: allowing the selected coefficients to co-adapt online remains beneficial versus no-op but is consistently weaker than leaving the selected Phase-A operator frozen. Panels B–C then compare each online endpoint with its own independently re-fit contemporaneous Fresh direction. Only the fixed-state coefficient-only ordinary-MTL control moves positively toward Fresh. The block-coordinate schedule changes the evolved base state, so its similarity to historical joint anti-tracking is not explained by the two schedules ending at an identical model state.

A. Historical continuation precursor: frozen selected action vs online-trainable coefficients

Metric	Phase A	Phase B	Phase B − Phase A (95% CI)
Next loss gain vs no-op	$+0.01149$	$+0.01034$	-0.001152 $[-0.001378, -0.000931]$
Next NDCG@10 gain vs no-op	$+0.00392$	$+0.00340$	-0.000516 $[-0.000922, -0.000155]$
Next HR@10 gain vs no-op	$+0.01007$	$+0.00855$	-0.001517 $[-0.002521, -0.000669]$
MTL mean NDCG gain	$+0.00132$	$+0.00114$	-0.000172 $[-0.000307, -0.000051]$

Phase-B Next-loss gain remains positive versus no-op (95% CI approximately $[0.00745, 0.01389]$), but the Phase-B minus Phase-A loss difference is negative in all seven seeds and all three checkpoint roles. This historical result motivated the contemporaneous-direction audit; it is not itself proof of an intrinsic objective mismatch.

B. Geometry relative to each endpoint's contemporaneous Fresh direction

Condition	Cosine	Dist. progress	Target proj.	Seed sign
Historical joint online MTL	-0.316871	-0.153148	-0.105758	7/7 negative
Online block-coordinate MTL	-0.317440	-0.152646	-0.105528	7/7 negative
Frozen-base coeff-only MTL	$+0.602$	$+0.213$	$+0.375$	7/7 positive

C. Behavioral and falsification contrasts

Contrast	Estimate	95% CI / support
Frozen coeff-MTL − historical joint, primary loss	$+0.010477$	$[0.008430, 0.012833]$
Fresh − frozen coeff-MTL, primary loss	$+0.002003$	$[0.001497, 0.002826]$
Fresh − frozen coeff-MTL, eligible utility	-0.001958	$[-0.006379, 0.001190]$
Fresh − block-coordinate, primary loss	$+0.012498$	$[0.010031, 0.015585]$; 7/7 positive
Block-coordinate − historical joint, cosine	-0.000569	$[-0.002623, 0.001332]$
Block-coordinate − historical joint, dist. progress	$+0.000503$	$[-0.000359, 0.001293]$
Block-coordinate − historical joint, target projection	$+0.000229$	$[-0.000600, 0.000971]$
Block-coordinate − original stale operator, primary loss	-0.002404	$[-0.002785, -0.002044]$; 7/7 negative

D. R1 schedule perturbation and horizon pattern (descriptive)

Quantity	Step / horizon 1	Step / horizon 2	Step / horizon 3
Fresh − block-coordinate primary gap	0.01195	0.01259	0.01296
Distance progress	-0.075	-0.154	-0.229
Target projection	-0.061	-0.109	-0.146

The R1 block-coordinate schedule also changes the evolved task tradeoff relative to historical joint endpoints (mean ranking-loss change $+0.01226$, Next-loss change $+0.00254$, Next-NDCG change -0.00735, and future-intent-loss change about -0.00924). Thus the near-identical anti-tracking geometry is not attributable to the schedules producing the same base state. Panel D is descriptive and is not used to estimate a universal tracking timescale.

procedure. SELECT chooses a dose from $\{0, 1/4, 1/2, 3/4, 1\}$ by minimum native loss among candidates satisfying the identity-relative quality floor. When none qualifies, the family anchor is retained with its failure recorded. A zero dose in rescue therefore returns the historical anchor, not identity. The local refit is projected Adam starting at that anchor; retained selected snapshots have $0, 1, 5, 10$, or 20 updates. The repaired records do not contain the exact original plan's learning rate or penalty coefficient, so those constants are not inferred from the later campaign.

The random-support control keeps the same active-coordinate count and contracts both endpoints to the smaller common norm before dose selection. Independently selected doses need not retain equal final norms. Three random controls and three coefficient-permutation controls are averaged, never

selected by best replicate. Permutation preserves the source value multiset before target masking; it does not by itself match the selected mask, dose or endpoint. Same and Fresh references retain their historical directions and have no new quality filter.

EVAL contains 600 lists for every rescue state and 600–610 for Foreign states. The reported FIT budget excludes SELECT labels and any historical training or anchor fitting. Exact original SELECT memberships and full search grids are unavailable in this repaired subset. The source-reported FIT/SELECT/EVAL separation and native execution checks are distinguished from the present row-level arithmetic audit in Appendix D.1.5.

Aggregation first averages control replicates within an action, actions within a state, and states within a target seed, then weights the seven seeds equally. States and donor actions share data and history. The original prespecified contrasts retain their 20,000-draw seed-bootstrap intervals; their seed vectors and means are independently reconstructed here. The retrospective rescue and swap summaries use 10,000 draws with RNG seed 20260925, and those intervals are also reconstructed. All are exploratory, conditional on the shared evaluation data and donor bank, without simultaneous-error control.

A.7.6 Behavioral failure of the archived online coefficients

Campaign 1 supplements the geometric comparison with a same-base behavioral test. The archived online coefficients have lower primary gain than the held coefficients in all $63/63$ evolved states: mean $dL = -0.002410508$, seed-bootstrap 95% interval $[-0.0027886, -0.0020541]$, with negative means in all seven seeds. The state differences range from -0.005366474 to -0.000305444. This is a relative failure: mean gain over identity remains $+0.026499$, compared with $+0.028910$ for the held anchor. It does not mean that every online endpoint is worse than identity, and it does not contradict a positive complete-system continuation gain measured against another base trajectory.

A SELECT-chosen local refit starting at the anchor raises loss gain by $+0.004837267$ (95% interval $[0.0031396, 0.0068265]$), with positive increments in all 63 states and all seven seed means. Mean NDCG improvement relative to the anchor is $+0.001420772$. The archived Fresh reference gains a further $+0.010069113$ over the anchor, but fails the identity quality floor in $6/63$ states. Both the held anchor and short refit fail that floor in $2/63$ states. A separate, retrospective incumbent-floor diagnostic finds three refit failures; the union is $4/63$. This diagnostic was not Campaign 1's SELECT rule. Positive mean quality and loss therefore do not certify each selected action.

The majority-sign mask yields a smaller $+0.000629423$ increment over the anchor (95% interval $[0.0002866, 0.0010628]$). Mean-sign masking yields $+0.000730319$. Scalar attenuation gives -0.000060668 and the averaged random-support control -0.000003957. These bounded controls establish useful local directions under the tested frozen-state procedures, without identifying a unique optimizer or a solution for live training. The complete rescue and nested-pilot tables retain all families, including negative controls.

A.7.7 Crossed bases do not reverse the coefficient deterioration

The early/late swap uses the archived learned coefficients at horizons 1 and 3 on both frozen bases, with a source-reported common native-input fingerprint. For the early base, replacing early with late coefficients changes primary gain by -0.002007367 (95% interval $[-0.0023697, -0.0016396]$); on the late base the change is -0.001843677 ($[-0.0021584, -0.0015289]$). Both contrasts are negative in all seven seed means and $20/21$ individual lineage-role records; they are not uniformly negative in all 21 records.

The interaction, late-base contrast minus early-base contrast, is $+0.000163690$ ($[0.0001099, 0.0002190]$). It reduces the mean harm on the later base without reversing it. This crossed readout tests a state-by-coefficient interaction. It does not establish beneficial coadaptation or causally explain which training component produced the coefficients or base states.

Optimization synthesis. The continuation and tracking experiments identify different objects. A previously selected bounded action can retain meaningful value through ordinary training, and exact pair/direction identity is not required for that persistence. At the same time, online adaptation of the

Table 21: All Campaign 1 rescue families, including the nested pilot. Positive dL and dQ favor the candidate over held coefficients on the same frozen base.

Cohort	Family	G_L	dL	dQ	Base fails
full	Held anchor	+0.028910	+0.000000	+0.000000	2/63
full	Short local refit	+0.033747	+0.004837	+0.001421	2/63
full	Archived Fresh	+0.038979	+0.010069	+0.001016	6/63
full	Majority mask	+0.029539	+0.000629	+0.000249	2/63
full	Matched majority (avg.)	+0.029523	+0.000614	+0.000242	6/189
full	Mean-sign mask	+0.029640	+0.000730	+0.000110	2/63
full	Identity	-0.000000	-0.028910	-0.010699	0/63
full	Random support (avg.)	+0.028906	-0.000004	+0.000036	6/189
full	Scalar attenuation	+0.028849	-0.000061	-0.000004	2/63
full	Archived online	+0.026499	-0.002411	-0.000802	2/63
pilot	Held anchor	+0.033410	+0.000000	+0.000000	0/8
pilot	Short local refit	+0.039812	+0.006402	+0.002760	0/8
pilot	Archived Fresh	+0.045220	+0.011810	+0.004682	0/8
pilot	Majority mask	+0.034341	+0.000931	+0.000296	0/8
pilot	Matched majority (avg.)	+0.034308	+0.000898	+0.000297	0/24
pilot	Mean-sign mask	+0.034340	+0.000930	+0.000280	0/8
pilot	Identity	-0.000000	-0.033410	-0.016962	0/8
pilot	Random support (avg.)	+0.033412	+0.000002	+0.000030	0/24
pilot	Scalar attenuation	+0.033410	+0.000000	+0.000000	0/8
pilot	Archived online	+0.030695	-0.002715	-0.000761	0/8

G_L is identity-relative primary gain; failures use identity-relative NDCG below -0.002. Full: 63 states, seven seeds; pilot: eight states, two nested seeds. Averaged random controls retain three raw replicas (189/24 rows). Incumbent-floor failures are reported as a separate diagnostic in the text. All effects are absolute units.

Table 22: Campaign 1 crossed readout: 21 lineage-role records within seven seeds.

Base	Coefficients	G_L	G_Q	Base pass
early	early	+0.028675	+0.010754	0.9524
early	late	+0.026668	+0.009664	0.9524
late	early	+0.026289	+0.010211	1.0000
late	late	+0.024445	+0.009398	1.0000

Each gain is relative to identity on that row's base, not to another base. The four aggregate cells and all paired state differences survive; original individual cell metrics are unavailable. Differences of these means reproduce the paired-effect means.

exposed coefficient coordinates does not move toward the contemporaneous Fresh reference under either tested ordinary MTL schedule. When the evolved representation is held fixed, however, the same task supervision drives those coordinates strongly toward Fresh and recovers most of the primary gap. These historical failures remain measured outcomes of the specified schedules. The later campaigns recover useful live benefit and associate the stronger deferred-versus-coupled endpoint comparison predominantly with the base path; fitting variability and objective-dependent quality remain separate constraints (Appendix B). No unique dynamical cause of the historical failure is identified.

A.8 Amortization, Bank Compatibility, and Deployment

Frozen-bank recovery uses pre-intervention information to select among previously acquired actions or abstain. Current-state fitting provides the complementary route, spending target supervision and computation to acquire a new action. Amazon exhibits a compatible-bank regime in which witnesses improve selection; Kuai reveals loss of available value before selection; Production separates secondary eligibility from primary usefulness.

Figure 3 locates this frozen-bank route alongside the separately paid target-fitting route examined in Appendix B.7.3. The two routes retain different information, candidate and fallback contracts.

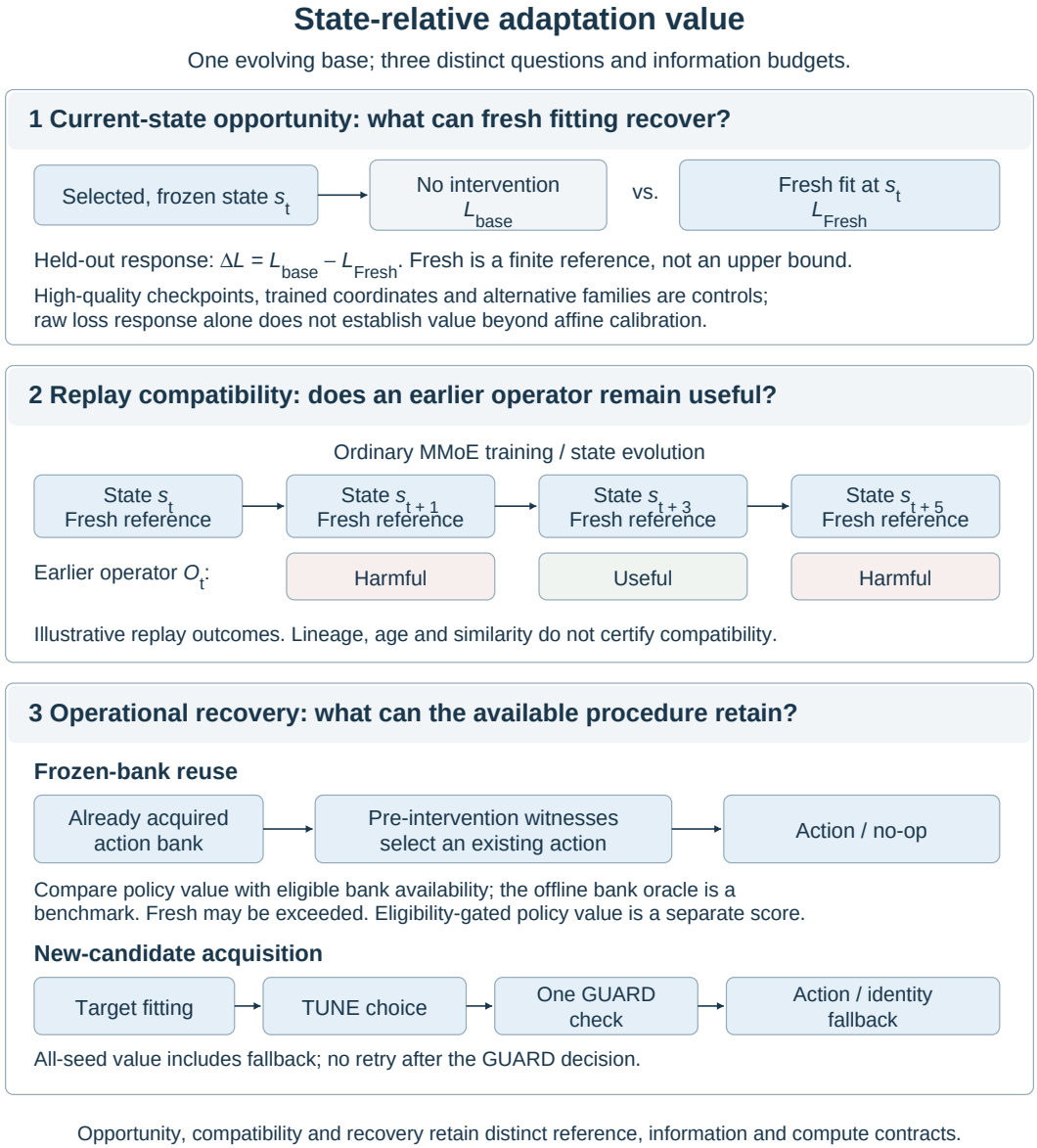

Figure 3: **Adaptation value under state evolution.** Current-state opportunity, replay compatibility, and operational recovery are distinct questions. Frozen-bank reuse and target-state candidate acquisition access different information and incur different computation costs; either route may withhold an intervention. Replay outcomes are illustrative.

