IS DENSE SUPERVISION WORTH ITS COST? A BUDGET-CONTROLLED STUDY OF SUPERVISION ALLOCATION IN MAML

Anonymous authors
Paper under double-blind review

ABSTRACT

Does supervising more adaptation steps produce a better meta-learner under a fixed training budget? We benchmark dense, randomly sampled and fixed-horizon supervision in MAML with exact and approximate differentiation, pairing task streams and evaluation episodes, measuring complete update cost, and evaluating the full adaptation path. Across five training seeds on CIFAR-FS and MiniImageNet, sampled and intermediate fixed-horizon supervision achieve lower mean full-path negative log-likelihood than dense supervision at a 1,800-second budget under exact Conv4 full-model adaptation. Fixed-5 achieves the lowest mean on both datasets. Both intermediate fixed horizons outperform full-budget dense supervision in mean loss at half the budget. Allocation and differentiation interact: larger relative gains can accompany worse absolute performance. Early and late prediction depths can favor different fixed targets. Later adaptation can improve accuracy while worsening NLL; temperature adjustment reduces this penalty while preserving fixed-5's mean advantage over sampled supervision. Our analysis separates training-target bias from differentiation error and gives a local quadratic cost–progress condition for a biased direction to outperform a same-target reference. These findings motivate joint selection of allocation and differentiation at the intended budget using depth-weighted probability scores alongside accuracy.

1 INTRODUCTION

A few-shot learner may be used after one adaptation step or after ten. Where along this path should query losses provide training supervision? *Dense Uniform supervision* averages losses after every support update. It trains for the whole path, but each additional loss consumes time that could fund more tasks and outer updates.

Can we retain the same target while spending less on each task? We use *single-horizon meta-learning (SHML)* as a sampled reference: it selects one depth uniformly per task and constructs only that prefix and query loss. Its expected loss equals the dense average. The comparison with Uniform therefore tests within-task averaging against its cost.

A further question is whether preserving that target helps at a deadline. *Fixed supervision* always uses the query loss at one chosen depth, changing both the target and update cost. Comparing it with Uniform and SHML asks whether concentrating supervision can improve the entire evaluation path. These procedures separate how much to average within a task from which target to optimize.

Our controlled benchmark crosses these allocations with exact, first-order and prefix-truncated differentiation in Conv4/full adaptation, plus a ResNet12/head comparison. It pairs task streams and evaluation episodes on CIFAR-FS and MiniImageNet under fixed-time and early equal-update budgets.

The main finding is a divergence between target fidelity and budgeted performance. At the exact Conv4/full 1,800-second endpoint, fixed-5 has the lowest mean full-path NLL on both datasets, improving over both references in all five seeds. On CIFAR-FS, fixed-5 improves NLL and accuracy over SHML at every depth in every seed. Rankings change with differentiation, adaptation setting and resource constraint.

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.

Full-path diagnostics connect these findings to prediction requirements: early and late depths can favor different fixed targets. Accuracy can improve after NLL begins to rise; temperature adjustment reduces much of fixed-3's late penalty while preserving fixed-5's mean gain over SHML. Our analysis separates target bias from derivative error and derives a local quadratic cost–progress condition organizing these comparisons.

Our contributions are:

- **Target fidelity does not determine budgeted performance.** A five-seed controlled benchmark finds that intermediate fixed targets improve mean full-path NLL over both references at the main endpoint. A cost–progress analysis separates the roles of target bias, derivative error, variance and cost.
- **Allocation and differentiation require joint comparison.** Their interaction changes absolute outcomes; a larger advantage over a reference can coexist with worse performance. Resource and model/adaptation settings also change the ranking.
- **Prediction depth and probability scale matter.** Early–late contrasts yield an empirical model-choice boundary under depth reweighting. Scale diagnostics reduce late NLL degradation while preserving the fixed-5/SHML mean gain, separating two observed effects.

2 Related Work

Controlled comparisons in few-shot learning. Implementation, backbone and evaluation choices affect few-shot comparisons (Chen et al., 2019); training and adaptation can also be studied as separate phases (Luo et al., 2023). Within MAML, restricted adaptation distinguishes rapid learning from feature reuse (Raghu et al., 2020), while BOIL examines body-only adaptation (Oh et al., 2021). We compare where query supervision is applied and how its meta-gradient is computed, using common forward-adaptation rules and a shared evaluation target.

Supervision and differentiation through adaptation. MAML differentiates through support updates to optimize an initialization (Finn et al., 2017); its original MiniImageNet experiments used five adaptation steps during training and ten at test time. Fixed-5 follows this training-horizon choice. Shared classifier initialization is a further design choice (Ye & Chao, 2022). MAML++ supervises intermediate states and anneals their loss weights toward the final step, and switches from first-order to second-order differentiation during training (Antoniou et al., 2019). Our dense reference instead keeps uniform weights and a fixed derivative rule within each run. MAML-stop samples a training depth from a model- and task-dependent oracle distribution and recomputes that path (Chen et al., 2020, Section 4.5 and Appendix B.2.2); SHML samples uniformly before adaptation, constructing only that prefix. SHOT regularizes curvature along adaptation (Lee et al., 2023), while Cooperative Meta-Learning augments gradients through a co-learner (Shin et al., 2024). Our allocation comparison holds the adaptation recipe fixed within each setting.

First-order approximations and truncated bilevel gradients reduce derivative computation (Nichol et al., 2018; Shaban et al., 2019). Unlike truncation retaining the final inner iterates, prefix-3 retains exact differentiation through the first three support-gradient computations. Memory-reduced meta-learning avoids retaining inner parameter history (Yang et al., 2025), and BinomGBML uses a truncated binomial expansion for meta-gradient estimation (Zhang et al., 2026). We cross derivative fidelity with supervision allocation to measure their joint budgeted effect.

Finite-budget optimization and randomized computation. Budgeted Training treats the optimization budget as part of the procedure (Li et al., 2020). Randomized telescoping estimators study cheaper stochastic computation and progress per cost (Beatson & Adams, 2019); MAID adapts lower-level and hypergradient accuracy with the outer step size under a solver-work budget (Salehi et al., 2025). Here SHML changes the estimator and its cost, whereas fixed supervision also changes the target. Measured training time and equal-update observations expose different aspects of that trade-off.

Evaluation as part of the comparison. DomainBed combines reference implementations with explicit conditions and selection rules (Gulrajani & Lopez-Paz, 2021). Our benchmark applies

this evaluation-centered approach to meta-training choices. Full-path evaluation also separates classification from probability quality: accuracy can improve while NLL worsens, and temperature scaling can diagnose score-scale effects (Guo et al., 2017). We locate these contrasts along adaptation and test whether supervision-allocation gains persist after scalar adjustment.

3 Problem Setup

We study how to allocate supervision along an adaptation path when meta-training time is fixed. Let $\tau = (S_\tau, Q_\tau)$ be a support–query task and θ a meta-initialization. Applying h support updates gives $U_h(\theta, S_\tau)$ and query NLL $\ell_{\tau,h}(\theta) = \ell(Q_\tau; U_h(\theta, S_\tau))$. For an adaptation-parameter subset A, each update uses

$$\phi_h = \phi_{h-1} - \alpha P_A C_\gamma(\nabla_A \ell(S_\tau; \phi_{h-1}))\,, \qquad C_\gamma(v) = v \min\{1, \gamma/\|v\|\}, \quad \phi_0 = \theta, \tag{1}$$

where P_A embeds the adapted coordinates and $C_\gamma(0) = 0$. Training augmentation is part of the sampled task; evaluation uses unaugmented held-out tasks.

For $H = 10$ and fixed, task-independent weights $p_h \geq 0$, $\sum_h p_h = 1$, define the training objective and evaluation risk separately:

$$F_p(\theta) = \mathbb{E}_{\tau\sim\mathcal{D}_{\text{tr}}} \sum_{h=1}^{H} p_h \ell_{\tau,h}(\theta), \qquad R_q(\theta) = \mathbb{E}_{\tau\sim\mathcal{D}_{\text{ev}}} \sum_{h=1}^{H} q_h \ell_{\tau,h}(\theta), \quad q_h = 1/H. \tag{2}$$

The benchmark fixes a uniform evaluation prior, $q_{10} = q$, and uses $F = F_q$ as its reference training target. Matching ∇F is training-target fidelity, not unbiased estimation of the gradient of unseen-class risk R_q. Uniform weighting assigns equal importance to each allowed stopping depth; the individual risks retain the information needed to inspect other priorities.

A procedure a specifies a supervision policy and a derivative rule. If its uth complete outer update costs $C_{a,u} > 0$, the saved time checkpoint is

$$U_a(T) = \min\left\{ u : \sum_{j=1}^{u} C_{a,j} \geq T \right\}, \qquad \theta_a(T) = \theta_{a,U_a(T)}. \tag{3}$$

Thus the recorded time can exceed T by the crossing update. The comparison is $R_q(\theta_a(T))$, together with each of its ten horizon-specific risks. Comparing $R_q(\theta_{a,u})$ at a common update count answers a different question by holding the number of sampled tasks constant.

4 From Dense to Sampled and Fixed Supervision

Uniform: average the full path. For each task, Uniform uses $H^{-1}\sum_{h=1}^{H} \ell_{\tau,h}$, constructing all ten support-update states and evaluating ten query losses. It computes the full horizon average for the reference training target $F = F_q$.

SHML: sample one term of the same average. For each task, SHML independently draws $J \sim q$, with $q_h = 1/H$, and uses $\ell_{\tau,J}$. Since $\mathbb{E}_J[\ell_{\tau,J} \mid \tau] = H^{-1}\sum_h \ell_{\tau,h}$, it preserves the expected loss while constructing only J states and one query loss. Its expected depth is 5.5, exchanging within-task averaging for cheaper updates and additional tasks.

Fixed: concentrate supervision at one depth. Fixed-k uses $\ell_{\tau,k}$ on every task, constructing k states and one query loss. Its point-mass target is F_{δ_k} rather than F_q. We compare $k \in \{3, 5, 10\}$: fixed-5 has similar expected depth to SHML, while fixed-10 shares Uniform's maximum depth but uses one query loss. All procedures are evaluated at every step from 1 to 10, including depths beyond a fixed training horizon.

Derivative fidelity. Exact differentiates every executed support-gradient computation in Equation 1, including its clipping scale. FO treats all support gradients as constants but retains the parameter identity path. Prefix-3 differentiates the first three support-gradient computations and stops the derivative through subsequent ones. These choices preserve the numerical forward adaptation for a common initialization and task. Exact fixed-3 and prefix-3 fixed-3 are identical and share a trajectory; FO fixed-3 remains a separate control.

Common optimizer and architecture. Each outer update averages four raw task directions, clips their global norm at 1, applies Adam, and projects classifier initialization rows to their mean. Optimizer moments remain row-specific; classifier rows are distinct during adaptation. Conv4/full adapts all parameters. ResNet12/head adapts the classifier while meta-training the encoder through differentiable cached features. Conv4 uses 64 channels per block. Both architectures use eight-group Group Normalization (Wu & He, 2018), inner learning rate 0.01, and global inner-gradient norm clipping at 5. The recipe is shared across policies within each setting (Appendix B).

5 Analysis: Target Fidelity and Progress per Unit Cost

Three comparisons organize the analysis. Uniform versus SHML isolates within-task averaging at the same target; fixed versus both references changes the target; crossing each allocation with derivative rules changes gradient fidelity. The analysis gives common-state conditions, while the benchmark measures the resulting complete training trajectories.

5.1 Dense versus sampled: the same mean at a different cost

Let $\widetilde{g}^r_{\tau,h}$ be the raw outer direction for horizon h under derivative rule r, before outer clipping and Adam. Write $g_{\tau,h} = \nabla \ell_{\tau,h}$ for the exact derivative. All quantities below are conditional on the same initialization; scalar variance means $\mathbb{E}\|X - \mathbb{E}X\|^2$, the trace of the covariance matrix.

Proposition 1 (Sampling identity and variance)**.** *At a fixed initialization, assume the raw directions have finite second moments. For any of the three derivative rules r in Section 4,*

$$\mathbb{E}_J[\widetilde{g}^r_{\tau,J} \mid \tau] = \widetilde{g}^r_{\tau,\mathrm{U}} := H^{-1} \sum_h \widetilde{g}^r_{\tau,h}. \tag{4}$$

For an average of N independent tasks, the scalar variances satisfy

$$v_{\mathrm{S},r} = \frac{V_{\mathrm{task},r} + V_{\mathrm{hor},r}}{N}, \qquad v_{\mathrm{U},r} = \frac{V_{\mathrm{task},r}}{N}, \quad V_{\mathrm{hor},r} = \mathbb{E}_\tau \mathbb{E}_J \|\widetilde{g}^r_{\tau,J} - \widetilde{g}^r_{\tau,\mathrm{U}}\|^2, \tag{5}$$

where $V_{\mathrm{task},r} = \mathbb{E}_\tau \|\widetilde{g}^r_{\tau,\mathrm{U}} - m_{q,r}\|^2$ and $m_{q,r} = \mathbb{E}_\tau \widetilde{g}^r_{\tau,\mathrm{U}}$. Under exact differentiation and interchange of expectation and differentiation, $m_{q,\mathrm{exact}} = \nabla F$.

Uniform removes conditional horizon variance by averaging within a task; SHML retains it at lower cost. Their fixed-time comparison asks whether the extra averaging is worth its cost.

5.2 Fixed targets and approximate derivatives change different terms

For fixed, task-independent weights p, the mean raw direction decomposes as

$$m_{p,r} = \nabla F + b_p + e_{p,r}, \quad b_p = \nabla F_p - \nabla F, \quad e_{p,r} = \mathbb{E}_\tau \sum_h p_h (\widetilde{g}^r_{\tau,h} - g_{\tau,h}). \tag{6}$$

Here b_p is training-target bias and $e_{p,r}$ is differentiation error. Exact Uniform and exact SHML have both terms zero. Fixed-k generally has $b_{\delta_k} \neq 0$, even with exact differentiation; FO and prefix-3 can introduce $e_{p,r}$ under any allocation. The two vectors can reinforce or cancel, and fixed supervision can raise or lower variance relative to Uniform or SHML. Its usefulness for the reference objective depends on the resulting alignment, curvature, and cost.

