LET THE TARGET SELECT FOR ITSELF: DATA SELECTION VIA TARGET-ALIGNED PATHS

Anonymous authors
Paper under double-blind review

ABSTRACT

Targeted data selection seeks training examples from a candidate pool that improve downstream task performance. While trajectory-based selectors effectively guide this process, existing approaches typically construct reference states by warming up on the candidate pool itself, thereby inheriting pool-dependent distributional biases. To address this coupling, we introduce Target-Aligned Candidate Selection (TACS), which constructs a short, capacity-constrained reference path using a compact target-validation proxy. Candidates are ranked by their normalized loss reduction along this trajectory, requiring only forward passes and allowing the path to be reused across distinct pools for a fixed model-target pair. Empirically, TACS achieves competitive downstream performance across controlled logistic, vision, and NLP benchmarks, while slashing selection compute by up to 69% and candidate storage to less than 10 MB in large-scale instruction tuning.

1 INTRODUCTION

Targeted data selection identifies training examples from an uncurated candidate pool to maximize downstream performance on a specific task (Xia et al., 2024; Killamsetty et al., 2021b;a). Many scalable selectors follow a two-stage paradigm: warming up a base model to construct reference checkpoints, then scoring candidate–target interactions along this trajectory (Pruthi et al., 2020; Xia et al., 2024; Jain et al., 2025; Min et al., 2026). These reference checkpoints fundamentally define the parameter states where candidate utility is judged.

Because existing warmups train on the candidate pool, the reference states evaluating candidates are themselves shaped by pool data. We find that this coupling introduces substantial variability: altering the warmup trajectory while holding candidate pool, target proxy, and scoring rule fixed markedly shifts rankings and subset composition. We term this phenomenon *reference-path dependence*: varying warmup data alone causes top-ranked subset overlap to drop to 7.0% and downstream F1 to fluctuate by over 5 points. Hence, reference trajectories act as active selection priors (Section 2.3).

To address these limitations, we propose Target-Aligned Candidate Selection (TACS), which grounds reference states entirely in the target task while treating the trajectory as an active design variable. Starting from the base model, TACS optimizes a short, capacity-constrained path on a compact target-validation proxy, calibrating its duration before scoring (Figure 1). Candidates are then ranked by forward-only normalized loss reduction along this path. Crucially, decoupling path construction from candidate data enables amortizing a single target path across multiple distinct pools.

Theoretically, a linearized model shows how reference-path uncertainty propagates through an attribution rule into selection variability and downstream risk. Pool warmup can provide useful pool-specific information, but also makes selection inherit variation in the warmup composition; a target-induced reference removes this source of pool-dependent uncertainty. We further characterize how scoring rules impose distinct geometric preferences over candidate distributions and how finite target paths spectrally emphasize faster-learned target directions.

Empirically, controlled experiments reveal behaviors consistent with this analysis: target-path selection produces more target-aligned retraining directions and more strongly filters corrupted candidates, while trajectory-depth ablations show that the useful extent of target adaptation varies substantially across tasks. At scale, across four instruction pools and three downstream tasks, TACS remains

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.

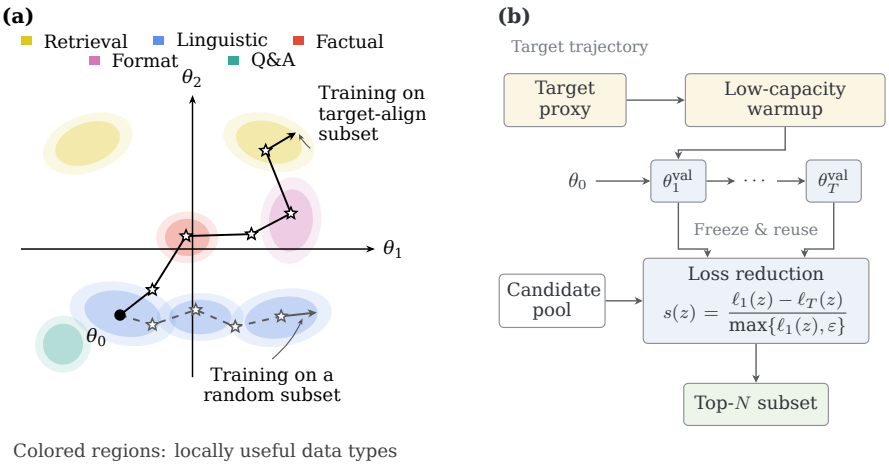


Figure 1: **Reference trajectories determine where candidates are evaluated. (a)** Parameter-space trajectories define the intermediate checkpoints where candidate utility is scored: a generic warmup path can miss task-relevant parameter regions, whereas a target-aligned path guides the model into states where target-critical data signals become visible. **(b)** TACS constructs this reference trajectory directly on a target-validation proxy, enabling forward-only candidate scoring and cross-pool path amortization.

competitive with LESS and ToV while reducing selection GPU-hours by 69% and 53%, respectively. Our main contributions are:

1. **Reference-path dependence**: We identify and quantify the sensitivity of trajectory-based data selection to reference construction, showing that changing the warmup data alone can substantially alter selected subsets and downstream performance (Sec. 2).
2. **TACS framework**: We introduce a target-guided, capacity-constrained reference path that decouples candidate evaluation from pool warmup, enabling forward-only scoring and path reuse across candidate pools (Sec. 3).
3. **Theoretical & empirical analysis**: We characterize how reference uncertainty propagates to selection, analyze the geometric preferences and learning dynamics of target-induced paths, and evaluate TACS across controlled tasks and instruction tuning (Sec. 4, 5).

2 Why the Reference Path Matters

2.1 Problem Setting

Targeted data selection seeks a subset $S \subseteq \mathcal{Z}_{\text{pool}}$ of budget $|S| = N$ from candidate pool $\mathcal{Z}_{\text{pool}}$ to minimize target risk $R(\theta(S)) = \mathbb{E}_{z\sim\mathcal{D}}[\ell(\theta(S); z)]$, where $\theta(S) \in \arg\min_{\theta\in\Theta} R_S(\theta)$ minimizes empirical risk $R_S(\theta) = \frac{1}{|S|}\sum_{z\in S} \ell(\theta; z)$. Because the target distribution $\mathcal{D}$ is inaccessible during selection, an empirical validation set $\mathcal{Z}_{\text{val}} \sim \mathcal{D}$ serves as a proxy via $R_{\text{val}}(\theta) = \frac{1}{|\mathcal{Z}_{\text{val}}|}\sum_{z\in\mathcal{Z}_{\text{val}}} \ell(\theta; z)$, while final generalization is evaluated on disjoint, held-out target test sets.

2.2 Trajectory-Conditioned Selection

For a subset S, consider retraining under gradient flow, $\dot{\theta}_t^S = -\eta_t \nabla R_S(\theta_t^S)$, initialized from a shared base model θ_0. Applying the chain rule to the validation risk yields

$$R_{\text{val}}(\theta_0) - R_{\text{val}}(\theta_T^S) = \frac{1}{|S|}\sum_{z\in S}\int_0^T \eta_t \left\langle \nabla R_{\text{val}}(\theta_t^S), \nabla \ell(\theta_t^S; z)\right\rangle \, \mathrm{d}t =: \frac{1}{|S|}\sum_{z\in S} F(z, \mathcal{T}_S). \quad (1)$$

Minimizing final validation risk thus translates to maximizing cumulative candidate–target gradient alignment along the retraining trajectory $\mathcal{T}_S = \{\theta_t^S\}_{t=0}^T$. In this ideal decomposition, the selected

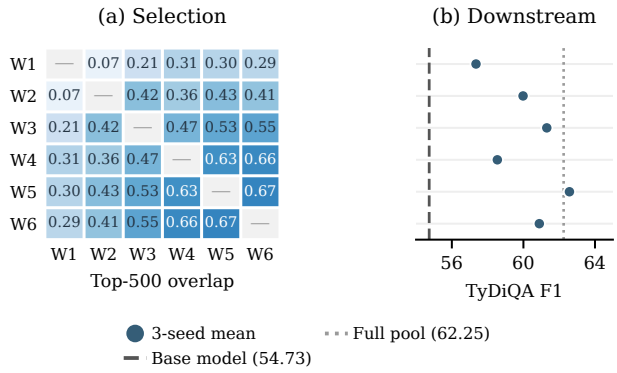


Figure 2: **Reference-path dependence in LESS.** Holding candidate pool, target proxy, and scoring rule fixed, varying only the warmup data subset yields divergent selections. **(a)** Pairwise top-500 candidate overlap. **(b)** Downstream TyDiQA F1 across warmup paths (lines denote base model and full-pool baselines).

subset S simultaneously dictates the examples evaluated and the trajectory along which evaluation occurs.

Because $\mathcal{T}_S$ is unknown prior to selection, practical selectors decouple this dependency by substituting $\mathcal{T}_S$ with an external reference path $\mathcal{T}_{\text{ref}}$. A Riemann sum over discrete checkpoints along $\mathcal{T}_{\text{ref}}$ approximates the integral $F(z, \mathcal{T}_{\text{ref}})$. For instance, LESS accumulates optimizer-adjusted gradient cosine similarities (Xia et al., 2024), while ToV measures localized loss changes induced by validation updates (Jain et al., 2025). Consequently, candidate rankings become explicitly conditioned on $\mathcal{T}_{\text{ref}}$. We refer to this sensitivity as *reference-path dependence*: altering the reference path shifts the evaluation states $\{\theta_t\}$, thereby altering candidate scores even with the pool, target proxy, and scoring rule held constant.

2.3 Observations

Selection is sensitive to the reference path. To quantify this sensitivity, we vary the warmup data across six trajectories in LESS while holding the pool, target proxy, and scoring rule constant. As shown in Figure 2, warmup variation causes severe divergence: pairwise top-500 selection overlap spans 0.070 to 0.674 (median 0.424), while full-pool rank correlations range from 0.100 to 0.843 (median 0.552). This divergence directly impacts downstream utility, with retrained TyDiQA F1 fluctuating between 57.35 and 62.57 (against 62.25 for full-pool training). Thus, reference paths act as active selection priors rather than inert scaffolds, establishing path construction as a critical design choice.

The target distribution as an idealized reference. To formalize the reference choice, consider an idealized setting that selects a training distribution Q from an admissible family $\mathcal{Q}$ (distributions capable of approximating $\mathcal{D}$). For any minimizer $\theta_Q^\star \in \arg\min_\theta \mathbb{E}_{z\sim Q}[\ell(\theta; z)]$, when $\mathcal{D} \in \mathcal{Q}$, direct target training attains the global lower bound:

$$R_{\mathcal{D}}(\theta_{\mathcal{D}}^\star) = \min_\theta R_{\mathcal{D}}(\theta) \le R_{\mathcal{D}}(\theta_Q^\star), \qquad \forall\, Q \in \mathcal{Q}.$$

Under the motivating premise that a sufficiently expressive pool contains an effective target surrogate, adaptation on $\mathcal{D}$ provides a canonical, pool-agnostic reference. This population inequality does not identify an optimal finite-time reference or a finite subset; it motivates leveraging the available target-validation proxy $\mathcal{Z}_{\text{val}}$ to construct target-induced reference trajectories for attribution.

3 Methodology

TACS separates two choices that pool-induced selectors couple: which model states define candidate utility, and which candidates are evaluated at those states. It first constructs a short reference path

using only the target proxy, completely decoupled from the candidate data. It then reuses two checkpoints along that path to score any candidate pool with forward-only evaluations. The complete procedure is outlined in Algorithm 1 (Appendix A).

3.1 Let the Target Define the Reference Path

Starting from base parameters θ_0, TACS takes a short sequence of updates on the target-validation risk:

$$\theta_t^{\text{val}} = \theta_{t-1}^{\text{val}} - \eta \nabla_\theta R_{\text{val}}(\theta_{t-1}^{\text{val}}), \qquad R_{\text{val}}(\theta) = \frac{1}{|\mathcal{Z}_{\text{val}}|} \sum_{z \in \mathcal{Z}_{\text{val}}} \ell(\theta; z). \tag{2}$$

These updates trace the reference path $\mathcal{T}_{\text{val}} = \{\theta_0^{\text{val}}, \ldots, \theta_T^{\text{val}}\}$. Because candidate data is not involved in path construction, a single trajectory can be amortized to score multiple candidate pools for a fixed model–target pair.

The path must traverse far enough to expose target-relevant candidate responses without overfitting the compact proxy. For language models, we regularize the warmup by restricting adaptation to a low-rank LoRA subspace (Hu et al., 2021). We jointly calibrate the learning rate η and horizon T by testing whether updates learned from one portion of the proxy generalize to held-out target instances beyond generic instruction signals. Specifically, across validation splits j, let g_{W_j} denote the endpoint gradient on the warmup split, g_{H_j} on held-out target examples, and g_P on a fixed, task-agnostic reference. We maximize the mean target specificity:

$$C_j(\eta, T) = \cos(g_{W_j}, g_{H_j}) - \cos(g_{W_j}, g_P). \tag{3}$$

The first term rewards target transfer, while the second discounts alignment with generic instruction following. Because g_P is fixed across all tasks and candidate sources, calibration does not reintroduce pool dependence.

3.2 Measure Candidate Response Along the Path

A useful candidate should become easier for the model as it adapts to the downstream task. We capture this response between the first saved checkpoint and the calibrated endpoint:

$$s(z) = \frac{\ell(\theta_1^{\text{val}}; z) - \ell(\theta_T^{\text{val}}; z)}{\max\{\ell(\theta_1^{\text{val}}; z), \varepsilon\}}. \tag{4}$$

We evaluate from θ_1^{val} rather than θ_0^{val} because the initial proxy pass may be dominated by generic, surface-level adaptation such as formatting alignment (Frankle et al., 2020). For token-level objectives, ℓ averages negative log-likelihood across prompt-masked response tokens. Normalizing by the initial loss expresses candidate improvement relative to its initial loss scale. Appendix G.4 examines the resulting selected-example lengths. Given budget N, TACS selects the top-N candidates: $S^* = \arg\max_{S \subseteq \mathcal{Z}_{\text{pool}}, |S|=N} \sum_{z \in S} s(z)$. Candidate scoring relies strictly on forward passes through the two frozen checkpoints, requiring zero candidate-gradient caching or backpropagation.

3.3 Why Loss Reduction Measures Path Alignment

The endpoint score summarizes the cumulative alignment along the trajectory rather than an isolated checkpoint update. Under continuous target-gradient flow, the raw loss reduction equates to:

$$\ell(\theta_1^{\text{val}}; z) - \ell(\theta_T^{\text{val}}; z) = \int_1^T \eta_t \left\langle \nabla \ell(\theta_t^{\text{val}}; z), \nabla R_{\text{val}}(\theta_t^{\text{val}}) \right\rangle \, \mathrm{d}t.$$

A candidate's raw loss decrease thus accumulates alignment with the target adaptation field across the path; normalization additionally weights that decrease by initial loss. With discrete optimizer steps, the endpoint difference measures response to the actual parameter displacements. Section 4 characterizes which candidate distributions an idealized normalized score prefers and how the path horizon filters target modes.

4 THEORETICAL ANALYSIS

To understand the mechanisms behind reference-path dependence, we study two complementary questions. First, how does variation in the warmup path propagate to the data selected by a trajectory-based attribution rule, and ultimately to downstream risk (§4.1)? Second, once the reference path is fixed to the target, what candidate distributions do different scoring rules favor, and what target information is captured along a finite path (§4.2)? We analyze both questions in a population model that replaces finite subsets with an admissible family of training distributions.

Setting. Motivated by linearized analyses of regularized LLM fine-tuning (Afzal et al., 2025), we consider a linearized regression model $f_w(x) = f_0(x) + \phi(x)^\top w$ with positive-definite covariance $H = \mathbb{E}[\phi(x)\phi(x)^\top] \succ 0$. A candidate distribution P_u corresponds to parameter displacement u, with risk $R_u(w) = \frac{1}{2}\|w - u\|_H^2$, where $\|z\|_H^2 = z^\top H z$. The target task P_q starts from initialization $w = 0$, and $D(u) = R_q(u)$ denotes downstream target risk. For a mixture of candidate distributions, let $\mu = \mathbb{E}[u]$ denote its mean optimum. Starting from $w = 0$, continuous gradient flow on this mixture reaches

$$w_t(\mu) = B_t\mu, \qquad B_t = I - e^{-Ht}.$$

4.1 REFERENCE-INDUCED SELECTION UNCERTAINTY

Section 2 showed that changing the warmup data alone can substantially alter selection and downstream performance. Following the distribution-selection formulation of Section 2.3, we model selection over an admissible family $\mathcal{Q} = \{Q_u : u \in \mathcal{U}\}$ and ask how variation in the warmup reference propagates to the selected distribution and its target risk. For a warmup distribution with mean optimum μ, let

$$s(u;\mu) = \mathcal{A}(u; B_{t_1}\mu, \ldots, B_{t_L}\mu, q), \qquad v(\mu) = \arg\max_{u\in\mathcal{U}} s(u;\mu), \tag{5}$$

where $\mathcal{A}$ is any smooth trajectory-based attribution score. Thus $v(\mu)$ is the distribution selected using the reference path generated by μ. Let the warmup mean be random with $\mathbb{E}[\widehat{\mu}] = \mu$ and $\mathrm{Cov}(\widehat{\mu}) = \Omega$. Write $J = D_\mu v(\mu)$ for the sensitivity of the selected distribution to the reference, and $g = J^\top H(v(\mu) - q)$.

