BUDGET-AWARE STEP SELECTION FOR FORWARD-ONLY RECOVERY OF ANALOG ACCELERATORS

Anonymous authors
Paper under double-blind review

ABSTRACT

Recovering a drifted analog accelerator with forward-only optimization costs queries both to update its parameters and to choose those updates. We study step selection under this combined cost. In a controlled testbed, measured single-step descent boundaries exceed whole-run failure boundaries by a geometric mean of $27.2\times$: initial progress alone is a poor guide to a full recovery run. We isolate accumulated readout noise, derive a finite-horizon displacement bound, and use its step scale to select a rate from donor-device measurements without searching on the target device. The resulting model captures observed noise and probe scaling, but overpredicts failure boundaries by $2.20\times$ in geometric mean; its budget dependence remains unresolved. On 36 held-out simulated crossbar targets, a selector with a fine rate grid and donors matched by device condition uses 93.62% fewer operational forward evaluations than target search, with a 0.399 percentage-point accuracy deficit. A transferred condition-specific fixed setting is more accurate at lower incremental cost, showing that avoiding target search need not require donor measurements. Calibration uses clean simulator readout, and the bound concerns noise-only displacement, not device reliability. These results support evaluating recovery methods by both accuracy and configuration cost, with fixed transfer as an explicit comparator. An anonymous project page is available at https://dordlark158.github.io/budget-aware-step-selection/.

1 INTRODUCTION

Analog accelerators encode neural-network parameters in physical states, such as conductances or optical phases, that drift after deployment. Hardware-aware training anticipates these imperfections offline (Rasch et al., 2023), but deployed devices may still need adaptation. Recovery must then justify the device queries it consumes, including those spent finding a suitable configuration before useful adaptation begins.

We study devices that expose writable parameters and forward evaluations of a scalar loss, but no backward pass for computing gradients. Unlike in-situ methods requiring backward signals and specialized update circuitry (Gokmen & Haensch, 2020; Gokmen, 2021; Wu et al., 2024; Rasch et al., 2024), forward-only methods estimate updates from parameter perturbations (Spall, 1992; Malladi et al., 2023; Liu et al., 2025), including on photonic hardware (Gu et al., 2020; Bandyopadhyay et al., 2024). Their usefulness depends on more than the accuracy of a selected run: searching for that run can cost many times its recovery budget. Shared calibration uses measurements from other devices to choose a configuration; fixed transfer reuses a previously selected setting. We therefore ask: *what accuracy–cost trade-off can shared calibration achieve relative to target-device search and fixed transfer?*

Why the recovery horizon matters. Two-point zeroth-order (ZO) optimization estimates a gradient from forward evaluations at two opposite perturbations of amplitude ε, then applies a step of size η. Small steps may exhaust the budget without recovery; large steps may cause numerical divergence or severe accuracy loss. A natural starting point is the step size that gives expected descent. Yet this is a single-step condition at one state, whereas recovery requires a useful trajectory over the entire budget. In our testbed, directly measured descent boundaries exceed whole-run failure boundaries by $27.2\times$ in geometric mean. This gap motivates a separate analysis of what accumulates over successive updates.

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.

Readout noise varies between evaluations, unlike persistent parameter drift. Its contribution persists even when the useful update becomes small. We isolate this contribution in a noise-only process and bound its finite-horizon displacement. The resulting scale relates the step size to readout noise, probe amplitude, direction count and query budget. We use measurements from other devices, called *donors*, to select the largest rate below a model-based threshold, or abstain. The bound applies to the idealized process; whether its scale predicts failure or selects useful recovery steps is an empirical question.

What shared calibration buys. We evaluate these questions separately. A controlled study freezes policies before recovery on 36 targets with disjoint donors and counts calibration, search, monitoring and validation. It shows that donor selection can substantially reduce queries relative to target search, but that a transferred fixed setting can be both cheaper and more accurate. Broader screens show why this comparator matters: changing probe and step selection changes optimizer rankings. The practical contribution is therefore an evaluated route to reducing configuration cost, with explicit tests of when the measurements themselves add value.

1.1 Contributions

1. **Separating initial descent from whole-run failure.** We measure their separation and test the scaling of a readout-noise displacement model, including its prediction errors and unresolved budget dependence (Section 4.1).
2. **A budget-aware selector and controlled cost comparison.** We derive the noise-only bound and turn its scale into a donor-calibrated rule. A finer-grid, cohort-specific extension incurs a 0.399 percentage-point (pp) deficit to target search with 93.62% fewer operational forwards, while stronger fixed-transfer controls delimit its benefit (Sections 3 and 4.3).
3. **Configuration-aware evaluation of recovery.** Joint probe–step selection changes comparisons between ZO, POEM (Ren & Luo, 2025) and PolyStep (Le, 2026). We distinguish shared acquisition, target search and recovery costs, and compare methods with comparable configuration screens (Section 4.4).

All hardware evaluations are simulations on small classification tasks. The controlled study separates device roles and final testing; the cross-substrate screens are retrospective. Clean radius calibration is simulator-assisted. These boundaries determine the claims we can make about deployment.

2 Problem setting and evaluation protocol

2.1 Forward-only recovery and its costs

We train a network offline, simulate degradation, and adapt exposed parameters $\theta \in \mathbb{R}^D$ while holding the others fixed. Recovery can write these parameters and query a scalar forward loss, but cannot request gradients. Labeled training and validation examples are available. The task is to choose a recovery configuration and recover accuracy within a declared optimization budget.

One unit of forward-evaluation expenditure (FEE) is a scalar training-minibatch loss query. A two-point update with m directions costs $2m$ queries; the controlled study fixes $B = 30{,}000$ FEE per run. This optimization budget determines the horizon in our analysis. To compare deployment policies, we separately count actual forward batches for calibration, configuration search, recovery, monitoring and validation, including failed candidates. A full-split evaluation may require multiple batches. For a fleet of N targets,

$$C_{\mathrm{op}}(N) = C_{\mathrm{shared}} + \sum_{j=1}^{N} C_{\mathrm{target},j}.$$

Shared calibration or fixed-policy acquisition is paid once for the fleet. Target search pays for every attempted candidate, including its selected run once. Previously acquired fixed settings incur only current operational cost; their earlier acquisition is reported separately. Final research-test and factory-preparation calls are also itemized separately (Appendix B).

2.2 SELECTION AND FAILURE ARE DIFFERENT ENDPOINTS

The controlled study uses disjoint donor and target devices and starts every policy from identical serialized target states. Donors supply either model coefficients or fixed settings; target search instead evaluates a declared rate grid on each target. All transfer policies are frozen before target recovery and all final choices before target testing.

Within each run, the final parameter state replaces the damaged baseline only if it improves mean accuracy over five independent noisy validation draws per state. Otherwise the policy returns the baseline. Separate monitoring at 31 budget positions flags mean validation accuracy at or below 15%; numerical failures are recorded when they occur. We call either kind of event an *escape*. Reverting to baseline does not erase an earlier event, and sampled monitoring cannot detect every intermediate failure. Primary performance is test accuracy with transient readout noise disabled and static damage retained (*clean readout*). Repeated noisy test evaluations are secondary measurements, not additional devices.

The mechanism testbed instead defines escape as non-finite loss. Retrospective deployment screens use single-draw checkpoint selection, test-guided configuration search and, in some crossbar comparisons, unmatched damaged states. We keep their diagnostics separate from the controlled study. Section 4 specifies the devices and comparators; Appendix C gives the full protocols. We next derive a step scale from the query budget, then test its predictive accuracy and value for configuration selection.

3 BUDGET-AWARE STEP SELECTION

A recovery step must produce useful progress without accumulating excessive noise over the run. We separate these requirements: a descent calculation describes one update, while a finite-horizon displacement bound supplies the budget dependence used by our donor rule.

3.1 ESTIMATOR NOISE AND SINGLE-STEP DESCENT

Let $f : \mathbb{R}^D \to \mathbb{R}$ be the clean-readout loss on a fixed minibatch. For the analysis, assume f is continuously differentiable and $f, \nabla f$ have at most polynomial growth. Independently draw $u_i \sim \mathcal{N}(0, I_D)$ and query $y_i^{\pm} = f(\theta \pm \varepsilon u_i) + \xi_i^{\pm}$, where $\varepsilon > 0$ and the errors $\xi_i^{\pm}$ are independent of one another and the directions, with mean zero and variance σ_{r}^2. The averaged two-point estimator is

$$\hat{g}_m = \frac{1}{m} \sum_{i=1}^{m} \frac{y_i^+ - y_i^-}{2\varepsilon} u_i.$$

Thus ε is the root-mean-square (RMS) probe displacement per coordinate; its RMS Euclidean length is $\varepsilon\sqrt{D}$. Gaussian integration by parts gives $\mu := \mathbb{E}\hat{g}_m = \nabla f_\varepsilon(\theta)$, where $f_\varepsilon(\theta) = \mathbb{E}_u f(\theta + \varepsilon u)$ (Nesterov & Spokoiny, 2017). Appendix F derives this identity.

To separate direction sampling from readout noise, write each summand as its noiseless finite difference plus $n_i = (\xi_i^+ - \xi_i^-)u_i/(2\varepsilon)$. Since $\mathbb{E}[n_i \mid u_i] = 0$ and $\mathbb{E}[n_i n_i^\top] = \sigma_{\mathrm{r}}^2 I_D/(2\varepsilon^2)$, the cross-covariance vanishes. Independence across directions then gives

$$\mathrm{Cov}(\hat{g}_m) = \frac{1}{m}\left(\Sigma_{\mathrm{search}} + \frac{\sigma_{\mathrm{r}}^2}{2\varepsilon^2} I_D\right), \qquad V_m = \frac{1}{m}\left(A_{\mathrm{search}} + \frac{D\sigma_{\mathrm{r}}^2}{2\varepsilon^2}\right), \tag{1}$$

where Σ_{search} is the noiseless single-direction covariance, $A_{\mathrm{search}} = \mathrm{tr}\,\Sigma_{\mathrm{search}}$, and $V_m = \mathbb{E}\|\hat{g}_m - \mu\|^2$ is the total variance. Averaging reduces both terms by $1/m$; it does not preferentially remove readout noise. At the testbed reference state the readout term is about $27\times$ the search term; their deployment ratio is unmeasured.

To relate this variance to progress, assume f has an L-Lipschitz gradient, $L > 0$. Define $G = \|\nabla f(\theta)\|$, $S = \nabla f(\theta)^\top \mu$, $M = \|\mu\|^2$ and $A = A_{\mathrm{search}} + D\sigma_{\mathrm{r}}^2/(2\varepsilon^2)$. The descent lemma gives

$$\mathbb{E}[f(\theta - \eta\hat{g}_m)] - f(\theta) \le -\eta S + \frac{L\eta^2}{2}(M + A/m).$$

Consequently, when $S > 0$, expected descent is guaranteed for

$$0 < \eta < \eta_{\mathrm{ED}}(m) := \frac{2mS}{L(mM + A)}. \tag{2}$$

Above this sufficient threshold the bound is inconclusive. The alignment S matters because finite probes estimate ∇f_ε, not ∇f: $\|\mu - \nabla f\| \le L\varepsilon\sqrt{D}$. Increasing the probe reduces the explicit readout variance but can increase smoothing bias. If $\mu \approx \nabla f$ and $D\sigma_{\mathrm{r}}^2/(2\varepsilon^2) \gg A_{\mathrm{search}} + mM$, then $\eta_{\mathrm{ED}} \approx 4m\varepsilon^2 G^2/(LD\sigma_{\mathrm{r}}^2)$.

Curvature as a diagnostic. Where $H = \nabla^2 f(\theta)$ exists, the local quadratic term is $\mathbb{E}[\hat{g}_m^\top H \hat{g}_m] = \mu^\top H \mu + \mathrm{tr}(H\Sigma_{\mathrm{search}})/m + \sigma_{\mathrm{r}}^2 \,\mathrm{tr}(H)/(2\varepsilon^2 m)$. Its readout contribution can be much smaller than the smoothness estimate using LD. We compare this trace approximation with measured descent and failure boundaries in Section 4.1. Substituting a local Hessian norm L_{loc} for L gives another diagnostic, not a smoothness bound along the update.

3.2 Accumulated noise over the recovery horizon

The descent condition controls expected loss change at one state; it does not bound how far repeated noisy updates can move the parameters. The readout term persists even when the gradient vanishes. To isolate its cumulative effect, retain only the readout-noise part of each update. Each averaged update costs $2m$ queries, so budget B permits $T = \lfloor B/(2m) \rfloor$ updates.

Proposition 1 (Noise-only displacement) *Let $D, m \ge 1$ and $T \ge 0$ be integers, $\varepsilon, R > 0$, $\sigma_{\mathrm{r}} \ge 0$, and $\eta_t \ge 0$ be deterministic steps. Let $n_t \in \mathbb{R}^D$ be independent, mean-zero vectors with covariance $\sigma_{\mathrm{r}}^2 I_D/(2\varepsilon^2 m)$. For $X_0 = 0$ and $X_{t+1} = X_t - \eta_t n_t$,*

$$Q_T := \mathbb{E}\|X_T\|^2 = \frac{D\sigma_{\mathrm{r}}^2}{2\varepsilon^2 m} \sum_{t=0}^{T-1} \eta_t^2, \qquad \mathbb{P}\left(\max_{0 \le t \le T} \|X_t\| \ge R \right) \le \min\{1, Q_T/R^2\}. \tag{3}$$

Proof. Put $c = D\sigma_{\mathrm{r}}^2/(2\varepsilon^2 m)$ and let $\mathcal{F}_t = \sigma(n_0, \dots, n_{t-1})$ denote the information available before update t. Expanding $X_T = X_t - \sum_{s=t}^{T-1} \eta_s n_s$ and using independence and zero means gives

$$\mathbb{E}[\|X_T\|^2 \mid \mathcal{F}_t] = \|X_t\|^2 + c \sum_{s=t}^{T-1} \eta_s^2 \ge \|X_t\|^2.$$

Taking $t = 0$ proves the identity for Q_T. Let $E_t = \{\|X_t\| \ge R,\ \|X_s\| < R \text{ for all } s < t\}$ be the first-exit event at t. These events are disjoint and $E_t \in \mathcal{F}_t$, so

$$Q_T \ge \sum_{t=0}^{T} \mathbb{E}[\mathbf{1}_{E_t} \|X_T\|^2] \ge \sum_{t=0}^{T} \mathbb{E}[\mathbf{1}_{E_t} \|X_t\|^2] \ge R^2 \sum_{t=0}^{T} \mathbb{P}(E_t).$$

Their union is the exit event. Dividing by R^2 and using probability at most one proves the bound, also for $T = 0$. □

For $B \ge 2m$ and $\sigma_{\mathrm{r}} > 0$, define $\eta_{\mathrm{RMS}} = (T^{-1} \sum_{t=0}^{T-1} \eta_t^2)^{1/2}$ and the step scale

$$\eta_{\mathrm{noise}}(m, B) := \frac{2R\varepsilon m}{\sigma_{\mathrm{r}}\sqrt{DB}}. \tag{4}$$

The ratio $\Pi_{\mathrm{esc}} := \eta_{\mathrm{RMS}}/\eta_{\mathrm{noise}}$ compares the scheduled RMS step with this scale. Since $2mT \le B$, the proposition gives $Q_T/R^2 \le \Pi_{\mathrm{esc}}^2$, with equality when $B = 2mT$. Thus $\eta_{\mathrm{RMS}} \le \sqrt{\delta}\,\eta_{\mathrm{noise}}$ bounds the noise-only exit probability by δ. The linear m factor combines lower per-update variance with fewer affordable updates; $B^{-1/2}$ compensates for accumulation over a longer horizon.

From displacement to a failure model. The scale η_{noise} corresponds to an RMS noise displacement of R when $B = 2mT$. We test whether it approximates $\eta_{\max}$, the largest tested escape-free rate. This approximation omits search variance and objective-driven motion and represents the usable region by one radius. To estimate R, we move along 24 random rays from the initial damaged

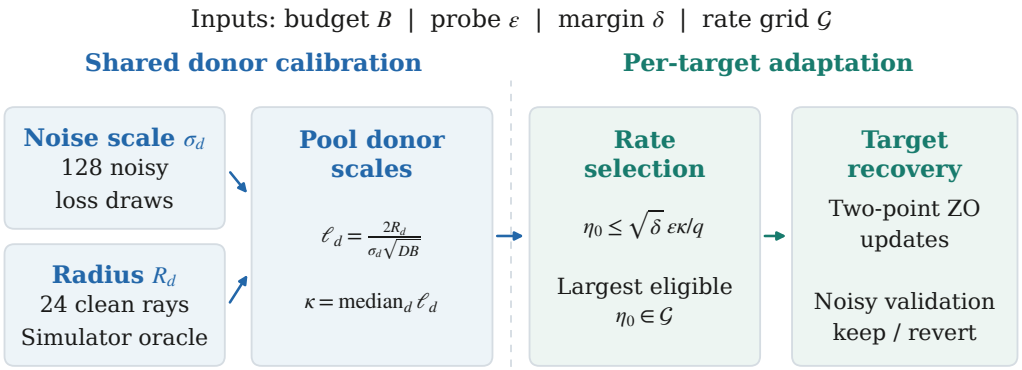


Figure 1: Donor-calibrated step selection for $m = 1$. Shared measurements determine a rate without target search. Here σ_d is donor loss-noise standard deviation, $\kappa = \text{median}_d\, \ell_d$, $\mathcal{G}$ is the candidate ladder and $q = \eta_{\text{RMS}}/\eta_0$. Clean radius measurements use a simulator oracle. Unsupported calibration or no eligible rate returns the damaged baseline.

state until the clean cross-entropy reaches $\log 10$, the loss of uniform ten-class prediction, and average the crossing distances. We estimate σ_{r} from repeated losses at that state. The simulator's weight perturbations need not induce unbiased, constant-variance scalar loss noise, and the mean radius need not enclose a failure-free ball. The proposition therefore bounds the stated noise process, not device failure. The failure model is motivated by the testbed observations; its coefficients are measured without fitting failure boundaries (Appendix H).

