WHEN DO CHEAP EVALUATIONS PRESERVE THE BEST COMPOSED PROGRAM?

Anonymous authors
Paper under double-blind review

ABSTRACT

Cheap evaluations are useful for searching over composed programs only when they identify candidates that remain strong after their continuations are optimized. We study this question for a finite library of first decisions. A least-squares decomposition separates each composed-value gap into a component aligned with the cheap ranking and a constrained continuation residual. The relative size of these components gives a one-sided condition for preserving the cheap winner and a sharp boundary beyond which a pairwise ordering cannot reverse. The same geometry extends to programs with multiple legs and yields certificates for retaining a small candidate set. Guided by this analysis, we evaluate an adaptive rule that reads continuations until the remaining candidates cannot challenge the incumbent, and a promotion rule that carries candidates through weak early fidelities. In simulated control archives, the geometric certificates agree with observed orderings and adaptive reads recover the best composed choice at reduced evaluation cost. On public multi-fidelity optimization benchmarks, the promotion rule improves final-fidelity regret against standard budget-matched schedules.

1 INTRODUCTION

Many decision problems are naturally compositional: a first choice determines which states, skills, or configurations become available to later choices. The quantity that ultimately matters is therefore not the cheap score of the first choice in isolation, but the value of the best continuation that can follow it. Evaluating every continuation at high fidelity is expensive, motivating the common strategy of ranking candidates cheaply and allocating additional computation only to the most promising ones. This principle appears in multi-fidelity optimization, successive halving, and Hyperband (Jamieson & Talwalkar, 2016; Li et al., 2018; Bischl et al., 2023).

The difficulty is that composition can change the ranking. Two candidates with similar cheap scores may lead to very different continuation opportunities, and a candidate that appears weak initially may become the best program after its continuation is optimized. Existing multi-fidelity procedures specify how to promote candidates or learn a mapping between fidelities (Falkner et al., 2018; Adriaensen et al., 2023). Random search and model-based configuration establish complementary search strategies (Bergstra & Bengio, 2012; Hutter et al., 2011; Bergstra et al., 2011; Snoek et al., 2012). We ask a structural question: *when is the cheap ranking itself guaranteed to survive optimized composition?*

Our starting point is a simple least-squares decomposition of optimized continuation values. For a fixed candidate library, it separates the composed score into a component aligned with the cheap ranking and an orthogonal candidate-specific residual. After standardization, every pairwise composed score difference takes the form

$$\text{aligned cheap gap} + \lambda \times \text{constrained residual gap}, \tag{1}$$

where a single scalar λ measures the strength of the residual relative to the aligned signal. This reduction turns cross-fidelity ranking preservation from a high-dimensional comparison into a one-dimensional perturbation problem.

The residual is not arbitrary. Ordinary least squares makes it orthogonal to both the intercept and the cheap-score vector, and our normalization fixes its radius. This geometry yields a sharp boundary

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

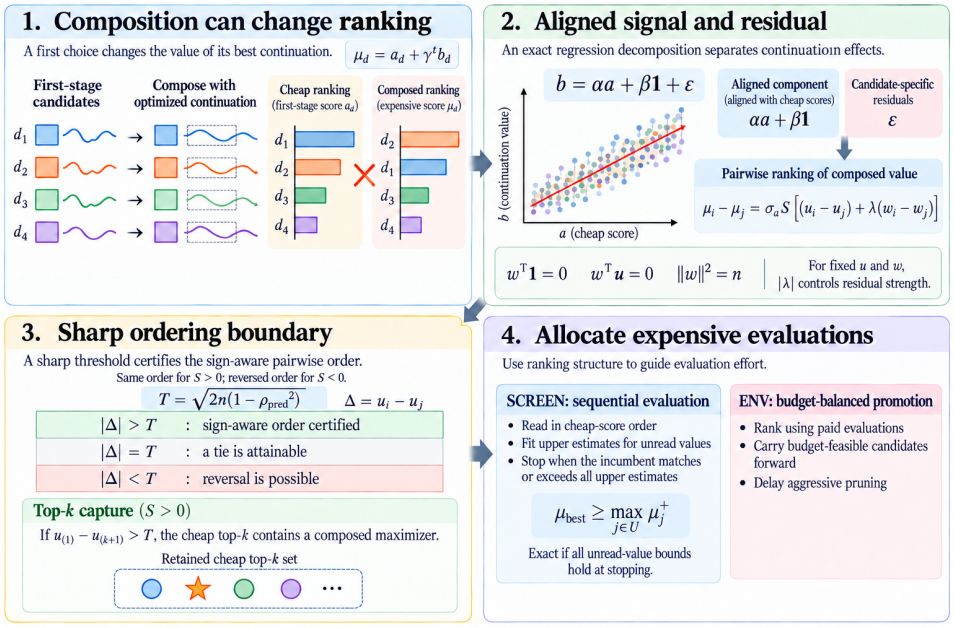


Figure 1: Overview. Optimized continuation can reorder cheap candidates; a least-squares decomposition reduces the shift to $u + \lambda w$, whose constrained residual geometry yields a sharp ordering certificate and motivates adaptive evaluation.

for pairwise rank reversals: once the standardized cheap-score gap exceeds $\sqrt{2n(1-\rho_{\text{pred}}^2)}$, no admissible continuation residual can change the sign-aware order. Below the boundary, an explicit admissible residual can reverse it. The same reduction and boundary persist for compositions of arbitrary depth because weighted sums of the later-leg residuals remain in the same orthogonal subspace.

These results suggest a direct evaluation strategy. When cheap gaps are large relative to the continuation uncertainty, only a small candidate set needs expensive continuation reads; when they are not, additional candidates should be carried forward. We instantiate this idea in two complementary settings: an adaptive stopping rule for composed-policy evaluation, and a budget-balanced promotion rule on public multi-fidelity optimization benchmarks. The first tests whether partial continuation reads can recover the best composed program at lower cost; the second tests whether delaying aggressive pruning helps when very early fidelities provide weak rankings.

Our contributions are threefold. First, we derive an exact scalar reduction of cheap-to-composed ranking under optimized continuation, including a one-sided preservation condition and a parameter-free description of the normalized residual geometry. Second, we prove a sharp, distribution-free ordering boundary, derive top-k capture certificates, and show that the same structure is closed under multi-leg composition. Third, we turn the characterization into adaptive evaluation and promotion rules, and validate the resulting predictions on control-task archives and public multi-fidelity benchmarks.

2 Fidelity Ranking for Composed Programs

We formalize the ranking problem before introducing the reduction. A program selects a first leg d from a library of n policies and then chooses a second leg e after the first leg has been executed. For rollout root r, first-leg horizon t, second-leg horizon h, and objective direction q, the archive records the response of each legal two-leg composition. The control experiments use $n = 16$, $t \in \{1, 2, 5, 10\}$, discount $\gamma = 0.98$, and three `dm_control` domains simulated with MuJoCo (Todorov et al., 2012; Tunyasuvunakool et al., 2020).

For a fixed (t, h, q) row, define the cheap first-leg score and optimized continuation score as

$$a_d = \mathbb{E}_r A_t[r, d, q], \qquad b_d = \mathbb{E}_r \max_e B_h[r, d, e, q], \qquad \mu_d = a_d + \gamma^t b_d. \tag{2}$$

The expectation-of-maximum convention follows the archive's per-root re-optimization. Per-root analyses use the same definitions without $\mathbb{E}_r$. The candidate selected by cheap evaluation is $d_1 = \arg\max_d a_d$, whereas the quantity of interest is the best composed program $\arg\max_d \mu_d$.

We study two increasingly permissive questions. The first asks when d_1 remains optimal after composition. The second asks when a retained top-k set under a is guaranteed to contain the maximizer of μ. These questions are about ranking rather than absolute value prediction: any common offset in the continuation scores is irrelevant, whereas candidate-specific continuation advantages can change the order.

This distinction leads directly to the decomposition in the next section. By regressing b on a, we isolate the continuation component that is aligned with the cheap ranking from the part that can actually perturb it. The normalized results assume $n \geq 3$, $\sigma_a > 0$, $\sigma_\varepsilon > 0$, and $S \neq 0$; degenerate rows are excluded. A zero residual gives an affine ranking; $S = 0$ leaves only the residual. Cheap-score ties use a fixed index order. Remaining optimal allows ties; being the unique maximizer does not. The cost-accounting details used by the adaptive evaluator are introduced only when we turn the characterization into a decision rule in Section 5.

3 A One-Sided Reduction of Cross-Fidelity Ordering

3.1 An exact scalar decomposition

For a fixed row, regress the n-vector of optimized continuation values b on the vector of cheap first-leg scores a with an intercept:

$$b = \alpha a + \beta \mathbf{1} + \varepsilon, \qquad \varepsilon \perp a, \quad \varepsilon \perp \mathbf{1}. \tag{3}$$

Let $\sigma_a = \|a - \bar{a}\mathbf{1}\|/\sqrt{n}$ and define

$$\begin{aligned} u &= (a - \bar{a}\mathbf{1})/\sigma_a, \qquad \sigma_\varepsilon = \|\varepsilon\|/\sqrt{n}, \qquad w = \varepsilon/\sigma_\varepsilon, \\ S &= 1 + \gamma^t \alpha, \qquad R = \gamma^t \sigma_\varepsilon/\sigma_a, \qquad \lambda = R/S. \end{aligned} \tag{4}$$

Substitution in (2) yields the exact pairwise identity

$$\mu(d') - \mu(d) = \sigma_a S\big[(u_{d'} - u_d) + \lambda(w_{d'} - w_d)\big]. \tag{5}$$

The component aligned with a is scaled by S; the remaining difference comes from w. In particular, $S > 0$ preserves the direction of the aligned ordering, whereas $S < 0$ reverses it. Ordinary least squares implies $\mathbf{1} \cdot w = u \cdot w = 0$ and $\|u\|^2 = \|w\|^2 = n$.

This decomposition is an identity for the observed library. The residual collects candidate-specific continuation effects, while discounting changes both S and R. After normalization, λ summarizes their relative strength and determines how far the scores move along the residual direction before the cheap order changes.

3.2 One-sided preservation of the cheap winner

Let $d_1 = \arg\max_d u_d$, $A_d = u_{d_1} - u_d \geq 0$, and $x_d = w_d - w_{d_1}$. Equation (5) reduces the question of whether d_1 also maximizes μ to a condition on λ.

Proposition 1 (One-sided preservation)**.** *Define*

$$\mathcal{R} = \{A_d/x_d : x_d > 0\}, \quad \mathcal{N} = \{A_d/x_d : x_d < 0\}, \quad \lambda_{\text{hi}} = \min \mathcal{R}, \tag{6}$$

with $\min \emptyset = +\infty$ *and* $\max \emptyset = -\infty$*. Let* λ_0 *denote the condition* $x_d = 0 \Rightarrow A_d = 0$*. Then* d_1 *remains optimal if and only if*

$$\begin{cases} \lambda \leq \lambda_{\text{hi}}, & S > 0, \\ \mathcal{R} = \emptyset,\ \lambda_0,\ \lambda \leq \min \mathcal{N}, & S < 0. \end{cases} \tag{7}$$

For $S > 0$*, intersect the bracket* $[\max \mathcal{N}, \lambda_{\text{hi}}]$ *with the reachable* $\lambda > 0$*. For* $S < 0$*, the reachable set is* $(-\infty, \min \mathcal{N}] \cap (-\infty, 0)$ *if* $\mathcal{R} = \emptyset$ *and* λ_0 *hold, and empty otherwise.*

On $S > 0$, the lower endpoint is nonpositive and $\lambda \geq 0$, so only the upper endpoint can bind. The negative-S condition accounts for an aligned continuation that reverses the cheap ordering. Both cases follow directly from nonnegative A_d; the proof is in Appendix B.10. Thus the optimality region is not a generic symmetric interval around zero: on the practically important $S > 0$ branch, only the upper endpoint can invalidate the cheap winner. This one-sided structure will be the bridge from the scalar reduction to a sharp worst-case ordering guarantee.

3.3 Residual geometry and survival scale

A probability model for the residual direction gives a survival function for the positive-S endpoint.

Proposition 2 (Survival identity). *For $\lambda > 0$, with $x_d = w_d - w_{d_1}$ indexed by $d \neq d_1$ and λ_{hi} from (6), $\mathbb{P}(\lambda_{\mathrm{hi}} \geq \lambda) = \mathbb{P}(x_d \leq A_d/\lambda$ for all $d)$.*

Corollary 3 (Tail response and radial factor). *If the joint distribution function F_x has a locally Lipschitz gradient near the origin in the positive orthant, then as $\lambda \to \infty$, $\mathbb{P}(\lambda_{\mathrm{hi}} \geq \lambda) = \mathbb{P}(x \leq 0) + c/\lambda + O(\lambda^{-2})$, where $c = \sum_d A_d g_d$ and g_d is a derivative of F_x at the origin. This expansion holds at a fixed design; the residual law itself can depend on that design. If $w = \varrho\theta$ with independent radius $\varrho > 0$ and direction θ, whose difference CDF satisfies the same regularity, and $\mathbb{E}[\varrho^{-2}] < \infty$, then $g_d = \mathbb{E}[1/\varrho] g_d^{\theta}$.*

The normalized residual has radius $\sqrt{n}$ by construction. A Gaussian on the same $(n-2)$-dimensional residual subspace has radius χ_{n-2} and the same uniform direction law. For $n = 16$, their predicted tail-coefficient ratio, when the directional coefficient is nonzero, is

$$\frac{c_{\mathrm{Gaussian}}}{c_{\mathrm{sphere}}} = \sqrt{n}\,\mathbb{E}[1/\chi_{n-2}] = 4\,\mathbb{E}[1/\chi_{14}] = 1.130920. \tag{8}$$

This comparison isolates the effect of radial normalization; no fitted parameter enters the ratio.

Because w is orthogonal to both $\mathbf{1}$ and u and has fixed norm, its admissible set is a sphere in an $(n-2)$-dimensional subspace. The survival identity makes the one-sided result measurable across rows: given a design A, a residual-direction law assigns a probability to the event that every competitor stays below the cheap winner. Its large-λ limit is the mass $\mathbb{P}(x \leq 0)$ at $\lambda_{\mathrm{hi}} = +\infty$, which may be zero; such directions preserve the winner at every positive residual scale. The first correction scales as $1/\lambda$ and depends linearly on the cheap gaps. Thus the parameter-free ratio in (8) tests the residual normalization separately from the particular cheap-score gaps in Walker and Cheetah.