5.3 Why fidelity alone does not order progress per unit cost

Consider an SGD surrogate at a common state with reference gradient $g = \nabla F$, estimator mean m_a, covariance Σ_a, step size η, and positive mean update cost c_a. For a quadratic reference objective with Hessian $\mathcal{H}$, expected one-update progress divided by mean cost is exactly

$$\mathcal{P}_a = \frac{\eta \langle g, m_a \rangle - \frac{\eta^2}{2} \left(m_a^\top \mathcal{H} m_a + \mathrm{tr}(\mathcal{H}\Sigma_a) \right)}{c_a}. \tag{7}$$

Write $G_a = c_a \mathcal{P}_a$ for expected one-update progress. If the reference has $G_{\mathrm{U}} > 0$, then $\mathcal{P}_a > \mathcal{P}_{\mathrm{U}}$ exactly when $G_a / G_{\mathrm{U}} > c_a / c_{\mathrm{U}}$: a cheaper direction can win while making less progress per update

when its cost reduction outweighs its progress reduction. For exact gradients on a convex quadratic, same-target sampling loses progress through horizon variance; fixed targets also change alignment. Appendix A.4 makes these conditions explicit, and Appendix A.5 separates target displacement from transient progress along a finite trajectory.

From fixed choices to sampling probabilities. A fixed nonuniform sampling distribution p offers a broader allocation family. Using the sampled loss directly optimizes F_p; weighting it by q_J/p_J preserves F_q when $p_h > 0$ wherever $q_h > 0$. These choices trade target bias against sampling variance and cost. For clipped, projected Adam, the relevant progress depends on the full parameter displacement (Appendix A.3); the reference comparisons below provide a starting point for studying that allocation problem.

6 Benchmark Construction and Protocol

The benchmark holds the evaluation path fixed, measures complete training updates, and crosses allocation with differentiation under a common recipe.

6.1 Datasets and comparison matrix

We use CIFAR-FS (Bertinetto et al., 2019) and MiniImageNet (Ravi & Larochelle, 2017), with 64/16/20-class train/validation/test splits, 5-way 1-shot tasks, and 15 query examples per class. Formal evaluation uses test classes; validation preflight checks feasibility. Methods share initialization and task/augmentation streams within each model setting; SHML uses a separate horizon RNG. Four tasks per update make equal-update comparisons equal in task exposure.

Five training seeds cover 190 trajectories and 810 observations, sharing a 600-episode test bank per dataset. Tables report seed means and sample SDs; contrasts pair methods within seed. Appendix B lists seeds, execution details and uncertainty calculations.

6.2 Resource and update controls

Requested time observations are 300, 900 and, for exact Conv4, 1,800 seconds. The synchronized timer includes sampling, augmentation, transfer, forward/backward computation, clipping, Adam and projection. Preparation, preflight, checkpoint I/O and evaluation are accounted for separately. Equation 3 defines threshold crossing; actual time and overshoot accompany each comparison. Checkpoints at 100 and 200 updates provide early equal-exposure controls. Training processes record peak allocated CUDA tensor memory; profiling and evaluation run separately. Jobs ran serially in shuffled order on the same NVIDIA L4 with shared numerical settings.

6.3 Metrics and comparisons

The primary metric, q_{10} NLL, averages query NLL equally over episodes and steps 1–10. Accuracy, Brier score and step-10 performance are secondary. The main endpoint is exact Conv4/full at 1,800 seconds. For dataset d and seed s, the primary comparison family includes every non-Uniform allocation:

$$\Delta^{a}_{d,s} = \widehat{R}^{a,\mathrm{exact}}_{d,s}(1800) - \widehat{R}^{\mathrm{U},\mathrm{exact}}_{d,s}(1800), \quad a \in \{\mathrm{S}, \mathrm{fixed3}, \mathrm{fixed5}, \mathrm{fixed10}\}. \tag{8}$$

We also report fixed-policy minus SHML contrasts, $\Delta^{\mathrm{fixed}k}_{d,s} - \Delta^{\mathrm{S}}_{d,s}$. Derivative and model/scope comparisons use 900 seconds. Prediction metrics come from saved logits and resource metrics from logs. Reweighting, half-budget, clipping, task-breadth, depth-minimum and scale diagnostics are exploratory analyses of these records, retaining training seeds as the unit of replication.

Selection and benchmark reuse. Checkpoints follow scheduled times or update counts; we report the complete grid under one recipe. Lowest mean denotes the best observed candidate, not validation-selected performance. Extensions should retain both Uniform and SHML, pair task and evaluation streams, and report allocation-selection cost separately (Appendix B.6).

Table 1: Reference suite on both datasets. All five denotes Uniform, SHML and fixed-3/5/10. Exact fixed-3 supplies the identical prefix-3 control. Budgets are training-update seconds, not total job elapsed time.

Model	Adaptation	Derivative	Policies	Budget (s)
Conv4	full	exact	All five	1,800
Conv4	full	FO	All five	900
Conv4	full	prefix-3	Uniform, SHML, fixed-5/10	900
ResNet12	head	exact	All five	900

7 Benchmark Results

At the main 1,800-second exact Conv4/full endpoint, sampled and intermediate fixed supervision achieve lower mean full-path NLL than dense Uniform, with fixed-5 attaining the lowest mean on both datasets. The following comparisons examine how the ranking changes with the derivative rule, update count, model and adaptation scope, and prediction depth.

7.1 Sampled supervision improves performance at equal time

At 1,800 seconds, SHML lowers mean q_{10} NLL relative to Uniform by 0.0476 on CIFAR-FS and 0.0681 on MiniImageNet, with mean accuracy gains of 2.61 and 3.66 percentage points (Table 2). NLL improves in all five seeds on both datasets. SHML completes 2.04 and 3.04 times as many updates, respectively. The mean NLL advantage is also present at 300 and 900 seconds (Figure 1). Recorded peak CUDA tensor allocation is also 67.5%/75.3% lower; fixed-10 achieves similar memory savings without shortening the support path (Appendix D.3).

7.2 Intermediate fixed targets lower full-path risk

Fixed-5 achieves the lowest mean NLL on both datasets at the main endpoint, improving over SHML by 0.0411 on CIFAR-FS and 0.0187 on MiniImageNet. Both fixed-3 and fixed-5 have lower aggregate NLL than Uniform in every seed. Fixed-5 also improves over SHML in every seed; fixed-3 does so in four of five seeds on each dataset. Mean accuracy favors fixed-5 on CIFAR-FS and fixed-3 on MiniImageNet. Fixed-5 lowers task-averaged path NLL relative to SHML on 67.5%/59.1% of episodes per seed on average, with positive median gains in every seed (Appendix D.1).

Half-budget checkpoints exceed the full-budget reference. At 900 seconds, both intermediate targets already have lower mean NLL than Uniform at 1,800 seconds. Each target improves NLL in 5/5 MiniImageNet seeds and 3/5 CIFAR-FS seeds; both also improve accuracy in all MiniImageNet

Table 2: Exact Conv4/full at the 1,800-second requested checkpoint. NLL and accuracy average steps 1–10 and report mean ± sample SD over five training seeds. Updates and actual seconds are seed means; actual seconds include threshold overshoot. Bold marks the best mean NLL and accuracy within each dataset.

Dataset	Policy	q_{10} NLL	q_{10} acc. (%)	Updates	Actual s
CIFAR-FS	Uniform	1.2410 ± 0.0211	51.16 ± 1.19	2,184	1800.54
	SHML	1.1934 ± 0.0207	53.78 ± 1.12	4,447	1800.22
	Fixed-3	1.1728 ± 0.0130	55.55 ± 0.57	7,619	1800.14
	Fixed-5	$\mathbf{1.1523} \pm 0.0097$	$\mathbf{55.84} \pm 0.25$	4,828	1800.23
	Fixed-10	1.2547 ± 0.0150	52.56 ± 0.85	2,548	1800.35
MiniImageNet	Uniform	1.4031 ± 0.0088	41.93 ± 0.53	1,233	1800.66
	SHML	1.3350 ± 0.0137	45.59 ± 0.46	3,743	1800.24
	Fixed-3	1.3221 ± 0.0145	$\mathbf{46.75} \pm 0.55$	5,743	1800.10
	Fixed-5	$\mathbf{1.3163} \pm 0.0069$	46.43 ± 0.28	4,020	1800.20
	Fixed-10	1.3679 ± 0.0214	44.34 ± 1.20	2,308	1800.41

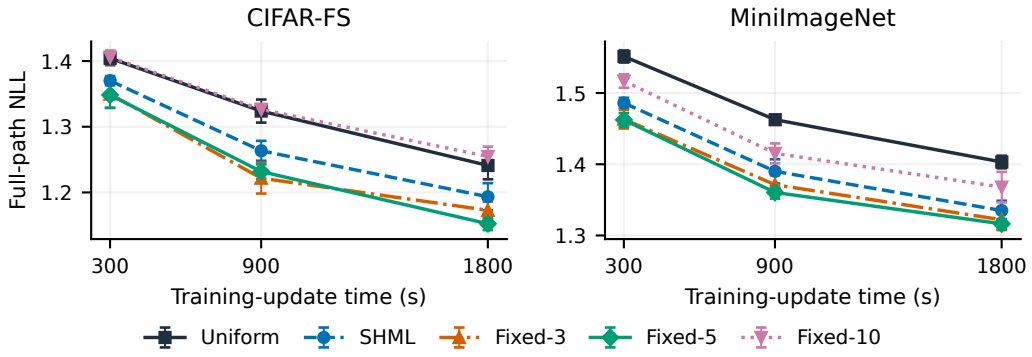

Figure 1: Allocation performance at equal training-time budgets. Exact Conv4/full; markers show five-seed means and bars show ± one sample SD. The 300/900/1,800-second requested observations are plotted at mean actual crossing times, in native seconds. Lines connect saved observations.

seeds. Table 19 compares all five policies against that reference, at the saved 900- and 1,800-second checkpoints.

The paired 95% t interval for fixed-5 minus SHML NLL is $[-0.0688, -0.0134]$ on CIFAR-FS and $[-0.0428, 0.0054]$ on MiniImageNet. The latter has a favorable direction in 5/5 seeds but an interval spanning zero. Appendix C.1 reports all seed values, paired intervals, and multiplicity-adjusted primary comparisons.

Fixed-5's seed-average NLL improves over SHML at all ten depths on CIFAR-FS and nine on MiniImageNet (Figure 2); mean accuracy improves at all depths on both. The CIFAR-FS advantages hold within every seed; MiniImageNet's individual depth profiles vary (Appendix C.7). Fixed-10, in contrast, has worse mean aggregate NLL than Uniform on CIFAR-FS but better NLL on MiniImageNet. Matching the full training target and achieving low held-out risk at a deadline are distinct optimization criteria (Section 5.2).

Depth reweighting tests the breadth of the advantage. For common, task-independent weights over steps 1–10, the worst seed-mean NLL gain equals the minimum depth-wise gain (Appendix A.6). Fixed-5's advantage over SHML persists under all such weights on CIFAR-FS; MiniImageNet's small step-1 deficit allows a reversal under sufficiently early-weighted evaluation. Table 15 reports all main pairs.

7.3 The derivative rule changes allocation contrasts

Table 3 reports absolute 900-second outcomes. The lowest observed mean is fixed-5 with FO on CIFAR-FS (1.1505) and fixed-5 with prefix-3 on MiniImageNet (1.3494). Fixed-5 has lower mean NLL than SHML under every derivative on both datasets, but paired seed win counts range from 2/5 to 5/5 (Appendix D.4).

FO improves every allocation over exact differentiation in all five CIFAR-FS seeds. On MiniImageNet it improves SHML's mean NLL by 0.0121, but worsens Uniform by 0.0783 and fixed-5 by 0.0162. Fixed-5's relative advantage over Uniform grows even while its absolute NLL worsens. Thus relative interactions and absolute outcomes answer different questions; joint selection requires the complete grid. Table 9 reports paired changes in allocation contrasts.

7.4 Early equal-update comparisons yield different rankings

At 100/200 updates under exact Conv4/full, Uniform has lower mean NLL than SHML, fixed-3 and fixed-5 on both datasets. Their ordering reverses at each fixed-time observation (Table 11 and Figure 1). On MiniImageNet, Uniform beats both intermediate targets at both early counts in every seed. Dense averaging can therefore be competitive per update early in training even when cheaper policies lead at a time deadline.

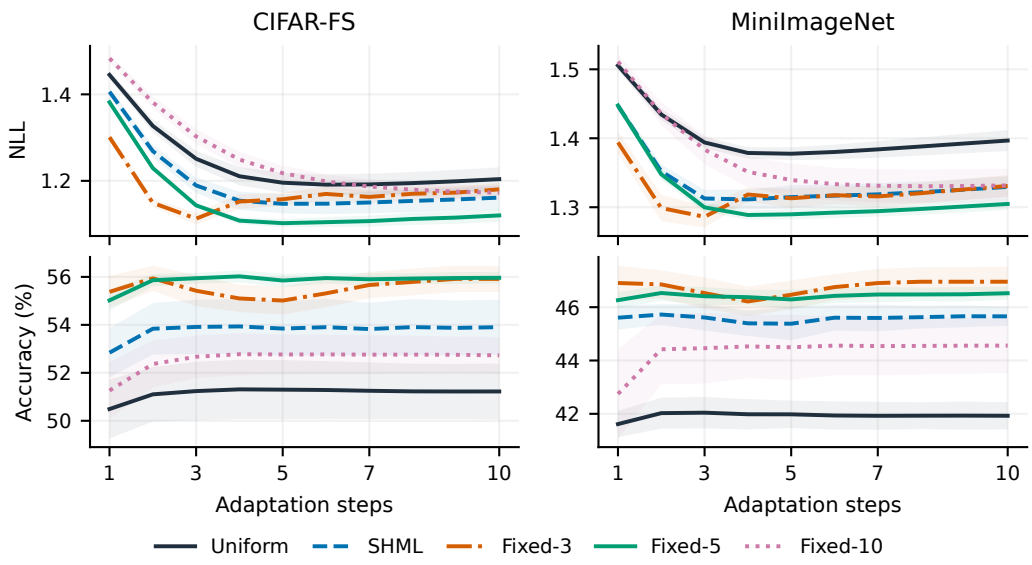

Figure 2: Full adaptation profiles at the exact Conv4/full 1,800-second checkpoint. Curves show five-seed means on a shared 600-episode test bank; shading is ± one sample SD across seeds.

Table 3: Conv4/full q_{10} NLL at 900 seconds. Seed mean on the first line, ± sample SD on the second; bold marks the lowest mean per dataset. Prefix-3 fixed-3 reuses the identical exact fixed-3 run.