Theorem 4.1 (Propagation of reference uncertainty)**.** Suppose s is smooth and $v(\mu)$ is a unique regular maximizer. For sufficiently concentrated warmup variation,

$$\begin{aligned} \frac{1}{2}\mathbb{E}\|v(\widehat{\mu}) - v(\widehat{\mu}')\|_H^2 &= \mathrm{tr}(HJ\Omega J^\top) + o(\|\Omega\|), \\ \mathrm{Var}[D(v(\widehat{\mu}))] &= g^\top \Omega g + o(\|\Omega\|), \end{aligned} \tag{6}$$

where $\widehat{\mu}'$ is an independent warmup. Moreover,

$$D(v(q)) - \mathbb{E}D(v(\widehat{\mu})) = D(v(q)) - D(\mathbb{E}v(\widehat{\mu})) - \frac{1}{2}\mathbb{E}\|v(\widehat{\mu}) - \mathbb{E}v(\widehat{\mu})\|_H^2. \tag{7}$$

Insight. Reference variation is transferred to selection through J. Pool warmup can therefore exploit useful pool-specific information while also introducing selection instability, whose magnitude depends jointly on warmup heterogeneity and the attribution rule's reference sensitivity. Only the target-relevant component of this variation propagates to downstream risk. Equation (7) further shows that pool warmup may improve average performance, but any such gain is offset by its induced dispersion. A fixed target reference removes the pool-composition source of this uncertainty.

4.2 SCORING GEOMETRY IN A LINEARIZED MODEL

We next examine the geometric inductive biases of candidate scoring rules along a fixed target trajectory. Continuous target gradient flow for duration τ reaches endpoint v_τ. For intuition, we decompose any candidate displacement into a target-aligned component and an orthogonal residual: $u = av + b$, where $v = v_\tau \neq 0$, $a \in \mathbb{R}$, and $b^\top Hv = 0$.

Theorem 4.2 (Geometric preferences of candidate scoring rules)**.** Under an inactive normalization floor, score candidates over $[0, v]$ with $u = av + b \neq 0$.

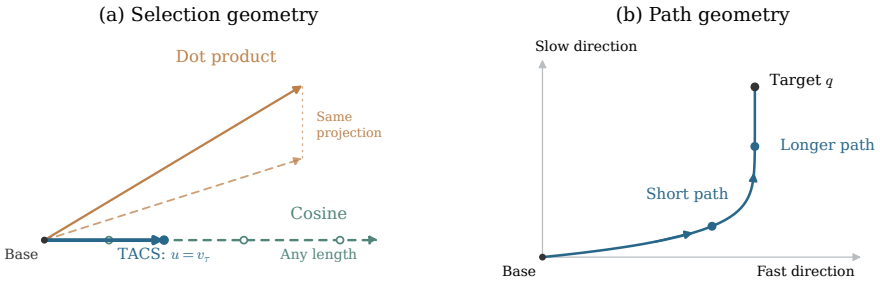


Figure 3: **Geometric and spectral characteristics of the linearized model. (a)** Scoring preferences in $H^{1/2}$ coordinates: raw loss reduction is insensitive to orthogonal components; cosine alignment leaves scale unconstrained; normalized loss reduction uniquely favors $u = v_\tau$. **(b)** Spectral mode evolution under target gradient flow: fast task modes ($\lambda_i \tau \gg 1$) contract rapidly, whereas slow modes are attenuated.

1. *Normalized loss drop (TACS)*: The score $s(u) = 1 - \|u - v\|_H^2 / \|u\|_H^2$ attains a unique global maximum at $u^* = v$, achieving downstream risk $D(u^*) = \frac{1}{2}\|v - q\|_H^2$. For $a > 1/2$, the score strictly penalizes orthogonal variance: $\partial s(u)/\partial \|b\|_H^2 < 0$.
2. *Cosine alignment*: $c(u, v) = \langle u, v \rangle_H / (\|u\|_H \|v\|_H)$ is scale-invariant and maximized by any point along the ray $u = \alpha v$ ($\alpha > 0$).
3. *Dot-product alignment*: $d(u, v) = \langle u, v \rangle_H = a\|v\|_H^2$ scales linearly with a and is invariant to the orthogonal component b. It induces the same candidate ranking as raw loss drop, since $R_u(0) - R_u(v) = d(u, v) - \frac{1}{2}\|v\|_H^2$ for fixed v.

Theorem 4.3 (Spectral bias of a finite target path)**.** Let $He_i = \lambda_i e_i$ with $\lambda_1 \geq \cdots \geq \lambda_d > 0$, and write $q = \sum_i q_i e_i$. Then target gradient flow for duration τ reaches

$$v_\tau = \sum_{i=1}^{d} \big(1 - e^{-\lambda_i \tau}\big) q_i e_i. \tag{8}$$

Thus larger-λ_i target modes are learned earlier, while smaller-λ_i modes remain attenuated along short paths. The relative target risk is $D(v_\tau)/R_q(0) = \sum_i \lambda_i q_i^2 e^{-2\lambda_i \tau} / \sum_i \lambda_i q_i^2$. Moreover, letting $C_q(k) = \sum_{i \leq k} \lambda_i q_i^2 / \|q\|_H^2$, if two targets share H and $C_{q_A}(k) \geq C_{q_B}(k)$ for all k, then $D(v_\tau^A)/R_{q_A}(0) \leq D(v_\tau^B)/R_{q_B}(0)$ for all $\tau > 0$.

Insight. Within this population framework, Theorems 4.2 and 4.3 formalize two complementary inductive biases illustrated in Figure 3. First, *direction and scale control*: raw loss reduction is insensitive to the orthogonal component b, while cosine alignment leaves candidate scale unconstrained. TACS uniquely favors $u = v_\tau$ and, for fixed $a > 1/2$, penalizes larger orthogonal components. Its positive-score superlevel sets are bounded, constraining candidate scale as well. Second, *task-dependent mode learning*: finite target paths capture fast modes while attenuating slow ones. Targets concentrated in fast directions can therefore be captured by short paths, whereas targets with substantial signal in slower directions require deeper adaptation. This provides a population-level explanation for task-dependent trajectory depth. In practice, held-out proxy calibration selects the horizon.

5 Empirical Study

5.1 Controlled Tasks: Mechanistic Validation

Target-induced paths prioritize task-aligned parameter updates. We evaluate TACS on a controlled logistic mixture with known target and distractor components. As shown in Figure 4a, the

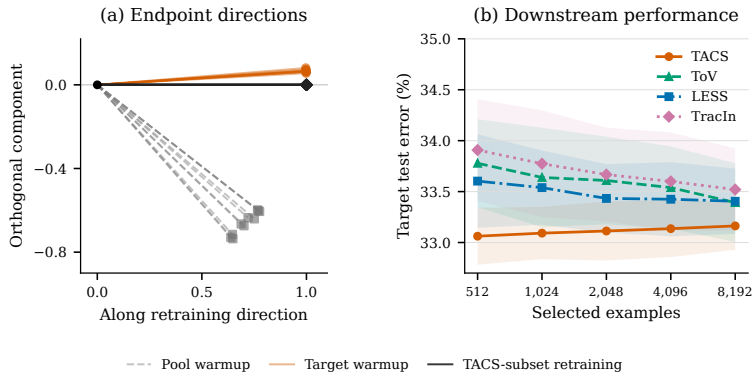


Figure 4: **Evaluation on a synthetic logistic mixture. (a)** Retraining parameter displacements for TACS subsets compared with target and pool warmup directions. **(b)** Downstream target test error across selection budgets (mean ± SD over ten seeds).

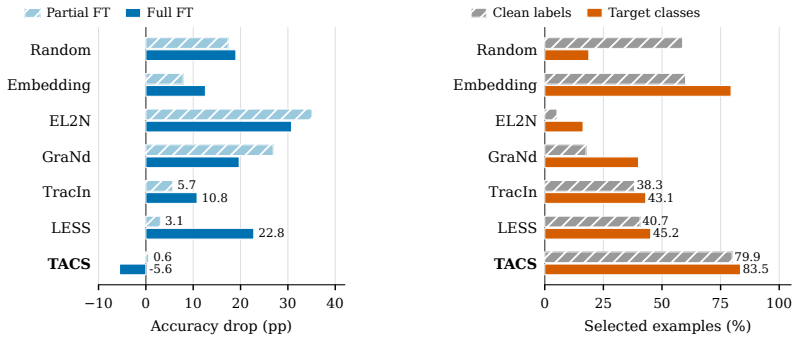


Figure 5: **Robustness under 40% label noise on CIFAR-10 (ResNet-18). (left)** Downstream accuracy drop from clean to noisy candidate pools (lower is better). **(right)** Proportion of clean and target-class examples within selected subsets from noisy candidate pool.

parameter displacement from retraining on a TACS-selected subset aligns substantially closer with the target-warmup path than with pool-warmed trajectories. This alignment yields consistently lower target test error across all selection budgets compared to ToV, LESS, and TracIn (Pruthi et al., 2020) (Figure 4b). This selective behavior is consistent with Theorem 4.2, where normalized loss reduction penalizes orthogonal distractor components, enabling TACS to prioritize target-aligned examples even in rare-target regimes.

Target trajectories filter label noise without pool contamination. We next examine selection robustness under distribution corruption using an ImageNet-pretrained `ResNet-18` on CIFAR-10 with 40% symmetric label noise. Under severe noise, TACS substantially outperforms matched gradient baselines (LESS under full fine-tuning and TracIn under partial finetuning), while matching them on clean pools (Figure 5). Structurally, pool-warmed references incorporate updates from corrupted candidates, whereas TACS keeps candidate-pool variation out of reference construction. This reflects the mechanism in Theorem 4.1: variation entering through a pool-dependent reference can propagate to both selection and downstream risk.

5.2 Instruction Tuning at Scale

We scale TACS to instruction tuning with `Llama-3.2-3B` across MMLU, TyDiQA, and BBH, selecting subsets from four instruction pools (COT, Dolly, Flan V2, and Oasst1) to retrain the base

model. We compare TACS against random selection, representative dynamic selectors (LESS (Xia et al., 2024) and ToV (Jain et al., 2025)), and full-pool training.

Table 1: **Downstream instruction-tuning performance across candidate pools and tasks.** Subset evaluation results (mean ± SD across three seeds; Llama-3.2-3B). Bold indicates the highest mean.

Pool	Task	Random	LESS	ToV	TACS	Full pool
COT	MMLU	54.28 ±0.23	55.15 ±0.53	54.94 ±0.71	**55.57** ±0.42	53.26
	TyDiQA	51.43 ±2.79	**62.63** ±2.60	55.56 ±1.21	62.12 ±1.41	47.45
	BBH	46.23 ±1.44	**48.12** ±0.37	47.44 ±0.81	46.11 ±1.18	43.06
Dolly	MMLU	56.36 ±0.03	**56.51** ±0.14	56.27 ±0.07	56.39 ±0.12	55.90
	TyDiQA	50.60 ±0.88	61.10 ±0.48	60.15 ±0.93	**63.21** ±0.99	56.95
	BBH	48.15 ±0.24	48.40 ±0.05	47.72 ±0.59	**48.46** ±0.96	48.24
Flan V2	MMLU	55.85 ±0.40	54.93 ±0.63	52.18 ±0.70	**56.02** ±0.19	53.75
	TyDiQA	57.47 ±1.92	**59.46** ±1.98	58.59 ±0.69	59.33 ±0.58	62.38
	BBH	47.50 ±0.61	46.20 ±0.42	47.72 ±1.13	**48.18** ±1.40	49.07
Oasst1	MMLU	**56.37** ±0.03	56.16 ±0.57	56.32 ±0.11	**56.37** ±0.06	54.78
	TyDiQA	32.78 ±2.93	55.58 ±1.77	55.52 ±0.68	**58.92** ±0.49	41.09
	BBH	47.72 ±1.02	47.47 ±0.51	48.58 ±0.56	**48.67** ±0.72	48.98
Average	MMLU	55.72 ±0.15	55.69 ±0.25	54.93 ±0.05	**56.08** ±0.10	54.42
	TyDiQA	48.07 ±0.03	59.69 ±0.42	57.45 ±0.13	**60.89** ±0.33	51.97
	BBH	47.40 ±0.50	47.55 ±0.19	**47.86** ±0.42	47.85 ±0.51	47.34
Base model (no training): MMLU 56.22; TyDiQA 42.00; BBH 48.61						

Downstream performance across tasks and pools. Table 1 shows that TACS is competitive across the evaluated pools and tasks, with the clearest average gains on TyDiQA. Two factors may explain these gains: diverse pools offer more target-aligned examples to select from, while tasks like TyDiQA leave more room for improvement through adaptation. The latter interpretation is consistent with the task-dependent learning dynamics in Theorem 4.3.

Compute and storage efficiency via path amortization. Table 2 highlights the efficiency of TACS across storage, compute, and scaling. Forward-only scoring eliminates candidate-gradient caching, reducing the scoring payload from over 35 GB in LESS to under 7 MB. More importantly, the target path is constructed once and reused across candidate pools, so adding a new pool requires only two forward evaluations per candidate. In our benchmark ($m = 4, k = 3$), this yields 53% lower selection compute than ToV and 69% lower than LESS. The complementary trade-off is that LESS amortizes more effectively across many targets for a fixed pool through reusable candidate gradients.

Table 2: **Stage-matched selection cost and scoring payload** on `Llama-3.2-3B` across $m = 4$ pools and $k = 3$ targets. GPU-hours encompass calibration, final reference construction, and candidate scoring under the Appendix G.5 protocol. Downstream retraining is excluded.

Method	Cost Model (GPU h)	Total GPU h	Scoring Storage
LESS	$46.38m + 0.15mk$	187.35	35480 MB
ToV	$1.06m + 10.04mk$	124.74	26.0 MB
TACS	$2.64k + 4.21mk$	58.46 (−69% / −53%)	6.5 **MB**

5.2.1 Ablation Studies

Low-capacity references can guide higher-capacity learners. Figure 6 evaluates LoRA adapter rank during reference construction while downstream retraining uses rank-128 LoRA. Increasing reference capacity yields no downstream gains, with ranks 64 and 128 degrading performance on TyDiQA. Thus, effective targeted selection does not require matching the adaptation capacity of the downstream learner: a rank-one reference can identify useful data for substantially higher-capacity retraining. The resulting reference is also exceptionally lightweight, making fine-grained trajectory checkpointing practical even under modest storage budgets, where retaining many high-rank or full-model states can quickly become costly.

Trajectory depth exhibits task-dependent trends. We find that useful trajectory depth is strongly task-dependent (Figure 6). While performance on MMLU and BBH remains stable across short horizons, TyDiQA benefits from deeper trajectories (improving up to 64 proxy epochs). This dependence persists with capacity-matched target warmup and unnormalized loss scoring (Appendix D). As characterized by Theorem 4.3, deeper target paths progressively expose slower target modes. This cross-task divergence establishes trajectory depth as an important design variable in target-reference construction, motivating the task-adaptive calibration used by TACS.

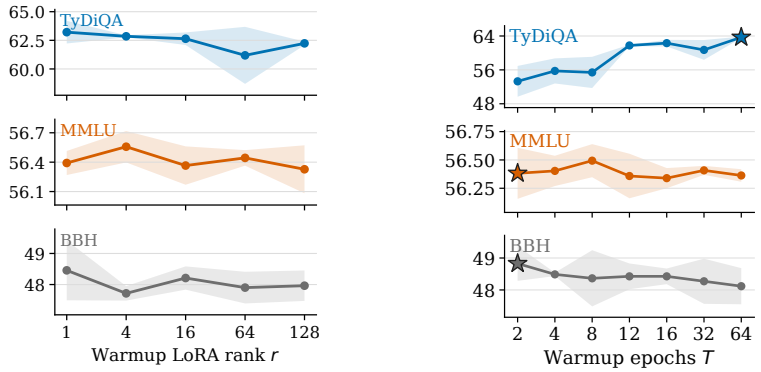


Figure 6: **Ablation of reference capacity and trajectory horizon on Dolly.** Downstream performance (three-seed mean ± SD) as warmup LoRA rank (**left**) and proxy-epoch depth T (**right**) vary. Stars denote calibration-selected depths.