The next step relates λ to fidelity correlation and derives a guarantee uniform over admissible residual directions.

3.4 Fidelity correlation is the same scalar

The scalar λ also determines the ordinary correlation between cheap and composed scores. This connects the reduction to the fidelity statistics commonly used to decide whether an early evaluation is informative.

Proposition 4 (Correlation and residual ratio). *With $S \neq 0$ and the definitions in (4),*

$$\rho_{\mathrm{pred}} = \frac{\operatorname{sign} S}{\sqrt{1+\lambda^2}}. \tag{9}$$

On $S > 0$, $\lambda \mapsto \rho_{\mathrm{pred}}$ is strictly decreasing; for example, $\rho_{\mathrm{pred}} \geq 0.9$ exactly when $\lambda \leq \sqrt{19}/9 \approx 0.4843$.

To separate continuation magnitude from its alignment with cheap scores, let $r = \operatorname{corr}(a, b)$, $s = \operatorname{sd}(b)/\operatorname{sd}(a)$, $k = \gamma^t s$ and $D = 1 + k^2 + 2kr$.

Proposition 5 (Shape of the correlation). *For $k > 0$, $|r| < 1$ and $D > 0$,*

$$\rho_{\mathrm{pred}} = \frac{1+kr}{\sqrt{D}}, \qquad \partial_k \rho_{\mathrm{pred}} = -\frac{k(1-r^2)}{D^{3/2}}, \qquad \partial_r \rho_{\mathrm{pred}} = \frac{k^2(k+r)}{D^{3/2}}. \tag{10}$$

When $k + r \neq 0$, a level curve has slope $dr/dk = (1-r^2)/[k(k+r)]$; its sign is positive when $k + r > 0$ and negative when $k + r < 0$.

The criterion depends on both the magnitude k and alignment r of the continuation. At fixed r, increasing k reduces correlation. The direction of the response to r depends on $k + r$, a condition that must be retained when interpreting fidelity schedules. On the archive's catalog-mean grid, all twelve comparisons that increase second-leg horizon reduce correlation, while all twelve that increase first-leg horizon raise it (Appendix B.4).

Equation (9) translates the geometric ranking test into a fidelity statistic. For positive S, a correlation threshold of 0.9 corresponds to $\lambda \leq \sqrt{19}/9$; as correlation weakens, the boundary in Theorem 6 widens continuously. A schedule can therefore be interpreted through the informativeness of its current fidelity before it commits to promoting a small subset. In the control archive, the 0.9 gate accepts 361 of 1056 eligible cells: median top-three capture is 0.885 in accepted cells and 0.270 elsewhere, while mean normalized regret is 0.0036 versus 0.0331 (Appendix B.4). These measurements connect the derived scalar to a decision-relevant ranking statistic.

Equation (9) expresses the worst-case ranking boundary directly through fidelity correlation and the standardized cheap-score gap.

4 Exact Ordering Guarantees under Composition

4.1 A sharp boundary for rank reversals

The previous section reduced composition to $u + \lambda w$ and showed that λ is a reparameterization of fidelity correlation. We now use the fixed residual geometry to determine exactly when no admissible continuation can reverse a pairwise order. For the score vectors in a row, define the Pearson correlation

$$\rho_{\text{pred}} = \operatorname{corr}(a, a + \gamma^t b). \tag{11}$$

The standardized residual w belongs to $V(u) = \operatorname{span}\{\mathbf{1}, u\}^{\perp}$ and has norm $\sqrt{n}$. Thus the admissible residual set is a sphere in a 14-dimensional subspace when $n = 16$. Equation (5) can be written as

$$\mu(d) - \text{const} = \sigma_a S(u_d + \lambda w_d), \qquad \|u\|^2 = \|w\|^2 = n, \quad u \cdot w = \mathbf{1} \cdot w = 0. \tag{12}$$

Theorem 6 (Sharp ranking boundary). *For $i \neq j$ with $a_i \neq a_j$ and $S \neq 0$, set $\Delta = u_i - u_j$ and $\mathcal{W} = \{w : \|w\|^2 = n,\ w \cdot \mathbf{1} = w \cdot u = 0\}$. If $|\Delta| > \sqrt{2n(1 - \rho_{\text{pred}}^2)}$, then for every $w \in \mathcal{W}$, $\operatorname{sign}(\mu_i - \mu_j) = \operatorname{sign}(\rho_{\text{pred}}) \operatorname{sign}(a_i - a_j)$. If the inequality is reversed, some $w \in \mathcal{W}$ reverses this sign-aware order. At equality an admissible w yields a tie and none yields a reversal. Both sides of the boundary are attainable.*

The extremal residual direction is proportional to the projection of $e_i - e_j$ onto $V(u)$. Its maximal effect on the pair is $\sigma_a R\sqrt{2n - \Delta^2}$, which yields the threshold in Theorem 6. The theorem holds for all admissible residuals; a distribution over directions is needed only to assign probabilities within the uncertified region (Appendix A).

The boundary has a direct geometric interpretation. Orthogonality removes the component of a pairwise contrast parallel to the cheap scores; only its projection onto $V(u)$ can be changed by the continuation residual. A larger standardized cheap-score gap leaves less room for such a change. Conversely, as $|\rho_{\text{pred}}|$ approaches one, the residual contribution shrinks and the uncertain band contracts. The sign of ρ_{pred} determines whether the certified order agrees with, or reverses, the cheap ranking.

Corollary 7 (Top-k capture certificate). *For $1 \leq k < n$ and $S > 0$, if the gap $\Delta_k = u_{d_1} - u_{d_{(k+1)}}$ exceeds $\sqrt{2n(1 - \rho_{\text{pred}}^2)}$, then $\arg\max_d \mu(d) \subseteq \operatorname{top}_k(u)$ for every $w \in \mathcal{W}$.*

Corollary 8 (Sharpness for the top candidate). *On $S > 0$ at $k = 1$, a strict gap beyond the boundary is necessary and sufficient for d_1 to be the unique maximizer for every admissible residual. At equality it remains a maximizer but can tie a rival; below the boundary some residual makes it suboptimal. For $k > 1$, the pairwise certificate is sufficient, without a claim of necessity for top-k capture.*

The design determines a critical residual scale: if $S > 0$ and $\lambda < \lambda^\star(A_{d_{(2)}})$, where $\lambda^\star(A) = A/\sqrt{2n - A^2}$, then d_1 survives for every admissible residual. The critical value uses the cheap gap; applying it still requires a bound on λ and knowledge that $S > 0$ (Appendix B.1).

For top-k search, the certificate compares the cheap winner with every candidate outside the retained set. A continuous geometric inequality thus provides a discrete search decision: excluded candidates cannot maximize the composed value under the stated residual geometry.

The argument above was stated for two legs only to keep notation minimal. Because its proof uses orthogonality rather than any special property of the second leg, the same construction survives deeper compositions, as shown next.

4.2 The reduction composes to arbitrary depth

Let $p = (p_1, \dots, p_L)$ be a chain of leg horizons with $\sum_i p_i = T$, $c_i = \sum_{j \le i} p_j$, $c_0 = 0$, and weights $g_i = \gamma^{c_{i-1}}$. Each leg defines a vector $\ell^{(i)}$ indexed by the initial choice d; the chain value is $\mu_d^{(L)} = \sum_i g_i \ell_d^{(i)}$. In the archive, any legal chain at the same total horizon has the same final readout:

$$\mu_d^{(L)}(q) = \mathbb{E}_r \operatorname{resp}[r, d, 0, d, \operatorname{hz}(T), q]. \tag{13}$$

This equality holds to the single-precision storage error over all 43 legal chains per domain. The chain decomposition therefore changes the representation of a fixed-horizon value, not the value being predicted.

Theorem 9 (Closure under composition)**.** *Let $a = \ell^{(1)}$ and use its standardized vector u. Assume $\sigma_a > 0$, $S^{(L)} \neq 0$, and $\|E^{(L)}\| > 0$. Regress each later leg on the first, $\ell^{(i)} = \alpha_i \ell^{(1)} + \beta_i \mathbf{1} + \varepsilon_i$. Let $S^{(L)} = 1 + \sum_{i \ge 2} g_i \alpha_i$, $E^{(L)} = \sum_{i \ge 2} g_i \varepsilon_i$, $\sigma_\varepsilon = \|E^{(L)}\|/\sqrt{n}$, $w = E^{(L)}/\sigma_\varepsilon$, $R^{(L)} = \sigma_\varepsilon/\sigma_a$, and $\lambda^{(L)} = R^{(L)}/S^{(L)}$. Then*

$$\mu^{(L)}(d') - \mu^{(L)}(d) = \sigma_a S^{(L)} \big[(u_{d'} - u_d) + \lambda^{(L)}(w_{d'} - w_d)\big], \tag{14}$$

with $\sum_d w_d = u \cdot w = 0$ and $\|w\|^2 = n$.

The weighted sum of least-squares residuals remains orthogonal to the intercept and first-leg score. Thus the one-sided condition and the sharp boundary apply at any chain depth under the corresponding chain-specific $S^{(L)}$ and $\lambda^{(L)}$. Longer chains increase the frequency of the $S < 0$ branch in the archive, and their residual legs show positive coherence (Appendix B.12).

Each later leg may have a different horizon and regression slope, but its weighted residual remains in the subspace orthogonal to the first-leg design. Their sum produces one residual vector and one scalar ratio at any depth.

The closure preserves the diagnostic's form, not the cost of optimizing a deeper program. The fixed-horizon identity is an archive property, not a claim that arbitrary optimized continuations have the same readout. The next section motivates partial evaluation and resource allocation when the full continuation vector is unavailable.

5 From Ordering Guarantees to a Fidelity Ladder

5.1 Adaptive evaluation of composed candidates

The theory identifies when cheap gaps are robust to continuation effects; an evaluator still needs a rule for spending expensive reads when the full residual is not yet known. We therefore use the same aligned-plus-residual view in a sequential form. When the cheap scores a_d are known, evaluate the continuation scores b_d in descending order of a_d. After $m \ge 3$ candidates have been read, fit a least-squares line on the observed pairs (a_d, b_d). Let $\hat{b}_j$ be its prediction for an unread candidate j, $\hat{\sigma}^2$ the residual sum of squares divided by $m - 2$, and $\bar{a}_m$ the mean cheap score of the read candidates. SCREEN assigns the upper estimate

$$b_j^+ = \hat{b}_j + \kappa \hat{\sigma} \sqrt{1 + \frac{1}{m} + \frac{(a_j - \bar{a}_m)^2}{\sum_{i \le m} (a_i - \bar{a}_m)^2}}, \tag{15}$$

and stops once

$$\max_{d \le m} (a_d + \gamma^t b_d) \ge \max_{j > m} (a_j + \gamma^t b_j^+). \tag{16}$$

If $b_j \leq b_j^+$ for all unread candidates, the selected candidate maximizes the true composed value. This implication is Proposition 10; we select the scale κ on FIT data and evaluate exact selection on held-out roots.

The procedure starts with the three highest-a candidates, so that the regression has a residual degree of freedom when the read scores vary. After each additional read it refits the line, updates the upper estimates of all unread candidates, and checks (16). The returned program is the best among the fully evaluated candidates. Its conditional guarantee follows immediately: on the stated upper-estimate event no unread candidate can exceed the best observed composed score. If the read scores have zero variance, continue reading until a fit is defined or all candidates have been evaluated (Appendix B.6).

The unit of cost is a rollout segment. On a fixed row, reading all cheap scores costs $16t$, and evaluating k second-leg blocks costs $16hk$; thus $\mathrm{SEL}(t, h, k) = 16t + 16hk$ and $\mathrm{FULL}(t, h) = 16t + 256h$. At the common $t = h = 10$ scale these become $160 + 160k$ and 2720. SCREEN estimates unread continuation values using the fitted line after its first m reads.

5.2 Promotion across a fidelity ladder

The same principle suggests a complementary rule when fidelity is a training resource rather than a continuation rollout: if the earliest rung gives a weak ranking, postpone aggressive elimination until a more informative rung. For public hyperparameter benchmarks, we evaluate a budget-balanced promotion rule, ENV. At each rung it promotes all candidates whose remaining evaluations fit within the current budget and rung schedule. Its schedule is selected on FIT tasks, and candidate selection uses only scores from paid evaluations. The rule carries more candidates through a weak first rung so that later, more informative rungs can distinguish them. This rule operationalizes the candidate-retention principle suggested by the ranking analysis.

All schedules share a cumulative-resource ledger. Reading a configuration again at an already paid fidelity is free; extending it to a higher fidelity costs the increment. ENV retains the budget needed to evaluate its selected output at the final fidelity, then admits promotions within the remaining budget. Competing methods receive the same budget and final-evaluation reserve. The comparison includes selected schedules and promotion rules; matched schedules isolate the promotion step (Appendix B.9).

These two procedures use different observations, but they instantiate the same decision principle: preserve candidate diversity while the current fidelity is too weak to support a confident order, and concentrate evaluation only after the ranking becomes informative. The experiments test the theoretical boundary first, then the resulting cost–quality trade-offs in both settings.

6 Experiments

The experiments follow the logic of the analysis. We first ask whether the sharp ordering boundary and the parameter-free residual prediction match the control archive. We then ask whether partial continuation reads reduce the cost of exact selection, and finally whether delaying aggressive pruning transfers to public multi-fidelity benchmarks. Details of the candidate pools, ledger, schedule grids, and FIT/TEST splits appear in Appendix B.

6.1 Does the theory predict cross-fidelity ordering?

Across Walker and Cheetah, 30 913 observed candidate pairs lie strictly outside Theorem 6's boundary and none violates the certified sign-aware order. Summed over $k \in \{1, 2, 3, 6\}$, the top-k certificate has 729 successful row–k checks across the tested values of k. The parameter-free tail-coefficient ratio 1.130920 compares with 1.128406 on Walker and 1.129423 on Cheetah. Residual directions also remain structured across horizons: their mean within-cell cosine is 0.69947 on Walker, versus zero for independent isotropic directions. Full radial and cross-horizon diagnostics are in Appendix B.2–B.3.