Dataset	Derivative	Uniform	SHML	Fixed-3	Fixed-5	Fixed-10
CIFAR-FS	Exact	1.3239 ±0.0176	1.2634 ±0.0151	1.2217 ±0.0234	1.2317 ±0.0111	1.3264 ±0.0080
	FO	1.2639 ±0.0158	1.1857 ±0.0168	1.1749 ±0.0241	**1.1505** ±0.0102	1.2286 ±0.0152
	Prefix-3	1.2642 ±0.0129	1.2284 ±0.0110	1.2217 ±0.0234	1.2055 ±0.0263	1.2436 ±0.0084
MiniImageNet	Exact	1.4625 ±0.0048	1.3901 ±0.0169	1.3712 ±0.0132	1.3604 ±0.0086	1.4153 ±0.0140
	FO	1.5408 ±0.0154	1.3780 ±0.0044	1.3766 ±0.0098	1.3766 ±0.0216	1.3968 ±0.0096
	Prefix-3	1.4795 ±0.0143	1.3603 ±0.0162	1.3712 ±0.0132	**1.3494** ±0.0130	1.3847 ±0.0083

Fixed-5 completes about 1.07–1.09 times as many updates as SHML at the main endpoint (Table 16). On MiniImageNet, its NLL is higher at 200 updates but lower at 1,800 seconds in every seed; CIFAR-FS has mixed early orderings. Fixed-3 completes 1.58 and 1.43 times as many updates as fixed-5 on CIFAR-FS and MiniImageNet, yet has higher mean endpoint NLL on both. Throughput alone does not order endpoint NLL; the early controls do not identify the source of the final fixed-5/SHML gap.

7.5 Model and adaptation scope change the allocation pattern

At 900 seconds in ResNet12/head, SHML's update-count ratio to Uniform is 1.063 on CIFAR-FS and 1.003 on MiniImageNet, down from 2.034 and 3.028 in Conv4/full. Its mean NLL differences from Uniform are +0.0259 and +0.0167, respectively (Table 13). The dense–sampled ranking therefore reverses in a setting with little realized cost advantage from sampling.

Fixed-3 also reverses its NLL contrast with Uniform in every seed on both datasets. Moving from Conv4/full to ResNet12/head, the mean differences change from -0.1022 to +0.0454 on CIFAR-FS and from -0.0913 to +0.0813 on MiniImageNet (Table 13).

Clipping rates, per-task counts and scaling under the shared recipe are reported in Appendix C.9.

7.6 Allocation shapes probability quality along adaptation

Full-path scores connect allocation to prediction depth. At the main endpoint, CIFAR-FS fixed-3 has lower NLL than fixed-5 at every step 1–3 in every seed, while fixed-5 is better at every step 4–10. The mean contrasts are -0.0640 and $+0.0567$ over these blocks (Figure 3). Assign probability ρ uniformly across steps 1–3 and $1-\rho$ uniformly across 4–10. Fixed-3 then has lower seed-mean NLL when $\rho > 0.470$, whereas uniform evaluation uses $\rho = 0.3$ and favors fixed-5. This selects between the existing models; individual-seed boundaries range from 0.341 to 0.677 (Appendix D.2).

Continuing fixed-3 from step 3 to 10 improves accuracy but worsens NLL and Brier in all five seeds on both datasets. Mean NLL rises by 0.0677/0.0447, alongside accuracy gains of 0.49/0.43 percentage points. This motivates a scale diagnostic (Guo et al., 2017): fit one positive temperature per model and depth on 300 episodes, score the other 300, then swap. The split is shared across methods and seeds; these are exploratory diagnostics within the existing test bank.

Temperature adjustment reduces fixed-3's late NLL penalty without changing predicted classes, while fixed-5 retains its full-path gain over SHML in every seed (Table 4). One shared temperature across the path preserves those gain signs, but minimum locations and the closer fixed-3/fixed-5 ranking are more scale-sensitive. Appendices E–F report minimum-depth ordering, the shared-path control and a centered-logit decomposition.

Configuration implications. Specify prediction-depth weights first. At the intended training budget, compare allocation–derivative pairs under one recipe on validation tasks, retaining dense and sampled references. Select using the weighted path score, inspect accuracy alongside probability quality, and report selection cost separately.

8 Scope and Limitations

The benchmark covers two datasets, one shared recipe and fixed candidates 3/5/10; five seeds quantify training variation on a common test bank. The Conv4/full–ResNet12/head comparison jointly changes architecture, adaptation scope and caching. Equal-update controls cover 100/200 updates, not the final endpoint. The analysis gives common-state conditions; gradient bias, curvature, variance and Adam displacements remain to be measured to test their explanatory role in the observed rankings. Depth reweighting uses empirical means with common, task-independent weights. Raw-score minima are restricted to steps 1–10, and temperature fits are exploratory diagnostics within the existing test bank.

9 Conclusion

Training-target fidelity alone does not determine fixed-budget performance: intermediate fixed targets improve mean full-path NLL over dense and sampled supervision in exact Conv4/full adaptation. Allocation and differentiation interact, so absolute outcomes of their combinations matter more than a contrast with one reference. Prediction-depth priorities can reverse model preference, while accuracy gains can coexist with probability degradation. The resulting practice is to specify the evaluation path, compare allocation–derivative pairs at the intended budget, and select on validation probability scores alongside accuracy.

Table 4: Two-fold temperature diagnostic at the main endpoint; five-seed mean NLL. Positive late changes mean step 10 is worse than step 3; positive method gains favor fixed-5. T is depth-specific.

	Fixed-3: step 10 minus 3		SHML minus fixed-5: path NLL	
Dataset	Raw	After T	Raw	After T
CIFAR-FS	$+0.06766$	$+0.00816$	$+0.04109$	$+0.04466$
MiniImageNet	$+0.04473$	$+0.00022$	$+0.01868$	$+0.02064$

Reproducibility Statement

Appendix A provides the analytical derivations. Appendix B details the experimental configuration, data processing, evaluation protocol, and timing procedure. Appendices D–F detail the prediction and scale diagnostics. Anonymized code and experiment configurations are provided as supplementary material with the OpenReview submission.

AI Use Statement

Generative AI tools assisted with translation, verification of experimental results, and code development and validation. They also supported literature review, experimental design, development of the analytical framework and derivations, result interpretation, and manuscript, table, and figure preparation. The authors reviewed and checked every AI-assisted step and its outputs and take full responsibility for the final methodology, implementation, claims, and manuscript.

References

Antreas Antoniou, Harrison Edwards, and Amos Storkey. How to train your MAML. In *International Conference on Learning Representations*, 2019. URL https://arxiv.org/abs/1810.09502.

Alex Beatson and Ryan P. Adams. Efficient optimization of loops and limits with randomized telescoping sums. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 534–543, 2019. URL https://proceedings.mlr.press/v97/beatson19a.html.

Luca Bertinetto, João F. Henriques, Philip H. S. Torr, and Andrea Vedaldi. Meta-learning with differentiable closed-form solvers. In *International Conference on Learning Representations*, 2019.

Wei-Yu Chen, Yen-Cheng Liu, Zsolt Kira, Yu-Chiang Frank Wang, and Jia-Bin Huang. A closer look at few-shot classification. In *International Conference on Learning Representations*, 2019.

Xinshi Chen, Hanjun Dai, Yu Li, Xin Gao, and Le Song. Learning to stop while learning to predict. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 1520–1530, 2020. URL https://proceedings.mlr.press/v119/chen20c.html.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1126–1135, 2017. URL https://proceedings.mlr.press/v70/finn17a.html.

Ishaan Gulrajani and David Lopez-Paz. In search of lost domain generalization. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=lQdXeXDoWtI.

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

JunHoo Lee, Jayeon Yoo, and Nojun Kwak. SHOT: Suppressing the hessian along the optimization trajectory for gradient-based meta-learning. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/c1cdf3236050ad902c6581458e55f0c5-Abstract-Conference.html.

Mengtian Li, Ersin Yumer, and Deva Ramanan. Budgeted training: Rethinking deep neural network training under resource constraints. In *International Conference on Learning Representations*, 2020.

Xu Luo, Hao Wu, Ji Zhang, Lianli Gao, Jing Xu, and Jingkuan Song. A closer look at few-shot classification again. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 23103–23123, 2023. URL https://proceedings.mlr.press/v202/luo23e.html.

Alex Nichol, Joshua Achiam, and John Schulman. On first-order meta-learning algorithms. *arXiv preprint arXiv:1803.02999*, 2018.

Jaehoon Oh, Hyungjun Yoo, ChangHwan Kim, and Se-Young Yun. BOIL: Towards representation change for few-shot learning. In *International Conference on Learning Representations*, 2021. URL https://openreview.net/forum?id=umIdUL8rMH.

Aniruddh Raghu, Maithra Raghu, Samy Bengio, and Oriol Vinyals. Rapid learning or feature reuse? towards understanding the effectiveness of MAML. In *International Conference on Learning Representations*, 2020.

Sachin Ravi and Hugo Larochelle. Optimization as a model for few-shot learning. In *International Conference on Learning Representations*, 2017.

Mohammad Sadegh Salehi, Subhadip Mukherjee, Lindon Roberts, and Matthias J. Ehrhardt. An adaptively inexact first-order method for bilevel optimization with application to hyperparameter learning. *SIAM Journal on Mathematics of Data Science*, 7(3):906–936, 2025. doi: 10.1137/24M1653513. URL https://epubs.siam.org/doi/10.1137/24M1653513.

Amirreza Shaban, Ching-An Cheng, Nathan Hatch, and Byron Boots. Truncated back-propagation for bilevel optimization. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 1723–1732, 2019. URL https://proceedings.mlr.press/v89/shaban19a.html.

Jongyun Shin, Seungjin Han, and Jangho Kim. Cooperative meta-learning with gradient augmentation. In *Proceedings of the Fortieth Conference on Uncertainty in Artificial Intelligence*, volume 244 of *Proceedings of Machine Learning Research*, pp. 3198–3210, 2024. URL https://proceedings.mlr.press/v244/shin24a.html.

Yuxin Wu and Kaiming He. Group normalization. In *Proceedings of the European Conference on Computer Vision*, pp. 3–19, 2018.

Honglin Yang, Ji Ma, and Xiao Yu. Memory-reduced meta-learning with guaranteed convergence. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 21938–21946, 2025. doi: 10.1609/aaai.v39i20.35501. URL https://ojs.aaai.org/index.php/AAAI/article/view/35501.

Han-Jia Ye and Wei-Lun Chao. How to train your MAML to excel in few-shot classification. In *International Conference on Learning Representations*, 2022.

Yilang Zhang, Abraham Jaeger Mountain, Bingcong Li, and Georgios B. Giannakis. Binomial gradient-based meta-learning for enhanced meta-gradient estimation. In *International Conference on Learning Representations*, 2026. URL https://proceedings.iclr.cc/paper_files/paper/2026/hash/a14ac7911df7d369919a687909a14a8b-Abstract-Conference.html.

A Derivations and Analytical Specializations

A.1 Mean-direction decomposition and sampling

All expectations in this subsection are over training tasks at the same initialization; differentiation is assumed to exist and commute with the expectation where used. For fixed, task-independent horizon weights p,

$$m_{p,r} = \mathbb{E}_\tau \sum_h p_h g_{\tau,h} + \mathbb{E}_\tau \sum_h p_h (\widetilde{g}^r_{\tau,h} - g_{\tau,h}) = \nabla F_p + e_{p,r}.$$

Adding and subtracting ∇F yields Equation 6. This is a decomposition of a direction field; it does not require FO or prefix-3 to be gradients of a separate scalar objective.

For Proposition 1, average over the finite uniform distribution of J at a fixed task. Linearity of the derivative rule through the outer sum gives Equation 4. Let $\mu_\tau = \widetilde{g}^r_{\tau,\mathrm{U}}$ and $\mu = \mathbb{E}_\tau \mu_\tau$. Expanding $\widetilde{g}^r_{\tau,J} - \mu = (\widetilde{g}^r_{\tau,J} - \mu_\tau) + (\mu_\tau - \mu)$ gives

$$\mathbb{E}_{\tau,J} \|\widetilde{g}^r_{\tau,J} - \mu\|^2 = \mathbb{E}_{\tau,J} \|\widetilde{g}^r_{\tau,J} - \mu_\tau\|^2 + \mathbb{E}_\tau \|\mu_\tau - \mu\|^2.$$

The cross term vanishes because $\mathbb{E}_J[\widetilde{g}^r_{\tau,J} - \mu_\tau \mid \tau] = 0$. Independence of tasks and horizon draws divides the variance of their average by N. The task draws include their augmentation randomness.

A.2 Clipped inner differentiation

For a support gradient v with norm $r_v = \|v\|$, away from the clipping boundary and zero, the clipping Jacobian is

$$J_{C_\gamma}(v) = \begin{cases} I, & r_v < \gamma, \\ \dfrac{\gamma}{r_v}\left(I - \dfrac{vv^\top}{r_v^2}\right), & r_v > \gamma. \end{cases} \tag{9}$$

Exact differentiation applies the chain rule through this Jacobian and the support-loss Hessian. FO suppresses the derivative of the entire clipped support gradient; prefix-3 suppresses it only after the first three updates. All rules preserve the direct identity derivative from ϕ_{h-1} to ϕ_h. The implementation stabilizes the norm with a squared-norm floor of 10^{-24}; at the configured threshold $\gamma = 5$, that floor lies in the unclipped branch. Nonsmooth boundaries use the automatic-differentiation convention and are not covered by a globally smooth guarantee.

A.3 One-update progress

At a fixed state, consider the SGD surrogate $\theta^+ = \theta - \eta \widehat{g}_a$, with $\mathbb{E}\widehat{g}_a = m_a = g + \delta_a$, $g = \nabla F(\theta)$, and $v_a = \mathbb{E}\|\widehat{g}_a - m_a\|^2$, where $\delta_a = b_p + e_{p,r}$. If F has L-Lipschitz gradient along the update segments, then

$$\mathbb{E}[F(\theta) - F(\theta^+)] \geq B_a := \eta \langle g, m_a \rangle - \frac{L\eta^2}{2}(\|m_a\|^2 + v_a), \tag{10}$$

$$B_a - B_0 = (\eta - L\eta^2)\langle g, \delta_a \rangle - \frac{L\eta^2}{2}\big(\|\delta_a\|^2 + v_a - v_0\big). \tag{11}$$

The reference estimator has mean g and variance v_0. Dividing B_a by mean update cost gives a cost-normalized lower bound; comparing such bounds alone does not order the actual decreases.

The descent lemma gives $F(\theta - \eta\widehat{g}_a) \leq F(\theta) - \eta\langle g, \widehat{g}_a \rangle + (L\eta^2/2)\|\widehat{g}_a\|^2$. Taking expectations and using $\mathbb{E}\|\widehat{g}_a\|^2 = \|m_a\|^2 + v_a$ proves Equation 10. Substitution of $m_a = g + \delta_a$ and subtraction of the zero-bias reference proves Equation 11. For $0 < \eta < 1/L$, alignment contributes positively to that difference.