6 Related Work

Targeted data selection identifies training examples aligned with downstream requirements, typically through bilevel optimization (Killamsetty et al., 2021b), gradient matching (Killamsetty et al., 2021a), or dynamic attribution signals computed along training paths (Xia et al., 2024; Jain et al., 2025; Min et al., 2026). More broadly, instance attribution methods estimate training example influence through gradients, inverse-Hessian approximations, or parameter perturbations during adaptation (Koh & Liang, 2017; Pruthi et al., 2020; Park et al., 2023).

Separately, conditional loss-reduction methods score candidates by contrasting losses against an adapted or holdout model (Moore & Lewis, 2010; Mindermann et al., 2022; Brandfonbrener et al., 2024). The unnormalized TACS score shares this loss-drop form. These methods motivate loss reduction as a selection signal, while TACS studies the reference states that define it. When candidate pools differ from the target and supervision is limited to a compact proxy, TACS controls reference capacity and calibrates trajectory depth to the target task. Appendix D provides detailed comparisons and additional depth ablations.

7 Conclusion.

We demonstrated that reference paths in trajectory-based selection serve as active selection priors. To eliminate pool dependence, TACS grounds reference checkpoints in a compact target proxy and scores candidates via forward-only normalized loss reduction. Across diverse benchmarks, TACS achieves competitive downstream utility while drastically reducing selection compute and storage. These benefits come with two practical boundaries.

Limitations. TACS depends on the target proxy faithfully representing the downstream task, as missing behaviors may receive insufficient selection signal. Its efficiency gains also favor settings with multiple candidate pools per target, whereas gradient-caching methods can amortize more effectively when a fixed pool serves many targets (see Appendix B for further discussion).

By establishing reference construction as an explicit design choice, our findings advance principled, resource-efficient data curation.

AI Use Statement

We used generative AI tools to write and test experiment and analysis code, to orchestrate cluster jobs, and to edit the manuscript text. AI-assisted code was tested before use, reported numbers are produced by scripts whose outputs we checked, and the authors reviewed all AI-assisted work. We also used these tools to assist with proof drafting and checking in the theory part. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

Ethics Statement

This work studies data selection for adapting pretrained models to a target task using public datasets, and releases no new pretrained model or scraped corpus. Appendix C discusses how selection inherits the values of the validation proxy and how auditing selected subsets mitigates this.

Reproducibility Statement

Section 3 and Algorithm 1 specify the method, and Appendix G.6 gives the calibration protocol. Appendix G lists datasets, models, retraining settings, evaluation protocols, and seeds for every experiment. Complete proofs of the formal results are in Appendices F and I.

References

Zahra Rahimi Afzal, Tara Esmaeilbeig, Mojtaba Soltanalian, and Mesrob I Ohannessian. Linearization explains fine-tuning in large language models. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=tdwRIP6NG2.

David Brandfonbrener, Hanlin Zhang, Andreas Kirsch, Jonathan Richard Schwarz, and Sham M. Kakade. Color-filter: Conditional loss reduction filtering for targeted language model pre-training. In *The Thirty-eighth Annual Conference on Neural Information Processing Systems*, 2024. URL https://openreview.net/forum?id=GUccmOMBv6.

Abdulkadir Canatar, Blake Bordelon, and Cengiz Pehlevan. Spectral bias and task-model alignment explain generalization in kernel regression and infinitely wide neural networks. *Nature Communications*, 12(1):2914, 2021. doi: 10.1038/s41467-021-23103-1. URL https://doi.org/10.1038/s41467-021-23103-1.

Jonathan H. Clark, Eunsol Choi, Michael Collins, Dan Garrette, Tom Kwiatkowski, Vitaly Nikolaev, and Jennimaria Palomaki. TyDi QA: A benchmark for information-seeking question answering in typologically diverse languages. *Transactions of the Association for Computational Linguistics*, 8:454–470, 2020. doi: 10.1162/tacl_a_00317. URL https://aclanthology.org/2020.tacl-1.30/.

Mike Conover, Matt Hayes, Ankit Mathur, Jianwei Xie, Jun Wan, Sam Shah, Ali Ghodsi, Patrick Wendell, Matei Zaharia, and Reynold Xin. Free dolly: Introducing the world's first truly open instruction-tuned llm, 2023. URL https://www.databricks.com/blog/2023/04/12/dolly-first-open-commercially-viable-instruction-tuned-llm.

Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, 2009. doi: 10.1109/CVPR.2009.5206848.

Jonathan Frankle, David J. Schwab, and Ari S. Morcos. The early phase of neural network training. In *International Conference on Learning Representations*, 2020. URL https://openreview.net/forum?id=Hkl1iRNFwS.

Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Alex Vaughan, Amy Yang, Angela Fan, Anirudh Goyal, Anthony Hartshorn, Aobo Yang, Archi Mitra, Archie Sravankumar, Artem Korenev, Arthur Hinsvark, Arun Rao, Aston Zhang, Aurelien Rodriguez, Austen Gregerson, Ava Spataru,

Baptiste Roziere, Bethany Biron, Binh Tang, Bobbie Chern, Charlotte Caucheteux, Chaya Nayak, Chloe Bi, Chris Marra, Chris McConnell, Christian Keller, Christophe Touret, Chunyang Wu, Corinne Wong, Cristian Canton Ferrer, Cyrus Nikolaidis, Damien Allonsius, Daniel Song, Danielle Pintz, Danny Livshits, Danny Wyatt, David Esiobu, Dhruv Choudhary, Dhruv Mahajan, Diego Garcia-Olano, Diego Perino, Dieuwke Hupkes, Egor Lakomkin, Ehab AlBadawy, Elina Lobanova, Emily Dinan, Eric Michael Smith, Filip Radenovic, Francisco Guzmán, Frank Zhang, Gabriel Synnaeve, Gabrielle Lee, Georgia Lewis Anderson, Govind Thattai, Graeme Nail, Gregoire Mialon, Guan Pang, Guillem Cucurell, Hailey Nguyen, Hannah Korevaar, Hu Xu, Hugo Touvron, Iliyan Zarov, Imanol Arrieta Ibarra, Isabel Kloumann, Ishan Misra, Ivan Evtimov, Jack Zhang, Jade Copet, Jaewon Lee, Jan Geffert, Jana Vranes, Jason Park, Jay Mahadeokar, Jeet Shah, Jelmer van der Linde, Jennifer Billock, Jenny Hong, Jenya Lee, Jeremy Fu, Jianfeng Chi, Jianyu Huang, Jiawen Liu, Jie Wang, Jiecao Yu, Joanna Bitton, Joe Spisak, Jongsoo Park, Joseph Rocca, Joshua Johnstun, Joshua Saxe, Junteng Jia, Kalyan Vasuden Alwala, Karthik Prasad, Kartikeya Upasani, Kate Plawiak, Ke Li, Kenneth Heafield, Kevin Stone, Khalid El-Arini, Krithika Iyer, Kshitiz Malik, Kuenley Chiu, Kunal Bhalla, Kushal Lakhotia, Lauren Rantala-Yeary, Laurens van der Maaten, Lawrence Chen, Liang Tan, Liz Jenkins, Louis Martin, Lovish Madaan, Lubo Malo, Lukas Blecher, Lukas Landzaat, Luke de Oliveira, Madeline Muzzi, Mahesh Pasupuleti, Mannat Singh, Manohar Paluri, Marcin Kardas, Maria Tsimpoukelli, Mathew Oldham, Mathieu Rita, Maya Pavlova, Melanie Kambadur, Mike Lewis, Min Si, Mitesh Kumar Singh, Mona Hassan, Naman Goyal, Narjes Torabi, Nikolay Bashlykov, Nikolay Bogoychev, Niladri Chatterji, Ning Zhang, Olivier Duchenne, Onur Çelebi, Patrick Alrassy, Pengchuan Zhang, Pengwei Li, Petar Vasic, Peter Weng, Prajjwal Bhargava, Pratik Dubal, Praveen Krishnan, Punit Singh Koura, Puxin Xu, Qing He, Qingxiao Dong, Ragavan Srinivasan, Raj Ganapathy, Ramon Calderer, Ricardo Silveira Cabral, Robert Stojnic, Roberta Raileanu, Rohan Maheswari, Rohit Girdhar, Rohit Patel, Romain Sauvestre, Ronnie Polidoro, Roshan Sumbaly, Ross Taylor, Ruan Silva, Rui Hou, Rui Wang, Saghar Hosseini, Sahana Chennabasappa, Sanjay Singh, Sean Bell, Seohyun Sonia Kim, Sergey Edunov, Shaoliang Nie, Sharan Narang, Sharath Raparthy, Sheng Shen, Shengye Wan, Shruti Bhosale, Shun Zhang, Simon Vandenhende, Soumya Batra, Spencer Whitman, Sten Sootla, Stephane Collot, Suchin Gururangan, Sydney Borodinsky, Tamar Herman, Tara Fowler, Tarek Sheasha, Thomas Georgiou, Thomas Scialom, Tobias Speckbacher, Todor Mihaylov, Tong Xiao, Ujjwal Karn, Vedanuj Goswami, Vibhor Gupta, Vignesh Ramanathan, Viktor Kerkez, Vincent Gonguet, Virginie Do, Vish Vogeti, Vítor Albiero, Vladan Petrovic, Weiwei Chu, Wenhan Xiong, Wenyin Fu, Whitney Meers, Xavier Martinet, Xiaodong Wang, Xiaofang Wang, Xiaoqing Ellen Tan, Xide Xia, Xinfeng Xie, Xuchao Jia, Xuewei Wang, Yaelle Goldschlag, Yashesh Gaur, Yasmine Babaei, Yi Wen, Yiwen Song, Yuchen Zhang, Yue Li, Yuning Mao, Zacharie Delpierre Coudert, Zheng Yan, Zhengxing Chen, Zoe Papakipos, Aaditya Singh, Aayushi Srivastava, Abha Jain, Adam Kelsey, Adam Shajnfeld, Adithya Gangidi, Adolfo Victoria, Ahuva Goldstand, Ajay Menon, Ajay Sharma, Alex Boesenberg, Alexei Baevski, Allie Feinstein, Amanda Kallet, Amit Sangani, Amos Teo, Anam Yunus, Andrei Lupu, Andres Alvarado, Andrew Caples, Andrew Gu, Andrew Ho, Andrew Poulton, Andrew Ryan, Ankit Ramchandani, Annie Dong, Annie Franco, Anuj Goyal, Aparajita Saraf, Arkabandhu Chowdhury, Ashley Gabriel, Ashwin Bharambe, Assaf Eisenman, Azadeh Yazdan, Beau James, Ben Maurer, Benjamin Leonhardi, Bernie Huang, Beth Loyd, Beto De Paola, Bhargavi Paranjape, Bing Liu, Bo Wu, Boyu Ni, Braden Hancock, Bram Wasti, Brandon Spence, Brani Stojkovic, Brian Gamido, Britt Montalvo, Carl Parker, Carly Burton, Catalina Mejia, Ce Liu, Changhan Wang, Changkyu Kim, Chao Zhou, Chester Hu, Ching-Hsiang Chu, Chris Cai, Chris Tindal, Christoph Feichtenhofer, Cynthia Gao, Damon Civin, Dana Beaty, Daniel Kreymer, Daniel Li, David Adkins, David Xu, Davide Testuggine, Delia David, Devi Parikh, Diana Liskovich, Didem Foss, Dingkang Wang, Duc Le, Dustin Holland, Edward Dowling, Eissa Jamil, Elaine Montgomery, Eleonora Presani, Emily Hahn, Emily Wood, Eric-Tuan Le, Erik Brinkman, Esteban Arcaute, Evan Dunbar, Evan Smothers, Fei Sun, Felix Kreuk, Feng Tian, Filippos Kokkinos, Firat Ozgenel, Francesco Caggioni, Frank Kanayet, Frank Seide, Gabriela Medina Florez, Gabriella Schwarz, Gada Badeer, Georgia Swee, Gil Halpern, Grant Herman, Grigory Sizov, Guangyi, Zhang, Guna Lakshminarayanan, Hakan Inan, Hamid Shojanazeri, Han Zou, Hannah Wang, Hanwen Zha, Haroun Habeeb, Harrison Rudolph, Helen Suk, Henry Aspegren, Hunter Goldman, Hongyuan Zhan, Ibrahim Damlaj, Igor Molybog, Igor Tufanov, Ilias Leontiadis, Irina-Elena Veliche, Itai Gat, Jake Weissman, James Geboski, James Kohli, Janice Lam, Japhet Asher, Jean-Baptiste Gaya, Jeff Marcus, Jeff Tang, Jennifer Chan, Jenny Zhen, Jeremy Reizenstein, Jeremy Teboul, Jessica Zhong, Jian Jin, Jingyi Yang, Joe Cummings, Jon Carvill, Jon Shepard, Jonathan McPhie, Jonathan Torres, Josh Ginsburg, Junjie Wang, Kai Wu, Kam Hou U, Karan Saxena, Kartikay Khandelwal, Katayoun Zand, Kathy Matosich, Kaushik

Veeraraghavan, Kelly Michelena, Keqian Li, Kiran Jagadeesh, Kun Huang, Kunal Chawla, Kyle Huang, Lailin Chen, Lakshya Garg, Lavender A, Leandro Silva, Lee Bell, Lei Zhang, Liangpeng Guo, Licheng Yu, Liron Moshkovich, Luca Wehrstedt, Madian Khabsa, Manav Avalani, Manish Bhatt, Martynas Mankus, Matan Hasson, Matthew Lennie, Matthias Reso, Maxim Groshev, Maxim Naumov, Maya Lathi, Meghan Keneally, Miao Liu, Michael L. Seltzer, Michal Valko, Michelle Restrepo, Mihir Patel, Mik Vyatskov, Mikayel Samvelyan, Mike Clark, Mike Macey, Mike Wang, Miquel Jubert Hermoso, Mo Metanat, Mohammad Rastegari, Munish Bansal, Nandhini Santhanam, Natascha Parks, Natasha White, Navyata Bawa, Nayan Singhal, Nick Egebo, Nicolas Usunier, Nikhil Mehta, Nikolay Pavlovich Laptev, Ning Dong, Norman Cheng, Oleg Chernoguz, Olivia Hart, Omkar Salpekar, Ozlem Kalinli, Parkin Kent, Parth Parekh, Paul Saab, Pavan Balaji, Pedro Rittner, Philip Bontrager, Pierre Roux, Piotr Dollar, Polina Zvyagina, Prashant Ratanchandani, Pritish Yuvraj, Qian Liang, Rachad Alao, Rachel Rodriguez, Rafi Ayub, Raghotham Murthy, Raghu Nayani, Rahul Mitra, Rangaprabhu Parthasarathy, Raymond Li, Rebekkah Hogan, Robin Battey, Rocky Wang, Russ Howes, Ruty Rinott, Sachin Mehta, Sachin Siby, Sai Jayesh Bondu, Samyak Datta, Sara Chugh, Sara Hunt, Sargun Dhillon, Sasha Sidorov, Satadru Pan, Saurabh Mahajan, Saurabh Verma, Seiji Yamamoto, Sharadh Ramaswamy, Shaun Lindsay, Shaun Lindsay, Sheng Feng, Shenghao Lin, Shengxin Cindy Zha, Shishir Patil, Shiva Shankar, Shuqiang Zhang, Shuqiang Zhang, Sinong Wang, Sneha Agarwal, Soji Sajuyigbe, Soumith Chintala, Stephanie Max, Stephen Chen, Steve Kehoe, Steve Satterfield, Sudarshan Govindaprasad, Sumit Gupta, Summer Deng, Sungmin Cho, Sunny Virk, Suraj Subramanian, Sy Choudhury, Sydney Goldman, Tal Remez, Tamar Glaser, Tamara Best, Thilo Koehler, Thomas Robinson, Tianhe Li, Tianjun Zhang, Tim Matthews, Timothy Chou, Tzook Shaked, Varun Vontimitta, Victoria Ajayi, Victoria Montanez, Vijai Mohan, Vinay Satish Kumar, Vishal Mangla, Vlad Ionescu, Vlad Poenaru, Vlad Tiberiu Mihailescu, Vladimir Ivanov, Wei Li, Wenchen Wang, Wenwen Jiang, Wes Bouaziz, Will Constable, Xiaocheng Tang, Xiaojian Wu, Xiaolan Wang, Xilun Wu, Xinbo Gao, Yaniv Kleinman, Yanjun Chen, Ye Hu, Ye Jia, Ye Qi, Yenda Li, Yilin Zhang, Ying Zhang, Yossi Adi, Youngjin Nam, Yu, Wang, Yu Zhao, Yuchen Hao, Yundi Qian, Yunlu Li, Yuzi He, Zach Rait, Zachary DeVito, Zef Rosnbrick, Zhaoduo Wen, Zhenyu Yang, Zhiwei Zhao, and Zhiyu Ma. The llama 3 herd of models, 2024. URL https://arxiv.org/abs/2407.21783.