6.2 Can partial evaluation reduce search cost?

SCREEN uses partial continuation reads to reduce evaluation cost. On Walker, $\kappa = 2$ achieves 92.59% exact selection at 36.8% of FULL cost; a fixed block of six candidates yields 92.66% at

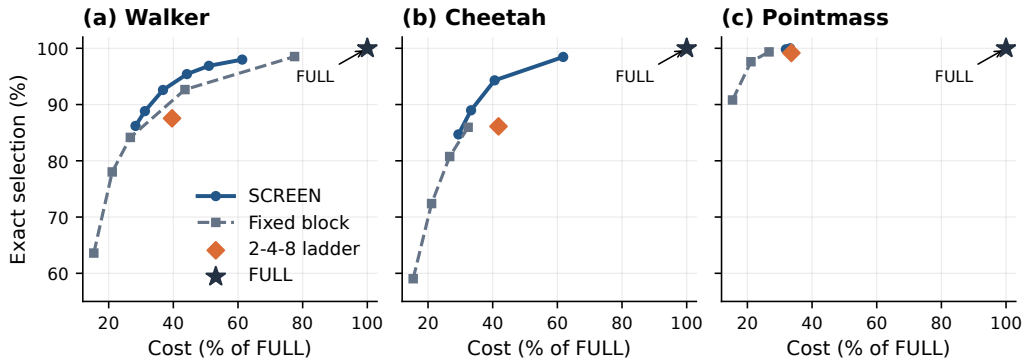


Figure 2: Exact selection versus evaluation cost on the control archive. SCREEN adapts continuation reads to each row; all methods use the same rollout-segment cost model.

43.6%. At $\kappa = 3$, SCREEN achieves 95.41% at 44.2% of FULL. On Cheetah, it reaches 94.30% at 40.6% of FULL for the configuration compared with the 2–4–8 certificate ladder. The full frontiers on all three control domains are in Appendix B.6.

Figure 2 shows the full selection–cost trade-off on the three control domains. On Walker, SCREEN nearly matches a fixed top-six block with 6.8 fewer percentage points of exhaustive cost. On Cheetah, the compared adaptive setting exceeds the 2–4–8 certificate ladder in exact selection while using less cost. On Pointmass, SCREEN reaches exact selection on all rows at 33.2% of FULL cost. SCREEN allocates further reads only when an unread upper estimate can still challenge the best observed composed score.

6.3 Does the allocation principle transfer beyond the archive?

LCBench has 34 tasks, a fixed pool of 512 configurations, and 51 epoch fidelities. We select schedules on 17 FIT tasks and report normalized final-fidelity regret on 17 TEST tasks at budgets 255–4080 epochs; lower is better. The first rung is weakly informative (median rank correlation 0.047 with epoch 51), rising to 0.685 at epoch 2. All methods use the same paid-evaluation ledger except that the marked BOHB baseline samples outside the fixed pool.

Figure 3(a) summarizes the LCBench comparison. ENV improves on successive halving and Hyperband at each of the five budgets, and improves on corrected random search at four of five. On the 85 paired TEST-task–budget cells, the mean regret gaps in favor of ENV are 0.0341 versus successive halving, 0.1158 versus Hyperband, and 0.0690 versus corrected random search. The full baseline grid and pool-controlled comparisons are in Appendix B.8.

In seven YAHPO `rbv2` scenarios, where fidelity is the training-set fraction, ENV has lower regret than Hyperband in 35/35 scenario–budget cells, than random search in 34/35, and than successive halving in 23/35. The first-rung correlation with final fidelity ranges from 0.400 to 0.862. Figure 3(b) makes the scenario-level pattern explicit: ENV wins against Hyperband at every tested budget and retains consistent advantages over random search, while its comparison with successive halving varies across scenarios. The pool sensitivity and full scenario-level results are in Appendix B.9.

7 Related Work

Resource allocation and fidelity models. Successive halving, Hyperband, and asynchronous allocation organize inexpensive evaluations into promotion schedules (Jamieson & Talwalkar, 2016; Li et al., 2018; 2020); BOHB and DEHB combine allocation with model-based or evolutionary search (Falkner et al., 2018; Awad et al., 2021). Multi-fidelity optimization models approximation bias, resource cost, or information shared across sources (Kandasamy et al., 2019; 2017; Klein et al., 2017; Sen et al., 2018; Poloczek et al., 2017). Related approaches share information across tasks, exploit evaluation traces, or extrapolate learning curves (Swersky et al., 2013; Wu et al., 2020; Adriaensen et al., 2023); scaling analyses characterize the effect of optimization budgets (Kadra et al., 2023).

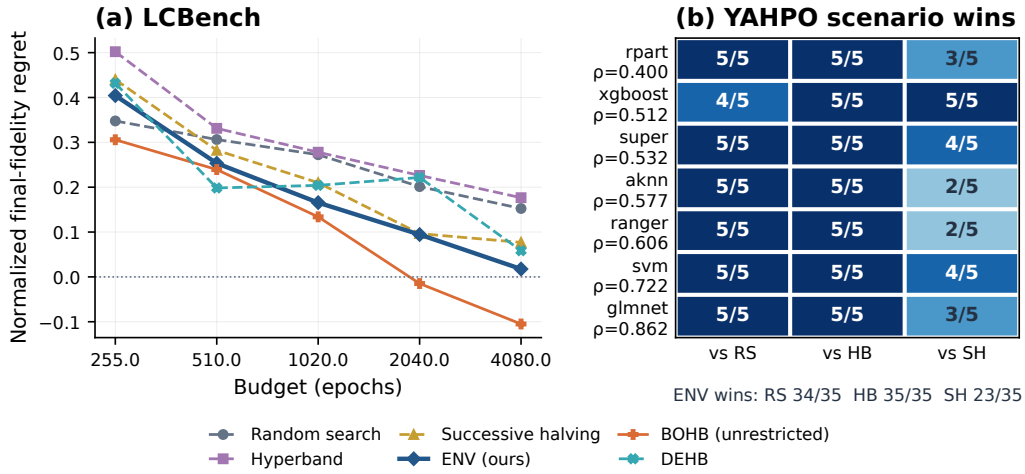


Figure 3: Multi-fidelity benchmark performance. (a) Normalized final-fidelity regret across five budgets (lower is better). (b) Budget wins for ENV against each baseline across seven scenarios; each cell counts wins out of five budgets.

Our contribution is an exact geometric condition on the ranking that these allocation decisions seek to exploit.

Identification and stopping. Pure exploration studies which candidate is best under limited observations (Bubeck et al., 2011; Karnin et al., 2013). Best-arm lower bounds, adaptive sampling, and confidence-based stopping characterize statistical identification costs (Kaufmann et al., 2016; Garivier & Kaufmann, 2016; Jamieson et al., 2014). Here the central object is a deterministic shift from cheap to composed scores: the sharp boundary quantifies which orders survive all residual directions allowed by a fixed design.

Temporal composition and transfer. Options and hierarchical policies represent reusable temporal structure (Sutton et al., 1999; Haarnoja et al., 2018). Successor features, generalized policy improvement, value-function composition, and policy caches support transfer through reusable policies or values (Barreto et al., 2017; 2018; Van Niekerk et al., 2019; Nemecek & Parr, 2021). We examine the complementary question of how optimizing the continuation changes the ranking of initial choices.

8 Conclusion

Optimized continuation turns cheap-to-expensive ranking into a structured perturbation problem. An exact least-squares reduction isolates one residual scale; orthogonality then yields a sharp pairwise boundary and top-k capture certificates that extend to multi-leg composition. These observations motivate adaptive continuation reads and budget-balanced promotion, connecting ranking geometry to practical evaluation decisions. In practice, the capture certificate compares continuation uncertainty with the gap between the cheap winner and the best excluded candidate. A small retained set can therefore contain the composed optimum even when the full ranking changes. This makes set retention a more attainable target than preserving every pairwise order and gives evaluation effort a concrete purpose.

Limitations. Three aspects delimit the results. The certificates concern a fixed finite library and require its correlation or residual scale. SCREEN's exactness is conditional on valid upper estimates, while ENV implements the retention principle as a budget-based heuristic. Empirical comparisons cover the reported benchmarks, pools, and resource ledgers; broader generalization requires additional tasks and evaluation settings.

ETHICS STATEMENT

The study uses simulated control trajectories and public optimization benchmarks; it involves no human subjects or personal data.

REPRODUCIBILITY STATEMENT

The experimental code will be made publicly available upon publication of this paper. Theoretical statements are proved in Appendix A; experimental protocols, additional tables, and supporting measurements are provided in Appendix B.

AI USE STATEMENT

Generative AI tools assisted with polishing the manuscript text and checking the correctness of mathematical derivations. In preparing this revision, they also assisted with the presentation of scientific figures and the verification of bibliographic metadata. The authors take responsibility for the final content, including the text, figures, references, and proofs.

REFERENCES

Steven Adriaensen, Herilalaina Rakotoarison, Samuel Müller, and Frank Hutter. Efficient Bayesian learning curve extrapolation using prior-data fitted networks. In *Advances in Neural Information Processing Systems*, volume 36, pp. 19858–19886, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/3f1a5e8bfcc3005724d246abe454c1e5-Abstract-Conference.html.

Noor Awad, Neeratyoy Mallik, and Frank Hutter. DEHB: Evolutionary Hyberband for scalable, robust and efficient hyperparameter optimization. In *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence*, pp. 2147–2153, 2021. doi: 10.24963/ijcai.2021/296.

Andre Barreto, Will Dabney, Remi Munos, Jonathan J. Hunt, Tom Schaul, Hado P. van Hasselt, and David Silver. Successor features for transfer in reinforcement learning. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/hash/350db081a661525235354dd3e19b8c05-Abstract.html.

Andre Barreto, Diana Borsa, John Quan, Tom Schaul, David Silver, Matteo Hessel, Daniel Mankowitz, Augustin Zidek, and Remi Munos. Transfer in Deep Reinforcement Learning Using Successor Features and Generalised Policy Improvement. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 501–510, 2018. URL https://proceedings.mlr.press/v80/barreto18a.html.

James Bergstra and Yoshua Bengio. Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, 13(10):281–305, 2012. URL http://jmlr.org/papers/v13/bergstra12a.html.

James Bergstra, Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. Algorithms for Hyper-Parameter Optimization. In *Advances in Neural Information Processing Systems*, volume 24, 2011. URL https://proceedings.neurips.cc/paper_files/paper/2011/file/86e8f7ab32cfd12577bc2619bc635690-Paper.pdf.

Bernd Bischl, Martin Binder, Michel Lang, Tobias Pielok, Jakob Richter, Stefan Coors, Janek Thomas, Theresa Ullmann, Marc Becker, Anne-Laure Boulesteix, Difan Deng, and Marius Lindauer. Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. *WIREs Data Mining and Knowledge Discovery*, 13(2), 2023. doi: 10.1002/widm.1484.

Sébastien Bubeck, Rémi Munos, and Gilles Stoltz. Pure exploration in finitely-armed and continuous-armed bandits. *Theoretical Computer Science*, 412(19):1832–1852, 2011. doi: 10.1016/j.tcs.2010.12.059.

Katharina Eggensperger, Philipp Müller, Neeratyoy Mallik, Matthias Feurer, Rene Sass, Aaron Klein, Noor Awad, Marius Lindauer, and Frank Hutter. HPOBench: A Collection of Reproducible Multi-Fidelity Benchmark Problems for HPO. In *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, volume 1, 2021. URL https://datasets-benchmarks-proceedings.neurips.cc/paper_files/paper/2021/file/93db85ed909c13838ff95ccfa94cebd9-Paper-round2.pdf.

Stefan Falkner, Aaron Klein, and Frank Hutter. BOHB: Robust and efficient hyperparameter optimization at scale. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 1437–1446, 2018. URL https://proceedings.mlr.press/v80/falkner18a.html.

Matthias Feurer, Aaron Klein, Katharina Eggensperger, Jost Springenberg, Manuel Blum, and Frank Hutter. Efficient and Robust Automated Machine Learning. In *Advances in Neural Information Processing Systems*, volume 28, 2015. URL https://proceedings.neurips.cc/paper_files/paper/2015/file/11d0e6287202fced83f79975ec59a3a6-Paper.pdf.

Aurélien Garivier and Emilie Kaufmann. Optimal Best Arm Identification with Fixed Confidence. In *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pp. 998–1027, 2016. URL https://proceedings.mlr.press/v49/garivier16a.html.

Tuomas Haarnoja, Kristian Hartikainen, Pieter Abbeel, and Sergey Levine. Latent Space Policies for Hierarchical Reinforcement Learning. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 1851–1860, 2018. URL https://proceedings.mlr.press/v80/haarnoja18a.html.

Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform Chernoff bounds via nonnegative supermartingales. *Probability Surveys*, 17, 2020. doi: 10.1214/18-ps321.

Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2), 2021. doi: 10.1214/20-aos1991.

Frank Hutter, Holger H. Hoos, and Kevin Leyton-Brown. Sequential Model-Based Optimization for General Algorithm Configuration. In *Learning and Intelligent Optimization*, Lecture Notes in Computer Science, pp. 507–523, 2011. doi: 10.1007/978-3-642-25566-3_40.

Kevin Jamieson and Ameet Talwalkar. Non-stochastic best arm identification and hyperparameter optimization. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pp. 240–248, 2016. URL https://proceedings.mlr.press/v51/jamieson16.html.

Kevin Jamieson, Matthew Malloy, Robert Nowak, and Sébastien Bubeck. lil' UCB : An Optimal Exploration Algorithm for Multi-Armed Bandits. In *Proceedings of The 27th Conference on Learning Theory*, volume 35 of *Proceedings of Machine Learning Research*, pp. 423–439, 2014. URL https://proceedings.mlr.press/v35/jamieson14.html.

Arlind Kadra, Maciej Janowski, Martin Wistuba, and Josif Grabocka. Scaling laws for hyperparameter optimization. In *Advances in Neural Information Processing Systems*, volume 36, pp. 47527–47553, 2023. doi: 10.52202/075280-2059. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/945c781d7194ea81026148838af95af7-Abstract-Conference.html.

Kirthevasan Kandasamy, Gautam Dasarathy, Jeff Schneider, and Barnabás Póczos. Multi-fidelity Bayesian optimisation with continuous approximations. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70 of *Proceedings of Machine Learning Research*, pp. 1799–1808, 2017. URL https://proceedings.mlr.press/v70/kandasamy17a.html.