3.3 Selecting a step from donor measurements

Shared donor measurements replace calibration on each target (Figure 1). Factoring εm out of Eq. 4 leaves a coefficient determined by the device and budget. Given a probe ε, a budget $B \ge 2m$, a margin $0 < \delta < 1$ and a donor pool $\mathcal{D}$, estimate this coefficient by

$$\ell_u = \frac{2R_u}{\sigma_{\text{r},u}\sqrt{D_u B}}, \qquad \kappa = \text{median}_{u\in\mathcal{D}}\, \ell_u.$$

The target is excluded from $\mathcal{D}$. An empty pool, a missing ray crossing, or any nonfinite or nonpositive radius or noise estimate causes abstention. For a prescribed schedule $\eta_t = \eta_0 a_t$, let $q = (T^{-1}\sum_{t=0}^{T-1} a_t^2)^{1/2} > 0$. On a finite positive rate ladder $\mathcal{G}$, select

$$\eta_0^* = \max\{\eta_0 \in \mathcal{G} : q\eta_0 \le \sqrt{\delta}\, \varepsilon m \kappa\}. \tag{5}$$

If the set is empty, retain the damaged state. Choosing the largest eligible rate favors movement within the model margin; it does not assert accuracy optimality. Neither the donor median nor the radius proxy provides a target-device risk bound.

Based on prior probe screens, the controlled study fixes $\varepsilon = 0.1$, $m = 1$; it uses $B = 30{,}000$ and $\delta = 0.1$. It varies two choices: the rate grid and the donor pool. The *coarse* grid is $\{10^{-3}, 10^{-2}, 10^{-1}\}$; the *fine* grid has 21 rates, $\{10^{k/4} : k = -24, \ldots, -4\}$. The *substrate* pool uses all 36 crossbar donors; the *cohort* pool uses the two donors matching the target's pretrained base model, noise level and drift duration. Labels such as *fine/cohort* specify both choices.

Each donor supplies 128 loss draws and clean-readout radius measurements. Cosine decay toward $\eta_0/100$ uses $a_t = 0.505 + 0.495\cos(\pi t/T)$, giving $q \approx 0.61447$ at $T = 15{,}000$; Appendix I derives the exact expression. All policies are fixed before target recovery. Calibration needs no gradients or Hessians, but obtaining its clean-readout measurements on physical devices remains untested.

4 Experiments

The experiments test the argument in three stages. We first compare the noise-displacement scale with measured failure boundaries, then evaluate the donor rule's accuracy and cost against search and

fixed transfer. Finally, we test whether configuration selection also changes comparisons between recovery algorithms.

Testbeds and devices. The mechanism testbed allows updates to all $101{,}770$ network parameters; its σ scales Gaussian weight noise relative to each tile's largest absolute weight. The deployment simulators restrict updates to $1{,}290$ output-layer parameters in resistive-memory (ReRAM) crossbars or $4{,}864$ optical phases in photonic meshes. Crossbars use programming noise and one- or two-day drift followed by global drift compensation; photonic meshes use fabrication offsets and thermal noise. Both classify MNIST digits, with 8×8 inputs for photonics. Appendix C gives architectures, simulators and settings.

Controlled evaluation. Three independently trained base models, called factories, each supply six ReRAM cohorts: read-noise settings $\{0.03, 0.1, 0.2\}$ crossed with one- and two-day drift. Two donors and two targets per factory–cohort give 36 of each. MNIST is split into $48{,}000$ training, $6{,}000$ development, $6{,}000$ validation and $10{,}000$ test examples. Factory epochs and configuration searches use development data. Primary policies hold $\varepsilon = 0.1$, $m = 1$. Controls distinguish three ways to obtain a step size: reuse the fixed default ($\eta_0 = 0.001$) or settings from a separate calibration population, search the donors, or search a 21-rate grid on each target. Donor search produces either one rate for all targets or a rate per cohort. Donor search and target search use clean development scoring, a simulator oracle. Results are descriptive means balanced across factories and cohorts; the 36 targets share three factories.

Broader method comparisons. We compare single-direction ZO, direction averaging, POEM and PolyStep. POEM adapts step and probe scales from an initial-distance parameter. PolyStep evaluates the vertices of a rotated polytope in a lower-dimensional subspace and moves toward their weighted mean, giving lower-loss vertices more weight. Its probe, step and weighting temperature jointly control the update (Appendix E). Deployment averaging uses m sequential microsteps of size η/m, unlike the same-state estimator analyzed above. Benchmark summaries use mean $\pm$ sample standard deviation (SD) and paired tests where device realizations match. Correction for multiple comparisons is specified in Appendix J. Supplementary tests also examine recovery of quantized weights in a spiking keyword classifier.

4.1 Separating descent from failure

We measure the largest tested rate with no escape over a full run. A logarithmic ladder brackets this boundary; model coefficients are measured separately at the initial damaged state (Appendix K).

Single-step descent boundaries are located by bisecting the sign of $\mathbb{E}[f(\theta - \eta\hat{g}_m)] - f(\theta)$, using 128 noisy estimates and clean objective evaluation. Across 27 comparisons, these boundaries exceed whole-run failure boundaries by $27.2\times$ in geometric mean. The trace approximation better locates descent than failure: its predicted/measured geometric mean is 0.834 on three seed-matched comparisons at $m = 1$, $\sigma = 0.2$ (Appendix L).

The noise-displacement model overpredicts failure boundaries by $2.20\times$ in geometric mean, with $15/21$ within grid resolution and $10/21$ within $2\times$ (Figure 2). Measured direction-count and readout-noise exponents, $+0.969$ and -1.059, support its m/σ_{r} dependence rather than the descent approximation's σ_{r}^{-2} scaling. Deterministic runs remain escape-free beyond the local $2/L_{\mathrm{loc}}$ reference. Increasing m from 1 to $3{,}189$ at matched queries raises the largest tested escape-free rate by $1{,}000\times$, without a resolved increase in the fraction of lost accuracy recovered across three seeds (Appendix R).

4.2 Budget dependence and transfer diagnostics

Agreement in noise and direction scaling does not establish the budget dependence needed by the selector. The predicted $B^{-1/2}$ scaling remains unresolved. Escape-time fits imply an exponent -0.411 with 95% interval $[-0.69, -0.29]$; direct budget-boundary measurements give -0.282 with interval $[-0.43, -0.13]$, excluding $-1/2$. The former pools three seeds at one noise level; the latter uses one seed. Nearly budget-independent escape time, when failure precedes the budget limit, is

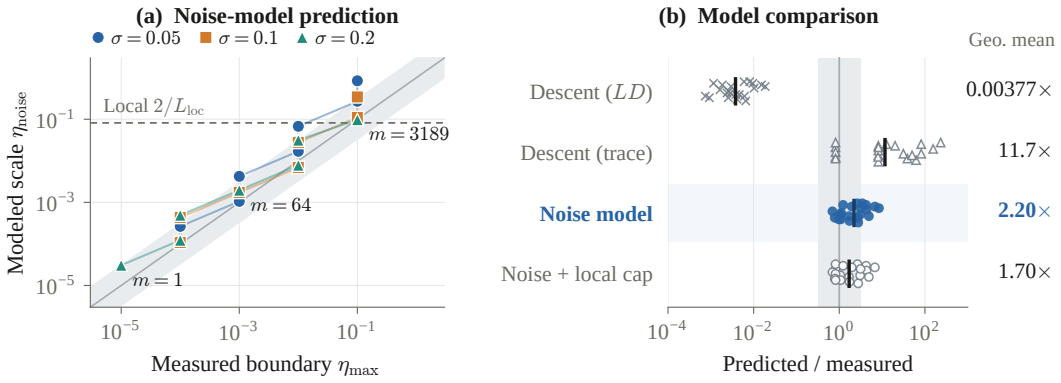


Figure 2: Predicting whole-run failure across 21 testbed settings. (a) Noise-model predictions; labels mark direction counts at $\sigma = 0.2$, and the dashed line marks $2/L_{\text{loc}}$. (b) Prediction/measurement ratios; vertical marks and the right column show geometric means. Gray bands denote $\sqrt{10}$ grid resolution, not statistical uncertainty. Appendix H defines the comparison diagnostics.

consistent with both free diffusion and restoring drift. The available measurements do not establish a replacement model (Appendices H and L).

Transfer across devices. We next test whether the scale transfers from the mechanism testbed to restricted device interfaces. Five crossbar and five photonic realizations, each with two estimators and 5×3 probe–step grids, provide 300 outcomes. Predicting escape when $\Pi_{\text{esc}} > 1$ classifies $277/300$ outcomes correctly with device-specific coefficients and $280/300$ with median donor coefficients. Locating the boundary is harder: for each fixed probe, three rates define a boundary interval. Device-specific predictions locate $74/100$ intervals within grid resolution, versus $44/100$ for always predicting the most frequent category, and overpredict bracketed boundaries by $3.23\times$ in geometric mean. The probe exponent estimate $+0.90$ has a model-conditional interval $[+0.76, +1.08]$; synthetic checks find coverage below the nominal level. These counts reuse devices, and one crossbar screen/calibration pair has mismatched damage (Appendix O).

Retrospective recovery across substrates. At $\delta = 0.1$, none of the 20 selected runs across ten devices has a recorded escape. Mean deficits to per-device 15-configuration test-grid maxima are 0.025 pp on crossbars and 4.26 pp on photonics. The latter is concentrated in averaging, with mean/worst gaps of $6.65/11.47$ pp. All selections are reproduced by fixed transfer: $(\varepsilon, \eta_0) = (0.1, 0.001)$ for $m = 1$, $(0.01, 0.01)$ for averaging. Target devices are excluded from their donor coefficient pools but occur in prior probe screens. These diagnostics cannot show whether donor measurements improve on fixed transfer; Table 8 and margin sensitivity appear in Appendix I.

4.3 Controlled donor transfer: accuracy and operational cost

The retrospective results leave two questions: what accuracy is lost by avoiding target search, and do donor measurements improve on a fixed setting? The controlled study compares these policies on disjoint targets with every required operational query counted. From mean damaged accuracy 93.35%, fine/cohort selection reaches 95.075%, versus 95.474% for target search. It requires 1,217,354 rather than 19,080,779 operational forwards across the fleet, a 93.62% reduction (Figure 3). The 0.399 pp accuracy deficit exceeds the prespecified 0.10 pp tolerance; this is a descriptive accuracy–cost trade-off, not statistical evidence of non-inferiority.

Does measuring donors add value? Coarse/substrate selection chooses $\eta_0 = 0.001$ on all 36 targets, exactly matching the fixed default's individual outcomes at 94.521%. Calibration adds 63,482 calls to the 1,153,872 calls both policies spend on recovery. The fine-grid, substrate-pooled policy gives the same selection. Fine/cohort improves on this default by 0.554 pp, but remains 0.316 pp below the transferred cohort-fixed policy's 95.391% and costs 5.50% more. Avoiding target search therefore does not require measurement when a suitable fixed setting is already available. Against

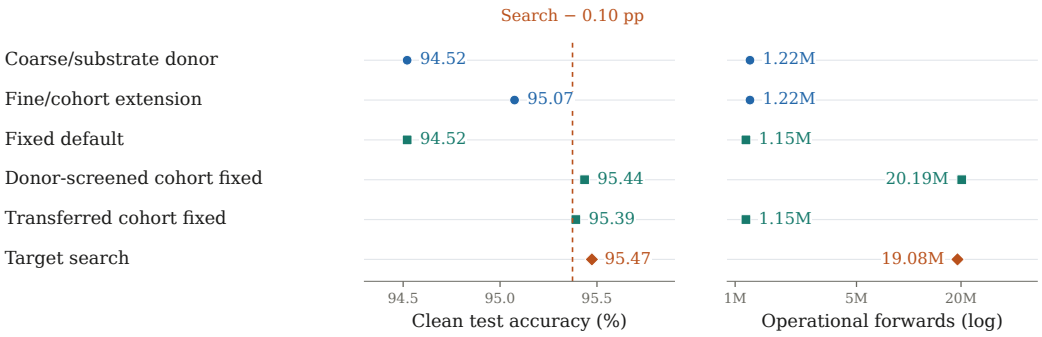


Figure 3: Recovery on 36 crossbar targets from three factories, at $\varepsilon = 0.1$, $m = 1$. Donor-screened fixed policies include acquisition on the donors; transferred fixed policies reuse settings from a separate calibration population. Costs include calibration, monitoring, validation and failed candidates. The dashed line marks the prespecified 0.10 pp tolerance below target search, not a confidence interval. Appendix B gives all policies, secondary metrics and the cost breakdown.

newly donor-screened cohort-fixed transfer, fine/cohort loses 0.361 pp and uses 93.97% fewer calls. The transferred setting's earlier acquisition remains a separate, incompletely costed expense.

Abstention, failures and probe sensitivity. Coarse/cohort abstains on $14/36$ targets and reaches 94.601%; abstained devices remain in the mean. Selected recovery runs have no recorded escape under any primary policy, including fixed controls. In contrast, $215/838$ target candidates across the primary grid and secondary probe tests have at least one such event; their costs and outcomes remain counted after reversion. At the secondary $\varepsilon = 0.03$, fine/cohort and transferred cohort-fixed reach 95.833% and 95.837%. This sparse sensitivity has no full donor-screening or target-search grid at that probe and cannot replace the primary cost comparison (Appendix B).

4.4 Configuration selection changes method comparisons

The controlled study holds the optimizer fixed to isolate step selection. We now ask whether configuration selection also changes comparisons between optimizers. Joint probe–step selection improves PolyStep over fixed configurations by 1.42–2.57 pp on four device-paired analog settings. On thermal meshes, changing the probe adds 2.15 pp beyond the plateau under step-only selection; the selected crossbar probe stays unchanged (Appendix G.1). Thus probe sensitivity depends on the substrate.

To make the search effort comparable, we examine 15-configuration thermal screens for each method. PolyStep's mean accuracy is 2.39 pp below single-direction ZO and 0.56 pp above averaging across ten chips at seed 42. These are descriptive paired comparisons. Screening uses 450,000 optimization FEE per ZO method and 446,175 for PolyStep, excluding calibration and validation. These test-guided screens do not establish independent holdout performance or equal total cost (Appendix C).

These comparisons show why the configuration protocol belongs in the recovery result. A method can look preferable at a fixed probe amplitude and lose that advantage after both probe and step are searched. Appendix S reports further benchmark results, averaging controls, POEM initialization sweeps and noisy-readout comparisons. They provide protocol sensitivity checks rather than a single ranking of recovery methods.

5 Related work

Forward-only learning and hardware calibration. NoiseZO (Wang et al., 2025) uses intrinsic resistive-memory write noise to generate perturbations for forward-only training. Our analysis concerns independent readout errors accumulated during recovery. DOCTOR (Lu et al., 2024) and

VeRA+ (Dong et al., 2026) restore hardware accuracy through calibration. Their measurement and actuation interfaces differ from the loss-only interface studied here.

Descent, stability and accumulated noise. Step-size limits depend on both curvature and the noise model. Song et al. (2026) bound the mean-square stability threshold of Gaussian ZO on an exact-oracle, positive-semidefinite quadratic using its Hessian trace and largest eigenvalue. Our model instead isolates readout noise that persists near stationary points. Two-point convergence guarantees also depend on the oracle: convex shared-realization analyses (Duchi et al., 2015), dimension-dependent iteration bounds (Nesterov & Spokoiny, 2017), and noisy-evaluation complexity (Jamieson et al., 2012) do not directly cover our nonconvex recovery setting with independent readout errors.

Adaptive scales and configuration cost. D-Adaptation (Defazio & Mishchenko, 2023), Prodigy (Mishchenko & Defazio, 2024) and DoG (Ivgi et al., 2023) adapt first-order step sizes online. POEM (Ren & Luo, 2025) adapts both the step and smoothing radius from an initial-distance parameter. Its bounded-domain guarantee assumes shared randomness between paired evaluations and projection; our implementation uses independent readout errors and unprojected updates. Selecting gains from noise measurements and an iteration budget has a long history in SPSA (Spall, 1998). Following work on optimizer tunability (Sivaprasad et al., 2020), we account separately for shared calibration and target-device search. Our contribution is the measured descent–failure separation and the controlled accuracy–cost evaluation of a displacement-based selector against search and fixed transfer. The covariance and second-moment arguments are standard mathematical tools used to construct that selector. Appendix A gives further comparisons.

6 Limitations and conclusion

Model and experimental scope. The analysis isolates independent readout noise and omits the objective's restoring dynamics. Its radius estimate does not define a verified safe region, and the experiments do not resolve the predicted budget exponent. Deployment averaging uses sequential updates rather than the same-state estimator analyzed in Section 3.1. All hardware results are simulations on small classification tasks. Physical hardware and larger models remain untested. The photonic loss is discontinuous at phase-wrap boundaries, and the spiking experiments have only three seeds and do not reach an accuracy plateau.

Selection and generalization. The controlled study separates donor and target devices, but uses only three factories and a dataset already used during method development. Radius calibration and development scoring require clean simulator readout; obtaining these measurements from noisy physical devices remains untested. Sampled monitoring does not certify a failure probability. The retrospective screens use test data for pretraining and configuration selection, reuse devices, omit some overhead costs and sometimes compare different damaged crossbar states (Appendix C). Incomplete acquisition costs prevent a full lifecycle comparison or a measured break-even fleet size. Query counts also do not establish hardware energy savings.