A.8.1 ORIGINAL AMAZON BANK CONFIGURATION

The original Amazon supervision exporter defines the eligibility rule: NDCG gain ≥ -0.002. The controller uses three run-grouped inner out-of-fold folds on outer-training data, then isotonic probability calibration on action-family $\times$ intervention-scope surfaces, with pooled fallback for sparse surfaces. Downstream policy simulation selects the cutoff maximizing eligibility-gated realized gain subject to observed violations among acted panels $\leq 5\%$. The scorer is refitted on all outer-training data; the calibrator and cutoff are frozen for outer-test use. These settings define the original frozen-bank controller, independently of the later TUNE/GUARD acquisition protocol.

The original `calibration_transfer.csv` analysis records the fitted temporal-forward cutoffs:

Scorer	Epoch 2	Epoch 3
Witness + action HistGB	0.85	0.90
Action-only HistGB	0.95	0.95
Witness + action MLP	0.95	0.85

On the same 122 temporal-forward panels, panel-weighted original fold-policy rows give raw loss gain 0.0368247736, eligibility-gated policy value 0.0359399954, and eligible bank value 0.0477095157. Coverage is $102/122$ (83.61%), with three violations among 102 acted panels (2.94%) and 75.33% bank capture. The archived one-sided 95% iid-panel risk upper bound is 7.43%: the point-estimate budget is met, but statistical certification is not established. Fresh was not evaluated on these exact panels.

The original v1.0.2 controller, fold manifests, model records, and isotonic calibration manifests identify this execution. All seven result members in the original final manifest pass their recorded hash checks, including candidate metrics, policy summaries, threshold selection, candidate predictions, out-of-fold calibration predictions, policy decisions, and the run report. The source population contains 3,384 candidates in 183 states from 15 runs and seven seeds. Temporal-forward folds train on 61 states and test on 61 at epoch 2, then train on 122 and test on 61 at epoch 3, producing the 122 held-out panels. The policy probabilities are calibrated for eligibility; this is distinct from the output-logit affine controls in Appendix B.7.2.

A.8.2 WITNESS CONSTRUCTION AND ACTION DESCRIPTORS

Let $w_{s,k,p}$ denote the pre-intervention descriptor of target state s, task k, and pair p. It contains the observational families of Appendix A.3, symmetric pair transforms, within-panel relative features, and past-only temporal context. Let $q(a)$ contain only the candidate action's pre-execution semantics and provenance. The scorer estimates action-conditioned outcomes such as

$$\widehat{\Delta L} = f_{\text{benefit}}(w, q(a)), \tag{28}$$

$$\widehat{p}_{\text{elig}} = f_{\text{elig}}(w, q(a)). \tag{29}$$

This is consequence prediction for a candidate action, not a learned scalar structural "badness" objective.

The historical calibrator sequence is retained because it discovered the transfer boundary. On the original two-seed corpus, grouped-run and run/regime transfer were strong, while leave-seed transfer collapsed. A frozen curriculum trained on 540 pair states from 30 checkpoints and 22 runs obtained within-panel Spearman 0.5808 and pair NDCG 0.9284 on completely unseen runs, whereas the original two-seed leave-seed audit obtained Spearman only 0.0316. The later seven-seed invariant redesign recovered useful transfer, but all final-gate-passing autonomous policies were Next-only. These results motivate the mature split-specific evaluation below rather than a claim of seed-canonical witness coordinates.

A.8.3 AMAZON GROUPED, ROLE-SHIFT, AND TEMPORAL-FORWARD RESULTS

Amazon provides the positive amortization regime. Across grouped-run, leave-role, and temporal-forward evaluation, adding current-state witnesses to the same candidate action descriptor increases realized eligible value relative to action-only HistGB. The paired eligibility-gated gain

lifts are $+0.00897$ (95% CI $[0.00583, 0.01218]$), $+0.00776$ ($[0.00416, 0.01119]$), and $+0.03360$ ($[0.02866, 0.03861]$), respectively.

The temporal-forward regime is the cleanest operational result. Mean within-panel action-ranking Spearman rises from 0.431 to 0.919; the witness-conditioned policy acts on 83.6% of panels and realizes 0.03594 eligible primary gain, capturing 75.3% of the best eligible value present in that finite bank. The corresponding action-only policy acts on 20.5% of panels and captures only 4.9% of bank value. This establishes useful witness-conditioned selection within the compatible Amazon bank; the contrasting Kuai regime below identifies its transfer boundary.

A.8.4 KUAI LIFECYCLE, TASK-PAIR PRIOR, AND PERMUTATION CONTROLS

Kuai provides the transfer boundary. Rich witnesses carry global opportunity signal, but fine-grained within-panel action ordering transfers weakly across independent seeds. In the initial audit, HistGB global loss-gain Spearman is roughly 0.69–0.70 while within-panel Spearman is only about 0.08–0.14. More decisive controls compare the full witness against baselines that preserve increasingly much transferable structure. Full witnesses improve capture over a weak context-only control by $+0.1130$ with positive support in all eight seeds, but do not robustly beat a strong task–pair prior (difference -0.0079). A lifecycle-preserving within-panel witness permutation changes capture by only $+0.0159$ with an interval crossing zero. Fine-grained witness identity is therefore not established as a transferable pair-specific controller coordinate.

The final frozen-bank label-free policy remains a negative deployment replication. Mean held-out primary gains for HistGB are -6.29×10^{-5} (context), -5.18×10^{-5} (static), -1.15×10^{-5} (temporal), and -1.90×10^{-5} (static+temporal). No-op therefore remains better than the tested learned frozen-bank policies. This does not negate Fresh headroom; it locates the failure downstream, in transferred action availability and selection.

A.8.5 BANK RETENTION IS AN UPSTREAM COMPATIBILITY VARIABLE

A selector can choose only among actions that remain useful in its available bank. We therefore separate

$$G_{\mathrm{fresh}} \longrightarrow G^{\star}_{\mathcal{B}} \longrightarrow G_{\pi}, \tag{30}$$

where G_{fresh} is the protocol-defined current-state reference, $G^{\star}_{\mathcal{B}}$ is the best eligible realized value available in the frozen bank with no-op allowed, and G_{π} is the policy's mean observed eligibility-gated value. Ineligible executed actions contribute zero to this score; the gate is retrospective, not an executable rejection rule. Raw policy loss can differ. The original Amazon fold-policy results separately report the ungated mean (Appendix A.8.1).

Kuai makes the first bottleneck explicit. Across 48 held-out panels, $G_{\mathrm{fresh}} = 0.00123993$ but the best eligible frozen cross-seed bank contains only $G^{\star}_{\mathcal{B}} = 0.00030970$, a retention ratio of 0.24977. Fresh-minus-bank is positive in all eight seed means with 95% CI approximately $[0.0007815, 0.0010970]$; Fresh-versus-frozen pair-rank correlation is only 0.11, and the same best pair occurs in about 25% of panels. The learned policy then realizes $G_{\pi} \approx -0.00001899$.

Amazon supplies the complementary compatible-bank regime. On the exact 122 prespecified temporal-forward panels, $G^{\star}_{\mathcal{B}} = 0.0477095$ and $G_{\pi} = 0.0359400$, giving bank-to-policy capture 0.75331 with policy coverage 0.8361. A population-matched Fresh value is intentionally not identified on this exact 122-panel population, so no Amazon Fresh-to-bank ratio is imputed.

A.8.6 CALIBRATION, ABSTENTION, AND SELECTED-TAIL SECONDARY QUALITY

Candidate ranking, coverage, primary usefulness, and selected-tail secondary quality are distinct quantities. For the witness-conditioned HistGB policy, observed unsafe-among-acted rates are 10.3% under grouped-run evaluation, 10.8% under leave-role, and 2.94% under temporal forward. The corresponding one-sided 95% iid-panel Clopper–Pearson reference bounds are approximately 17.4%, 16.9%, and 7.4%. Temporal-forward has 3 violations among 102 acted panels (of 122), giving an exact one-sided reference bound 7.426%. Repeated states/panels can be dependent; this is not a cluster-aware coverage guarantee. Temporal-forward therefore satisfies the predeclared 5% budget as a point estimate, but does not statistically certify that budget at the stated confidence level.

Table 23: Amortized candidate-value prediction is regime-dependent. Panel A records the historical Amazon transfer boundary that motivated invariant/relative redesign. Panel B reports the mature confirmatory Amazon HistGB comparison using the same candidate bank and action descriptors, with or without current-state witnesses. Panel C reports the Kuai leave-one-seed-out boundary: full witnesses beat a weak context control, but not a strong task–pair prior or a lifecycle-preserving within-panel permutation, and the final transferred frozen-bank policy remains nonpositive. These are action-selection results conditional on the available bank, not evidence that the bank itself preserves Fresh value.

A. Historical Amazon transfer boundary

Evaluation / corpus	Within-panel ρ	Pair NDCG	Boundary exposed
Original grouped-run (87 checkpoints)	0.6407	0.9454	strong in-corpus calibration
Temporal extension (540 pair states)	0.5662	0.9346	temporal transfer retained
Completely unseen runs (540 pair states)	0.5808	0.9284	run transfer retained
Original two-seed leave-seed	0.0316	0.7384	seed transfer collapses
Five external seeds, legacy view	≈ 0.099	–	failure persists externally

B. Mature Amazon confirmatory scorer

Split	Scorer	ρ_{within}	Elig. AUC	Coverage	Gated gain	Bank capture
Grouped-run	Action only	0.308	0.727	24.6%	0.00798	16.4%
	Witness+action	**0.699**	**0.814**	47.5%	**0.01695**	**34.9%**
Leave-role	Action only	0.470	0.783	45.9%	0.01369	28.2%
	Witness+action	**0.728**	**0.818**	60.7%	**0.02145**	**44.2%**
Temporal fwd.	Action only	0.431	0.873	20.5%	0.00234	4.9%
	Witness+action	**0.919**	**0.983**	83.6%	**0.03594**	**75.3%**

Gated gain is U_ϵ, not ungated executed loss gain. Paired witness-minus-action lifts: grouped $+0.00897$ (95% CI $[0.00583, 0.01218]$), leave-role $+0.00776$ ($[0.00416, 0.01119]$), temporal-forward $+0.03360$ ($[0.02866, 0.03861]$). Development leave-seed is kept separate from these confirmatory splits.

C. Kuai transfer boundary

Comparison / policy	Estimate	Support	Reading
Full witness − context-only capture	$+0.1130$	8/8 seeds	improves over weak context
Full witness − task–pair prior capture	-0.0079	–	no advantage over strong prior
Full witness − within-panel permutation	$+0.0159$	CI crosses 0	fine identity not established
Initial global vs within-panel ranking	$\rho \approx 0.69$–0.70 vs 0.08–0.14	LOSO audit	global signal $\neq$ local ordering
Context HistGB policy	-6.29×10^{-5}	12.5% coverage	nonpositive
Static HistGB policy	-5.18×10^{-5}	37.5% coverage	nonpositive
Temporal HistGB policy	-1.15×10^{-5}	12.5% coverage	nonpositive
Static+temporal HistGB policy	-1.90×10^{-5}	37.5% coverage	nonpositive

Production gives an independent reason not to collapse eligibility into usefulness. At $\epsilon = 0.002$, Foreign replay is 98.0% eligible on the secondary metric while its eligibility-gated primary gain is -0.000052. At $\epsilon = 0.010$, Same replay is 97.2% eligible while its gated primary gain is -0.000364. Fresh is 100% eligible throughout the reported tolerance curve. Eligibility is therefore a constraint on candidate value, not evidence that primary value is positive.