Kirthevasan Kandasamy, Gautam Dasarathy, Junier Oliva, Jeff Schneider, and Barnabás Póczos. Multi-fidelity Gaussian Process Bandit Optimisation. *Journal of Artificial Intelligence Research*, 66:151–196, 2019. doi: 10.1613/jair.1.11288.

Zohar Karnin, Tomer Koren, and Oren Somekh. Almost Optimal Exploration in Multi-Armed Bandits. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pp. 1238–1246, 2013. URL https://proceedings.mlr.press/v28/karnin13.html.

Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the Complexity of Best-Arm Identification in Multi-Armed Bandit Models. *Journal of Machine Learning Research*, 17(1):1–42, 2016. URL http://jmlr.org/papers/v17/kaufman16a.html.

Aaron Klein, Stefan Falkner, Simon Bartels, Philipp Hennig, and Frank Hutter. Fast Bayesian Optimization of Machine Learning Hyperparameters on Large Datasets. In *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 528–536, 2017. URL https://proceedings.mlr.press/v54/klein17a.html.

Liam Li, Kevin Jamieson, Afshin Rostamizadeh, Ekaterina Gonina, Jonathan Ben-tzur, Moritz Hardt, Benjamin Recht, and Ameet Talwalkar. A System for Massively Parallel Hyperparameter Tuning. In *Proceedings of Machine Learning and Systems*, volume 2, pp. 230–246, 2020. URL https://proceedings.mlsys.org/paper_files/paper/2020/file/a06f20b349c6cf09a6b171c71b88bbfc-Paper.pdf.

Lisha Li, Kevin Jamieson, Giulia DeSalvo, Afshin Rostamizadeh, and Ameet Talwalkar. Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research*, 18(185):1–52, 2018. URL https://jmlr.org/papers/v18/16-558.html.

Francesco Mezzadri. How to Generate Random Matrices from the Classical Compact Groups. *Notices of the American Mathematical Society*, 54(5):592–604, 2007. URL https://www.ams.org/notices/200705/fea-mezzadri-web.pdf.

Mark Nemecek and Ronald Parr. Policy Caches with Successor Features. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 8025–8033, 2021. URL https://proceedings.mlr.press/v139/nemecek21a.html.

Florian Pfisterer, Lennart Schneider, Julia Moosbauer, Martin Binder, and Bernd Bischl. YAHPO gym - an efficient multi-objective multi-fidelity benchmark for hyperparameter optimization. In *Proceedings of the First International Conference on Automated Machine Learning*, volume 188 of *Proceedings of Machine Learning Research*, pp. 3/1–39, 2022. URL https://proceedings.mlr.press/v188/pfisterer22a.html.

Sebastian Pineda Arango, Hadi Jomaa, Martin Wistuba, and Josif Grabocka. HPO-B: A Large-Scale Reproducible Benchmark for Black-Box HPO based on OpenML . In *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, volume 1, 2021. URL https://datasets-benchmarks-proceedings.neurips.cc/paper_files/paper/2021/file/ec8956637a99787bd197eacd77acce5e-Paper-round2.pdf.

Matthias Poloczek, Jialei Wang, and Peter Frazier. Multi-Information Source Optimization. In *Advances in Neural Information Processing Systems*, volume 30, 2017. URL https://proceedings.neurips.cc/paper_files/paper/2017/file/df1f1d20ee86704251795841e6a9405a-Paper.pdf.

Rajat Sen, Kirthevasan Kandasamy, and Sanjay Shakkottai. Multi-Fidelity Black-Box Optimization with Hierarchical Partitions. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, pp. 4538–4547, 2018. URL https://proceedings.mlr.press/v80/sen18a.html.

Jasper Snoek, Hugo Larochelle, and Ryan Adams. Practical Bayesian Optimization of Machine Learning Algorithms. In *Advances in Neural Information Processing Systems*, volume 25, 2012. URL https://proceedings.neurips.cc/paper_files/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf.

Richard S. Sutton, Doina Precup, and Satinder Singh. Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning. *Artificial Intelligence*, 112(1–2):181–211, 1999. doi: 10.1016/S0004-3702(99)00052-1.

Kevin Swersky, Jasper Snoek, and Ryan Adams. Multi-Task Bayesian Optimization. In *Advances in Neural Information Processing Systems*, volume 26, 2013. URL https://proceedings.neurips.cc/paper_files/paper/2013/file/f33ba15effa5c10e873bf3842afb46a6-Paper.pdf.

Emanuel Todorov, Tom Erez, and Yuval Tassa. MuJoCo: A physics engine for model-based control. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5026–5033, 2012. doi: 10.1109/IROS.2012.6386109.

Saran Tunyasuvunakool, Alistair Muldal, Yotam Doron, Siqi Liu, Steven Bohez, Josh Merel, Tom Erez, Timothy Lillicrap, Nicolas Heess, and Yuval Tassa. dm_control: Software and tasks for continuous control. *Software Impacts*, 6:100022, 2020. doi: 10.1016/j.simpa.2020.100022.

Benjamin Van Niekerk, Steven James, Adam Earle, and Benjamin Rosman. Composing Value Functions in Reinforcement Learning. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 6401–6409, 2019. URL https://proceedings.mlr.press/v97/van-niekerk19a.html.

Jian Wu, Saul Toscano-Palmerin, Peter I. Frazier, and Andrew Gordon Wilson. Practical Multi-fidelity Bayesian Optimization for Hyperparameter Tuning. In *Proceedings of The 35th Uncertainty in Artificial Intelligence Conference*, volume 115 of *Proceedings of Machine Learning Research*, pp. 788–798, 2020. URL https://proceedings.mlr.press/v115/wu20a.html.

Lucas Zimmer, Marius Lindauer, and Frank Hutter. Auto-PyTorch: Multi-fidelity metalearning for efficient and robust AutoDL. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43 (9):3079–3090, 2021. doi: 10.1109/TPAMI.2021.3067763.

APPENDIX CONTENTS

A. Mathematical proofs

Proofs of the theoretical results 15

B. Supporting measurements

Ordering boundary and top-k capture 18

Residual survival and tail diagnostics 19

Horizon coherence and residual sampling 19

Fidelity gates and leave-one-out validation 19

Cheap-score pruning by sign branch 20

Adaptive continuation reads and stopping bounds 20

Matched-budget control comparisons 22

LCBench protocols, baselines, and sensitivity analyses 24

YAHPO scenarios and transfer diagnostics 26

Reachable one-sided bracket 28

Full-ordering and capture certificates 28

Composition-chain coherence 29

Reachable-state separation 29

State-restoration checks 29

A PROOFS

The proofs use the nondegeneracy assumptions in Section 2. Strict order, weak optimality, and boundary ties are distinguished explicitly.

Theorem 6: the band.

Proof. $\operatorname{sign}(\rho_{\text{pred}}) = \operatorname{sign} S$ by Proposition 4, and $\operatorname{sign}(a_i - a_j) = \operatorname{sign} \Delta$ because u is a centred and divided by a positive constant, so by (5) the question is the sign of $\Delta + \lambda(w_i - w_j)$. The residual part is the linear functional $e_i - e_j$ evaluated at w, and on the radius-$\sqrt{n}$ sphere of $V(u) = \{u, \mathbf{1}\}^{\perp}$ a linear functional with unit vector v has extrema $\pm\sqrt{n}$. With P the orthogonal projector onto $V(u)$,

$$\|P(e_i - e_j)\|^2 \;=\; \|e_i - e_j\|^2 - \frac{(u_i - u_j)^2}{n} \;=\; 2 - \frac{\Delta^2}{n},$$

the first term because P removes only the u component and $e_i - e_j$ is already orthogonal to $\mathbf{1}$. Hence $|\lambda(w_i - w_j)| \le |\lambda|\sqrt{2n - \Delta^2}$ for every $w \in \mathcal{W}$, with both signs attained at $w^\star = \pm\sqrt{n}\, P(e_i - e_j)/\|P(e_i - e_j)\|$, which is admissible when $P(e_i - e_j) \neq 0$. If the projection vanishes, the residual contrast is zero and the nonzero cheap gap retains its sign. The design's sign is then forced for every admissible w if and only if $|\lambda|\sqrt{2n - \Delta^2} < |\Delta|$, that is $\Delta^2(1 + \lambda^2) > 2n\lambda^2$, and $\lambda^2/(1 + \lambda^2) = 1 - \rho^2_{\text{pred}}$ from Proposition 4 turns that into $\Delta^2 > 2n(1 - \rho^2_{\text{pred}})$. At equality the bound $|\lambda(w_i - w_j)| \le |\Delta|$ is attained at $w^\star$ with the sign that cancels Δ, so the pair can be *tied* there; and since the bound is an upper bound at every w, it can never be reversed. □

Corollary 7: a capture beyond the band.

Proof. For $d \notin \operatorname{top}_k(u)$ we have $A_d = u_{d_1} - u_d \ge \Delta_k > 0$, so the pair (d_1, d) is beyond the band with $\Delta > 0$. By the closed-form extremes of Section 4.1 the gap of that pair is at least $\sigma_a\,[S\Delta - R\sqrt{2n - \Delta^2}]$ for every admissible w, and that bracket is positive because $S > 0$ and Δ is beyond the band: the inequality $S\Delta > R\sqrt{2n - \Delta^2}$ is $\Delta^2(S^2 + R^2) > 2nR^2$, which is exactly $\Delta^2 > 2n(1 - \rho^2_{\text{pred}})$. So $\mu(d_1) > \mu(d)$ for every w and every d outside $\operatorname{top}_k(u)$, and the maximiser of μ attains at least $\mu(d_1)$, hence is not any of them. □

Corollary 8: unique and weak optimality.

Proof. On $S > 0$, the smallest cheap gap from d_1 is $\Delta_1 = u_{d_1} - u_{d_{(2)}}$. Beyond the boundary, every rival has a strictly smaller composed score for every admissible residual. Below it, an extremal residual makes $d_{(2)}$ strictly better than d_1. At equality no rival can exceed d_1, but one can tie it. The strict condition characterizes uniform uniqueness; the non-strict condition characterizes uniform weak optimality. For $k > 1$, making an excluded candidate beat d_1 need not make it beat every retained candidate. Thus the pairwise test is sufficient, not shown necessary, for top-k capture. □

The extremes over the admissible set.

Proof. By (12), $\mu(i) - \mu(j) = \sigma_a\,[S\Delta + R(w_i - w_j)]$, since $S\lambda = R$ by definition of λ. The map $w \mapsto w_i - w_j$ is the linear functional $e_i - e_j$, whose extremes on the radius-$\sqrt{n}$ sphere of $V(u)$ are $\pm\sqrt{n}\,\|P(e_i - e_j)\| = \pm\sqrt{2n - \Delta^2}$ by the computation in the proof of Theorem 6, attained at $w^\star_{ij}$, and $R \ge 0$ selects the sign that gives the minimum. For the zero, the lower bracket vanishes when $S\Delta = R\sqrt{2n - \Delta^2}$, that is when $\Delta^2(S^2 + R^2) = 2nR^2$ together with $S\Delta \ge 0$, that is when $\Delta^2 = 2n(1 - \rho^2_{\text{pred}})$ on that branch, since $\rho^2_{\text{pred}} = S^2/(S^2 + R^2)$ by Proposition 4. On $S\Delta \ge 0$ the lower bracket reads $|S|\,|\Delta| - R\sqrt{2n - |\Delta|^2}$, on which substituting $|\Delta|$ for the argument leaves $S\Delta$ written as $|S||\Delta|$ and gives the derivative $|S| + R|\Delta|/\sqrt{2n - |\Delta|^2} > 0$ for $|\Delta| < \sqrt{2n}$; monotonicity extends to the endpoint by continuity. □

The other branch certifies an inversion. On $S < 0$, a positive gap from d_1 to any excluded candidate beyond the band certifies that d_1 is not optimal. If the gap between the two lowest cheap scores exceeds the band, the lowest-score candidate is the unique composed maximizer for every admissible residual.

Proof. For an excluded candidate d whose gap $\Delta = u_{d_1} - u_d > 0$ is beyond the band, the upper extreme of the closed-form extremes of Section 4.1 is $\sigma_a\,[S\Delta + R\sqrt{2n - \Delta^2}]$. With $S < 0$ that is negative exactly when $|S|\Delta > R\sqrt{2n - \Delta^2}$, which is the same inequality as before with both sides' signs taken over, so it holds under the same hypothesis. Hence $\mu(d) > \mu(d_1)$ for every w, and the maximiser of μ attains more than $\mu(d_1)$, so it is not d_1. For the second claim put $i = d_{(n)}$, so that $\Delta = u_{d_{(n)}} - u_d < 0$ and $\min_w(\mu(d_{(n)}) - \mu(d)) = \sigma_a\,[|S|\,|\Delta| - R\sqrt{2n - \Delta^2}]$, positive exactly when $|\Delta|$ is beyond the band. That bracket is increasing in $|\Delta|$ on this branch and $|\Delta|$ is minimised over $d \neq d_{(n)}$ at $d = d_{(n-1)}$, so the narrowest bottom gap is the binding pair. □

A design-dependent critical residual scale.

Proof. $\lambda^\star$ is increasing, and $A_d = u_{d_1} - u_d$ is minimised over $d \neq d_1$ at $d = d_{(2)}$, so $\lambda < \lambda^\star(A_{d_{(2)}})$ implies $\lambda < \lambda^\star(A_d)$ for every $d \neq d_1$, which is the band condition for the pair (d_1, d); Theorem 6 then gives $\mu(d_1) > \mu(d)$ for every admissible w. A fidelity at which every admissible w keeps d_1 on top is a fidelity below the endpoint, whence the critical residual scale of Section 4.1 (at the endpoint itself the statement is weak, by Theorem 6). □

What the whole ordering costs. Let $d_{(i)}$ index the cheap scores in decreasing order, and let $g = \min_{i<n}(u_{d_{(i)}} - u_{d_{(i+1)}})$. Write $A_{(2)} = u_{d_1} - u_{d_{(2)}}$ and $\lambda^\star(A) = A/\sqrt{2n - A^2}$. For strictly ordered cheap scores, the sign-aware strict ordering is preserved for every admissible residual exactly when g exceeds the band.