Conclusion. A step size that yields initial progress can fail over the full budget. The noise-displacement model turns this distinction into a budget-aware step rule. On the controlled targets, shared calibration reduces queries relative to target search at a measurable accuracy cost, while a transferred fixed setting performs better at lower incremental cost. The evidence supports evaluating calibration as a deployment expense whose value must be tested against fixed transfer. Together with the changes in optimizer rankings under joint probe–step selection, this makes configuration choice part of the recovery method and its cost.

AI USE STATEMENT

An AI assistant was used for editorial checks. The authors take full responsibility for all methods, mathematical investigations, experiments, and conclusions.

ETHICS STATEMENT

All hardware evaluations use simulation and existing classification datasets; this study recruits no new human participants. The model-based thresholds are not reliability certificates for safety-critical deployment. Dataset and checkpoint redistribution remain subject to their licenses; the spiking checkpoint is not distributable.

REPRODUCIBILITY STATEMENT

Section 3 proves the displacement bound; Appendices F and I derive the estimator identities, descent condition and schedule conversion. Appendices B and C specify the experimental protocols, data splits, replication units, configuration choices and query costs. The accompanying code provides experiment drivers, processed results, analysis scripts, figure generation and environment specifications. Some underlying simulator files are unavailable; data-availability limits and differences between crossbar realizations are described in Appendices C and J. An anonymous project page, `https://dordlark158.github.io/budget-aware-step-selection/`, presents the method and the two main experiments interactively from the same stored results.

REFERENCES

Farshid Ashtiani, Mohamad Hossein Idjadi, and Kwangwoong Kim. Integrated photonic neural network with on-chip backpropagation training. *Nature*, 651(8107):927–932, 2026. doi: 10.1038/s41586-026-10262-8.

Saumil Bandyopadhyay, Alexander Sludds, Stefan Krastanov, Ryan Hamerly, Nicholas Harris, Darius Bunandar, Matthew Streshinsky, Michael Hochberg, and Dirk Englund. Single-chip photonic deep neural network with forward-only training. *Nature Photonics*, 18:1335–1343, 2024. doi: 10.1038/s41566-024-01567-z.

Dmitriy Bystrov, Daniil Medyakov, Dmitry Bylinkin, and Aleksandr Beznosikov. Zero-order parameter-free optimization for LMO-based methods: Novel approach for efficient fine-tuning. *arXiv preprint arXiv:2606.14970*, 2026.

Jeremy M. Cohen, Simran Kaur, Yuanzhi Li, J. Zico Kolter, and Ameet Talwalkar. Gradient descent on neural networks typically occurs at the edge of stability. In *International Conference on Learning Representations (ICLR)*, 2021.

George E. Dahl, Frank Schneider, Zachary Nado, Naman Agarwal, Chandramouli Shama Sastry, Philipp Hennig, Sourabh Medapati, Runa Eschenhagen, Priya Kasimbeg, Daniel Suo, Juhan Bae, Justin Gilmer, Abel L. Peirson, Bilal Khan, Rohan Anil, Mike Rabbat, Shankar Krishnan, Daniel Snider, Ehsan Amid, Kongtao Chen, Chris J. Maddison, Rakshith Vasudev, Michal Badura, Ankush Garg, and Peter Mattson. Benchmarking neural network training algorithms. *arXiv preprint arXiv:2306.07179*, 2023.

Hassan Dbouk, Nidham Gazagnadou, Matthias Reisser, and Christos Louizos. On adaptivity in zeroth-order optimization. *arXiv preprint arXiv:2605.03869*, 2026.

Aaron Defazio and Konstantin Mishchenko. Learning-rate-free learning by D-Adaptation. In *International Conference on Machine Learning (ICML)*, volume 202 of *Proceedings of Machine Learning Research*, pp. 7449–7479, 2023.

Weirong Dong, Kai Zhou, Zhen Kong, Zhengke Yang, Quan Cheng, Haoyuan Li, Junkai Huang, Jun Lan, Yida Li, Masanori Hashimoto, and Longyang Lin. VeRA+: Vector-based lightweight digital compensation for drift-resilient RRAM in-memory computing. *arXiv preprint arXiv:2603.26016*, 2026.

John C. Duchi, Michael I. Jordan, Martin J. Wainwright, and Andre Wibisono. Optimal rates for zero-order convex optimization: The power of two function evaluations. *IEEE Transactions on Information Theory*, 61(5):2788–2806, 2015. doi: 10.1109/TIT.2015.2409256.

Tanmay Gautam, Youngsuk Park, Hao Zhou, Parameswaran Raman, and Wooseok Ha. Variance-reduced zeroth-order methods for fine-tuning language models. In *International Conference on Machine Learning (ICML)*, volume 235 of *Proceedings of Machine Learning Research*, pp. 15180–15208, 2024.

Tayfun Gokmen. Enabling training of neural networks on noisy hardware. *Frontiers in Artificial Intelligence*, 4:699148, 2021. doi: 10.3389/frai.2021.699148.

Tayfun Gokmen and Wilfried Haensch. Algorithm for training neural networks on resistive device arrays. *Frontiers in Neuroscience*, 14:103, 2020. doi: 10.3389/fnins.2020.00103.

Tayfun Gokmen and Yurii Vlasov. Acceleration of deep neural network training with resistive cross-point devices: Design considerations. *Frontiers in Neuroscience*, 10:333, 2016. doi: 10.3389/fnins.2016.00333.

Jiaqi Gu, Zheng Zhao, Chenghao Feng, Wuxi Li, Ray T. Chen, and David Z. Pan. FLOPS: EFficient on-chip learning for OPtical neural networks through stochastic zeroth-order optimization. In *Design Automation Conference (DAC)*, pp. 1–6, 2020. doi: 10.1109/DAC18072.2020.9218593.

Jiaqi Gu, Hanqing Zhu, Chenghao Feng, Zixuan Jiang, Ray T. Chen, and David Z. Pan. L2ight: Enabling on-chip learning for optical neural networks via efficient in-situ subspace optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, pp. 8649–8661, 2021.

Maor Ivgi, Oliver Hinder, and Yair Carmon. DoG is SGD's best friend: A parameter-free dynamic step size schedule. In *International Conference on Machine Learning (ICML)*, volume 202 of *Proceedings of Machine Learning Research*, pp. 14465–14499, 2023.

Kevin G. Jamieson, Robert D. Nowak, and Benjamin Recht. Query complexity of derivative-free optimization. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 25, 2012.

An T. Le. Training non-differentiable networks via optimal transport. *Transactions on Machine Learning Research*, 2026. doi: 10.48550/arXiv.2605.01928.

An T. Le, Georgia Chalvatzaki, Armin Biess, and Jan Peters. Accelerating motion planning via optimal transport. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, pp. 78453–78482, 2023. doi: 10.52202/075280-3430.

Zhiyuan Li, Sadhika Malladi, and Sanjeev Arora. On the validity of modeling SGD with stochastic differential equations (SDEs). In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 34, pp. 12712–12725, 2021.

Wei Lin, Qingyu Song, and Hong Xu. The multi-query paradox in zeroth-order optimization. *arXiv preprint arXiv:2509.15552*, 2025.

Sijia Liu, Yicheng Lang, Soumyadeep Pal, Changsheng Wang, Yancheng Huang, Chongyu Fan, James Diffenderfer, Bhavya Kailkhura, and Yihua Zhang. Position: Zeroth-order optimization in deep learning is underexplored, not underpowered. In *International Conference on Machine Learning (ICML)*, 2026. doi: 10.48550/arXiv.2605.15622.

Yong Liu, Zirui Zhu, Chaoyu Gong, Minhao Cheng, Cho-Jui Hsieh, and Yang You. Sparse MeZO: Less parameters for better performance in zeroth-order LLM fine-tuning. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, pp. 21139–21169, 2025. doi: 10.52202/085713-0712.

Haotian Lu, Sanmitra Banerjee, and Jiaqi Gu. DOCTOR: Dynamic on-chip temporal variation remediation toward self-corrected photonic tensor accelerators. *Journal of Lightwave Technology*, 42(22):8042–8050, 2024. doi: 10.1109/JLT.2024.3409208.

Sadhika Malladi, Tianyu Gao, Eshaan Nichani, Alex Damian, Jason D. Lee, Danqi Chen, and Sanjeev Arora. Fine-tuning language models with just forward passes. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, pp. 53038–53075, 2023. doi: 10.52202/075280-2308.

Konstantin Mishchenko and Aaron Defazio. Prodigy: An expeditiously adaptive parameter-free learner. In *International Conference on Machine Learning (ICML)*, volume 235 of *Proceedings of Machine Learning Research*, pp. 35779–35804, 2024.

Yurii Nesterov and Vladimir Spokoiny. Random gradient-free minimization of convex functions. *Foundations of Computational Mathematics*, 17(2):527–566, 2017. doi: 10.1007/s10208-015-9296-2.

Malte J. Rasch, Charles Mackin, Manuel Le Gallo, An Chen, Andrea Fasoli, Frédéric Odermatt, Ning Li, S. R. Nandakumar, Pritish Narayanan, Hsinyu Tsai, Geoffrey W. Burr, Abu Sebastian, and Vijay Narayanan. Hardware-aware training for large-scale and diverse deep learning inference workloads using in-memory computing-based accelerators. *Nature Communications*, 14:5282, 2023. doi: 10.1038/s41467-023-40770-4.

Malte J. Rasch, Fabio Carta, Omobayode Fagbohungbe, and Tayfun Gokmen. Fast and robust analog in-memory deep neural network training. *Nature Communications*, 15:7133, 2024. doi: 10.1038/s41467-024-51221-z.

Kunjie Ren and Luo Luo. A parameter-free and near-optimal zeroth-order algorithm for stochastic convex optimization. In *International Conference on Machine Learning (ICML)*, volume 267 of *Proceedings of Machine Learning Research*, pp. 51450–51470, 2025.

Zhaolin Ren, Yujie Tang, and Na Li. Escaping saddle points in zeroth-order optimization: The power of two-point estimators. In *International Conference on Machine Learning (ICML)*, volume 202 of *Proceedings of Machine Learning Research*, pp. 28914–28975, 2023.

Prabhu Teja Sivaprasad, Florian Mai, Thijs Vogels, Martin Jaggi, and François Fleuret. Optimizer benchmarking needs to account for hyperparameter tuning. In *International Conference on Machine Learning (ICML)*, volume 119 of *Proceedings of Machine Learning Research*, pp. 9036–9045, 2020.

Minhak Song, Liang Zhang, Bingcong Li, Niao He, Michael Muehlebach, and Sewoong Oh. Zeroth-order optimization at the edge of stability. In *International Conference on Machine Learning (ICML)*, 2026. doi: 10.48550/arXiv.2604.14669.

James C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, 37(3):332–341, 1992. doi: 10.1109/9.119632.

James C. Spall. Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3):817–823, 1998. doi: 10.1109/7.705889.

Alexandra Suvorikova, Igor Pavlov, Artem Vasin, Georgii Bychkov, Anastasia Antsiferova, Darina Dvinskikh, and Alexander Gasnikov. Wall-clock complexity for zeroth-order optimization with tunable oracle fidelity. *arXiv preprint arXiv:2605.31346*, 2026.

Shuqi Wang, Zhengwu Liu, Chenchen Ding, Chen Zhang, Taiqiang Wu, Jiajun Zhou, and Ngai Wong. NoiseZO: RRAM noise-driven zeroth-order optimization for efficient forward-only training. In *Design Automation Conference (DAC)*, pp. 1–7, 2025. doi: 10.1109/DAC63849.2025.11132557.

Lei Wu, Chao Ma, and Weinan E. How SGD selects the global minima in over-parameterized learning: A dynamical stability perspective. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 31, 2018.

Zhaoxian Wu, Tayfun Gokmen, Malte J. Rasch, and Tianyi Chen. Towards exact gradient-based training on analog in-memory computing. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 37, pp. 37264–37304, 2024. doi: 10.52202/079017-1176.

Zhaoxian Wu, Quan Xiao, Tayfun Gokmen, Omobayode Fagbohungbe, and Tianyi Chen. Analog in-memory training on general non-ideal resistive elements: The impact of response functions. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, pp. 105439–105492, 2025. doi: 10.52202/085713-3522.

Zeke Xie, Issei Sato, and Masashi Sugiyama. A diffusion theory for deep learning dynamics: Stochastic gradient descent exponentially favors flat minima. In *International Conference on Learning Representations (ICLR)*, 2021.

A Further related work

Analog and photonic adaptation. Analog in-situ training uses backward signals and specialized updates, from rank-update SGD (Gokmen & Vlasov, 2016) to Tiki-Taka and its successors (Gokmen & Haensch, 2020; Gokmen, 2021; Rasch et al., 2024), with theory developed by Wu et al. (2024; 2025). Hardware-aware training improves robustness to modeled non-idealities offline (Rasch et al., 2023); we study residual degradation after deployment. Calibration is another relevant comparator: VeRA+ (Dong et al., 2026) compensates RRAM drift digitally without on-chip retraining, and DOCTOR (Lu et al., 2024) studies training-free photonic drift calibration. Such alternatives should be assessed with their measurement costs.

NoiseZO (Wang et al., 2025) uses device write noise to generate perturbations for forward-only RRAM training. Our focus is the usable step under independent readout noise. FLOPS (Gu et al., 2020) and single-chip forward-only training (Bandyopadhyay et al., 2024) demonstrate photonic learning. L2ight (Gu et al., 2021) combines calibration and subspace learning with backward signals; Ashtiani et al. (2026) demonstrate on-chip backpropagation. These systems differ from our phase-mode simulator, which has no automatic differentiation through its phase-to-unitary conversion. The restriction belongs to the simulator rather than to photonic hardware in general.

Zeroth-order estimators and oracle assumptions. SPSA (Spall, 1992) estimates a direction from two evaluations at any dimension; Spall (1998) gives practical settings that fix the perturbation size from the measurement-noise standard deviation and the gain sequence from the intended early-iteration displacement and the iteration budget. Our rule shares that goal of avoiding target search but uses a whole-run displacement model with coefficients measured on donor devices. Its risk bound is conditional on that model, not a guarantee for the simulator. MeZO (Malladi et al., 2023) implements in-place optimization at scale; Sparse-MeZO (Liu et al., 2025) restricts the perturbed subset, and MeZO-SVRG (Gautam et al., 2024) spends additional queries on variance reduction. Probe settings depend on the tuning mode: MeZO uses 10^{-3} for full-parameter tuning and RoBERTa LoRA, 10^{-2} for OPT LoRA, and 10^{-1} for prefix tuning (Tables 15–16). Sparse-MeZO also reports 10^{-3}. Our initial probe amplitude, including POEM's initial radius, was 10^{-3}; POEM's published experiments use 10^{-2}.

The two-point guarantees of Duchi et al. (2015) use a convex shared-realization oracle, whereas our paired queries have independent read errors. The dimension-dependent bounds of Nesterov & Spokoiny (2017), noisy-evaluation complexity of Jamieson et al. (2012), and state-dependent-noise analysis of Ren et al. (2023) also require their stated assumptions. None supplies a direct guarantee for our nonconvex simulator. Lin et al. (2025) analyze limitations of simple averaging under convex stochastic assumptions; Appendix R tests recovery separately. Suvorikova et al. (2026) study query-count and fidelity trade-offs under explicit oracle and cost models. These results motivate accounting for query costs without proving that a particular averaging choice is optimal here. Liu et al. (2026) argue that perceived ZO limitations partly reflect estimator-centric design, while Dbouk et al. (2026) study limits of coordinatewise adaptation and propose scalar-adaptive MEAZO. These approaches address complementary aspects of estimator design and adaptation.

Stability and stochastic escape. Song et al. (2026) extend the edge-of-stability perspective of Cohen et al. (2021) to Gaussian ZO with an exact oracle on a positive-semidefinite quadratic. Their mean-square stability threshold obeys $2/(\mathrm{tr}(H) + 2\lambda_{\max}) \le \eta^*_{\mathrm{ms}} \le 2/\mathrm{tr}(H)$, where $\lambda_{\max}$ is the largest Hessian eigenvalue. At our testbed reference state, $2/\mathrm{tr}(H) = 8.4 \times 10^{-3}$, while noisy runs at $m = 1$, $\sigma_{\mathrm{r}} = 0.272$ diverge near 10^{-5}. This difference motivates examining readout noise; the local quadratic approximation does not provide a global simulator stability limit. Their Appendix

B.3 includes direction averaging, whose threshold approaches the first-order $2/\lambda_{\max}$ limit. In our readout-dominated approximation, direction count also changes the number of steps affordable at fixed budget, producing the linear m factor in Eq. 4. That scaling cannot be extrapolated beyond its noise-dominated regime.

Wu et al. (2018) study local dynamical stability and minima accessible to SGD. Xie et al. (2021) analyze escape rates with minibatch-gradient noise, and Li et al. (2021) give testable necessary conditions for SGD diffusion approximations. Our model instead assumes an isotropic per-evaluation readout floor that persists near a minimum. We hold the minibatch size fixed and vary the forward-query budget B. Curvature may affect escape, but the RMS displacement criterion does not solve the first-passage problem.

Adaptive scales and selection cost. D-Adaptation, Prodigy and DoG adapt first-order scales online (Defazio & Mishchenko, 2023; Mishchenko & Defazio, 2024; Ivgi et al., 2023). POEM (Ren & Luo, 2025) adapts both ZO step size and smoothing radius from an initial-distance parameter. Its shared-randomness and projection assumptions differ from our adaptation (Appendix Q). AdaNAGED (Bystrov et al., 2026) uses a linear minimization oracle and is not evaluated here. Sivaprasad et al. (2020) and AlgoPerf (Dahl et al., 2023) explicitly account for configuration selection. Repeated trials on a deployed accelerator may be costly or unavailable; our donor rule studies that constraint. PolyStep (Le, 2026), a randomized direct-search method in the Sinkhorn Step family (Le et al., 2023), provides the fixed-configuration case study.