If F is quadratic, Taylor expansion is exact and

$$\mathbb{E}[\widehat{g}_a^\top \mathcal{H} \widehat{g}_a] = m_a^\top \mathcal{H} m_a + \mathrm{tr}(\mathcal{H}\Sigma_a),$$

which proves Equation 7. Its normalization is expected progress divided by expected cost, not the expectation of a random progress/cost ratio.

For clipped, projected Adam, condition on the full parameter and optimizer state z and define the actual displacement $D_a = \theta - \theta^+$, its mean $\bar{D}_a$, and covariance S_a. Replacing $\eta\widehat{g}_a$ with D_a gives

$$\mathbb{E}[F(\theta) - F(\theta^+) \mid z] \geq \langle g, \bar{D}_a \rangle - \frac{L}{2}\left(\|\bar{D}_a\|^2 + \operatorname{tr} S_a\right). \tag{12}$$

Clipping, Adam's state, and classifier projection are all included in D_a. Equality of raw direction means does not imply equality of $\bar{D}_a$. These calculations concern the training objective; held-out risk also depends on generalization.

For example, in the isotropic quadratic case $\mathcal{H} = LI$, equal costs, and $0 < \eta < 1/L$, the biased estimator provides greater expected one-update progress than the unbiased reference exactly when

$$(1 - L\eta)\langle g, \delta_a \rangle > \frac{L\eta}{2}\left(\|\delta_a\|^2 + v_a - v_0\right). \tag{13}$$

Unequal costs require comparing Equation 7 instead. This condition concerns a state and an update rule, not a horizon index alone.

A.4 A COST–PROGRESS THRESHOLD FOR SAMPLING AND FIXED TARGETS

For the quadratic calculation in Equation 7, compare an exact Uniform reference with an action a at the same state, step size, and reference objective. Let $m_{\mathrm{U}} = g$, $\delta_a = m_a - g$, and $G_a = c_a\mathcal{P}_a$. Expanding the quadratic gives

$$\begin{aligned} G_a - G_{\mathrm{U}} = \eta\langle g, \delta_a \rangle - \eta^2 g^\top \mathcal{H}\delta_a \\ - \frac{\eta^2}{2}\left(\delta_a^\top \mathcal{H}\delta_a + \operatorname{tr}\big[\mathcal{H}(\Sigma_a - \Sigma_{\mathrm{U}})\big]\right). \end{aligned} \tag{14}$$

Since both costs are positive,

$$\mathcal{P}_a > \mathcal{P}_{\mathrm{U}} \quad \Longleftrightarrow \quad G_a - G_{\mathrm{U}} > \left(\frac{c_a}{c_{\mathrm{U}}} - 1\right) G_{\mathrm{U}}. \tag{15}$$

If $G_{\mathrm{U}} > 0$, this is the progress-retention condition $G_a/G_{\mathrm{U}} > c_a/c_{\mathrm{U}}$ in the main text. Equation 14 identifies what the cost saving must compensate: changed alignment and curvature terms, as well as stochastic variation. For exact fixed-k, $\delta_a = b_{\delta_k}$; derivative approximation additionally contributes $e_{p,r}$.

For exact SHML, $\delta_{\mathrm{S}} = 0$ and the matrix law of total covariance gives $\Sigma_{\mathrm{S}} - \Sigma_{\mathrm{U}} = C_{\mathrm{hor}}$, where

$$C_{\mathrm{hor}} = \frac{1}{N}\mathbb{E}_\tau \operatorname{Cov}_J(g_{\tau,J} \mid \tau) \succeq 0.$$

Thus $G_{\mathrm{S}} = G_{\mathrm{U}} - \Omega$, with $\Omega = (\eta^2/2)\operatorname{tr}(\mathcal{H}C_{\mathrm{hor}})$. If $\mathcal{H} \succeq 0$, this is a nonnegative averaging penalty, and

$$\mathcal{P}_{\mathrm{S}} > \mathcal{P}_{\mathrm{U}} \quad \Longleftrightarrow \quad \Omega < \left(1 - \frac{c_{\mathrm{S}}}{c_{\mathrm{U}}}\right) G_{\mathrm{U}}. \tag{16}$$

At equal costs, sampling cannot improve this local quadratic SGD comparison; as its relative cost decreases, a larger variance penalty can be tolerated when $G_{\mathrm{U}} > 0$. For indefinite curvature, Ω need not be nonnegative, although the algebraic identity still holds.

This threshold suggests measuring both relative update cost and retained progress, rather than selecting a horizon from bias or cost alone. The benchmark records trajectory-averaged costs and held-out outcomes, not G_a at a shared state. Its cached-head comparison (Table 13) supplies an observed small-cost-gap regime, not an estimate of Ω or a test of the local SGD identity for Adam.

A.5 AN EXACT FINITE-BUDGET SPECIALIZATION

Consider the scalar reference risk $F(x) = \lambda x^2/2$, $\lambda > 0$. An action uses the noisy gradient of a displaced quadratic, $\widehat{g}_{a,u} = \lambda(x_{a,u} - d_a) + \epsilon_{a,u}$, with independent, zero-mean noise of variance σ_a^2, deterministic update cost $c_a > 0$, and a shared deterministic initialization x_0. Take $0 < \eta\lambda < 2$

and $r = 1 - \eta\lambda$, so $|r| < 1$. For a constant action, the benchmark's first-crossing clock gives $n_a(T) = \lceil T/c_a \rceil$ for $T > 0$. Then

$$\mathbb{E}F(x_{a,n_a(T)}) = \frac{\lambda}{2}\left\{\left[r^{n_a(T)}x_0 + (1 - r^{n_a(T)})d_a\right]^2 + \frac{\eta^2\sigma_a^2(1 - r^{2n_a(T)})}{1 - r^2}\right\}. \tag{17}$$

To prove the identity, unroll $x_{a,u+1} = rx_{a,u} + (1 - r)d_a - \eta\epsilon_{a,u}$. Its mean is $r^u x_0 + (1 - r^u)d_a$ and variance is $\eta^2\sigma_a^2 \sum_{j=0}^{u-1} r^{2j}$. Using $\mathbb{E}x^2 = (\mathbb{E}x)^2 + \mathrm{Var}(x)$ yields Equation 17.

The formula separates three effects: $n_a(T)$ controls how much initial-state error is contracted; d_a changes the target and can either offset or amplify the transient mean; and σ_a^2 accumulates stochastic variation. An inexpensive biased action can therefore have lower finite-budget risk, although its limiting risk is

$$\lim_{T\to\infty} \mathbb{E}F(x_{a,n_a(T)}) = \frac{\lambda}{2}\left(d_a^2 + \frac{\eta^2\sigma_a^2}{1 - r^2}\right).$$

Whether a crossover occurs depends on the action-specific values of c_a, d_a, and σ_a^2. Applying this expression to horizon selection would require estimating their counterparts along the training trajectory. The expression compares constant actions; adaptive allocation additionally requires accounting for how each action changes future states.

A.6 Reweighting a fixed evaluation path

Fix two trained models a, b and a common bank of evaluation tasks. Let $\widehat{L}_h^a$ be model a's mean NLL at depth h and define the improvement $\delta_h^{a,b} = \widehat{L}_h^b - \widehat{L}_h^a$. A common, task-independent stopping distribution $w \in \Delta_H$ gives $\widehat{R}_w^b - \widehat{R}_w^a = \sum_h w_h \delta_h^{a,b}$. Since this is a convex combination,

$$\min_h \delta_h^{a,b} \le \widehat{R}_w^b - \widehat{R}_w^a \le \max_h \delta_h^{a,b}.$$

Both endpoints are attained by putting all mass on an extremizing depth. Consequently, the empirical ranking favors a strictly for every w exactly when $\min_h \delta_h^{a,b} > 0$. If the minimum is negative and the maximum is positive, two stopping distributions give opposite rankings. The same argument applies to accuracy with improvement defined as $\widehat{A}_h^a - \widehat{A}_h^b$. It also applies to profiles averaged across training seeds, by linearity. Dominance of that average does not imply dominance within every seed.

This calculation reweights the existing depth means; it does not retrain either model under a different prior. The extrema are empirical margins, not confidence bounds on population risk. Task-adaptive stopping requires task-conditioned outcomes and is outside this common-weight calculation.

A.7 What the saved benchmark observations identify

Table 5: Connection between analysis and measurement. *Not recorded* denotes an absent measurement.

Quantity or question	Available comparison or artifact	Interpretation
Final held-out utility	q_{10} NLL and full 1–10 curves	Primary measured outcome
Stopping-depth reweighting	Extrema of paired depth-mean differences	Empirical ranking across common depth weights
Update cost and exposure	Actual seconds, updates, tasks, support steps, query counts	Realized allocation and throughput
Allocation–derivative interaction	Common 900 s paired differences-in-differences	Change in the allocation contrast
Early update-level behavior	100/200 updates with paired task streams	Equal early task exposure
Cumulative clipping exposure	Inner-step and outer-update counts and scale sums (Table 17)	Observed operation of shared thresholds along each trajectory
$b_p, e_{p,r}, \Sigma_a, \mathcal{H}$	Not recorded at formal checkpoints	No direct mechanism estimate
Full Adam displacement distribution	Not recorded over repeated tasks at a fixed state	No optimizer-level unbiasedness test

B Full Protocol and Accounting

B.1 Shared configuration

Table 6: Executed common recipe.

Item	Value
Datasets and class split	CIFAR-FS, MiniImageNet; 64/16/20
RGB image resolution	CIFAR-FS: 32×32; MiniImageNet: 84×84
Episode	5-way, 1-shot, 15 query examples/class
Training support steps	Uniform: 10; SHML: sampled $J \in \{1, \dots, 10\}$; fixed-k: k
Training query-loss depths	Uniform: 1–10; SHML: J only; fixed-k: k only
Fixed-horizon candidates	$k \in \{3, 5, 10\}$
Evaluation depths	$h \in \{1, \dots, 10\}$ for every trained model
Meta-batch	4 tasks
Inner learning rate / norm clip	0.01 / 5
Outer Adam learning rate / norm clip	0.001 / 1
Adam defaults	$\beta_1 = 0.9,\ \beta_2 = 0.999,\ \epsilon = 10^{-8}$; no weight decay
Time observations	300/900/1,800 s as allowed by trajectory budget
Update observations	100, 200
Evaluation bank	600 test episodes per dataset, shared across methods
Training seeds	3403801132, 4001970766, 1818641017, 3515040331, 3205086188
Evaluation seeds	CIFAR-FS: 2257402719; MiniImageNet: 300878434
Queue seed	3906170466
Precision	float32
Hardware	NVIDIA L4, four CPU threads
Runtime	Python 3.12.3, PyTorch 2.9.1+cu129, CUDA 12.9

Evaluation seeds fix the dataset-specific episode banks shared across methods and training seeds. The queue seed, combined with each training seed, randomizes the execution order of training jobs.

Episodes sample five classes without replacement and, within each class, one support and fifteen disjoint query images without replacement. RGB values are converted from uint8 to float32 and divided by 255. Channel-wise normalization uses CIFAR-FS mean $(0.5071, 0.4865, 0.4409)$ and standard deviation $(0.2673, 0.2564, 0.2762)$; MiniImageNet uses mean $(0.485, 0.456, 0.406)$ and standard deviation $(0.229, 0.224, 0.225)$.

Before normalization, training applies reflection-padded random crops and horizontal flips, with padding 4 for CIFAR-FS and 8 for MiniImageNet. Each image is augmented once per task and reused throughout adaptation. Evaluation uses no random augmentation. Conv4 has four 64-channel blocks; ResNet12 has widths 64/160/320/640. Both use stateless eight-group Group Normalization. Shared-classifier-row projection and fixed learning rates are common to all policies. The same inner learning rate is used at every adaptation step, including evaluation beyond a fixed training horizon.

Inner clipping uses the global norm over adapted parameters at every support step in training and evaluation. Outer clipping uses the global norm after averaging four task directions, before Adam. Exact differentiation includes the inner clipping scale.

B.2 Observation counts and resource use

Per dataset, the matrix contains five exact Conv4 trajectories, five FO, four prefix-3, and five exact ResNet12/head trajectories, totaling 19. Each exact Conv4 trajectory has three time labels and two update labels; other trajectories have two of each. Thus $25 + 20 + 16 + 20 = 81$ observations per dataset give 38 training trajectories and 162 observations per seed. Across five seeds this is 190 trajectories and 810 observations. Every observation replays 600 episodes, for 486,000 checkpoint–episode evaluations. A shared episode or checkpoint is not an independent training run.

The nominal training budget per seed is $2[5(1800) + (5+4+5)(900)] = 43{,}200$ seconds. Across five seeds, recorded training time is 216,051.87 seconds, including final threshold overshoots. Evaluation

uses 37,337.70 seconds and recorded checkpoint writes 834.39 seconds. Preparation, profiling, and process startup are additional costs, not charged to the training-update axis.

B.3 PREFLIGHT AND EXECUTION

Preflight profiles training updates on training tasks and evaluation on validation tasks. Its feasibility check is $200c_{\max} \leq 0.8B_{\min}$, where $c_{\max}$ is the slowest profiled update within a dataset/model/scope group and $B_{\min}$ is that group's shortest training budget. The training and evaluation jobs completed without a reported failure. The 1,190 recorded profiling, training, and evaluation process intervals do not overlap, including across seeds. Job order is shuffled using a stream separate from training; task exposure is paired by update number, not by wall-clock time.

B.4 RESULT VERIFICATION

We independently recomputed NLL, accuracy, Brier score, confidence, and prediction turnover from each saved logits array, with shape $600 \times 10 \times 75 \times 5$. Labels, predictions, horizon indices, episode indices, bank hashes, and support/query disjointness were checked. All per-depth values and available aggregate summaries agree within 3.8×10^{-15} absolute error. Each evaluation's update count and training time also match its saved training record. Source files and episode banks match across the five exports; resolved configurations differ only in training seed and output path, and the recorded hardware and runtime agree. All five configured training repetitions are included.

The export contains the executed configuration, environment, source snapshot, episode banks, prediction arrays, and training/evaluation records. It does not contain model weights or image pixels, and its evaluation identity does not bind a checkpoint-content hash. This verification therefore establishes consistency of exported predictions and metrics, not fresh inference from the original checkpoint.

