HQ-RECYCLE: EFFICIENT BILEVEL OPTIMIZATION THROUGH ADJOINT RECYCLING

Anonymous authors
Paper under double-blind review

ABSTRACT

We introduce Hq-Recycle, a quasi-Newton method for bilevel optimization that reduces repeated work in implicit hypergradient estimation by recycling adjoint estimates across outer iterations. The method tracks the lower-level solution using limited-memory quasi-Newton updates and retains the previous adjoint estimate. A fresh directional finite-difference curvature probe produces a one-dimensional Galerkin initialization, followed by preconditioned corrections only when the estimated residual exceeds a prescribed tolerance. This procedure limits refinement to a small probe budget while accounting for changes in the current adjoint system. For exact symmetric positive-definite systems, we establish an energy-minimization property of the recycled initialization and an error bound governed by the previous residual and inter-iteration system drift. We evaluate the four-probe configuration on three label-corrupted linear classification tasks and grouped ridge regression against qNBO, SHINE, PZOBO, BOME, HOAG, and AID-CG, using validation-based tuning and five fresh evaluation seeds. In classification ablations, recycling reduces the mean number of adjoint probes from four to approximately 1.07 per outer iteration. On grouped ridge regression, Hq-Recycle achieves a 2.44-fold speedup over tuned qNBO in mean time to a prespecified validation target and lowers mean test loss by 20.2% under a 0.5-second solver budget.

1 INTRODUCTION

Bilevel optimization learns upper-level variables through their effect on a lower-level solution. In data reweighting, for example, sample weights are chosen to improve validation performance after training the classifier. Implicit differentiation computes hypergradients through an adjoint linear solve, but each solve adds to the cost of lower-level optimization (Pedregosa, 2016; Ji et al., 2021; Fang et al., 2025). Under a short time budget, the challenge is to adapt the previous adjoint to the current system with few additional computations.

Quasi-Newton methods already retain curvature information from lower optimization. SHINE shares an inverse estimate between the lower solve and hypergradient computation, and qNBO uses quasi-Newton updates for lower tracking and inverse-action approximation (Ramzi et al., 2022; Fang et al., 2025). The previous adjoint contains complementary information: it reflects both curvature and the preceding right-hand side. When successive systems are related, adapting this vector offers a compact way to reduce repeated work.

We introduce Hq-Recycle, which *refreshes, rescales, and corrects* one retained adjoint. A one-dimensional Galerkin solve rescales the retained adjoint using a fresh curvature action in the current system. A residual test then determines whether additional preconditioned corrections are needed. The evaluated configuration uses at most four probes, including the refresh, and stores one adjoint vector alongside the lower L-BFGS memory. Lower-gradient stopping, deferred preconditioning, and shared directional computations reduce the cost per outer step.

After an expanded outer-rate search, ten-seed ridge confirmation shows a $2.42\times$ speedup over qNBO in mean time to a fixed validation target and a 10.98% reduction in mean test loss within 0.5 seconds.

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.

We make three contributions:

- A compact adjoint-recycling solver that combines current-system adaptation and residual-controlled refinement with quasi-Newton lower tracking, under an explicit four-probe budget.
- A local analysis quantifying the energy improvement over the retained estimate and the effect of system drift. Matched real-system replay and controlled drift experiments identify when rescaling improves local hypergradient accuracy or reduces correction work.
- An empirical study combining established baselines, component ablations, and independently tuned initialization variants. Independent lower-response evaluation and repeated timing distinguish solution quality, implementation efficiency, and work to a fixed target.

The abstract reports the initial five-seed evaluation, preserved in Appendix L. A separate five-seed study evaluates implementation changes and six baselines; the ten-seed confirmation adds broader rate searches and mechanism controls. We report each cohort separately. The abstract's figure of approximately 1.07 classification probes per iteration is a whole-trajectory average; Section 4.4 also examines work before target attainment.

Related work. Implicit methods such as HOAG and approximate implicit differentiation use iterative inverse-Hessian actions (Pedregosa, 2016; Ji et al., 2021), whereas PZOBO and BOME develop zeroth-order and first-order alternatives (Sow et al., 2022; Liu et al., 2022). SHINE, qNBO, and curvature-aware hypergradient approximation exploit lower-level curvature (Ramzi et al., 2022; Fang et al., 2025; Dong et al., 2025). For related adjoint systems, LancBiO constructs dynamic Krylov subspaces (Yang et al., 2025), and Ehrhardt et al. (2024) develop recycled Krylov methods with Ritz generalized singular vectors and hypergradient-oriented stopping. Building on these reuse and projection ideas, we study how to adapt a single retained adjoint within a four-probe budget.

2 HQ-RECYCLE

2.1 PROBLEM AND NOTATION

We consider

$$\min_{x\in\mathbb{R}^{d_x}} \Phi(x) = F(x, y^\star(x)), \qquad y^\star(x) = \arg\min_{y\in\mathbb{R}^{d_y}} f(x, y). \tag{1}$$

When the lower objective is smooth and strongly convex, implicit differentiation gives

$$\nabla\Phi(x) = \nabla_x F(x, y^\star) - C(x, y^\star)q^\star, \quad A(x, y^\star)q^\star = b(x, y^\star), \tag{2}$$

where $A = \nabla^2_{yy} f \in \mathbb{R}^{d_y\times d_y}$, $b = \nabla_y F$, and $C = [\partial(\nabla_y f)/\partial x]^\top \in \mathbb{R}^{d_x\times d_y}$. The vector $q^\star$ is the *adjoint*, or inverse-Hessian–vector product. At outer iteration k, we approximately minimize $f(x_k, \cdot)$ starting from y_k to obtain the tracked lower state $\bar{y}_k$. We write A_k, b_k, C_k for the quantities evaluated at $(x_k, \bar{y}_k)$. The exact local adjoint $q_k^\dagger = A_k^{-1} b_k$ is distinct from the adjoint evaluated at $y^\star(x_k)$.

2.2 LOWER-LEVEL TRACKING AND REUSABLE CURVATURE

Starting from y_k, we perform up to $T_{\max}$ limited-memory BFGS (L-BFGS) steps on $f(x_k, \cdot)$. After at least one step, tracking stops if the lower-gradient norm is at most ϵ_{low}. We use $\epsilon_{\text{low}} = 10^{-4}$ for classification; the initial implementation uses a fixed step count. Each accepted curvature pair is

$$s_t = y_{t+1} - y_t, \qquad w_t = \nabla_y f(x_k, y_{t+1}) - \nabla_y f(x_k, y_t). \tag{3}$$

We retain at most m pairs satisfying a positive relative-curvature test. These pairs define an inverse L-BFGS operator B_k that supplies lower-level directions and, once the memory is frozen after tracking, preconditions the adjoint solve. A cached compact representation implements the equivalent two-loop recursion; Appendix B specifies the operator and safeguards. Adjoint probes do not update this memory. The lower-level pairs approximate inverse curvature; the retained adjoint captures information about the previous right-hand side.

2.3 REFRESH, RESCALE, AND CORRECT

When a usable adjoint is available, we reserve one probe to refresh it in the current system. We then correct the residual within the remaining budget, with the tracked lower state and curvature memory fixed throughout the solve. For a nonzero probe direction p, approximate the Hessian–vector product (HVP) $A_k p$ by the forward finite-difference (FD) action

$$\mathcal{D}_k(p) = \frac{\nabla_y f(x_k, \bar{y}_k + h_p p) - \nabla_y f(x_k, \bar{y}_k)}{h_p}, \quad h_p = \frac{\delta_k}{\|p\|}, \quad \delta_k = 10^{-3} \max\{1, \|\bar{y}_k\|\}. \tag{4}$$

The generic implementation caches the anchor gradient from lower tracking, so each action requires one additional lower-gradient evaluation without forming A_k. Mixed-derivative actions $C_k p$ use automatic differentiation (AD). For weighted linear softmax objectives, the structured backend reuses directional logits for both finite-difference and mixed actions. The forward-difference approximation is unchanged (Appendix B.3).

If the previous estimate q_{k-1} is usable, normalize $v = q_{k-1}/\|q_{k-1}\|$, compute $a = \mathcal{D}_k(v)$, and set

$$\alpha = \frac{v^\top b_k}{v^\top a}, \qquad q_k^{(0)} = \alpha v, \qquad \widehat{r}_k^{(0)} = b_k - \alpha a. \tag{5}$$

We require $v^\top a$ to be finite and positive. Otherwise the initialization remains zero, and the attempted refresh still counts toward the probe budget. With exact actions, Eq. (5) is a one-dimensional Galerkin solve. We recompute a to refresh curvature information along the retained direction.

If $\left\|\widehat{r}_k^{(0)}\right\| \le \tau \|b_k\|$, we accept the recycled estimate. Otherwise we apply preconditioned conjugate-gradient (PCG) recurrences from $q_k^{(0)}$, with B_k fixed and $\mathcal{D}_k$ supplying the actions. These are PCG-style corrections because finite differences need not define an exact linear operator. The optimized implementation constructs or applies the frozen adjoint preconditioner only when correction requires it. Each call to $\mathcal{D}_k$, including the refresh, consumes one probe. The first solve starts from zero. A zero right-hand side gives $q_k = 0$ without probing. We set $P = 4$ and $\tau = 0.05$; a refresh leaves at most three correction probes. Algorithm 1 summarizes the procedure. The generic outer update is

$$\widehat{g}_k = \nabla_x F(x_k, \bar{y}_k) - C_k q_k, \qquad x_{k+1} = x_k - \eta \widehat{g}_k, \qquad y_{k+1} = \bar{y}_k. \tag{6}$$

Weight-coordinate update for classification. Write the lower objective as $\widetilde{f}(w, y)$ with $w = \sigma(x) \in (0,1)^n$ and define $\widetilde{C} = [\partial(\nabla_y \widetilde{f})/\partial w]^\top$. The validation objective in our classification tasks does not depend directly on w. The approximate weight hypergradient is therefore $\widehat{g}^w = -\widetilde{C} q$. The revised configuration takes a mirror step with negative Bernoulli entropy $\psi(w) = \sum_i [w_i \log w_i + (1 - w_i)\log(1 - w_i)]$. Since $\nabla\psi(w) = \operatorname{logit}(w) = x$, its dual update is

$$x_{k+1} = x_k - \eta \widehat{g}_k^w = x_k + \eta \widetilde{C}_k q_k, \qquad y_{k+1} = \bar{y}_k. \tag{7}$$

We update the logits using the weight hypergradient. The objective function is unchanged; the update geometry differs. Ridge and the initial implementation use Eq. (6). Appendix J details the configuration and derivative-backend ablations.

Computational cost. The retained adjoint requires $O(d_y)$ persistent storage beyond the $O(md_y)$ curvature memory. Applying the cached compact operator costs $O(md_y + m^2)$, and rebuilding it after memory changes costs $O(m^2 d_y)$. Each adjoint solve uses at most P fresh curvature probes, including the refresh, plus vector operations and any preconditioner applications. The structured backend temporarily stores current probabilities and a mixed-action accumulator. Wall-clock measurements include lower tracking, compact construction, and mixed differentiation.

3 LOCAL ANALYSIS OF ADJOINT RECYCLING

We characterize the recycled initialization through energy minimization, inter-iteration drift, and local hypergradient error. We first consider exact positive-definite systems, then account for the residual gap introduced by finite-difference actions.

Algorithm 1 Hq-Recycle (evaluated with $P = 4$, $\tau = 0.05$)

Require: x_0, y_0; lower cap $T_{\max}$ and tolerance ϵ_{low}; memory size m; probe cap P; tolerance τ

Initialize lower curvature memory $\mathcal{M} = \varnothing$, previous adjoint $q_{-1} = 0$
for $k = 0, 1, \ldots$ **do**
  Track $\bar{y}_k$ with up to $T_{\max}$ lower steps; update accepted pairs in $\mathcal{M}$
  Freeze memory defining B_k; form b_k and current action $\mathcal{D}_k$
  Set $q = 0$, $\widehat{r} = b_k$, probe count $j = 0$
  if $b_k = 0$ **then**
    Set $q_k = 0$ and proceed to the outer update
  else
    if q_{k-1} is finite and nonzero **then**
      $v = q_{k-1} / \|q_{k-1}\|$; $a = \mathcal{D}_k(v)$; $j \leftarrow j + 1$
      if $v^\top a$ is finite and positive **then**
        $\alpha = (v^\top b_k)/(v^\top a)$; $q = \alpha v$; $\widehat{r} = b_k - \alpha a$
      end if
    end if
    Correct $(q, \widehat{r})$ using B_k on demand with PCG recurrences while $\|\widehat{r}\| > \tau \|b_k\|$ and $j < P$; count each action and stop on breakdown
    Store $q_k = q$
  end if
  Update (x_{k+1}, y_{k+1}) by Eq. (6) or Eq. (7)
end for

Assumption 1 (Local exact system)**.** *The current lower Hessian A is symmetric and satisfies $A \succeq \mu I$ for $\mu > 0$. For correction steps, actions are exact and the inverse preconditioning operator B is fixed and symmetric positive definite (SPD).*

Write $q^\dagger = A^{-1}b$, $J(q) = \frac{1}{2}q^\top Aq - b^\top q$, and $\|z\|_A = (z^\top Az)^{1/2}$. For $u \neq 0$, we normalize the retained vector as $v = u/\|u\|$. This leaves the Galerkin estimate unchanged. We apply the standard energy-minimization property of a one-dimensional Galerkin solve to the retained approximate adjoint, then bound the effect of system drift. Appendix A provides the proofs.

Proposition 1 (Optimal rescaling of a retained direction)**.** *Under Assumption 1, for any nonzero retained vector u, the refreshed estimate*

$$q^{(0)} = \frac{u^\top b}{u^\top Au}\, u \tag{8}$$

minimizes J over $\operatorname{span}\{u\}$*. Consequently,*

$$\left\| q^{(0)} - q^\dagger \right\|_A \leq \min\{\left\| u - q^\dagger \right\|_A, \left\| q^\dagger \right\|_A\}. \tag{9}$$

The improvement over using u unchanged has the exact form

$$\left\| u - q^\dagger \right\|_A^2 - \left\| q^{(0)} - q^\dagger \right\|_A^2 = \left(1 - \frac{u^\top b}{u^\top Au}\right)^2 \|u\|_A^2 . \tag{10}$$

The energy gain equals the squared deviation of the optimal coefficient from one, multiplied by the retained vector's squared energy norm. Exact PCG corrections further minimize energy error over their growing affine search spaces. Energy error and the Euclidean stopping residual measure different aspects of the approximation (Appendix A.2); Section 4.4 examines both under the fixed budget.

Proposition 2 (Effect of inter-iteration drift)**.** *Let u be the previous approximate adjoint, with residual $r_{k-1} = b_{k-1} - A_{k-1}u$. Set $\Delta A_k = A_k - A_{k-1}$ and $\Delta b_k = b_k - b_{k-1}$. If $A_k \succeq \mu I$ and $u \neq 0$, the refreshed initialization satisfies*

$$\left\| q_k^{(0)} - q_k^\dagger \right\|_{A_k} \leq \frac{\|r_{k-1}\| + \|\Delta b_k\| + \|\Delta A_k\|\,\|u\|}{\sqrt{\mu}}. \tag{11}$$

Here the matrix norm is the Euclidean operator norm.

The bound links initialization accuracy to the previous residual and changes in the right-hand side and Hessian. Small residuals and limited drift favor accurate reuse. The current-system residual then determines whether further correction is needed.

Proposition 3 (Residual and local hypergradient error)**.** *If $A_k \succeq \mu I$ and the* true *residual satisfies $\|b_k - A_k q_k\| \le \tau \|b_k\|$, then*

$$\left\| q_k - q_k^\dagger \right\| \le \frac{\tau \|b_k\|}{\mu}, \qquad \|\widehat{g}_k - g(x_k, \bar{y}_k)\| \le \frac{\|C_k\| \tau \|b_k\|}{\mu}, \tag{12}$$

where $g(x,y) = \nabla_x F(x,y) - C(x,y)A(x,y)^{-1}b(x,y)$. If $g(x,\cdot)$ is L_g-Lipschitz along the segment joining $\bar{y}_k$ and $y^\star(x_k)$, then

$$\|\widehat{g}_k - \nabla\Phi(x_k)\| \le \frac{\|C_k\| \tau \|b_k\|}{\mu} + L_g \|\bar{y}_k - y^\star(x_k)\|. \tag{13}$$

Finite differences and a fixed budget. The implementation monitors a recursively updated estimated residual. If the lower Hessian is locally L_A-Lipschitz, the forward-difference action in Eq. (4) satisfies $\|\mathcal{D}_k(p) - A_k p\| \le (L_A \delta_k/2)\|p\|$ in exact arithmetic. Finite-difference actions need not be additive. Together with rounding, this can create a gap between the estimated residual and the true residual in Proposition 3. Appendix A derives the gap explicitly. The four-probe rule limits computation. The error bound also depends on accumulated action error. The same local analysis applies in weight coordinates with C_k replaced by $\widetilde{C}_k$ and the corresponding Lipschitz assumption.