Ernst Hairer, Syvert Norsett, and Gerhard Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*, volume 8. 01 1993. ISBN 978-3-540-56670-0. doi: 10.1007/978-3-540-78862-1.

Jie Hao, Rui Yu, Wei Zhang, Huixia Wang, Jie Xu, and Mingrui Liu. BLISS: A lightweight bilevel influence scoring method for data selection in language model pretraining. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=cFWgTYm1hW.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016. doi: 10.1109/CVPR.2016.90.

Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. Measuring massive multitask language understanding, 2021. URL https://arxiv.org/abs/2009.03300.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. Lora: Low-rank adaptation of large language models, 2021. URL https://arxiv.org/abs/2106.09685.

Ayush Jain, Andrea Montanari, and Eren Sasoglu. Train on validation (tov): Fast data selection with applications to fine-tuning, 2025. URL https://arxiv.org/abs/2510.00386.

Krishnateja Killamsetty, Durga Sivasubramanian, Ganesh Ramakrishnan, Abir De, and Rishabh Iyer. Grad-match: Gradient matching based data subset selection for efficient deep model training, 2021a. URL https://arxiv.org/abs/2103.00123.

Krishnateja Killamsetty, Durga Sivasubramanian, Ganesh Ramakrishnan, and Rishabh Iyer. Glister: Generalization based data subset selection for efficient and robust learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35:8110–8118, 05 2021b. doi: 10.1609/aaai.v35i9.16988.

Pang Wei Koh and Percy Liang. Understanding black-box predictions via influence functions. In Doina Precup and Yee Whye Teh (eds.), *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1885–1894. PMLR, 06–11 Aug 2017. URL https://proceedings.mlr.press/v70/koh17a.html.

Alex Krizhevsky and Geoffrey Hinton. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009.

Andreas Köpf, Yannic Kilcher, Dimitri von Rütte, Sotiris Anagnostidis, Zhi-Rui Tam, Keith Stevens, Abdullah Barhoum, Nguyen Minh Duc, Oliver Stanley, Richárd Nagyfi, Shahul ES, Sameer Suri, David Glushkov, Arnav Dantuluri, Andrew Maguire, Christoph Schuhmann, Huu Nguyen, and Alexander Mattick. Openassistant conversations – democratizing large language model alignment, 2023. URL https://arxiv.org/abs/2304.07327.

Shayne Longpre, Le Hou, Tu Vu, Albert Webson, Hyung Won Chung, Yi Tay, Denny Zhou, Quoc V Le, Barret Zoph, Jason Wei, and Adam Roberts. The flan collection: Designing data and methods for effective instruction tuning. In Andreas Krause, Emma Brunskill, Kyunghyun Cho, Barbara Engelhardt, Sivan Sabato, and Jonathan Scarlett (eds.), *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 22631–22648. PMLR, 23–29 Jul 2023. URL https://proceedings.mlr.press/v202/longpre23a.html.

Guanghui Min, Tianhao Huang, Ke Wan, and Chen Chen. GIST: Targeted data selection for instruction tuning via coupled optimization geometry. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=K2KeBuKAtu.

Sören Mindermann, Jan M Brauner, Muhammed T Razzak, Mrinank Sharma, Andreas Kirsch, Winnie Xu, Benedikt Höltgen, Aidan N Gomez, Adrien Morisot, Sebastian Farquhar, and Yarin Gal. Prioritized training on points that are learnable, worth learning, and not yet learnt. In Kamalika Chaudhuri, Stefanie Jegelka, Le Song, Csaba Szepesvari, Gang Niu, and Sivan Sabato (eds.), *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 15630–15649. PMLR, 17–23 Jul 2022. URL https://proceedings.mlr.press/v162/mindermann22a.html.

Baharan Mirzasoleiman, Jeff Bilmes, and Jure Leskovec. Coresets for data-efficient training of machine learning models. In *Proceedings of the 37th International Conference on Machine Learning*, ICML'20. JMLR.org, 2020.

Robert C. Moore and William Lewis. Intelligent selection of language model training data. In Jan Hajič, Sandra Carberry, Stephen Clark, and Joakim Nivre (eds.), *Proceedings of the ACL 2010 Conference Short Papers*, pp. 220–224, Uppsala, Sweden, July 2010. Association for Computational Linguistics. URL https://aclanthology.org/P10-2041/.

Sung Min Park, Kristian Georgiev, Andrew Ilyas, Guillaume Leclerc, and Aleksander Mądry. Trak: attributing model behavior at scale. In *Proceedings of the 40th International Conference on Machine Learning*, ICML'23. JMLR.org, 2023.

Mansheej Paul, Surya Ganguli, and Gintare Karolina Dziugaite. Deep learning on a data diet: finding important examples early in training. In *Proceedings of the 35th International Conference on Neural Information Processing Systems*, NIPS '21, Red Hook, NY, USA, 2021. Curran Associates Inc. ISBN 9781713845393.

Geoff Pleiss, Tianyi Zhang, Ethan Elenberg, and Kilian Q. Weinberger. Identifying mislabeled data using the area under the margin ranking. In *Proceedings of the 34th International Conference on Neural Information Processing Systems*, NIPS '20, Red Hook, NY, USA, 2020. Curran Associates Inc. ISBN 9781713829546.

Garima Pruthi, Frederick Liu, Satyen Kale, and Mukund Sundararajan. Estimating training data influence by tracing gradient descent. In H. Larochelle, M. Ranzato, R. Hadsell, M.F. Balcan, and H. Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 19920–19930. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper_files/paper/2020/file/e6385d39ec9394f2f3a354d9d2b88eec-Paper.pdf.

Shivalika Singh, Freddie Vargus, Daniel D'souza, Börje F. Karlsson, Abinaya Mahendiran, Wei-Yin Ko, Herumb Shandilya, Jay Patel, Deividas Mataciunas, Laura O'Mahony, Mike Zhang, Ramith Hettiarachchi, Joseph Wilson, Marina Machado, Luisa Moura, Dominik Krzemiński, Hakimeh Fadaei, Irem Ergun, Ifeoma Okoh, Aisha Alaagib, Oshan Mudannayake, Zaid Alyafeai, Vu Chien, Sebastian Ruder, Surya Guthikonda, Emad Alghamdi, Sebastian Gehrmann, Niklas Muennighoff, Max Bartolo, Julia Kreutzer, Ahmet Üstün, Marzieh Fadaee, and Sara Hooker. Aya dataset: An open-access collection for multilingual instruction tuning. In Lun-Wei Ku, Andre Martins, and Vivek Srikumar (eds.), *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 11521–11567, Bangkok, Thailand, August 2024. Association for Computational Linguistics. doi: 10.18653/v1/2024.acl-long.620. URL https://aclanthology.org/2024.acl-long.620/.

Mirac Suzgun, Nathan Scales, Nathanael Schärli, Sebastian Gehrmann, Yi Tay, Hyung Won Chung, Aakanksha Chowdhery, Quoc Le, Ed Chi, Denny Zhou, and Jason Wei. Challenging BIG-bench tasks and whether chain-of-thought can solve them. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki (eds.), *Findings of the Association for Computational Linguistics: ACL 2023*, pp. 13003–13051, Toronto, Canada, July 2023. Association for Computational Linguistics. doi: 10.18653/v1/2023.findings-acl.824. URL https://aclanthology.org/2023.findings-acl.824/.

Swabha Swayamdipta, Roy Schwartz, Nicholas Lourie, Yizhong Wang, Hannaneh Hajishirzi, Noah A. Smith, and Yejin Choi. Dataset cartography: Mapping and diagnosing datasets with training dynamics. In Bonnie Webber, Trevor Cohn, Yulan He, and Yang Liu (eds.), *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 9275–9293, Online, November 2020. Association for Computational Linguistics. doi: 10.18653/v1/2020.emnlp-main.746. URL https://aclanthology.org/2020.emnlp-main.746/.

Mariya Toneva, Alessandro Sordoni, Remi Tachet des Combes, Adam Trischler, Yoshua Bengio, and Geoffrey J. Gordon. An empirical study of example forgetting during deep neural network learning, 2019. URL https://arxiv.org/abs/1812.05159.

Fangxin Wang, Peyman Baghershahi, Langzhou He, Henry Peng Zou, Sourav Medya, and Philip S. Yu. Filter-then-weight: Online data selection and reweighting for llm fine-tuning, 2026a. URL https://arxiv.org/abs/2604.00001.

Shaobo Wang, Xuan Ouyang, Tianyi Xu, Yuzheng Hu, Jialin Liu, Guo Chen, Tianyu Zhang, Junhao Zheng, Kexin Yang, Xingzhang Ren, Dayiheng Liu, and Linfeng Zhang. OPUS: Towards efficient and principled data selection in large language model pre-training in every iteration. In *Forty-third International Conference on Machine Learning*, 2026b. URL https://openreview.net/forum?id=lcLxGrk5N9.

Yizhong Wang, Hamish Ivison, Pradeep Dasigi, Jack Hessel, Tushar Khot, Khyathi Raghavi Chandu, David Wadden, Kelsey MacMillan, Noah A. Smith, Iz Beltagy, and Hannaneh Hajishirzi. How far can camels go? exploring the state of instruction tuning on open resources. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA, 2023. Curran Associates Inc.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H. Chi, Quoc V. Le, and Denny Zhou. Chain-of-thought prompting elicits reasoning in large language models. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, Red Hook, NY, USA, 2022. Curran Associates Inc. ISBN 9781713871088.

Mengzhou Xia, Sadhika Malladi, Suchin Gururangan, Sanjeev Arora, and Danqi Chen. Less: selecting influential data for targeted instruction tuning. In *Proceedings of the 41st International Conference on Machine Learning*, ICML'24. JMLR.org, 2024.

Ningyuan Yang, Weihua Du, Weiwei Sun, Sean Welleck, and Yiming Yang. Gradalign: Gradient-aligned data selection for llm reinforcement learning, 2026. URL https://arxiv.org/abs/2602.21492.

A ALGORITHM

Algorithm 1 summarizes the procedure of Section 3.

Algorithm 1 **T**arget-**A**ligned **C**andidate **S**election (TACS).

1: **Input:** Val set $\mathcal{Z}_{\text{val}}$, Pool $\mathcal{Z}_{\text{pool}}$, Budget N
2: **Initialize:** Base parameters θ_0
3: Select learning rate η and epochs T via cross-validation (App. G.6)
4: Generate target-aligned warmup trajectory on $\mathcal{Z}_{\text{val}}$, $\mathcal{T}_{\text{val}} = \{\theta_0^{\text{val}}, ..., \theta_T^{\text{val}}\}$, via Eq. 2
5: **for** each candidate $z \in \mathcal{Z}_{\text{pool}}$ **do**
6: Compute normalized candidate score $s(z)$ via Eq. 4
7: **end for**
8: **return** Subset $S \subset \mathcal{Z}_{\text{pool}}$ maximizing total score

B LIMITATIONS

Dependence on the validation proxy. TACS guides its reference trajectory using the validation proxy $\mathcal{Z}_{\text{val}}$; consequently, candidate scoring reflects the task aspects captured within this proxy. If a proxy is unrepresentative or omits key facets of the target distribution, candidates addressing the missing behaviors may receive suboptimal scores. Effective proxy curation therefore remains a natural complement to target-aligned selection.

Capacity regularization and calibration. Restricting the reference path to a low-rank LoRA subspace acts as an explicit capacity constraint, moderating rapid memorization of small proxies. However, adaptation duration and step size must still be calibrated to prevent proxy-specific over-adaptation. While cross-validation (Appendix G.6) provides an automated selection rule, calibration over very compact proxies (such as the nine-example TyDiQA proxy) involves few held-out instances per fold, which can increase sensitivity to proxy partition choices.

Scope of the theoretical analysis. The theoretical derivations in Section 4 analyze candidate preferences and reference perturbations under continuous gradient flow and linearized population dynamics. In practice, the implementation uses discrete AdamW updates within a low-rank parameterization. Furthermore, our analysis characterizes the instantaneous preference rule for individual candidate responses to a finite warmup endpoint; it does not model post-selection retraining dynamics or guarantee the subset that globally minimizes downstream target risk.

Amortization trade-offs across pools and targets. TACS offers substantial efficiency when a fixed target trajectory is amortized to screen multiple candidate pools. However, this efficiency profile introduces an operational tradeoff: adapting to each new downstream target requires constructing a new reference path, whereas gradient-caching selectors (such as LESS) can amortize precomputed candidate gradients across multiple target proxies. TACS is therefore most advantageous when curating diverse candidate sources for a focused set of downstream tasks.

C BROADER IMPACTS

Targeted data selection substantially reduces the computational overhead and environmental footprint associated with fine-tuning large-scale foundation models by identifying high-utility candidate subsets. Furthermore, scoring candidates against explicit reference checkpoints provides an interpretable audit trail of data attribution for specific downstream behaviors.

However, because the selection trajectory is guided directly by the target validation proxy $\mathcal{Z}_{\text{val}}$, the curated data distribution necessarily reflects the inductive biases and distributional properties of this proxy. If $\mathcal{Z}_{\text{val}}$ encodes societal biases, skewed demographic representations, or toxic content, the selection mechanism may preferentially retain candidates that reinforce these undesirable traits. Moreover, the efficiency gains offered by reusable reference paths could lower the technical barrier for adapting models toward malicious or dual-use tasks. Responsible deployment therefore requires

rigorous provenance documentation for validation proxies and thorough data auditing of selected subsets prior to full downstream retraining.

D Extended Related Work

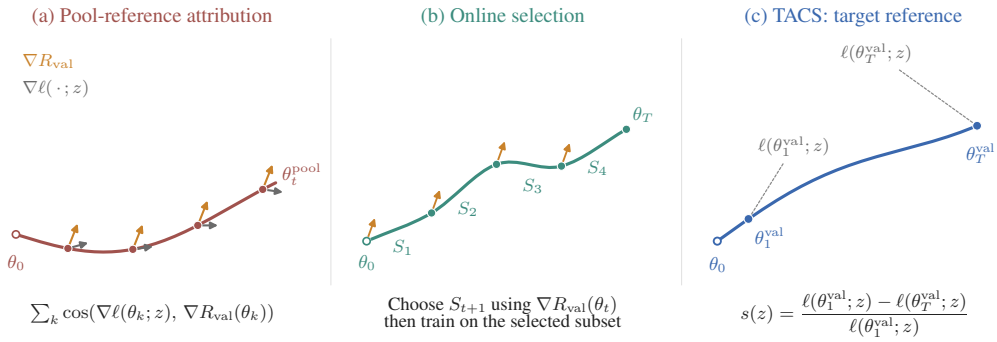


Figure 7: **Taxonomy of training dynamics in targeted data selection. (a)** Pool-reference attribution warms up on candidate data and measures gradient alignment with the target proxy across checkpoints. **(b)** Online selection iteratively alternates subset filtering and model retraining. **(c)** TACS constructs an early, regularized target-proxy trajectory and evaluates candidates via normalized loss response along this path. Paths are illustrative.

Pool-reference attribution. A prominent family of dynamic selectors evaluates candidate utility at parameter states visited during a candidate-pool warmup. LESS (Xia et al., 2024) calculates projected cosine gradient similarities across intermediate checkpoints of a pool-trained model. ToV (Jain et al., 2025) evaluates candidates via empirical loss improvements induced by one-step target perturbations around pool checkpoints, while GIST (Min et al., 2026) projects gradients into a low-rank target subspace. Although varying in their local scoring metrics, these methods share a foundational dependency: the reference parameter states where candidates are judged are fundamentally dictated by the candidate pool itself. As established in Section 2.3 and Proposition F.3, reference construction can couple candidate scoring to pool statistics. TACS departs fundamentally from this paradigm by constructing the reference trajectory exclusively from the target proxy (Figure 7c), entirely decoupling reference states from candidate data.

Online selection and coreset optimization. Coreset and online selection algorithms optimize subsets during the actual training process. CRAIG (Mirzasoleiman et al., 2020), GradMatch (Killamsetty et al., 2021a), and GLISTER (Killamsetty et al., 2021b) iteratively match batch gradients to full-pool or validation objectives. Dynamic bilevel selectors (Wang et al., 2026a;b; Hao et al., 2026) continuously re-estimate candidate utility as training proceeds, and GradAlign (Yang et al., 2026) enforces directional alignment during reinforcement learning fine-tuning. While effective, online methods incur substantial retraining overhead because data selection is tightly interleaved with model optimization. In contrast, TACS performs a single, forward-only selection pass from a static target trajectory prior to retraining, decoupling selection from downstream model execution.