B Controlled crossbar evaluation and query accounting

B.1 Protocol and policy definitions

The controlled crossbar study uses three independently trained factory models, each with six conditions: readout-noise parameters $\{0.03, 0.1, 0.2\}$ crossed with drift durations of one and two days. Every factory–condition stratum has two donor and two target devices, giving 36 donors and 36 targets. The roles and device identities are disjoint; donors and targets share the three factory clusters. Every recovery configuration for a given device starts from the same serialized damaged state, including the compensated readout state. MNIST uses $48{,}000$ training, $6{,}000$ development, $6{,}000$ validation and $10{,}000$ test examples. Recovery writes the output layer ($D = 1{,}290$); loss queries use training minibatches of 512 examples. The primary experiment fixes $\varepsilon = 0.1$, $m = 1$, $\delta = 0.1$, optimizer seed 42 and $B = 30{,}000$ optimization FEE.

Calibration and selection. Each donor supplies 128 noisy losses on one frozen 512-example batch, with sample standard deviation $\sigma_{\mathrm{r},u}$ (divisor 127), and clean-loss measurements on 24 fixed isotropic rays. Clean bisection estimates the chance-loss crossing at $\log 10$; R_u is the arithmetic mean of all 24 crossing midpoints. An unsupported ray or zero noise scale makes the donor coefficient unsupported; the selector does not silently drop such donors. Clean readout remains a simulator oracle. These measurements do not establish a physical noisy-only calibration procedure.

For each declared pool, the donor rule computes $\ell_u = 2R_u/(\sigma_{\mathrm{r},u}\sqrt{D_u B})$ and takes the median of these levels, rather than a ratio of separately pooled components. It selects the largest candidate with $\eta_{\mathrm{RMS}} \le \sqrt{\delta}\,\varepsilon m\ \mathrm{median}_u\,\ell_u$, or abstains when the pool is unsupported or no candidate qualifies. Cosine decay to $0.01\eta_0$ uses the exact finite-schedule RMS over $T = B/(2m) = 15{,}000$ updates. The coarse ladder is $\{10^{-3}, 10^{-2}, 10^{-1}\}$. The fine extension uses the 21 rates $\{10^{k/4} : k = -24, \ldots, -4\}$.

The substrate pool contains all 36 same-substrate donors. The cohort extension instead uses the two donors from the target's factory and noise–drift condition. Changing the pool, the rate ladder, or both defines separately reported extensions. All donor measurements and transfer choices are frozen before any target recovery. Donor-screened global and cohort-fixed controls select a rate from the donor recovery grid, using clean development accuracy and the same declared global or factory–cohort pools. A candidate must complete the budget without numerical, monitoring or selection errors, or a sampled threshold event, on every donor in that pool; exact accuracy ties choose the smaller rate. Transferred global and condition-matched fixed controls use frozen noisy-development

choices made on a separate calibration population. The fixed default uses $\eta_0 = 0.001$. No target recovery or test outcome selects these transfer settings.

Recovery, monitoring and test access. Within a run, selection compares only the damaged baseline and at most one distinct terminal state, each with five noisy validation draws. The terminal state is retained only for a strict increase in mean validation accuracy; a tie, numerical failure or terminal-selection error returns the baseline. Monitoring is a separate operation: five noisy validation draws at the initial state and every $1{,}000$ optimization FEE detect a sampled event when their mean accuracy is at most 0.15. These monitoring checkpoints are not additional candidates for state selection. An earlier sampled event remains recorded even if the terminal state or the subsequently selected baseline scores well.

Targets execute the full primary 21-rate grid for controlled comparisons. The target-search reference chooses among eligible runs using only clean development accuracy of their validation-selected states, with smaller-rate tie breaking. It is a paid search comparator, not a test-set grid maximum. The complete final-evaluation roster is then frozen before any test read. Each unique selected state receives one clean test evaluation and 20 noisy test evaluations; identical state references shared by policies are evaluated once. Transfer settings are fixed before target recovery, all cohorts are included, and failed candidates are not retried. At secondary $\varepsilon = 0.03$, only the unique cells required by the four fixed sensitivity policies are run.

B.2 All policy outcomes

Tables 1 and 2 retain all 15 policy views, including the baseline. Means weight factories equally, conditions equally within factory and targets equally within each stratum. The three factory clusters do not constitute a powered noninferiority study: the 0.10-pp accuracy margin is a descriptive criterion. Clean test accuracy is primary; noisy test accuracy measures a different deployment endpoint. Abstentions are evaluated at the same damaged baseline and remain in the means.

The coarse/substrate rule and the fine/substrate extension both choose $\eta_0 = 0.001$ on all 36 targets, sharing the fixed default's selected states and test scores. Calibration therefore adds $63{,}482$ forwards (5.50% over that fixed control) without changing recovery. The coarse/cohort extension abstains on 14 targets; its smaller operational count reflects 22 recovery attempts, not cheaper execution of 36 recoveries. The fine/cohort extension trails primary target search by 0.399167 pp, exceeding the 0.10-pp criterion despite lower operational cost. Stronger global and cohort-fixed controls remain in the comparison.

Across all 838 target runs (756 primary-grid cells and 82 unique secondary cells), there are 173 numerical failures, 115 sampled threshold events, 215 runs in their union and 239 baseline reversions. These categories overlap: reversion does not erase a trajectory event, and a failed run remains paid. Every selected recovery policy has zero recorded numerical failures or sampled events among its chosen runs; fixed and searched controls share this result. The coarse/substrate rule, fine/substrate extension and fixed default each have six baseline reversions. Zero selected events is not an advantage unique to donor measurement and does not establish an unseen-device failure probability. The 15 policies yield 540 target-policy outcomes. Tables 1 and 2 include every policy, with abstentions evaluated at baseline.

B.3 Operational costs and total experimental expenditure

One root-model forward is a forward pass of the complete network on one minibatch. The policy cost totals count each required calibration, screening or recovery run once. They include paid monitoring, validation and development scoring. One complete selected recovery per target costs $36 \times 32{,}052 = 1{,}153{,}872$ root-model forwards. The 36 donor calibrations cost $63{,}482$, so the nonabstaining donor policies require $1{,}217{,}354$ operational forwards. The donor screening grid costs $19{,}038{,}821$; adding selected target recoveries gives $20{,}192{,}693$ for either donor-screened fixed policy. If its fixed setting is already available, the alternative operational view is $1{,}153{,}872$. This is a different acquisition scenario, not a reason to remove screening from the newly acquired fixed-control comparison.

The primary target-search view costs $19{,}080{,}779$, including unsuccessful grid cells and the selected recovery already contained in that grid; it does not charge the selected recovery twice. Transferred

Table 1: Controlled crossbar evaluation, primary $\varepsilon = 0.1$: all ten recovery policies and the damaged baseline on the same 36 targets. Accuracy means are in percent; means and costs are rounded to three decimal places. *Abst.* counts targets on which no recovery is attempted. Operational cost is in millions of root-model forwards and includes the prerequisites for that policy. Donor-screened fixed controls include donor screening; transferred fixed controls exclude their separately retained sunk acquisition costs. Final research test evaluation and common factory preparation are itemized separately below.

Policy	Clean	Noisy	Abst.	Op. forwards (10^6)
Damaged baseline	93.349	86.454	0	0.000
Fixed default	94.521	89.114	0	1.154
Coarse/substrate	94.521	89.114	0	1.217
Fine/substrate extension	94.521	89.114	0	1.217
Coarse/cohort extension	94.601	88.319	14	0.769
Fine/cohort extension	95.075	92.312	0	1.217
Transferred global fixed	95.343	92.876	0	1.154
Transferred cohort fixed	95.391	92.876	0	1.154
Donor-screened global fixed	95.406	92.685	0	20.193
Donor-screened cohort fixed	95.436	91.904	0	20.193
Target search	95.474	91.714	0	19.081

Table 2: Controlled crossbar evaluation, secondary $\varepsilon = 0.03$: all four frozen sensitivity policies on the same 36 targets. Column definitions and costing conventions match Table 1; means and costs are rounded to three decimal places. This sparse sensitivity evaluation has no corresponding full target-search arm at $\varepsilon = 0.03$; its outcomes do not replace the primary $\varepsilon = 0.1$ comparison.

Policy	Clean	Noisy	Abst.	Op. forwards (10^6)
Coarse/cohort extension	94.298	87.458	24	0.448
Fine/cohort extension	95.833	92.841	0	1.217
Transferred global fixed	95.670	92.612	0	1.154
Transferred cohort fixed	95.837	92.895	0	1.154

fixed settings likewise cost 1,153,872 in current target operation, while their calibration-population grids, calibration and preparation remain sunk prerequisites. They were not free to discover. Acquisition on the separate calibration population includes 9,497,315 additional forwards and a reused $\varepsilon = 0.03$ grid costing 8,313,512 forwards. Preparation and earlier noisy calibration are not fully costed. These overlapping acquisition costs are separate from controlled-evaluation expenditure.

Each policy view requires $15{,}120 = 36 \times (1 + 20) \times 20$ final research-test forwards: 20 test minibatches per readout draw. These are separate from operational cost. Shared selected-state references reduce the actual campaign's final-test total to 132,300; summing policy views would count shared evaluations repeatedly. Table 3 gives the total experimental expenditure. The campaign total of 40,945,236 already includes the separately itemized 1,590 factory/preparation forwards. Direct compensation probes have separate tile-call and input-vector units and are not added to root-model forwards. None of these software counts establishes physical latency, energy, or a minimum-cost hardware implementation.

B.4 Measurement diagnostics and cross-substrate donor evaluation

The separate noisy-radius replication reuses 18 exposed development devices with optimizer seeds 101, 202 and 303. At its primary $\varepsilon = 0.03$, the interpolated noisy mean-radius policy trails cohort-fixed noisy development accuracy by 0.30463 pp overall and 0.76111 pp in high-noise conditions. The clean-bisection control trails cohort-fixed by 0.04198 pp overall. These are additional optimizer-seed comparisons on development devices, not held-out target evaluation, and clean-oracle performance does not establish noisy-only measurement feasibility.

Table 3: Unique root-model forwards executed in the controlled crossbar evaluation. The target entry includes the full primary grid and sparse secondary cells. These executed costs are distinct from overlapping deployment policy views; failed attempts remain included.

Executed stage	Root-model forwards
Factory training and preparation	1,590
Donor calibration	63,482
Donor recovery grid	19,038,821
Target grid and sparse sensitivity	21,709,043
Deduplicated final test evaluation	132,300
Total	40,945,236

A further retrospective donor-only diagnostic compares optimistic eligible coarse-grid and fine-grid recovery envelopes: their mean gap is 0.859722 pp, with $26/36$ development devices above 0.10 pp. The envelope is a candidate-resolution diagnostic, not a deployable selector or a new target-policy evaluation. Cohort-matching and measurement-masking analyses use these same devices and do not isolate measurement value over condition-matched fixed settings.

For completeness, Table 8 reports the retrospective donor evaluation on five crossbar and five photonic devices. Its 20 selections share devices, donor pools and prior probe screens. No selected setting escapes, but the photonic deficit averages 4.26 pp and the worst averaging deficit is 11.47 pp. The target is excluded from its donor coefficient pool, while the probe screens already include those devices. All devices receive the same setting within each estimator class. This comparison therefore does not identify added accuracy from donor coefficients over transferring those fixed settings. Its 15-configuration test-screen maximum is different from the development-selected search reference used in the controlled crossbar evaluation. Appendix I retains threshold sensitivities and abstentions.

C Experimental details

This section describes the mechanism and retrospective cross-substrate benchmarks. The controlled crossbar study has separate data splits, selection rules and a complete operational cost breakdown, specified in Appendix B.

Analog deployment tasks. Both analog settings use MNIST with a fixed split of $54{,}000$ training and $6{,}000$ validation examples (split seed 12345), and the $10{,}000$-example test set. Optimization queries evaluate mean cross-entropy on minibatches of at most 512 training examples. The crossbar uses a 784–128–10 ReLU MLP and adapts its output layer; the photonic model pools inputs to 8×8 and converts a 64–64–10 ReLU MLP to phase-mode MZI layers. Offline pretraining selects the best test-scored epoch (five epochs for crossbar, eight for photonic). Thus recovery checkpoint selection uses validation, while pretraining and configuration selection also use the test set.

Cost accounting. PolyStep evaluates complete polytopes, leaving 0.85%–10.40% of the nominal budget unused in the compared settings. For each estimator and target, the donor rule spends one $30{,}000$-query recovery budget; the retrospective grid maximum uses 15 runs, totaling $450{,}000$ optimization queries. The selected grid run is already among those 15 runs. A further confirmation run would add its own cost. Both protocols also incur validation, checkpoint-scoring and test-evaluation queries, which are excluded from FEE. Donor noise measurement, clean-loss ray bisection and prior probe screens add shared calibration costs. Their complete totals are not recorded. For N targets, shared calibration contributes its total cost divided by N to each target's amortized cost; its omission prevents a measured break-even point or total-query speedup claim. Loss queries on minibatches and accuracy evaluation on full splits also process different numbers of examples. Thus configuration counts and optimization FEE should not be read as wall-clock, energy or complete device-query costs.

Device realizations. Some crossbar experiments use different damaged states under the same seed labels. At seeds $42/123/456$, the baseline grid, seed-42 screens and radius-selected POEM use

Table 4: Selection and evaluation stages for the retrospective analog benchmarks. Optimization FEE counts training-minibatch loss queries only. Shared calibration and total evaluation costs are not fully measured.

Stage	Data and selection criterion	In optimization FEE?
Pretraining checkpoint	Test accuracy across training epochs	No
Donor calibration	Repeated loss draws and clean-loss ray bisections	No; total not logged
Recovery checkpoint	Validation accuracy; damaged state included	No; only loss queries are counted
Configuration screen	Clean-readout test accuracy across runs	Each run's loss queries only
Final evaluation	Clean test accuracy or 20 noisy test draws	No

states with accuracies $91.00/94.45/91.93\%$. The fixed-radius comparisons, PolyStep probe–step searches and additional screening seeds use states with accuracies $88.54/94.85/91.30\%$. The corresponding offsets, $-2.46/+0.40/-0.63$ pp, are comparable to several method differences. Matching seed labels or subtracting baseline accuracy therefore does not establish pairing. The PolyStep configuration gains in Section 4.4 use the same damaged realizations and five shared seeds for both configurations.

For seed 42, the crossbar screen and coefficient measurements also use different realizations: $\sigma_{\mathrm{r}} = 0.142$ at the grid state and 0.160 at the calibration state. With $R = 5.83$, the latter changes the predicted η/ε boundary from 0.0215 to 0.0190; both values lie within the measured bracket. The source of the realization difference is unresolved. Photonic comparisons share chip-specific damaged-state accuracies. The simulator wraps phases before applying fabrication-noise scaling, making loss discontinuous at wrap boundaries.

Configurations in Table 10. The upper block uses ZO fixed step sizes with transferred noisy-readout probes (0.1 for $m = 1$, 0.01 for averaging; deterministic probe 0.001), PolyStep's two-axis screen winners, and analog POEM's bracketed $r_\varepsilon = 0.1$ (spiking: initial radius). It uses 3–10 crossbar seeds, ten photonic chips or three spiking seeds. These selection protocols differ; Table 5 instead compares screen-selected ZO with PolyStep's fixed configuration.

Replicate sets in Tables 10 and 5. Photonic means normally use ten chips at one optimizer seed. Thermal $m = 1$ and PolyStep in Table 5 instead pool 30 chip–seed observations (seeds 42–44). Crossbar means use seed families: the reference damaged-state accuracy is 93.10% over five seeds, 93.17% over POEM's ten, and 92.46% over the three-seed subset for the exact-gradient baseline and fixed PolyStep configuration. The two crossbar reference rows use the same state. PolyStep's screen-selected entries in Table 10 use the second set of damaged realizations (reference mean 92.57%). The read-noise averaging entry (94.63%) combines realizations at seeds $123/456$, with $\varepsilon = 10^{-2}$ and the fixed step size. Thermal PolyStep at the fixed configuration is 85.68% over ten chips and 85.46% in the 30-observation pool. A fixed affine calibration reaches 77.62% on the deterministic photonic setting without a configuration search. On the crossbar Monte Carlo deployment metric, PolyStep at its fixed configuration has the lowest within-run readout SD among the compared methods (0.21 against 0.22, 0.37 and 0.45) and the smallest drop from selected-checkpoint to deployed accuracy (1.66 pp against 2.08, 3.78 and 4.24). This readout robustness accompanies its $+1.87$ pp deployment lead in Appendix G.1 but does not identify its cause.

Testbed accuracy is a running test maximum unless explicitly labeled final and unselected. A validation-selection control finds up to 0.25 pp inflation for PolyStep in the tested conditions; it does not bound bias in pairwise method differences.

Data availability. Per-run results and stored coefficients support the main accuracy tables and boundary analyses. The testbed radius is stored for seed 42; radii for other seeds and new deployment coefficients require simulator measurements. Individual outcomes from the 13-configuration

Table 5: Mean clean-readout test accuracy (%) of the selected checkpoint. Noisy analog ZO configurations are selected by a designated-unit 5×3 (ε, η) screen; deterministic analog ZO uses a rate bracket at the fixed probe amplitude. The photonic $m = 1$ winner is selected on chip 0 and transferred across chips. Spiking ZO uses its clean-readout 5×3 screen ($m > 1$: $\varepsilon = 10$ least significant bits, LSB). Analog direction averaging uses $m = 250$; the spiking direction count follows its query-matched protocol. PolyStep uses its fixed configuration; analog POEM uses $r_\varepsilon = 10^{-1}$ selected on three units, and spiking POEM uses its initial radius. Thermal $m = 1$/PolyStep pool 30 chip–seed observations; other replicate sets and realization differences are given above. "No adapt." is a reference damaged-state mean. Bold denotes the largest non-oracle raw mean, without a statistical-superiority claim.