A.8.7 Regret decomposition

For the final bounded Amazon HistGB scorer, residual safe regret is not dominated by unsafe action selection. Across grouped-run, leave-role, and temporal-forward evaluation, approximately 50–57% of remaining regret comes from abstaining when a safe positive oracle action exists, about 33–38% from selecting a safe but suboptimal action, and only about 9–11% from unsafe selections. In the 122-panel temporal-forward split, only 18 panels fall into the "abstain-despite-oracle" category, yet they account for 53.2% of residual safe regret. Conservative abstention is therefore useful but expensive; the next selection boundary is not reducible to making the safety classifier more aggressive.

Table 24: Deployment decomposes into opportunity availability, finite-bank selection, abstention, and secondary-quality eligibility. Panel A keeps the Fresh → bank → policy stages separate and intentionally leaves Amazon Fresh blank on the exact 122-panel temporal population because it was not population-matched there. Panel B distinguishes observed selected-tail risk from uncertainty-aware certification. Panel C decomposes residual safe regret for the final Amazon witness-conditioned HistGB policy. Panel D shows the full Production tolerance curve, where high secondary eligibility can coexist with nonpositive primary value. The final strip records a historical negative result: frozen dynamic routing was nearly action-invariant across matched histories.

A. Opportunity, compatible bank, and realized policy

Setting	G_{fresh}	$G^{\star}_{\mathcal{B}}$	G_{π}	Availability / selection
Amazon temporal, exact 122 panels	–†	0.0477095	0.0359400	capture = 0.75331; coverage = 0.8361
Kuai cross-seed, 48 panels	0.00123993	0.00030970	−0.00001899	bank/Fresh = 0.24977

Values in A use observed eligibility-gated primary gain, not ungated executed-policy gain. B uses iid-panel reference bounds, not cluster-aware certification. †Population-matched Amazon Fresh is intentionally unidentified on this exact 122-panel population and is not imputed. For Kuai, Fresh-minus-bank is positive in 8/8 seed means with 95% CI approximately $[0.0007815, 0.0010970]$; Fresh-versus-frozen pair-rank correlation is 0.11, and exact best-pair agreement is about 25%.

B. Amazon selected-tail risk: point estimate vs certification

Split	Coverage	Unsafe / acted	One-sided 95% upper	5% budget
Grouped-run	47.5%	10.3%	17.4%	not met
Leave-role	60.7%	10.8%	16.9%	not met
Temporal-forward	83.6%	2.94%	7.4%	point met; not certified

C. Amazon residual safe-regret decomposition

Split	Abstain despite oracle	Safe but suboptimal	Unsafe selection
Grouped-run	56.9%	32.5%	10.6%
Leave-role	50.2%	38.4%	11.4%
Temporal-forward	53.2%	37.9%	8.9%

Temporal-forward contains 18/122 abstain-despite-oracle panels, 56/122 safe-but-suboptimal panels, and 3/122 unsafe selections; the first category alone contributes 53.2% of residual safe regret.

D. Production secondary eligibility does not certify primary usefulness

ϵ	Fresh elig.	Same elig.	Foreign elig.	Same gated gain	Foreign gated gain
0	100%	41.3%	44.2%	+0.001207	+0.000110
.001	100%	47.2%	95.4%	+0.001167	−0.000013
.002	100%	69.8%	98.0%	+0.000794	−0.000052
.005	100%	88.1%	100%	+0.000265	−0.000094
.010	100%	97.2%	100%	−0.000364	−0.000094
.020	100%	100%	100%	−0.000664	−0.000094

Historical frozen dynamic-routing boundary

Matched comparison	Same-action fraction	Support / boundary
No-op branch vs selected branch	≈ 0.979	630 matched windows
Selected history vs matched selected history	1.000	630 matched windows

This historical controller behaved mainly as checkpoint-context act/no-op gating. It does not establish high-frequency lifecycle-responsive pair switching or a solved closed-loop refresh policy.

A.8.8 Historical frozen dynamic-routing boundary

An earlier frozen dynamic-routing implementation was intended to test whether the controller would respond to high-frequency lifecycle changes. A matched-history audit showed that interpretation to be too strong. The same-action fraction is approximately 0.979 between no-op and selected branches and 1.000 between matched selected continuation histories over 630 windows each. The controller therefore behaved mainly as checkpoint-context act-versus-no-op gating, not as evidence of high-frequency lifecycle-responsive pair switching. Static bank compatibility, static selection, and refresh cadence remain separate control problems.

A.8.9 MEASURED FITTING AND TRANSPORT COMPUTATION

Fresh and the target-gradient mask have different computational scopes. On three profiled Kuai target states, the complete 12-action Fresh search took $28{,}896$–$53{,}520$ backward operations and 15.9–16.6 minutes on a V100. The corresponding 12 GradFix derivations used 240 backward evaluations and no target optimization steps; those counts exclude replay EVAL. In Amazon, isolated sign derivation averaged 16.43 ms per action. Historical Fresh-worker clocks include additional stages and are not matched to that isolated derivation: 378 archived Amazon Fresh jobs record a median 22.94 s per complete historical worker, including checkpoint loading, probing, fitting, selection, and evaluation. The Amazon port records stage-specific median CUDA allocated peaks of about 333 MiB for cold/probe, 139 MiB for sign derivation, and 211 MiB for Raw/GradFix evaluation; median CUDA reserved peak is about 424 MiB, and median process peak RSS is 3.008 GiB. These are neither incremental mask memory nor a matched Fresh-versus-port memory comparison. Across 63 target-state jobs, median complete worker time through result upload is 17.88 s. No continuous GPU-utilization trace was recorded.

The learned controller also needs online witness extraction, bank scoring, and selected-action execution. These stages have not been jointly profiled against a Fresh-search baseline. The measurements establish that repeated Fresh fitting is costly under the profiled protocol and that masking can be cheap to derive; they do not establish an end-to-end latency or memory advantage for a prospective deployment policy.

Deployment synthesis. A.8 separates four quantities that should not be collapsed. Fresh measures current-state opportunity. Bank compatibility determines how much of that opportunity is even available to a deployable selector. Selection and abstention determine how much available value the policy realizes. Secondary-quality eligibility constrains which actions are acceptable but does not certify positive primary value. Amazon shows that state conditioning can strongly amortize search when the bank remains compatible; Kuai shows that bank transfer can dominate the loss before selection; Production shows that high eligibility can coexist with null or negative primary value. The current evidence does not establish a globally reusable controller or a closed-loop refresh policy.

A.9 STRUCTURED ACTION-SPACE CLOSURE

The reference experiments deliberately use a small bounded action family, but the scientific object is richer than a pair index plus a coefficient vector. A fitted action also has support, scope, dose, composition, and task-allocation semantics, and its numerical coefficients are meaningful only in the representation frame in which they were fitted. Appendix A.5.3 already separates pair support from reach and coefficient capacity. This final Level-II block closes the remaining questions: whether individually useful actions compose additively, whether a fixed intervention budget should be shared across tasks, and whether the fitted coefficient bank admits a compact reusable representation.

The closure is deliberately asymmetric. Composition and the codec provide bounded Amazon evidence; the cross-task allocation panel is exploratory because it contains one seed only. None of these studies identifies a global raw-coordinate operator manifold.

A.9.1 COMPOSITION AND REDUNDANCY

Combining individually useful local actions does not yield an additive utility law. The final composition audit reconciles every required singleton dependency before forming the joint-action interaction. At fixed total nominal dose $\sum_m \alpha_m = 1$, both tested constructions remain behaviorally beneficial and secondary-eligible on the five-state panel, but their joint benefit is below the sum of their constituent benefits. Best-safe disjoint matching has mean primary interaction -0.002914 and NDCG interaction -0.003143; all-safe consensus has primary interaction -0.004738.

The expanded-dose rows sharpen the boundary rather than reversing it. Giving each component full dose increases raw loss, NDCG, and HR gains, but also increases total nominal dose from 1 to 2 for matching and from 1 to 5 for consensus. Dose-adjusted utility falls from 0.4669 to 0.4297 for matching and from 0.4642 to 0.3050 for consensus, while the interaction terms become more negative. The composed realized displacement is also about 13.3% below the sum of constituent

displacements across these regimes. Thus negative interaction here means *redundancy/subadditivity*, not that the composed intervention itself is harmful. The data support sparse bounded composition as an available action primitive; they do not support a constructive-additivity assumption.

A.9.2 Cross-task allocation

A seed-13 Amazon panel asks a different question: with the same total nominal budget, should one shared source-task operator be applied across tasks, or should the budget be allocated through task-specific operators? Across nine matched best-primary states, task-specific fixed-budget allocation uses slightly *less* realized dose on average (0.1215 versus 0.1352) while improving mean task loss gain (0.01904 versus 0.00955), mean task NDCG (0.00705 versus 0.00536), and minimum-task NDCG ($+0.00205$ versus -0.00160). The paired task-specific-minus-shared contrasts are $+0.009490$, $+0.001692$, and $+0.003650$, respectively, and all nine matched mean-task-loss differences are positive.

An independent-per-task arm reaches still larger absolute value, but uses three units of nominal dose and is therefore an unmatched higher-budget reference, not evidence for the equal-budget comparison. The equal-budget result itself remains exploratory: it is one seed and nine states, so it establishes a concrete counterexample to task-agnostic allocation, not population-level superiority of one allocation rule.

A.9.3 Compact frame-local operator representation

The fitted Amazon coefficient bank contains compact structure within a common representation frame. The strongest tested compact representation, the frame-local M2-r32 codec, has mean coefficient relative error 7.06% on the target-eligible population while preserving average behavior closely: reconstructed replay minus the original target action changes primary loss gain by -5.5×10^{-5} and NDCG by approximately -2×10^{-6}.

Average fidelity does not certify selected-tail eligibility. About 1.61% of target-eligible actions cross the secondary-quality boundary after reconstruction. An exploratory $+0.003$ target-side margin removes the observed crossings while retaining about 85.4% of the originally target-eligible population, but this is a post-hoc margin diagnostic rather than a calibrated deployment guarantee. The codec therefore supports a bounded statement: optimized interventions admit a compact *frame-local* low-rank representation whose average behavior is stable, while small boundary crossings remain.

A.9.4 Frame-local structure and cross-model transport

The codec result is intentionally frame-local. It does not identify coordinates that can be transported unchanged between independently trained models. The independent transport evidence points in the opposite direction: raw Foreign replay is weak, family-preserving permutation alignment closes only a minority of the Amazon gap (Appendix A.6.5), and the tested functional or representation proximity variables do not rescue observed Foreign replay (Appendix A.6.6). On Kuai, the Same and Foreign similarity supports do not even overlap sufficiently for the nearest-Foreign audit to identify a residual lineage effect.

Composition redundancy, behavioral near-ties, and frame-local compression show that distinct coefficient vectors can have similar measured behavior. Extending this relation across models requires an explicit identification map and a held-out transport test. The present codec and replay experiments separate those within-frame and cross-model questions.

Action-space synthesis. Across A.5–A.9, the empirical object is therefore an action-conditioned, state-dependent response rather than a privileged coordinate vector. Support and reach matter; dose changes both benefit and secondary risk; useful actions are not additive under naive composition; task allocation can matter even at fixed nominal budget; and within one frame the coefficient bank is compressible without being universally transportable. These closures motivate the finite bounded candidate action formulation used by the deployment experiments rather than a combinatorial or globally generated action space.

Table 25: Structured action-space closure. Panel A separates absolute composition value from interaction relative to constituent singleton actions: all reported compositions remain beneficial, but interaction is predominantly redundant/subadditive and becomes more negative when total dose is expanded. Panel B is the exploratory seed-13 task-allocation panel; only the first two rows are equal-budget, while independent-per-task is an explicitly higher-budget reference. Panel C shows that the strongest tested compact codec preserves average behavior within a common representation frame but does not eliminate secondary-quality boundary crossings. These results delimit bounded action semantics; they do not identify a globally reusable raw-coordinate operator manifold.

A. Multi-action composition: positive absolute value, subadditive interaction

Regime / construction	Total α	Utility / dose	ΔL	ΔNDCG	ΔHR	Mean disp.
Fixed-budget matching	1	0.4669	0.0678	0.0382	0.0607	0.1453
Fixed-budget consensus	1	0.4642	0.0661	0.0388	0.0610	0.1426
Full-dose matching	2	0.4297	0.1286	0.0531	0.0908	0.2993
Full-dose consensus	5	0.3050	0.2262	0.0713	0.1308	0.7418

Interaction relative to singleton sum	Loss	NDCG	HR
Fixed-budget matching	-0.00291	-0.00314	-0.00262
Fixed-budget consensus	-0.00474	-0.00532	-0.00164
Full-dose matching	-0.01390	-0.01918	-0.02230
Full-dose consensus	-0.13091	-0.11312	-0.15607

Five-state composition panel. The expanded audit reconciles 70/70 singleton dependencies across fixed-budget and full-dose constructions. Every composed gain above remains positive; negative interaction means the joint gain is below the sum of its constituent singleton gains. Composed realized displacement is approximately 13.3% below the constituent-displacement sum across the reported strategy/dose regimes.

B. Cross-task allocation (Amazon seed 13; nine matched states)

Allocation	Total α	Realized dose	Mean ΔL	Mean ΔNDCG	Min-task ΔNDCG	Eligible
Shared single-task operator	1	0.1352	0.00955	0.00536	-0.00160	7/9
Task-specific fixed budget	1	0.1215	0.01904	0.00705	$+0.00205$	9/9
Independent per task†	3	0.3781	0.05517	0.01748	$+0.00618$	9/9

Equal-budget task-specific minus shared: mean-task loss $+0.009490$, mean-task NDCG $+0.001692$, minimum-task NDCG $+0.003650$, and realized-dose difference -0.013706; all 9/9 mean-task-loss differences favor task-specific allocation. †The independent-per-task arm uses three units of nominal dose and is not a matched-budget comparator. This entire panel is exploratory single-seed evidence.