Loss-based filtering versus trajectory-guided attribution. Static loss heuristics identify noisy or outlier instances (Swayamdipta et al., 2020; Toneva et al., 2019; Pleiss et al., 2020; Paul et al., 2021) without downstream task awareness. In contrast, conditional loss-difference methods evaluate candidate utility via an adapted or holdout model: Moore–Lewis selection (Moore & Lewis, 2010) contrasts domain-specific against general language models, RHO-LOSS (Mindermann et al., 2022) uses an independently trained holdout model, and CoLoR-Filter (Brandfonbrener et al., 2024) measures the unnormalized loss drop from downstream adaptation. Motivated by this broad principle, TACS connects closely to conditional loss filtering: without scale normalization, its endpoint score reduces to an unnormalized target-induced loss drop.

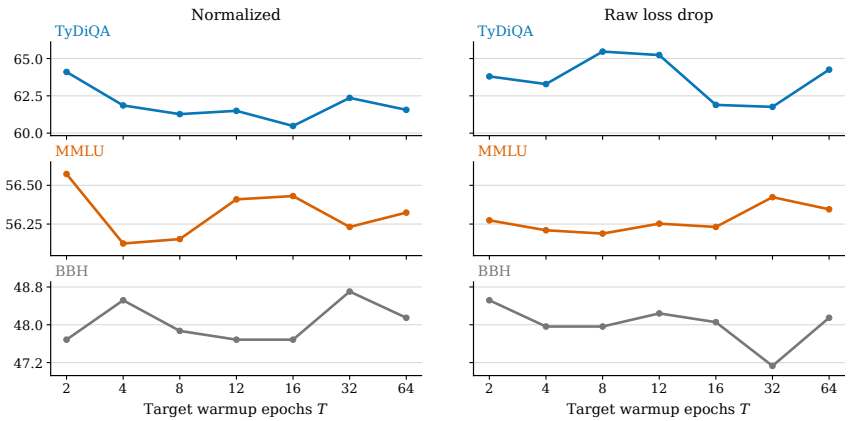


Figure 8: **Reference-depth sensitivity with capacity-matched warmup.** Target warmup uses the same model and LoRA configuration as downstream retraining. Curves show downstream performance of subsets selected using normalized (left) and raw (right) endpoint loss reduction across target warmup depths.

Our targeted instruction-tuning setting uses compact target proxies to construct reference paths. Figure 8 examines capacity-matched target warmup under normalized and raw loss reduction. Performance varies non-monotonically with depth, and the observed best depth differs across tasks under both scoring rules. These observations motivate task-specific depth calibration. Together with the reference-capacity ablation in Figure 6, they support capacity control and horizon calibration when constructing reference paths from limited target data.

E Notation Table

Table 3: Notation used throughout the paper. We reserve calligraphic $\mathcal{Z}$ for finite datasets or pools and calligraphic $\mathcal{D}$ for population distributions.

Symbol	Meaning
$\mathcal{Z}_{\text{pool}}$ or $\mathcal{Z}$	finite candidate pool from which examples are selected
$\mathcal{Z}_{\text{val}}$	finite target validation/proxy set used to construct the warmup trajectory
$\mathcal{D}$, $\mathcal{D}_{\text{tar}}$	target population distribution underlying $\mathcal{Z}_{\text{val}}$ and test data
z	one candidate or validation example; for SFT, an input-output sequence pair
$S \subseteq \mathcal{Z}_{\text{pool}}$	selected subset used for downstream retraining
N	selection budget, i.e., $\lvert S\rvert = N$
$\theta_0, \theta_t, \theta_T$	initial, intermediate, and final model parameters along a trajectory
θ_t^{val}, θ_t^{pool}	validation-induced and pool-induced trajectory states
$\mathcal{T} = \{\theta_t\}_{t=0}^{T}$	optimization trajectory used for attribution or scoring
η_t	learning rate or continuous-time step size at time t
$\ell(\theta; z)$	per-example loss of parameters θ on example z
$R(\theta)$, $R_{\text{val}}(\theta)$, $R_S(\theta)$	target, validation empirical, and selected-subset risks
$F(z, \mathcal{T})$	path alignment of candidate z along $\mathcal{T}$
$\bar{g}, \Sigma$	mean and covariance of candidate gradients at a reference state
$\delta_v(z)$, $\phi(z)$	score change under reference error v; pool typicality
$\Delta\ell(z)$	endpoint loss drop $\ell(\theta_1^{\text{val}}; z) - \ell(\theta_T^{\text{val}}; z)$
$s(z)$	normalized TACS score, $\Delta\ell(z)/\max\{\ell(\theta_1^{\text{val}}; z), \varepsilon\}$
r, α	LoRA rank and scaling factor used in adapter warmups
M, K	number of cross-validation folds and trajectory-depth candidates

F TRAJECTORY SCORES AS PATH INTEGRALS AND REFERENCE DYNAMICS

This appendix provides the formal continuous-time foundations for the trajectory-alignment formulation introduced in Section 2, details the perturbation analysis of reference paths, and connects finite-time trajectory attribution to classical influence functions.

F.1 TRAJECTORY-BASED ATTRIBUTION UNDER CONTINUOUS FLOW

Let R_0 denote the reference objective governing model optimization, and let R denote the target validation objective. Consider the learning-rate-weighted gradient flow trajectory:

$$\frac{d\theta_t}{dt} = -\eta_t \nabla R_0(\theta_t), \qquad \theta_0 \text{ fixed}, \tag{9}$$

where $\eta_t \geq 0$. For any candidate example z, we define the finite-time trajectory attribution score over horizon T as:

$$\mathcal{I}^T_{R,R_0}(z) := \int_0^T \eta_t \left\langle \nabla R(\theta_t), \nabla_\theta \ell(\theta_t; z) \right\rangle \, dt. \tag{10}$$

Lemma F.1 (Convergence of discrete checkpoint attribution)**.** Let checkpoints follow discrete gradient updates $\theta_{k+1} = \theta_k - (\tau_{k+1} - \tau_k)\eta(\tau_k)\nabla R_0(\theta_k)$ across a partition $0 = \tau_0 < \tau_1 < \cdots < \tau_K = T$. Suppose ∇R_0, ∇R, and $\nabla_\theta \ell(\cdot; z)$ are Lipschitz continuous and bounded on a compact neighborhood of the trajectory, and η_t is continuous on $[0, T]$. As the mesh size $h = \max_k(\tau_{k+1} - \tau_k) \to 0$, the discrete weighted attribution sum converges to $\mathcal{I}^T_{R,R_0}(z)$:

$$\lim_{h \to 0} \sum_{k=0}^{K-1} (\tau_{k+1} - \tau_k)\, \eta(\tau_k) \left\langle \nabla R(\theta_k), \nabla_\theta \ell(\theta_k; z) \right\rangle = \mathcal{I}^T_{R,R_0}(z). \tag{11}$$

Proof. Under the stated Lipschitz and boundedness conditions, explicit Euler discretization of Eq. (9) guarantees uniform trajectory error $\max_k \|\theta_k - \theta(\tau_k)\| = \mathcal{O}(h)$ (Hairer et al., 1993). Define the instantaneous alignment function $\psi(\theta) = \langle \nabla R(\theta), \nabla_\theta \ell(\theta; z) \rangle$. Because ∇R and $\nabla_\theta \ell$ are Lipschitz and bounded, $\psi(\theta)$ is Lipschitz on the trajectory domain. Consequently, substituting continuous states $\theta(\tau_k)$ for discrete checkpoints θ_k shifts the weighted sum by $\mathcal{O}(h)$. The resulting quantity $\sum_k (\tau_{k+1} - \tau_k)\eta(\tau_k)\psi(\theta(\tau_k))$ is a standard Riemann sum of the continuous integrand $\eta_t \psi(\theta_t)$, which converges directly to the Riemann integral $\mathcal{I}^T_{R,R_0}(z)$ as $h \to 0$. □

Equation (10) cleanly decouples two roles: the target objective R provides the directional gradient field, whereas the reference objective R_0 determines the trajectory manifold over which this field is evaluated.

F.2 REFERENCE-PATH DEPENDENCE

Definition F.2 (Path-induced attribution gap)**.** Let R_a and R_b be two distinct reference objectives generating trajectories $\theta_t^{(a)}$ and $\theta_t^{(b)}$, respectively. For candidate z, the path-induced attribution gap is:

$$\Gamma_T(z; R, R_a, R_b) := \mathcal{I}^T_{R,R_a}(z) - \mathcal{I}^T_{R,R_b}(z). \tag{12}$$

When candidate-pool and target distributions diverge, a pool-induced trajectory ($R_0 = R_{\text{pool}}$) and a target-induced trajectory ($R_0 = R_{\text{val}}$) traverse divergent parameter submanifolds. Consequently, the target gradient field $\nabla R(\theta_t)$ is evaluated at fundamentally different coordinates, altering both the magnitude and relative ordering of candidate scores. TACS resolves this divergence by fixing $R_0 = R_{\text{val}}$.

F.3 FIRST-ORDER REFERENCE ERRORS AND DIRECTIONAL BIAS

This subsection proves Proposition F.3, examines common-displacement properties, and analyzes the corresponding bias in gradient-alignment metrics.

Proposition F.3 (First-order reference errors and pool typicality)**.** Let $s(z;\theta)$ be differentiable at the reference state θ, and set $a_z = \nabla_\theta s(z;\theta)$. A small reference displacement v changes the score to first order by $\delta_v(z) = \langle a_z, v\rangle$. For any candidate statistic $\phi(z)$, $\mathrm{Cov}_z(\delta_v, \phi) = \langle v, \mathrm{Cov}_z(a_z, \phi)\rangle$. If v is independent of the candidate pool and has mean zero, the expected covariance is zero. For a raw endpoint loss drop perturbed at its final checkpoint, let g_z be the candidate gradient, $\bar{g} = \mathbb{E}_z g_z$, $\Sigma = \mathrm{Cov}_z(g_z)$, and $\phi(z) = \langle g_z, \bar{g}\rangle$. A pool-descent displacement $v = -\lambda\bar{g}$ with $\lambda > 0$ then gives $\mathrm{Cov}_z(\delta_v, \phi) = \lambda\bar{g}^\top\Sigma\bar{g} \geq 0$.

Proof of Proposition F.3. Let $s(z;\theta)$ denote a candidate scoring function evaluated at reference state θ. A first-order perturbation v shifts the score by $\delta_v(z) = \langle a_z, v\rangle$, where $a_z = \nabla_\theta s(z;\theta)$ is the candidate response vector. For any candidate statistic $\phi(z)$, bilinearity of covariance yields:

$$\mathrm{Cov}_z\left(\delta_v(z), \phi(z)\right) = \langle v, c_\phi\rangle, \qquad \text{where } c_\phi = \mathrm{Cov}_z\left(a_z, \phi(z)\right). \tag{13}$$

The vector c_ϕ is fixed by the candidate pool and the unperturbed state. If perturbation v is statistically independent of the candidate pool with $\mathbb{E}[v] = 0$, taking the expectation over v yields:

$$\mathbb{E}_v\left[\mathrm{Cov}_z(\delta_v, \phi)\right] = \langle \mathbb{E}[v], c_\phi\rangle = 0. \tag{14}$$

For the raw endpoint loss difference $s(z) = \ell(\theta_a; z) - \ell(\theta_b; z)$, a deviation at endpoint θ_b gives response vector $a_z = -\nabla_\theta\ell(\theta_b; z) =: -g_z$. Evaluating the covariance against candidate pool typicality $\phi(z) = \langle g_z, \bar{g}\rangle$ with pool covariance $\Sigma = \mathrm{Cov}_z(g_z)$ gives $c_\phi = -\Sigma\bar{g}$. Thus:

$$\mathrm{Cov}_z(\delta_v, \phi) = -v^\top\Sigma\bar{g}. \tag{15}$$

Under a pool-descent step $v = -\lambda\bar{g}$ ($\lambda > 0$), this evaluates to:

$$\mathrm{Cov}_z(\delta_v, \phi) = \lambda\,\bar{g}^\top\Sigma\bar{g} \geq 0, \tag{16}$$

which is non-negative because covariance matrix Σ is positive semi-definite. □

Joint endpoint perturbations. For the raw loss drop $s(z) = \ell(\theta_a; z) - \ell(\theta_b; z)$, if both reference checkpoints undergo an identical displacement v, the first-order variation is:

$$\delta_v(z) = \left[\nabla_\theta\ell(\theta_a; z) - \nabla_\theta\ell(\theta_b; z)\right]^\top v. \tag{17}$$

By the mean value theorem, this can be written as $\delta_v(z) = (\theta_a - \theta_b)^\top\bar{H}_z v$, where $\bar{H}_z = \int_0^1 \nabla^2_\theta\ell((1-t)\theta_b + t\theta_a; z)\,dt$ represents the path-averaged candidate Hessian. Thus, common displacements do not vanish identically in general non-convex settings; rather, they scale with candidate curvature and the trajectory displacement $\|\theta_a - \theta_b\|$.

Bias in gradient-alignment selectors. Consider local gradient-alignment metrics of the form $s(z;\theta) = \langle\nabla R_{\mathrm{val}}(\theta), \nabla\ell(\theta; z)\rangle$. Differentiating with respect to θ yields the response vector $a_z = H_{\mathrm{val}}(\theta)g_z + H_z(\theta)\nabla R_{\mathrm{val}}(\theta)$, where H_{val} and H_z denote the validation and candidate Hessians. Under locally homogeneous candidate curvature ($H_z \approx \bar{H}$), the second term is approximately common across candidates. Hence $c_\phi \approx H_{\mathrm{val}}\Sigma\bar{g}$, and a pool-descent perturbation $v = -\lambda\bar{g}$ gives

$$\mathrm{Cov}_z(\delta_v, \phi) \approx -\lambda\,\bar{g}^\top H_{\mathrm{val}}\Sigma\bar{g}. \tag{18}$$

This couples score changes to the pool statistics $\bar{g}, \Sigma$ and target curvature H_{val}, with the covariance sign determined by their joint geometry.

Empirical endpoint sensitivity diagnostic. Figure 9 evaluates the sensitivity of TACS rankings to imposed reference displacements: $\theta_T \leftarrow \theta_T + m\,\|\theta_T - \theta_1\|\,u$, where u is a unit vector and $m \in \{0.25, 0.5, 1, 2, 4\}$. Across model–task settings, structured displacements along pool-gradient or validation-gradient directions induce substantially larger ranking shifts than isotropic perturbations, corroborating the directional sensitivity characterized by Proposition F.3.

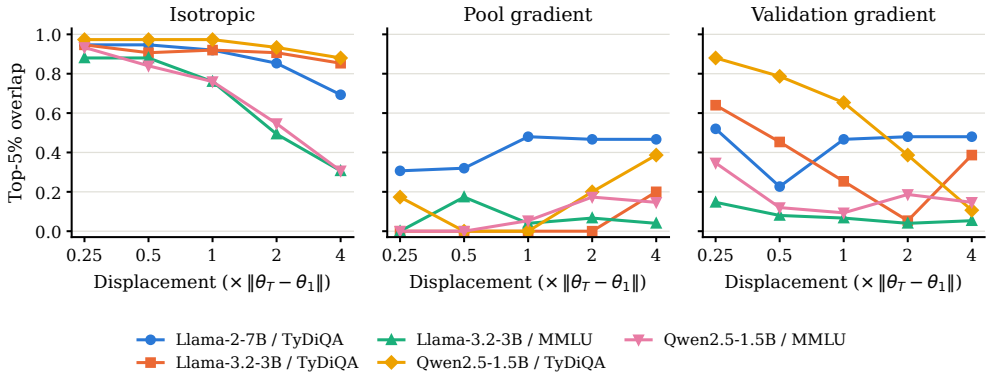


Figure 9: **Ranking sensitivity under structured reference perturbations.** Selection overlap (top-5%) with the unperturbed TACS subset as the reference endpoint is displaced along isotropic, pool-gradient, and validation-gradient directions across varying perturbation scales (1,500 Dolly candidates).