4 Experiments

We first evaluate the complete method against established baselines, then use component ablations and matched initialization variants to explain its efficiency. Real-system replay and controlled drift isolate the effect of rescaling.

4.1 Tasks and evaluation protocol

Tasks and metrics. We evaluate grouped ridge regression with 256 lower variables and 16 regularization hyperparameters, and three linear classification tasks with 50% corrupted training labels: CIFAR-10 with frozen ImageNet-pretrained ResNet-18 features, Fashion-MNIST, and MNIST (Krizhevsky, 2009; He et al., 2016; Xiao et al., 2017; LeCun et al., n.d.). Classification uses $5{,}000/1{,}000/10{,}000$ train/validation/test examples. The primary confirmation metric is validation loss after an independent lower solve for hyperparameters learned within 0.5 seconds. We select configurations by validation loss and report test metrics as secondary outcomes. Loss is half MSE for ridge and cross-entropy for classification (Appendix C).

Baselines and HQ variants. The six-baseline study compares qNBO, SHINE, PZOBO, BOME, HOAG, and AID-CG. Confirmation includes qNBO, SHINE, HOAG, and warm-start AID-CG on ridge, and warm-start AID-CG on classification. Two matched variants of Hq-Recycle (HQ, also denoted G) test initialization: *without rescaling* (W, Warm-PCG) keeps the preceding adjoint unchanged; *without adjoint reuse* (Z, Zero-PCG) starts from zero. Both preserve HQ's lower tracking, preconditioner, derivative backend, update geometry, tolerance, and four-probe cap. G and W count the current-system refresh toward this cap. Archived AID-CG results use a cold-start adapter and are labeled accordingly; confirmation uses the corrected warm-start adapter with its native CG budget (Appendix D).

Selection, repeats, and timing. On ridge, we evaluate all seven configurations at outer rates $\{30, 100, 300, 900, 2700\}$ with other settings fixed. We select each rate by mean one-second independent-response validation loss on development seeds $601, 602, 1302$. CIFAR-10 uses three candidate rates per method; Fashion-MNIST and MNIST retain frozen settings. Confirmation uses ten new seeds with three timing repeats each. We average repeats within seed before computing means, sample SDs, and paired 95% bootstrap intervals. We run configurations sequentially on one RTX 4090 D in float32 with TF32 disabled. CUDA-synchronized timing includes complete optimizer steps and excludes data loading, discarded warmup, evaluation, and checkpoint transfer. Each endpoint is the last complete state within its budget. Appendices D and G give all settings and distinguish the three evaluation cohorts.

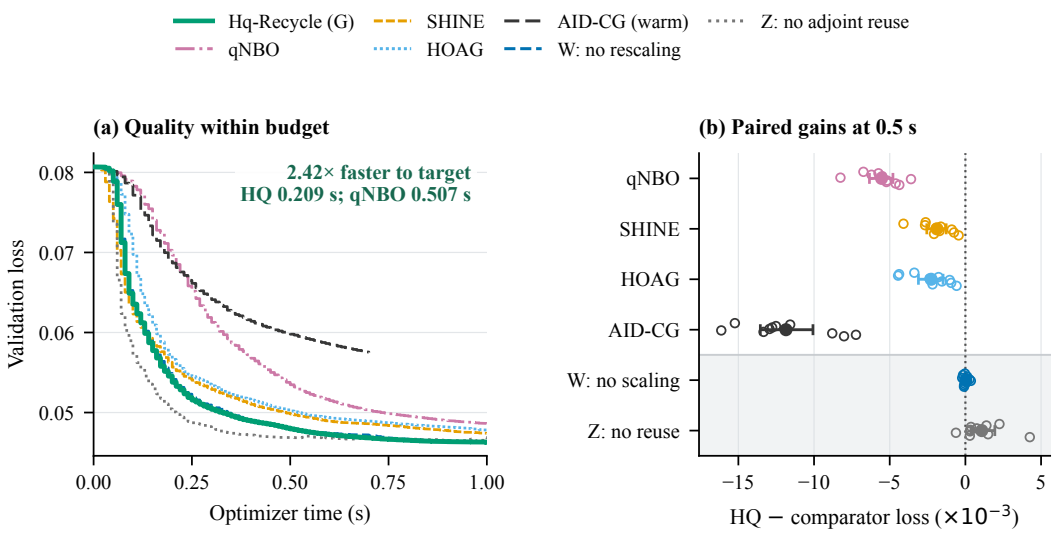

Figure 1: **Ridge efficiency after an expanded outer-rate search.** Left: mean independent-response validation loss versus optimizer time; the callout reports mean first-observed time to the development-fixed target. Right: paired 0.5-second loss differences, HQ minus comparator; negative values favor HQ. Open dots are ten seed-level differences; filled markers and bars show means and paired 95% bootstrap intervals. W and Z are matched HQ variants. All methods have all 30 primary endpoints; curves stop when full seed/repeat coverage is lost.

Table 1: **Ridge confirmation at 0.5 seconds.** Independent-response losses, mean (sample SD) across ten seeds after averaging three repeats per seed. All rows have 30/30 endpoints. Rates are selected on development validation loss. Matched HQ variants are grouped separately; bold values mark the lowest mean across all rows.

Method / variant	Outer rate	Validation loss	Test loss
Hq-Recycle	900	0.04793 (0.00315)	0.05129 (0.00273)
qNBO	900	0.05347 (0.00273)	0.05762 (0.00389)
SHINE	900	0.04981 (0.00317)	0.05305 (0.00237)
HOAG	900	0.05019 (0.00330)	0.05332 (0.00242)
AID-CG (warm)	300	0.05979 (0.00350)	0.06386 (0.00534)
Matched HQ variants			
No rescaling (W)	900	0.04791 (0.00324)	0.05133 (0.00273)
No adjoint reuse (Z)	2700	**0.04685** (0.00321)	**0.04989** (0.00314)

4.2 Short-budget efficiency against established baselines

Faster progress with a small probe budget. For the development-fixed 75% validation target, Hq-Recycle achieves a 2.42× speedup over qNBO in mean time to target, with a paired 95% interval of $[2.15, 2.70]$. Mean first-observed times are 0.209 seconds for HQ and 0.507 for qNBO; SHINE and HOAG take 0.273 and 0.291 seconds. All four methods attain the target in every timing repeat. At 0.5 seconds, HQ reduces validation loss by 10.36%, 3.78%, 4.51%, and 19.83% relative to qNBO, SHINE, HOAG, and warm-start AID-CG, respectively. Each comparison favors HQ in all ten paired seeds (Figure 1, Table 1). Test loss is 10.98% below qNBO's.

The original six-baseline results, including PZOBO and BOME, remain in Appendix M; independent classification lower-response checks appear in Appendix K. We retain all valid early endpoints from the confirmation runs. At three seconds, qNBO and warm-start AID-CG have 12/30 and 27/30 failed ridge trajectories under the selected configurations. Appendix D.4 records these failures and the endpoint-persistence correction. Curves stop when full seed/repeat coverage is lost.

Table 2: **Component contributions to throughput.** Full HQ divided by each ablated variant at one second, mean (sample SD) over five paired seeds. Values above one favor full HQ. All variants keep the outer rate fixed. These are iteration-throughput ratios; work to a quality target is analyzed separately. Ridge uses a fixed single lower step. Paired loss changes appear in Appendix J.

Ablated variant	CIFAR-10	Fashion	MNIST	Ridge
No adjoint reuse	2.12 (0.30)	1.81 (0.23)	1.95 (0.07)	1.27 (0.09)
No lower stopping	2.59 (0.33)	1.80 (0.20)	2.62 (0.08)	–
No residual stopping	1.66 (0.21)	1.57 (0.15)	1.66 (0.23)	1.20 (0.08)
Eager preconditioning	1.08 (0.12)	1.07 (0.09)	1.08 (0.06)	0.98 (0.06)
Two-loop kernel	1.18 (0.19)	1.11 (0.10)	1.18 (0.08)	1.17 (0.04)
Identity adjoint precond.	1.06 (0.14)	0.99 (0.08)	1.02 (0.03)	1.01 (0.10)

Table 3: **Classification confirmation at** 0.5 **seconds.** Validation cross-entropy (↓) and test accuracy (%, ↑) after independent lower optimization, averaged over ten seeds. Three repeats are averaged within seed. W removes rescaling; Z removes adjoint reuse. Full SDs and test losses appear in Appendix D.3.

	CIFAR-10		Fashion-MNIST		MNIST	
Method / variant	Val. CE	Acc.	Val. CE	Acc.	Val. CE	Acc.
Hq-Recycle	1.51203	74.293	1.17930	71.696	1.12915	83.301
No rescaling (W)	1.51203	74.292	1.17930	71.697	1.12914	83.303
No adjoint reuse (Z)	1.51276	74.166	1.18016	71.714	1.12988	83.107
AID-CG (warm)	1.51704	74.328	1.18865	72.021	1.13679	83.211

4.3 Component contributions and matched HQ variants

Which components reduce computation? Single-component ablations in the five-seed implementation study hold all other settings fixed. Table 2 reports throughput, and Appendix J gives paired validation losses for every intervention. Lower stopping increases classification throughput by 1.80–$2.62\times$; deferred preconditioning adds 6.6–8.2%. On classification, retaining the adjoint increases throughput by 1.81–$2.12\times$ relative to the fixed-rate no-reuse variant, with lower validation loss in all 15 task–seed pairs. Disabling residual stopping slightly improves classification validation loss but requires more computation. Together, the implementation and configuration changes yield 2.33–$3.28\times$ the classification throughput of the initial implementation, with small accuracy changes.

What does initialization contribute after rate selection? For ten-seed confirmation, we also select rates for G, W, and Z independently on ridge and CIFAR-10. On ridge, W reaches the 75% target in 0.218 seconds and attains a validation loss close to HQ: the paired difference, HQ minus W, is 2.10×10^{-5}, with interval $[-8.04 \times 10^{-5}, 1.30 \times 10^{-4}]$. Z's selected rate is 2700, the grid's upper boundary. It reaches the target in 0.144 seconds and has lower 0.5-second validation loss than HQ in nine of ten seeds. Fixed-rate ablations isolate a component's contribution within a chosen configuration; independently tuned variants compare performance after rate selection.

On classification, G and W attain similar mean validation losses, both below those of Z and warm-start AID-CG after the common lower solve (Table 3). Test accuracy rankings vary by task. All 360 trajectories complete, and the maximum lower-gradient norm after independent optimization is 1.57×10^{-8}. The independent solves evaluate the learned sample weights through their lower responses.

4.4 Why and when does rescaling help?

Local accuracy on identical real systems. We collect 100 states from each G/W trajectory on ridge and CIFAR-10 for three development seeds. At 40 anchors per trajectory, G, W, Z, and a residual-minimizing scalar variant R solve the same current (A, b) with a common frozen preconditioner and historical approximate adjoint. All have a four-probe budget including refresh. We

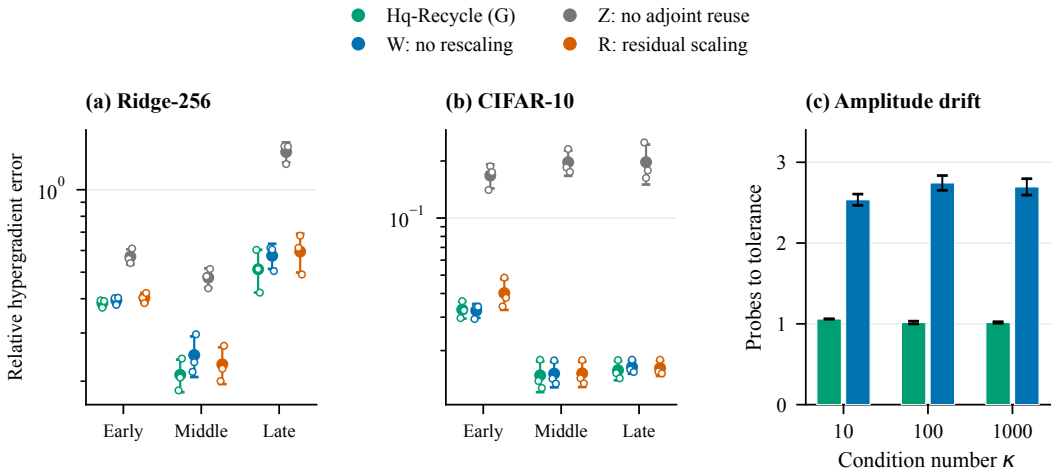

Figure 2: **Accuracy and work saved by adapting the retained direction.** (a,b) Final local hypergradient relative error on identical real-system anchors with production finite differences under the four-probe cap; points average both source trajectories within each seed, and bars show mean and sample SD across three seeds. Early/middle/late denote steps 1–10/11–50/51–100. Classification uses sample-weight coordinates. (c) Probes to true residual tolerance 0.05 under controlled amplitude drift, mean and sample SD across five seeds, conditioned on joint G/W attainment. The three conditions retain $250/250$, $242/250$, and $238/250$ states. W/Z remove rescaling/adjoint reuse; R uses residual-minimizing scaling.

measure error against high-accuracy reference adjoints at the tracked lower state. Appendix E also reports exact-HVP and sequential-history replay.

On early/middle ridge anchors, G reduces final relative local hypergradient error versus W by 8.37%, 7.73%, and 12.11% across the three seeds; final direction cosines also improve (Figure 2). CIFAR-10 differences are smaller and mixed. Positive rescaling alone preserves direction: the final directional gains arise after correction and stopping. These local accuracy gains coexist with similar G/W end-to-end losses. Across $13{,}440$ solves, every tolerance-based stop also meets the true relative residual threshold of 0.05. R reduces the estimated initialization residual but worsens final hypergradient accuracy. We therefore retain Galerkin scaling under the predetermined selection criterion (Appendix E).

Amplitude adaptation reduces correction work. In 128-dimensional SPD sequences, amplitude drift favors rescaling: G reaches the true residual tolerance in approximately one probe, compared with 2.54–2.74 for W, at condition numbers 10, 100, and 1000 (Figure 2c). Direction drift reduces this saving. Because G stops earlier, its energy error can remain higher than W's even when both meet the tolerance. Appendix F reports unconditional attainment, the two-parameter drift grid, and real-trajectory coverage.

Work before reaching a quality target. On ridge, G uses 171.1 cumulative probes versus W's 180.6 before the 90% target, a 5.3% reduction across the nine seeds for which both methods attain the target in every timing repeat. Before the classification targets, G/W/Z all use four probes per update; the approximately 1.07 whole-trajectory average primarily reflects cheaper iterations after target attainment. On CIFAR-10, G and W reach the target in eight cumulative probes versus Z's twelve; Fashion-MNIST and MNIST counts are essentially equal. Figure 3 in Appendix D.5 separates these stages.

5 Discussion and limitations

Hq-Recycle adapts a retained adjoint within a fixed probe budget. Lower tracking supplies reusable curvature, and a fresh action adapts the adjoint to the current system. When the refreshed estimate passes the residual test, the solver avoids further corrections and preconditioner applications. The complete method improves short-budget ridge performance against the evaluated qNBO, SHINE, HOAG, and AID-CG configurations. Component ablations and matched variants show how these savings depend on initialization, stopping, and outer-step selection.

The experiments cover strongly convex linear lower problems on one GPU, with outer-rate searches that hold other method settings fixed. Future work should compare against larger recycled-space methods such as LancBiO and evaluate nonconvex or stochastic lower problems. The analysis characterizes local adjoint error; convergence of the outer method with adaptive stopping and finite-difference actions remains open.