C. Compact frame-local operator representation

Codec	Coeff. rel. error	Replay$-$target ΔL	Replay$-$target ΔNDCG	Eligible$\rightarrow$ineligible	Margin diagnostic
M2-r32	7.06%	-0.000055	-0.000002	1.61%	$+0.003$: 0 observed crossings; 85.4% retained

Codec statistics are evaluated on the target-eligible population in a common Amazon representation frame. The $+0.003$ margin is exploratory and is not an uncertainty-aware safety certificate. Leave-seed raw-coordinate transport is unavailable in this codec analysis; the result supports frame-local compression only.

B RECOVERY, TRAINING TRAJECTORIES, AND OBJECTIVE-DEPENDENT VALUE

The following experiments extend the original opportunity, compatibility, and recovery study. They ask whether stronger current-state fitting changes the historical failure, whether the base trajectory or endpoint coefficients account for later differences, and which measured gains survive attribution, calibration, and quality controls. The source campaign numbers identify provenance; they are not independent replications of one unchanged procedure. Earlier measurements remain valid for their original comparisons even when a later control narrows their interpretation.

B.1 EXPERIMENTAL CONTRACTS AND EVIDENCE LOCATIONS

References and units. An absolute loss gain is reference loss minus candidate loss; a quality change is candidate quality minus reference quality. In the public Amazon experiments, the code identifier `trx` denotes the *Next* task. Identity on the same frozen base, a held incumbent, the complete fixed-incumbent continuation, and the complete ordinary-MTL continuation are different references. A positive increment beyond an incumbent must not be added again to a gain already measured relative to identity. Amazon loss uses the native masked item BCE, averaged within a list and then equally across lists; ranking uses NDCG@10 and HR@10. Kuai uses its native binary-example loss and whole-partition AUC. Twenty Amazon FIT lists contain different information from twenty Kuai binary examples.

Comparator key.

Term	What is evaluated or held fixed
Identity	The stated base with no applied operator; it is not another trained base.
Held incumbent	An earlier fitted operator, reused without new fitting at the target.
Ordinary / fixed	Base continuation without an operator / with the incumbent held fixed.
Periodic / interleaved	Live base continuation with the specified coefficient refresh schedule.
Deferred	New coefficient fitting at a saved final base, without changing that base.
Shadow	Sequential fitting across saved states; new coefficients do not alter the saved base path.
Common anchor	Shared initialization, constraint center and penalty center; historical information remains.

Independent and repeated observations. The model seed is the unit for the reported seed intervals. States, task/pair units, candidate endpoints, calibration splits, and repeated evaluations are dependent observations within seeds. Unless a table states otherwise, the later campaign intervals are pointwise 95% percentile bootstrap intervals from 5,000 resamples of seed means, conditional on shared data and the exploratory research sequence. They are not simultaneous confidence intervals over the campaign search. Selection locked before a campaign's new EVAL readout does not remove adaptive use of earlier results. Three-seed controls and five-seed fresh results retain their own sample sizes.

Candidate access and fallback. Original amortized bank selection (Appendix A.8) predicts the value of previously available actions. Later refit-and-guard experiments incur new target fitting and evaluate different candidate banks. Their costs and information access are separate. Where a later rule selects once on TUNE, it checks that winner once on GUARD; rejection does not trigger a retry. Reported policy means include the actual fallback. Fallback to an incumbent can inherit that incumbent's quality failure, and fallback to a complete fixed checkpoint restores both its base and operator. It is generally different from zeroing an operator on the rejected candidate's base.

The original 122-panel Amazon temporal-forward results remain in Appendix A.8.3 and Table 24. Their reported eligible bank capture is not an uncensored executed-policy loss mean. The original fold-policy records report that distinct raw mean (Appendix A.8.1); newer all-seed policy readouts must not be substituted into the old population.

Historical information and constraints. The common-anchor experiments retain an earlier fitted vector as initialization, radius center, and penalty center. Equal new update counts therefore do

not equalize historical fitting costs or accessible intervention sets. The inherited box-then-radial retraction is not asserted to be the exact Euclidean projection onto the intersection. During frozen coefficient fitting the base and routing are held fixed; during live alternating training the operator has a forward effect on base updates. A statement about freezing during fitting does not mean the live base trajectory is frozen.

Table 37 links these question-led sections to the original source panels.

B.2 How much does local fitting recover, and does added complexity help?

Current-state recovery. Campaign 2 studies seven Amazon seeds, 28 refresh states, 28 mechanism states, and 14 archived trajectories. Its completed CORE study compares current-state refitting with the held incumbent after the historical recipe failed. Warm fitting with FIT20 and a selectable 20-update cap produces mean incremental loss gain $+0.015316$ with interval $[0.010699, 0.020365]$, positive for all seven seed means. Mean NDCG increases by 0.002699, but the union of the recorded quality failures contains 4 of 28 states. Thus useful average primary and secondary responses coexist with individual-state violations. These are selected frozen-state results, with LR and intermediate-endpoint access; they are not the outcomes of an unselected single update.

Complexity and the comparator. Source-gradient masks improve some replay comparisons, but shuffled controls remain competitive. TIES-style merges, learned nonnegative mixtures, task-sign/GradFix-style masks, consensus rules, and ranking projections do not provide a decisive, consistent advantage over strong simple target fits in this operator-space experiment. The merge study has two donors, many one-donor groups, and 64 coefficients; it is not a benchmark of full-network merging. An exploratory donor-magnitude construction beats the weaker target-vote restriction by 0.004019 raw and 0.003357 aligned, yet trails searched target Adam by 0.013726 and 0.014983. This observation motivated the matched control in Appendix B.3.1; it does not establish donor specificity.

Data and update budgets. Increasing FIT from 20 to 100 complete 20-item lists changes the key means little, while expanding the selectable update cap from 20 to 80 can worsen ranking. The paths actually execute 80 updates with shared prefixes: a 20-update selection cap is not a claim of only 20 incurred updates. SELECT contains 101–102 additional lists. Fixed target-gradient endpoints illustrate the objective conflict, with mean NDCG changes -0.024351 raw and -0.031975 aligned in the corresponding FIT20 comparison. Sequential carry reaches $+0.008026$ incremental loss gain at horizon 3, but its movement and cumulative update budgets differ from hold/reset controls. The matched schedules below replace a strong causal tracking interpretation with a more limited one.

B.3 Is fitted-operator motion evidence of representation motion?

Repeatability of response versus coefficients. Campaign 3 uses 42 frozen Amazon states from seven seeds, three deterministic FIT20 resamples, three LR levels, and two radii. Within-state resampling variation exceeds between-state same-resample variation on all seven seed averages at all six LR/radius settings, for coefficient distance, per-list loss RMS, NDCG RMS, and changed-ranking-list frequency. Functional comparisons evaluate both endpoints on the same final-state EVAL population. At LR 0.003 and radius 1, the within/between values are $0.39235/0.19462$ for coefficient distance, $0.000593/0.000352$ for loss RMS, $0.01993/0.01313$ for NDCG RMS, and $2.209\%/1.138\%$ for changed-list frequency. These resamples can overlap; this comparison is not a formal variance-component decomposition.

All 756 own-state refits improve incumbent-relative loss. At the same LR/radius setting, the mean gain is approximately 0.012789, with 13 quality violations among its 126 repeated endpoints. These are dependent outcomes, not 756 independent experiments. The joint observation is that the local loss opportunity repeats while the fitted coefficient representation varies. Coefficient distance therefore cannot, by itself, quantify behavioral drift caused by model evolution, and a finite Fresh fit is not a unique optimum.

Matching fitting schedules. Fourteen saved trajectories permit moving, frozen-final, and reordered fitting schedules that end on the same base and share the original anchor, FIT rows, total

radius, and total updates. Adam-carry moving-minus-frozen mean gains range from -0.000212793 to $+0.000082707$; across all methods the range is approximately $[-0.000563547, +0.000233814]$. Some intervals exclude zero, so this is evidence of relatively small schedule differences, not equivalence. Frozen-final fitting uses future-state access and is a diagnostic comparator. The historical coupled base training is not rerun by this saved-state experiment. It narrows the proposed large generic moving-state explanation without erasing the historical endpoint gap.

Local signal and projection audit. At each of the forced doses 0.01 and 0.05, descent improves primary loss in all 42 states and ascent worsens it in all 42. At dose 0.05, the respective mean gains are approximately $+0.000924$ and -0.000921. No-op selection cannot explain these forced responses, but quality still requires separate evaluation. All 42 audited projection records have rank 64, with condition numbers 4.7255–5.6879 and pseudoinverse norms 1.9655–2.4738. This audit does not certify unexamined historical projections or imply equal realized displacements across parameterizations.

B.3.1 Matched donor controls and candidate quality

The donor control shares exact support, target-vote signs, and full-dose ℓ_2 norm across candidate magnitude families. Across 2,016 matched groups, reconstructed norm differences are at numerical precision. At FIT20, donor-minus-constant loss gains are -0.000591954 raw (seven negative seed differences) and -0.000860468 aligned (six negative). Fixed donor-versus- permutation contrasts give no consistent advantage. These controls remove the robust donor-magnitude interpretation of the earlier motivating result. They do not establish that all donor information is useless: support, candidate availability, and norm budget remain source-conditioned.

Target-gradient magnitudes under the common vote signs improve raw fixed loss more than donor magnitudes (0.020071 versus 0.013376), while worsening mean NDCG (-0.007768 versus $+0.000134$). The sign rule means this action is not the unmodified negative mean gradient. With dose/action selection and separate guarding, target magnitudes yield loss/NDCG gains $0.013626/0.006887$, active on 18/28 states with no observed violations; donor magnitudes yield $0.010212/0.003587$, active on 19/28 with none. The paired operational loss-difference interval spans zero. Fixed-action averages and selected-bank outcomes are different estimands.

The full-SELECT target bank is active on 28/28, gains 0.020973, and has two quality violations, with worst NDCG change -0.011441. The split-guard procedure reduces both gain and coverage to the 18/28 result above. This compares complete selection rules, rather than isolating one validation check. Its unchanged fallback must retain inherited failures where present. The Campaign 3 target-Adam comparator is only LR 0.003, 20 updates, and one endpoint; its mean displacement is about 0.47, versus 1.15–1.17 for the matched magnitude families. It is neither the stronger searched Campaign 2 comparator nor a member of the norm-matched contrast.

B.4 Does the fitting remedy survive stronger and cross-domain controls?

A conditional Amazon optimizer result. Campaign 3 calibrates each SGD rate using state-1 FIT to match Adam's first unprojected proposal norm. This matches neither direction nor subsequent path length. In all 12 moving-schedule settings, SGD improves the primary gain over Adam carry in every seed mean; differences range from 0.000500 to 0.002882. At radius 1 and 60 updates, separate TUNE/GUARD yields gains 0.006641 versus 0.002863, with 8/14 versus 4/14 changed operators. The paired contrast is 0.003779 with interval $[0.000337, 0.006952]$. Moment reset and task-sign filtering do not produce comparable consistent remedies. Proposed, filtered, and post-constraint steps remain distinct; first-order sign counts alone do not identify the magnitude of harmful motion.

Campaign 4 expands Adam search over 28 frozen states: two source roles and two horizons within seven seeds. Both optimizers share the coefficient box, original-anchor radii 0.5 and 1, penalty, FIT data, and LR/snapshot access. Adam LR levels $0.0003, 0.003, 0.03$ calibrate the SGD rates. Endpoints at 5, 20, and 80 updates create max-20 and max-80 banks. All 126 seed-by-setting averages favor SGD across the 18 fixed settings; they are 126 dependent contrasts from seven seeds. Table 26 shows that the selected advantage persists, while guarded superiority at radius 1 remains uncertain. An identity-floor failure under guarded SGD is inherited from the fallback incumbent, illustrating why active and fallback outcomes must both be retained.

Table 26: Stronger Amazon fitting comparator, Campaign 4. Gains are beyond the held incumbent, averaged within each of seven seeds over four frozen states. The selected radius-1/max-80 comparisons use matched search access; the interval is for the paired SGD-minus-Adam contrast. Full SELECT and split guard are different complete procedures. Sources: V4_04–09, V4_15.

Procedure	SGD gain	Adam gain	Paired difference [95% interval]
Radius 1, full SELECT	0.013169	0.010835	0.002334 [0.000795, 0.003815]
Radius 1, split guard	0.007372	0.005636	0.001736 [−0.000491, 0.003255]

At radius 1/max-80, full SELECT changes 27/28 SGD and 26/28 Adam operators; split guard changes 16/28 and 15/28. SGD split guard has zero incumbent-floor violations but one inherited identity-floor violation under fallback. Adam retains two incumbent-floor and one identity-floor violations under both procedures. Radius-0.5/max-80 guarded SGD-minus-Adam is 0.002547 [0.001565, 0.003658], with 7/7 positive seed contrasts; radius choice remains part of the comparison.

Which optimizer component is implicated? On the diagnostic horizon-3/radius-1/LR-level-0.003 slice, removing diagonal scaling while retaining exponential momentum tracks SGD; removing first momentum while retaining diagonal scaling tracks Adam. At 20 updates, the mean gains for Adam, Adam without first momentum, momentum without scaling, and SGD are respectively 0.00558162, 0.00555685, 0.00792012, and 0.00790996. At 80 updates they are 0.01240870, 0.01245076, 0.01386053, and 0.01384387. This pattern implicates coordinate scaling in the tested constrained problem, while subsequent displacements, regularization, and boundary contacts remain unmatched. It neither uniquely identifies the historical failure mechanism nor proposes a new optimizer.