F.4 Extension to Adaptive Optimizers

The trajectory-conditioning effect extends directly to adaptive optimizers such as AdamW. In the small-step limit, Adam induces a preconditioned gradient flow:

$$\frac{d\theta_t}{dt} = -\eta_t P_t^{-1} \widetilde{\nabla} R_0(\theta_t), \tag{19}$$

where P_t incorporates running second-moment scaling and $\widetilde{\nabla} R_0$ reflects momentum smoothing. The corresponding trajectory attribution score becomes:

$$\mathcal{I}^T_{R,R_0,\mathrm{opt}}(z) := \int_0^T \eta_t \left\langle \nabla R(\theta_t), P_t^{-1} \widetilde{g}(z,\theta_t) \right\rangle dt, \tag{20}$$

where $\widetilde{g}(z,\theta_t)$ represents the momentum-smoothed candidate gradient. While optimizer preconditioning rescales the local metric tensor, the evaluation points θ_t remain strictly governed by reference objective R_0.

F.5 Connection to Classical Influence Functions

For a sample-weighted empirical objective $L(\theta; w) = \sum_j w_j \ell(\theta; z_j)$, differentiating continuous gradient flow with respect to weighting w_i yields the transported parameter sensitivity:

$$\frac{\partial \theta_T}{\partial w_i} = -\int_0^T \Phi(T,s)\, \eta_s \nabla_\theta \ell(\theta_s; z_i)\, ds, \tag{21}$$

where the transition matrix $\Phi(T,s) = \mathcal{T} \exp\left(-\int_s^T \eta_u \nabla^2_\theta L(\theta_u)\, du\right)$ transports earlier perturbations through subsequent trajectory Hessians.

Trajectory-based attribution differs from classical influence in two operational respects:

1. *Omission of recursive Hessian transport*: Trajectory attribution measures instantaneous gradient alignment at each visited state, bypassing the computationally prohibitive computation of $\Phi(T,s)$.
2. *Finite-time transient regime*: Classical influence functions evaluate asymptotic sensitivities at an isolated local minimum θ^*, yielding $-H(\theta^*)^{-1}\nabla_\theta \ell(\theta^*; z_i)$. In contrast, modern targeted adaptation operates within early, non-converged parameter regimes where the geometric structure of the reference path itself constitutes an active component of data selection.

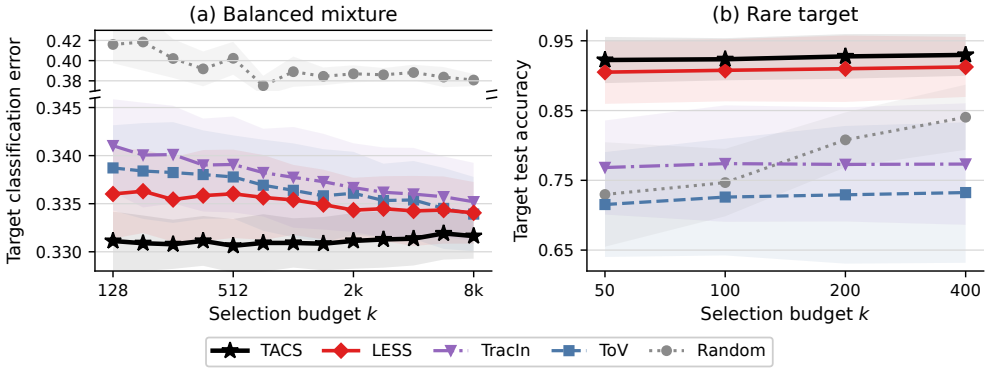


Figure 10: **Controlled logistic regression metrics across budgets** (mean ± std over 10 seeds). **(a)** Balanced mixture target error across budgets. **(b)** Rare-target test accuracy.

G Experimental Details

G.1 Controlled Logistic Regression Details

To isolate candidate selection geometry under exact distributions, we instantiate synthetic logistic models:

$$x \sim \mathcal{N}(0, I_d), \qquad \Pr(y = 1 \mid x) = \sigma(\langle x, \theta^* \rangle), \qquad \sigma(t) = \frac{1}{1 + \exp(-t)}. \tag{22}$$

Trajectories use full-batch gradient descent with linear learning-rate decay, retraining selected subsets from the shared initialization. Configurations are summarized in Table 4, with complete budget curves shown in Figure 10.

Table 4: **Controlled logistic regression configurations.** The balanced pool is an equal mixture of target and orthogonal-distractor environments; the rare-target pool contains 5% target data and six distractors at 80° to the target.

Configuration	Balanced Mixture	Rare-Target Stress Test
Dimension d	10	48
Pool / Validation / Test sizes	131,072 / 1,024 / 10,000	16,000 / 8 / 2,000
Pool reference schedule	4,096 rows, 4 steps, LR 0.5	2,000 rows, 36 steps, LR 0.55
Selection budgets	128–8,192 ($\times\sqrt{2}$)	50, 100, 200, 400
Retraining schedule	120 steps, LR 0.7	120 steps, LR 0.7
Seeds	10	10

Endpoint projection coordinates. Figure 4a projects retraining $(\theta_T^{(r)} - \theta_0^{(r)})$, validation-warmup $(\theta_T^{(v)} - \theta_0^{(v)})$, and pool-warmup displacements onto an orthonormal basis defined by the retraining direction $e_1 = (\theta_T^{(r)} - \theta_0^{(r)})/\|\theta_T^{(r)} - \theta_0^{(r)}\|_2$ and the residual orthogonal component $e_2 = u/\|u\|_2$, where $u = \theta_T^{(v)} - \theta_0^{(v)} - \langle \theta_T^{(v)} - \theta_0^{(v)}, e_1 \rangle e_1$.

Scoring protocol. TACS scores candidates by normalized loss reduction between initial and endpoint checkpoints along the validation path: $s_i^{\mathrm{TACS}} = [\ell_i(\theta_{\mathrm{ref}}) - \ell_i(\theta_{\mathrm{last}})]/\max\{\ell_i(\theta_{\mathrm{ref}}), \varepsilon\}$. Step count and learning rate are selected via validation AUROC on held-out target samples. Baselines (LESS, TracIn, and ToV) share pool checkpoints, candidate budgets, and retraining schedules. LESS scores by cosine gradient similarity across checkpoints, TracIn by unnormalized gradient inner products, and ToV by candidate loss drop under one-step validation perturbations ($\alpha = 0.1$).

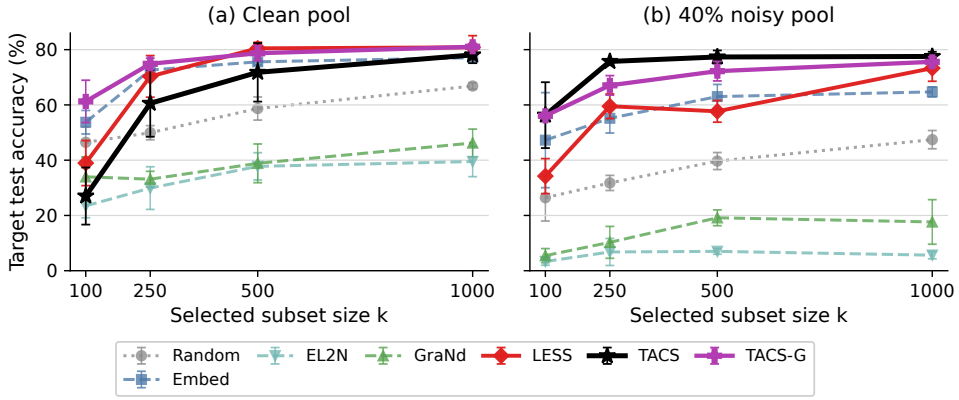


Figure 11: **Binary CIFAR-10 full fine-tuning target accuracy across budgets** under clean and 40% label-noise pools (mean ± std over 5 clean / 3 noisy seeds). TACS-G denotes class-grouped trajectories; LESS scores by cosine alignment.

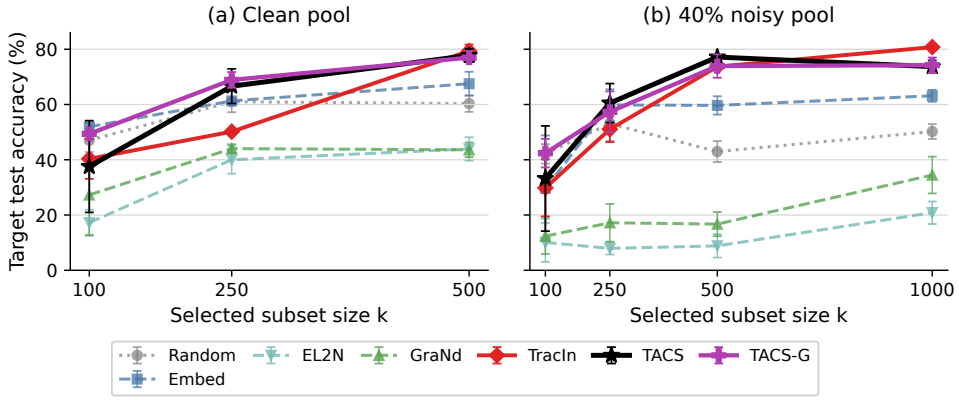


Figure 12: **Binary CIFAR-10 partial fine tuning accuracy across budgets** under clean and 40% label-noise pools (mean ± std over 2 clean / 3 noisy seeds).

G.2 Details for Computer Vision Benchmarks

Task setup and training protocols. We evaluate binary cat-versus-dog classification on CIFAR-10 (Krizhevsky & Hinton, 2009) with an ImageNet-pretrained (Deng et al., 2009) `ResNet-18` (He et al., 2016) (96×96 resolution). The target proxy contains 100 clean images. Candidate pools contain 10,000 images for full fine-tuning and 4,000 images for partial fine tuning on `layer4` and the classification head. The noisy regime injects 40% symmetric label noise into candidate pools while keeping proxy and test sets clean.

TACS generates reference trajectories via Adam (LR 5×10^{-5}, batch size 32), scoring candidates via normalized loss reduction between passes 1 and 4. Baselines include pool-reference gradient alignment (LESS under full fine-tuning, TracIn under partial fine tuning), EL2N, GraNd (Paul et al., 2021), and pretrained feature retrieval. Retraining runs for 8 epochs (full fine-tuning) or 4 epochs (partial fine tuning) at LR 10^{-4}. Performance sweeps across budgets are shown in Figures 11 and 12.

Class-grouped target trajectories (TACS-G). When target validation data covers multiple classes, joint warmups follow their mean gradient. The class-grouped variant, TACS-G, runs one trajectory per target class g matching the step schedule of the joint path, scoring each candidate by its maximum normalized response: $s_{\mathrm{G}}(z) = \max_g s_g(z)$. Grouping sharpens target class boundaries in clean pools while exhibiting a slight trade-off in label-noise resilience (Figure 11).

G.3 Details for Instruction Tuning

G.3.1 Dataset Statistics

Candidate pools and target proxies are detailed in Table 5. All data are standardized into the unified Tulu chat format (Wang et al., 2023).

Table 5: **Dataset statistics** for candidate pools ($\mathcal{Z}_{\text{pool}}$) and target proxies ($\mathcal{Z}_{\text{val}}$).

Candidate Training Pools		Target Proxies		
Source	**Size**	**Task**	**Size**	**Prompt Format**
Flan V2 (Longpre et al., 2023)	100,000	MMLU (Hendrycks et al., 2021)	285	5-shot
COT (Wei et al., 2022)	100,000	BBH (Suzgun et al., 2023)	81	1-shot
Oasst1 (Köpf et al., 2023)	55,668	TyDiQA (Clark et al., 2020)	9	1-shot
Dolly (Conover et al., 2023)	15,011			

G.3.2 Model and Training Configurations

All instruction tuning experiments use `Llama-3.2-3B` (Grattafiori et al., 2024), maximum sequence length 2,048, bfloat16 precision, and LoRA on attention projections with AdamW (zero weight decay). Checkpoint selection evaluates candidate checkpoints on validation data prior to test evaluation, reporting the optimal validation checkpoint alongside the final epoch-four checkpoint. Hyperparameters across stages are summarized in Table 6.

Table 6: **Instruction-tuning training settings.** Source-specific learning rates for pool warmup and retraining are selected on random subsets from $\{1, 2, 5\} \times 10^{-5}$ and shared across methods.

Stage	Data Allocation	Adapter Configuration	Optimization Schedule
LESS / ToV warmup	Random 5% pool subset	LoRA $r = 128, \alpha = 512$, drop 0.1	4 epochs, AdamW, source LR
TACS warmup	Full target proxy $\mathcal{Z}_{\text{val}}$	LoRA $r = 1, \alpha = 4$, drop 0.1	AdamW, LR 10^{-5}, $T = 2/2/64$ for MMLU/BBH/TyDiQA
Subset retraining	Selected 5% pool subset	LoRA $r = 128, \alpha = 512$, drop 0.1	4 epochs, effective batch 32, linear warmup

Baseline specifications. LESS follows its official implementation ($d_{\text{proj}} = 8192$), projecting Adam-preconditioned gradients and aggregating cosine similarities across 4 pool warmup checkpoints. ToV applies validation perturbations ($\alpha = 0.1$) around pool checkpoints, scoring candidates by maximum loss improvement across checkpoints.

G.3.3 Paired Clean-Run Differences

Table 7 reports TACS minus baseline in display points. Intervals use paired evaluation-item resampling and condition on the three fitted seed runs.

G.3.4 Score-Band Monotonicity Diagnostic

To evaluate whether TACS scores correlate monotonically with downstream utility across the candidate distribution, we sort the Dolly candidate pool by TACS score on TyDiQA and partition it into seven disjoint 5% bands. Each band is retrained across three seeds (3, 7, 42). As shown in Table 8, downstream TyDiQA performance aligns closely with score percentiles across both checkpoint selection criteria ($\rho = 0.964$, exact two-sided permutation $p = 0.0028$), confirming that the scoring function acts as a reliable proxy for downstream transfer.

G.3.5 Ablation Settings

Ablation experiments use the clean Dolly pool with `Llama-3.2-3B` across three seeds. Figure 6a evaluates LoRA rank $r \in \{1, 4, 16, 64, 128\}$ ($\alpha = 4r$) under the calibrated schedule ($\eta^* = 10^{-5}$,

Table 7: Paired clean-run differences. A positive value favors TACS; brackets are 95% paired evaluation-item bootstrap intervals conditional on the three fitted seed runs.

Pool	Task	vs. LESS: mean [95% CI]	vs. ToV: mean [95% CI]
COT	MMLU	0.42 [0.03, 0.80]	0.63 [0.15, 1.11]
	TyDiQA	-0.51 [-1.37, 0.34]	6.56 [5.40, 7.66]
	BBH	-2.01 [-3.73, -0.28]	-1.33 [-2.96, 0.34]
Dolly	MMLU	-0.12 [-0.31, 0.05]	0.12 [-0.08, 0.32]
	TyDiQA	2.12 [1.28, 2.94]	3.07 [2.14, 3.96]
	BBH	0.06 [-1.14, 1.23]	0.74 [-0.43, 1.94]
Flan V2	MMLU	1.09 [0.62, 1.55]	3.83 [3.27, 4.37]
	TyDiQA	-0.13 [-0.96, 0.72]	0.73 [-0.24, 1.70]
	BBH	1.98 [0.22, 3.77]	0.46 [-1.33, 2.28]
Oasst1	MMLU	0.21 [-0.09, 0.51]	0.04 [-0.33, 0.41]
	TyDiQA	3.34 [2.44, 4.24]	3.40 [2.57, 4.22]
	BBH	1.20 [-0.43, 2.78]	0.09 [-1.57, 1.70]

Table 8: **Dolly→TyDiQA score-band diagnostic on `Llama-3.2-3B`.** Disjoint 5% candidate bands retrained across seeds 3, 7, and 42, evaluated on 1,713 TyDiQA questions.

Score Percentile	Mean TACS Score	Fixed-Final F1	Validation-Selected F1
0–5% (Highest)	+0.03048	63.60	63.27
15–20%	+0.00799	58.24	57.34
30–35%	+0.00278	55.47	54.61
47.5–52.5%	+0.00031	43.81	44.03
65–70%	−0.00142	37.54	37.31
80–85%	−0.00386	42.34	40.53
95–100% (Lowest)	−0.01381	27.33	30.27

$T^* = 2$ for MMLU/BBH, $T^* = 64$ for TyDiQA). Figure 6b evaluates trajectory horizons $T \in \{2, 4, 8, 12, 16, 32, 64\}$.

G.3.6 Reference-Path Sensitivity Diagnostic

To empirically isolate reference-path dependence, we evaluate LESS selections on `Qwen2.5-1.5B` across structured warmups (k-means clusters, single-source warmups) versus random pool warmups on 750 reservoir examples (4 epochs, LoRA $r = 128$). Table 9 reports selection overlaps and total variation distances (TVD). In all cases, structured pool warmups induce substantially larger selection divergence than random draws, demonstrating that trajectory-based attribution is actively steered by the warmup distribution.