B.5 METRIC AND CONTRAST DEFINITIONS

NLL uses natural logarithms; accuracy is reported in percent and differences in percentage points. Multiclass Brier sums squared probability errors over classes. All metrics average queries within an episode, then episodes; q_{10} gives equal weight to depths 1–10, while h_{10} uses only depth 10. For a paired comparison, the reported effect is

$$d_s^{a,b} = \frac{1}{600} \sum_{e=1}^{600} \left[\frac{1}{10} \sum_{h=1}^{10} (\ell_{s,e,h}^a - \ell_{s,e,h}^b) \right].$$

We average these differences across the five independent training seeds. Tables report seed means and sample standard deviations (denominator $S - 1$, $S = 5$). Paired 95% Student-t intervals use $\bar{d} \pm t_{0.975,4}\, s_d/\sqrt{5}$, with $t_{0.975,4} \approx 2.776$. These intervals describe training-seed variation conditional on the fixed test episode bank, not population uncertainty over new tasks; their calibration assumes approximately normal seed-level differences.

For the eight primary NLL contrasts in Equation 8 (four policies times two datasets), we additionally report Holm-adjusted two-sided paired t-test p values as an uncertainty summary. The displayed 95% intervals are marginal, not simultaneous. Fixed-policy versus SHML contrasts, derivative interactions, and depth reweighting are supporting analyses, without multiplicity-adjusted significance claims. Episode or depth counts are never used as the training replication count.

B.6 REUSING AND EXTENDING THE BENCHMARK

Comparison unit. A benchmark setting comprises the dataset and class split, episode shape, model and adaptation scope, common optimizer recipe, derivative rule, evaluation path, and runtime environment. A supervision-allocation comparison holds these fixed and changes the allocation policy. The reference suite in Table 1 supplies Uniform, SHML, and fixed-step comparisons; extensions should retain both same-target references rather than selecting a reference after observing the new method's score.

We report all configured candidates (3/5/10) under the common recipe, with uniform evaluation weights and scheduled checkpoints. Validation preflight checks feasibility rather than allocation quality.

Resource and selection rules. Use the synchronized training-update timer and first-crossing checkpoint rule in Section 6. Any additional computation used to choose an allocation during training belongs inside that timer. Report setup, profiling, checkpoint writing, and evaluation separately. If an extension tunes a schedule or controller, specify its validation-only selection rule and report the search cost separately from the final training run. Fix the evaluation episode bank and scoring rule before examining test outcomes.

Comparable runs. Pair initialization and task/augmentation streams by training seed and update number within a model setting. Keep allocation-specific randomness separate from task sampling. When hardware, software, or numerical precision changes, rerun the reference procedures in that environment and report it as a separate setting; an equal number of seconds on two GPUs is not a matched compute comparison. Changes to the shared optimizer recipe, adaptation scope, or depth range likewise define a new setting.

Minimum result record. For each observation, report the dataset, model/scope, allocation, derivative, training seed, requested time or update count, actual accumulated training seconds, and completed updates. Retain task, support-step, and query-loss counts to describe resource allocation. Evaluation records should contain the bank identity, episode and depth indices, labels, predictions or logits, and the depth-wise and aggregate metrics defined above. Report all reference procedures at the same requested observations; preserve per-seed records when aggregating independent training runs.

Implementation and artifact boundaries. In the executed source snapshot, `experiment.py` defines the recipe and comparison matrix, `src/model.py` implements adaptation and derivative rules, `src/worker.py` handles training and evaluation, and `src/benchmark.py` coordinates execution. Adding an allocation requires extending these definitions and their tests while preserving the shared evaluation path. The observed numerical identities include identical forward adaptation across derivative rules and exact/prefix-3 equivalence for fixed-3. Section B.4 describes the available result export; replication from model weights additionally requires retaining the original checkpoints and binding their content hashes to evaluation records.

C Supplementary Results

C.1 Training-seed results and paired uncertainty

Table 7: Individual training-seed q_{10} NLL: exact Conv4/full, 1,800 seconds. Every row uses the same dataset-specific 600-episode bank.

Dataset	Training seed	Uniform	SHML	Fixed-3	Fixed-5	Fixed-10
CIFAR-FS	3403801132	1.23952	1.19224	1.16165	1.15700	1.25218
	4001970766	1.27156	1.22481	1.16531	1.15421	1.24761
	1818641017	1.21474	1.19940	1.19245	1.14327	1.26527
	3515040331	1.24828	1.17130	1.16478	1.14210	1.23511
	3205086188	1.23112	1.17931	1.17991	1.16503	1.27326
MiniImageNet	3403801132	1.40176	1.34033	1.32225	1.31396	1.36166
	4001970766	1.41539	1.35544	1.33746	1.30677	1.35443
	1818641017	1.39599	1.32020	1.29835	1.31557	1.39884
	3515040331	1.39416	1.32570	1.32841	1.32489	1.34501
	3205086188	1.40836	1.33341	1.32379	1.32049	1.37933

Table 8: Paired NLL contrasts at the main endpoint; candidate minus reference. Intervals use five seed differences and are marginal 95% Student-t intervals. Favorable counts denote negative differences. Holm adjustment covers only the eight primary contrasts against Uniform; other rows are supporting comparisons.

Dataset	Pair	Mean Δ	95% interval	Favorable	Holm p
CIFAR-FS	SHML / Uniform	-0.04763	[-0.07485, -0.02041]	5/5	0.0332
	Fixed-3 / Uniform	-0.06822	[-0.10832, -0.02813]	5/5	0.0332
	Fixed-5 / Uniform	-0.08872	[-0.11629, -0.06115]	5/5	0.0043
	Fixed-10 / Uniform	+0.01364	[-0.02709, +0.05438]	2/5	0.4051
	Fixed-3 / SHML	-0.02059	[-0.05130, +0.01012]	4/5	–
	Fixed-5 / SHML	-0.04109	[-0.06880, -0.01338]	5/5	–
	Fixed-10 / SHML	+0.06127	[+0.02976, +0.09279]	0/5	–
MiniImageNet	SHML / Uniform	-0.06811	[-0.07726, -0.05897]	5/5	0.0003
	Fixed-3 / Uniform	-0.08108	[-0.09542, -0.06673]	5/5	0.0007
	Fixed-5 / Uniform	-0.08679	[-0.10464, -0.06894]	5/5	0.0010
	Fixed-10 / Uniform	-0.03527	[-0.06548, -0.00506]	4/5	0.0633
	Fixed-3 / SHML	-0.01296	[-0.02518, -0.00075]	4/5	–
	Fixed-5 / SHML	-0.01868	[-0.04279, +0.00543]	5/5	–
	Fixed-10 / SHML	+0.03284	[-0.00507, +0.07076]	1/5	–

The primary contrasts for SHML, fixed-3, and fixed-5 on both datasets have Holm-adjusted $p < 0.05$; neither fixed-10 contrast does. This differs from asking whether the observed direction agrees across seeds: SHML and both intermediate targets have lower NLL than Uniform in 5/5 seeds on both datasets. The fixed-5–SHML MiniImageNet interval includes zero, despite a favorable direction in every seed.

C.2 Direct allocation–derivative comparisons

Table 3 in the main text gives the absolute NLL grid. At the common 900-second point, define the interaction for allocation a and $r \in \{\mathrm{FO}, \mathrm{prefix3}\}$ by

$$I_{d,s}^{a,r} = (\widehat{R}_{d,s}^{a,r} - \widehat{R}_{d,s}^{\mathrm{U},r}) - (\widehat{R}_{d,s}^{a,\mathrm{exact}} - \widehat{R}_{d,s}^{\mathrm{U},\mathrm{exact}}). \tag{18}$$

A positive value makes the allocation's NLL contrast with Uniform less favorable under r. Absolute NLL and accuracy rank the complete procedures; an interaction instead describes a change in the reference contrast.

Table 9: Allocation–derivative interactions at 900 seconds (Equation 18); mean $\pm$ sample SD over five seeds. Positive values indicate a less favorable NLL contrast with Uniform than under exact differentiation.

Dataset	Derivative	SHML	Fixed-3	Fixed-5	Fixed-10
CIFAR-FS	FO	-0.018 ± 0.028	0.013 ± 0.039	-0.021 ± 0.015	-0.038 ± 0.020
	Prefix-3	0.025 ± 0.016	0.060 ± 0.014	0.033 ± 0.035	-0.023 ± 0.011
MiniImageNet	FO	-0.090 ± 0.026	-0.073 ± 0.023	-0.062 ± 0.024	-0.097 ± 0.023
	Prefix-3	-0.047 ± 0.024	-0.017 ± 0.016	-0.028 ± 0.029	-0.048 ± 0.026

Table 10: q_{10} accuracy (%) for the same 900-second comparison; mean $\pm$ sample SD across five seeds.

Dataset	Derivative	Uniform	SHML	Fixed-3	Fixed-5	Fixed-10
CIFAR-FS	Exact	46.90 ± 1.02	50.23 ± 0.83	52.77 ± 1.02	51.79 ± 0.55	48.38 ± 0.73
	FO	49.56 ± 0.84	53.85 ± 0.72	56.05 ± 1.03	55.88 ± 0.49	52.74 ± 0.92
	Prefix-3	49.70 ± 0.60	51.85 ± 0.59	52.77 ± 1.02	53.18 ± 1.39	52.24 ± 0.53
MiniImageNet	Exact	38.40 ± 0.31	42.52 ± 0.76	44.24 ± 0.85	44.03 ± 0.61	41.57 ± 0.65
	FO	33.09 ± 1.25	43.13 ± 0.33	43.95 ± 0.43	43.42 ± 0.93	42.46 ± 0.58
	Prefix-3	37.50 ± 0.89	44.17 ± 0.43	44.24 ± 0.85	44.81 ± 0.38	43.20 ± 0.47

C.3 Early equal-update controls

Table 11: q_{10} NLL at equal early update counts; seed mean on the first line, $\pm$ sample SD on the second. Within each setting, methods share initialization and task exposure.

Dataset	Derivative	Updates	Uniform	SHML	Fixed-3	Fixed-5	Fixed-10
CIFAR-FS	Exact	100	1.4721	1.4764	1.5017	1.4752	1.4864
			$\pm$0.0181	$\pm$0.0135	$\pm$0.0191	$\pm$0.0161	$\pm$0.0134
		200	1.4391	1.4489	1.4757	1.4423	1.4516
			$\pm$0.0120	$\pm$0.0201	$\pm$0.0101	$\pm$0.0117	$\pm$0.0073
	FO	100	1.4907	1.4986	1.5218	1.5024	1.5057
			$\pm$0.0063	$\pm$0.0117	$\pm$0.0121	$\pm$0.0055	$\pm$0.0078
		200	1.4664	1.4742	1.4919	1.4727	1.4684
			$\pm$0.0151	$\pm$0.0113	$\pm$0.0138	$\pm$0.0160	$\pm$0.0095
	Prefix-3	100	1.4732	1.4809	1.5017	1.4770	1.4868
			$\pm$0.0150	$\pm$0.0074	$\pm$0.0191	$\pm$0.0150	$\pm$0.0032
		200	1.4368	1.4536	1.4757	1.4346	1.4541
			$\pm$0.0112	$\pm$0.0054	$\pm$0.0101	$\pm$0.0061	$\pm$0.0104
MiniImageNet	Exact	100	1.5647	1.5804	1.6401	1.5844	1.5646
			$\pm$0.0124	$\pm$0.0073	$\pm$0.0067	$\pm$0.0148	$\pm$0.0087
		200	1.5462	1.5567	1.6110	1.5673	1.5420
			$\pm$0.0080	$\pm$0.0055	$\pm$0.0219	$\pm$0.0070	$\pm$0.0062
	FO	100	1.6037	1.6024	1.6558	1.6148	1.6001
			$\pm$0.0084	$\pm$0.0117	$\pm$0.0318	$\pm$0.0123	$\pm$0.0046
		200	1.5914	1.6008	1.6494	1.6004	1.5961
			$\pm$0.0189	$\pm$0.0145	$\pm$0.0248	$\pm$0.0095	$\pm$0.0049
	Prefix-3	100	1.5835	1.5966	1.6401	1.5925	1.6053
			$\pm$0.0142	$\pm$0.0085	$\pm$0.0067	$\pm$0.0178	$\pm$0.0145
		200	1.5764	1.5931	1.6110	1.5784	1.5982
			$\pm$0.0105	$\pm$0.0068	$\pm$0.0219	$\pm$0.0124	$\pm$0.0110

C.4 SECOND MODEL AND ADAPTATION SETTING

Table 12: ResNet12/head at the requested 900-second checkpoint. NLL and accuracy average steps 1–10 (mean ± sample SD over five seeds). Updates and actual seconds are seed means; seconds include overshoot.

Dataset	Policy	q_{10} NLL	q_{10} acc. (%)	Updates	Actual s
CIFAR-FS	Uniform	1.1568 ± 0.0292	55.83 ± 1.52	3,463	900.13
	SHML	1.1827 ± 0.0322	54.46 ± 1.47	3,683	900.12
	Fixed-3	1.2022 ± 0.0195	55.00 ± 0.75	3,742	900.13
	Fixed-5	1.1497 ± 0.0053	56.12 ± 0.27	3,704	900.08
	Fixed-10	1.1469 ± 0.0227	56.98 ± 1.18	3,648	900.14
MiniImageNet	Uniform	1.4899 ± 0.0162	37.50 ± 0.98	682	900.67
	SHML	1.5067 ± 0.0114	36.53 ± 0.80	684	900.95
	Fixed-3	1.5712 ± 0.0130	35.89 ± 0.71	685	900.74
	Fixed-5	1.5146 ± 0.0307	36.63 ± 1.56	684	900.55
	Fixed-10	1.4816 ± 0.0321	37.66 ± 2.23	683	900.53

C.5 MATCHED-TIME COST AND TARGET CONTRASTS

Table 13 pairs each policy with Uniform within the same dataset, training seed, and model/scope setting at 900 seconds. It retains all four non-Uniform policies. Update ratios are calculated within each seed before averaging; NLL differences use the paired seed outcomes.

Table 13: Matched 900-second comparisons under exact differentiation. Each policy is compared with Uniform in the same row's model/scope setting. Values are mean ± sample SD across five seeds; favorable counts denote negative NLL differences. Architecture, adaptation scope, and feature caching change jointly between settings.