6 Conclusion

We introduced Hq-Recycle, a compact quasi-Newton method that refreshes, rescales, and corrects one retained adjoint. In ten-seed ridge confirmation, it achieves a $2.42\times$ speedup over qNBO in mean time to target. Component ablations identify substantial reductions in iteration cost, and local theory and controlled experiments clarify how rescaling affects accuracy and correction work. Adapting a single adjoint thus offers a practical, low-storage approach to bilevel optimization under tight computational budgets.

AI use statement

We used AI to assist with language polishing and presentation, literature retrieval and the identification of related work, the verification and refinement of mathematical derivations and proofs, and minor revisions to experimental code. The authors independently reviewed and validated all AI-assisted outputs and take full responsibility for the final content of this work, including its methods, theoretical claims, experimental results, and references.

Reproducibility statement. The appendix specifies objectives, data construction, implementation details, search grids, selected settings, failure accounting, and proofs. The experiment package contains raw metrics, checkpoints, frozen configurations, data and source hashes, and reproduction scripts. Figures and tables use these recorded outputs.

REFERENCES

Youran Dong, Junfeng Yang, Wei Yao, and Jin Zhang. Efficient curvature-aware hypergradient approximation for bilevel optimization. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 14231–14259. PMLR, 2025. URL https://proceedings.mlr.press/v267/dong25m.html.

Matthias J. Ehrhardt, Silvia Gazzola, and Sebastian J. Scott. Efficient gradient-based methods for bilevel learning via recycling krylov subspaces. *arXiv preprint arXiv:2412.08264*, 2024. doi: 10.48550/arXiv.2412.08264. URL https://arxiv.org/abs/2412.08264.

Sheng Fang, Yongjin Liu, Wei Yao, Chengming Yu, and Jin Zhang. qNBO: quasi-newton meets bilevel optimization. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/6d44b1617ca7001a5ba7ea4f28039d41-Paper-Conference.pdf.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, 2016. URL https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html.

Kaiyi Ji, Junjie Yang, and Yingbin Liang. Bilevel optimization: Convergence analysis and enhanced design. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 4882–4892. PMLR, 2021. URL https://proceedings.mlr.press/v139/ji21c.html.

Alex Krizhevsky. Learning multiple layers of features from tiny images. Technical report, University of Toronto, 2009. URL https://www.cs.toronto.edu/~kriz/cifar.html.

Yann LeCun, Corinna Cortes, and Christopher J. C. Burges. The MNIST database of handwritten digits, n.d. URL https://yann.lecun.org/exdb/mnist/index.html.

Bo Liu, Mao Ye, Stephen Wright, Peter Stone, and Qiang Liu. BOME! bilevel optimization made easy: A simple first-order approach. In *Advances in Neural Information Processing Systems*, volume 35, pp. 17248–17262, 2022. doi: 10.52202/068431-1254. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/6dddcff5b115b40c998a08fbd1cea4d7-Paper-Conference.pdf.

Fabian Pedregosa. Hyperparameter optimization with approximate gradient. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48, pp. 737–746, 2016. URL https://proceedings.mlr.press/v48/pedregosa16.html.

Zaccharie Ramzi, Florian Mannel, Shaojie Bai, Jean-Luc Starck, Philippe Ciuciu, and Thomas Moreau. SHINE: SHaring the INverse Estimate from the forward pass for bi-level optimization and implicit models. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=-ApAkox5mp.

Daouda Sow, Kaiyi Ji, and Yingbin Liang. On the convergence theory for hessian-free bilevel algorithms. In *Advances in Neural Information Processing Systems*, volume 35, pp. 4136–4149, 2022. doi: 10.52202/068431-0299. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/1a82986c9f321217f2ed407a14dcfa0b-Paper-Conference.pdf.

Han Xiao, Kashif Rasul, and Roland Vollgraf. Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747*, 2017. URL https://arxiv.org/abs/1708.07747.

Yan Yang, Bin Gao, and Ya-xiang Yuan. LancBiO: Dynamic lanczos-aided bilevel optimization via krylov subspace. In *International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/file/54ba04e78dbd019300f6731bdd46ca16-Paper-Conference.pdf.

A PROOFS AND THE FINITE-DIFFERENCE RESIDUAL GAP

All norms without a subscript are Euclidean norms or their induced operator norms. The analysis concerns the local adjoint system at the current tracked lower state.

A.1 OPTIMAL RESCALING

Proof of Proposition 1. For $u \neq 0$, positive definiteness gives $u^\top Au > 0$. Restricting J to tu yields

$$J(tu) = \tfrac{1}{2}t^2 u^\top Au - tu^\top b,$$

whose unique scalar minimizer is $t = (u^\top b)/(u^\top Au)$. Also,

$$J(q) - J(q^\dagger) = \tfrac{1}{2}\left\|q - q^\dagger\right\|_A^2.$$

Both u and zero belong to $\mathrm{span}\{u\}$, so comparing their objectives with the minimizer proves Eq. (9). The residual is orthogonal to u because $u^\top(b - Aq^{(0)}) = 0$. For Eq. (10), write $q^{(0)} = cu$. Galerkin orthogonality gives $\left\langle u, q^{(0)} - q^\dagger \right\rangle_A = 0$, where $\langle s, t\rangle_A = s^\top At$. Expanding $u - q^\dagger = (1-c)u + (q^{(0)} - q^\dagger)$ therefore yields the stated squared-error difference. □

The scalar coefficient can be negative. Accordingly, the minimization is over a linear span, not a nonnegative ray. Normalizing u changes neither the minimizing vector nor this exact-system argument.

A.2 ENERGY ERROR AND THE STOPPING RESIDUAL

Take $A = \mathrm{diag}(1, 2)$, $b = (3, 1)^\top$, and a retained vector $u = (1, 1)^\top$. Then $q^\dagger = (3, 1/2)^\top$ and Galerkin rescaling gives $q^{(0)} = (4/3, 4/3)^\top$. The squared energy error decreases from $9/2$ to $25/6$, while the squared Euclidean residual increases from 5 to $50/9$. The stopping rule can therefore make a different decision from the energy comparison, even with exact SPD actions.

At a fixed anchor with no direct upper derivative, rescaling u to αu also rescales its mixed hypergradient by α. For positive α and no correction, the direction is preserved and the magnitude changes. Across iterations, rescaling also affects residual decisions and subsequent states. The matched warm-PCG control isolates its empirical contribution from that of retaining an adjoint.

A.3 DRIFT BETWEEN SYSTEMS

Proof of Proposition 2. The residual of the unchanged old estimate in the current system is exactly

$$\begin{aligned} b_k - A_k u &= b_{k-1} - A_{k-1}u + (b_k - b_{k-1}) - (A_k - A_{k-1})u \\ &= r_{k-1} + \Delta b_k - \Delta A_k u. \end{aligned}$$

Writing this residual as r, we have

$$\left\|u - q_k^\dagger\right\|_{A_k}^2 = r^\top A_k^{-1} r \leq \frac{\|r\|^2}{\mu}.$$

The triangle inequality and submultiplicativity bound $\|r\|$ by $\|r_{k-1}\| + \|\Delta b_k\| + \|\Delta A_k\|\,\|u\|$. Applying Proposition 1 in the current system gives the result for $q_k^{(0)}$. □

The bound gives a sufficient condition for accurate reuse in terms of the previous residual and system drift.

A.4 RESIDUAL CONTROL AND LOWER TRACKING

Proof of Proposition 3. For $r = b_k - A_k q_k$, $q_k - q_k^\dagger = -A_k^{-1} r$ and $\left\|A_k^{-1}\right\| \leq 1/\mu$. Hence $\left\|q_k - q_k^\dagger\right\| \leq \|r\|/\mu$. Subtracting the two local hypergradients gives $\widehat{g}_k - g(x_k, \bar{y}_k) = -C_k(q_k - q_k^\dagger)$, which proves Eq. (12). Finally,

$$\|\widehat{g}_k - \nabla\Phi(x_k)\| \leq \|\widehat{g}_k - g(x_k, \bar{y}_k)\| + \|g(x_k, \bar{y}_k) - g(x_k, y^\star(x_k))\|.$$

Use the assumed Lipschitz property in the second term to obtain Eq. (13). □

The decomposition separates adjoint approximation error from lower-state tracking error; controlling both terms is the key requirement for an outer convergence analysis.

A.5 A DIRECTIONAL FINITE-DIFFERENCE BOUND

Fix (x, y) and suppose $A(y) = \nabla^2_{yy} f(x, y)$ is L_A-Lipschitz on the probed segments. For $h > 0$ and $p \neq 0$, the fundamental theorem of calculus gives

$$\frac{\nabla_y f(x, y + hp) - \nabla_y f(x, y)}{h} - A(y)p = \int_0^1 [A(y + thp) - A(y)]p \, dt.$$

Taking norms bounds this by $(L_A h/2) \|p\|^2$. With $h = \delta/\|p\|$, the error is at most $(L_A \delta/2)\|p\|$. This argument is conditional on local Hessian Lipschitz continuity and exact arithmetic; it does not control cancellation in float32.

An explicit residual gap quantifies the difference between the estimated and true residuals. Let $e(p) = \mathcal{D}(p) - Ap$ at a fixed anchor. If the initialization is $q^{(0)} = c_0 v$, followed by updates $q^{(j+1)} = q^{(j)} + \alpha_j p_j$, then the recursive residual after s corrections is

$$\widehat{r} = b - c_0 \mathcal{D}(v) - \sum_{j=0}^{s-1} \alpha_j \mathcal{D}(p_j).$$

The true residual obeys

$$r - \widehat{r} = c_0 e(v) + \sum_{j=0}^{s-1} \alpha_j e(p_j), \qquad \|r - \widehat{r}\| \leq \frac{L_A \delta}{2} \left(|c_0| \, \|v\| + \sum_{j=0}^{s-1} |\alpha_j| \, \|p_j\| \right). \tag{14}$$

The initial term is absent for a zero start. Replacing $\tau \|b\|$ in Proposition 3 with $\|\widehat{r}\| + \|r - \widehat{r}\|$ yields a conditional bound for approximate actions. Equation (14) quantifies how directional errors accumulate in the recursive residual. The implementation monitors $\widehat{r}$; it does not estimate L_A or the accumulated gap during training.

B IMPLEMENTATION DETAILS

B.1 TRACKING, CURVATURE ACCEPTANCE, AND COMPACT APPLICATION

We use a memory size of $m = 10$, a curvature threshold of $\epsilon = 10^{-10}$, and no periodic resets. A pair is accepted when $\|s\| > \epsilon$, $\|w\| > \epsilon$, and $s^\top w > \epsilon \|s\| \|w\|$. A newly accepted pair is immediately available to the next inner step. The lower search direction is clipped to have norm at most 10. Backtracking tests up to ten halvings for a nonincreasing lower objective; if none passes, the implementation takes the final reduced step without a further monotonicity guarantee. The revised nominal lower step size is 1 on all four tasks; the initial ridge implementation uses 0.1.

For nonempty history, let $S = [s_1, \ldots, s_m]$, $Y = [w_1, \ldots, w_m]$, and

$$R = \mathrm{triu}(S^\top Y) + \epsilon I, \qquad D = \mathrm{diag}(R), \qquad \gamma = \max\left\{ \frac{s_m^\top w_m}{w_m^\top w_m + \epsilon}, 10^{-4} \right\}.$$

Here D denotes the diagonal matrix with the diagonal entries of R. The implemented operator applied to v is

$$a = R^{-1} S^\top v, \qquad z = \gamma(v - Ya), \qquad Bv = z + SR^{-\top}(Da - Y^\top z). \tag{15}$$

Triangular solves and matrix–vector products replace the sequential two-loop application. Empty history gives the identity. The factorization data are cached until the history changes, and construction time is included in solver timing. Numerical trajectory checks verify agreement with the corresponding stabilized two-loop implementation.

B.2 Correction recurrences and guards

Given $q, \widehat{r}$, form $z = B\widehat{r}$, $p = z$, and $\rho = \widehat{r}^\top z$. For each available probe, compute $a = \mathcal{D}(p)$ and perform

$$\alpha = \rho/(p^\top a), \qquad q \leftarrow q + \alpha p, \qquad \widehat{r} \leftarrow \widehat{r} - \alpha a.$$

If the residual test fails and budget remains, set

$$z = B\widehat{r}, \quad \rho_{\text{new}} = \widehat{r}^\top z, \quad p \leftarrow z + (\rho_{\text{new}}/\rho)p, \quad \rho \leftarrow \rho_{\text{new}}.$$

Correction stops if a preconditioned residual product or directional curvature is non-finite or nonpositive. A zero right-hand side yields a zero adjoint estimate without probing. All actions, including a discarded refresh, count against the four-probe cap. No additional exact-residual recomputation is performed. The finite-difference anchor stays fixed during each solve, as does B.

B.3 Paired finite-difference and mixed actions

For the weighted softmax objective, let $\mathcal{L}$ map classifier parameters to the $n \times c$ training-logit matrix, including biases, and let $\mathcal{L}^*$ be its Euclidean adjoint. At the fixed anchor, write $Z = \mathcal{L}(\bar{y})$, $\Pi = \mathrm{softmax}(Z)$, and $E = \Pi - Y$, where Y contains one-hot labels. With rowwise broadcasting of $w = \sigma(x)$, the primary backend computes

$$\mathcal{D}(p) = \mathcal{L}^*\left(\frac{w \odot [\mathrm{softmax}(Z + h_p\mathcal{L}(p)) - \Pi]}{nh_p}\right) + 0.002p, \tag{16}$$

$$(\widetilde{C}p)_i = \frac{1}{n}\left\langle E_i, \mathcal{L}(p)_i \right\rangle. \tag{17}$$

These expressions share $\mathcal{L}(p)$ and represent the same objective as generic differentiation. The first is still a forward difference. Within a solve, accumulated mixed actions track the linear combination defining q because $\widetilde{C}$ is linear at that anchor; this does not assume additivity of $\mathcal{D}$. No Hessian action is carried across outer iterations. Current probabilities and mixed accumulators are temporary working storage, not extra retained adjoint directions.

The exact-HVP extension replaces the softmax difference quotient by its Jacobian action, $\Pi \odot [\mathcal{L}(p) - s\mathbf{1}_c^\top]$, where $s_i = \langle \Pi_i, \mathcal{L}(p)_i \rangle$. It is labeled separately and excluded from the finite-difference main table. Ridge uses the generic finite-difference backend and automatic differentiation for the mixed action.

B.4 Deferred preconditioning and lower stopping

The lower solve updates its curvature memory as usual. The adjoint solve uses this current memory and constructs or applies the compact operator only when residual correction requires it. A successful refresh followed by residual acceptance avoids an adjoint-preconditioner application. The eager control additionally evaluates $B_k b_k$ before checking whether the retained adjoint suffices.

Classification takes at least one lower step and stops after a completed step when the lower-gradient norm is at most 10^{-4}. The caps are 5, 3, and 5 steps for CIFAR-10, Fashion-MNIST, and MNIST, respectively. Ridge takes one lower step. The identity-adjoint ablation uses the identity operator only in the adjoint solve; lower directions still use L-BFGS. Integration tests cover finite-difference and exact backends, deferred construction, generic objectives with a nonzero direct upper derivative, budget guards, and this ablation boundary.

C Tasks and data construction

C.1 Label-corrupted linear classification

For each seed, a CPU random permutation selects 6,000 examples from the official training split. The first 5,000 form the lower training set and the next 1,000 the clean upper validation set. Of the lower labels, exactly 2,500 are selected without replacement and shifted by a uniformly sampled integer in $\{1, \ldots, 9\}$ modulo ten. Hence every selected label changes. The official test set is untouched and is never used for configuration selection.

CIFAR-10 uses precomputed 512-dimensional features from an ImageNet-pretrained ResNet-18 (Krizhevsky, 2009; He et al., 2016); Fashion-MNIST and MNIST use flattened 784-dimensional pixels (Xiao et al., 2017; LeCun et al., n.d.). Inputs are converted to float32, divided by 255, and normalized rowwise by their Euclidean norm, with a denominator floor of 10^{-12}. The same transformation is applied to train, validation, and test inputs. The scalar division before normalization is inherited by the feature-input loader as well. Classifier parameters $y = (W, c)$ comprise weights and bias for ten classes, so $d_y = 5{,}130$ for CIFAR-10 and $7{,}850$ otherwise. Both $x \in \mathbb{R}^{5000}$ and y start at zero.