How much fitting is necessary? At radius 1, guarded SGD gain is 0.007361 for max-20 and 0.007372 for max-80; the increment is 0.00001098 with interval [−0.00001909, 0.00005031]. The max-80 bank can still choose an early endpoint, so this is saturation within the search ladder, not proof of 20-step optimality. A one-gradient ray gives guarded gain 0.003509 with 15/28 active, compared with iterative SGD's 0.007372 and 16/28. The ray contracts at its first box boundary and has a smaller candidate search; it does not match accessible paths or compute. Most gain is available early, but a useful initial direction does not account for all the measured fitted response.

B.4.1 KUAI LIMITS THE REMEDY, WHILE PRESERVING THE ORIGINAL OPPORTUNITY QUESTION

Campaign 4 repeats frozen fitting on eight Kuai seeds, two tasks, six expert pairs, and horizons 1 and 3 after the epoch-11 anchor (192 state/task/pair units). FIT contains 20 balanced binary examples; SELECT has 2,025 examples, split into 1,215 TUNE and 810 GUARD, and EVAL has 11,835. All four primary SGD-minus-Adam seed intervals span zero (Table 27). Small engagement gains occur, while long-view incremental gains remain inconclusive. No full live Amazon pipeline is replicated by this frozen comparison.

Table 27: Kuai frozen-refit boundary, Campaign 4. Radius 1/max-80 split guard, with eight seeds and six pairs per task/horizon. Gains are incremental beyond the held incumbent. Every primary paired SGD-minus-Adam interval spans zero. All eight rows of the two-optimizer comparison have zero observed AUC-floor violations, although mean AUC changes are small and negative. Sources: V4_30–40.

Task / horizon	Adam gain	SGD gain	Adam active	SGD active
Engagement / 1	0.00005548	0.00005195	29/48	16/48
Engagement / 3	0.00006831	0.00007541	24/48	22/48
Long view / 1	−0.00000638	−0.00001314	5/48	7/48
Long view / 3	−0.00000525	0.00000394	6/48	8/48

Campaign 5 extends Kuai to nested FIT20/100/500 support and prior weighting. All 24 raw 80-step SGD-minus-Adam means are negative; guarded intervals have mixed signs. Larger FIT support does not consistently rescue selected performance. For FIT500 long view, weighting improves the

raw SGD comparison by 0.000828 and 0.000928 at horizons 1 and 3, but the weighted absolute gains remain negative. Weighting also changes the SGD norm calibration, so it does not isolate prior mismatch. All 6,140 FIT-pool labels are accessed for support/prior construction; a small fitting subset is not the total label-access count.

These outcomes distinguish optimization from generalization. Campaign 4's radius-1/LR-level-0.003 SGD fits improve stored pre-update FIT loss by approximately 0.023280 for engagement and 0.033021 for long view, while 80-step EVAL gains are -0.000686 and -0.001645. The FIT trace is one update short of the EVAL endpoint and is not an exact generalization-gap estimator. Lower projected task-ascent frequency on Kuai does not repair this boundary. The prior incumbent still has useful identity-relative loss gains, and passing an AUC tolerance does not imply an AUC increase. The result restricts the proposed optimizer remedy; it does not negate the earlier multi-domain opportunity measurements.

B.5 Which training trajectory supports the stronger endpoint?

B.5.1 Useful live correction and its finite-horizon persistence

Campaign 4 continues each of seven Amazon source checkpoints for one native epoch under ordinary MTL, fixed incumbent, interleaved Adam, interleaved SGD, periodic SGD, and target-only specialization. Every arm processes 767,453 examples in the same recorded order over 93 base updates. The three adaptive arms each perform 60 coefficient updates, 20 per third of the epoch; the final milestone is predeclared. Native base Adam uses the restored state, LR 10^{-5}, weight decay 10^{-4}, and epsilon 10^{-7}. Base clipping is disabled (`grad_clip=None`); coefficient clipping is 5. The base updates with the current operator held fixed, and coefficient steps freeze the base. Projection coordinates remain fixed.

All three adaptive procedures improve final Next loss over fixed continuation in all seven seeds (Table 28). Periodic exceeds interleaved SGD by 0.000566, interval $[0.000165, 0.001025]$, with six positive paired seeds. The source-rate selection can differ across optimizers, including seed 41; the Adam–SGD comparison is between configured procedures. Periodic evaluation follows a refresh immediately, whereas interleaved coefficients have affected intervening model updates. Freshness, timing, and subsequent base response are not separately identified. Target-only specialization does not supply a successful simple remedy under its tested one-epoch recipe.

Table 28: Live Amazon benefit, Campaign 4. All means compare the final complete adaptive system with the time-matched fixed-incumbent system. Each row has seven positive loss contrasts. The final column counts seeds with any recorded all-head quality-floor failure; it is separate from the later admission decision. Source: V4_17–22.

Procedure	Loss gain	95% seed interval	ΔNDCG	Failures
Interleaved Adam	0.004428	[0.002368, 0.006615]	0.000257	2/7
Interleaved SGD	0.005973	[0.003981, 0.007877]	0.001089	2/7
Periodic SGD	0.006539	[0.004318, 0.008824]	0.001353	2/7

The initial locked live guard admits only 1/7 seeds per adaptive arm. It requires loss non-worsening versus fixed and all three quality floors against both ordinary and fixed; rejection returns the complete fixed checkpoint. Four rejected cases per arm retrospectively pass EVAL and two fail, so hindsight does not justify simply loosening the guard. The combined system can improve while its own unassisted base worsens: interleaved SGD's 0.005973 gain equals a -0.009404 base-identity component plus a 0.015377 operator-benefit component. This is exact readout accounting, not causal mediation or evidence of absorption into the base.

Campaign 5 trains the base for a further epoch while either retaining the operator or continuing refresh. Periodic no-refresh retains loss advantage 0.005536, interval $[0.003683, 0.007306]$, positive in 7/7 seeds: 85.47% of the first-epoch mean advantage remains. Periodic refresh adds 0.000554 $[0.000125, 0.001055]$; interleaved refresh adds 0.000457, also positive in 7/7. The operator remains applied. The no-refresh periodic readout decomposes into -0.008914 for base identity and $+0.014450$ for operator benefit. Persistence is a finite-horizon property of the complete system; guarded differences between branches also include admission changes.

B.5.2 DEFERRED FITTING AND CROSSED ENDPOINT READOUTS

Campaign 6 introduces no new base training. It fits coefficients at saved fixed-incumbent endpoints and compares them with the two-epoch periodic trajectory. The primary 120-update comparison matches base exposure and the total coefficient-update count. It retains FIT20, the inherited SGD rate, original-anchor radius 1 and penalty 0.01, coefficient box $[-0.5, 0.5]$, clipping 5, and float32 arithmetic. Gradient acquisition states, base paths, elapsed time, and complete historical compute are not matched. Deferred120 beats periodic120 by 0.008397 with interval $[0.005983, 0.010873]$, positive in 7/7. The 60-update comparison gives 0.007790 $[0.005325, 0.010039]$, also 7/7.

Table 29: Within-seed crossed endpoint readouts, Campaign 6. Entries are mean Next losses over seven seeds; lower is better. The two coefficient vectors share the pair and parameterization. Changing the base has a much larger measured effect than exchanging these endpoint vectors on a common base. Sources: V6_07, V6_33, V6_36.

Trained base	Deferred120 coefficients	Periodic120 coefficients
Fixed-incumbent base	0.322687	0.322654
Periodic120 base	0.331141	0.331084

Periodic-coefficient advantage on the fixed base: 0.000033148 $[-0.000027194, 0.000128572]$; on the periodic base: 0.000056569 $[-0.000004400, 0.000150376]$. Fixed-base advantages are 0.008453435 with deferred coefficients and 0.008430013 with periodic coefficients. Source contrasts are evaluated before table rounding.

The crossed readouts (Table 29) associate most of the gap with the base path. On either common base, exchanging endpoint coefficients changes mean loss by only 0.000033 or 0.000057; the corresponding intervals lie inside the declared ± 0.0006 margin. Holding either coefficient fixed, changing the base gives 0.008453 or 0.008430, favoring fixed continuation in every seed. These within-lineage exchanges are different from arbitrary Foreign transfer. They are consistent with compensatory adaptation or another change in the multi-task path, but do not identify a unique mechanism. Five of seven deferred trajectories end on the radius boundary, and recorded constrained steps are not uniformly task-descent steps; no convergence claim follows.

At budget 8,192 and calibration resample 0, guarded deferred120 gains 0.012242 over the full fixed checkpoint, with 6/7 active and no observed quality violations. Its paired advantage over guarded periodic120 is 0.006151 $[0.001141, 0.010257]$. Deferred60 is active on 7/7 and gains 0.013326, but admits one quality failure. A larger mean policy gain is therefore not synonymous with the stronger quality-constrained outcome. Deferred20's favorable comparison with periodic120 is post hoc; the EVAL-selected economical budget is not promoted to a new primary policy.

B.5.3 NATIVE ORDERING AND COMMON-ANCHOR FITTED ORDERING

Campaign 7 fits the ordinary, fixed, and periodic bases using a common historical anchor, rate, penalty center, and radius. It also tests periodic warm initialization from that path's terminal coefficient and an identity-centered ordinary fit. The primary added budget is 120 updates; 20 and 240 are prespecified sensitivities. The ordinary-common arm is a planned secondary control. Its native-versus-refitted ordering is displayed in Table 30; expanded pairwise and guarded interpretations remain descriptive follow-through, not retrospectively primary hypotheses.

Table 30: Native and refitted endpoint ordering, Campaign 7. Native loss includes the operator used during that base's training. All common fits initialize and constrain around the same inherited anchor, with 120 new updates. Guarded gain uses the complete fixed checkpoint as reference and includes fallback. Seven seeds; lower loss is better. Ordinary common is a planned secondary control; expanded contrasts are descriptive. Sources: V7_02–04.

Base / fit	Native loss	Fitted loss	Guarded gain	Active / failures
Ordinary / common	0.350821	0.307231	0.025299	6/7; 0
Fixed / common	0.337175	0.322687	0.012242	6/7; 0
Periodic / common	0.331084	0.331137	0.006038	7/7; 0
Periodic / warm	0.331084	0.330648	0.006526	7/7; 0

Base / fit	Native loss	Fitted loss	Guarded gain	Active / failures
Ordinary / identity	0.350821	0.332318	0.004857	7/7; 0

Raw common-refit quality failures: ordinary 1/7, fixed 1/7, periodic 0/7; periodic warm and ordinary identity have 0/7. Identity-start fitting changes the feasible center, penalty center, and calibrated LR. The cross-base ordering has not been subjected to matched affine calibration.

Ordinary common beats fixed common by 0.015457 $[0.012996, 0.018812]$, with seven positive paired seed contrasts. After the primary point guard, it gains 0.025299 $[0.014444, 0.035454]$ over the complete fixed checkpoint, active on 6/7 with no observed failure. The ordinary base's native loss is worse, but its endpoint after the specified common-anchor fit is better. This is a protocol-defined comparison, not history-free ordinary fine-tuning. All common paths inherit information; the ℓ_2 norm of the anchor is 2.217–2.404, so radius-one balls centered at identity and anchor are disjoint in every seed. The identity comparison changes initialization, feasible set, penalty center, and FIT-calibrated rate. Campaign 8 supplies same-base zero-step attribution, but does not calibrate the ordering across these three bases.

The fixed-common advantage survives a stronger periodic comparator: 0.008449 $[0.005868, 0.010939]$ over periodic common and 0.007961 $[0.004637, 0.010698]$ over periodic warm. Native periodic continuation still improves over native fixed, so the new comparison does not erase the earlier useful live result. The ordering is stable at 240 updates, but small mean increments do not prove convergence. At that budget the fixed-common point guard admits seed 31 and incurs a quality failure that the 120-update guard rejected. Ordinary common's seed-31 raw ranking change is -0.010315 versus ordinary, with interval $[-0.013584, -0.007045]$, despite its positive loss gain; primary guarding abstains on that case.

B.5.4 Shadow fitting separates chronological updates from base feedback

The shadow procedure carries coefficients through six saved fixed-base states, taking 20 FIT updates at each, and evaluates on the common final base. It does not train an independent shadow base and its new coefficients never alter the saved base trajectory. Shadow exceeds deferred fitting by 0.000036 $[-0.0000068, 0.0001010]$, inside the declared ± 0.0006 loss margin, and exceeds coupled periodic by 0.008433 $[0.005860, 0.010886]$. The first milestone is identical before the paths diverge. Thus sequential fitting itself is compatible with the deferred endpoint benefit. Coupled coefficients were already detached during base updates; the distinction is forward feedback, not correction of a missing detach operation. Similar terminal loss within a margin does not imply identical vectors, a shared optimum, or universal schedule equivalence.

B.6 What limits admission, attenuation, and the value of more data?

B.6.1 Admission information for fixed candidates

Campaign 5 varies calibration evidence while holding candidates and quality floors fixed. Within this campaign, increasing from 101 to 8,192 lists raises periodic activity from 11/35 to 33/35 and mean gain from 0.002019 to 0.006012 $[0.003947, 0.008331]$, with no observed all-head quality violations. The 35 records are seven seeds times five dependent resamples. No coefficient refitting occurs in this block, and all raw candidates pass the larger EVAL population. The V4-to-V5 EVAL population changes, so the clean calibration- budget contrast is within V5; changing failure counts across releases cannot be attributed solely to more calibration data.

The three-arm TUNE-selected bank gains 0.006014, essentially the same mean as periodic-only admission's 0.006012; it does not average candidates. The total gain over ordinary continuation is 0.020352, consisting of 0.014340 already in the fixed system plus the 0.006012 increment. Direct selection of the incumbent is inactive even without a GUARD rejection. Reducing the fitted radius to 0.25 or 0.5 sacrifices gain and does not order safety: radius 0.25 introduces an admitted failure in seed 31 on both second-epoch branches. Those two records share one seed.