Proof. Every pairwise gap $u_{d_{(i)}} - u_{d_{(j)}}$, $i < j$, is the sum of the adjacent gaps between the two arms, hence at least g, hence beyond the band; by Theorem 6 the sign of μ on each pair is the design's, with $\operatorname{sign}\rho_{\mathrm{pred}} = +1$ on $S > 0$ and -1 on $S < 0$, so the two orderings agree in the first case and are mirror images in the second. If g is below the threshold, an extremal residual reverses a smallest adjacent pair; at equality it ties that pair but cannot reverse it. Thus the strict condition characterizes preservation of all strict orders. $\lambda^\star$ is increasing in A and $g \leq A_{(2)}$, the top gap being one of the adjacent gaps, so $\lambda^\star(g) \leq \lambda^\star(A_{(2)})$ and the ordering holds on a strictly smaller interval of $|\lambda|$ whenever the two gaps differ. □

Proposition 1: one-sidedness.

Proof. For a fixed d the constraint $S(A_d - \lambda x_d) \geq 0$ reads $A_d \geq \lambda x_d$ when $S > 0$ and $A_d \leq \lambda x_d$ when $S < 0$. In the first case the solution set in λ is

$$x_d > 0:\ \Big(-\infty,\ \tfrac{A_d}{x_d}\Big], \qquad x_d = 0:\ \mathbb{R}, \qquad x_d < 0:\ \Big[\tfrac{A_d}{x_d},\ \infty\Big).$$

Since $A_d \geq 0$, every right-hand endpoint is ≥ 0 and every left-hand endpoint is ≤ 0, so the intersection is $[\max\mathcal{N}, \lambda_{\mathrm{hi}}]$; when either set is empty the corresponding end is $\mp\infty$. In the second case it is

$$x_d > 0:\ \Big[\tfrac{A_d}{x_d},\ \infty\Big), \qquad x_d = 0:\ \big\{\lambda : A_d \leq 0\big\}, \qquad x_d < 0:\ \Big(-\infty,\ \tfrac{A_d}{x_d}\Big].$$

The first of these is a lower bound of the form $\lambda \geq A_d/x_d \geq 0$, so for a *negative* λ it is satisfiable only if it is vacuous, i.e. only if $\mathcal{R} = \emptyset$; the third is $\lambda \leq \min\mathcal{N}$; the second forces $A_d = 0$. Finally, writing $c = \gamma^t$, $\lambda = c\sigma_\varepsilon/(\sigma_a S)$ with $c, \sigma_a, \sigma_\varepsilon > 0$ has the sign of S, so on an $S > 0$ row only $\lambda > 0$ occurs, which discards $[\max\mathcal{N}, 0)$ without loss, and on an $S < 0$ row only $\lambda < 0$ occurs, which is why the $[\max\mathcal{N}, \lambda_{\mathrm{hi}}]$ branch cannot be read there. □

Proposition 2: the survival identity.

Proof. $\lambda_{\mathrm{hi}} < \lambda$ iff there is d with $x_d > 0$ and $A_d/x_d < \lambda$. Since $A_d \geq 0$ and $\lambda > 0$, the condition $A_d/x_d < \lambda$ with $x_d > 0$ is equivalent to $x_d > A_d/\lambda$. Hence $\lambda_{\mathrm{hi}} < \lambda$ iff $\exists d : \; x_d > A_d/\lambda$, and the complement is the stated event. The term $d = d_1$ contributes $x_{d_1} = 0 \leq A_{d_1}/\lambda = 0$ and never triggers. □

Corollary 3, first part: linear response.

Proof. $\mathbb{P}(x_d \leq A_d/\lambda \; \forall d)$ is the distribution function of x evaluated at the point A/λ, which tends to 0 from above. The locally Lipschitz gradient assumed in the corollary gives the expansion $\mathbb{P}(x \leq 0) + \sum_d (A_d/\lambda)\, \partial_d \mathbb{P}(x \leq c)|_{c=0} + O(\|A/\lambda\|^2)$, and $\|A/\lambda\|^2 = O(\lambda^{-2})$. □

Corollary 3, second part: the harmonic factor.

Proof. By the exact homogeneity of the survival function the survival function of λ_{hi} under a random radius is $\mathbb{E}_\varrho[\mathbb{P}_\theta(\lambda_{\mathrm{hi}}(\theta) \geq \varrho\lambda)]$. Taking $\lambda \to \infty$ and using Corollary 3 term by term,

$$\mathbb{E}_\varrho\Big[\mathbb{P}_\theta(x^\theta \leq 0) + \frac{c^\theta}{\varrho\lambda}\Big] = \mathbb{P}(x \leq 0) + \frac{\mathbb{E}[1/\varrho]}{\lambda} \sum_d A_d g_d^\theta,$$

since $\mathbb{P}_\theta(x^\theta \leq 0)$ does not depend on ϱ. To justify the remainder, write $t = 1/\lambda$ and split at $\varrho = t/\delta$, where the directional CDF has a quadratic Taylor remainder for arguments up to δA. On $\varrho \geq t/\delta$, its expected remainder is bounded by $Ct^2\mathbb{E}[\varrho^{-2}]$. On the complement, boundedness of a CDF gives $\mathbb{P}(\varrho < t/\delta) \leq (t/\delta)^2\mathbb{E}[\varrho^{-2}]$; the omitted linear term is also $O(t^2)$ because $t\mathbb{E}[\varrho^{-1}\mathbf{1}_{\{\varrho<t/\delta\}}] \leq (t^2/\delta)\mathbb{E}[\varrho^{-2}]$. Thus the mixed expansion has an $O(\lambda^{-2})$ remainder, and the stated coefficient follows. □

Theorem 9: the reduction composes.

Proof. Substituting the leg regressions of Theorem 9 into the composite of Theorem 9 and collecting the coefficient of $\ell^{(1)}$ gives $\mu^{(L)} = S^{(L)}\ell^{(1)} + E^{(L)} + \text{const}$, and $\ell^{(1)} = \bar{\ell}^{(1)}\mathbf{1} + \sigma_a u$ turns this into $\mu^{(L)} = \sigma_a S^{(L)} u + E^{(L)} + \text{const}$. Differencing at d and d' removes the constant and gives (14). Each ε_i is orthogonal to $\mathbf{1}$ and to $\ell^{(1)}$, hence to every element of $\mathrm{span}\{\mathbf{1}, \ell^{(1)}\} \ni u$, and both orthogonality conditions are linear in the residual, so $E^{(L)} \perp \mathbf{1}$ and $E^{(L)} \perp u$; dividing by $\sigma_\varepsilon = \|E^{(L)}\|/\sqrt{n}$ gives the three identities. Nothing in the argument uses $L = 2$, and nothing uses the weights being geometric: any nonnegative later-leg weights, with the first-leg weight fixed at one, would do. □

The band is a theorem about composition.

Proof. Theorem 9 supplies the two inputs the band theorem and its proof use, and no others: the gap identity (14) and the three identities, that is, $w = E^{(L)}/\sigma_\varepsilon$ of norm $\sqrt{n}$ orthogonal to $\mathbf{1}$ and to u. With the first-leg vector held fixed, its design u is the same at every depth, so $\mathcal{W}$ is the same set and $P_{V(u)}$ is the same projector. For attainability, put the whole residual in one leg: for a unit vector $\hat{e} \in V(u)$, setting $\varepsilon_i = c\hat{e}$ and $\varepsilon_j = 0$ for $j \neq i$ gives $E^{(L)} = g_i c\hat{e}$, hence $w = \sqrt{n}\,\hat{e}$ for $g_i c > 0$, and every direction of $V(u)$ — $w^\star$ included — is reached. Note that the argument is not that a chain's residual is large; it is that the support of the residual does not depend on L at all. □

The split of the tail is invisible. For archive chains with the same total horizon and the same first-leg vector, the decomposition parameters are invariant to splitting the later legs.

Proof. By (13) $\mu^{(L)}$ depends only on T, so it is constant on the fibre, and $\ell^{(1)}$ is fixed by hypothesis. Let Π_V project onto $V(u) = \mathrm{span}\{\mathbf{1}, u\}^\perp$. Then $\Pi_V \ell^{(1)} = 0$ and $\Pi_V \varepsilon_i = \varepsilon_i$, so

$$E^{(L)} \;=\; \Pi_V \mu^{(L)}, \tag{17}$$

a function of $(\mu^{(L)}, u)$ alone. Accounting for the intercept gives, with $a_c = \ell^{(1)} - \overline{\ell^{(1)}}\mathbf{1}$, $S^{(L)} = \langle a_c, \mu^{(L)} - \overline{\mu^{(L)}}\mathbf{1}\rangle/\|a_c\|^2$, and $R^{(L)}$, $\lambda^{(L)}$ and $\mathrm{corr}(\ell^{(1)}, \mu^{(L)})$ are functions of $(\mu^{(L)}, \ell^{(1)})$. □

The rotationally invariant null is closed under composition. Conditionally on the nonnegative weights $c_i = g_i\|\varepsilon_i\|$, suppose the joint law of the nonzero residual directions is invariant under every common orthogonal transformation of $V(u)$. Then, conditional on $E^{(L)} \neq 0$, its direction is uniform and independent of its radius. Coordinate exchangeability alone is not sufficient for this conclusion.

Proof. Write $E^{(L)} = \sum_{i\geq 2: c_i>0} c_i Z_i$ with $Z_i = \varepsilon_i/\|\varepsilon_i\|$ and $c_i = g_i\|\varepsilon_i\| \geq 0$. For any orthogonal O acting on $V(u)$ — and fixing $\ell^{(1)}$, which lies in $V(u)^\perp$ — the joint law of $(OZ_i)_i$ is that of $(Z_i)_i$, conditionally on the weights c_i; hence so is the law of $OE^{(L)}$, which is therefore O-invariant. An invariant law on $V(u)$ is a mixture over radii of uniform laws on the spheres of $V(u)$, so its direction, conditional on a nonzero radius, is uniform and independent of its radius. Neither the weights, nor the leg lengths, nor L appears in the conclusion. □

Coherence preserves each row's marginal law. Let $(e_1, \ldots, e_r)$ be a uniform orthonormal frame in $V(u)$, with $r \leq n-2$. Let $C \succeq 0$ have unit diagonal and rank at most r, choose $G \in \mathbb{R}^{H\times r}$ with $GG^\top = C$, and put $v_h = \sum_k G_{hk} e_k$. Then (i) each v_h is uniform on the unit sphere, and (ii) $\langle v_h, v_{h'}\rangle = C_{hh'}$. The normalized residuals are $w_h = \sqrt{n}\, v_h$, so their cosines equal C.

Proof. (ii) is the expansion $\langle v_h, v_{h'}\rangle = \sum_{k,k'} G_{hk}G_{h'k'}\langle e_k, e_{k'}\rangle = \sum_k G_{hk}G_{h'k} = (GG^\top)_{hh'} = C_{hh'}$. For (i), $\mathrm{diag}(C) = 1$ makes each row g of G a unit vector, so there is an orthogonal O whose first column is g, that is, with $O^\top g = f_1$, where f_1 is the first coordinate vector in $\mathbb{R}^r$. Then $\sum_k g_k e_k = \sum_j (f_1)_j \tilde{e}_j$ with $\tilde{e}_j = \sum_k O_{kj} e_k$, and $\tilde{E}$ is again a uniform frame because the Stiefel manifold is invariant under the right action of O. So v_h has the law of e_1, which does not mention C. □

Proposition 4: the criterion is a bijection of λ.

Proof. Write $c = \gamma^t$ and use population standard deviations. Let $\kappa = c\,\mathrm{sd}(b)/\mathrm{sd}(a)$ and $v = \mathrm{corr}(a,b)$, so that $\alpha = v\,\mathrm{sd}(b)/\mathrm{sd}(a)$ and $1 + \kappa v = 1 + c\alpha = S$. Then $\mathrm{cov}(a, a+cb) = \mathrm{sd}(a)^2 S$ and $\mathrm{Var}(a+cb) = \mathrm{sd}(a)^2(1+\kappa^2+2\kappa v)$, and

$$1 + \kappa^2 + 2\kappa v = (1+\kappa v)^2 + \kappa^2(1-v^2) = S^2\Big(1 + \big(\kappa\sqrt{1-v^2}/S\big)^2\Big).$$

Since $\|\varepsilon\|^2 = n\,\mathrm{sd}(b)^2(1-v^2)$ and $\sigma_a = \mathrm{sd}(a)$, the fraction inside the square is λ exactly. Dividing by $\sqrt{\mathrm{Var}(a+cb)} = \mathrm{sd}(a)|S|\sqrt{1+\lambda^2}$ leaves the sign of S in the numerator. □

Proposition 5: the criterion's shape.

Proof. Direct covariance calculation gives $\rho = (1+kr)D^{-1/2}$, with $D = (k+r)^2 + 1 - r^2 > 0$ under $|r| < 1$, including when $1+kr = 0$. Then $\partial_k \rho = [rD - (1+kr)(k+r)]/D^{3/2}$, and the bracket is $rk^2 + 2kr^2 - k - k^2 r - kr^2 = -k(1-r^2)$; likewise $\partial_r \rho = k[D - 1 - kr]/D^{3/2} = k^2(k+r)/D^{3/2}$. Implicit differentiation of $\rho(k,r) = \text{const}$ gives $dr/dk = -\partial_k\rho/\partial_r\rho = (1-r^2)/(k(k+r))$ when $k + r \neq 0$. Its sign is the sign of $k+r$, as stated in the proposition. □

B SUPPORTING MEASUREMENTS

This appendix documents the geometric checks, evaluation protocols, and robustness analyses behind the main results. Tables retain the archived measurements; headline comparisons use the corrected random-search baseline. Cellwise t statistics are descriptive because budgets repeat within tasks. The reproducibility statement specifies the materials available in the package.