Writing $\ell_i(y)$ for cross-entropy and σ for the sigmoid, we use

$$f(x,y) = \frac{1}{5000}\sum_{i=1}^{5000} \sigma(x_i)\ell_i(y) + 10^{-3}\left\|y\right\|^2, \qquad F(x,y) = \frac{1}{1000}\sum_{j=1}^{1000} \ell_j^{\mathrm{val}}(y). \tag{18}$$

Biases are regularized too, giving $\nabla^2_{yy} f \succeq 0.002I$. The upper validation set supplies the optimization objective. Test metrics are evaluated at the tracked lower state or after an independent lower solve, as specified for each experiment. The primary classification confirmation uses the latter.

C.2 Grouped ridge regression

For each CUDA seed, we sample Gaussian feature matrices with $d = 256$ columns whose standard deviations are logarithmically spaced from 0.1 to 1. A teacher vector has independent $N(0, 1/32)$ entries in its first 32 coordinates and zeros elsewhere. Targets equal the linear response plus independent Gaussian noise of standard deviation 0.3. Train, validation, and test sizes are 512, 256, and $1{,}024$.

Let $g(j) = \lfloor 16j/256 \rfloor$ for zero-indexed feature j and $\lambda_j(x) = 10^{-4} + \mathrm{softplus}(x_{g(j)})$. Then

$$f(x,y) = \frac{\left\|X_{\mathrm{tr}}y - t_{\mathrm{tr}}\right\|^2}{2n_{\mathrm{tr}}} + \frac{1}{2}\sum_{j=1}^{256}\lambda_j(x)y_j^2, \qquad F(x,y) = \frac{\left\|X_{\mathrm{val}}y - t_{\mathrm{val}}\right\|^2}{2n_{\mathrm{val}}}. \tag{19}$$

We initialize the 16 upper variables to -6 and the lower vector to zero. Training uses the equivalent precomputed Gram formulation. At evaluation, we independently solve the current ridge system in float64, cast the solution to float32, and require a relative linear residual of at most 10^{-5}. Reported primary validation and test losses use this lower response. The reference solve is excluded from solver time, and the learned x comes from each method's tracked trajectory. This controlled synthetic task measures early progress in hyperparameter optimization.

D Ten-seed confirmation protocol and complete results

We conduct confirmation on seeds 26092601–26092610, which had not appeared in the preceding experiments, under the frozen selection protocol below. The initial cohort (610–614), implementation cohort (1301–1305), and confirmation cohort are distinct. Seed 1302, whose earlier result motivated further diagnosis, serves only as a development seed in this study. We select configurations without consulting confirmation test results.

D.1 Search grids and selected configurations

For ridge, we evaluate each of seven methods at outer rates $30, 100, 300, 900, 2700$ on development seeds $601, 602, 1302$, with a one-second budget. Among settings that complete all three development runs, we select the rate with the lowest mean independent-response validation loss. Of 105 attempted development trajectories, 99 complete. Warm-start AID-CG fails at rates 900 and 2700 on all three seeds. The selected rates are G/W/qNBO/SHINE/HOAG: 900; Z: 2700; warm-start AID-CG: 300. All other settings remain fixed: qNBO retains Q4-BFGS with its inherited recipe, SHINE and HOAG retain recipe 0, and AID-CG retains recipe 1 (Appendix G). The methods thus receive the same additional outer-rate search, with their earlier development histories preserved.

For CIFAR-10, we evaluate G/W/Z at rates $3000, 10000, 30000$ and warm-start AID-CG at rates $30000, 100000, 300000$. The grids apply the same three multipliers to previously selected centers

for methods with different update geometries. We select by mean 0.5-second validation loss after independent float64 lower optimization on the same three development seeds. The selected rates are G/W: 30000; Z: 10000; AID-CG: 100000. For Fashion-MNIST and MNIST, we retain the previously fixed G/W/Z rate 10000 and AID-CG rate 100000 without task-specific retuning. G/W/Z retain the lower caps and stopping settings in Table 10.

Corrected warm-start AID-CG. The AID-CG adapter in the initial and implementation studies starts its adjoint at zero each iteration. The corrected adapter retains the preceding approximate adjoint and recomputes its residual under the current Hessian before CG correction. We count this extra HVP. Its inherited lower-gradient descent and native CG cap remain unchanged; the four-probe restriction applies only to the matched G/W/Z controls. The labels AID-CG (warm) and AID-CG (cold) distinguish these implementations throughout the results.

D.2 Timing repeats, primary endpoints, and uncertainty

We run each selected method three times per data seed, starting each repeat with a fresh optimizer and state. We randomize method order within seed and execute all timed jobs serially. Two separate warmup steps are discarded. Complete CUDA-synchronized optimizer steps count toward the budget; evaluation, reference solves, and checkpoint writes do not. A budget endpoint is the last completed state with cumulative measured time no greater than the budget. For curves, we carry the last observation forward on a common grid and stop when full seed/repeat coverage is lost. We record first target attainment from the original observations, without interpolation, at an approximately 0.01-second logging interval.

The primary comparison is the 0.5-second independent-response validation loss. The endpoints at 0.2, 1, and 3 seconds and all test metrics are auxiliary. Timing repeats are averaged within a seed; sample SDs summarize these ten seed means. To compute paired percentile bootstrap intervals, we resample data seeds jointly across methods, with $10{,}000$ resamples and random seed 260926. These intervals describe variation across the sampled data seeds and timing runs under the stated implementation and hardware. Classification seeds share the official test set. A non-significant difference is not evidence of equivalence.

Table 4: All ridge confirmation endpoints: mean (sample SD) independent-response losses. Superscripts identify fewer than ten seeds with all three timing repeats available. These incomplete late-budget means are conditional on completion and are not full-cohort rankings; "–" SD denotes only one seed. Every 0.5-second entry has all ten seeds.

Method	0.2 s	0.5 s	1 s	3 s
Validation loss				
Hq-Recycle	0.0537 (0.0031)	0.0479 (0.0032)	0.0463 (0.0030)	0.0461 (0.0030)
Warm-PCG	0.0540 (0.0035)	0.0479 (0.0032)	0.0463 (0.0030)	0.0461 (0.0030)
Zero-PCG	0.0500 (0.0032)	0.0469 (0.0032)	0.0465 (0.0030)	0.0464 (0.0031)
qNBO	0.0697 (0.0046)	0.0535 (0.0027)	0.0486 (0.0028)	$0.0480\ (0.0018)^{6/10}$
SHINE	0.0557 (0.0028)	0.0498 (0.0032)	0.0474 (0.0031)	0.0464 (0.0030)
HOAG	0.0566 (0.0030)	0.0502 (0.0033)	0.0478 (0.0032)	0.0463 (0.0030)
AID-CG (warm)	0.0687 (0.0053)	0.0598 (0.0035)	$0.0553\ (0.0029)^{9/10}$	$0.0506\ (–)^{1/10}$
Test loss				
Hq-Recycle	0.0568 (0.0039)	0.0513 (0.0027)	0.0494 (0.0031)	0.0494 (0.0031)
Warm-PCG	0.0568 (0.0035)	0.0513 (0.0027)	0.0494 (0.0030)	0.0493 (0.0031)
Zero-PCG	0.0530 (0.0030)	0.0499 (0.0031)	0.0495 (0.0032)	0.0494 (0.0031)
qNBO	0.0718 (0.0053)	0.0576 (0.0039)	0.0520 (0.0031)	$0.0489\ (0.0032)^{6/10}$
SHINE	0.0592 (0.0041)	0.0531 (0.0024)	0.0509 (0.0028)	0.0495 (0.0030)
HOAG	0.0602 (0.0053)	0.0533 (0.0024)	0.0511 (0.0027)	0.0494 (0.0032)
AID-CG (warm)	0.0712 (0.0057)	0.0639 (0.0053)	$0.0592\ (0.0052)^{9/10}$	$0.0516\ (–)^{1/10}$

D.3 Classification lower-response checks

The classification study comprises 360 trajectories: three tasks, four methods, ten seeds, and three repeats. All complete and yield 1440 raw budget endpoints. At the primary 0.5-second endpoint,

we additionally optimize the lower objective for all 360 states with the common float64 procedure in Appendix K; the maximum lower-gradient norm is 1.5653×10^{-8}. Table 5 reports these independent-response results. The auxiliary endpoints in Table 6 evaluate tracked classifiers, as only the primary endpoint receives the additional lower solve.

Table 5: Classification confirmation at 0.5 seconds after independent lower optimization: mean (sample SD) over ten data seeds, after averaging three timing repeats within seed. Losses are cross-entropies; accuracy is in percent.

Task	Method	Validation loss	Test loss	Test accuracy
CIFAR-10	Hq-Recycle	1.51203 (0.01013)	1.52378 (0.00366)	74.293 (0.699)
	Warm-PCG	1.51203 (0.01014)	1.52380 (0.00366)	74.292 (0.695)
	Zero-PCG	1.51276 (0.01016)	1.52359 (0.00345)	74.166 (0.623)
	AID-CG (warm)	1.51704 (0.01021)	1.52808 (0.00357)	74.328 (0.659)
Fashion-MNIST	Hq-Recycle	1.17930 (0.01534)	1.19372 (0.00264)	71.696 (1.024)
	Warm-PCG	1.17930 (0.01534)	1.19369 (0.00265)	71.697 (1.030)
	Zero-PCG	1.18016 (0.01531)	1.19337 (0.00260)	71.714 (1.096)
	AID-CG (warm)	1.18865 (0.01556)	1.20193 (0.00276)	72.021 (1.143)
MNIST	Hq-Recycle	1.12915 (0.01593)	1.13056 (0.00262)	83.301 (0.700)
	Warm-PCG	1.12914 (0.01593)	1.13062 (0.00262)	83.303 (0.684)
	Zero-PCG	1.12988 (0.01593)	1.13015 (0.00269)	83.107 (0.628)
	AID-CG (warm)	1.13679 (0.01598)	1.13839 (0.00274)	83.211 (0.653)

Table 6: Classification confirmation validation losses at all four budgets, mean (sample SD), all ten seeds. The starred 0.5-second column uses independent lower responses; other columns use tracked classifiers. Values across these columns therefore have different evaluation states.

Method	0.2 s	0.5 s*	1 s	3 s
CIFAR-10				
Hq-Recycle	1.51205 (0.01014)	1.51203 (0.01013)	1.51203 (0.01013)	1.51203 (0.01013)
Warm-PCG	1.51208 (0.01013)	1.51203 (0.01014)	1.51203 (0.01014)	1.51203 (0.01014)
Zero-PCG	1.51376 (0.01017)	1.51276 (0.01016)	1.51256 (0.01014)	1.51247 (0.01014)
AID-CG (warm)	1.72766 (0.03665)	1.51704 (0.01021)	1.51608 (0.01025)	1.51284 (0.01010)
Fashion-MNIST				
Hq-Recycle	1.17941 (0.01535)	1.17930 (0.01534)	1.17929 (0.01534)	1.17930 (0.01534)
Warm-PCG	1.17944 (0.01534)	1.17930 (0.01534)	1.17930 (0.01534)	1.17930 (0.01534)
Zero-PCG	1.18061 (0.01553)	1.18016 (0.01531)	1.18004 (0.01532)	1.17998 (0.01534)
AID-CG (warm)	1.33618 (0.04413)	1.18865 (0.01556)	1.18670 (0.01557)	1.18298 (0.01519)
MNIST				
Hq-Recycle	1.12929 (0.01593)	1.12915 (0.01593)	1.12914 (0.01593)	1.12914 (0.01592)
Warm-PCG	1.12928 (0.01594)	1.12914 (0.01593)	1.12914 (0.01593)	1.12914 (0.01593)
Zero-PCG	1.13063 (0.01591)	1.12988 (0.01593)	1.12976 (0.01588)	1.12967 (0.01591)
AID-CG (warm)	1.32846 (0.03467)	1.13679 (0.01598)	1.13420 (0.01579)	1.13093 (0.01579)

D.4 Failure accounting and endpoint-recording correction

The final ridge collection consists of 210 trajectories: seven methods, ten seeds, and three repeats. All methods retain $30/30$ primary endpoints. At one second, warm-start AID-CG has $27/30$ endpoints; every other method has $30/30$. At three seconds, qNBO has $18/30$ and AID-CG has $3/30$; the remaining five methods retain all 30. We preserve every numerical failure and report it for the implementation and configuration that produced it.

The initial confirmation recorder wrote endpoints to disk only after a trajectory finished, so a later failure could erase earlier valid observations. We corrected endpoint persistence without changing optimizer code, configurations, timing boundaries, seeds, or selection. We repeated all 60 qNBO/AID-CG trajectories, including initially successful runs, and use this replacement collection in the final analysis. The original records and source version are retained. Recovery runs are not additional independent samples. The merged final collection contains 798 of 840 possible ridge budget endpoints, including all 210 primary endpoints.

Untimed fixed-step replays diagnose representative failures. AID-CG seed 26092604 reaches an upper-variable magnitude of roughly 2.6×10^{34} by step 102, and qNBO seed 26092605 reaches roughly 1.25×10^{21} by step 165. Double-precision reference residuals also fail in these states, so these cases are not explained solely by the float32 evaluation threshold. Confirmation seeds are not used to retune the configurations after observing these failures.

D.5 Work and time to fixed quality

The confirmation reuses both development-fixed thresholds from Appendix H. Classification targets use tracked-state validation loss; ridge targets use independent-response validation loss. Attainment counts include all attempted trajectories. Conditional time and work summaries use only seeds for which all three repeats attain the target; this denominator is reported. They are not estimates for nonattaining runs. In particular, ridge G/W/Z/qNBO/SHINE/HOAG each attain the 90% target in $27/30$ trajectories, spanning nine complete seeds.

Table 7: Ridge confirmation: observed target attainment and conditional work. Runs count attained timing trajectories out of 30; seeds count those with all three repeats attained out of 10. Time and probe means use only these complete seeds. A dash indicates that a comparable probe counter is not available for that implementation. No failed or unobserved target is assigned zero cost.

Method	Runs	Seeds	Time (s)	Probes	Probes/update
75% development-fixed target					
Hq-Recycle	30/30	10/10	0.2091	104.5	3.98
Warm-PCG	30/30	10/10	0.2182	107.9	3.97
Zero-PCG	30/30	10/10	0.1436	70.3	4.00
qNBO	30/30	10/10	0.5065	–	–
SHINE	30/30	10/10	0.2731	–	–
HOAG	30/30	10/10	0.2910	–	–
AID-CG (warm)	24/30	8/10	1.1892	595.8	10.98
90% development-fixed target					
Hq-Recycle	27/30	9/10	0.3809	171.1	3.35
Warm-PCG	27/30	9/10	0.3936	180.6	3.53
Zero-PCG	27/30	9/10	0.2357	115.6	4.00
qNBO	27/30	9/10	0.8417	–	–
SHINE	27/30	9/10	0.6227	–	–
HOAG	27/30	9/10	0.6968	–	–
AID-CG (warm)	6/30	2/10	1.7315	824.0	10.99

Before the 90% classification target, G/W/Z all use four probes per update. Their cumulative counts are $8/8/12$ on CIFAR-10, approximately 10.93 each on Fashion-MNIST, and approximately 10.40 each on MNIST. All three methods reach these targets within a few updates, at times close to the logging resolution, which limits the precision of time rankings. On ridge, G's mean probes per update decline from 3.99 in 0–0.2 seconds to 2.39 in 0.2–0.5 seconds and 1.28 in 1–3 seconds; W's corresponding means are 4.00, 2.82, and 1.35. These phase-specific counts distinguish the cost of reaching a quality target from the average cost over a complete trajectory.

E Real-system replay and the scaling diagnostic

We collect 100 consecutive states from each G/W trajectory for two tasks (ridge and CIFAR-10) and three development seeds $(601, 602, 1302)$: 1200 states in total. Forty anchors per trajectory comprise steps 1–10, even steps 12–50, and steps $55, 60, \ldots, 100$. Counterfactual replay fixes the current upper and lower states, right-hand side, preconditioner, and previous approximate adjoint. It compares Z/W/G/R independently at the same anchor. Sequential replay traverses all 100 states with each solver carrying its own approximate adjoint history. Every solve has $P = 4$ including refresh and $\tau = 0.05$. We evaluate exact float64 HVPs (exact64) and production float32 finite differences (FD32) separately, for a total of $13{,}440$ solves and $29{,}062$ recorded initialization/correction events.