B.6.2 COMPLETE FAMILY STRESS TEST AND NEAR-BOUNDARY ERRORS

Campaign 6 tests seven candidate families at calibration budgets 2,048 and 8,192, using 60% TUNE and 40% GUARD, with five fixed resamples. It requires Next loss non-worsening against fixed and NDCG at least -0.002 relative to ordinary and fixed on each of three heads. The stress bank chooses once among six candidate families and hold. Table 31 preserves every family/budget row. A finite-EVAL point failure among active decisions is an observed false accept under this contract, not a known population error.

Table 31: Complete Campaign 6 admission stress summary. Each row contains 35 correlated decisions (seven seeds, five calibration resamples). FA is an active candidate below the finite-EVAL point floor; FR is a rejected positive, point-feasible candidate, defined for the stress bank's TUNE winner. Gain is all-seed, averaged over splits within seed, relative to the complete fixed checkpoint. All FA records concern seed 31. Source: V6_13.

Family	Lists	Active / 35	FA	FR	Gain
Deferred120	2,048	33	3	0	0.013589
Deferred120	8,192	30	0	0	0.012242
Periodic120	2,048	31	0	4	0.005409
Periodic120	8,192	34	0	1	0.005902
Permuted	2,048	28	4	1	0.002180
Permuted	8,192	29	4	0	0.002184
Radius .25, frozen	2,048	33	5	2	0.001854
Radius .25, frozen	8,192	35	5	0	0.001932
Radius .25, refresh	2,048	33	4	1	0.001940
Radius .25, refresh	8,192	35	5	0	0.002048
Reversed	2,048	0	0	0	0
Reversed	8,192	0	0	0	0
Stress bank	2,048	33	3	1	0.013589
Stress bank	8,192	32	1	0	0.012484

Larger calibration removes observed deferred120 false accepts but leaves small-radius failures and introduces one for the stress bank. All Block B false accepts involve seed 31 near the ordinary-relative floor: fixed Next NDCG changes by -0.001921 versus a -0.002 tolerance, and deferred120 by -0.002685. The quality intervals of false-accept cases straddle the floor. Reversed correction gives mean loss gain -0.014732, harms all seven seeds, and is rejected in all 70 screenings. A single permuted correction per seed retains partial utility (0.002020 mean gain, 6/7 loss-positive and 5/7 jointly point-feasible), but post-retraction norms are only approximately matched at 98.26–100% of deferred norms. Rejecting gross harm and resolving a narrow quality boundary are different capabilities.

B.6.3 COMMON-SET INITIALIZATION AND WEAKER ACTIONS

Campaign 6's three-seed challenge places both held and identity starts in the same identity-centered radius-one family, with the same penalty and FIT-calibrated rate. Warm advantage falls from 0.008345 at 20 updates to 0.00000667 at 80, interval $[-0.000139, 0.000145]$, within the declared ± 0.0006 margin. Warm initialization accelerates early fitting here; the experiment does not establish a superior warm optimum. Guarded 80-step means are 0.001785 and 0.001787, active on 2/3 with no observed violations. This shared-set test is different from the distinct-center Campaign 7 comparisons.

The rejected seed-17 full action gains 0.021855 in loss relative to its held reference while losing 0.026409 NDCG versus identity, with ranking interval $[-0.029463, -0.023354]$. Campaign 7's strength bank finds that $1/16$ strength yields 0.001276 loss gain and -0.001185 NDCG change. A harmful full action can contain a useful weaker action, without every direction becoming safe under shrinkage. Ordinary common is not included in this Campaign 7 challenge bank. The primary shadow point rule admits 7/7 with one failure; its LCB counterpart admits 6/7 with none. Across resamples, two LCB admissions at strength $1/4$ remain slightly below the EVAL point floor. These records do not establish population harm where intervals straddle the floor, but they preclude a claim of zero observed point failures across the sensitivity study. Additional A-block LCB comparisons retain their post hoc sensitivity status.

B.6.4 TRAINING SCALE VERSUS EXPOSURE

Campaign 5's three-seed side study compares 10% and full training data under constant-LR fresh training, a fixed pair, and both equal-pass and equal-step views. The 10%-minus-full Fresh gain is 0.007559 at equal passes and 0.000479 at equal steps; both intervals include zero. On the same 3,130-list support, the corresponding means are 0.007340 and 0.000514. Conditioning on seen users and candidate items does not match all distributional factors. The full-data views share a checkpoint, and equal steps need not equal exactly the same number of example presentations. The evidence does not establish a data-size law or explain Amazon–Kuai differences causally.

Refresh in this side study uses a historical center, whereas Fresh uses identity; refresh identity norms often reach 1.65–1.86 against Fresh's radius 1. Larger refresh gain is not an initialization-only comparison. The source quality ledger preserves the scale study's -0.002011 failure, the Kuai long-view harm and AUC failure, and the small-radius persistence failures. Threshold crossings are evaluated before display rounding and are not removed for being numerically small.

B.7 WHERE DOES FINAL GAIN COME FROM, AND WHAT REPEATS ON FRESH WEIGHTS?

B.7.1 TWO COHORTS AND THEIR TRAINING AND EVALUATION HISTORIES

The final experiment has two cohorts. Core comprises seeds $\{13, 17, 23, 29, 31, 37, 41\}$ and reuses Campaign 7 ordinary-MTL bases and previously selected per-seed anchors/pairs. Fresh comprises $\{43, 47, 53, 59, 61\}$, with new initializations, 12 ordinary-MTL epochs, fixed pair $[0, 1]$ and direction $i \to j$. Each fresh trajectory fits its own epoch-6 anchor using the reconstructed historical recipe; a core operator is not transported into an unrelated initialization. Both target Next with one visible intervention unit per seed. Each cohort has 8,192 EVAL lists and 8,192 observed unique users. Core uses calibration budgets 2,048 and 8,192 with five splits; fresh uses 8,192 and one split.

Fresh EVAL overlaps core EVAL on 2,586 users/rows and the disjoint core calibration reservoir on a further 5,255: 7,841 of 8,192 fresh EVAL rows have appeared in these earlier roles. Within-cohort FIT/TUNE/GUARD/EVAL separation and locked selection remain distinct from adaptive reuse across campaigns. The remaining 351 rows are not certified externally untouched. Fresh weights repeat a recipe and its same-base response measurements; they do not independently replicate every earlier trajectory comparison. The cohorts differ in histories, pair choice, and rate procedures and are not pooled into a homogeneous twelve-seed interval.

B.7.2 SAME-BASE ATTRIBUTION AND THE AFFINE ALTERNATIVE

On the same frozen base and EVAL lists, let z_0 be the historical anchor and z_{120} the refit. The accounting identity is

$$H = \mathcal{L}(0) - \mathcal{L}(z_0), \qquad F = \mathcal{L}(z_0) - \mathcal{L}(z_{120}), \qquad T = H + F. \tag{31}$$

Here zero is identity. The new fit uses FIT20 and 120 SGD updates, with radius 1 and penalty 0.01 about its own path center, box $[-0.5, 0.5]$, and clipping 5. Core intact and identity-start endpoints are imported from Campaign 7; the three core permuted fits are new. Table 32 gives the same-base results and calibrated comparisons. Both H and F are positive in every seed of each cohort. The inherited fractions 69.6% and 72.9% are ratios of cohort means; mean per-seed fractions instead are 70.6% and 71.9%. This accounting does not identify causal representation drift or the origin of historical benefit. It separates zero-step retained value from new fitting for the final ordinary-base result.

Table 32: Same-base raw attribution and affine controls, Campaign 8. All values are absolute Next-loss reductions; intervals are exploratory 95% seed bootstrap intervals conditional on the observed data. Core and fresh are distinct cohorts; displayed values are rounded independently. Cal fits separate FIT-only positive affine transforms to each endpoint. H, total, and identity-start gains use same-base identity; F uses the anchor; calibrated contrasts name both endpoints. Sources: final V8 P02, P04 and consolidated T03, T06.

Quantity	Core (7 seeds)	Fresh (5 seeds)
Inherited anchor H	0.030342 $[0.024671, 0.037379]$	0.051348 $[0.038732, 0.064333]$

Quantity	Core (7 seeds)	Fresh (5 seeds)
Additional refit F	0.013248 $[0.009340, 0.017496]$	0.019088 $[0.014356, 0.023238]$
Total $T = H + F$	0.043591 $[0.034280, 0.054692]$	0.070436 $[0.056629, 0.083922]$
Identity-start refit	0.018504 $[0.015160, 0.023194]$	0.034564 $[0.024277, 0.044851]$
H/T, ratio of means	69.6%	72.9%
Identity calibration alone	0.166739	0.290854
$\mathrm{Cal}(z_{120})$ over $\mathrm{Cal}(0)$	0.001598 $[0.000590, 0.002522]$	-0.003544 $[-0.008271, 0.001183]$
$\mathrm{Cal}(z_{120})$ over $\mathrm{Cal}(z_0)$	0.000457 $[0.000093, 0.000768]$	-0.001427 $[-0.004662, 0.000427]$
Raw refit over $\mathrm{Cal}(0)$	-0.123148 $[-0.164955, -0.086284]$	-0.220418 $[-0.301091, -0.152657]$

Each identity, anchor, and refit endpoint receives a separate FIT-only positive affine logit transform $a\ell + b$. Fitting uses native float32 logits with float64 L-BFGS-B, $a \in [10^{-4}, 10^{4}]$, and $b \in [-100, 100]$. All 36 fits converge; core solutions are interior, while fresh seed 61's anchor/refit slopes reach the lower bound. The archived rank arrays remain unchanged. This is a rank-preserving BCE control, not evidence about ECE, Brier score, or real-world probability calibration. It is a two-parameter alternative action family, not a matched historical-compute comparison.

Affine calibration of identity gains 0.166739 in core and 0.290854 in fresh and beats uncalibrated full refit BCE in every seed. Calibrating both endpoints leaves a small positive core residual (0.001598), while the fresh residual is -0.003544 with an interval spanning zero. Calibrated refit beyond calibrated anchor adds 0.000457 in core and -0.001427 in fresh. It is incorrect either to attribute all raw gain to distinctive feature adaptation or to say all useful signal disappears. Raw-minus-calibrated differences are not causal mediation fractions. Other-base alternatives and permuted endpoints lack this same affine panel; the cross-base ordering and coordinate mechanism remain unisolated by calibration.

B.7.3 Ranking and selective recovery with explicit coverage

Full refit has positive average NDCG in core but a negative fresh mean whose seed interval spans zero (Table 33). Point-tolerance failures occur in core seed 31 and fresh seeds 47, 53, and 61. Fresh seed 61's NDCG goes from 0.391633 at identity to 0.221673 at anchor and 0.145427 after refit while BCE improves. For these three fresh failure seeds, no source LR candidate qualified at epoch 6 and the prespecified middle-rate fallback supplied the anchor; the two later successes had qualifying source rates. This descriptive pattern weakens an explanation based solely on later representation motion but does not identify a validated classifier of success.

Table 33: Raw quality and primary selective recovery, Campaign 8. Same-base identity is the reference unless an anchor-relative contrast is named. The primary rule uses the incremental bank, 8,192 calibration lists, resample 0, one TUNE winner, one GUARD check, and identity fallback. Policy means include all seeds. Intervals are exploratory seed intervals; fresh intervals reaching zero and the active denominators remain explicit. Sources: final V8 P03, P06–07, T04–05.

Outcome	Core (7 seeds)	Fresh (5 seeds)
Raw refit ΔNDCG@10	0.013338 $[0.003566, 0.022781]$	-0.062320 $[-0.160004, 0.011670]$
Additional refit ΔNDCG@10 (over anchor)	0.003315 $[-0.000203, 0.006199]$	-0.026614 $[-0.056904, 0.003507]$
Raw refit ΔHR@10	0.031006 $[0.017561, 0.044120]$	-0.061963 $[-0.180811, 0.023145]$
Raw quality failures	1/7	3/5
Policy active	6/7	2/5
Policy all-seed loss gain	0.036746 $[0.021925, 0.051492]$	0.022730 $[0, 0.050785]$
Policy all-seed ΔNDCG@10	0.014812 $[0.007393, 0.023069]$	0.007952 $[0, 0.016984]$
Policy all-seed ΔHR@10	0.029384 $[0.014334, 0.043701]$	0.013379 $[0, 0.028955]$
Observed active violations	0/6	0/2

The primary policy is the same-base incremental bank, calibration budget 8,192, resample 0, and cluster-LCB quality rule. TUNE selects one candidate, GUARD checks it once, and rejection returns same-base identity without retry. Quality uses an approximate user-cluster t lower bound with Bonferroni adjustment over three heads; target loss uses point non-worsening. It is not a loss-confidence guarantee or a simultaneous campaign-wide guarantee. All admitted primary candidates are full refits; the four failure seeds select identity at TUNE. Hence primary point and LCB rules choose exactly the same endpoints in both cohorts, and the primary gain cannot be attributed specifically to confidence-aware guarding.

The core policy is active on 6/7 and fresh on 2/5, with no observed active violations. All-seed means include identity fallbacks; the fresh intervals reach zero. The policy gain itself comprises inherited/additional means $0.025560/0.011186$ in core and $0.015352/0.007378$ in fresh. A post hoc calculation calibrates these already-selected active sets, giving residual all-seed gains 0.001722 $[0.000903, 0.002534]$ and 0.000701 $[0, 0.001514]$. This is an attribution diagnostic, not a prospectively tested calibration-aware policy. The core fixed-reference continuity result, 0.025299 loss and 0.010459 NDCG gain at 6/7 coverage, uses a different complete-system reference and must not replace the same-base policy values.