Dataset	Model/scope	Policy	Update ratio	Δ NLL	Favorable
CIFAR-FS	Conv4/full	SHML	2.034 ± 0.022	-0.0605 ± 0.0245	5/5
		Fixed-3	3.485 ± 0.025	-0.1022 ± 0.0337	5/5
		Fixed-5	2.212 ± 0.036	-0.0921 ± 0.0127	5/5
		Fixed-10	1.167 ± 0.007	0.0026 ± 0.0150	3/5
	ResNet12/head	SHML	1.063 ± 0.003	0.0259 ± 0.0316	1/5
		Fixed-3	1.080 ± 0.002	0.0454 ± 0.0260	0/5
		Fixed-5	1.070 ± 0.003	-0.0071 ± 0.0246	3/5
		Fixed-10	1.053 ± 0.001	-0.0099 ± 0.0454	2/5
MiniImageNet	Conv4/full	SHML	3.028 ± 0.036	-0.0725 ± 0.0191	5/5
		Fixed-3	4.660 ± 0.022	-0.0913 ± 0.0096	5/5
		Fixed-5	3.259 ± 0.014	-0.1021 ± 0.0081	5/5
		Fixed-10	1.871 ± 0.012	-0.0473 ± 0.0128	5/5
	ResNet12/head	SHML	1.003 ± 0.003	0.0167 ± 0.0163	1/5
		Fixed-3	1.004 ± 0.004	0.0813 ± 0.0201	0/5
		Fixed-5	1.003 ± 0.002	0.0246 ± 0.0281	1/5
		Fixed-10	1.001 ± 0.003	-0.0083 ± 0.0402	3/5

SHML's large update-count advantage in Conv4/full accompanies lower NLL than Uniform in all five seeds on both datasets. In ResNet12/head, update ratios are close to one, and SHML's mean NLL contrast is positive on both datasets. Fixed-3 reverses its NLL contrast in every seed on both datasets. These settings provide distinct cost/quality regimes for future allocation procedures. Separately, at the 1,800-second Conv4 endpoint, fixed-3 completes 1.58 and 1.43 times as many updates as fixed-5 on CIFAR-FS and MiniImageNet, respectively, but has higher mean NLL on both. Thus the ranking of mean utility is not monotone in throughput even within a model/scope setting.

C.6 SECONDARY ENDPOINT METRICS

Table 14: Exact Conv4/full at 1,800 seconds. Brier averages all ten depths; the other columns evaluate only depth 10. Values are mean ± sample SD across five seeds. These endpoint metrics are distinct from the primary q_{10} NLL.

Dataset	Policy	q_{10} Brier	h_{10} NLL	h_{10} acc. (%)
CIFAR-FS	Uniform	0.6338 ± 0.0096	1.2040 ± 0.0271	51.22 ± 1.16
	SHML	0.6101 ± 0.0102	1.1612 ± 0.0281	53.90 ± 1.15
	Fixed-3	0.5969 ± 0.0064	1.1803 ± 0.0144	55.92 ± 0.54
	Fixed-5	0.5906 ± 0.0039	1.1201 ± 0.0150	55.96 ± 0.24
	Fixed-10	0.6381 ± 0.0069	1.1725 ± 0.0162	52.73 ± 0.76
MiniImageNet	Uniform	0.7100 ± 0.0040	1.3967 ± 0.0151	41.93 ± 0.51
	SHML	0.6800 ± 0.0050	1.3295 ± 0.0154	45.66 ± 0.36
	Fixed-3	0.6754 ± 0.0053	1.3308 ± 0.0158	46.96 ± 0.57
	Fixed-5	0.6727 ± 0.0025	1.3047 ± 0.0087	46.52 ± 0.24
	Fixed-10	0.6944 ± 0.0096	1.3317 ± 0.0244	44.56 ± 1.03

C.7 DEPTH REWEIGHTING AT THE MAIN ENDPOINT

Table 15 gives the empirical seed-average improvement range over common, task-independent depth weights. A positive minimum preserves the ranking; a range containing both signs permits a reversal (Appendix A.6). For fixed-5 versus SHML, every CIFAR-FS seed improves NLL and accuracy at all ten depths. On MiniImageNet, seeds 1818641017 and 3515040331 improve NLL at eight and five depths, and accuracy at four and ten, respectively; seed 3205086188 improves NLL at eight depths and accuracy at ten. The first two seeds improve both metrics at all depths.

Table 15: Exact Conv4/full, 1,800 seconds; pairs are candidate / reference. NLL gain is reference minus candidate, and accuracy gain the reverse (pp). Values are extrema of ten seed-average depth means, not confidence intervals or extrema across individual seeds.

		NLL gain		Accuracy gain (pp)	
Dataset	Pair	Min	Max	Min	Max
CIFAR-FS	SHML / Uniform	+0.03907	+0.06167	+2.35	+2.74
	Fixed-3 / Uniform	+0.02152	+0.17797	+3.72	+4.89
	Fixed-5 / Uniform	+0.06329	+0.10775	+4.53	+4.76
	Fixed-10 / Uniform	-0.05418	+0.03151	+0.77	+1.53
	Fixed-3 / SHML	-0.02255	+0.11963	+1.17	+2.54
	Fixed-5 / SHML	+0.02422	+0.04611	+2.00	+2.18
	Fixed-10 / SHML	-0.11335	-0.01125	-1.58	-1.07
MiniImageNet	SHML / Uniform	+0.05844	+0.08231	+3.40	+3.99
	Fixed-3 / Uniform	+0.06023	+0.13563	+4.23	+5.29
	Fixed-5 / Uniform	+0.05798	+0.09417	+4.31	+4.65
	Fixed-10 / Uniform	-0.00572	+0.06503	+1.13	+2.63
	Fixed-3 / SHML	-0.00708	+0.05332	+0.82	+1.33
	Fixed-5 / SHML	-0.00046	+0.02494	+0.66	+0.98
	Fixed-10 / SHML	-0.08343	-0.00214	-2.86	-0.87

C.8 UPDATE EXPOSURE AND FIXED-5 VERSUS SHML

Mean update cost is accumulated training time divided by completed updates, including in-timer task preparation. Table 16 contrasts early equal-update observations with the final equal-time comparison.

Table 16: Exact Conv4/full. Ratios and update costs use the 1,800-second endpoint; F5 / S denotes fixed-5 / SHML. Entries average per-seed quantities; NLL differences are fixed-5 minus SHML. Only the 100/200-update columns match update counts.

Dataset	Update ratio	s/update (F5 / S)	Δ_{100}	Δ_{200}	Δ_{1800s}
CIFAR-FS	1.086	0.373 / 0.405	-0.00119	-0.00659	-0.04109
MiniImageNet	1.074	0.448 / 0.481	+0.00403	+0.01053	-0.01868

C.9 Clipping exposure under the shared recipe

For each trajectory, let M be the total number of executed support updates and U the number of outer updates. If s_i^{in} and s_u^{out} are the applied clipping coefficients, the reported quantities are

$$f_{\text{in}} = \frac{1}{M}\sum_{i=1}^{M} \mathbf{1}\{s_i^{\text{in}} < 1\}, \qquad \bar{s}_{\text{in}} = \frac{1}{M}\sum_{i=1}^{M} s_i^{\text{in}},$$

$$f_{\text{out}} = \frac{1}{U}\sum_{u=1}^{U} \mathbf{1}\{s_u^{\text{out}} < 1\}, \qquad \bar{s}_{\text{out}} = \frac{1}{U}\sum_{u=1}^{U} s_u^{\text{out}}.$$

Scale means include unclipped events with coefficient one. Inner scales use the global norm over adapted parameters, with threshold 5; outer scales use the norm of the four-task averaged direction before Adam, with threshold 1 and the implementation's 10^{-6} denominator offset. We compute each ratio within a seed and then report the equally weighted seed mean and sample SD. The counters accumulate over training through the complete update crossing 1,800 seconds.

Table 17: Cumulative clipping diagnostics for exact Conv4/full, 1,800 seconds. Entries are mean $\pm$ sample SD across five training seeds. Inner rates count executed support updates; outer rates count meta-updates. Scale means include both clipped and unclipped events.

Dataset	Policy	Inner clipped (%)	Inner scale	Outer clipped (%)	Outer scale
CIFAR-FS	Uniform	24.0 ± 0.4	0.9426 ± 0.0016	98.9 ± 0.4	0.5568 ± 0.0200
	SHML	43.7 ± 0.2	0.8542 ± 0.0009	99.2 ± 0.3	0.4485 ± 0.0071
	Fixed-3	96.6 ± 0.1	0.5666 ± 0.0020	99.6 ± 0.1	0.3081 ± 0.0057
	Fixed-5	55.5 ± 0.4	0.8056 ± 0.0026	99.6 ± 0.1	0.4136 ± 0.0120
	Fixed-10	19.1 ± 0.5	0.9660 ± 0.0012	99.7 ± 0.2	0.5080 ± 0.0141
MiniImageNet	Uniform	29.4 ± 0.7	0.8895 ± 0.0015	99.9 ± 0.1	0.4454 ± 0.0153
	SHML	48.4 ± 0.3	0.7733 ± 0.0028	99.9 ± 0.0	0.3390 ± 0.0067
	Fixed-3	99.0 ± 0.3	0.5011 ± 0.0016	99.9 ± 0.1	0.2322 ± 0.0055
	Fixed-5	59.1 ± 0.4	0.7306 ± 0.0022	99.8 ± 0.3	0.3299 ± 0.0106
	Fixed-10	27.7 ± 0.3	0.9090 ± 0.0011	100.0 ± 0.0	0.4054 ± 0.0120

Fixed-3's mean inner scales are 0.567 and 0.501 on CIFAR-FS and MiniImageNet, compared with 0.943 and 0.890 for Uniform. Outer clipping is frequent for all five policies, while its average scale also differs across allocations.

Normalizing by tasks instead of executed steps. Let n_{task} count sampled training tasks, including repeated draws. The same counters give mean executed depth, clipped steps per task, and the cumulative coefficient deficit per task:

$$\bar{h} = M/n_{\text{task}}, \qquad c_{\text{task}} = \frac{1}{n_{\text{task}}}\sum_{i=1}^{M} \mathbf{1}\{s_i^{\text{in}} < 1\} = \bar{h} f_{\text{in}},$$

$$d_{\text{task}} = \frac{1}{n_{\text{task}}}\sum_{i=1}^{M} (1 - s_i^{\text{in}}) = \bar{h}(1 - \bar{s}_{\text{in}}).$$

These identities hold within each seed; we compute task-normalized values before averaging, rather than multiplying seed means. The deficit is a dimensionless sum of coefficient reductions, not a removed gradient norm or parameter displacement.

Table 18: Task-normalized inner clipping for the same trajectories as Table 17; mean ± sample SD across five seeds. $\bar{h}$ counts executed support steps per task, c_{task} counts clipped steps per task, and d_{task} sums $1 - s_i^{\text{in}}$ per task.

Dataset	Policy	$\bar{h}$	c_{task}	d_{task}
CIFAR-FS	Uniform	10.000 ± 0.000	2.397 ± 0.039	0.574 ± 0.016
	SHML	5.494 ± 0.018	2.402 ± 0.011	0.801 ± 0.006
	Fixed-3	3.000 ± 0.000	2.898 ± 0.004	1.300 ± 0.006
	Fixed-5	5.000 ± 0.000	2.775 ± 0.019	0.972 ± 0.013
	Fixed-10	10.000 ± 0.000	1.908 ± 0.050	0.340 ± 0.012
MiniImageNet	Uniform	10.000 ± 0.000	2.938 ± 0.065	1.105 ± 0.015
	SHML	5.493 ± 0.025	2.656 ± 0.010	1.245 ± 0.012
	Fixed-3	3.000 ± 0.000	2.971 ± 0.008	1.497 ± 0.005
	Fixed-5	5.000 ± 0.000	2.956 ± 0.022	1.347 ± 0.011
	Fixed-10	10.000 ± 0.000	2.769 ± 0.028	0.910 ± 0.011

On MiniImageNet, fixed-3 and Uniform have very different step-level clipping rates but similar counts per task (2.97 versus 2.94). On CIFAR-FS, the counts are 2.90 versus 2.40. Fixed-k traverses steps 1–k, Uniform traverses 1–10, and SHML traverses 1–J: shorter paths can have larger clipped fractions without proportionally more clipped events per task. Coefficient deficits retain information about scaling in addition to counts. Both normalizations combine different visited depths and model states, so they do not isolate an effect at a matched depth or state. Exact gradients already include inner clipping, preserving the raw sampling identity in Proposition 1; interpreting Adam motion additionally requires its moments and the final parameter projection.

C.10 Half-budget checkpoints versus full-budget Uniform

This comparison uses the saved exact Conv4/full checkpoints: each policy at 900 requested seconds is compared with Uniform at 1,800 requested seconds on the same dataset and seed. The reference mean NLL is 1.2410 on CIFAR-FS and 1.4031 on MiniImageNet. Table 19 retains all five policies, including Uniform's own earlier checkpoint. Individual actual-time ratios range from 0.49971 to 0.50063, reflecting complete-update threshold crossing. Differences are formed within each seed before averaging. The ratios describe these checkpoints, not interpolated time-to-quality speedups.

Table 19: All policies at 900 seconds versus Uniform at 1,800 seconds; exact Conv4/full, mean ± sample SD across five seeds. Deltas are candidate minus reference. Favorable counts indicate lower NLL and higher accuracy, respectively; accuracy differences are percentage points.

Dataset	Policy	NLL	Δ NLL	Δ acc. (pp)	Seeds (NLL, acc.)
CIFAR-FS	Uniform	1.3239 ± 0.0176	0.0828 ± 0.0174	-4.26 ± 0.95	0/5, 0/5
	SHML	1.2634 ± 0.0151	0.0223 ± 0.0327	-0.93 ± 1.78	1/5, 1/5
	Fixed-3	1.2217 ± 0.0234	-0.0193 ± 0.0310	1.60 ± 1.65	3/5, 4/5
	Fixed-5	1.2317 ± 0.0111	-0.0093 ± 0.0175	0.63 ± 1.03	3/5, 3/5
	Fixed-10	1.3264 ± 0.0080	0.0854 ± 0.0198	-2.79 ± 0.97	0/5, 0/5
MiniImageNet	Uniform	1.4625 ± 0.0048	0.0594 ± 0.0115	-3.53 ± 0.54	0/5, 0/5
	SHML	1.3901 ± 0.0169	-0.0131 ± 0.0209	0.59 ± 1.03	4/5, 4/5
	Fixed-3	1.3712 ± 0.0132	-0.0319 ± 0.0205	2.31 ± 1.27	5/5, 5/5
	Fixed-5	1.3604 ± 0.0086	-0.0427 ± 0.0083	2.10 ± 0.71	5/5, 5/5
	Fixed-10	1.4153 ± 0.0140	0.0121 ± 0.0155	-0.36 ± 0.89	1/5, 2/5

Table 20: Individual-seed NLL differences for Table 19. $s_1 = 3403801132$, $s_2 = 4001970766$, $s_3 = 1818641017$, $s_4 = 3515040331$, and $s_5 = 3205086188$. Negative differences favor the half-budget candidate.