Evaluation reference and coordinates. High-accuracy adjoints solve the current exact system at the tracked lower state, not at a newly optimized lower solution. The stored float32 right-hand side is

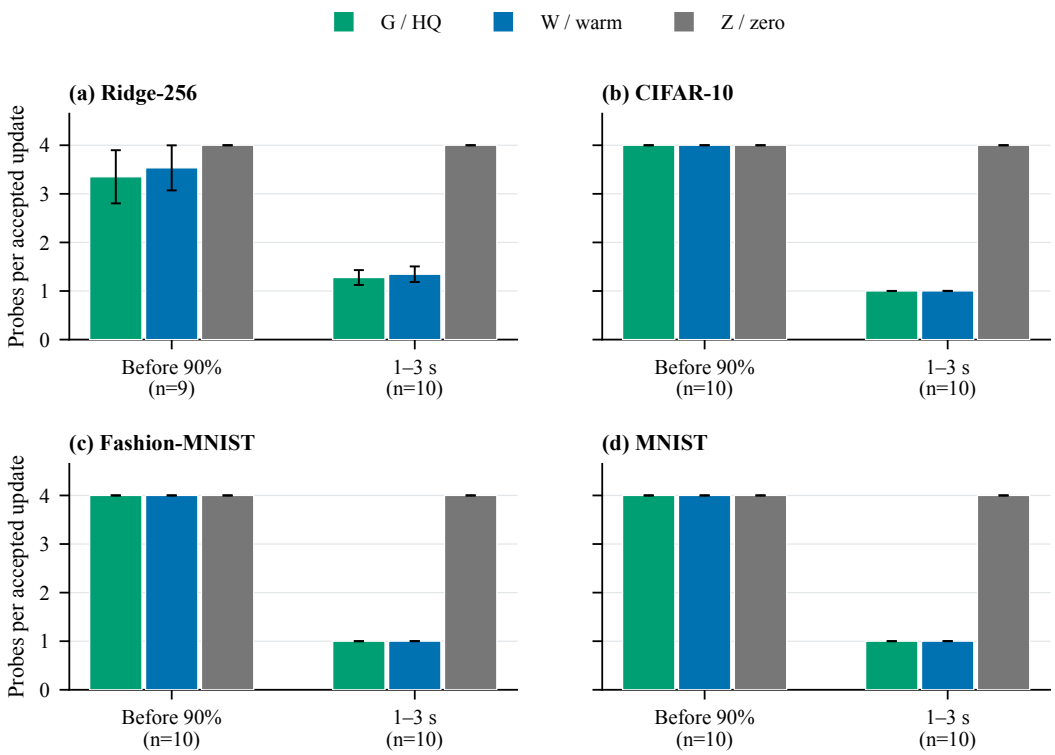

Figure 3: Mean probes per accepted update before reaching the development-fixed 90% validation target and during 1–3 seconds. Error bars are sample SDs across seed-level means. Target-conditioned values use ten complete seeds for classification and nine for ridge; later values use all ten. Counts include refresh. Classification targets evaluate tracked classifiers; ridge targets evaluate independent lower responses. W and Z are HQ variants without rescaling and without adjoint reuse.

promoted to float64 and held fixed across solvers. Reference adjoints only evaluate error; they never initialize a tested solver or affect its stopping. The largest reference relative residual is 9.99×10^{-12}. We record true residual, recursive residual, relative energy error, relative local hypergradient error, direction cosine, probe count, and false tolerance acceptance. No reference hypergradient norm is below 10^{-12}.

Ridge uses the original grouped-regularizer coordinates. CIFAR-10 uses the sample-weight hypergradient $-\widetilde{C}q$, matching Eq. (7); no additional sigmoid derivative is applied. Thus the classification diagnostic is not the Euclidean logit-gradient norm. All initializations within a task share the same mixed action. For summaries, we average anchors within each task/seed/source before averaging the two source trajectories and then seeds. Table 8 pools the early and middle anchors, with 30 anchors per source and seed.

Table 8: Paired production-FD replay on early/middle real-system anchors. Errors are relative local hypergradient errors after the budgeted solve; reduction is $(\mathrm{W}-\mathrm{G})/\mathrm{W}$. The last column is G minus W direction cosine. Each row averages both trajectory sources within one development seed.

Task	Seed	G error	W error	Reduction	Δ cosine
Ridge-256	601	0.25494	0.27823	8.37%	+0.00488
Ridge-256	602	0.26110	0.28298	7.73%	+0.01029
Ridge-256	1302	0.29195	0.33220	12.11%	+0.01030
CIFAR-10	601	0.01910	0.01925	0.77%	+0.00002
CIFAR-10	602	0.02407	0.02319	-3.78%	-0.00006
CIFAR-10	1302	0.01932	0.02027	4.67%	+0.00004

E.1 Energy rescaling versus residual rescaling

Let $v = u/\|u\|$ and $a = \mathcal{D}(v)$. The diagnostic candidate R replaces the Galerkin coefficient with

$$\alpha_R = \frac{a^\top b}{a^\top a}, \qquad q_R^{(0)} = \alpha_R v. \tag{20}$$

For a finite positive denominator, this is the standard scalar least-squares minimizer of $\|b - \alpha a\|^2$. Its *estimated* initialization residual cannot exceed that of G, W, or zero, using the same refreshed a. This introduces no additional probe. However, with exact actions, local hypergradient error satisfies

$$h(q) - h(q^\dagger) = CA^{-1}(b - Aq),$$

so minimizing an unweighted residual need not minimize hypergradient error. With no direct upper derivative, positive rescaling alone also preserves the initial hypergradient direction; correction and stopping determine subsequent directional changes.

We froze the retention rule before replay: on FD32 early/middle counterfactual states, R must reduce mean relative hypergradient error by at least 1% or raise true-tolerance attainment by at least five percentage points relative to G; no task may incur more than 5% error degradation, and exact64 must not contradict the benefit. R increases error by 6.91% on ridge and 12.87% on CIFAR-10, with only 0.28 percentage points greater overall true-tolerance attainment. The exact64 results support the same decision. R therefore remains a diagnostic comparison; we retain G for the full optimization experiments.

Across 468 valid exact64 counterfactual initializations, G has no greater energy error than W within 10^{-10} numerical tolerance. Across 936 valid initializations from both backends, R has no greater estimated residual than G/W/zero. The 348 early/middle G/W pairs with valid initialization cosines differ by at most 3.03×10^{-8}. These checks confirm the energy, residual, and direction properties at initialization. The replay results separately measure performance after correction. No false tolerance acceptance occurs in the full replay collection.

F Controlled amplitude and direction drift

We construct exact 128-dimensional SPD systems in float64, using five seeds 701–705 and condition numbers $\kappa \in \{10, 100, 1000\}$. A random orthogonal matrix Q and a geometrically spaced spectrum from 1 to κ define A. For orthonormal random vectors e, z, prescribe

$$q_t^\dagger = \exp(a\sin(0.35t))\left[\cos(\theta\sin(0.31t))e + \sin(\theta\sin(0.31t))z\right], \qquad b_t = A_t q_t^\dagger,$$

for $t = 0, \ldots, 49$. The static, amplitude, and direction cases use $(a, \theta) = (0, 0), (0.6, 0), (0, 0.6)$ with fixed A. The mixed case uses $(0.6, 0.6)$, rotates the first two columns of Q by $0.5\sin(0.23t)$, and multiplies all eigenvalues by $\exp(0.2\sin(0.29t))$. The condition number stays fixed. All methods use the identity preconditioner, carry their own approximate history, and share $P = 4, \tau = 0.05$; exact truth only evaluates error. The first solve starts from zero.

An additional 6×6 grid uses $a \in \{0, 0.03, 0.1, 0.3, 0.8, 1.5\}$ and $\theta \in \{0, 0.01, 0.03, 0.1, 0.3, 0.8\}$ at $\kappa = 100$, with five seeds and 25 states each. Together the two experiments contain 7500 systems and $30{,}000$ solves. They isolate how prescribed system changes affect local error and probe cost.

Table 9: Amplitude-drift sequences. G/W probe means condition on both methods meeting the true relative residual tolerance 0.05; the joint-state count and each method's unconditional attainment are also reported. Five seeds supply 250 states per condition number. Conditional probe counts average within seed before averaging across seeds.

κ	G probes	W probes	Joint attainment	G attainment	W attainment
10	1.060	2.536	250/250	100.0%	100.0%
100	1.016	2.744	242/250	98.0%	96.8%
1000	1.017	2.695	238/250	97.2%	95.2%

Under amplitude drift, G reaches this tolerance with fewer probes (Table 9). W continues correcting and attains smaller average energy error. Direction drift largely removes the probe saving; G has

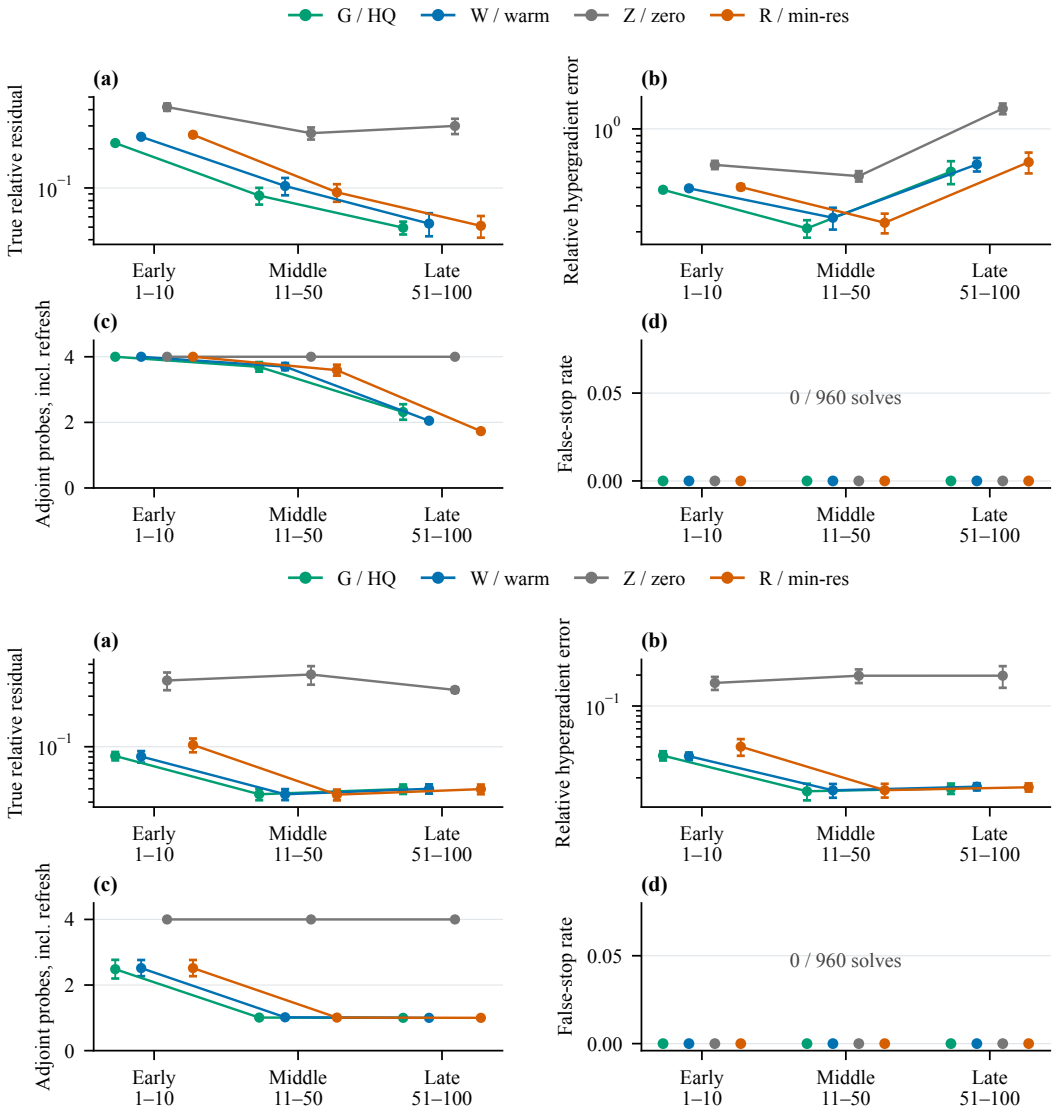

Figure 4: Exact64 replication of the real-system counterfactual replay: ridge (top) and CIFAR-10 (bottom). Each method has at most four probes including refresh. Reference residuals and local hypergradients are evaluated at the tracked state. Panels report final true residual, relative hypergradient error, probes, and false-stop frequency; means and SDs aggregate within each seed first. Within each seed, the two trajectory sources are averaged before aggregating across the three seeds.

lower mean energy error in $16/36$ parameter-grid cells. The benefit therefore depends on the type of drift and the chosen accuracy criterion.

We distinguish the prescribed sinusoidal parameters from observed consecutive-solution drift. The latter is

$$d_t^{\mathrm{amp}} = \left| \log \frac{\left\| q_t^\dagger \right\|}{\left\| q_{t-1}^\dagger \right\|} \right|, \qquad d_t^{\mathrm{angle}} = \arccos \frac{(q_t^\dagger)^\top q_{t-1}^\dagger}{\left\| q_t^\dagger \right\| \left\| q_{t-1}^\dagger \right\|}.$$

The real-system collection supplies 1188 adjacent reference pairs. The grid's observed drift reaches approximately 0.516 in absolute log amplitude and 0.246 radians in angle; 25 real ridge points exceed its angular coverage. Figure 6 includes those points and leaves empty synthetic bins blank.

The comparison establishes coverage in these two drift coordinates; real and synthetic systems have different operators and preconditioners.

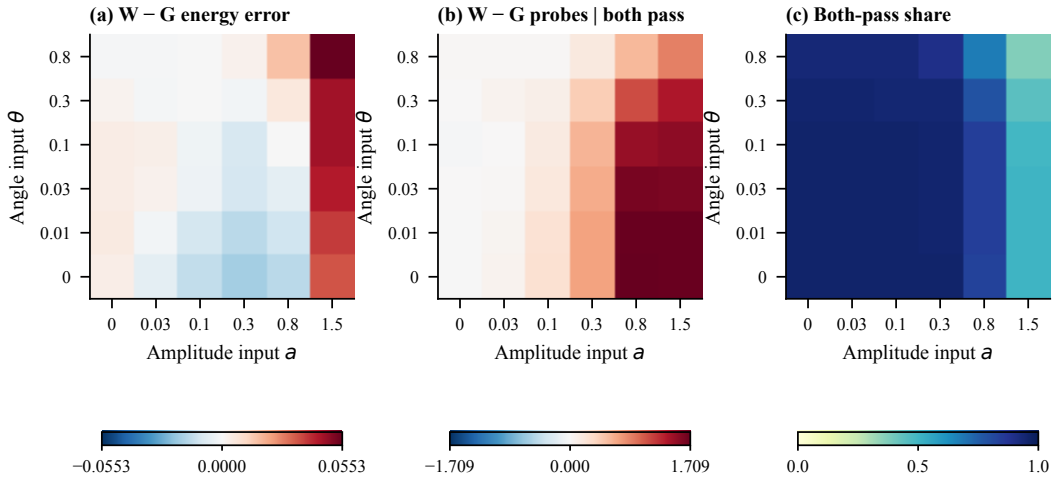

Figure 5: Prescribed amplitude–direction parameter grid at $\kappa = 100$. Panels show W minus G relative energy error, W minus G probe count conditioned on both attaining the true tolerance 0.05, and joint attainment share. Positive differences favor G; the joint share ranges from 36.8% to 96.0%. The parameters control bounded sinusoidal sequences; they are not consecutive-state drift measurements.