B.1 ORDERING BOUNDARY AND TOP-k CAPTURE

Beyond-band ordering is checked by counting rather than by sampling: over the 512 rows (t, h, root) of each domain, 12 318 of walker's 61 440 pairs and 18 595 of cheetah's lie strictly beyond the band, and of those 30 913 pairs *none* is ordered differently by the composition. Sharpness is checked

by construction: in a separate extremal-direction check, choosing the adverse direction from $\pm w^\star$ reverses every one of the 16 605 walker and 13 716 cheetah pairs inside the band, and none of the 2 595 and 5 484 outside, so the threshold is the boundary rather than a sufficient condition that happens to hold. The certificate of Corollary 7 fires in 729 row–k checks summed across k — 19, 35, 56 and 128 of walker's 512 at $k = 1, 2, 3, 6$ against 51, 95, 129 and 216 of cheetah's — and every check has capture 1. A row can contribute at several k, so this is not a distinct-row count. At $k = 1$ the strict condition characterizes uniform unique optimality, not every observed survival. All 729 checks lie on the $S > 0$ branch, consistent with the fact that the archive's $S < 0$ rows are exactly where the band is *widest*, 37 of them on walker and 12 on cheetah, with $|\rho_{\text{pred}}|$ a median 0.24 and 0.19.

B.2 Residual survival and tail diagnostics

The constrained residual null distinguishes composition effects from an unconstrained isotropic model. Under an isotropic draw in $\mathbb{R}^{16}$, $\mathbb{P}(\lambda_{\text{hi}} = \infty) = 1/16$. On the radius-$\sqrt{n}$ sphere in $V(u)$, coordinate j instead has variance $(n - 1 - u_j^2)/(n - 2)$, so the design changes coordinate exchangeability and the survival atom. Centering alone preserves the isotropic winner probability; orthogonality to u changes it.

The constraint reduces the null atom by 30.6% on Walker and 26.1% on Cheetah. Per catalog, the observed and constrained-null atoms are 0.0437 versus 0.0430 on Walker and 0.0184 versus 0.0450 on Cheetah. Walker survival is above the null at the reported finite thresholds, with ratios 1.002 at $\tau = 0.25$ and 1.102 at $\tau = 1$. The domain-dependent calibration supports using the constrained geometry rather than a universal exchangeable null.

Survival calibration uses the $S > 0$ branch: applying its endpoint test to $S < 0$ would pass automatically because $\lambda < 0 \le \lambda_{\text{hi}}$. Negative-$S$ rows require the separate branch in Proposition 1.

B.3 Horizon coherence and residual sampling

Ordered by the horizon *gap* the walker cosines are not monotone, while ordered by the horizon *ratio* they are. What holds on both domains without exception is that the two endpoints enter separately: raising the *lower* rung of a substitution raises the cosine and raising the *upper* rung lowers it, 360 of 360 over the walker catalog entries, 150 of 150 over their distinct names and 36 of 36 on cheetah. A two-parameter fit $c_0 + c_1 h + c_2 h'$ recovers the six means at $R^2 = 0.957$ on walker with $c_1 = +0.080$, $c_2 = -0.050$ and $R^2 = 0.974$ on cheetah with $c_1 = +0.032$, $c_2 = -0.019$, the same signs in both.

A coherent null is built by drawing four orthonormal directions in $V(u)$, taking any factor G with $GG^\top = C$, where $C \succeq 0$ has unit diagonal, and setting $w_h = \sqrt{n} \sum_k G_{hk} e_k$; the resulting cosine matrix equals C exactly, and each residual has norm $\sqrt{n}$.

The coherent and independent nulls were compared using paired Haar frames. The rank-histogram differences were 0.00011 on Walker and 0.00055 on Cheetah, against redraw floors of 0.00016 and 0.00017. Sampling uses the orthogonal polar factor of a Gaussian. QR sampling is equivalent only with the appropriate diagonal-sign correction (Mezzadri, 2007); each normalized residual has radius $\sqrt{n}$.

B.4 Fidelity gates and leave-one-out validation

The 0.90 correlation gate accepts 361 of 1056 eligible cells. Median top-three capture is 0.885 in accepted cells versus 0.270 elsewhere; mean normalized regret is 0.0036 versus 0.0331. Increasing the second-leg horizon reduces catalog-mean correlation in all twelve tested comparisons; increasing the first-leg horizon raises it in all twelve, consistent with Section 3.4.

Leave-one-basis-out validation selects the same three cells in all 69 folds: $(t, h) = (10, 1), (10, 2), (5, 1)$. All three are favorable on the held-out base in 66/69 folds, including 60/60 on Walker and 6/6 on Cheetah; their worst single captures are $+0.211$ and $+0.925$, respectively. Reward-conditioned and task-free Walker cells have similar median criteria (0.786 and 0.802) and pass rates (0.333 and 0.331), indicating that the diagnostic concerns the score pair rather than the semantic label of the objective direction.

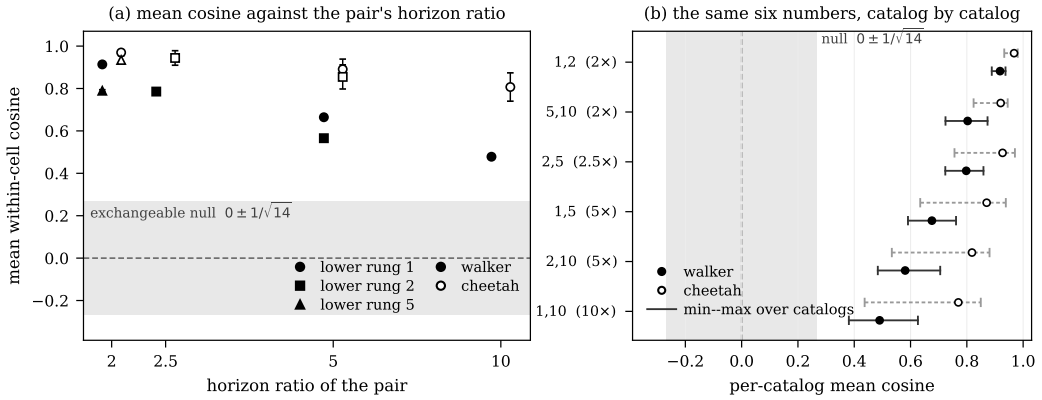


Figure 4: The six within-cell cosines of a 4-horizon cell, both domains. Every number drawn here is read from the archived run log of `bl47`, which is its producer; `bl71` parses that log, reproduces each value to the digits the log printed, and draws. (a) Each pair is placed at its horizon *ratio*; the marker gives the pair's lower rung, filled for walker and open for cheetah, at a slight horizontal offset within a ratio group so that the bars do not overprint; the bars are ± 1 standard error across catalogs. The grey strip is the exchangeable null, $0 \pm 1/\sqrt{14}$: every mark drawn lies above it, the closest being walker's $(1, 10)$ bar at 1.43 times the strip's half-width, and the ordering is by ratio rather than by gap — ordered by gap, $(1, 5)$ at gap 4 sits *below* $(5, 10)$ at gap 5 on walker, and on cheetah one pair inverts across ratios 2 and 2.5. (b) The same six numbers catalog by catalog: the dot is the mean over catalogs of the per-catalog mean and the bar is the min–max of that same quantity, so one functional appears throughout (the pooled cell mean of panel (a), which differs from the dot by up to 0.037, is not drawn). The two domains' bar widths are not comparable, cheetah carrying 6 catalogs against walker's 60.

The gate is a screening diagnostic: 14 accepted cells have zero top-three capture, and 507 rejected cells would have been favorable. Raising it to 0.94 eliminates these false acceptances while rejecting 149 additional cells. Folds share training bases, so their counts are not independent trials.

B.5 Cheap-score pruning by sign branch

The sign-aware reduction identifies when cheap-score pruning is useful. For $k \in \{1, 2, 4, 8\}$, the ratio of random-subset regret to cheap-top-k regret is 8.00, 12.14, 20.37, and 36.08 on Walker's $S > 0$ branch, and 9.11, 12.68, 19.85, and 37.94 on Cheetah's. Thus cheap-score retention improves substantially on random subsets in this regime. On $S < 0$, the ratios fall below one (0.878 to 0.669 on Walker and 0.837 to 0.505 on Cheetah). Reversing the already observed cheap ranking then yields lower regret than the random baseline in all eight cells, consistent with the theorem's sign-aware interpretation.

B.6 Adaptive continuation reads and stopping bounds

A row is a query over sixteen first-leg candidates. Reading all cheap scores costs $16t$; reading one optimized continuation costs $16h$. Thus a block of k candidates costs $16t + 16hk$, compared with FULL $= 16t + 256h$. SCREEN reads candidates in decreasing cheap-score order and chooses when to stop using the continuation values already observed.

Partial-block gates. With two distinct cheap scores and $S_2 \neq 0$, a two-candidate fit has no residual degree of freedom and gives $\rho_2 = \operatorname{sign} S_2$. The sign gate accepts 75.99%, 72.89%, and 80.47% of rows on Walker, Cheetah, and Pointmass; exactness among accepted rows is 83.70%, 80.96%, and 98.97%. The reported 0.90 gate's accepted-row exactness increases to 96.68% for $k = 3$ and 99.41% for $k = 4$. The single-k escalating planner reads the top k, stops if the gate passes, and otherwise reads all sixteen. The 2–4–8 ladder uses the first accepted rung. These are empirical partial-block gates, distinct from the full-library geometric certificate.

Predictive upper estimates. The block is m pairs (a_d, b_d), so it carries a regression and not only its sign, and a regression fitted on the drives read can be extrapolated to the drives not yet read. Index the sixteen drives $1, \dots, 16$ in decreasing a, so that a planner reading the top m has read exactly the initial segment $\{1, \dots, m\}$. Fit $b = \bar{b} + \hat{\alpha}(a - \bar{a})$ by least squares on those m drives, let $\hat{\sigma}^2$ be its residual mean square, and put for each unread drive j

$$b_j^+ = \hat{b}_j + \kappa\,\hat{\sigma}\sqrt{1 + \tfrac{1}{m} + \tfrac{(a_j - \bar{a})^2}{\sum_{i \le m}(a_i - \bar{a})^2}}, \qquad \hat{b}_j = \bar{b} + \hat{\alpha}(a_j - \bar{a}). \tag{18}$$

The rule reads drives in decreasing a and stops at the first m with

$$\max_{d \le m}\big(a_d + \gamma^t b_d\big) \;\ge\; \max_{j > m}\big(a_j + \gamma^t b_j^+\big), \tag{19}$$

returning the best drive among those it read. With $\kappa = 0$ the bound is the fitted value; with $\kappa > 0$ the rule pays for extrapolation, both for the fit's residual and for how far outside the read block the unread drive sits.

Proposition 10 (Soundness of the stopping rule)**.** *If $b_j \le b_j^+$ at every unread drive j, then firing (19) at m implies that a maximizer among $\{1, \dots, m\}$ is also a global maximizer. Unread candidates may tie it.*

Proof. For any unread j, the first inequality below is the hypothesis and the second is (19): $a_j + \gamma^t b_j \le a_j + \gamma^t b_j^+ \le \max_{j' > m}(a_{j'} + \gamma^t b_{j'}^+) \le \max_{d \le m}(a_d + \gamma^t b_d)$, so no unread drive beats the incumbent. □

The guarantee is conditional on unread-value coverage. Under an independent homoscedastic Gaussian linear model at fixed m, the usual prediction error uses T_{m-2} when $\hat{\sigma}^2 = \mathrm{SSE}/(m-2)$. A fixed-$m$ union bound is $\min\{1, (16-m)\mathbb{P}(T_{m-2} > \kappa)\}$. Sequential coverage requires a time-uniform argument, as in the confidence sequence literature (Howard et al., 2020; 2021); SCREEN here uses FIT-selected κ and reports empirical TEST exactness. Archived T_{m-1} comparisons are not used as calibrated coverage claims.

Operational form of SCREEN. The following restates the rule in Section 5; it is not a replacement for the missing experimental implementation.

1. Sort candidates by decreasing a_d and read the first three continuation scores. Use a fixed candidate-index order for cheap-score ties.
2. At each $m < n$, fit an intercept and slope to the m read pairs and compute $\hat{\sigma}^2 = \mathrm{SSE}/(m-2)$ and the unread estimates in (18).
3. If (19) holds, return a maximizer among read candidates. Otherwise read the next candidate and repeat. At $m = n$, return a maximizer over the fully evaluated library.

The normalized analysis assumes nondegenerate fits. For a deployment implementation, a safe extension is to continue reading whenever the observed design has zero variance or a fitted quantity is non-finite; no experimental result for this extension is claimed here. The archived implementation's variance divisor and tie handling must be checked against its producer before reproduction.

Selection–cost frontier. SCREEN adapts its effort to the query rather than paying for the same block on every row. On Walker, $\kappa = 2$ gives exactness 0.9259 at 36.8% of FULL, compared with 0.9266 at 43.6% for a fixed top-six block. At $\kappa = 3$ it reaches 0.9541 at 44.2%; at $\kappa = 6$, 0.9799 at 61.3%. Relative to the cheapest fixed block reaching those accuracy levels, savings are 6.8, 10.7, and 16.2 percentage points of FULL. The corresponding Cheetah savings are 3.0, 29.9, and 15.7 points.

The 2–4–8 ladder is dominated by a SCREEN setting in each domain: 0.8755 at 39.6% versus 0.9259 at 36.8% on Walker; 0.8612 at 41.8% versus 0.9430 at 40.6% on Cheetah; and 0.9917 at 33.6% versus 1.0000 at 33.2% on Pointmass. On Pointmass, the last setting is exact on all 2176 rows. Increasing SCREEN to $\kappa = 8$ gives 0.9838 and 0.9870 exactness at 67.98% and 67.69% on Walker and Cheetah. Fixed blocks remain effective at low budgets; exhaustive-fallback gates reach higher exactness at greater cost. Table 1 retains the complete reported non-dominated frontier.