B.7.4 A HARMFUL ANCHOR CHANGES WHAT ATTENUATION CAN RECOVER

Incremental shrinkage follows $z_0 + \alpha(z_{120} - z_0)$, whereas whole-vector shrinkage follows αz_{120}, for $\alpha \in \{0, 1/16, 1/8, 1/4, 1/2, 1\}$. The first returns to the anchor; the second returns to identity. Every incremental strength fails the point floor for core 31 and fresh 47/53/61. Core 31's anchor itself has NDCG change -0.004306. Its whole-vector half-strength action gains 0.024558 loss with -0.001250 NDCG change, but the GUARD LCB is -0.005070 and its EVAL interval crosses the floor, so LCB admission abstains. Whole-point selection covers 7/7 core with no point failure, but 4/5 fresh with a seed-53 violation. Fresh seed 61 still loses 0.004926 NDCG at the smallest positive whole strength $1/16$. No positive dose in this grid is point-safe for all five fresh seeds; this is a candidate-family boundary, not proof that every alternative direction would fail.

Across all core sensitivity contracts, 221 active LCB decisions have no observed breach; 260 active point decisions include two records of the same seed/candidate/split failure under two references. Core same-base incremental LCB coverage increases from 26/35 at budget 2,048 to 30/35 at 8,192; the fixed-reference contract increases from 23/35 to 30/35. These correlated records do not estimate an independent safety rate. The guard can distinguish some near-boundary cases without certifying the fallback or ensuring universal monotonic improvement with budget.

B.7.5 SIGNED-COORDINATE PLACEMENT UNDER DIFFERENT FEASIBLE CENTERS

The intact anchor beats the mean of three signed-coordinate permutations in raw BCE before and after local refitting for every seed. Before-refit advantages are 0.029275 $[0.023122, 0.036889]$ core and 0.048947 $[0.035669, 0.061976]$ fresh; after-refit advantages are 0.025472 $[0.020436, 0.032070]$ and 0.037085 $[0.027548, 0.047367]$. Signed-coordinate reassignment therefore changes raw loss under this bounded protocol. Each permutation has its own feasible center, and the intact endpoint is outside the permuted radius-one neighborhoods. Historical anchors also lie outside identity-start radius-one balls. The result does not equalize reach or historical compute and does not reverse the matched donor-magnitude falsification in Appendix B.3.1.

The corresponding after-refit NDCG contrasts are 0.004574 $[-0.001604, 0.010615]$ core and -0.044390 $[-0.105179, 0.005841]$ fresh. Calibrated permutation controls are absent, so the raw-loss advantage does not identify a transferable semantic basis. Identity-start fitting has smaller mean loss benefit, but zero core point-ranking failures; it fails on the same three fresh failure seeds. This limits a universal warm-start interpretation and reinforces the separation between the objective and quality of an intervention.

C Interpretation and Scope

C.1 Interpretation of State-Relative Adaptation

The experiments establish a state-conditioned response under explicit action, information, and quality contracts. Three separations organize the interpretation: current-state value versus replay compatibility; the training path versus the coefficients applied at its endpoint; and raw objective gain versus calibrated and ranking-qualified recovery. The foundational studies establish the first separation across domains, while the later Amazon controls resolve the second and quantify the third.

C.1.1 State-dependent high-value intervention region

The empirical results are easier to interpret as a state-conditioned response landscape than as recovery of one unique target vector. Let $\mathcal{A}_s$ denote an explicitly specified *tested* action family at frozen state s, after applying the same action-validity and secondary-quality contract used by that experiment. Define the best observed eligible value within that finite protocol,

$$\widehat{U}^{\max}_{\epsilon,\mathcal{A}}(s) = \max\left\{0, \max_{a\in\mathcal{A}_s} U_\epsilon(s,a)\right\}, \tag{32}$$

and the corresponding η-high-value set

$$\mathcal{O}_{\eta,\mathcal{A}}(s) = \left\{a \in \mathcal{A}_s : U_\epsilon(s,a) \geq \widehat{U}^{\max}_{\epsilon,\mathcal{A}}(s) - \eta\right\}. \tag{33}$$

The subscript $\mathcal{A}$ is essential: this is a finite protocol-relative set, not a global action optimum or mathematical ceiling.

This language accommodates several otherwise awkward observations without strengthening them. Near-tied actions can differ in pair or direction while having nearly indistinguishable held-out value; the rotated-basis eligible-value contrast remains unresolved while a higher-capacity rank-1 family reaches stronger response under the reported displacement control; and the frame-local codec can change coefficients noticeably while preserving average behavior. Fresh is therefore best treated as the state-specific representative returned by the prespecified Fresh fitting protocol for the reference family, not as the unique element of $\mathcal{O}_{\eta,\mathcal{A}}(s)$ and not as a global optimum.

C.1.2 Local regularity and non-monotonic compatibility

Replay makes compatibility a relation between an action and the state on which it is applied. Amazon and Kuai contain regimes in which an earlier operator retains substantial value on nearby descendants, making a finite replay bank empirically plausible. That observation is local: it does not license a monotone age-to-compatibility law or a same-lineage transport theorem.

Production supplies the decisive boundary. Same-lineage replay changes sign over the tested horizons even though Fresh remains useful, while raw Foreign replay remains weak. The Amazon alignment-only control and measured-similarity audits leave a substantial replay gap; target-aware dose-selected masking recovers additional value under a different protocol (Appendix A.6.8), but they also do not identify lineage itself as causal; the Kuai similarity supports do not overlap sufficiently for that contrast. The surviving object is therefore *current-state compatibility*. Lineage, elapsed horizon, pair identity, direction, and measured similarity are contextual variables, not certificates.

C.1.3 Response rather than structural energy

Appendix A.4 already establishes the empirical point: measured structural quantities can be real and informative without providing a monotonic task-utility objective. The interpretive consequence is that the primitive label in AMIYC is *behavioral response to a specified action from a specified state*. Structural measurements may condition prediction of that response, as they do in the compatible Amazon amortization regime, but the measurement itself need not be a potential to minimize.

This also limits how much should be read into normalized notions such as "susceptibility per unit intervention." Such a quantity would require a predeclared action-distance definition and validation that the denominator is behaviorally meaningful across scopes and parameterizations. The present

work measures realized response and realized displacement separately; it does not identify a universal response-per-distance law.

C.1.4 Coordinate controls and output calibration

The positive diagonal initially made a native-axis mis-scaling explanation plausible. K1-D retains residual Fresh response when bounded diagonal parameters are present and trained. The fixed rotated-basis family has stronger raw primary response while its eligibility-gated contrast remains unresolved; this supports native-axis nonnecessity in this one-basis test, without establishing eligible equivalence or universal rotational invariance. The rank-1 control shows that the reference diagonal is not the only useful intervention family, but differs in parameter count even under a realized-displacement comparison. K1-N leaves response after normalization while changing selected base quality, so a quality-matched normalization explanation remains unresolved. These construct controls neither isolate a native-axis defect nor resolve output-logit calibration.

Campaign 8 supplies the stronger output-affine comparison (Appendix B.7.2). FIT-only positive affine calibration of identity beats uncalibrated refit BCE in every core and fresh seed while preserving ranking. Calibrating both endpoints leaves a small positive core residual and an unresolved fresh-cohort residual. Raw BCE response therefore does not by itself identify representation improvement; useful selected ranking outcomes remain a separate result. The diagonal provides a low-dimensional, replayable reference for measuring these objective-dependent responses.

C.1.5 Behavioral geometry of intervention coordinates

The action-space closures also separate raw-coordinate identity from behavioral identity. Pair and direction can differ while endpoint value is nearly tied; composition is strongly subadditive; and a compact frame-local codec reconstructs operators with small average behavioral error despite nonzero coefficient error. At the same time, the codec produces a small secondary-boundary tail, and raw cross-lineage transport remains weak. Euclidean coefficient distance is therefore neither established as a sufficient behavioral metric nor shown to be irrelevant.

A useful future object would be a response-induced metric or equivalence relation validated directly against held-out primary value and distance to the secondary-quality boundary. A cross-model claim would additionally require an explicit identification map between representation frames and a held-out transport test. Neither object is estimated here. The present result is only that behavior can be more stable than raw action identity within a frame, not that a global operator manifold has been recovered.

C.1.6 Moving target as a tracking problem

The tracking comparisons separate coefficient-fitting chronology from feedback into base training. Within the fixed positive-diagonal reference family, let z_t^{Fresh} denote the coefficients returned by the Fresh protocol after the evolved state s_t is frozen. For any online coefficient endpoint z_t, the coordinate diagnostic

$$e_t = z_t - z_t^{\mathrm{Fresh}} \tag{34}$$

records one notion of endpoint tracking error; its behavioral relevance must still be checked on EVAL.

Appendix A.7 constrains the interpretation in two directions. When the evolved representation is held fixed, ordinary MTL supervision moves the same coefficient coordinates toward the contemporaneous Fresh reference and recovers substantial primary value. Under both tested online ordinary-MTL schedules, the endpoint update remains anti-aligned with its own contemporaneous Fresh direction. The block-coordinate control changes the base trajectory without rescuing tracking. This rejects rescue by that specific schedule; it does not isolate update ordering or state motion as the cause.

These comparisons support historical recipe failure rather than an identified state-motion cause. Repeated fitting can vary more than nearby-state changes (Appendix B.3); actual live correction can improve the complete system (Appendix B.5.1); and crossed and shadow comparisons associate the stronger endpoint gap mainly with the base path (Appendices B.5.2 and B.5.4). The crossed endpoint and shadow readouts provide the behavioral evidence for this base-path association.

C.1.7 PRIMARY VALUE AND SECONDARY ELIGIBILITY ARE DISTINCT SURFACES

The deployment results make a final interpretive separation. The surface that orders primary response is not the same object as the secondary-quality eligibility boundary. More expressive adapters can improve the primary objective while degrading ranking; a frame-local codec can preserve mean behavior while crossing the eligibility boundary in a small tail; and Production shows that high secondary eligibility can coexist with nonpositive primary value.

Consequently, a future behavioral geometry for deployment would need to represent at least three objects jointly: current-state primary response, distance or margin to the secondary-quality boundary, and the subset of actions actually reachable from the available bank. Appendix A.8 measures these quantities operationally; it does not identify a differential geometry or reduce them to one scalar "safety" coordinate.

C.1.8 REMAINING SCOPE

The original lifecycle and adapter-family boundaries remain: the larger raw post-peak gap is quality-coupled, and the archived adapter ladder lacks an all-family matched-capacity, matched-displacement comparison. Later results add further limits. Fresh weights reuse examined data; affine-adjusted cross-base ordering and calibrated permutation contrasts are untested (Appendix B). Table 34 records extensions beyond the current evidence, not experiments claimed here.

Interpretive synthesis. AMIYC establishes measurable, state-relative adaptation value and distinguishes how it is retained or acquired. The appropriate description includes the base state, historical operator information, fitting family, evaluation objective, and quality constraint. Together these quantities explain why native quality, coefficient motion, and action availability answer different questions. Appendix C.2 summarizes the controls supporting this interpretation.

Table 34: Prospective extensions beyond the current evidence. These are unexecuted comparisons that could broaden generality or identify mechanism; the current claims retain their stated boundaries.

Open axis	Current boundary	Minimal evidence that would materially extend it	Role
Post-specialization saturation	Fresh is measured after ordinary joint-MTL training and checkpoint selection. Prior work already shows that multi-task models, including sparse MoEs, can recover additional task-specific value through subsequent specialization (Liu et al., 2019; Gupta et al., 2022); the present experiments do not establish that the bounded Fresh response persists after such refinement.	Continue a selected MMoE on one target task to a validation-selected specialized endpoint, freeze that state, and rerun the same Fresh FIT–SELECT–EVAL audit.	interpretation / construct boundary
MMoE specificity	Current-state bounded response is established in three recommendation/ranking MMoE settings; whether the same state-relative response/replay structure depends on shared-expert MMoE architecture is not identified.	Run one matched dense shared-bottom multi-task baseline on the same tasks/data under a prespecified bounded intervention and the same FIT–SELECT–EVAL contract.	scope extension
Quality-matched normalization	K1-N leaves residual Fresh response but materially changes selected base quality.	Use a normalization design whose ordinary selected base quality is matched closely enough that residual Fresh response can be compared without the present quality confound.	construct validity
Compatibility coordinates	Same replay can persist, disappear, and reappear; tested similarity variables do not certify compatibility, and Kuai lacks overlapping Same/Foreign support for an exact matched comparison.	Evaluate pre-intervention matching or prospective compatibility prediction on genuinely overlapping support using untouched target-state replay outcomes.	identification
Tracking mechanism	The historical online schedules fail, whereas later live correction helps relative to fixed continuation. Crossed and shadow controls associate most of the stronger endpoint gap with the base path; mechanism and calibrated cross-base ordering remain unresolved.	Isolate a proposed feedback mechanism and compare bases under the same affine controls on an untouched population.	mechanism
Adapter equivalence	Several families access useful response, but unresolved rotated/global contrasts and different capacities do not establish general equivalence.	Predeclare a response-induced distance or equivalence rule and test whether it predicts held-out primary equivalence and secondary-boundary crossings better than raw coefficient distance.	representation
Bank-regime prediction	Amazon and Kuai show that selection quality and bank compatibility are separable; current-state opportunity can disappear before candidate ranking begins.	Evaluate a label-free pre-replay predictor of finite-bank retention or incompatibility prospectively before target replay outcomes are exposed.	deployment
Broader model classes	Current evidence is restricted to recommendation/ranking MMoEs.	Apply the same frozen-state response/replay contract to a genuinely different jointly trained architecture, sparse MoE, or language-model MoE.	external scope

C.2 What the Controls Distinguish

C.2.1 Scientific comparisons

The controls separate explanations for observed adaptation value. Each comparison holds a declared part of the system fixed and measures the resulting behavior; the estimates and uncertainty remain in their experimental sections.