G.4 Ablations on Scoring Variants and Sequence Length

Length distribution across scoring criteria. Figure 13 illustrates length distributions of subsets selected by raw versus normalized endpoint drops across all instruction-tuning pools. Normalization effectively recalibrates candidate loss scales, confirming that the metric operates as an explicit loss-scale normalizer.

Scoring formulation ablation. Figure 14 evaluates normalization and baseline checkpoint choices ($a \in \{0, 1\}$) across synthetic and vision tasks:

$$s_a^{\text{raw}}(z) = \ell(\theta_a^{\text{val}}; z) - \ell(\theta_T^{\text{val}}; z), \qquad s_a^{\text{norm}}(z) = \frac{\ell(\theta_a^{\text{val}}; z) - \ell(\theta_T^{\text{val}}; z)}{\max\{\ell(\theta_a^{\text{val}}; z), \varepsilon\}}. \tag{23}$$

In noisy CIFAR-10, scale normalization substantially improves clean target sample concentration, particularly under restricted budgets, showing that relative loss reduction actively penalizes high-loss corrupted instances.

Table 9: **Reference-path sensitivity of LESS on `Qwen2.5-1.5B`.** Overlap measures the pairwise shared fraction among top-5% selections (chance = 0.05). TVD ratio compares composition distances between structured and random warmups.

				Top-5% overlap		
Pool	**Rank**	**Warmup family**	**Target**	Structured	Random	**TVD ratio**
Dolly	128	k-means	TyDiQA	0.236	0.654	5.5
	128	Diversity-matched	TyDiQA	0.307	0.654	4.4
	1	k-means	TyDiQA	0.610	0.898	6.0
	1†	k-means	TyDiQA	0.475	0.847	3.8
Flan V2	128	k-means	TyDiQA	0.307	0.456	–
Four sources	128	Source-isolated	TyDiQA	0.210	0.509	4.0
	128	Source-isolated	MMLU	0.110	0.406	3.1
	128	Source-isolated	BBH	0.097	0.417	3.7

All warmups use LoRA; rank denotes r. † Norm-matched warmup.

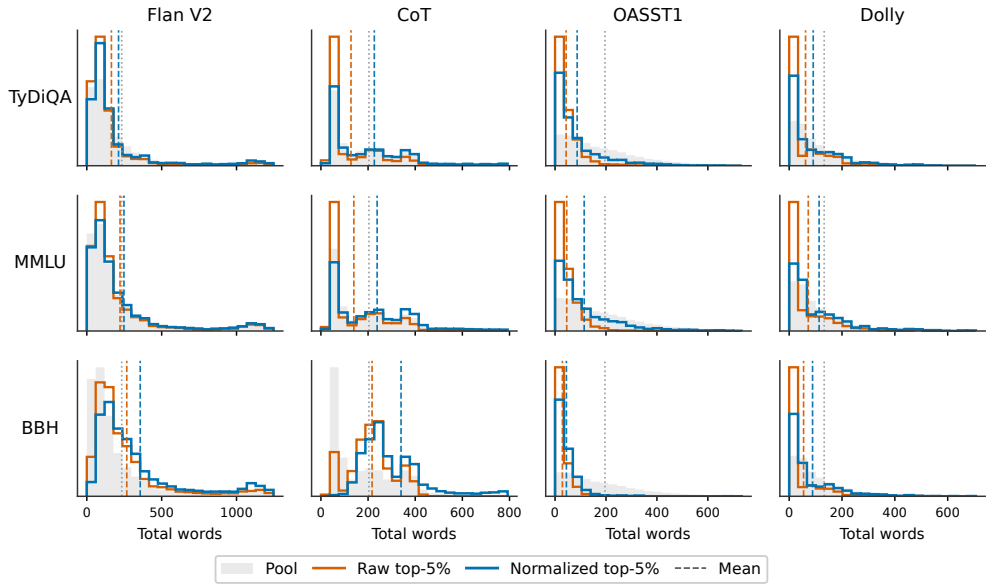


Figure 13: **Selected-length distributions under raw and normalized endpoint loss differences.** Full candidate pools (gray) versus selected top-5% subsets under raw and normalized TACS metrics.

G.5 Selection Cost Accounting Breakdown

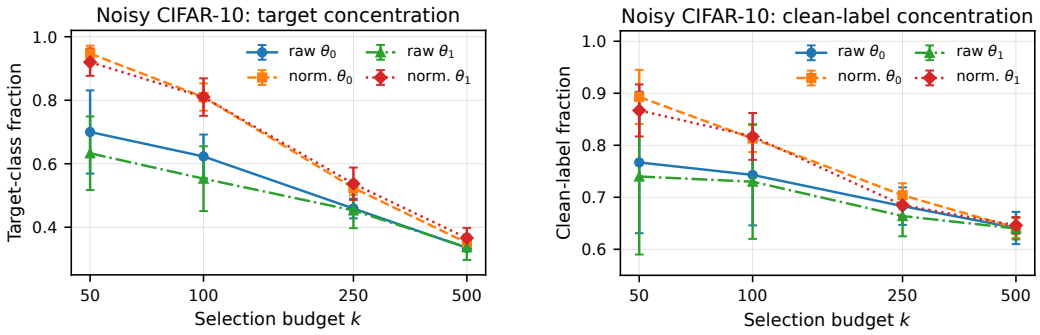
Table 2 and Figure 15 report selection costs in the seed-3 study, excluding downstream retraining and evaluation. Storage counts candidate scoring payloads transferred to disk.

The end-to-end selection compute decomposes as:

$$C_{\text{selection}} = C_{\text{warmup/calib}} + C_{\text{reference}} + C_{\text{scoring}}. \tag{24}$$

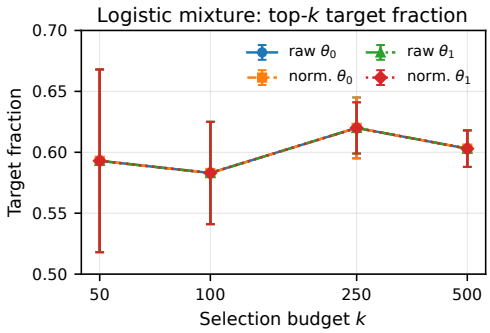
For LESS, selection totals 187.35 GPU-hours (185.53 h for pool warmups and candidate gradient extraction, 1.83 h for pairwise scoring). For ToV, selection totals 124.74 GPU-hours (4.24 h for pool warmups, 120.50 h for validation perturbations and candidate scoring). For TACS, selection totals 58.46 GPU-hours (7.54 h for the 3-fold calibration grid, 0.37 h for final target references, and 50.55 h for candidate scoring).

Figure 16 evaluates the empirical cost models across varying numbers of candidate pools (m) and target tasks (k). TACS is the most efficient selector in multi-pool regimes, while LESS becomes

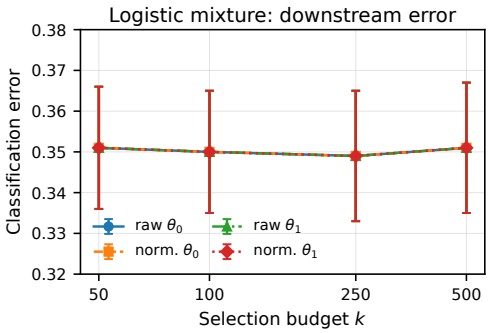


(a) CIFAR target-class fraction

(b) CIFAR clean-label fraction



(c) Logistic target fraction



(d) Logistic downstream error

Figure 14: **Score-variant ablation in CIFAR-10 and logistic regression.** Evaluating clean and target-class sample purity under raw vs. normalized and θ_0 vs. θ_1 baseline checkpoints.

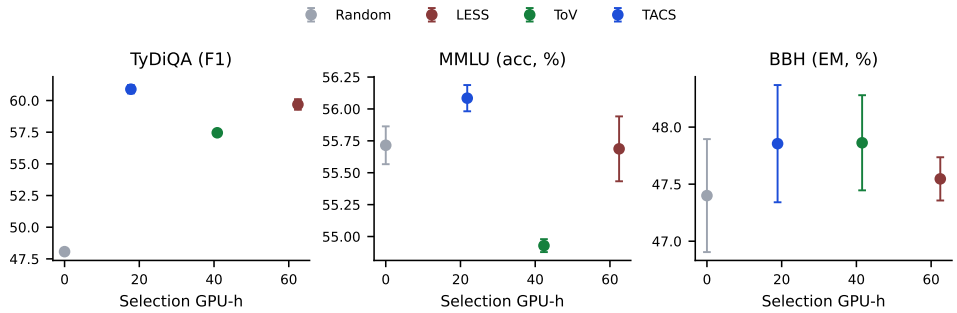


Figure 15: **Selection compute versus downstream utility** in instruction tuning ($m = 4$ pools, $k = 3$ targets). Points denote source-averaged three-seed means; error bars denote sample standard deviations. Common retraining and evaluation are excluded.

advantageous when a static pool is amortized across an exceptionally large number of downstream targets.

G.6 Trajectory Calibration Protocol

We calibrate learning rate η and horizon T via 3-fold cross-validation on proxy $\mathcal{Z}_{\text{val}}$ using rank-one LoRA ($r = 1, \alpha = 4$). Across candidate grids $\eta \in \{1, 2, 5\} \times 10^{-5}$ and $T \in \{2, 4, 8, 12, 16, 32, 64\}$, we record endpoint gradients on the warmup split (g_{W_j}), held-out split (g_{H_j}), and a generic reference sample P (g_P), maximizing mean specificity:

$$C_j(\eta, T) = \cos(g_{W_j}, g_{H_j}) - \cos(g_{W_j}, g_P). \tag{25}$$

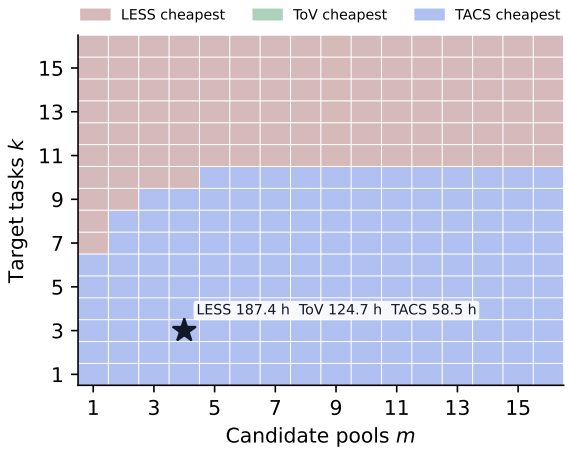


Figure 16: **Empirical cost crossover regimes.** Extrapolated compute boundaries as candidate pools (m) and downstream targets (k) scale under empirical cost models. The star marks the evaluated setup ($m = 4, k = 3$).

Ties break in favor of shorter trajectories and smaller learning rates (Algorithm 2). All tasks select $\eta^* = 10^{-5}$, with $T^* = 2$ for MMLU and BBH, and $T^* = 64$ for TyDiQA. Generic reference P consists of 100 fixed Latin-script examples from Aya (Singh et al., 2024), verified to have zero string overlap with candidate pools.

Algorithm 2 M-Fold Trajectory Calibration

Input: Target validation set $\mathcal{Z}_{\text{val}}$, generic reference sample P, grids H_η, H_T
Partition $\mathcal{Z}_{\text{val}}$ into $M = 3$ disjoint folds
for each candidate learning rate $\eta \in H_\eta$ **do**
  for each fold $j = 1, \ldots, M$ **do**
    Set held-out split $H_j \leftarrow$ fold j, warmup split $W_j \leftarrow \mathcal{Z}_{\text{val}} \setminus H_j$
    Train rank-1 LoRA on W_j with step size η up to $\max H_T$ epochs
    for each horizon $T \in H_T$ **do**
      Compute endpoint gradients g_{W_j}, g_{H_j}, g_P at checkpoint $\theta_{j,T}$
      $C_j(\eta, T) \leftarrow \cos(g_{W_j}, g_{H_j}) - \cos(g_{W_j}, g_P)$
    end for
  end for
  Compute mean specificity across folds: $\bar{C}(\eta, T) \leftarrow \frac{1}{M} \sum_{j=1}^{M} C_j(\eta, T)$
end for
return $(\eta^*, T^*) = \arg\max_{\eta,T} \bar{C}(\eta, T)$ (breaking ties in favor of smaller T, then smaller η)

H Proof of Reference-Induced Selection Uncertainty

We prove Theorem 4.1 for a fixed scored family $\mathcal{Q} = \{Q_u : u \in \mathcal{U}\}$. The warmup reference is random; the candidate family, target, checkpoint schedule, and scoring rule are fixed. Assume $\mathcal{U}$ is compact and $s(u; \xi)$ is continuous on $\mathcal{U}$ and C^3 in the reference and local feasible coordinates. On an open neighborhood of the compact convex reference set $\mathcal{N}$, suppose every score has a unique maximizer $v(\xi)$ in a C^3 feasible chart $u = \psi(a)$ with a negative-definite Hessian in a. Active constraints must have a stable active set. These conditions include regular smooth boundary maxima. Assume all empirical warmup means lie in $\mathcal{N}$ and have finite fourth moments.

Warmup covariance and selector response. For task mean ξ, averaging $R_u(w) = \frac{1}{2}\|w - u\|_H^2$ gives gradient $H(w - \xi)$. Solving $\dot{w} = -H(w - \xi)$ from zero yields $w_t = B_t \xi$. Conditional on composition C, $\mathbb{E}[\widehat{\mu}_m \mid C] = \mu_C$ and $\mathrm{Cov}(\widehat{\mu}_m \mid C) = \Sigma_C/m$. Total covariance therefore gives

$$\mathbb{E}\widehat{\mu}_m = \mu, \qquad \mathrm{Cov}(\widehat{\mu}_m) = \Omega_c + \Sigma_w/m. \tag{26}$$

For uniform sampling without replacement from $M_C > 1$ tasks, with $\Sigma_C = M_C^{-1} \sum_j (U_j - \mu_C)(U_j - \mu_C)^\top$, the conditional term becomes $(M_C - m)\Sigma_C/[m(M_C - 1)]$ by expanding inclusion-indicator covariances.

In local coordinates let $\widetilde{s}(a;\xi) = s(\psi(a);\xi)$, $A_\xi = -\nabla^2_{aa}\widetilde{s}(a_\xi;\xi) \succ 0$, and $C_\xi = \nabla^2_{a\xi}\widetilde{s}(a_\xi;\xi)$. The implicit function theorem applied to $\nabla_a \widetilde{s}(a_\xi;\xi) = 0$ gives a C^2 stationary branch. Negative curvature preserves local maximality; compactness and uniqueness ensure that the branch is the global maximizer locally and that branches agree on overlaps. A finite cover of $\mathcal{N}$ bounds its first two derivatives. Differentiation of stationarity yields

$$D_\xi a_\xi = A_\xi^{-1} C_\xi, \qquad J(\xi) = D\psi(a_\xi) A_\xi^{-1} C_\xi. \tag{27}$$

For an interior optimum this is $J = -s_{uu}^{-1} s_{u\xi}$. For $s(u;\xi) = \mathcal{A}(u; B_{t_1}\xi, \dots, B_{t_L}\xi, q)$, the chain rule gives $C_\xi = \sum_\ell \nabla^2_{aw_\ell} \widetilde{\mathcal{A}}(a_\xi; B_{t_1}\xi, \dots, B_{t_L}\xi, q) B_{t_\ell}$, with $\widetilde{\mathcal{A}}(a;\cdot) = \mathcal{A}(\psi(a);\cdot)$. Thus warmup dynamics, the score's mixed response, and inverse selection curvature jointly determine J. A strictly increasing smooth score transformation $h(s(u;\xi),\xi)$ preserves J: at stationarity its negative Hessian and mixed derivative are both multiplied by $h_s > 0$.