Setting	no adapt.	oracle	m=1	m>1	POEM	PolyStep
Crossbar, deterministic	93.10	96.46	96.50	**96.56**	96.42	94.60
Crossbar, read noise	93.10	96.11	95.31	**95.88**	95.76	93.80
Photonic, deterministic	61.84	n/a	88.20	88.48	**89.64**	85.43
Photonic, thermal	61.84	n/a	**90.20**	87.27	89.38	85.46
Spiking SoC	27.29	n/a	**38.13**	35.44	27.27	28.11

search used to choose the fixed photonic configuration are unavailable. Some supplementary statistical comparisons are available only as summaries, limiting independent reanalysis (Appendix J).

Comparable final screening on the thermal mesh. Each method receives a 5×3 screen on chip 0 at thermal noise 0.05 and nominal optimization budget $30{,}000$ per configuration, followed by confirmation on ten chips. The final-screen totals are $450{,}000$ FEE per ZO method and $446{,}175$ for PolyStep ($29{,}745$ per run after step rounding). Table 6 uses optimizer seed 42, fabrication noise 0.02, and the final selected checkpoints. Methods share the recorded undamaged and damaged accuracies on each chip; identity of the underlying parameter arrays has not been verified.

PolyStep uses projection rank 2 and one transport solve per three update steps; averaging uses 250 sequential microsteps (Appendix E). PolyStep minus single-direction ZO is -2.390 pp ($0/10$ positive pairs), and minus averaging is $+0.560$ pp ($8/10$). Excluding the screening chip gives $-2.323/+0.642$ pp ($0/9$, $8/9$); also excluding chips 2 and 8 used to select the fixed configuration gives $-2.184/+0.874$ pp ($0/7$, $7/7$). These exclusions are retrospective sensitivity checks, not prospective holdouts. Fixed-configuration selection, grid design, amortization searches, and validation/test queries are excluded from the totals. Averaging's winner is at its upper tested rate. Thus the comparison concerns observed screen winners, without establishing global optima, equal total calibration cost, or a corresponding comparison under noisy deployment readout.

Sensitivity of the thermal-mesh comparison. The thermal comparison at fixed probe amplitude favors PolyStep's fixed configuration over the best ZO baseline at pinned $\varepsilon = 10^{-3}$ by $+10.13$ pp. Transferring the chip-0 probe–rate screen winner gives single-direction ZO $90.20 \pm 0.28\%$ against the fixed configuration's $85.46 \pm 1.17\%$ across ten chips and three optimizer seeds. PolyStep's difference is -4.73 pp, with ZO ahead on all 30 chip–seed observations ($t = 16.15$, $p = 5.9 \times 10^{-8}$ at the chip unit). This differs from the comparable final-screen evaluation in Section 4.4. Single-direction ZO leads by 4.23 pp on the seven chips excluded from both configuration screens, by 4.39 pp against PolyStep's best single tested setting, and by 3.83 pp against a per-chip oracle that

Table 6: Winners of shared 15-configuration thermal-mesh screens. Accuracy is mean ± sample SD of clean-readout test accuracy (%) on ten chips at seed 42. PolyStep reports its probe and step scale settings (fixed configuration: $(1, 10)$); ZO reports probe amplitude and initial rate.

Method	Probe setting	Step setting	Accuracy (%)
PolyStep	$\rho_\mathrm{p} = 0.3$	$\rho_\mathrm{s} = 10$	87.834 ± 0.779
Single-direction ZO	$\varepsilon = 0.1$	$\eta = 0.01$	90.224 ± 0.281
Direction averaging	$\varepsilon = 0.01$	$\eta = 0.1$	87.274 ± 0.546

takes the best of PolyStep's nine tested settings on each chip. With the step size held fixed and only the transferred probe radius $\varepsilon = 10^{-1}$, $m = 1$ reaches $88.28 \pm 0.42\%$ on the ten chips, 2.60 pp above PolyStep's fixed configuration ($t = 8.40$, $10/10$ chips). These comparisons use different replicate pools, so their differences cannot isolate the contributions of probe and step selection. On the noisy crossbar, PolyStep minus $m = 1$ changes from an unresolved $+0.74$ pp at the pinned probe to -1.51 pp after ZO screening ($0/5$ positive pairs, $p = 0.0033$, directional), and to $+0.57$ pp with PolyStep screening; the last comparison mixes damaged realizations, and baseline subtraction cannot restore pairing. Together with POEM against PolyStep's fixed configuration on the thermal mesh ($-10.72 \rightarrow +3.70$ pp between its initial and screened radius, Appendix Q), these are three comparisons whose sign changes after additional configuration selection. The POEM radius changes both initial probe and step scales, so its reversal cannot be assigned to the probe alone.

D EXACT REVERSION

Crossbar conductance targets and photonic phase settings are floating-point parameters without a state quantizer. Their exact reversion occurs when checkpoint selection retains the damaged baseline; equality alone does not show whether intermediate states moved. Instrumentation confirms movement but no candidate improvement at $\varepsilon = 10^{-3}$ on every tested noisy crossbar seed and thermal chip. At $\varepsilon = 10^{-1}$, six to seven candidates beat the baseline on $5/5$ crossbar seeds and $10/10$ photonic chips. Reversion is therefore specific to the initial probe setting in these comparisons.

The int8 spiking SoC stores a floating-point copy of each weight, called its shadow state, and quantizes it once per forward evaluation. Dither perturbs the initial shadow within a quantization bin. Updates smaller than one least significant bit (LSB) can therefore accumulate. For an unsaturated coordinate, let a_j, b_j be the initial shadow state's distances to its lower and upper rounding boundaries. Away from ties, its final quantized value is unchanged exactly when $-a_j < \Delta_j(T) < b_j$, where $\Delta_j(T) = -\sum_t \eta_t \hat{g}_{t,j}$. The symmetric condition $|\Delta_j(T)| < \rho_j = \min(a_j, b_j)$ is sufficient, not necessary. Dither reaches 0.499 LSB, making some ρ_j as small as 10^{-3} LSB; the observed 0.0017 LSB net shadow displacement must be interpreted against these directional distances. A per-step sub-LSB condition cannot determine the final written state.

E POLYSTEP

PolyStep (Le, 2026) is randomized direct search with an entropic optimal-transport (OT) barycenter, without a finite-difference gradient. On piecewise-constant objectives, either PolyStep or finite differences can obtain signal when probes cross region boundaries; neither is guaranteed a signal when all queried costs coincide.

For writable parameters $\theta \in \mathbb{R}^D$, one transport solve proceeds as follows. We use PolyStep's one-sided softmax solver, which computes the weights in step 4 in closed form without Sinkhorn iterations.

1. **Project.** A rank-r per-layer random projection P, with orthonormal columns where tall, defines $\theta = \theta_{\text{base}} + Pc$, $c \in \mathbb{R}^S$. Partition each layer's coordinates into particles of dimension $d = 8$ (padding the final particle where needed). Projection rank changes S, while the particle template remains fixed.
2. **Rotate.** Draw a uniform rotation $Q \in \mathrm{SO}(d)$ and rotate a mean-centered, unit-radius orthoplex with $V = 2d = 16$ vertices $u_j \in \{\pm e_1, \ldots, \pm e_d\}$.
3. **Probe.** For each particle $z \in \mathbb{R}^d$, evaluate the vertices on an interior scale ladder with mean $\bar{s} = 1/2$. To first order, $z_j = z + \bar{s}\rho_{\text{p}} Q u_j$. The cost is $V \sum_\ell \lceil S_\ell/d \rceil$ forward evaluations per solve, where S_ℓ is the number of projected coordinates in layer ℓ.
4. **Weight and update.** For costs C_j, form $w_j \propto \exp[-(C_j - \min_k C_k)/\gamma]$ and update $z \leftarrow z + \rho_{\text{s}} \sum_j w_j Q u_j$. Here γ is the OT temperature in cost units; ρ_{p} and ρ_{s} are the physical probe and step radii after applying the chosen schedules and scale factors.

Between transport solves, two additional momentum updates reuse the estimated direction without new loss queries. The momentum exponential moving average uses coefficient 0.7. For locally

smooth noiseless costs, a first-order expansion of the softmax weights gives

$$\Delta\theta \approx -\frac{\rho_{\mathrm{s}}\rho_{\mathrm{p}}}{\gamma k}\, g, \tag{6}$$

where g is the projected gradient and $k = d/\bar{s} = 16$. To see this, let $v_j = Qu_j$ and let h be the loss gradient with respect to one particle. The orthoplex satisfies $\sum_j v_j = 0$ and $V^{-1}\sum_j v_j v_j^\top = I_d/d$. At fixed $\gamma > 0$, the cost expansion $C_j = C_0 + \bar{s}\rho_{\mathrm{p}} h^\top v_j + o(\rho_{\mathrm{p}})$ gives

$$w_j = \frac{1}{V} - \frac{\bar{s}\rho_{\mathrm{p}}}{V\gamma} h^\top v_j + o(\rho_{\mathrm{p}}), \qquad \rho_{\mathrm{s}} \sum_j w_j v_j = -\frac{\bar{s}\rho_{\mathrm{s}}\rho_{\mathrm{p}}}{\gamma d} h + o(\rho_{\mathrm{p}}).$$

Mapping this displacement back through P gives Eq. 6 for fixed ρ_{s}. The product $\rho_{\mathrm{s}}\rho_{\mathrm{p}}/\gamma$ acts as an effective step size. The ratio ρ_{p}/γ sets linearized cost contrast, but is not a signal-to-noise ratio because readout error also scales with $1/\gamma$. Outside this approximation, the probe radius changes sampled regions and curvature.

We use two fixed configurations across substrate classes, selected offline. The rank-8 orthoplex with cosine schedules on $(\gamma, \rho_{\mathrm{s}}, \rho_{\mathrm{p}})$ serves the crossbar, testbed and spiking SoC; the latter uses a unit-standard-deviation parameter-scale adapter. A flat-schedule SoC control differs by $+0.23$ pp ($p = 0.11$). The photonic fixed configuration uses rank 2 and $10\times$ the effective step. It has $S = 1{,}138$ and costs 2,288 evaluations per solve, allowing 13 solves at the matched budget, versus $S = 3{,}200$, 6,400 evaluations and four solves at rank 8. Its offline screen used three of the ten confirmation chips. The seven remaining chips score higher descriptively, which does not exclude selection bias or population differences. The shared calibration cost is separate from target-device recovery. Additional searches vary the effective step over $10^4\times$ and the probe scale separately.

E.1 Component ablations

These controls vary PolyStep's components at matched FEE and, where specified, matched vertex count, on the mechanism testbed and the crossbar (Appendix G). All configurations retain an effective step size. Testbed accuracies are running maxima, so a pre-divergence peak must be distinguished from successful final recovery.

Geometry. At $V = 16$, orthoplex, cube and centered-random templates reach respectively $95.4/94.3/91.3$, $96.2/95.2/90.2$, and $95.2/94.1/90.8\%$ at $\sigma = 0/0.1/0.2$. The difference in degradation slope between structured and random templates is 0.30 pp, below three times the variation across seeds (0.58 pp). Orthoplex minus cube is -0.23 pp ($p = 0.11$); orthoplex minus centered-random is $+0.28$ pp ($p = 0.014$ at the seed unit), an exploratory comparison that does not survive correction. At $\sigma = 0.2$, simplex-$d4$ ($V = 5$) and orthoplex-$d4$ ($V = 8$) diverge; their $85.8/88.9\%$ entries are pre-divergence peaks. Simplex-$d8$ ($V = 9$) remains at 90.0%. The approximately 5 pp vertex-count ordering at this noise level therefore includes failed runs.

Centering and weighting. The uncentered random template at the same vertex count deteriorates $87.7 \to 15.5 \to 12.4\%$ across the three noise levels, with paired difference -54.85 pp ($p = 3.1 \times 10^{-5}$), significant after Bonferroni correction. Replacing softmax-OT weights with top-k or greedy cost-ranked selection yields 12.70–23.28% on the clean task versus 95.4% for softmax, across the tested step-radius range; each solver's best point is confirmed on three seeds. This supports the weighting within the tested schedule envelope, without excluding other ranked-solver configurations. Convex weights bound each particle's displacement by ρ_{s} for unit vertices, but the probes use a separate radius, and greedy/top-k weights share this bound. Bounded displacement alone does not explain softmax's advantage.

Schedule-only control. Plain centered averaging uses PolyStep's schedules with the measured template constant $k = 16$. With this scale matching, it diverges on $3/3$ noisy testbed seeds and exactly reverts on $3/3$ noisy crossbar seeds, where PolyStep recovers. Deterministically it beats PolyStep (97.34 ± 0.04 versus $95.36 \pm 0.15\%$ on the testbed) and gains 1.34 pp above its own crossbar baseline instead of reverting. The difference between the control and PolyStep therefore depends on readout noise.

F TWO-POINT ESTIMATION AND DESCENT

We derive the mean, covariance and descent condition used in Section 3.1. Fix $\theta \in \mathbb{R}^D$, $\varepsilon > 0$ and a minibatch. Assume f is continuously differentiable and $f, \nabla f$ have at most polynomial growth. Let $u_i \sim \mathcal{N}(0, I_D)$ be independent, and let $\xi_i^{\pm}$ be mutually independent errors, independent of the directions, with mean zero and variance σ_{r}^2. The estimator is

$$\hat{g}_m(\theta) = \frac{1}{m} \sum_{i=1}^{m} \frac{f(\theta + \varepsilon u_i) - f(\theta - \varepsilon u_i) + \xi_i^+ - \xi_i^-}{2\varepsilon} u_i. \tag{7}$$

Mean and smoothing bias. Write $\phi(u) = (2\pi)^{-D/2} e^{-\|u\|^2/2}$ for the Gaussian density. Symmetry under $u \mapsto -u$ removes the negative probe, and $\partial_j \phi = -u_j \phi$ gives, for each coordinate j,

$$\frac{1}{\varepsilon} \int f(\theta + \varepsilon u) u_j \phi(u)\, du = -\frac{1}{\varepsilon} \int f(\theta + \varepsilon u) \partial_j \phi(u)\, du = \int \partial_j f(\theta + \varepsilon u) \phi(u)\, du.$$

Polynomial growth makes the boundary term vanish and permits differentiation under the integral. Therefore, for $f_\varepsilon(\theta) = \mathbb{E}_u f(\theta + \varepsilon u)$, $\mu := \mathbb{E}\hat{g}_m = \nabla f_\varepsilon(\theta)$. If ∇f is L-Lipschitz, then

$$\|\mu - \nabla f(\theta)\| \le \mathbb{E}\|\nabla f(\theta + \varepsilon u) - \nabla f(\theta)\| \le L\varepsilon \mathbb{E}\|u\| \le L\varepsilon \sqrt{\mathbb{E}\|u\|^2} = L\varepsilon\sqrt{D}.$$

Thus larger probes reduce explicit readout amplification but can increase smoothing bias. In particular, writing $G = \|\nabla f(\theta)\|$, the alignment obeys $S := \nabla f(\theta)^\top \mu \ge G(G - L\varepsilon\sqrt{D})$. The condition $G > L\varepsilon\sqrt{D}$ is sufficient for $S > 0$, but is not required.

Covariance. Write $z_i = [f(\theta + \varepsilon u_i) - f(\theta - \varepsilon u_i)] u_i/(2\varepsilon)$ and $n_i = (\xi_i^+ - \xi_i^-) u_i/(2\varepsilon)$. Conditional centering gives $\mathbb{E}[n_i \mid u_i] = 0$, hence $\mathbb{E}[(z_i - \mu) n_i^\top] = 0$. Since $\mathbb{E}(\xi_i^+ - \xi_i^-)^2 = 2\sigma_{\mathrm{r}}^2$ and $\mathbb{E}[u_i u_i^\top] = I_D$,

$$\mathbb{E}[n_i n_i^\top] = \frac{2\sigma_{\mathrm{r}}^2}{4\varepsilon^2} \mathbb{E}[u_i u_i^\top] = \frac{\sigma_{\mathrm{r}}^2}{2\varepsilon^2} I_D.$$

Let $\Sigma_{\mathrm{search}} = \mathrm{Cov}(z_i)$ and $A_{\mathrm{search}} = \operatorname{tr} \Sigma_{\mathrm{search}}$. Independence across i yields

$$\mathrm{Cov}(\hat{g}_m) = \frac{1}{m} \left(\Sigma_{\mathrm{search}} + \frac{\sigma_{\mathrm{r}}^2}{2\varepsilon^2} I_D \right), \qquad V_m = \frac{1}{m} \left(A_{\mathrm{search}} + \frac{D\sigma_{\mathrm{r}}^2}{2\varepsilon^2} \right). \tag{8}$$

For $A_{\mathrm{search}} > 0$, the inflation $V_m(\sigma_{\mathrm{r}})/V_m(0) = 1 + D\sigma_{\mathrm{r}}^2/(2\varepsilon^2 A_{\mathrm{search}})$ is independent of m. The measured variance exponents are -0.97 to -1.02 in 13 noise-level series spanning 3.5 decades of m. Correlated readout errors need not obey this identity.