Table 35: Scientific distinctions supported by the controlled comparisons.

Question	Control	Finding
State description versus objective	Structural sweeps and intervention responses (A.3–A.4)	Witnesses describe state and condition response prediction; their scalar ordering differs from intervention value.
Availability of coordinates	K1-D and alternative same-boundary adapters (A.5)	Residual response remains on trained coordinates and across reference families; composite bounds and parameter counts specify the comparison.
Compatibility of an old operator	Same/Foreign replay, target-gradient transport and matched donor controls (A.6, B.3)	Behavior depends on the target state. Lineage and coefficient movement alone do not determine replay value.
Chronology versus base feedback	Deferred, crossed and shadow fitting (B.5)	The tested endpoint gap is predominantly associated with the base path; shadow and deferred fitting give similar endpoints.
Inherited versus new gain	Zero-step anchor and common-state refit (B.7.2)	Most final raw gain is inherited; the additional refit contribution is measured separately.
Objective versus ranking response	Positive affine calibration and ranking readouts (B.7.2)	A small positive calibrated residual remains in the core cohort; fresh-weight uncertainty spans zero, while selective ranking recovery remains measurable.
Availability versus selection	Frozen-bank oracle, witness policy and paid acquisition (A.8, B.7.3)	Bank transfer and action selection impose separate limits; acquisition expands candidates under a different information and compute contract.

C.2.2 Using these comparisons

The findings support a common empirical framework while retaining the conditions of each result. Amazon's optimizer and recovery gains sit alongside Kuai's transfer boundary; original bank reuse sits alongside separately paid candidate acquisition; and raw loss sits alongside affine and ranking readouts. This organization preserves both the shared phenomenon and its domain-specific limits.

C.3 Scope of the Current Conclusions

Integrated interpretation. Current-state fitting demonstrates usable response at frozen checkpoints. Replay evaluates the value of retaining an earlier operator, and selection measures how much of an available bank can be recovered. The later controls add a training-path comparison: useful live correction is possible, while deferred fitting and shadow updates expose the cost of feedback into the base under the tested coupled schedule.

Objective and recovery contract. Historical information accounts for most final raw Amazon gain. Positive affine calibration separates a rank-preserving loss response from intermediate intervention effects. Core and fresh-weight results therefore report raw gain, calibrated residual, ranking, admission, and abstention separately. Persistence is measured for one additional epoch with the operator retained. Cross-base ordering after matched affine calibration and generalization to untouched external data remain prospective tests.

D Reproducibility and Evidence Navigation

D.1 Reproducibility and Experimental Artifacts

The evidence links each reported quantity to its cohort, protocol, model state, action identity, and analysis output. Candidate counts and independent seed/lineage units are recorded separately.

D.1.1 Protocol contract

FIT, SELECT, EVAL, and where applicable ALIGN have distinct roles. Intervention fitting and executable decisions use their designated information; retrospective best-pair and bank maxima are labeled evaluation summaries. Base-checkpoint selection may share the parent development population, and cross-campaign population reuse is reported in Appendix B.7.1. Each fitted operator is stored with its action, coefficients, selected setting, and split identity. Replay applies that operator without target refitting. Versioned protocols preserve the identities of corrected implementations.

Table 36: Experimental artifacts and their inferential units. The registry supplies stable source identifiers and version bindings.

Result family	Protocol / artifacts	Evaluation unit
Original opportunity	Frozen intervention backend; checkpoint and split manifests, action summaries and stored operators.	Amazon original 360-diagonal panel; Kuai eight seeds; Production nine targets in three lineages. Expanded Amazon rows are separate.
Construct controls	K1 architecture and learned-parameter audits; rotated-basis and rank-1 adapter protocols.	Seed-level contrasts, with selected-base quality, parameter count and displacement recorded.
Replay and transport	Fresh/Same/Foreign state manifests; shared hard-mask core, 666 paired action rows and sign-sufficient records.	Kuai eight seeds and Amazon seven seeds for the target-gradient port; Production three lineages.
Continuation and tracking	Matched continuation endpoints, contemporaneous refits, crossed and shadow controls.	Paired seed-level outcomes on the specified base trajectories and readout contracts.
Frozen-bank selection	Original Amazon v1.0.2 controller; OOF and outer-fold manifests, predictions, policy decisions and final analysis.	183 source states, seven seeds; 122 temporal-forward held-out panels. Kuai retains its own 48-panel contract.
Structured action space	Canonical structured-closure analyzer and supervision exporter; singleton, allocation, composition and codec outputs.	210 catalogued states and 2,619 operators; each control retains its matched support or frame.
Attribution and admission	Campaign protocols, committed job outputs, scalar/vector readouts and 36 affine fits.	Seven core and five fresh training seeds; fresh evaluations reuse previously examined data.

D.1.2 Experimental sources and analysis entry points

The separate scientific archive contains source snapshots, campaign protocols, result tables, and verification inputs. In its historical source tree, `supplementary/audit_v3_2/reproduce.py` checks retained transport and closure summaries; `supplementary/audit_v3_2/forward_and_mask_source.py` provides the forward/mask source. The campaign source subtree `evidence/campaign_composite/campaigns/C08/sources/reproduce/executed_code/` contains the V8 controller and supporting implementation. Source-era paths in the registry identify this archive; the clean manuscript package contains the active TeX sources, figures, PDFs, and current evidence registry.

D.1.3 DATA ACCESS AND ANONYMIZED RELEASE

The public-data protocols identify the processed datasets, model dimensions, splits, seeds and evaluation roles. Production uses restricted inputs; its model and population contracts, aggregate outputs and analysis interfaces support assessment of the reported experiments. The AI-use and reproducibility statements precede the bibliography.

D.1.4 CURRENT EVIDENCE REGISTRY

The manuscript builds from `main.tex`. The accompanying `evidence_registry/` contains `claims.csv`, `evidence.csv`, `sources.csv`, `effective_claim_status.csv`, `issues.csv`, and `table_inventory.csv`. Stable claim and evidence identifiers link results to protocols and manuscript locations. The latest source bindings record the original Amazon configuration and bank-manifest verification; `bank_recovery/` supplies the matched numerical reconciliation. Historical source views are preserved separately for traceability. The build script compiles the combined document, produces paper and supplement splits, and records the active inputs.

D.1.5 INITIAL DIAGNOSIS AND AGGREGATE RECONSTRUCTION

The initial diagnosis archive supports reconstruction of 81,370 non-swap EVAL rows into 52,718 action, 2,272 state and 288 seed aggregates. The analysis checks 320 final mean cells, 120 prespecified contrast vectors, 16 retrospective rescue summaries and six swap summaries. These checks concern the retained numerical records; the original execution separately supplies native parity records. Later studies retain their own source and cohort identities.

D.2 ADDITIONAL-STUDY PROVENANCE AND MEASURED WORK

Released evidence and completeness. The companion evidence composite preserves complete campaign reports, worksheets, numbered source panels, and the claim-evolution ledger. Its imported statements retain their original statuses and are linked to current manuscript locations in the registry. The final V8 export is the union of the core and fresh exports: 966 inherited scientific and provenance members are byte-identical, and two additional files index the union. All V8 scientific blocks are complete. Export count is not cohort count, and repeated V5 endpoints in V6 or V7 endpoints in V8 are not new replications.

Audits and runtime. Archive/vector arithmetic checks verify reported arrays, selection logic, and scalar calculations within their documented scope; cached/native parity remains recorded execution evidence. None is described as an independent rerun of all model training. Campaigns 2 and 3 record about 64 minutes and 51m47s respectively with 16 workers, four per V100; Campaign 4 records about 58m47s and uses one live worker per GPU. Campaign 5 records 262 committed jobs over about 5h02m35, including actual Kuai coefficient fits and roughly 109 minutes of scale-study training. The Campaign 6 profile favors four workers per GPU for its particular workload, with throughput ratios 1.943 for fitting and 1.584 for metric work relative to one worker on the representative GPU. These workloads and overlapping stage windows do not define a controlled cross-campaign speedup or deployed-controller latency.

V8 spans 72.99 minutes from the first to last scientific job and sums 5.37 worker-hours over 86 records. New work comprises 5,460 base updates, 46,047,180 example presentations, 7,320 coefficient updates, and 36 affine fits. Of the coefficient updates, 5,520 fit current states, 1,500 fit fresh source anchors, and 300 search fresh source rates; 1,680 imported prior endpoint updates are not newly incurred. Core-anchor acquisition, setup, profiling, and export are outside the incremental workload. Fresh source projection rank is audited on five source states, not every historical checkpoint. These distinctions preserve the information and compute costs of the result.

Table 37: Campaign crosswalk for the additional evidence. Source identifiers resolve under `evidence/campaign_composite/`; `campaigns/C0x/CHECKPOINT.md` records the qualified claim and links to the original report and full numerical panels. Campaigns 2–8 are exploratory continuations; within-run primary and planned controls retain their original status.

Study	Question and experimental unit	Original panels and manuscript home
C2	Local rescue; 7 Amazon seeds, 28 states and 14 trajectories	R1–R5, M1–M5, B1, T1–T3; Section B.2. All 45 master panels remain in the source report.
C3	Refit variation and matched controls; 7 seeds, 42 states	V3_01–14; Sections B.3–B.3.1. Donor study: 28 states; schedules: 14 trajectories.
C4	Stronger fitting, live benefit, Kuai boundary	V4_01–46; Sections B.4, B.4.1, B.5.1. Amazon: 7 seeds; Kuai: 8.
C5	Admission, persistence, failed repairs, scale	V5_01–46; Sections B.6, B.5.1, B.4.1, B.6.4. Amazon 7, Kuai 8, scale 3 seeds.
C6	Deferred fit, crossed readouts, shared-set start	V6_01–38; Sections B.5.2, B.6.2, B.6.3. A/B: 7 seeds; C: 3.
C7	Endpoint ordering, shadow fit, attenuation	V7_01–37; Sections B.5.3, B.5.4, B.6.3. Main cohort 7; challenge 3 seeds.
C8	Attribution, affine control, fresh weights, policy	Final report panels 1–16; P01–P09 and source T tables; Sections B.7–B.7.5. Core 7 and fresh 5 are distinct cohorts.

D.3 Cohort and Source Bindings

D.3.1 Original experiment records

Opportunity cohorts. The original Amazon Phase-II report supplies the 360-outcome Table 1A headline. The later 540-outcome diagonal snapshot has its own stored operators and is reported separately in Appendix A.5. Kuai's original opportunity, lifecycle, and trained-coordinate controls retain their respective cohorts. The K1-D all-pair mean is descriptive; its larger composite feasible set is stated alongside that control.

Original Amazon bank. The original controller source, calibration and fold records establish the eligibility floor, probability cutoffs and matched 122-panel results. The original final manifest verifies all seven listed result members. Appendix A.8.1 reports the measured raw and gated policy values, coverage and violations. This bank population has no population-matched Fresh measurement, so its capture ratio is policy relative to bank value.

Checkpoints and objective controls. Source checkpoints and split records are retained for Amazon, eight Kuai trajectories, and the three Production lineages. Production provides 12 FIT caches and frozen FIT/SELECT/EVAL population definitions. Its pairwise logistic objective differs from pointwise BCE: a common affine bias cancels from the pairwise logit difference. The completed V8 output-affine comparisons concern their declared Amazon cohorts; they do not supply matched affine outcomes for the original opportunity populations.

Evaluation independence. Final Amazon confirmation uses new training initializations. Of 8,192 fresh EVAL rows, 7,841 overlap earlier core EVAL or calibration roles. Reported uncertainty consequently describes fresh-weight variation on the stated data, with external-population generalization retained as a separate question.

D.4 Scientific Evidence Map

Table 38 indexes the scientific questions and their experimental sections. Stable claim and evidence identifiers join the detailed CSV registry described in Appendix D.1.4.

Table 38: Scientific map of opportunity, compatibility and recovery. Detailed estimates remain in the indicated experimental sections.

Question	Measurement	Evidence location
What state is exposed?	Task-visible additive boundary, structural witnesses and direct intervention readouts.	A.3–A.4; FND and FAL registry records.
Does bounded opportunity remain?	Fresh response near selected checkpoints, trained diagonal coordinates and alternative adapters.	A.5; INT and STAT records.
When does replay retain value?	Same/Foreign replay, lifecycle motion, target-gradient transport and matched controls.	A.6; LIF and TRN records.
What changes during training?	Continuation, frozen-state supervision, coupled, deferred, crossed and shadow readouts.	A.7, B.5; OPT and Campaign 1–7 records.
What can a bank recover?	Bank oracle, action-only and witness-conditioned selection, coverage, risk and regret.	A.8; AMO and SAFE records.
How is action value structured?	Scope, dose, composition, task allocation and frame-local compression.	A.9; ACT records.
What survives stronger recovery controls?	Inherited/new gain, matched output calibration, ranking, admission and fresh weights.	B.6–B.7; Campaign 5–8 records.

The initial diagnosis is indexed below; Table 37 maps the subsequent controls to their source studies.

Table 39: Campaign 1 additions to the scientific evidence map. Existing baseline IDs remain unchanged.

Claim	Supported observation	Boundary and evidence
OPT-06	Archived online coefficients lose to held coefficients at 63/63 fixed readout states.	Same-base comparison, seven seeds; not failure versus identity. C1-E01/E02/E09.
OPT-07	Later coefficients reduce mean value on both early and late bases.	Negative in 20/21 records per base; interaction does not identify a unique cause. C1-E03.
OPT-08	Short local refitting recovers additional frozen-state value.	Quality failures and historical/selection cost remain. C1-E02/E06.
TRN-08	Target-aware masking improves weak Foreign replay.	Shuffled controls are competitive; action average is not bank-policy value. C1-E04–E09.