Dataset	Policy	s_1	s_2	s_3	s_4	s_5
CIFAR-FS	Uniform	+0.06627	+0.08110	+0.10692	+0.06700	+0.09286
	SHML	+0.01411	-0.01602	+0.07434	+0.01683	+0.02242
	Fixed-3	+0.01192	-0.05751	-0.01763	-0.04321	+0.00972
	Fixed-5	-0.01911	-0.02902	+0.00616	-0.01648	+0.01192
	Fixed-10	+0.09149	+0.06459	+0.11209	+0.06688	+0.09189
MiniImageNet	Uniform	+0.06067	+0.04860	+0.06488	+0.07508	+0.04767
	SHML	-0.01399	-0.02969	+0.02290	-0.01965	-0.02492
	Fixed-3	-0.02678	-0.05307	-0.02873	-0.00182	-0.04906
	Fixed-5	-0.04659	-0.04527	-0.04729	-0.02793	-0.04650
	Fixed-10	-0.00798	+0.00971	+0.01496	+0.03523	+0.00874

Table 21: Individual-seed accuracy differences (pp) for Table 19. Seed order follows Table 20; positive values favor the half-budget candidate.

Dataset	Policy	s_1	s_2	s_3	s_4	s_5
CIFAR-FS	Uniform	-3.47	-4.21	-5.33	-3.20	-5.10
	SHML	-0.72	+1.35	-3.64	-0.68	-0.96
	Fixed-3	+0.29	+3.82	+1.34	+2.70	-0.13
	Fixed-5	+1.19	+1.81	-0.35	+1.06	-0.56
	Fixed-10	-3.39	-1.96	-4.15	-1.88	-2.55
MiniImageNet	Uniform	-3.22	-3.02	-4.11	-4.11	-3.18
	SHML	+0.60	+1.65	-1.07	+0.58	+1.18
	Fixed-3	+2.30	+3.80	+1.74	+0.53	+3.16
	Fixed-5	+2.53	+2.43	+2.34	+0.84	+2.35
	Fixed-10	+0.64	+0.27	-0.61	-1.67	-0.42

D Prediction and Resource Diagnostics

These exploratory analyses use exported predictions and training counters from the existing runs. All 810 prediction files were independently recomputed from their logits in float64. The maximum absolute discrepancy against saved NLL was 4.44×10^{-16}; accuracy, Brier score and confidence agreed to 2.23×10^{-16}. Episode indices, labels, horizons and bank hashes align across methods and seeds within each dataset. Task-breadth and adaptation-trajectory diagnostics use exact Conv4/full at 1,800 seconds; the final subsection compares derivatives at 300/900 seconds. Summaries first average within each seed; the 600 shared episodes are not additional independent training replicates.

D.1 Breadth of improvement across paired tasks

For episode e and seed s, let $D_{s,e} = L^{\mathrm{ref}}_{s,e} - L^{\mathrm{cand}}_{s,e}$, where L averages query NLL over all ten depths. A task win means $D_{s,e} > 0$. Every seed has a positive median gain in Table 22. Symmetrically trimming the largest and smallest 10% of gains leaves mean gains close to the untrimmed values; the advantage describes more than a few favorable episodes.

Table 22: Task-level breadth on the paired test banks. Win percentages are mean ± sample SD across five seeds; the next two columns are across-seed ranges. Trimmed gains average after removing 10% at each end within each seed. Positive NLL gains favor the candidate, listed first.

Dataset	Candidate / reference	Wins (%)	Win range	Median gain range	Trimmed gain
CIFAR-FS	SHML / Uniform	68.7 ± 8.7	56.0–79.8	0.0138–0.0786	0.0482
	Fixed-5 / SHML	67.5 ± 6.8	58.0–74.8	0.0199–0.0725	0.0417
MiniImageNet	SHML / Uniform	78.3 ± 3.5	74.7–82.8	0.0614–0.0761	0.0692
	Fixed-5 / SHML	59.1 ± 9.4	51.0–73.3	0.0010–0.0488	0.0189

Winning tasks vary with training randomness. For fixed-5 versus SHML, 17.0% of CIFAR-FS tasks and 8.2% of MiniImageNet tasks improve in all five seeds; 79.3% and 66.3% improve in at least three. Removing only the largest 10% of gains, a deliberately unfavorable sensitivity calculation, leaves seed-average gains of 0.0216 and 0.0036. They remain positive in four CIFAR-FS seeds and two MiniImageNet seeds. These summaries describe the fixed bank and do not replace full-bank NLL or seed-level uncertainty.

D.2 Early and late adaptation favor different supervision depths

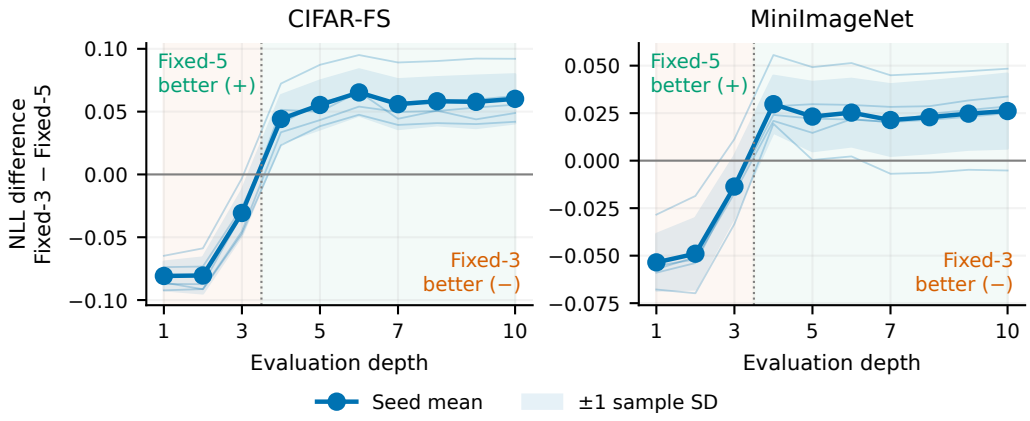

Figure 3: Paired NLL of fixed-3 minus fixed-5 at each adaptation depth. Thin lines are training seeds, the thick line their mean, and shading one sample SD of paired differences. Negative values favor fixed-3. The early–late reversal holds at every depth in its respective block in all five CIFAR-FS seeds.

On CIFAR-FS, the mean contrast is $d_E = -0.06398$ for steps 1–3 and $d_L = 0.05670$ for steps 4–10. If evaluation places probability ρ uniformly within the first block and $1 - \rho$ uniformly within the second, the seed-mean contrast is $\rho d_E + (1-\rho) d_L$. It favors fixed-3 when $\rho > d_L/(d_L - d_E) = 0.470$, compared with $\rho = 0.3$ under q_{10}. Individual-seed crossing weights range from 0.341 to 0.677. This reweights saved risks under the stated block weights, rather than learning a stopping rule. On MiniImageNet, all seeds favor fixed-3 over steps 1–3 on average; four of five favor fixed-5 over steps 4–10.

Table 23: Changes along adaptation; seed means. Fixed-3 uses steps 3 to 10; other rows use 5 to 10. Accuracy and confidence changes are percentage points. The last column counts seeds with increasing NLL.

Dataset	Policy	Steps	ΔNLL	ΔAcc.	ΔConf.	ΔBrier	NLL ↑
CIFAR-FS	Uniform	5–10	+0.0083	−0.08	+8.94	+0.0045	4/5
	SHML	5–10	+0.0143	+0.06	+7.75	+0.0061	5/5
	Fixed-3	3–10	+0.0677	+0.49	+12.97	+0.0224	5/5
	Fixed-5	5–10	+0.0180	+0.11	+7.65	+0.0078	5/5
	Fixed-10	5–10	−0.0454	−0.04	+9.15	−0.0208	0/5
MiniImageNet	Uniform	5–10	+0.0191	−0.05	+6.15	+0.0083	5/5
	SHML	5–10	+0.0148	+0.28	+5.00	+0.0057	5/5
	Fixed-3	3–10	+0.0447	+0.43	+9.40	+0.0173	5/5
	Fixed-5	5–10	+0.0150	+0.23	+5.59	+0.0066	5/5
	Fixed-10	5–10	−0.0078	+0.06	+6.51	−0.0034	0/5

Late NLL degradation also appears under SHML and Uniform; it is not exclusive to fixed supervision. Fixed-3's minimum NLL occurs at step 3 in every dataset–seed pair. Fixed-5 minimizes at step 5 on CIFAR-FS and steps 4–5 on MiniImageNet; fixed-10 at step 10 and steps 8–9, respectively.

Partition fixed-3 queries by correctness at step 3 and keep membership fixed when comparing steps 3 and 10. For correct group C and wrong group W,

$$\Delta L = \Pr(C)\mathbb{E}[\Delta\ell \mid C] + \Pr(W)\mathbb{E}[\Delta\ell \mid W]. \tag{19}$$

Contributions are normalized over all queries within each seed before averaging. On CIFAR-FS, $0.06766 = -0.09655 + 0.16421$; on MiniImageNet, $0.04473 = -0.06821 + 0.11294$. The initially wrong group has increasing mean NLL in every seed (ranges 0.3468–0.3934 and 0.1919–0.2265), while the initially correct group has decreasing mean NLL in every seed. Small net accuracy gains coexist with larger probability losses on initially misclassified queries.

From steps 5 to 10, 18.6% of CIFAR-FS queries change predicted label at least once under fixed-3, compared with 7.7% under fixed-5. Prediction changes can cancel in aggregate accuracy. NLL and Brier are proper probability scores; neither they nor mean confidence alone isolate calibration error or a training mechanism.

D.3 Memory costs of dense and single-horizon supervision

Training logs record `torch.cuda.max_memory_allocated()` in each isolated training process, with no peak reset. It includes setup, training and checkpoint stages; profiling and evaluation are separate processes. This measures peak allocated CUDA tensor storage, not reserved allocator memory or total device usage. MiB denotes 2^{20} bytes.

Table 24: Peak allocated CUDA memory (MiB), exact differentiation; five-seed means for full 1,800-second Conv4/full and 900-second ResNet12/head runs. SD is zero except MiniImageNet Conv4/full SHML (3.46 MiB).

Dataset	Setting	Uniform	SHML	Fixed-3	Fixed-5	Fixed-10
CIFAR-FS	Conv4/full	926.39	301.18	239.56	257.17	301.18
	ResNet12/head	814.05	813.95	813.75	813.80	813.95
MiniImageNet	Conv4/full	5672.43	1401.43	1007.53	1120.43	1405.16
	ResNet12/head	3486.80	3486.69	3486.49	3486.55	3486.69

Uniform and fixed-10 both execute ten support steps in Conv4/full, yet fixed-10 uses 67.5%/75.2% less peak allocated memory. The implementation retains ten supervised query branches for Uniform and one for fixed-10, consistent with the storage difference. SHML's maximum depth is ten, explaining its peak's proximity to fixed-10 rather than fixed-5. In ResNet12/head, cached encoder features are shared across depths. Uniform exceeds SHML by 0.106 MiB on each dataset (0.013%/0.003%,

respectively, using unrounded byte counts). Graph recomputation or caching changes can change this implementation-specific cost profile.

Table 25: Actual accumulated training-update times at 1,800 requested seconds. Across-seed ranges include the crossing update.

Dataset	Policy	Mean seconds	Across-seed range
CIFAR-FS	Uniform	1800.537	1800.390–1800.676
	SHML	1800.219	1800.025–1800.454
	Fixed-3	1800.145	1800.034–1800.202
	Fixed-5	1800.227	1800.164–1800.289
	Fixed-10	1800.351	1800.149–1800.630
MiniImageNet	Uniform	1800.660	1800.280–1801.293
	SHML	1800.240	1800.098–1800.399
	Fixed-3	1800.100	1800.012–1800.205
	Fixed-5	1800.200	1800.030–1800.343
	Fixed-10	1800.411	1800.291–1800.486

D.4 Absolute outcomes and paired contrasts across derivative rules

Table 3 reports absolute 900-second outcomes. Table 26 adds contrasts with SHML. A lower mean need not imply consistent paired gains: CIFAR-FS FO fixed-3 wins in only one seed, with a -0.0674 difference in that seed. MiniImageNet FO fixed-5 wins in two seeds. These are descriptive comparisons of the complete grid.

Table 26: 900-second NLL differences versus SHML under the same derivative: mean difference and favorable paired seed count. Negative differences favor the fixed policy. Prefix-3 fixed-3 reuses exact fixed-3; its reference uses prefix-3.

Dataset	Derivative	Fixed-3	Fixed-5	Fixed-10
CIFAR-FS	Exact	-0.0417 (5/5)	-0.0316 (5/5)	$+0.0631$ (0/5)
	FO	-0.0109 (1/5)	-0.0353 (5/5)	$+0.0429$ (0/5)
	Prefix-3	-0.0067 (3/5)	-0.0229 (3/5)	$+0.0152$ (0/5)
MiniImageNet	Exact	-0.0188 (4/5)	-0.0296 (5/5)	$+0.0252$ (1/5)
	FO	-0.0014 (3/5)	-0.0014 (2/5)	$+0.0189$ (0/5)
	Prefix-3	$+0.0109$ (2/5)	-0.0109 (4/5)	$+0.0244$ (0/5)

At 900 seconds on MiniImageNet, changing exact to FO raises Uniform's mean NLL by 0.07834 and lowers SHML's by 0.01209. Of the 0.09042 expansion in their contrast, 86.6% is the arithmetic contribution of Uniform's degradation. At 300 seconds, FO worsens both Uniform ($+0.03538$) and SHML ($+0.02691$) in all seeds, while making SHML's relative contrast more favorable by 0.00847. A favorable interaction can coexist with worse absolute performance for both procedures. This decomposition describes outcomes, not a gradient-level mechanism.

E Supervision Depth and Observed Probability-Score Minima

For every dataset, setting, seed and checkpoint, we minimize the mean 600-episode NLL over ten prediction depths; Brier is treated analogously. Prefix-3 fixed-3 reuses exact fixed-3. The time grid has 18 conditions: 300/900 seconds in four settings on two datasets, plus the exact Conv4/full 1,800-second endpoint on both datasets. Five seeds give 90 descriptive triplets. The 100/200-update grid has 16 conditions and 80 triplets. Repeated checkpoints and reused controls do not increase the independent seed count.

NLL minima satisfy $h_3^* \le h_5^* < h_{10}^*$ in all 90 time triplets, with 87 strictly ordered. Brier gives 90 nondecreasing and 88 strict triplets. At equal early updates, NLL gives 77 nondecreasing and 71 strict triplets; Brier gives 77 and 72. Table 27 provides the settings shown in Figure 4.