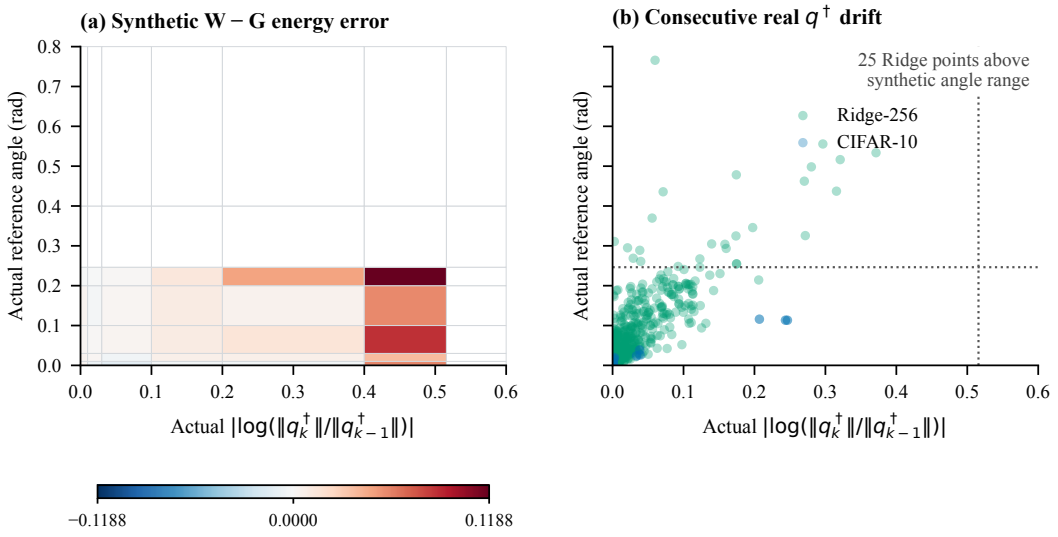

Figure 6: Observed consecutive-reference drift. Left: synthetic W minus G relative energy error, averaged within seed and then across seeds; positive values favor G, and empty bins remain blank. Right: 1188 real-system pairs with dotted lines at the synthetic coordinate maxima; all 25 out-of-range ridge points remain visible. Matching coordinates does not imply matching operators or preconditioners.

G Five-seed implementation study: settings and accounting

G.1 Adopted settings for the five-seed implementation study

Table 10 specifies the revised main configuration. All tasks use a memory size of ten, an adjoint probe cap of four, a residual tolerance of 0.05, at least one lower step, and no periodic reset. Classifi-

cation uses the paired finite-difference backend and mirror update; ridge uses generic differentiation and a Euclidean outer update. These settings were adopted from earlier exploration and fixed before running seeds 1301–1305. The methods share evaluation budgets, while their development histories and tuning effort differ.

Table 10: Revised Hq-Recycle settings, fixed before the five-seed implementation study. Lower and outer step columns give nominal learning rates. Ridge uses a fixed single lower step.

Task	$T_{\max}$	Lower step	Outer step	ϵ_{low}	Outer rule
CIFAR-10	5	1	10000	10^{-4}	Mirror
Fashion-MNIST	3	1	10000	10^{-4}	Mirror
MNIST	5	1	10000	10^{-4}	Mirror
Ridge-256	1	1	900	–	Euclidean

G.2 Initial development grids

Initial development seeds are 601 and 602, and initial evaluation seeds are 610–614. Each method–task pair has nine configurations, each receiving one second on each development seed. The outer learning-rate multiplier is one of $\{0.3, 1, 3\}$ relative to a previously selected center. Hq-Recycle (initial) crosses these rates with $T \in \{1, 3, 5\}$; $m = 10$, $P = 4$, and $\tau = 0.05$ are fixed. Other methods cross rates with the three recipes below, except qNBO, which crosses rates with its three variants using each variant's previously selected recipe. Full grid centers and chosen settings are tabulated in Tables 11 and 12.

- SHINE: inner-step counts of 5, 10, or 20, with a memory size of ten. We retain the local implementation of outer-problem awareness (OPA) and its compact inverse kernel.
- AID-CG (cold): $(T, \text{lower step}, \text{CG cap}) = (5, 1, 5), (20, 1, 10), (20, 5, 20)$.
- HOAG: $(T, \text{lower step}, \text{CG cap}) = (5, 0.1, 5), (10, 1, 10), (20, 1, 20)$, with CG tolerance 10^{-6}.
- BOME: $(T, \text{lower step}) = (5, 1), (20, 1), (20, 5)$; $\eta = 0.1$, value-based constraint mode, and the joint lower-variable outer step size equals the lower step size.
- PZOBO: the tuples $(T, \text{lower step}, \text{directions}, \text{smoothing})$ are $(5, 1, 1, 0.1)$, $(20, 1, 5, 0.1)$, and $(20, 5, 5, 1)$.
- qNBO recipes: $(T, \gamma, h_0) = (5, 0.1, 1), (5, 1, 1), (10, 1, 10)$. Shared settings are $P = 1$, $\beta = 0.1$, a memory size of ten, and $\xi = 1$. Variants use Q1-BFGS, Q4-BFGS, or Q4-SR1.

The search budget is nine configurations per method–task pair. qNBO uses the validated local direction-pair convention (`paper_literal_direction_pair=False`). SHINE uses the compact kernel, and AID-CG (cold)/HOAG cache the fixed Hessian differentiation graph within each solve. The artifact includes the evaluated source implementations and configurations for all methods.

G.3 Timing and failure accounting

Two warmup steps use a separate optimizer and state, both discarded before the timed trajectory. CUDA synchronization brackets each full optimizer step; cumulative measured step time determines endpoints. The last complete state within each budget is retained. Refresh, compact construction/application, lower backtracking, and mixed differentiation occur inside the timed step. Data preparation, evaluation, state transfer, and warmup do not. Method order is permuted with a fixed per-seed rule. Runs are sequential, with OpenMP and Intel Math Kernel Library (MKL) thread counts set to one; fixed GPU clocks are not assumed. The environment uses PyTorch 2.8.0+cu128 and an RTX 4090 D (24 GB).

The expanded main stage attempts 195 trajectories: seven principal methods and two controls on four tasks, plus an exact-HVP extension on three classification tasks, each with five seeds. It completes 189 trajectories and saves 756 budget states. The ablation stage adds 140 successful trajectories and 420 budget states. Main one-second states provide matched ablation controls. AID-CG (cold) fails on ridge seeds 1301, 1302, and 1304; PZOBO fails on 1302, 1304, and 1305. A

Table 11: Outer learning-rate search centers. Each center is multiplied by 0.3, 1, or 3. The three qNBO rows have their own inherited centers and recipes. C/F/M/R abbreviate CIFAR-10, Fashion-MNIST, MNIST, and ridge.

Method / variant	C	F	M	R
Initial implementation	100000	100000	100000	100
qNBO Q1-BFGS	100000, r1	100000, r2	100000, r2	100, r0
qNBO Q4-BFGS	10000, r1	10000, r1	10000, r1	100, r1
qNBO Q4-SR1	100000, r2	10000, r1	100000, r2	1, r1
SHINE	100000	100000	100000	100
PZOBO	10000	1000	1000	100
BOME	10000	10000	100000	100
HOAG	100000	100000	100000	100
AID-CG (cold)	100000	100000	100000	100

Table 12: Initially selected settings; baseline rows are reused in the five-seed implementation study. Each cell gives the outer learning rate and recipe index (0-based); Hq-Recycle (initial) cells give the outer learning rate and lower-step count T. qNBO variants are stated separately in the text.

Method	C	F	M	R
Initial implementation	100000, $T = 5$	100000, $T = 3$	100000, $T = 5$	300, $T = 3$
qNBO	100000, r2	100000, r2	100000, r2	300, r1
SHINE	100000, r0	100000, r0	100000, r0	300, r0
PZOBO	30000, r2	300, r2	300, r2	100, r0
BOME	10000, r2	30000, r2	30000, r2	100, r1
HOAG	100000, r1	100000, r1	100000, r1	300, r0
AID-CG (cold)	100000, r2	100000, r2	100000, r2	300, r1

failed trajectory contributes no endpoint summary, even if an earlier budget was reached internally. Curvature-check failures do not imply that the mathematical ridge objective is nonconvex. All failures remain in the records, and target summaries account for every attempt.

The experiment artifact preserves raw curves, failures, all $1{,}176$ formal budget states, configurations, source versions, and reproduction instructions. Main and ablation stages have identical paired data metadata and executed-source hashes. The declared budgets sum to 725 solver seconds across attempts; warmup, failed work, overbudget steps, and independent evaluation are additional computation.

G.4 Statistics

Curve bands and table parentheses are sample SDs, not confidence intervals. Paired comparisons use the same task seed. A throughput ratio divides accepted steps per measured solver second, then averages those per-seed ratios; it is not a ratio of mean iteration counts or a time-to-quality estimate. Classification seeds vary the split and corruption but share official test examples. Budget endpoints on one trajectory are correlated. We make no general significance or equivalence claim from five seeds.

H Development-fixed targets

Let L_0 be the mean initial development validation loss and L_1 the best mean one-second development loss among selected configurations. Targets are $L_0 - \rho(L_0 - L_1)$ for $\rho \in \{0.75, 0.90\}$. The original thresholds are reused unchanged in the five-seed implementation study; rounded values are shown below:

Task	75% target	90% target
CIFAR-10	1.689746	1.567179
Fashion-MNIST	1.444893	1.273355
MNIST	1.402684	1.222703
Ridge	0.053877	0.049552

Hitting time is the first logged accepted state at or below the threshold, without interpolation below the approximately 0.05-second sampling resolution. Nonattainment and numerical failure receive the three-second cap in capped means. For a fixed absolute target, attainment also depends on variation in the achievable losses across data seeds. Classification targets are often reached near the sampling resolution; small time differences there are not precise speedup estimates. Figure 7 shows the ridge attainment times.

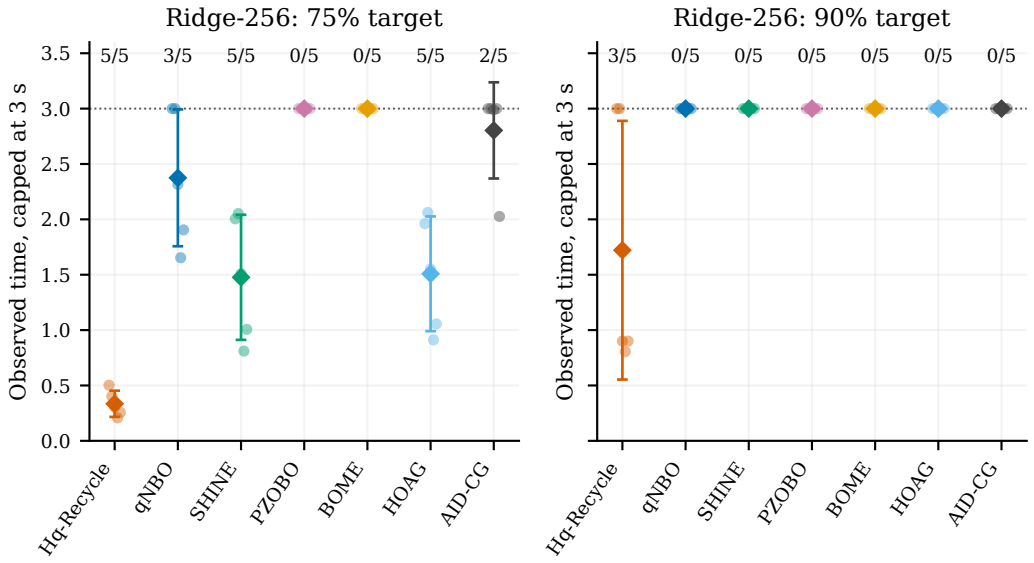

Figure 7: Ridge time to development-fixed validation targets. Dots show all five observed or capped times; diamonds and whiskers show mean and sample SD. Counts indicate target attainment. Nonattainment and numerical failure receive the three-second cap.

I Direct implementation comparison and matched controls

The initial-implementation control reruns the original solver and selected parameters on implementation-study seeds under the common timing protocol. Figure 8 shows the paired throughput gains. Table 13 also reports warm-PCG, which shares the revised lower tracking, geometry, finite-difference backend, deferred preconditioning, probe cap, and residual tolerance. Warm-PCG tests the retained adjoint with unit scaling. Its throughput ratios near one and small loss differences indicate comparable performance to Galerkin rescaling on these tasks.

Table 13: Three-second matched comparisons. Throughput ratios are mean (sample SD) over five pairs. The last column is the mean raw test-loss difference, warm-PCG minus Hq-Recycle; a positive value favors Hq-Recycle.

Task	HQ / initial	HQ / warm-PCG	$L_{\mathrm{warm}} - L_{\mathrm{HQ}}$
CIFAR-10	3.281 (0.126)	0.930 (0.099)	1.69×10^{-5}
Fashion-MNIST	2.333 (0.079)	1.014 (0.025)	-1.35×10^{-4}
MNIST	3.280 (0.082)	1.048 (0.042)	1.39×10^{-4}
Ridge-256	1.545 (0.065)	0.968 (0.057)	5.25×10^{-5}

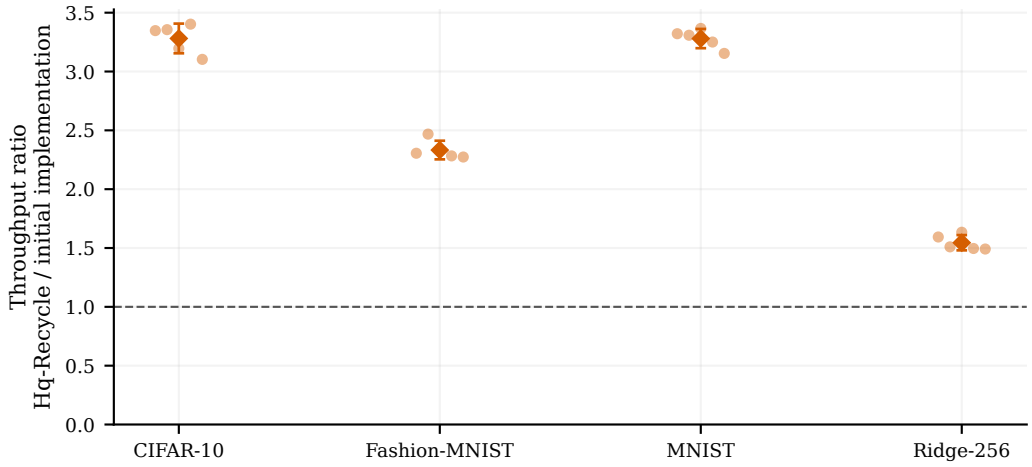

Figure 8: Iteration throughput of optimized Hq-Recycle relative to the initial implementation at three seconds. Dots show five paired ratios; diamonds show mean and sample SD. Ratios measure the combined implementation and configuration changes.

Exact-operator extension. Table 14 reports results for an exact softmax Hessian action with otherwise matched settings. Exact/FD throughput ratios are computed per seed before averaging. This structured-operator extension is reported separately from the main finite-difference comparisons.

Table 14: Classification with exact Hessian–vector products at three seconds. FD denotes finite differences. Entries are mean (sample SD) across five seeds; losses are tracked-state test cross-entropy.

Task	Exact / FD throughput	FD test loss	Exact test loss
CIFAR-10	1.030 (0.046)	1.52110 (0.00270)	1.52104 (0.00270)
Fashion-MNIST	1.026 (0.052)	1.19350 (0.00128)	1.19338 (0.00124)
MNIST	1.009 (0.058)	1.13067 (0.00188)	1.13066 (0.00182)

J Component and configuration ablations

The results in this section use only seeds 1301–1305. All AID-CG labels in the archived figures denote the cold-start adapter. These experiments hold ablated configurations fixed and do not perform the independent rate selection used for Zero-PCG in the ten-seed confirmation.

Table 15: Five-seed implementation study at one second: means of per-trajectory probe counts, including refresh. The last row is the mean paired throughput ratio, full method divided by no recycling. Counts average the entire trajectory, including iterations after the quality targets.