Descent. For an L-Lipschitz gradient and any displacement h, integration along the segment gives

$$f(\theta + h) - f(\theta) - \nabla f(\theta)^\top h = \int_0^1 [\nabla f(\theta + sh) - \nabla f(\theta)]^\top h\, ds \le \frac{L}{2} \|h\|^2.$$

Set $h = -\eta \hat{g}_m$, $M = \|\mu\|^2$ and $A = A_{\mathrm{search}} + D\sigma_{\mathrm{r}}^2/(2\varepsilon^2)$. Since $\mathbb{E}\|\hat{g}_m\|^2 = M + A/m$,

$$\mathbb{E}[f(\theta - \eta \hat{g}_m)] - f(\theta) \le -\eta S + \frac{L\eta^2}{2}(M + A/m).$$

When $S > 0$, the right-hand side is negative exactly for $0 < \eta < 2mS/[L(mM + A)]$, proving Eq. 2. It tends to $2S/(LM)$ as $m \to \infty$, reducing to $2/L$ when $S = M$. This sufficient condition does not locate a failure boundary.

For a Hessian H at θ, expanding $\hat{g}_m = \mu + (\hat{g}_m - \mu)$ instead gives $\mathbb{E}[\hat{g}_m^\top H \hat{g}_m] = \mu^\top H \mu + \operatorname{tr}(H \operatorname{Cov} \hat{g}_m)$. Substitution of Eq. 8 gives the trace diagnostic in the main text; it uses curvature at one state rather than a bound along the update. These statements concern the idealized scalar-loss oracle. Fresh weight perturbations in the simulator need not produce unbiased readout noise with constant variance.

G TESTBED AND SUPPLEMENTARY EXPERIMENTS

The mechanism testbed separates three quantities: estimator variance at a fixed state, expected progress after one update, and failure over a complete run. We vary direction count, readout noise and query budget to compare their scaling. Deployment experiments then assess transfer to restricted parameter interfaces and the accuracy cost of shared configuration selection. PolyStep ablations use matched query budgets and, where specified, equal vertex counts.

Boundary measurements. At the initial damaged state, autograd gives G^2, deterministic-readout estimator sampling gives A_{search}, Hessian-vector power iteration estimates L_{loc}, and Hutchinson sampling estimates $\mathrm{tr}(H)$. Power iteration returns the dominant eigenvalue magnitude, without its sign or a global smoothness bound. Isotropic bisection along 24 rays measures the chance-loss radius R. The seed-42 coefficient point uses $G^2 = 3.44$, $L_{\mathrm{loc}} = 24.4$, $\mathrm{tr}(H) = 238.9 \pm 4.5$, $A_{\mathrm{search}} = 3.4 \times 10^5$, $\sigma_{\mathrm{r}} = 0.272$, $\varepsilon = 0.02$, $D = 101{,}770$, $B = 2 \times 10^5$, and $R = 29.7 \pm 0.6$; the two uncertainties are standard errors of the trace and radius estimates. Each forward redraws Gaussian weight perturbations scaled by the current per-tile maximum absolute weight. Thus σ_{r} is measured loss variability at that state, and both loss mean and variance can change during optimization. Boundary coefficients are measured without fitting the observed thresholds. Expected descent is measured separately by bisection on the sign of $\mathbb{E}[f(\theta - \eta\hat{g}_m)] - f(\theta)$ using common random numbers.

Configuration searches. Step-size searches use at least three candidate values and extend the range when the best observed setting lies at an edge. Joint searches evaluate 5×3 probe–step grids and transfer the selected configuration across the stated devices. PolyStep uses a fixed configuration for each substrate class, with separate searches over effective step and probe scale. POEM uses a search over its initial-distance parameter (Appendix Q).

G.1 ADDITIONAL RESULTS

Query budget and single-direction recovery. For $m = 1$, final-iterate test accuracy at each budget's best step size on seed 42 is 40.22, 69.42, 83.25, 91.22 and 93.82% at 1, 10, 100, 1,000 and 3,189 times a 3,540-query budget. Repeating the two largest budgets with seeds 123/456 gives $94.04 \pm 0.19\%$ at the exact query match, compared with $94.15 \pm 0.10\%$ for direction averaging on the same three seeds. The paired difference is -0.11 pp and is unresolved at $n = 3$ (Appendix S). The direct budget exponent is -0.282 $[-0.43, -0.13]$; its estimation is detailed in Appendix N.

PolyStep configuration sensitivity. Varying only the effective step gives approximately zero gain on the photonic mesh and 1.5–2.3 pp on the crossbar. Varying the probe as well gives gains of 2.57/2.15 pp under deterministic/thermal photonic readout. The crossbar gains are 1.42 pp under deterministic readout ($t = 5.55$, 5/5 positive pairs) and 2.21 pp under read noise ($t = 6.28$, 5/5), measured against the fixed configuration on the same damaged realizations. Selected-checkpoint accuracy also becomes less variable: its sample standard deviation falls from 0.44 to 0.23 pp under crossbar read noise and from 1.12 to 0.78 pp on the thermal mesh. These results show sensitivity to both configuration axes.

Recovery near the crossbar accuracy plateau. Across five seeds, single-direction ZO differs from the configuration with the largest observed mean accuracy by -0.07 pp ($p = 0.21$). The experiment does not resolve a winner among these settings.

Noisy deployment accuracy. Across 20 crossbar seeds, PolyStep's fixed configuration exceeds direction averaging at its fixed configuration by 1.87 pp (20/20 positive pairs, $t = 5.95$, $p = 9.9 \times 10^{-6}$). This is below the Bonferroni threshold 5.0×10^{-4} for 100 comparisons. Each selected checkpoint is evaluated on 20 fresh noisy readout draws; seeds, rather than readout draws, are the replication units. This comparison concerns fixed configurations rather than winners of a common configuration search.

Allocating queries to repeated readout. At matched query cost, averaging two readouts per PolyStep cost evaluation halves the number of transport updates and reduces thermal-mesh accuracy by 0.97 pp ($p = 0.012$, uncorrected). This setting favors spending queries on additional updates.

H ACCURACY AND LIMITS OF THE DISPLACEMENT MODEL

Prediction accuracy. The geometric mean of predicted/measured boundary ratios is 2.20. Individual errors vary: $10/21$ predictions lie within $2\times$ and $15/21$ within the decade grid's $\sqrt{10}$ resolution. Imposing a $2/L_{\mathrm{loc}}$ cap changes these values to $1.70\times$, $13/21$, and $17/21$, respectively, but does not follow from the displacement model. A sufficient descent condition cannot supply a necessary cap. Deterministic runs survive through $\eta = 1$, beyond $2/L_{\mathrm{loc}} = 0.082$; this does not falsify the sufficient condition. At $\sigma_{\mathrm{r}} = 0$ the readout-dominated approximation is inapplicable: the readout term imposes no finite restriction, while curvature and search variance can still constrain steps. We therefore use the uncapped prediction throughout.

Although coefficients are measured without boundary fitting, the displacement model's form was chosen after inspection of the initial grid. That grid is not an independent test of the model form. Additional seeds and deployment analyses provide further measurements, while the controlled crossbar study evaluates policies fixed before recovery on disjoint target devices (Appendix B). It does not test physical hardware or a new task distribution.

In Figure 2, the sufficient-descent and trace forms, capped at $2/L_{\mathrm{loc}}$, have predicted/measured geometric means 0.00377 and 11.7, with $0/21$ and $5/21$ within grid resolution. The uncapped escape ratio has geometric SD 2.01; the figure's direction labels follow the $\sigma = 0.2$ series. The three-seed mean descent boundary at $m = 1$, $\sigma = 0.2$ is 3.35×10^{-4}. Comparing it with the seed-42 trace prediction 3.13×10^{-4} gives approximately 7% agreement, but this mixes a three-seed measurement with one seed's coefficients. The seed-matched geometric ratio 0.834 is the main-text comparison.

Radius and geometry. The radius R averages chance-loss crossings along 24 isotropic rays on one clean batch. It is neither the distance to non-finite loss nor a verified safe-region boundary. The fixed state is not stationary ($G^2 = 3.44$). Using this radius assumes approximately spherical geometry, monotone ray crossings, and a chance-loss surface tracking first exit; none is validated. The curvature ratio $L_{\mathrm{loc}}/(\mathrm{tr}(H)/D) \approx 10^4$ indicates substantial anisotropy. If the minimum radius of the same surface controlled exit, using its mean would raise the predicted boundary because $\eta_{\max} \propto R$. Neither that controlling radius nor its relation to failure is established, so this diagnostic does not identify the sign or magnitude of the prediction error.

Departures from free diffusion. The model omits drift, although single steps still descend well beyond the whole-run failure boundary. At $m = 16$, three-seed local log–log escape-time slopes across one decade of η are $-4.03, -2.55, -2.14, -1.08$, rather than the constant -2 required by free diffusion. This curvature is compatible with restoring drift but does not identify it. A descriptive barrier form is $\mathrm{FEE}_{\mathrm{esc}} = A\eta^{-p}e^{K/\eta}$, motivated by escape-time analyses of SGD (Xie et al., 2021). On the five-point ladder, fixed $p = 1/2$ gives a coefficient of determination 0.998, versus 0.920 for free diffusion and 0.977 for a barrier with $p = 2$. These descriptive fits do not resolve the budget exponent: the fitted escape-time exponent 2.43 ± 0.31 does not exclude 2.

Alternative escape-time models. Let $s^2 = D\sigma_{\mathrm{r}}^2/(2\varepsilon^2 m)$ denote total estimator-noise variance. We compare four descriptions of the query count at escape. Free diffusion (M1) predicts $\mathrm{FEE}_{\mathrm{esc}} = 2mR^2/(\eta^2 s^2)$ using independently measured coefficients. A fitted power law (M2) uses $\mathrm{FEE}_{\mathrm{esc}} = A\eta^{-b}m^c$. The Ornstein–Uhlenbeck approximation (M3) adds linear restoring drift: $X_{t+1} = (1 - \eta\kappa)X_t + \eta n_t$, with $X_0 = 0$, $\kappa > 0$, and the noise increments of Proposition 1. Writing $\rho = 1 - \eta\kappa$ and assuming $0 < \rho < 1$, iteration gives $X_T = \eta \sum_{j=0}^{T-1} \rho^j n_{T-1-j}$. Independence removes cross terms, so

$$\mathbb{E}\|X_T\|^2 = \eta^2 s^2 \sum_{j=0}^{T-1} \rho^{2j} = \eta^2 s^2 \frac{1 - \rho^{2T}}{1 - \rho^2}.$$

M3 predicts escape when this expression reaches R^2, with query count $2mT$; it predicts no escape if the limiting second moment is below R^2. This variance-threshold approximation is not a first-

passage calculation. The barrier model (M4) uses $A\eta^{-p}e^{K/\eta}$ at fixed m; M4m substitutes $K = K_1 m$ when fitting across direction counts. The latter models are exploratory fits.

Fits minimize squared log escape-time residuals. Table 7 compares Akaike's information criterion (AIC) and its small-sample correction (AICc) on two point sets. The preferred model depends on the correction: for example, the fixed-$p = 1/2$ barrier improves AIC over free diffusion by 13.70 on five points, while AICc favors free diffusion by 8.97. These small samples do not establish an alternative model.

Interpretation and censored observations. The barrier coefficient K is fitted phenomenologically. Relating it to a curvature requires the variance along an escape direction; total variance s^2 and coordinate variance s^2/D differ by $D = 101{,}770$. These fits therefore do not identify physical curvature. The value $p = 1/2$ is imposed rather than derived.

Twelve runs survive to their budget limit and provide right-censored escape times. The OU model predicts escape too early for three of these runs, compared with four for free diffusion. Its predicted budget exponent, -0.258, is close to the measured -0.282, but similar agreement holds over a broad $40\times$ range of fitted drift strengths. The fitted power law also agrees with that exponent. Thus this agreement does not identify restoring drift. In this variance-threshold surrogate, sufficiently small steps keep the limiting second moment below R^2; this is not a first-passage or device-stability guarantee.

With fixed $p = 1/2$, M4m violates one of the twelve censoring constraints. Fitting p removes those violations and gives AIC -15.59, but AICc is $+4.41$, compared with -8.56 for free diffusion. Its predictive performance on separate budget measurements has not been established. The measured boundary levels $0.47/0.61/0.55$ across an $8.6\times$ range of σ_{r} also do not distinguish radius error from restoring drift. Escape times that precede the budget limit are consistent with either explanation.

Query fidelity. For independent zero-mean read errors, averaging q draws divides variance by q at q times the query cost. Suvorikova et al. (2026) analyze this cost–fidelity trade-off under explicit oracle and cost models. Our $q = 2$ control did not improve recovery at the tested budgets.

I SELECTION RULE

We first derive the schedule conversion used by all donor policies, then describe the retrospective cross-substrate evaluation. The controlled crossbar protocol and its grid and pooling extensions are specified in Appendix B.

Scheduled steps and the selection margin. Proposition 1 is proved in the main text. Its second-moment identity makes the schedule conversion explicit. For $T = \lfloor B/(2m) \rfloor \geq 1$, define $\eta^2_{\mathrm{RMS}} =$

Table 7: Escape-time models fitted to the five-point $m = 16$ ladder and the seven uncensored configurations. These point sets must be compared separately. The parameter count excludes residual variance; $q = \mathrm{params} + 1$ includes it. Lower AIC and $\mathrm{AICc} = \mathrm{AIC} + 2q(q+1)/(n-q-1)$ are better. A common additive constant is omitted from both criteria. AICc is undefined when $n \leq q + 1$.

Point set	Candidate	params	q	n	AIC	AICc
m=16 ladder	M1: free diffusion	0	1	5	-5.55	-4.22
	M4: barrier, $p = \frac{1}{2}$	2	3	5	-19.25	$+4.75$
	M4: barrier, $p = 2$	2	3	5	-7.69	$+16.31$
	M4: barrier, p free	3	4	5	-21.54	undefined
7 uncensored	M1: free diffusion	0	1	7	-9.36	-8.56
	M2: free power law	3	4	7	-7.75	$+12.25$
	M3: OU variance threshold	1	2	7	-7.74	-4.74
	M4m: barrier, $p = \frac{1}{2}$	2	3	7	-10.63	-2.63
	M4m: barrier, p free	3	4	7	-15.59	$+4.41$

$T^{-1}\sum_{t=0}^{T-1}\eta_t^2$ and $\Pi_{\mathrm{esc}} = \eta_{\mathrm{RMS}}\sigma_{\mathrm{r}}\sqrt{DB}/(2R\varepsilon m)$. Then

$$\mathbb{P}\left(\max_{t\le T}\|X_t\| \ge R\right) \le \frac{Q_T}{R^2} = \frac{2mT}{B}\Pi_{\mathrm{esc}}^2 \le \Pi_{\mathrm{esc}}^2. \tag{9}$$

Thus $\Pi_{\mathrm{esc}} \le \sqrt{\delta}$ is sufficient for the stated model margin. Equality $2mT = B$ recovers the divisible-budget formula; unused queries only loosen the bound. At $\Pi_{\mathrm{esc}} = 1$ the upper bound is trivial. A device-failure guarantee would also require control of the objective-driven motion and a ball wholly contained in a failure-free region. Neither the mean ray radius nor donor transfer establishes these properties.

Exact cosine conversion. Write $\eta_t = \eta_0[a + b\cos(\pi t/T)]$, with $a = 101/200$, $b = 99/200$ and $t = 0, \ldots, T-1$. For $T \ge 2$, pairing t with $T - t$ gives $\sum_{t=0}^{T-1}\cos(\pi t/T) = 1$. The finite geometric sum of $e^{2\pi i t/T}$ is zero; taking real parts and using $\cos^2 x = (1 + \cos 2x)/2$ gives $\sum_{t=0}^{T-1}\cos^2(\pi t/T) = T/2$. Hence

$$q^2 := \frac{\eta_{\mathrm{RMS}}^2}{\eta_0^2} = a^2 + \frac{b^2}{2} + \frac{2ab}{T} = \frac{30203}{80000} + \frac{9999}{20000T}.$$

For $T = 1$, the only step is η_0 and $q = 1$. At $T = 15{,}000$ and 60, respectively, $q = 0.61447$ and 0.62118; the continuous approximation 0.61444 drops the $1/T$ term. Substituting $q\eta_0$ into Eq. 9 and replacing the device coefficient $2R/(\sigma_{\mathrm{r}}\sqrt{DB})$ by the donor median yields Eq. 5. The substitution defines the transfer rule; it is not a consequence of the probability bound.

Retrospective evaluation protocol. The rule takes δ as input. We fix ε by estimator class (10^{-1} for m=1, 10^{-2} for direction averaging), using prior screens (Section 4.4). For each donor unit, compute $\ell_u = 2R_u/(\sigma_{\mathrm{r},u}\sqrt{D_u B})$. Take the median of these levels over the *other* units of the same substrate. We do not take separate coefficient medians. Write their median as ℓ_{donor}. On the rate ladder $\eta_0 \in \{10^{-3}, 10^{-2}, 10^{-1}\}$, choose the largest initial rate whose donor-estimated escape ratio $\widehat{\Pi}_{\mathrm{esc}} = \eta_{\mathrm{RMS}}/(\ell_{\mathrm{donor}}\varepsilon m)$ is at most $\sqrt{\delta}$, using the schedule factor derived above. The target's coefficients and screen outcomes do not enter this selection step. We then read its saved outcome to evaluate the chosen setting. The probe-radius choices were informed by prior screens, so this is not a fully held-out design. Zero per-device search excludes shared calibration and the recovery run. The per-device (ε, η) grid argmax requires 15 budgeted runs. The rule requires a nonempty donor pool with finite, positive radius and noise measurements; otherwise it abstains. Outcome evaluation requires a final accuracy and checkpoint observations. All 300 evaluated configurations have finite checkpoint observations; missing observations would be treated as unknown rather than escape-free.