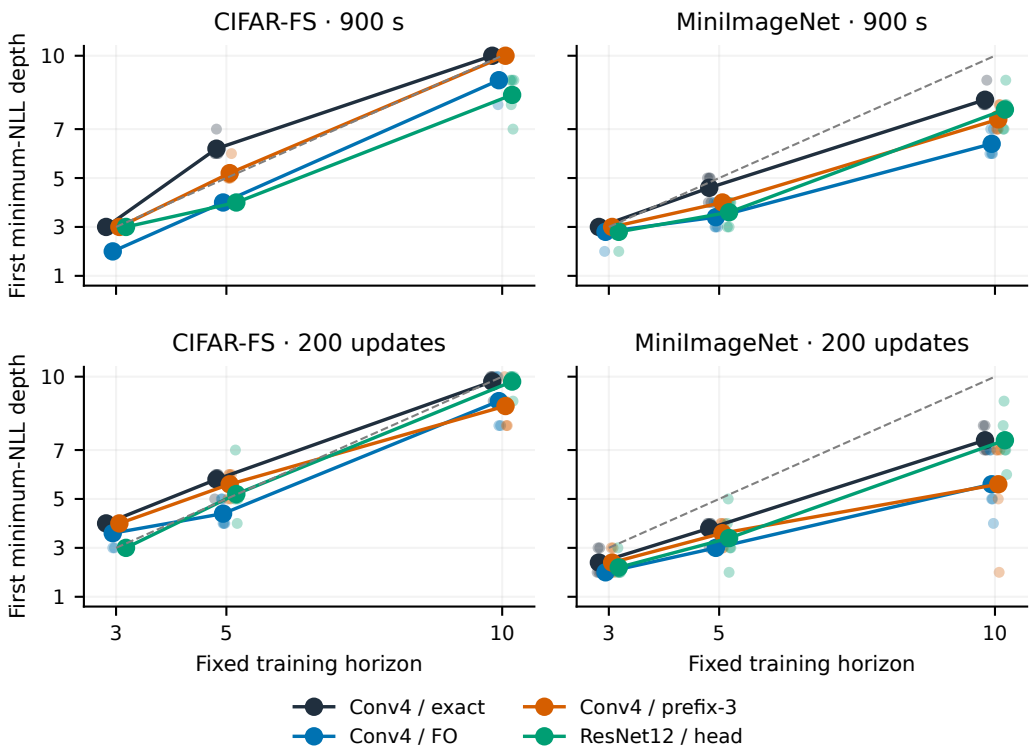

Figure 4: Supervision depth and observed NLL-minimizing prediction depth. Points are individual seeds; lines their means. Top: equal training time; bottom: equal updates/task exposure. Dashed lines indicate equal training and test depths. Prefix-3 fixed-3 reuses exact fixed-3. Minima are restricted to steps 1–10; a boundary value does not establish an optimum beyond 10.

Table 27: Observed NLL-minimizing depths: ranges across five seeds. Final columns count nondecreasing and strictly increasing triplets. Prefix-3 fixed-3 reuses exact fixed-3.

Budget	Dataset	Setting	Fixed-3	Fixed-5	Fixed-10	$\leq$	$<$
900 s	CIFAR-FS	Conv4/full exact	3	6–7	10	5/5	5/5
		Conv4/full FO	2	4	8–10	5/5	5/5
		Conv4/full prefix-3	3	5–6	10	5/5	5/5
		ResNet12/head exact	3	4	7–9	5/5	5/5
	MiniImageNet	Conv4/full exact	3	4–5	8–9	5/5	5/5
		Conv4/full FO	2–3	3–4	6–7	5/5	3/5
		Conv4/full prefix-3	3	4	7–8	5/5	5/5
		ResNet12/head exact	2–3	3–4	7–9	5/5	4/5
200 updates	CIFAR-FS	Conv4/full exact	4	5–6	9–10	5/5	5/5
		Conv4/full FO	3–4	4–5	8–10	5/5	4/5
		Conv4/full prefix-3	4	5–6	8–10	5/5	5/5
		ResNet12/head exact	3	4–7	9–10	5/5	5/5
	MiniImageNet	Conv4/full exact	2–3	3–4	7–8	5/5	5/5
		Conv4/full FO	2	3	4–7	5/5	5/5
		Conv4/full prefix-3	2–3	3–4	2–7	4/5	4/5
		ResNet12/head exact	2–3	2–5	6–9	5/5	4/5

Early exceptions. All three reversals occur in MiniImageNet prefix-3 fixed-10 runs; NLL and Brier agree. At 100 updates, seeds 1818641017 and 3403801132 have minimum triplets $(2, 3, 2)$ and $(2, 4, 2)$; at 200 updates, seed 3515040331 has $(3, 3, 2)$. For fixed-10, excess NLL at fixed-5's minimizing depth is 0.02153, 0.00182 and 0.01240, respectively. These include meaningful

differences, not only numerical ties. The fixed-10 depths within 0.005 of its minimum are $\{2\}$, $\{2, 4\}$ and $\{2, 4\}$; near-minimum sets need not be contiguous.

Flatness and effect size. The ordinal pattern need not identify a unique practically preferable step. Excluding duplicate exact fixed-3 aliases, median best–second-best NLL gaps at time checkpoints are 0.01305, 0.00163 and 0.00058 for fixed-3/5/10. Median counts of depths within 0.005 NLL of the minimum are 1, 3 and 4; eight of 90 fixed-10 records have gaps below 10^{-4}. The tolerance describes sensitivity, not a selected stopping threshold.

Across all 90 time triplets, fixed-3's step-10 excess NLL over its own minimum is at least 0.03945, with median 0.08640; fixed-5's median is 0.02549. Fixed-10's minimum is at the right boundary in 32 triplets and internal in 58. A minimum at 10 does not identify the unrestricted optimum.

Scope of the accuracy contrast. The main endpoint's all-seed accuracy–NLL divergence is not universal. Among fixed-3's 130 distinct time/update observations, step-10 NLL exceeds step-3 NLL in all 130, Brier worsens in 127, and accuracy rises in 77. More adaptation does not always improve classification accuracy.

F Probability-Scale and Logit-Geometry Diagnostics

F.1 Two-fold scalar-temperature adjustment

The analysis uses all 50 exact Conv4/full endpoints at 1,800 seconds: five policies, five seeds and two datasets, at every depth 1–10. For each dataset, partition the shared 600-episode bank into halves using a fixed permutation with seed 20260922. All queries of an episode stay in one half. The split is shared across models and depths. For each model and depth, fit $\beta = 1/T > 0$ to minimize mean NLL of $\mathrm{softmax}(\beta z)$ on one half, evaluate the other, and swap. Scores average the 600 out-of-fold episodes within each training seed. Positive scaling preserves all class orderings and top-1 predictions. The depth-specific version fits ten scalars across the path. A shared-path control fits one scalar jointly to equally weighted depths 1–10 for each model and fitting half.

The inverse-temperature objective is convex. Its monotone first derivative is solved on $[10^{-6}, 10^{3}]$; all 1,100 fits have interior solutions. Absolute fitted derivatives are below 10^{-10}, and recomputed raw scores agree with saved metrics within 10^{-12}. Raw metrics remain the primary benchmark outcomes.

This is an exploratory diagnostic within the test bank used to discover the patterns. Separating fitting and scoring episodes avoids evaluating a scalar on its fitting episodes, but does not create an independent validation bank; folds share dataset classes and can reuse images. Temperatures are diagnostic parameters, not independently selected deployment settings.

Table 28: Five-seed mean NLL diagnostics. Positive late changes mean step 10 is worse than step 3; positive method gains favor fixed-5. Parentheses give positive paired seed counts out of five.

Dataset	Contrast	Raw	Depth-specific T	Shared-path T
CIFAR-FS	Fixed-3: $h_{10} - h_3$	$+.06766$ (5)	$+.00816$ (5)	$+.00437$ (4)
	SHML minus fixed-5	$+.04109$ (5)	$+.04466$ (5)	$+.04100$ (5)
	Fixed-3 minus fixed-5	$+.02050$ (5)	$+.01951$ (5)	$+.00898$ (2)
MiniImageNet	Fixed-3: $h_{10} - h_3$	$+.04473$ (5)	$+.00022$ (2)	$+.00129$ (3)
	SHML minus fixed-5	$+.01868$ (5)	$+.02064$ (5)	$+.01747$ (5)
	Fixed-3 minus fixed-5	$+.00572$ (4)	$-.00098$ (2)	$-.00683$ (1)

Fixed-3's fitted temperatures average 0.958 at step 3 and 1.477 at step 10 on CIFAR-FS, and 1.060/1.458 on MiniImageNet, averaging $T = 1/\beta$ over both fitting folds and all five seeds. Shared-path temperature reduces the late difference partly by increasing step-3 NLL: 1.11263 to 1.12900 on CIFAR-FS and 1.28612 to 1.29094 on MiniImageNet. Thus not every depth improves. Across 500 model–depth cells, out-of-fold NLL increases in 27 after depth-specific scaling (maximum 0.000713), and in 202 after shared-path scaling, which optimizes a compromise across depths. A remaining

adjusted gap can reflect class ranking, relative logit geometry, or distortions one scalar cannot correct; it is not a pure measure of discrimination.

The observed minimum also depends on adjustment. At these ten exact Conv4/full dataset–seed endpoints, depth-specific scaling retains nondecreasing fixed-3/5/10 minima in all ten cases (eight strict), whereas shared-path scaling does so in two. Raw minima are strictly ordered in all ten. The full-grid pattern concerns raw probability trajectories, not a calibration-invariant stopping rule.

F.2 A FOUR-CORNER OUTPUT DECOMPOSITION WITHOUT FITTED SCALARS

For each query, center the five logits at steps 3 and 10: $c_h = z_h - \mathrm{mean}(z_h)\mathbf{1}$, $r_h = \|c_h\|_2$, $u_h = c_h/r_h$. Centering preserves softmax probabilities. Score r_3u_3, $r_{10}u_{10}$, $r_{10}u_3$ and r_3u_{10}. For query NLL L,

$$D_{\mathrm{radius}} = \tfrac{1}{2}\{L(r_{10}u_3) - L(r_3u_3) + L(r_{10}u_{10}) - L(r_3u_{10})\}, \tag{20}$$

$$D_{\mathrm{direction}} = \tfrac{1}{2}\{L(r_3u_{10}) - L(r_3u_3) + L(r_{10}u_{10}) - L(r_{10}u_3)\}. \tag{21}$$

Their sum is exactly $L(z_{10}) - L(z_3)$. Average over queries and episodes within each seed. This symmetric algebraic decomposition changes observed output coordinates, not training. Radius exchanges are query-specific, unlike global temperature scaling. Direction contains all relative class-score geometry, not only argmax or discrete ranks. All inspected queries have nonzero centered radius.

For fixed-3, mean total/radius/direction changes are $0.06766 = 0.05263 + 0.01503$ on CIFAR-FS and $0.04473 = 0.03889 + 0.00584$ on MiniImageNet. Both contributions are positive in all seeds on both datasets. Mean radii increase from 2.341 to 3.415 and from 1.981 to 2.640, respectively. Radius change accounts for the larger part of the observed increase, while direction also worsens NLL. These are output-space contributions, not causal percentages of a training mechanism.

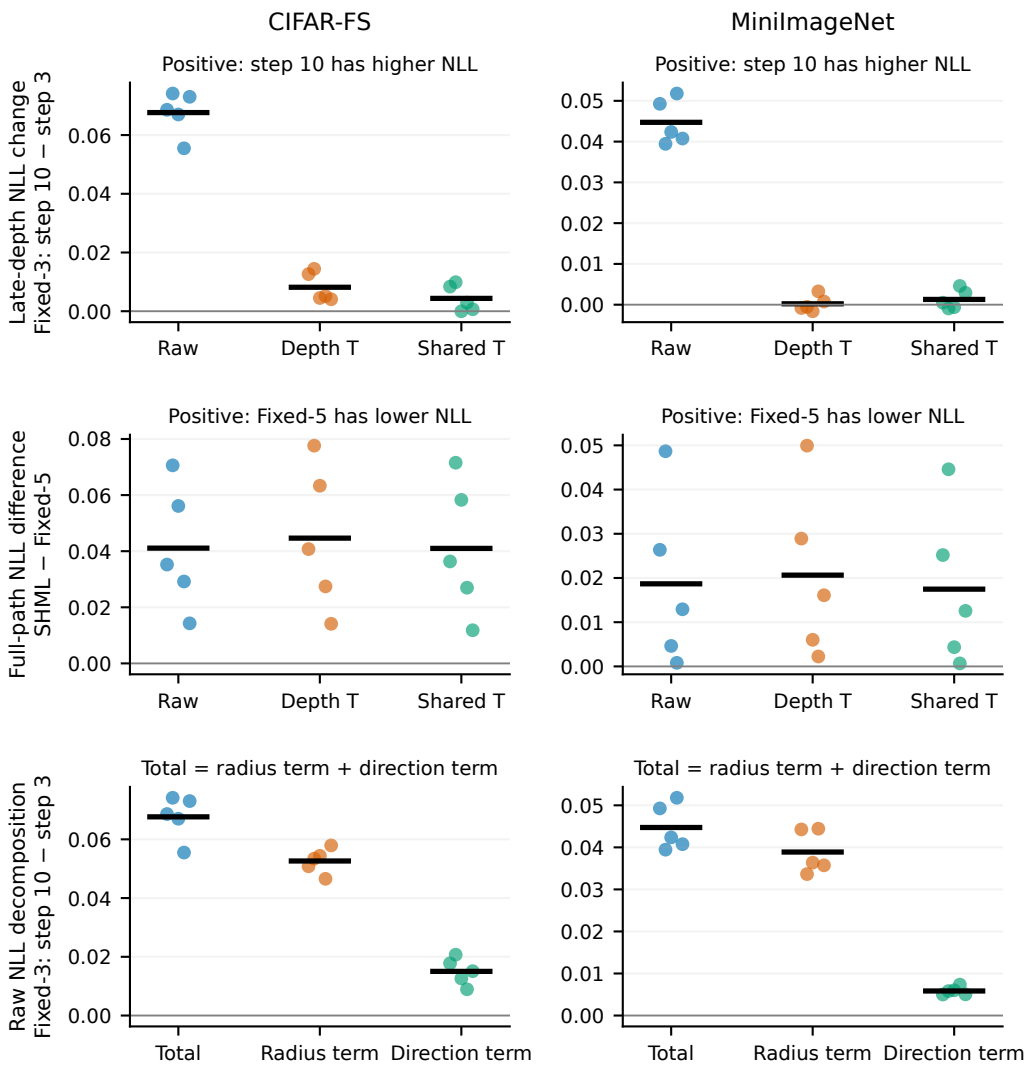

Figure 5: Exploratory scale diagnostics, with one column per dataset. Top: fixed-3's late NLL change. Middle: fixed-5's full-path gain over SHML. Bottom: the radius/direction decomposition above. Points are training seeds; short bars are means. Depth-specific T uses ten scalars; shared T uses one across the path.