Variant / statistic	CIFAR-10	Fashion	MNIST	Ridge
Hq-Recycle	1.07	1.09	1.07	2.38
No recycling	4.00	4.00	4.00	4.00
No residual stopping	4.00	4.00	4.00	4.00
Identity adjoint precond.	1.06	1.09	1.06	2.93
Two-loop kernel	1.09	1.10	1.08	2.42
Throughput (full / no recycling)	2.12	1.81	1.95	1.27

All ablations retain the revised settings except for the specified intervention. No recycling starts the adjoint at zero with the same four-probe budget. No residual stopping sets $\tau = 0$ while retaining numerical guards. Identity adjoint preconditioning preserves lower L-BFGS. Two-loop recursion

replaces compact application in both lower and adjoint computation. Eager preconditioning computes an initial inverse action before testing reuse. No lower stopping uses the full classification lower-step cap.

The original-outer classification control restores the old geometry, outer step size, and generic AD backend together. The original-lower ridge control restores three lower steps with a lower step size of 0.1; the old-outer ridge control changes only its outer step size from 900 to 300. The first two are configuration restorations rather than single-factor ablations. No variant is retuned. Table 15 and Figure 9 show probe use and paired loss differences; Tables 16–19 give the numerical comparisons.

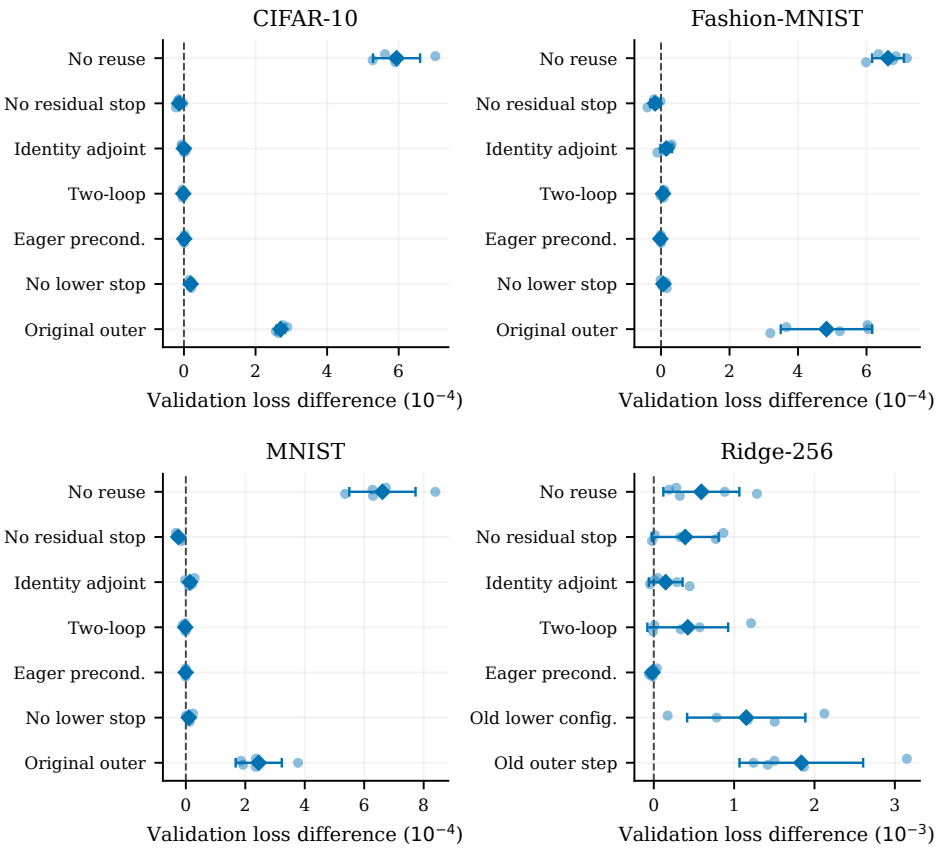

Figure 9: One-second validation loss, ablated minus full Hq-Recycle. Positive differences favor the full method. Dots show every paired seed difference; diamonds show mean and sample SD. Task-specific axis units are explicit and negative effects are retained. Configuration restorations change multiple settings at once, as specified above.

Table 16: CIFAR-10: one-second component comparisons. Throughput is full/ablated, mean (sample SD); ΔL is mean validation loss, ablated minus full. Positive differences favor the full method. The last column counts pairs in which the full method has lower validation loss.

Ablation	Throughput	ΔL	Full wins
Eager preconditioning	1.081 (0.121)	7.15×10^{-7}	4/5
Identity adjoint precond.	1.056 (0.140)	-2.15×10^{-7}	2/5
No residual stopping	1.656 (0.205)	-1.34×10^{-5}	0/5
Two-loop kernel	1.181 (0.190)	-1.91×10^{-6}	1/5
No lower stopping	2.589 (0.325)	1.86×10^{-5}	5/5
No recycling	2.120 (0.301)	5.94×10^{-4}	5/5
Original outer configuration	1.544 (0.158)	2.70×10^{-4}	5/5

Table 17: Fashion-MNIST: one-second component comparisons. Throughput is full/ablated, mean (sample SD); ΔL is mean validation loss, ablated minus full. Positive differences favor the full method. The last column counts pairs in which the full method has lower validation loss.

Ablation	Throughput	ΔL	Full wins
Eager preconditioning	1.066 (0.093)	-1.81×10^{-6}	1/5
Identity adjoint precond.	0.985 (0.076)	1.53×10^{-5}	4/5
No residual stopping	1.575 (0.150)	-1.75×10^{-5}	0/5
Two-loop kernel	1.113 (0.096)	5.79×10^{-6}	4/5
No lower stopping	1.798 (0.198)	6.94×10^{-6}	4/5
No recycling	1.815 (0.228)	6.63×10^{-4}	5/5
Original outer configuration	1.493 (0.126)	4.83×10^{-4}	5/5

Table 18: MNIST: one-second component comparisons. Throughput is full/ablated, mean (sample SD); ΔL is mean validation loss, ablated minus full. Positive differences favor the full method. The last column counts pairs in which the full method has lower validation loss.

Ablation	Throughput	ΔL	Full wins
Eager preconditioning	1.082 (0.059)	-5.48×10^{-7}	0/5
Identity adjoint precond.	1.017 (0.030)	1.38×10^{-5}	4/5
No residual stopping	1.664 (0.227)	-2.52×10^{-5}	0/5
Two-loop kernel	1.183 (0.081)	-2.00×10^{-6}	1/5
No lower stopping	2.616 (0.076)	1.02×10^{-5}	5/5
No recycling	1.953 (0.068)	6.61×10^{-4}	5/5
Original outer configuration	1.638 (0.108)	2.45×10^{-4}	5/5

Table 19: Ridge-256: one-second component comparisons. Throughput is full/ablated, mean (sample SD); ΔL is mean validation loss, ablated minus full. Positive differences favor the full method. The last column counts pairs in which the full method has lower validation loss.

Ablation	Throughput	ΔL	Full wins
Eager preconditioning	0.981 (0.057)	-1.66×10^{-5}	1/5
Identity adjoint precond.	1.006 (0.100)	1.48×10^{-4}	4/5
No residual stopping	1.197 (0.084)	3.91×10^{-4}	4/5
Two-loop kernel	1.173 (0.043)	4.23×10^{-4}	4/5
No recycling	1.275 (0.094)	5.92×10^{-4}	5/5
Original lower configuration	1.446 (0.104)	1.15×10^{-3}	5/5
Original outer step	0.908 (0.031)	1.84×10^{-3}	5/5

K UNIFORM LOWER-RESPONSE EVALUATION

For each saved classification state, we keep x fixed and initialize from its saved y. Float32-preprocessed inputs are converted to float64, preserving the objective's data. L-BFGS uses a learning rate of one, at most 150 iterations, a history size of 20, a gradient tolerance of 10^{-10}, a change tolerance of 10^{-16}, and a strong-Wolfe line search. The common acceptance threshold is a lower-gradient norm of 10^{-6}.

The 300 main states cover three classification tasks, ten configurations per task, five seeds, and two budgets (0.5 and 3 seconds). Their maximum residual is 1.6865×10^{-8}. The 105 one-second ablation states reach at most 1.785×10^{-8}, and 15 separately evaluated one-second Hq-Recycle controls reach at most 1.591×10^{-8}. The same procedure and evaluator are used throughout, and each budget is analyzed separately.

Tracked classifiers measure anytime prediction quality, while independently optimized lower responses evaluate the learned hyperparameters through the bilevel objective. Table 20 reports this additional, untimed evaluation of hyperparameters learned within three seconds; Figure 10 compares cross-entropy before and after the additional solve. Classification no-recycling loss differences

continue to favor the full method in all 15 pairs. The small improvements from disabling residual stopping also persist.

Table 20: Three-second learned hyperparameters evaluated after a common additional lower solve. Test accuracy (%), mean (sample SD), five seeds. This includes untimed evaluation optimization.

Method	CIFAR-10	Fashion-MNIST	MNIST
Hq-Recycle	74.55 (0.19)	71.78 (0.65)	83.54 (0.41)
qNBO	74.18 (0.47)	71.31 (0.95)	82.42 (0.62)
SHINE	74.14 (0.38)	71.68 (1.04)	82.88 (0.53)
PZOBO	70.83 (1.30)	68.63 (1.41)	80.06 (1.35)
BOME	73.56 (0.36)	71.44 (0.97)	82.40 (0.58)
HOAG	74.62 (0.18)	71.72 (0.74)	83.54 (0.41)
AID-CG (cold)	74.55 (0.20)	71.87 (0.94)	83.36 (0.43)

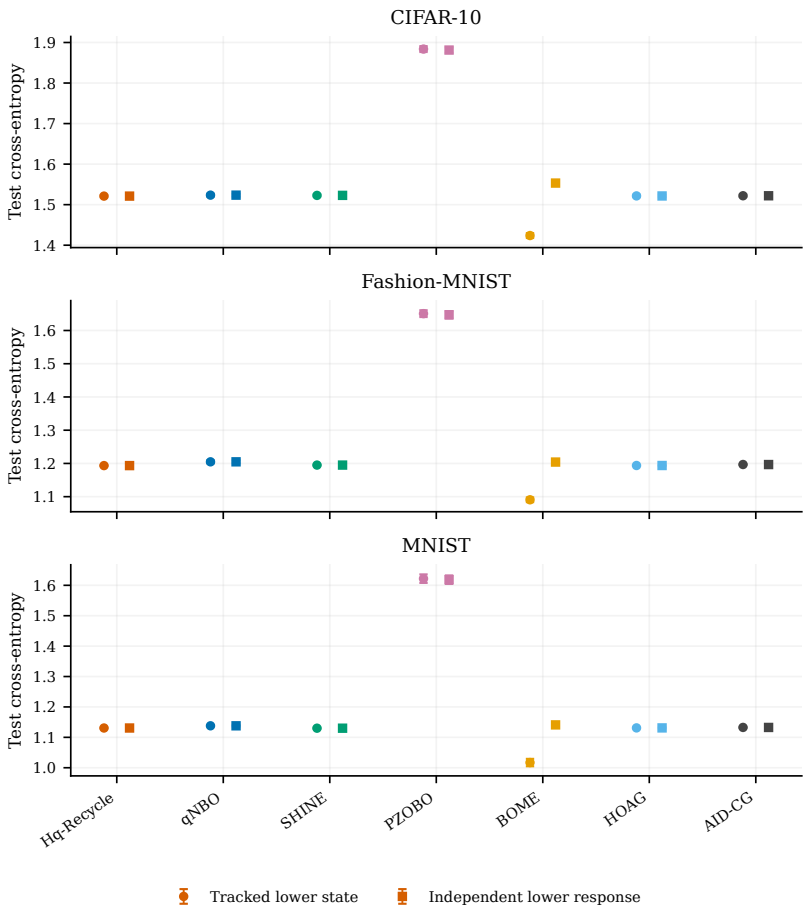

Figure 10: Test cross-entropy of tracked classifiers and independently optimized lower responses for three-second hyperparameters, mean and sample SD. Hq-Recycle changes little after lower optimization. BOME's tracked-state cross-entropy advantage does not persist at the lower response. The additional solve is excluded from solver time.

L Initial evaluation supporting the abstract

The initial evaluation uses the same tasks and objectives with seeds 610–614 and supplies the numerical results in the abstract.

On CIFAR-10, Fashion-MNIST, MNIST, and ridge, respectively, lower-step counts are 5, 3, 5, and 3; outer learning rates are 10^5, 10^5, 10^5, and 300; and lower learning rates are 1, 1, 1, and 0.1. This configuration uses fixed lower-step counts, the generic finite-difference/AD backend, Euclidean logit updates for classification, and eager adjoint preconditioning. The five-seed implementation study reruns it as the initial-implementation control.

The initial validation search completes $480/504$ attempts. The selected configurations complete $135/140$ main trajectories, with 405 endpoints at 0.5, 1, and 3 seconds. All 80 ablation trajectories succeed. Ridge AID-CG (cold) fails on seeds 610, 612, 613, and 614, and PZOBO on 612.

Both Hq-Recycle and qNBO reach the 75% ridge target in all five initial runs (Table 21). Their mean first-observed-time ratio is 2.4427, rounded to 2.44; the mean 0.5-second test-loss reduction is 20.2053%, rounded to 20.2%. Table 22 reports the initial classification probe counts.

Table 21: Initial ridge evaluation, seeds 610–614. Target time and test loss are mean (sample SD); both methods attain the target in all five runs.

Method	Target time (s)	Attainment	Test loss at 0.5 s
Initial implementation	0.465 (0.251)	5/5	0.05893 (0.00397)
qNBO	1.136 (0.524)	5/5	0.07386 (0.00308)

Table 22: Initial classification ablations: mean probes per outer iteration over three-second trajectories, averaged across five seeds. Expanded one-second counts appear in Table 15.

Variant	CIFAR-10	Fashion-MNIST	MNIST
Initial implementation	1.075	1.070	1.068
No recycling	4.000	4.000	4.000

M Complete five-seed implementation-study results

All results in this section use seeds 1301–1305 and the originally selected baseline settings. They are distinct from the expanded-rate, ten-seed confirmation. In every archived figure below, AID-CG denotes the cold-start implementation. The previously reported 27.8% ridge test-loss reduction at 0.5 seconds uses this cohort and rate search; the new confirmation reports 10.98% after expanding the rate grid.

Table 23: Five-seed implementation study: three-second test performance. Classification uses tracked-state accuracy (%, ↑); ridge uses independent-response half MSE (↓). Mean (sample SD); superscripts mark only two completed runs. These conditional results preserve the earlier evaluation and are not pooled with confirmation.

Method	CIFAR-10	Fashion	MNIST	Ridge-256
Hq-Recycle	74.55 (0.19)	71.78 (0.65)	83.54 (0.41)	**0.0489** (0.0023)
qNBO	74.18 (0.48)	71.31 (0.95)	82.43 (0.61)	0.0526 (0.0021)
SHINE	74.12 (0.38)	71.67 (1.03)	82.89 (0.52)	0.0509 (0.0026)
PZOBO	70.91 (1.10)	68.88 (1.46)	79.97 (1.45)	0.0570 (0.0037)$^{n=2}$
BOME	73.29 (1.25)	71.61 (1.49)	**83.71** (0.89)	0.0783 (0.0038)
HOAG	**74.62** (0.18)	71.73 (0.75)	83.55 (0.42)	0.0505 (0.0024)
AID-CG (cold)	74.55 (0.21)	**71.85** (0.94)	83.35 (0.45)	0.0553 (0.0051)$^{n=2}$

Tables 24–27 report all four predeclared budgets for ridge and the three classification tasks. Columns show each method's progression over time. Figure 11 shows validation trajectories, and Figure 12 compares test endpoints. Means with two successful runs are marked explicitly.