Accuracy and margin sensitivity. We evaluate the rule at $\delta = 0.1$, over the 20 (substrate, unit, estimator) settings the screens cover. No selected run escaped ($0/20$). The mean accuracy deficit from the per-device grid maximum is 0.025 pp on the crossbar and 4.26 pp on the photonic mesh. Across both substrates, the mean deficit is 2.14 pp, the median is 0.85 pp, and the largest is 11.47 pp (Table 8). It exceeds the best rate at pinned $\varepsilon = 10^{-3}$ on each target by $+8.70$ pp on average ($19/20$ positive differences). Increasing δ to 0.2 reduces the mean deficit to 0.48 pp, with no observed escapes. At $\delta = 0.02$ and 0.05, five settings have no admissible rate; we record abstentions, not safe outcomes.

Interpretation. The 20 settings share 10 devices and overlapping donor pools; they are not independent risk trials. A fixed-transfer comparator sets $(\varepsilon, \eta_0) = (0.1, 0.001)$ for $m = 1$ and $(0.01, 0.01)$ for $m = 250$, on both substrates. It reads no donor coefficients and reproduces all 20 default selections exactly, including every accuracy and escape outcome. These settings are identified retrospectively, so this comparator is not an independently chosen prospective baseline. It demonstrates that the existing outcomes cannot attribute a benefit to donor measurements. The three-rate ladder also limits resolution. The observations support a search-cost and accuracy comparison on the existing screens. They do not calibrate $\mathbb{P}_{\mathrm{device}}(\text{escape}) \le \delta$ or establish safety on unseen devices. A median donor level is not a conservative bound for every target, so the model inequality does not transfer directly to device failure.

Table 8: Donor-rule selections at $\delta = 0.1$ and $B = 30{,}000$ FEE. Each row contains five devices; $m = 250$ denotes sequential deployment averaging. Accuracy is mean clean-readout test accuracy (%); gaps are each device's 15-configuration grid maximum minus the selected accuracy (pp). Fixed-setting transfer reproduces every row without donor measurements. Devices are shared across estimators.

Setting	(ε, η_0)	Accuracy	Mean gap	Worst gap	Escapes
Crossbar, $m = 1$	$(0.1, 0.001)$	95.20	0.01	0.06	0/5
Crossbar, $m = 250$	$(0.01, 0.01)$	95.73	0.04	0.15	0/5
Photonic, $m = 1$	$(0.1, 0.001)$	88.38	1.86	2.43	0/5
Photonic, $m = 250$	$(0.01, 0.01)$	80.74	6.65	11.47	0/5

J STATISTICAL TREATMENT

The tests below concern the benchmark comparisons. The controlled crossbar study instead uses descriptive factory/cohort-balanced summaries (Appendix B) and is not included in this multiple-comparison family.

Replication units are matched device realizations: seeds for the crossbar and spiking SoC, and chip indices for the photonic mesh. We use two-sided paired t tests, first averaging conditions within each replicate when the same device is evaluated under several conditions. Bonferroni correction over 100 comparisons gives a significance threshold of $0.05/100 = 5 \times 10^{-4}$. Uncorrected p-values are exploratory and do not establish family-level significance. Selection protocols differ across experiments (Section 4.4 and Appendix S), and reuse of screening devices can bias comparisons. Small replicate counts limit precision.

Uncertainty at three replicates. Percentile-bootstrap intervals use 10^4 resamples. At $n = 3$, the all-minimum and all-maximum resamples each have probability $3^{-3} = 1/27 > 0.025$, so the population bootstrap $2.5/97.5$ percentiles equal the sample range of the three paired differences. For independent continuous differences symmetric about their mean, this range covers that mean with probability $1 - 2(1/2)^3 = 75\%$. Excluding zero is only three-way sign agreement (two-sided sign-test $p = 0.25$). We therefore identify these intervals as sample ranges, not 95% confidence intervals, and provide Student-t intervals separately. Family corrections use the paired-test p-values, so this interval qualification does not change their outcomes. The budget-exponent intervals use ordinary least squares (OLS) with three residual degrees of freedom and $t_{.975,3} = 3.182$.

Repeated conditions and available statistics. Multiple conditions evaluated on one device are repeated measurements, not independent replicates. For example, the post-drift comparison averages four conditions within each of three seeds, giving $p = 0.093$. For sparse versus standard ZO, the mean difference over all four shared noise levels is 6.58 pp; restricting the conditions changes that mean and is not a separate replication. Some supplementary comparisons are available only as summary statistics. Seven comparisons across multiple conditions lack tests at the declared device-level replication unit and are therefore treated as descriptive. None supports a claim of significance after family correction. Per-run analyses for the main noisy-deployment comparison use the full 20 matched seeds.

Projection-seed sensitivity. Crossbar clean accuracy changes from 94.60 to 94.74% as the spread widens from 0.43 to 0.65 pp. Across ten photonic chips, same-chip projection differences are $+0.52$ pp ($p = 0.14$, deterministic) and $+0.15$ pp ($p = 0.49$, thermal). The testbed reversal holds on $12/12$ shared-seed pairs with similar interval widths. These checks support the examined comparisons, without establishing projection robustness for every reported interval.

K BOUNDARY CONVENTION

A laddered boundary can be summarized by its largest divergence-free rate (*last-stable*) or by the geometric midpoint to the first-diverging rate. We use last-stable throughout. Midpoints multiply measured levels by the ladder's half-step: $\sqrt{10} \approx 3.16$ for decade spacing and $10^{1/8} \approx 1.33$ for

quarter-decade spacing. This changes prediction/measurement ratios, but not within-grid log–log slopes such as $\sigma_{\mathrm{r}}^{-1.059}$ and $m^{0.969}$.

Quantity	last-stable	geometric midpoint
Escape law over 21 points, uncapped	$2.20\times$, $15/21$ within grid	$0.70\times$
the same points with the $2/L_{\mathrm{loc}}$ cap	$1.70\times$, $17/21$ within grid	$0.54\times$, $14/21$
Descent vs divergence (geom. mean)	$27.2\times$ (range 7.9–76.8)	$8.6\times$ (range 2.5–24.3)
Quarter-decade noise ladder	$0.54\times$ $(0.47/0.61/0.55)$	$0.40\times$ $(0.35/0.46/0.41)$

The uncapped decade-grid and quarter-decade prediction/measurement ratios differ by $4.07\times$ under last-stable $(2.20/0.54)$, versus $1.75\times$ under midpoints $(0.70/0.40)$. Resolution explains part, but not all, of their difference. The additional cap changes the prediction independently of this choice of measured reference.

L Exponent measurements

The noise exponent -1.059 uses the quarter-decade ladder at $m = 1$; the averaging exponent 0.969 uses the decade grid. The sufficient descent form $\eta_{\mathrm{ED}} = 2mS/[L(mM + A)]$ saturates with m. The separately measured descent boundary has exponent 0.87 ± 0.09, distinct from the escape boundary's 0.969. At $m = 1$, $\sigma = 0.2$, the trace-based descent prediction/measurement ratio has geometric mean 0.834 over three seed pairs.

A diagnostic ratio of the two approximations. In the readout-dominated, uncapped limit, the trace-based expected-descent approximation is $\eta_{\mathrm{ED}}^{\mathrm{tr}(H)} \simeq 4m\varepsilon^2 G^2/[\mathrm{tr}(H)\sigma_{\mathrm{r}}^2]$. Dividing by the modeled noise-displacement scale gives

$$\frac{\eta_{\mathrm{ED}}^{\mathrm{tr}(H)}}{\eta_{\mathrm{noise}}} \simeq \frac{2\varepsilon G^2\sqrt{DB}}{\mathrm{tr}(H)\sigma_{\mathrm{r}}R}. \tag{10}$$

The ratio is independent of m and falls as $1/\sigma_{\mathrm{r}}$ when all other coefficients are fixed. This cancellation is conditional on the approximations, not a general property of the measured boundaries. The measured σ_{r} spans $8.6\times$ across the points, while the measured boundary ratio spans $9.7\times$, a difference of 13%. This retrospective range diagnostic does not separately validate m independence or control variation in the other coefficients.

The three noise-level ratios $0.47/0.61/0.55$ have a geometric standard deviation of 1.11–$1.14\times$; their $1.29\times$ maximum/minimum ratio is a range, not a dispersion estimate.

Budget resolution. The $-1/2$ budget exponent remains unresolved because the available step-size grids are coarse and the escape-time relation is curved. Finer boundary measurements at common budgets, with replication across the same damaged states, would distinguish these sources of uncertainty. A direct fit should retain both bracketed boundaries and runs that survive to the budget limit, rather than infer a budget exponent solely by inverting an escape-time slope.

M Readout noise and quantized updates

The ratio $r_{\mathrm{noise}} = \sigma_{\mathrm{r}}/\varepsilon$ describes readout-noise amplification but omits device geometry and dimension. At $\varepsilon = 10^{-3}$ it is 53.1 on the mesh and 160.4 on the crossbar, whose D and R differ by factors of 3.8 and 3.7. The displacement model includes these quantities through Π_{esc}. A separate scale is relevant to quantized weights:

$$\Pi_{\mathrm{esc}} = \frac{\eta\sigma_{\mathrm{r}}\sqrt{DB}}{2R\varepsilon m}, \qquad \Pi_{\mathrm{LSB}} \sim \frac{2\eta}{\Delta}\left(G_\infty + \frac{\sigma_{\mathrm{r}}}{\varepsilon\sqrt{m}}\right), \tag{11}$$

where Δ is the quantization step and $G_\infty = \|\nabla f\|_\infty$ for a smooth local surrogate. The second expression compares a heuristic single-update scale with half a quantization step. It omits accumulation in the floating-point shadow and the distance of each coordinate to its rounding boundary.

The cumulative rounding condition in Appendix D determines actual state changes. Only Π_{esc} is quantitatively evaluated on the analog grids; these two ratios do not establish a universal recovery criterion across analog and quantized devices.

N BOUNDARY PREDICTION ACROSS DEVICES

The deployment screens use $\varepsilon \in \{10^{-3}, 10^{-2}, 10^{-1}, 0.3, 1\}$, $\eta_0 \in \{10^{-3}, 10^{-2}, 10^{-1}\}$, $B = 30{,}000$ per configuration, and $m \in \{1, 250\}$. Figure 5 illustrates the prediction on one realization per substrate; the full evaluation uses five realizations per substrate.

Boundary intervals and censoring. For a fixed device, estimator and probe amplitude, the three tested step sizes form a grid column. Escape is monotone in step size in these experiments. A transition between two adjacent steps brackets the boundary; if all three steps survive or all three fail, the boundary lies beyond one end of the grid and is censored. Each column therefore describes one boundary, and columns sharing a device are dependent.

On the two illustrated realizations, the model locates $17/20$ boundaries within the grid's resolution, compared with $9/20$ for the most frequent boundary category. Eight boundaries are bracketed, with a predicted/measured geometric mean of $2.78\times$; the remaining twelve are censored, including nine above the grid. All three configuration-classification errors have $\Pi_{\text{esc}} \in [0.91, 3.04]$, within the grid-resolution band.

Across all ten realizations, boundary placement is correct for $74/100$ columns under the same resolution tolerance, compared with $44/100$ for the majority category. There are 41 bracketed boundaries, with a predicted/measured geometric mean of $3.23\times$. Without the resolution tolerance, placement counts are $15/20$ and $68/100$ for the two and ten realizations, respectively. Across all realizations, prediction ratios are $1.85\times$ on the photonic mesh and $5.79\times$ on the crossbar. The measured crossbar noise spans 0.0349–0.1604; prediction errors are largest at the smallest measured noise levels. Because the prediction itself contains σ_{r}^{-1}, this association does not independently identify a noise-scaling error.

Schedule conversion and sequential updates. We use the exact cosine RMS factor derived in Appendix I. A sensitivity check using these values leaves the self/donor classifications and all default selector choices unchanged. The continuous approximation gives a resolved-boundary prediction/measurement geometric mean of $2.78\times$. The direction-averaged class takes m sequential η_t/m microsteps instead of one same-state averaged step. Under independent readout noise with constant variance, one block contributes

$$m\left(\frac{\eta_t}{m}\right)^2 \frac{D\sigma_{\text{r}}^2}{2\varepsilon^2} = \eta_t^2 \frac{D\sigma_{\text{r}}^2}{2\varepsilon^2 m}$$

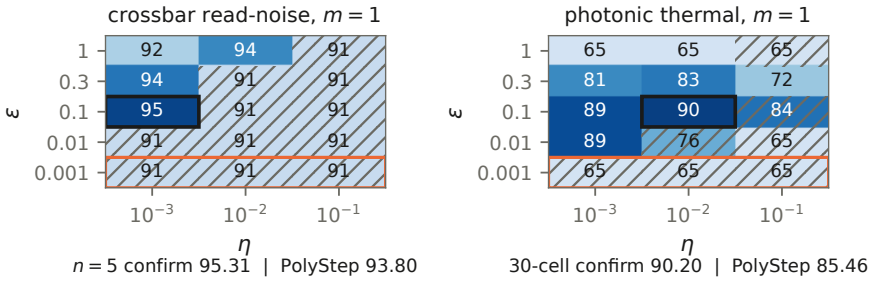


Figure 4: Single-direction ZO (ε, η) screens on one noisy crossbar realization and one thermal photonic chip. Color shows clean-readout test accuracy (%) of the selected checkpoint, normalized separately by panel; hatching denotes escape. The orange outline marks the initial $\varepsilon = 10^{-3}$ radius and the black outline the screen maximum. Labels below give transferred confirmation results and PolyStep's fixed-configuration accuracy. The photonic confirmation pools 30 chip–seed observations across ten chips. Screens select shared configurations; they are not repeated for every device.

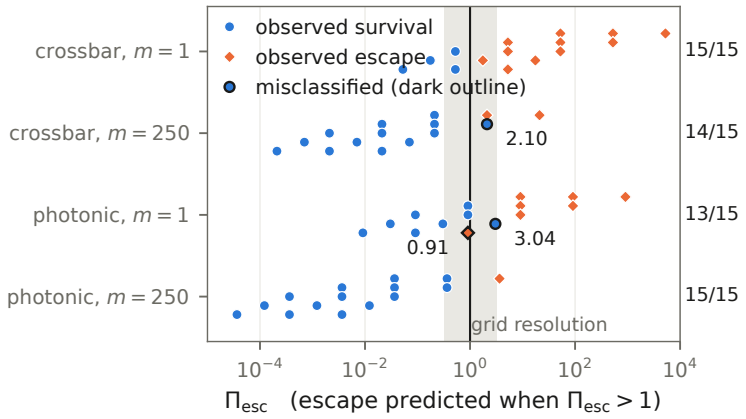


Figure 5: Escape classification on one crossbar and one photonic realization: 30 configurations per substrate from two 5×3 probe/rate screens. Markers show observed outcomes; $\Pi_{\text{esc}} > 1$ predicts escape, defined as a checkpoint accuracy ≤ 0.15. The band represents the rate grid's $3.16\times$ resolution; dark outlines mark the three errors. Coefficients are measured without boundary fitting. Crossbar: $D = 1{,}290$, $\sigma_{\text{r}} = 0.160$, $R = 5.83 \pm 0.21$, $29/30$ correct. Photonic: $D = 4{,}864$, $\sigma_{\text{r}} = 0.053$, $R = 21.6 \pm 1.05$, $28/30$ correct. Radius uncertainties are standard errors over 24 isotropic rays. The crossbar seed-42 screen and coefficients use different damaged-state realizations (Appendix C); pooled accuracy is $57/60$.

to the noise-only second moment. Summing over blocks recovers Q_T. Applying the same argument to all mT microsteps also bounds noise-only exits between blocks. The actual search terms are evaluated at changing states, so neither the estimator covariance identity nor equality of optimization trajectories follows for sequential updates.

Escape-time and budget sampling. At $m = 16$, $\sigma = 0.2$, $\varepsilon = 0.02$ and constant rates $\eta \in \{0.001, 0.00178, 0.00316, 0.00562, 0.01\}$, escape-time analysis uses seeds $42/123/456$. Repeated budgets are geometrically averaged within seed, then across seeds at each rate. The networks share damage seed 12345 but have different initializations. OLS over the five rate means gives slope -2.43 ± 0.31 (standard error; three residual degrees of freedom); inverting its 95% Student-t interval gives the budget exponent -0.411 $[-0.69, -0.29]$. At the two shared escaped rates $0.00562/0.01$, the $m = 64$ versus $m = 16$ log-ratio slopes are $+2.45/+1.76$, bracketing the free-diffusion prediction $+2$. These are two-rate diagnostics, not independent exponent replications.

The $1.10\times$ budget-invariance check is seed 42 at $\eta = 0.00178$: escape occurs at FEE $= 10{,}592/9{,}728/9{,}632/9{,}632/9{,}632$ for $B = 25{,}000/50{,}000/100{,}000/200{,}000/400{,}000$. All escape times precede the smallest budget. Budget independence is consistent with both free diffusion and restoring drift; it does not distinguish them. A decreasing escape-time function still implies a budget-dependent boundary through $\text{FEE}_{\text{esc}}(\eta_{\max}) = B$.

The direct budget slope uses five seed-42 boundaries. The query budgets are $3{,}540$, $35{,}400$, $354{,}000$, $3{,}540{,}000$ and $11{,}289{,}060$. The largest escape-free step sizes are 3×10^{-4}, 10^{-4}, 10^{-4}, 3×10^{-5} and 3×10^{-5}, respectively. OLS gives -0.282 $[-0.43, -0.13]$ with three residual degrees of freedom. Additional seeds $123/456$ replicate accuracy at the top two rungs but do not enter this slope fit. Its coarse boundaries and the curved escape-time relation leave the $-1/2$ budget exponent unresolved.