Table 1: Non-dominated selection–cost points on the 256-root tree. Exactness is the fraction of rows returning a true maximizer. Cost is the mean of cost/FULL at each row's (t, h), with FULL $= 16t + 256h$. The floor assumes an oracle stopping at the true maximizer in cheap-score order. "Certificate" denotes the empirical single-k gate with exhaustive fallback, not the full-library geometric guarantee. Dominance requires at least one strict improvement.

domain	rule	exact	cost (% of FULL)
walker, 38 720 rows, floor 23.0			
	fixed block $m = 1, 2, 3$	0.6360, 0.7802, 0.8416	15.4, 21.1, 26.7
	screening $\kappa = 0, 1, 2$	0.8620, 0.8883, 0.9259	28.3, 31.2, 36.8
	fixed block $m = 6$	0.9266	43.6
	screening $\kappa = 3, 4, 6$	0.9541, 0.9688, 0.9799	44.2, 51.0, 61.3
	fixed block $m = 12$	0.9853	77.5
	certificate $k = 4, 6, 8$	0.9941, 0.9994, 1.0000	82.0, 86.6, 89.3
cheetah, 3840 rows, floor 24.5			
	fixed block $m = 1, 2, 3$	0.5904, 0.7240, 0.8076	15.4, 21.1, 26.7
	screening $\kappa = 0$	0.8469	29.4
	fixed block $m = 4$	0.8596	32.4
	screening $\kappa = 1, 2, 3, 4, 6$	0.8898 to 0.9846	33.3 to 61.8
	certificate $k = 4, 6$	0.9959, 1.0000	80.3, 84.6
pointmass, 2176 rows, floor 16.3			
	fixed block $m = 1, 2, 3$	0.9081, 0.9761, 0.9936	15.4, 21.1, 26.7
	screening $\kappa = 0, 3$	0.9982, 1.0000	31.9, 33.2

B.7 MATCHED-BUDGET CONTROL COMPARISONS

Protocol. The control comparison uses the same rollout-segment ledger and the grid $t, h \in \{1, 2, 5, 10\}$. It evaluates 15 families and 113 configurations, including fixed and random subsets, uniform allocation, UCB-style scores, successive halving, Hyperband, bias-budget elimination, partial-block gates, and SCREEN. Configurations are selected on a FIT half of root groups and evaluated on the disjoint TEST half over eight balanced partitions. The query populations have 60 catalogs and 2420 rows on Walker, 6 and 240 on Cheetah, and 3 and 136 on Pointmass. Two full-table-sign rules, REV and BRANCH, are diagnostic references, excluded from budgeted competitors because observing their sign would require exhaustive continuation reads.

Results. On the 256-root tree, SCREEN is the lowest-loss family at budgets 1000 and 1400 on Walker and Cheetah, improving mean normalized regret by 0.00052 and 0.00171, respectively. At budget 400 it improves on the best competing family by 0.00143 on Walker and 0.00031 on Pointmass. Table 2 retains all budgets, both root scales, realized costs, and paired differences. SCREEN is infeasible at 150 and 250 because its initial fit is too costly; at 650 the coarse-grid comparison favors the competing family. The differences are small and should be read with their costs and catalog counts.

Sensitivity to the SCREEN grid. The original search uses six κ values, compared with twenty UCB settings and eighteen partial-block settings. Refining κ to steps of 0.25, with values up to 8, improves the two Cheetah cells constrained by the original grid. At budget 400, the selected $\kappa = 0.75$ costs 386 and lowers TEST loss from 0.01839 to 0.01449; at 650, $\kappa = 2.5$–2.75 costs 638 and lowers loss from 0.00653 to 0.00484. Both improvements occur in all eight partitions. The paired comparisons change from $+0.00109$ to -0.00281 and from $+0.00010$ to -0.00158. Walker at 650 improves from a $+0.00070$ to a $+0.00018$ gap. The 32-root tree shows the same direction for the two Cheetah cells: 0.01846 to 0.01611 at 400, and 0.00856 to 0.00616 at 650.

At matched realized spend, the refined Cheetah gaps are $+0.00004$ and -0.00151, which separates the rule's benefit from the size of its budget. On Pointmass, gains saturate and the remaining gaps are at most 0.00004 above budget 400. Mean FIT-cost caps need not equal realized TEST spend; both are distinguished in the table. The vectorized sweep and literal stopping rule agree on all 715 776 comparisons across 44 736 cell-rows and sixteen tested κ values.

Table 2: Matched-budget control comparisons on two root scales. Each configuration is selected on FIT root groups and evaluated on disjoint TEST groups over eight balanced partitions. Loss is mean normalized regret; cost uses the rollout-segment ledger. The paired difference d averages the eight partition means; negative favors SCREEN. The reported catalog-clustered z uses only the first partition, so it need not have the sign of d. Full-table-sign oracle references are excluded from the eleven competing families. "Infeasible" means no SCREEN setting satisfies the mean FIT-cost cap.

	best affordable family				the criterion's rule			paired	
budget	family	configuration	loss	cost	$\hat{\kappa}$	loss	cost	d	z
published 32-root tree									
walker, 60 catalogs, 2420 rows									
150	SEL	k=1	0.09497	144		rule infeasible at this budget			
250	SEL	k=2	0.04713	216		rule infeasible at this budget			
400	SEL	k=4	0.01947	360	κ=1	0.01790	383	-0.00157	-4.98
650	SEL	k=8	0.00561	648	κ=3	0.00811	607	+0.00250	+12.77
1000	MF1	k=8	0.00336	872	κ=6	0.00331	900	-0.00005	-1.39
1400	MF1	k=8	0.00336	872	κ=6	0.00331	900	-0.00005	-1.39
cheetah, 6 catalogs, 240 rows									
150	UCB	k=1, β=0.5	0.08265	144		rule infeasible at this budget			
250	SEL	k=2	0.04145	216		rule infeasible at this budget			
400	SEL	k=4	0.01900	360	κ=0	0.01846	370	-0.00054	-4.72
650	SEL	k=8	0.00648	648	κ=2	0.00856	540	+0.00207	+3.74
1000	MF1	k=8	0.00405	872	κ=6	0.00272	890	-0.00133	+0.54
1400	MF1	k=8	0.00405	872	κ=6	0.00272	890	-0.00133	+0.54
pointmass, 3 catalogs, 136 rows									
150	SEL	k=1	0.00849	144		rule infeasible at this budget			
250	SEL	k=2	0.00231	216		rule infeasible at this budget			
400	SEL	k=4	0.00049	360	κ=3	0.00011	389	-0.00038	-1.02
650	SEL	k=8	0.00005	648	κ=6	0.00004	467	-0.00001	-1.00
1000	MF1	k=8	0.00004	872	κ=6	0.00004	467	+0.00000	+1.00
1400	MF1	k=8	0.00004	872	κ=6	0.00004	467	+0.00000	+1.00
256-root tree									
walker, 60 catalogs, 2420 rows									
150	UCB	k=1, β=0	0.07282	144		rule infeasible at this budget			
250	SEL	k=2	0.03470	216		rule infeasible at this budget			
400	UCB	k=4, β=1	0.01422	360	κ=1	0.01278	374	-0.00143	-5.66
650	UCB	k=8, β=2	0.00424	648	κ=3	0.00493	606	+0.00070	+2.24
1000	MF1	k=8	0.00267	872	κ=6	0.00216	854	-0.00052	-0.44
1400	MF1	k=8	0.00267	872	κ=6	0.00216	854	-0.00052	-0.44
cheetah, 6 catalogs, 240 rows									
150	SEL	k=1	0.08430	144		rule infeasible at this budget			
250	UCB	k=2, β=0	0.04230	216		rule infeasible at this budget			
400	UCB	k=4, β=1	0.01730	360	κ=0	0.01839	336	+0.00109	-0.19
650	SEL	k=8	0.00643	648	κ=2	0.00653	554	+0.00010	+1.33
1000	MF1	k=8	0.00353	872	κ=6	0.00182	892	-0.00171	-1.79
1400	MF1	k=8	0.00353	872	κ=6	0.00182	892	-0.00171	-1.79
pointmass, 3 catalogs, 136 rows									
150	UCB	k=1, β=0	0.00607	144		rule infeasible at this budget			
250	SEL	k=2	0.00186	216		rule infeasible at this budget			
400	SEL	k=4	0.00040	360	κ=4	0.00010	392	-0.00031	-0.55
650	SEL	k=8	0.00004	648	κ=6	0.00007	448	+0.00003	+0.50
1000	MF1	k=8	0.00003	872	κ=6	0.00007	448	+0.00004	+0.50
1400	MF1	k=8	0.00003	872	κ=6	0.00007	448	+0.00004	+0.50

B.8 LCBENCH PROTOCOLS, BASELINES, AND SENSITIVITY ANALYSES

Benchmark and protocol. Automated configuration and learning systems motivate evaluating both search and allocation (Feurer et al., 2015; Zimmer et al., 2021). Public suites such as HPOBench, HPO-B, and YAHPO support reproducible comparisons across tasks (Eggensperger et al., 2021; Pineda Arango et al., 2021; Pfisterer et al., 2022). We use the YAHPO LCBench surrogate for 34 OpenML tasks, drawing a fresh 512-configuration pool per task and materializing 51 epoch fidelities into a $34 \times 512 \times 51$ array. The target is validation balanced accuracy at epoch 51. This pool is distinct from the original LCBench sample of 2000 configurations.

The ledger charges resource increments, makes re-reads free, and reserves $H = 51$ epochs to evaluate the returned configuration at the target. Budgets are $B \in \{255, 510, 1020, 2040, 4080\} = 51 \times \{5, 10, 20, 40, 80\}$. Normalized regret is $(\text{best} - \text{found})/(\text{best} - \text{median})$ at the target, with best and median taken over the task's fixed pool. Even-indexed tasks form the 17 FIT tasks; odd-indexed tasks form the 17 TEST tasks. All free parameters are selected on FIT.

Random search varies its read epoch; ladder grids use $\eta \in \{2, 3, 4\}$ and $r_0 \in \{1, 2, 3, 4\}$; Hyperband uses $\eta \in \{2, 3\}$ and $R \in \{27, 51\}$. ENV promotes candidates using paid scores and the number of remaining evaluations affordable under its schedule. ORC uses the same counts with target-ranked survivors. The producer is required to reproduce the exact population recurrence and tie rules.

Primary comparisons. ENV has lower TEST regret than Hyperband and successive halving at all five budgets and than corrected random search at four. The paired mean regret gaps over 85 task–budget cells are $+0.1158$, $+0.0341$, and $+0.0690$, respectively. Table 3 includes model-based samplers and candidate-set controls; Table 4 gives paired counts. Unrestricted BOHB has lower regret than ENV, with a mean gap of -0.0752; the pool-constrained BOHB variants have gaps of $+0.0031$ and -0.0149. These controls distinguish improvements within a fixed candidate set from the benefit of generating new candidates.

Table 3: LCBench: FIT-selected mean TEST regret under the shared resource ledger. The first random-search row is the historical baseline; current comparisons use "axis live," which repairs its fidelity selection. Unrestricted BOHB and DEHB may propose configurations outside the fixed pool; negative regret means exceeding that pool's best value. Pool-snapped and pool-redrawn BOHB isolate candidate-set effects. ENV wins $14/15$ budget comparisons against corrected random search, Hyperband, and successive halving combined; ORC uses privileged target rankings.

	normalised regret at $H = 51$, FIT-selected, TEST tasks					
family	B=255	B=510	B=1020	B=2040	B=4080	selected configuration
random search	0.7142	0.4060	0.3131	0.2425	0.1875	e=51 throughout
Hyperband	0.5021	0.3312	0.2780	0.2262	0.1766	η=2, $R \in \{51, 27\}$
successive halving	0.4399	0.2823	0.2100	0.0965	0.0769	$\eta \in \{4, 2, 3, 3, 3\}$
random search, axis live	0.3477	0.3063	0.2724	0.2013	0.1525	$e \in \{5, 13, 13, 8, 13\}$
random search, fresh	0.6007	0.5101	0.3168	0.2505	0.1330	e=51, fresh draws, no pool
random search, capped	0.6952	0.3771	0.3046	0.2488	0.1666	e=51, fresh draws, capped at 512
BOHB (model-based)	0.3058	0.2393	0.1339	-0.0148	-0.1050	(η, R): (2, 51), (2, 51), (3, 27), (3, 27), (2, 27)
BOHB, pool-snapped	0.4173	0.2681	0.1315	0.0877	0.0461	(η, R): (3, 27), (3, 51), (2, 27), (2, 51), (2, 51)
BOHB, pool-redrawn	0.3536	0.2510	0.1258	0.0968	0.0334	(η, R): (2, 51), (2, 51), (3, 51), (2, 27), (2, 51)
DEHB (model-based)	0.4318	0.1983	0.2041	0.2217	0.0585	(η, R): (3, 51), (2, 51), (3, 51), (3, 51), (2, 27)
ENV (this paper)	0.4043	0.2534	0.1656	0.0944	0.0176	$\eta \in \{2, 3, 2, 4, 2\}$, $r_0 \in \{1, 1, 1, 2, 1\}$
ORC (oracle)	0.3274	0.2309	0.1624	0.0554	0.0000	same counts as ENV, target-ranked survivors

Table 4: Paired LCBench comparisons over 85 TEST task–budget cells. The gap is B's regret minus A's; positive favors ENV. Ties remain in the cellwise descriptive t statistic. Current random-search claims use the "axis live" row. The BOHB controls separate unrestricted sampling from pool-constrained search, and ORC quantifies the remaining promotion opportunity.

		paired cells			mean regret gap	
A	B	A better	tie	B better	gap	t
ENV	random search, as archived	64	14	7	$+0.1856$	$+7.12$
ENV	random search, axis live	51	26	8	$+0.0690$	$+3.82$
ENV	Hyperband	63	6	16	$+0.1158$	$+4.91$
ENV	successive halving	27	53	5	$+0.0341$	$+3.76$
ENV	BOHB (model-based)	29	0	56	-0.0752	-3.26
ENV	BOHB, pool-snapped	37	17	31	$+0.0031$	$+0.15$
ENV	BOHB, pool-redrawn	27	24	34	-0.0149	-0.78
ENV	DEHB (model-based)	48	0	37	$+0.0358$	$+1.57$
ENV	ORC	0	64	21	-0.0318	-3.76
ENV	CPR, composed	10	69	6	$+0.0132$	$+2.00$
ENV	CPR, uncomposed	8	74	3	$+0.0094$	$+1.96$

Coverage and promotion. The first epoch is weakly informative: median rank correlation with epoch 51 is 0.047, increasing to 0.685 at epoch 2 and 0.828 at epoch 4. These rank correlations describe the benchmark and are not substituted for the Pearson correlation in the theorem.

rung (epoch)	1	2	4	8	16	32
median corr with H	$+0.047$	$+0.685$	$+0.828$	$+0.912$	$+0.961$	$+0.989$

At budget 4080, ENV covers all 512 candidates at its first rung, retains 20 to the target, and reads the pool optimum with probability 0.882. The FIT-selected successive-halving schedule begins with 243 candidates at epoch 4, retains approximately 61 to the target, and captures the optimum with probability 0.412. Their spent-budget fractions are similar. ORC retains ENV's counts and captures the optimum on every task, identifying an opportunity for better promotion decisions within the same budget.