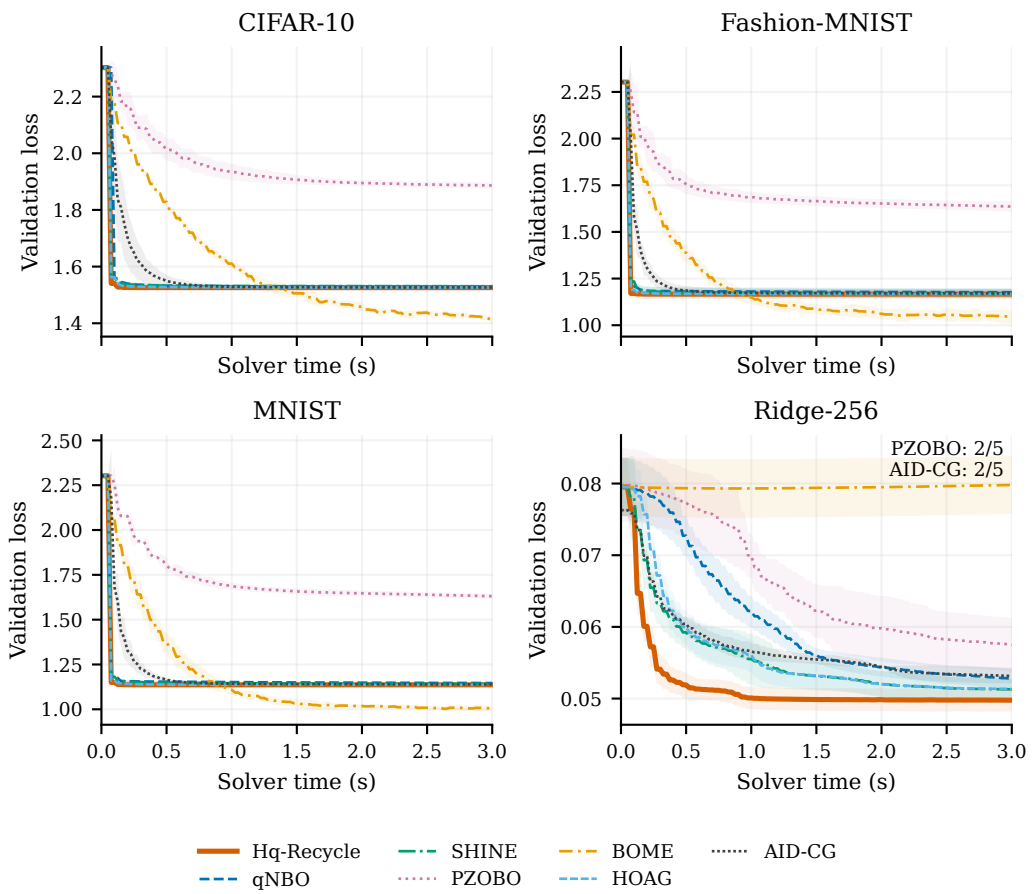

Figure 11: Five-seed implementation study: validation curves with mean and sample SD. Classification evaluates tracked states; ridge evaluates independent lower responses. Ridge PZOBO and AID-CG (cold) each have only $2/5$ complete trajectories; all others have five. AID-CG in this archived figure denotes the cold-start adapter.

Table 24: Ridge-256: all implementation-study budgets. Mean (sample SD), five successful runs unless marked. Losses are MSE$/2$ at an independently solved lower response.

Method	Metric	0.2 s	0.5 s	1 s	3 s
Hq-Recycle	Val. loss	0.0577 (0.0026)	0.0515 (0.0017)	0.0500 (0.0016)	0.0498 (0.0016)
	Test loss	0.0573 (0.0031)	0.0509 (0.0023)	0.0495 (0.0023)	0.0489 (0.0023)
qNBO	Val. loss	0.0784 (0.0041)	0.0718 (0.0026)	0.0617 (0.0021)	0.0528 (0.0015)
	Test loss	0.0768 (0.0035)	0.0705 (0.0041)	0.0609 (0.0018)	0.0526 (0.0021)
SHINE	Val. loss	0.0661 (0.0035)	0.0588 (0.0022)	0.0553 (0.0036)	0.0513 (0.0015)
	Test loss	0.0644 (0.0036)	0.0573 (0.0032)	0.0544 (0.0034)	0.0509 (0.0026)
PZOBO$^{n=2}$	Val. loss	0.0790 (0.0049)	0.0769 (0.0049)	0.0686 (0.0035)	0.0575 (0.0039)
	Test loss	0.0796 (0.0053)	0.0774 (0.0056)	0.0698 (0.0059)	0.0570 (0.0037)
BOME	Val. loss	0.0794 (0.0041)	0.0793 (0.0041)	0.0793 (0.0041)	0.0798 (0.0040)
	Test loss	0.0778 (0.0037)	0.0777 (0.0037)	0.0777 (0.0036)	0.0783 (0.0038)
HOAG	Val. loss	0.0718 (0.0015)	0.0594 (0.0024)	0.0555 (0.0023)	0.0513 (0.0014)
	Test loss	0.0712 (0.0050)	0.0585 (0.0028)	0.0543 (0.0033)	0.0505 (0.0024)
AID-CG (cold)$^{n=2}$	Val. loss	0.0676 (0.0016)	0.0601 (0.0015)	0.0565 (0.0012)	0.0531 (0.0010)
	Test loss	0.0712 (0.0093)	0.0625 (0.0059)	0.0583 (0.0055)	0.0553 (0.0051)

Table 25: CIFAR-10: all implementation-study budgets. Mean (sample SD), five successful runs unless marked. Raw tracked-state cross-entropy and test accuracy.

Method	Metric	0.2 s	0.5 s	1 s	3 s
Hq-Recycle	Val. loss	1.5259 (0.0065)	1.5258 (0.0065)	1.5258 (0.0065)	1.5258 (0.0065)
	Test loss	1.5212 (0.0028)	1.5211 (0.0027)	1.5211 (0.0027)	1.5211 (0.0027)
	Acc. (%)	74.55 (0.18)	74.60 (0.19)	74.60 (0.24)	74.55 (0.19)
qNBO	Val. loss	1.5398 (0.0075)	1.5325 (0.0069)	1.5304 (0.0068)	1.5285 (0.0066)
	Test loss	1.5348 (0.0032)	1.5275 (0.0036)	1.5256 (0.0036)	1.5235 (0.0033)
	Acc. (%)	73.87 (0.53)	74.05 (0.49)	74.09 (0.48)	74.18 (0.48)
SHINE	Val. loss	1.5374 (0.0067)	1.5327 (0.0062)	1.5314 (0.0060)	1.5305 (0.0061)
	Test loss	1.5286 (0.0043)	1.5242 (0.0029)	1.5233 (0.0026)	1.5228 (0.0028)
	Acc. (%)	73.92 (0.70)	73.96 (0.45)	74.09 (0.41)	74.12 (0.38)
PZOBO	Val. loss	2.1563 (0.0609)	2.0122 (0.0387)	1.9335 (0.0300)	1.8866 (0.0092)
	Test loss	2.1550 (0.0597)	2.0095 (0.0370)	1.9305 (0.0272)	1.8838 (0.0073)
	Acc. (%)	55.07 (4.29)	65.24 (2.35)	68.96 (2.54)	70.91 (1.10)
BOME	Val. loss	2.0340 (0.0037)	1.8152 (0.0095)	1.5974 (0.0221)	1.4111 (0.0117)
	Test loss	2.0464 (0.0028)	1.8287 (0.0057)	1.6172 (0.0197)	1.4238 (0.0060)
	Acc. (%)	58.56 (3.43)	67.79 (3.25)	70.42 (3.79)	73.29 (1.25)
HOAG	Val. loss	1.5308 (0.0067)	1.5275 (0.0066)	1.5268 (0.0065)	1.5262 (0.0065)
	Test loss	1.5265 (0.0030)	1.5230 (0.0027)	1.5222 (0.0029)	1.5216 (0.0029)
	Acc. (%)	74.42 (0.24)	74.56 (0.19)	74.61 (0.21)	74.62 (0.18)
AID-CG (cold)	Val. loss	1.6946 (0.1097)	1.5486 (0.0215)	1.5291 (0.0073)	1.5265 (0.0066)
	Test loss	1.6904 (0.1098)	1.5437 (0.0191)	1.5244 (0.0043)	1.5220 (0.0030)
	Acc. (%)	72.68 (1.00)	74.12 (0.19)	74.46 (0.18)	74.55 (0.21)

Table 26: Fashion-MNIST: all implementation-study budgets. Mean (sample SD), five successful runs unless marked. Raw tracked-state cross-entropy and test accuracy.

Method	Metric	0.2 s	0.5 s	1 s	3 s
Hq-Recycle	Val. loss	1.1667 (0.0241)	1.1666 (0.0241)	1.1666 (0.0241)	1.1666 (0.0241)
	Test loss	1.1934 (0.0014)	1.1935 (0.0014)	1.1935 (0.0013)	1.1935 (0.0013)
	Acc. (%)	71.81 (0.74)	71.79 (0.65)	71.76 (0.62)	71.78 (0.65)
qNBO	Val. loss	1.1835 (0.0212)	1.1821 (0.0223)	1.1803 (0.0223)	1.1765 (0.0224)
	Test loss	1.2076 (0.0043)	1.2076 (0.0041)	1.2074 (0.0041)	1.2049 (0.0043)
	Acc. (%)	71.66 (1.07)	71.43 (0.90)	71.34 (0.89)	71.31 (0.95)
SHINE	Val. loss	1.1830 (0.0247)	1.1763 (0.0250)	1.1736 (0.0247)	1.1703 (0.0246)
	Test loss	1.2055 (0.0036)	1.1997 (0.0029)	1.1976 (0.0020)	1.1950 (0.0018)
	Acc. (%)	72.21 (0.79)	71.92 (0.91)	71.86 (1.03)	71.67 (1.03)
PZOBO	Val. loss	1.9662 (0.1017)	1.7565 (0.0545)	1.6857 (0.0300)	1.6361 (0.0273)
	Test loss	1.9733 (0.0987)	1.7674 (0.0472)	1.6983 (0.0189)	1.6507 (0.0103)
	Acc. (%)	60.81 (3.35)	66.41 (0.93)	68.58 (1.00)	68.88 (1.46)
BOME	Val. loss	1.7136 (0.0528)	1.3757 (0.0470)	1.1413 (0.0373)	1.0486 (0.0344)
	Test loss	1.7421 (0.0564)	1.4143 (0.0392)	1.1843 (0.0195)	1.0906 (0.0071)
	Acc. (%)	62.02 (1.31)	67.83 (1.55)	71.02 (1.01)	71.61 (1.49)
HOAG	Val. loss	1.1715 (0.0237)	1.1684 (0.0241)	1.1675 (0.0242)	1.1669 (0.0241)
	Test loss	1.1982 (0.0013)	1.1954 (0.0011)	1.1944 (0.0011)	1.1938 (0.0013)
	Acc. (%)	71.81 (0.88)	71.76 (0.89)	71.77 (0.83)	71.73 (0.75)
AID-CG (cold)	Val. loss	1.2973 (0.0549)	1.1847 (0.0241)	1.1738 (0.0233)	1.1705 (0.0234)
	Test loss	1.3183 (0.0599)	1.2096 (0.0084)	1.1996 (0.0021)	1.1968 (0.0015)
	Acc. (%)	69.83 (1.68)	71.66 (1.01)	71.86 (0.98)	71.85 (0.94)

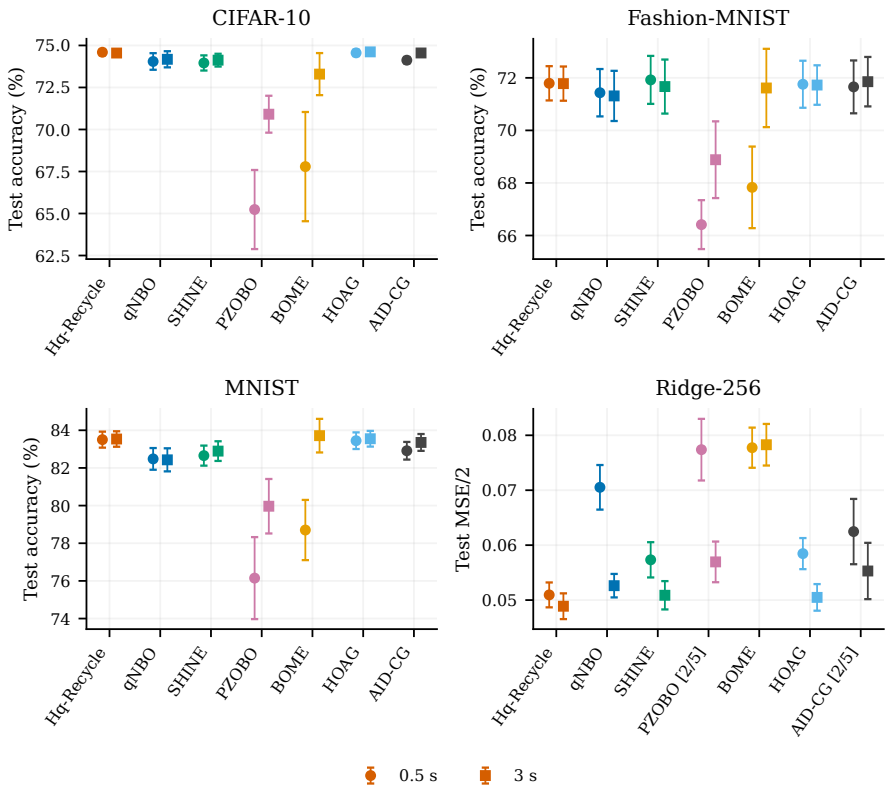

Figure 12: Five-seed implementation study test endpoints at 0.5 and 3 seconds, mean and sample SD. Classification reports accuracy, and ridge reports independent-response MSE/2. Labels marked [2/5] denote two completed trajectories out of five attempts; all other groups have five. Test metrics were not used to select these configurations. AID-CG denotes the cold-start adapter.

Table 27: MNIST: all implementation-study budgets. Mean (sample SD), five successful runs unless marked. Raw tracked-state cross-entropy and test accuracy.

Method	Metric	0.2 s	0.5 s	1 s	3 s
Hq-Recycle	Val. loss	1.1366 (0.0047)	1.1365 (0.0047)	1.1365 (0.0047)	1.1365 (0.0047)
	Test loss	1.1305 (0.0018)	1.1306 (0.0019)	1.1307 (0.0019)	1.1307 (0.0019)
	Acc. (%)	83.54 (0.41)	83.50 (0.43)	83.50 (0.39)	83.54 (0.41)
qNBO	Val. loss	1.1566 (0.0059)	1.1539 (0.0046)	1.1517 (0.0042)	1.1465 (0.0047)
	Test loss	1.1431 (0.0015)	1.1425 (0.0026)	1.1422 (0.0036)	1.1379 (0.0021)
	Acc. (%)	82.68 (0.57)	82.48 (0.58)	82.30 (0.59)	82.43 (0.61)
SHINE	Val. loss	1.1533 (0.0081)	1.1488 (0.0068)	1.1455 (0.0061)	1.1418 (0.0049)
	Test loss	1.1374 (0.0029)	1.1339 (0.0019)	1.1317 (0.0017)	1.1301 (0.0023)
	Acc. (%)	82.48 (0.45)	82.65 (0.53)	82.72 (0.51)	82.89 (0.52)
PZOBO	Val. loss	2.0779 (0.0052)	1.8055 (0.0436)	1.6875 (0.0192)	1.6310 (0.0146)
	Test loss	2.0744 (0.0056)	1.7975 (0.0434)	1.6785 (0.0189)	1.6219 (0.0141)
	Acc. (%)	66.99 (3.35)	76.15 (2.18)	79.18 (1.86)	79.97 (1.45)
BOME	Val. loss	1.7558 (0.0724)	1.3545 (0.0785)	1.1078 (0.0231)	1.0053 (0.0107)
	Test loss	1.7694 (0.0665)	1.3737 (0.0762)	1.1265 (0.0255)	1.0170 (0.0123)
	Acc. (%)	73.61 (5.69)	78.70 (1.60)	82.14 (0.89)	83.71 (0.89)
HOAG	Val. loss	1.1410 (0.0043)	1.1381 (0.0046)	1.1373 (0.0047)	1.1367 (0.0047)
	Test loss	1.1353 (0.0021)	1.1325 (0.0021)	1.1316 (0.0019)	1.1310 (0.0020)
	Acc. (%)	83.38 (0.48)	83.45 (0.44)	83.53 (0.45)	83.55 (0.42)
AID-CG (cold)	Val. loss	1.3274 (0.0655)	1.1627 (0.0106)	1.1423 (0.0040)	1.1384 (0.0044)
	Test loss	1.3196 (0.0631)	1.1561 (0.0074)	1.1364 (0.0014)	1.1326 (0.0021)
	Acc. (%)	81.09 (0.99)	82.91 (0.47)	83.25 (0.44)	83.35 (0.45)