O Donor transfer and probe-amplitude scaling

We evaluate transfer by predicting each target's outcomes using coefficients from other devices. We also estimate the probe-amplitude exponent from the configuration grids. Both analyses are retrospective: the screens informed their design and reuse the same devices. The separate controlled crossbar study fixes its policies before recovery on disjoint targets (Appendix B).

Calibration from other devices. Within a setting, D and B are fixed. The prediction therefore depends on the unit only through the boundary level $\ell = 2R/(\sigma_{\mathrm{r}}\sqrt{DB})$. That level spans $5.27\times$ over the five crossbar units and $2.94\times$ over the five photonic ones. We score each unit's 5×3 screen against every *other* unit's coefficients. This gives $1{,}104/1{,}200$ donor scores (92.0%), reusing 300 points, against $277/300$ (92.3%) using target-device coefficients. Boundary-placement counts are $300/400$ and $74/100$, respectively. Taking the median donor level gives $280/300$ correct classifications and $72/100$ boundary placements, excluding each target from its donor pool. This tests a shared coefficient estimate without reading the target's coefficients; it does not validate a target-device risk guarantee.

The pooled geometric mean provides no independent evidence of donor transfer. It averages log ratios, while leave-one-out permutes ℓ within each setting; a symmetric permutation therefore leaves the mean almost unchanged. The similar self-scored and leave-one-out values, $3.23\times$ and $3.21\times$, largely reflect this invariance. Classification and column-placement scores are threshold comparisons and do not share that invariance. Per-unit errors show variation obscured by the pooled mean. Level error runs 3.15–$16.59\times$ across the crossbar units and 1.06–$3.65\times$ across the photonic ones. It tracks σ_{r} *inversely*: the largest over-prediction, $16.59\times$, sits on the smallest measured σ_{r}, 0.035. The error ratio itself contains σ_{r}^{-1}. That association cannot separately identify a noise exponent, radius error or other coefficient mismatch. We report it as a level-transfer diagnostic.

Probe-amplitude scaling. The two forms differ in ε when everything else is fixed. The escape form gives $\eta_{\max} \propto \varepsilon^{+1}$, and the $\mathrm{tr}(H)$ descent form gives ε^{+2}. The screens sweep ε over three decades at a fixed unit and method. We estimate a common exponent conditional on the interval model below. This analysis is retrospective.

Resolved-only least-squares slopes are diagnostic here. The rate ladder has three rungs, so a resolved column's last-stable boundary takes at most two values, one decade apart. Four crossbar groups have slope 1.01 and one has 0.98; three photonic groups have 0.50 and two have 0. The observed endpoint span divided by the probe-radius span is not a strict bound on an OLS slope with unevenly spaced probes. We use all columns, including censored ones, in the fit below.

Let b be the failure boundary and g index a substrate, estimator and device. We assume $\log_{10} b = a_g + \beta \log_{10} \varepsilon + e$, with residual $e \sim \mathcal{N}(0, s^2)$, and fit all 100 columns by interval-censored maximum likelihood. A separate intercept a_g absorbs each group's scale, while β describes the common dependence on probe amplitude. Each column contributes the probability that its boundary lies in the interval its ladder outcome implies. The 59 censored columns enter as one-sided constraints on the boundary. The estimated exponent is

$$\beta \;=\; +0.90 \;\; [+0.76,\, +1.08] \quad \text{(likelihood-ratio interval)}. \tag{12}$$

The likelihood-ratio (LR) interval contains the noise model's $+1$ (LR $\chi^2_1 = 1.2$) and excludes the descent approximation's $+2$ (LR $\chi^2_1 = 46.8$), with residual standard deviation 0.33 decades. We assess the estimator on synthetic data drawn on the same grid and ε set. Over $\beta \in \{0.5, 0.75, 1, 1.5, 2\}$ its maximum absolute median bias is about 0.12, so it resolves the two candidates. Its interval covers the true exponent in 17–23 of 25 replicates, below its nominal 95% coverage; Eq. 12 therefore understates uncertainty. This estimate concerns ε in the deployment settings, whereas the testbed measures the σ_{r} exponent (Section 4.1). Together these observations support the proposed scaling on different hardware, subject to the model and coverage qualifications above.

P SPIKING SOC BUDGET SENSITIVITY

The Xylo Audio 2 system-on-chip (SoC), evaluated with XyloSim, classifies eight keywords. Recovery adapts 3,072 int8 weights, accumulating updates in a floating-point shadow before quantization (Appendix D).

At the fixed configuration, PolyStep changes the int8 written state on two of three seeds, POEM on one of six configurations, and the ZO baselines on none; screened ZO does recover. PolyStep's accuracy gain is unresolved: $+0.82$ pp ($t = 1.36$, $p = 0.31$, $n = 3$; per-seed gains $2.00/0.47/0.00$ pp). The initial 1-LSB probe is already $m = 1$'s deterministic optimum. Under read noise the optimum moves to 3 LSB and one-third the rate.

A single-seed diagnostic at $16\times$ the deployment budget (958,464 FEE) reaches 35.13%, gaining 8.93 pp above its own 26.20% baseline and recovering 94% of its 9.47 pp lost accuracy. Recovery rises monotonically, including another 4.40 pp in the final budget interval; the $4\times$ point gains 3.00 pp. Despite spending $16\times$ more, it remains below screened $m = 1$ on the same seed (36.60%). A separate diagnostic removing the subspace cap at $8\times$ budget exactly reverts. Because write-set size and budget both change, that failure does not isolate the effect of the cap.

Q POEM

POEM (Ren & Luo, 2025) adapts step size and smoothing radius through $r_t = \max(r_{t-1}, \|x_t - x_0\|)$, $\mu_t = r_t\sqrt{d/(t+1)}$, and $\eta_t = r_t/\sqrt{\sum_{k=0}^{t}\|g_k\|^2}$. Here x_t is the writable parameter vector, $r_{-1} = r_\varepsilon > 0$, and the update is $x_{t+1} = x_t - \eta_t g_t$. We evaluate it at matched optimization FEE with the same checkpoint selection, first at an initial radius and then at a screened radius. Its published analysis uses shared-randomness paired queries. The bounded-domain algorithm we implement also requires projection and $r_\varepsilon \in (0, D_{\mathcal{X}}]$ for domain diameter $D_{\mathcal{X}}$. Our independent readout errors and unprojected updates are outside that guarantee. We do not use the paper's separate unbounded-domain modification, which changes schedules and requires a Lipschitz estimate. POEM stores a D-float copy of x_0; optimizer memory is not benchmarked here.

Estimator and initialization. For a unit-sphere direction $u_t \in \mathbb{R}^d$, POEM uses $g_t = d(y_t^+ - y_t^-)u_t/(2\mu_t)$. Here d is the writable dimension, denoted by D in the main text. At initialization, $\mu_0 = r_\varepsilon\sqrt{d}$, so the per-coordinate RMS probe amplitude is r_ε. We evaluate an initial value of 10^{-3} and select a second value by a shared search. POEM's published experiments use 10^{-2}; 10^{-3} is a comparison setting rather than a universal default.

Initial-radius results. On the crossbar's ten shared seeds, POEM exceeds PolyStep by 1.36 pp under deterministic readout ($t = 11.85$, $p = 8.6 \times 10^{-7}$, $10/10$ positive pairs), but the read-noise difference is -0.93 pp ($t = -1.74$, $p = 0.12$, $3/10$). Under read noise, its gain over its own damaged state is 0.14 ± 0.28 pp, versus PolyStep's 1.06 pp, and $8/10$ seeds remain within 0.05 pp of baseline. Accuracy proximity does not prove bit-identical reversion. The clean gains are 3.14 versus 1.79 pp. On ten thermal photonic chips, POEM reaches $74.96 \pm 5.23\%$, below screened $m = 1$ (90.20%) and fixed-configuration PolyStep (85.68%) on every chip. Disabling thermal noise gives 75.70%: poor recovery persists, without establishing general noise independence. On the SoC, POEM matches baseline accuracy on $5/6$ configurations and differs by -0.07 pp on the sixth (mean -0.01 pp). One run changes the selected written state by 6 LSB, demonstrating movement without an accuracy gain.

Radius selection. A 10^{-5}–10^{0} ladder on three units per analog substrate finds an interior optimum at $r_\varepsilon = 10^{-1}$ for both. Recovery over each unit's own damaged state at $r_\varepsilon = 10^{-5}, \ldots, 10^{0}$ is $0.00/-0.01/-0.04/+3.17/+3.35/+2.12$ pp on the noisy crossbar and $+11.7/+11.9/+11.7/+14.5/+26.3/+0.7$ pp on the thermal mesh. Its recovery spans 3.39 pp on the noisy crossbar and 25.6 pp on the thermal mesh. Confirmation at 10^{-1} gives $95.76/96.42\%$ on noisy/deterministic crossbar readouts and $89.38/89.64\%$ on thermal/deterministic photonic readouts. On the thermal mesh, it gains 14.42 pp over its initial radius ($10/10$, $p = 6.7 \times 10^{-6}$) and 1.54 pp over the two-axis PolyStep screen ($10/10$, $p = 2.2 \times 10^{-4}$). The corresponding noisy-crossbar difference is -0.57 pp after subtracting each method's damaged-state accuracy ($p = 0.29$), subject to the realization caveat in Appendix C.

Against screened $m = 1$, POEM differs by $+0.53$ pp on the noisy crossbar ($5/5$, $p = 0.020$), -0.05 pp on the deterministic crossbar ($p = 0.17$), and -0.85 pp on the thermal mesh ($0/10$, $p = 9.5 \times 10^{-4}$). On noisy deployment accuracy it exceeds fixed-configuration PolyStep by 1.27 pp ($19/20$ seeds). These comparisons require selecting one scalar rather than separate probe and step scales. Because r_ε initializes both μ_t and η_t, the ladder cannot assign the improvement to either axis alone. The shared observed optimum could be calibrated across devices, but still reflects a configuration search.

Interpretation. The initial radius affects both probe and step schedules, so the sweep establishes sensitivity to initialization without isolating either mechanism. Moreover, r_t is nondecreasing, and the probe radius need not remain at its initial value. Separating the two effects would require runs that vary one schedule while holding the other fixed.

R AVERAGING AND RECOVERY

The grid tests seven constant rates 10^{-6} through 10^{0} for each m, with same-base direction averaging, matched candidate counts and three independently seeded damaged starting states, each shared across configurations. Divergence-free operation and recovery are separate endpoints: a stable run can recover no accuracy.

Deterministic runs remain stable to the grid's upper edge. Under noise, the largest stable rate rises by three decades from $m = 1$ to $m = 3189$, but the recovered fraction varies within its seed spread. Thus averaging widens the tested rate range without a resolved increase in recovery. At $m = 1$, rate 10^{-5} already recovers 84% of lost accuracy, contradicting an intrinsic single-direction failure. The convex stochastic averaging limits of Lin et al. (2025) provide related context, not a proof for these nonconvex recovery experiments with independent readout errors.

S CONFIGURATION-DEPENDENT BENCHMARK COMPARISONS

These comparisons supplement Section 4.4. Table 10 separates fixed settings chosen by different protocols from comparable two-axis thermal screens. Each comparison specifies its replicate pool and selection protocol.

Sensitivity to the selection protocol. At fixed $\varepsilon = 10^{-3}$, the thermal comparison favors PolyStep's fixed configuration by 10.13 pp. After joint probe–step selection, transferred single-direction ZO leads that same configuration by 4.73 pp across ten chips and three optimizer seeds. Different replicate pools prevent these contrasts from isolating probe and step effects. Crossbar comparisons also change sign, but some use different damaged realizations (Appendix C).

Direction averaging at equal query cost. At equal update counts, $m = 3{,}189$ averaging costs $3{,}189\times$ as many queries as $m = 1$. At equal query budgets and separately selected rates, final-iterate accuracy is $94.04 \pm 0.19\%$ for $m = 1$ versus $94.15 \pm 0.10\%$ for averaging on the same three seeds. The difference is unresolved: single-direction failure does not persist when budgets are matched and sufficiently small rates are considered (Appendix G.1).

Adaptive scales still depend on initialization. POEM's initial distance r_ε sets both its probe scale and initial step coefficient $r_\varepsilon/\|g_0\|$, where g_0 is the first gradient estimate. On ten thermal chips,

Table 9: Direction averaging at matched FEE (2×10^5 forward evaluations). Top: largest tested rate with zero divergences across three seeds; deterministic boundaries remain right-censored at the upper rate 10^0. Bottom: mean $\pm$ sample SD across seeds of the fraction of lost accuracy recovered, $100(A_{\text{selected}} - A_{\text{damaged}})/(A_{\text{undamaged}} - A_{\text{damaged}})$. All accuracies use the test set with clean readout. Each m uses one rate maximizing mean accuracy gain across seeds, rather than a separate optimum per seed. Values above 100% mean the selected test accuracy exceeds the undamaged reference.

m	1	4	16	64	256	1024	3189
Largest divergence-free rate							
deterministic	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}
stochastic	10^{-5}	10^{-4}	10^{-4}	10^{-3}	10^{-2}	10^{-2}	10^{-2}
Headroom recovered at that m's best rate (%)							
deterministic	99 ± 1	100 ± 2	98 ± 2	101 ± 1	96 ± 2	95 ± 3	92 ± 4
stochastic	84 ± 7	81 ± 10	81 ± 11	83 ± 9	79 ± 13	83 ± 10	75 ± 14

Table 10: Mean clean-readout test accuracy (%) after validation selection with revert, at matched optimization FEE. Upper: ZO uses fixed step sizes and transferred noisy-readout probes, PolyStep uses screen winners, and analog POEM uses a screened radius (spiking: initial radius). Analog direction averaging uses $m = 250$; spiking follows its query-matched protocol. Replicate sets and crossbar damaged states differ, so "no adapt." is a reference mean. Lower: shared 15-configuration thermal screens, evaluated on ten chips at seed 42. Bold marks the largest mean, not statistical superiority; "(=n)" denotes accuracy equal to no adaptation. Appendix C reports selection, replication, and thermal-screen settings and SDs.

Setting	no adapt.	m=1	m>1	POEM	PolyStep
Fixed settings chosen by different protocols					
Crossbar, deterministic	93.10	**96.48**	94.14	96.42	95.90
Crossbar, read noise	93.10	95.31	94.63	95.76	**95.88**
Photonic, deterministic	61.84	88.20	75.65	**89.64**	88.00
Photonic, thermal	61.84	88.28	75.49	**89.38**	87.83
Spiking SoC	27.29	27.29 (=n)	27.29 (=n)	27.27	**28.11**
Winners of comparable two-axis screens					
Photonic, thermal	—	**90.22**	87.27	—	87.83

changing r_ε from 10^{-3} to the screen-selected 10^{-1} raises accuracy from 74.96% to 89.38%. This establishes initialization sensitivity without isolating probe and step effects (Appendix Q).

Noisy deployment of fixed configurations. Over 20 noisy test draws per selected crossbar checkpoint, PolyStep's fixed configuration leads fixed averaging by 1.87 pp ($20/20$ positive seeds, significant after correction). Screened-radius POEM leads the same PolyStep configuration by 1.27 pp ($19/20$ positive seeds). These compare fixed configurations, not all full-screen winners (Appendices C and Q). PolyStep ablations and quantized-actuation diagnostics remain in Appendices E.1 and P.

T NOTATION

Symbol	Meaning	Measured or set
η, η_t	constant or scheduled step size	set per method
$\eta_{\max}$	empirical failure boundary	largest tested escape-free rate
η_{noise}	noise-displacement scale	Eq. 4; approximation to failure boundary
η_{ED}	sufficient descent threshold	Eq. 2; distinct from measured descent boundary
ε	per-coordinate RMS probe amplitude	set per method; transfer tested on analog screens
m	directions averaged per step	set per method
B	optimization budget, in FEE	set by the protocol; excludes overhead
C_{op}, C_{shared}	operational and shared-acquisition calls	controlled crossbar costs (Appendix B)
$T = \lfloor B/(2m) \rfloor$	complete updates in budget	derived
D	writable dimension	substrate-determined
R	radius proxy	mean ray distance to uniform-prediction loss (Appendix H)
σ	nominal injected noise level	the label of a testbed grid point
σ_{r}	readout noise standard deviation	measured at the initial damaged state
$r_{\text{noise}} = \sigma_{\text{r}}/\varepsilon$	readout-to-probe ratio	derived; omits substrate scale
L, L_{loc}, $\text{tr}(H)$	smoothness bound; local spectral-norm estimate; trace	L assumed; L_{loc} by power iteration; trace by Hutchinson
G^2, G_∞	squared gradient norm; $\|\nabla f\|_\infty$	autograd, testbed only
A_{search}, Σ_{search}	estimator variance and covariance at $\sigma_{\text{r}} = 0$	deterministic-readout estimator variance
V_m	trace variance of $\hat{g}_m$ at m	Eq. 1
μ, S, M	estimator mean; alignment; squared mean norm	∇f_ε; $\nabla f^\top \mu$; $\|\mu\|^2$
q, κ, δ	schedule RMS factor; donor coefficient; model margin	Eq. 5 and Appendix I
X_t, Q_T	noise-only walk; terminal second moment	Proposition 1
Π_{esc}, Π_{LSB}	escape and actuation groups	Eq. 11
$\widehat{\Pi}_{\text{esc}}$	donor-estimated escape ratio	Appendix I
ρ_{s}, ρ_{p}, γ	PolyStep step scale, probe scale, OT temperature	fixed configuration or screen-selected, per protocol
k, V	template constant $d/\bar{s}$; vertex count	16 for the deployed orthoplex
r_ε, μ_t	POEM's initial-distance parameter; its probe radius at t	bracketed once (Appendix Q)
Δ	quantizer step	int8 SoC only
FEE	one scalar loss evaluation on the device	the budget unit