Finite-perturbation variance lemma. Let $F : \mathcal{N} \to \mathbb{R}^r$ be C^2, $\widehat{\xi} = \xi + \delta$, $\mathbb{E}\delta = 0$, $\mathrm{Cov}(\delta) = \Omega$, and $\|D^2F(x)[a,b]\| \le L_F\|a\|\|b\|$ along every relevant segment. With $L = DF(\xi)$, Taylor's formula gives $F(\xi+\delta) = F(\xi) + L\delta + r(\delta)$ and $\|r(\delta)\| \le L_F\|\delta\|^2/2$. Expanding centered squared norms and bounding the cross and remainder terms gives

$$\left| \mathrm{tr}\,\mathrm{Cov}(F(\widehat{\xi})) - \mathrm{tr}(L\Omega L^\top) \right| \le \|L\|_{\mathrm{op}} L_F \mathbb{E}\|\delta\|^3 + \frac{L_F^2}{4}\mathbb{E}\|\delta\|^4. \tag{28}$$

For independent V, V', $\frac{1}{2}\mathbb{E}\|V - V'\|_H^2 = \mathrm{tr}(H\,\mathrm{Cov}(V))$. Apply Eq. (28) to $F(\xi) = H^{1/2}v(\xi)$, $\delta = \widehat{\mu}_m - \mu$, and $\Omega = \Omega_c + \Sigma_w/m$. Apply it again to the scalar $f(\xi) = D(v(\xi))$, whose gradient is $g = J^\top H(v(\mu) - q)$. Since $\mathbb{E}\|\delta\|^3 \le \epsilon_m^3$ and $\mathbb{E}\|\delta\|^4 = \epsilon_m^4$, these two applications prove Eq. (6). The leading dispersion is positive exactly when $H^{1/2}J\Omega^{1/2} \neq 0$; the leading risk variance is positive exactly when $\Omega^{1/2}g \neq 0$. At $v(\mu) = q$, risk variance can be fourth order while selection dispersion is second order.

Exact limits and mean identity. Construct each warmup from an infinite conditional iid sequence. The conditional strong law gives $\widehat{\mu}_m \to \mu_C$ almost surely. Continuity of v, compactness of $\mathcal{U}$, and dominated convergence for first and second moments imply

$$I_m \to \mathrm{tr}(H\,\mathrm{Cov}_C(v(\mu_C))), \qquad \mathrm{Var}\, D(\widehat{v}_m) \to \mathrm{Var}_C\, D(v(\mu_C)). \tag{29}$$

These limits are positive precisely when the selected optimum and its target risk, respectively, are nonconstant almost surely across compositions. For any square-integrable selected optimum V, expanding $V - q = (\mathbb{E}V - q) + (V - \mathbb{E}V)$ gives $\mathbb{E}D(V) = D(\mathbb{E}V) + \frac{1}{2}\mathrm{tr}(H\,\mathrm{Cov}(V))$. Taking $V = \widehat{v}_m$ proves Eq. (7). The deterministic shift also satisfies $v(\mu) - v(q) = \int_0^1 J(q + t(\mu - q))(\mu - q)\,dt$ on a regular connecting segment.

Finite target proxy. For an iid n-sample target proxy $\widehat{q}_n$ with covariance Σ_D/n, finite fourth moments give $\mathbb{E}\|\widehat{q}_n - q\|^4 = O(n^{-2})$ by coordinatewise fourth-moment expansion and third absolute moments $O(n^{-3/2})$ by Hölder's inequality. Applying Eq. (28) at q gives

$$\begin{aligned} \mathrm{tr}(H\,\mathrm{Cov}(v(\widehat{q}_n))) &= n^{-1}\,\mathrm{tr}(HJ_q\Sigma_D J_q^\top) + O(n^{-3/2}), \\ \mathrm{Var}\, D(v(\widehat{q}_n)) &= n^{-1} g_q^\top \Sigma_D g_q + O(n^{-3/2}). \end{aligned} \tag{30}$$

where $J_q = Dv(q)$ and $g_q = J_q^\top H(v(q) - q)$. For local regularity, restrict Taylor expansion to a regular ball: fourth moments make excursion probability $O(n^{-2})$, while compactness bounds the omitted outcomes. A fixed target proxy has zero reference variation conditional on that proxy. If the same proxy enters both path and scoring target, use $v(\xi, r)$ and replace J_q by $\partial_\xi v(q,q) + \partial_r v(q,q)$. This completes the proof of Theorem 4.1.

H.1 Distributional disagreement and score scope

If Q_u has density p_u and $\xi \mapsto p_{v(\xi)}$ is twice differentiable as an L^1-valued map, write $h(x) = \nabla_\xi p_{v(\xi)}(x)$. For independent centered Gaussian Z, Z' with covariance Γ, an L^1 Taylor expansion and $\mathrm{TV}(P,Q) = \frac{1}{2}\|p-q\|_1$ give

$$\mathbb{E}\,\mathrm{TV}(Q_{v(\xi+\varepsilon Z)}, Q_{v(\xi+\varepsilon Z')}) = \frac{\varepsilon}{\sqrt{\pi}} \int \sqrt{h(x)^\top \Gamma h(x)}\, d\lambda(x) + O(\varepsilon^2). \tag{31}$$

The remainder follows from the bounded second derivative and reverse triangle inequality; Gaussian excursions outside the regular neighborhood are negligible with a probability-valued extension. For iid mean perturbations, $\varepsilon = m^{-1/2}$ gives order $m^{-1/2}$ disagreement when the coefficient is nonzero. A sample-mean CLT gives the same coefficient under uniform integrability. For category proportions $\pi(u)$, letting $b_k^\top$ be the rows of $D(\pi \circ v)(\xi)$, the same argument gives $\mathbb{E}\,\mathrm{TV}_{\mathrm{cat}} = \varepsilon \sum_k \sqrt{b_k^\top \Gamma b_k}/\sqrt{\pi} + O(\varepsilon^2)$. If square-root densities are C^2 as L^2 maps on common support, squared Hellinger distance has leading term $\varepsilon^2\,\mathrm{tr}(F_Q(v(\xi))J\Gamma J^\top)/4$, where F_Q is Fisher information; the remainder is $O(\varepsilon^3 + \varepsilon^4)$.

The score class includes normalized checkpoint alignments $\sum_\ell \alpha_\ell \langle H(w_\ell - u), H(w_\ell - q)\rangle / [\|H(w_\ell - u)\|^\rho \|H(w_\ell - q)\|^\nu]$, $w_\ell = B_{t_\ell}\xi$, on domains with nonzero denominators (including $0 \le \rho \le 2$, $\nu \ge 0$). It also includes smooth denominator regularizations, endpoint loss changes, nonlinear checkpoint aggregations, and preconditioned bilinear scores. Each still requires a regular optimum. A maximum across checkpoint scores is smooth on a stable unique active branch; ties require directional analysis. A reference-independent score has $J = 0$.

For a constrained example, take $\mathcal{U} = \{u : \|u\| = R\}$ and $B = \sum_\ell \alpha_\ell B_{t_\ell}$ with $\sum_\ell \alpha_\ell = 1$. Unnormalized checkpoint alignment reduces, up to a candidate-common term, to $s(u;\xi) = u^\top b(\xi)$ with $b(\xi) = H^2(q - B\xi)$. For $b \neq 0$, $v(\xi) = Rb/\|b\|$ and $J(\xi) = -R(I - ee^\top)H^2B/\|b\|$, where $e = b/\|b\|$. The tangent Hessian is $-\|b\|I/R$, proving regularity. Sensitivity increases as the residual ranking direction weakens, subject to remaining away from $b = 0$.

I Distribution Selection in the Linearized Model

We analyze the population preferences behind Theorems 4.2 and 4.3 in the linearized model of Section 4.2. The shared covariance controls how quickly target directions are learned, while candidate and target signals are represented by u and q. This setting separates direction, scale, and orthogonal components without the combinatorics of finite-subset selection; optimizer corrections and checkpoint aggregation remain outside the theorem. It is also closely related to linearized analyses of language-model fine-tuning (Afzal et al., 2025).

I.1 Detailed Analysis and Proofs

Which candidate distributions are preferred? Let target gradient flow reach v_τ after warmup time $\tau > 0$. Write a candidate optimum as $u = av_\tau + b$, where $b^\top H v_\tau = 0$. With an inactive denominator floor, normalized endpoint scoring gives, for $u \neq 0$,

$$s(u; v_\tau) = 1 - \frac{\|u - v_\tau\|_H^2}{\|u\|_H^2} = \frac{(2a-1)\|v_\tau\|_H^2}{a^2\|v_\tau\|_H^2 + \|b\|_H^2}. \tag{32}$$

For any $0 < c < 1$, completing the square gives

$$s(u; v_\tau) \ge c \quad \Longleftrightarrow \quad \left\| u - \frac{v_\tau}{c} \right\|_H^2 \le \frac{1-c}{c^2}\|v_\tau\|_H^2.$$

Thus, every positive-score superlevel set is bounded and shrinks to $\{v_\tau\}$ as $c \uparrow 1$.

The unique maximizer over this task family is $u = v_\tau$. At fixed $a > 1/2$, increasing $\|b\|_H$ lowers the score. TACS therefore favors candidate distributions whose optima approach the target warmup endpoint and penalizes orthogonal components in the geometry of the loss. Both the preferred direction and scale are determined by the target trajectory.

Dot-product alignment and raw loss drop. For the fixed target endpoint v_τ, dot-product alignment is $d(u, v_\tau) = u^\top H v_\tau = a\|v_\tau\|_H^2$. Expanding the endpoint loss drop gives

$$\begin{aligned}\Delta(u; v_\tau) &= R_u(0) - R_u(v_\tau) = \tfrac{1}{2}\|u\|_H^2 - \tfrac{1}{2}\|u - v_\tau\|_H^2 \\ &= d(u, v_\tau) - \tfrac{1}{2}\|v_\tau\|_H^2 = (a - \tfrac{1}{2})\|v_\tau\|_H^2.\end{aligned}$$

The subtracted term is constant across candidates for fixed v_τ, so both scores induce identical rankings, including ties. Both are independent of b at fixed a and tend to infinity as $a \to \infty$. These preferences hold over the unrestricted candidate family; a finite pool or a constraint on candidate signal magnitude can limit the attainable scores. For TACS, the same increase in orthogonal signal raises the normalization denominator while leaving the raw loss drop unchanged, explaining the penalty when $a > 1/2$.

Cosine alignment to the same endpoint is

$$c(u, v_\tau) = \frac{u^\top H v_\tau}{\|u\|_H \|v_\tau\|_H}.$$

By Cauchy–Schwarz it is at most one, with equality precisely when $u = \alpha v_\tau$, $\alpha > 0$. It therefore fixes the direction but leaves the candidate scale undetermined.

Connection to trajectory-integrated gradient alignment. Along target gradient flow $\dot{w}_t = -\nabla R_q(w_t)$, $w_0 = 0$, the chain rule gives the exact identity

$$\int_0^\tau \langle \nabla R_u(w_t), \nabla R_q(w_t) \rangle \, dt = R_u(0) - R_u(v_\tau) = d(u, v_\tau) - \tfrac{1}{2}\|v_\tau\|_H^2.$$

Thus integrated gradient dot products along this target path and endpoint dot-product alignment give the same ranking in the quadratic model. The instantaneous gradient inner product uses the ordinary parameter-space metric; d is the resulting endpoint expression in the H metric. This identity does not equate an arbitrary weighted checkpoint sum on a pool-trained path with an endpoint score. Cosine-normalized gradients also do not satisfy this loss-difference identity.

Actual single-checkpoint gradient cosine is a related, distinct comparison. At a reference checkpoint $w = \gamma q + \xi \neq q$, where $\xi^\top H q = 0$, the candidate and target gradients are $H(w - u)$ and $H(w - q)$. Since H is invertible, their cosine reaches one precisely when

$$u = w + \alpha(q - w) = [\gamma + \alpha(1 - \gamma)]q + (1 - \alpha)\xi, \qquad \alpha > 0. \tag{33}$$

Thus, equally scoring candidates can carry different multiples of the reference residual ξ. This family includes the target optimum at $\alpha = 1$; the cosine criterion alone leaves the scale undetermined. The calculation describes a single-checkpoint gradient cosine score, before the optimizer adjustments and aggregation used by LESS. Its preferred ray is generally different from the endpoint-alignment ray in Theorem 4.2: v_τ need not be parallel to q.

These preferences matter through the update induced by the selected tasks. For a uniformly weighted collection of candidate distributions S, write its mean optimum as $\bar{u}_S = \bar{a}_S q + \bar{b}_S$, with $\bar{b}_S^\top H q = 0$. Training on their average risk converges to $\bar{u}_S$, whose target risk is

$$R_q(\bar{u}_S) = \frac{1}{2}(\bar{a}_S - 1)^2 \|q\|_H^2 + \frac{1}{2}\|\bar{b}_S\|_H^2. \tag{34}$$

Target performance therefore depends on both the aligned scale and the mean orthogonal residual. TACS penalizes residuals relative to v_τ, whereas downstream risk is measured relative to q. Understanding which data it favors thus also requires understanding how well the warmup endpoint captures the target.

Which target components guide selection? Let $He_i = \lambda_i e_i$, with orthonormal e_i and $\lambda_1 \geq \cdots \geq \lambda_d > 0$, and write $q_i = e_i^\top q$. Starting from zero, target gradient flow reaches

$$v_\tau = (I - e^{-H\tau})q = \sum_{i=1}^{d} (1 - e^{-\lambda_i \tau}) q_i e_i \tag{35}$$

at time τ.

Thus, the distribution maximizing TACS matches a spectrally filtered target. Components with $\lambda_i \tau \gg 1$ are largely preserved, while those with $\lambda_i \tau \ll 1$ are attenuated. TACS favors candidate distributions matching the target components learned during warmup. A component can be relevant to the target yet receive little support from a short reference trajectory if it lies in a slowly learned direction.

Training to convergence on the score-maximizing distribution P_{v_τ} yields $w = v_\tau$, with relative target risk

$$\frac{R_q(v_\tau)}{R_q(0)} = \frac{\sum_{i=1}^d \lambda_i q_i^2 e^{-2\lambda_i \tau}}{\sum_{i=1}^d \lambda_i q_i^2}. \tag{36}$$

To characterize when this preferred distribution captures the target well, define the cumulative target power in the leading eigendirections,

$$C_q(k) = \frac{\sum_{i=1}^k \lambda_i q_i^2}{\sum_{i=1}^d \lambda_i q_i^2}. \tag{37}$$

For a fixed H, if one target has at least as large a $C_q(k)$ at every k as another, its relative risk in Equation (36) is no larger at any warmup duration. This follows by summation by parts, since $e^{-2\lambda_i \tau}$ increases with i. The condition expresses task–model alignment through cumulative spectral power (Canatar et al., 2021): target signal concentrated in rapidly learned directions makes the score-preferred distribution a closer approximation to the target. Concentration in a few directions alone is insufficient if those directions are learned slowly.

The score's penalty on orthogonal components is therefore relative to the warmup endpoint, which inherits the target trajectory's spectral preference. All target components in this model are representable and eventually learned; the attenuation arises from finite warmup. The score-maximizing distribution need not minimize downstream target risk, which also depends on the aggregate update from selected tasks. These results characterize population-level distribution preferences; individual-example scoring with cross-entropy introduces additional loss-scale and curvature effects.

Proof of Theorem 4.2. The population gradient is $\nabla R_u(w) = H(w - u)$. Consequently, target gradient flow solves $\dot{v} = -H(v - q)$, $v(0) = 0$, yielding Equation (35). Substituting the population risks into the normalized endpoint score and using $b^\top H v_\tau = 0$ gives Equation (32). The score is at most one, with equality only at $u = v_\tau$; its numerator in the orthogonal decomposition is positive for $a > 1/2$, so increasing $\|b\|_H^2$ strictly decreases it.

For cosine alignment, Cauchy–Schwarz in the H inner product gives $c(u, v) \leq 1$, with equality exactly when $u = \alpha v$, $\alpha > 0$. For dot-product alignment, $b^\top H v = 0$ gives $d(u, v) = a\|v\|_H^2$, independent of the orthogonal residual and unbounded as $a \to \infty$. Expanding the quadratic risks gives $R_u(0) - R_u(v) = d(u, v) - \frac{1}{2}\|v\|_H^2$. For fixed v, the subtracted term is independent of u, proving the claimed ranking equivalence with raw loss drop. Writing $q = cv + r$ and $u = av + b$, with r, b both H-orthogonal to v, gives $D(u) = \frac{1}{2}(a - c)^2\|v\|_H^2 + \frac{1}{2}\|b - r\|_H^2$. For a mixture of candidate distributions, the average gradient is $H(w - \bar{u}_S)$, so gradient flow converges to $\bar{u}_S$. Expanding $\|\bar{u}_S - q\|_H^2$ in its target-aligned and orthogonal components gives Equation (34).

Proof of Theorem 4.3. The gradient-flow solution above gives the preferred optimum. Substitution into R_q gives the stated relative risk. Let $p_i = \lambda_i q_i^2 / \|q\|_H^2$ and $h_i = e^{-2\lambda_i \tau}$. The relative risk in Equation (36) is $\sum_i p_i h_i$. Since $C_q(k) = \sum_{i \leq k} p_i$ and $C_q(d) = 1$, summation by parts gives

$$\sum_{i=1}^d p_i h_i = h_d - \sum_{k=1}^{d-1} C_q(k)(h_{k+1} - h_k).$$

Each difference $h_{k+1} - h_k$ is nonnegative. Thus, for fixed H, larger cumulative target power at every k implies no larger relative risk at every warmup duration.