Table 5: Candidate coverage at $B = 4080$ on LCBench. ENV initially covers all 512 pool candidates and later evaluates 20 at the target, capturing the pool optimum more often than the FIT-selected successive-halving schedule. ORC uses the same promotion counts and target rankings. Mean spent-budget fractions are 0.984 for ENV and 0.980 for SH.

	candidates still held, averaged over TEST tasks								
	at epoch							H-reads	
family, B=4080	1	2	4	8	16	32	51	(ARGMAX rate)	
successive halving	243.0	243.0	243.0	81.0	61.1	61.1	61.0	61.0	(0.412)
ENV	512.0	512.0	327.0	163.0	81.0	41.0	20.0	20.0	(0.882)
ORC	512.0	512.0	327.0	163.0	81.0	41.0	20.0	20.0	(1.000)

Promotion-rule comparisons. The directly applied control ranking band retains 0.5–9% of the pool across rungs. On common schedules, ENV has lower regret than a fixed standardized gap in 18 cells versus 5, and than a flat band in 22 versus 6. Flat and composed bands split their 40 paired cells 10/20/10 (wins/ties/losses). ORC reaches regret 0.0000 at the largest budget, compared with 0.0176 for ENV. These observations support retaining budget-feasible candidates until later evaluations distinguish them.

Model-based sampler controls. DEHB uses the same four (η, R) settings and fidelity rungs as BOHB under the shared ledger. Its paired gap to ENV is $+0.0358$, with ENV lower at four budgets;

DEHB leads at $B = 510$ (0.1983 versus 0.2534). Two fresh-process runs agree on all 680 regrets, recommendations, call counts, and realized spends. Giving DEHB its best TEST configuration reduces the mean gap to $+0.0116$. Its mean search spend is 0.912 of budget.

The unrestricted BOHB comparison is stable across eight sampler-seed offsets: gaps range from -0.1053 to -0.0525 (mean -0.0840). DEHB's paired gap remains positive on all eight offsets, $+0.0358$ to $+0.0720$ (mean $+0.0447$); selection on TEST changes its sign on three. ENV also improves on the two fresh-draw random-search controls at every budget on all eight offsets, with mean gaps $+0.1658$ and $+0.1602$.

Candidate-set controls. Pool-snapped BOHB maps proposals to their nearest pool configuration in range-normalized coordinates, with lower-index tie-breaking. Pool-redrawn BOHB rejects repeated proposed pool candidates when possible. Both retain the published sampler and shared ledger. Their differences from unrestricted BOHB are $+0.0783$ and $+0.0603$ mean regret, showing that the candidate set materially changes the comparison. Cache hits are free for all methods. The snapped form repeats an entry-rung candidate in 36.05% of proposals; the redrawn form makes $622\,056$ extra proposal draws. These are protocol characteristics, not additional resource charges in the reported ledger.

Pool-size and pool-draw sensitivity. The 4096-configuration pool contains the published 512 configurations exactly. On this larger search problem, paired ENV gains remain positive against Hyperband ($+0.1173$) and successive halving ($+0.0269$), compared with $+0.1158$ and $+0.0341$ on the smaller pool. The historical random-search pool-size control uses the pre-repair baseline and is not used to claim robustness against corrected random search.

Splitting the 4096 candidates into eight disjoint pools gives eight within-pool comparisons. ENV's gap is positive on all eight against Hyperband ($+0.1158$ to $+0.1640$, mean $+0.1326$) and successive halving ($+0.0318$ to $+0.0644$, mean $+0.0457$). The same signs hold when each baseline receives its best TEST configuration. Against the historical three schedule baselines combined, ENV leads at $36/40$ pool–budget pairs; individual budget winners can change even when average gains persist. Each pool uses its own normalization, so levels are not compared across pools.

For samplers that generate configurations outside the pool, changing the pool changes only the normalization: with best values b_k and scales d_k, $\mathrm{nr}_k = \mathrm{nr}_0(d_0/d_k) + (b_k - b_0)/d_k$. Under this re-normalization, unrestricted BOHB remains ahead on all eight pools, while DEHB's paired gap remains positive on all eight. The confined BOHB controls require rerunning the search. At the three smallest budgets, unrestricted BOHB is further ahead of ENV than either confined form on all eight pools: differences range from 0.0344 to 0.0936 for snapping and from 0.0033 to 0.0704 for redrawing. This separates candidate-set effects from a favorable single pool draw.

Implementation checks and scope. Three archived defects were corrected: random search's inactive low-fidelity recommendation branch, indexing an outside-pool sampler as a pool row, and sharing the pool's random stream in a fresh-draw control. Current headline claims use the repaired readouts. Two random-search repair implementations agree on all 40 checked cells; the on-demand ledger and materialized table agree on the checked recommendations and costs. Historical rows are labeled in the tables to distinguish them from current baselines.

The SCREEN variants were also evaluated with a separate cheap-pass-then-descend allocation (SH2). On LCBench, standalone successive halving is better across the tested budgets; the stopping step itself improves on SH2 at four of five budgets, with a paired gap of -0.0121 and descriptive $t = -0.82$. This motivates evaluating ENV's retention policy separately from the control stopping rule. Cross-scenario results are given next.

B.9 YAHPO SCENARIOS AND TRANSFER DIAGNOSTICS

Protocol. Seven additional YAHPO scenarios use training-set fraction as fidelity: `rpart`, `xgboost`, `super`, `aknn`, `ranger`, `svm`, and `glmnet`. Each scenario has a fresh 512-configuration pool and nineteen fidelity levels from 0.05 to 1, stored in an $n_{\mathrm{inst}} \times 512 \times 19$ array. The target is 1 and budgets are $\{5, 10, 20, 40, 80\}$ full-evaluation units. The ledger, target-evaluation reserve, normalized regret, and index-based FIT/TEST split follow the LCBench protocol.

Ladder grids use $\eta \in \{2, 3, 4\}$ and $r_0 \in \{0.05, 0.1, 0.2, 0.4\}$; Hyperband uses $\eta \in \{2, 3, 4\}$ and $R \in \{0.5, 1\}$. Random search uses $\{0.05, 0.0625, 0.1, 0.125, 0.2, 0.3, 0.5, 1\}$. Each scenario's grid contains 278 rows. All main comparisons select configurations on FIT and evaluate on TEST.

Table 6: Seven YAHPO scenarios under the shared ledger. Positive gaps favor ENV and average over 290, 295, 255, 295, 295, 265, and 285 TEST instance–budget cells in row order. Each count lists wins out of five budgets against RS/HB/SH. "Paired" selects all configurations on FIT; "generous" gives each baseline its best TEST configuration. The bottom row aggregates scenario–budget wins. The rung statistic is rank correlation, distinct from the theorem's Pearson correlation.

scenario	ρ at r_1	mean paired regret gap			cells ENV better	
		vs RS	vs Hyperband	vs SH	paired	generous
rbv2_rpart	$+0.400$	$+0.0970$	$+0.1327$	$+0.0149$	$5/5/3$	$5/5/3$
rbv2_xgboost	$+0.512$	$+0.0606$	$+0.0719$	$+0.0162$	$4/5/5$	$4/5/5$
rbv2_super	$+0.532$	$+0.0859$	$+0.0906$	$+0.0060$	$5/5/4$	$5/5/4$
rbv2_aknn	$+0.577$	$+0.0363$	$+0.1452$	$+0.0104$	$5/5/2$	$5/5/1$
rbv2_ranger	$+0.606$	$+0.0982$	$+0.1754$	$+0.0092$	$5/5/2$	$5/5/2$
rbv2_svm	$+0.722$	$+0.1073$	$+0.0610$	$+0.0063$	$5/5/4$	$5/5/1$
rbv2_glmnet	$+0.862$	$+0.0484$	$+0.0348$	$+0.0047$	$5/5/3$	$5/5/3$
		$34/35$	$35/35$	$23/35$	$92/105$	$88/105$

Cross-scenario performance. ENV has lower regret than Hyperband in $35/35$ scenario–budget cells, than random search in $34/35$, and than successive halving in $23/35$. These aggregate to $92/105$ paired comparisons and $88/105$ when baselines receive their best TEST configurations. Mean paired gaps are positive against all three baselines in every scenario. The sole random-search budget exception is xgboost at $B = 5$, with gap -0.0173.

Split and pool sensitivity. Re-drawing the instance split 2000 times at seed 20260922 gives positive pooled gaps against Hyperband on every draw in every scenario. The same holds against random search except aknn, which is positive on $1994/2000$. The following table retains the complete successive-halving split diagnostics and the descriptive cellwise statistics.

	rpart	xgboost	super	aknn	ranger	svm	glmnet
ρ at r_1	$+0.400$	$+0.512$	$+0.532$	$+0.577$	$+0.606$	$+0.722$	$+0.862$
t, vs RS	9.63	4.05	6.91	3.84	8.21	5.17	3.87
t, vs Hyperband	10.48	5.38	7.36	8.43	12.21	4.43	3.48
t, vs SH	2.66	2.84	1.43	1.38	1.50	0.49	1.35
splits with SH gap > 0	1936	2000	1863	960	1298	1332	1711
median SH gap over draws	$+0.0112$	$+0.0165$	$+0.0053$	-0.0004	$+0.0033$	$+0.0048$	$+0.0048$
median generous cells	13	13	14	11	12	11	10

On the first pool, scenarios with weaker first-rung transfer tend to show more stable gains over successive halving. Across these seven scenarios and LCBench, Spearman correlations are -0.833 for archived gaps and -0.738 for split medians (exact permutation $p = 0.0154$ and 0.0458). This is a descriptive association across chosen scenarios.

A second pool draw at seed 771111 preserves all $35/35$ Hyperband wins, $33/35$ random-search wins, and $19/35$ successive-halving wins; paired aggregate wins are $87/105$ and generous wins $84/105$. The first-rung association does not replicate (eight-scenario correlation $+0.095$, $p = 0.8401$). The robust observation is the measured allocation performance, especially against Hyperband, rather than a universal mechanism determined by first-rung correlation alone.

Separating allocation from stopping. The SH2 allocation and its stopping-rule extension were evaluated on the same scenarios and ledger. Table entries below retain the full comparison: the first gap is SH2 regret minus SH regret, and the second is rule regret minus SH2 regret. The share divides the allocation gap by SH's mean regret.

scenario	ρ at r_1	SH2 − SH		rule − SH2		SH lower
		gap	share	gap	t	of 5
lcbench	+0.047	+0.0574	0.313	−0.0121	−0.82	5/5
rbv2_rpart	+0.400	+0.0807	0.531	−0.0358	−2.98	5/5
rbv2_xgboost	+0.512	+0.0488	0.317	−0.0346	−3.73	4/5
rbv2_super	+0.532	+0.0797	0.627	−0.0665	−4.90	4/5
rbv2_aknn	+0.577	+0.2269	0.820	−0.2509	−6.64	1/5
rbv2_ranger	+0.606	+0.1020	0.480	−0.0913	−6.63	4/5
rbv2_svm	+0.722	+0.0420	0.256	−0.0347	−2.53	3/5
rbv2_glmnet	+0.862	+0.0244	0.286	−0.0322	−3.92	0/5

The stopping step improves mean regret relative to SH2 on seven of the eight scenarios, with descriptive t values from -2.53 to -6.64; LCBench has the smaller -0.0121 gap at $t = -0.82$. Standalone SH is lower than the filled stopping variant in six scenarios; the rule leads on aknn and glmnet. The SH2 allocation deficit scales with scenario regret, not first-rung transfer (Pearson -0.0553 against transfer versus $+0.8815$ against mean regret). These controls separate the value of a stopping decision from the allocation schedule to which it is attached.

Matched-schedule promotion. When FIT selects the same (η, r_0) for ENV and SH, the comparison holds the schedule fixed. On rpart, all five budgets match: ENV wins three and ties two, with mean gap $+0.0149$. On xgboost, all three matched budgets favor ENV (mean $+0.0155$). On super, ENV wins three of four (mean $+0.0063$); on glmnet, the mean gap across three matched budgets is $+0.0040$. The matched comparisons favor SH on ranger and aknn (means -0.0026 and -0.0005), while svm has no matched schedule. Thus promotion itself improves four of the six scenarios with matched comparisons, with schedule choices also contributing to the aggregate results.

Fidelity ladder. Median rank correlation between a rung and target fidelity increases monotonically in all seven scenarios. The complete ladder is retained below.

rung (training-set fraction)	0.05	0.0625	0.1	0.125	0.2	0.3	0.5	1
rbv2_rpart	+0.400	+0.460	+0.618	+0.670	+0.806	+0.908	+0.963	1.000
rbv2_xgboost	+0.512	+0.541	+0.611	+0.659	+0.737	+0.808	+0.895	1.000
rbv2_super	+0.532	+0.566	+0.653	+0.722	+0.807	+0.865	+0.911	1.000
rbv2_aknn	+0.577	+0.623	+0.672	+0.711	+0.756	+0.785	+0.827	1.000
rbv2_ranger	+0.606	+0.661	+0.727	+0.761	+0.845	+0.892	+0.929	1.000
rbv2_svm	+0.722	+0.752	+0.794	+0.818	+0.871	+0.910	+0.947	1.000
rbv2_glmnet	+0.862	+0.874	+0.884	+0.892	+0.915	+0.928	+0.932	1.000

These correlations describe the fidelity structure of the reported pool. The protocol compares methods on paid observations and measures robustness through separate split, pool, and matched-schedule checks.

B.10 Reachable one-sided bracket

The reduction's bracket is written for both signs of S before either is used: with $B_d = -x_d$, $\mathrm{br}(A, B) = \{\lambda \in \mathbb{R} : \sigma_a S(A_d + \lambda B_d) \geq 0 \ \forall d \neq d_1,\ \lambda S > 0\}$ equals $\left[\max \mathcal{N},\ \lambda_{\mathrm{hi}}\right] \cap (0, \infty)$ on $S > 0$; equals $\left(-\infty,\ \min \mathcal{N}\right] \cap (-\infty, 0)$ on $S < 0$ when $\mathcal{R} = \emptyset$ and λ_0 holds; and is empty otherwise. Since $\lambda = \gamma^t \sigma_\varepsilon / (\sigma_a S)$ carries the sign of S, only the part of the interval on that side is reachable, which is what leaves (7) with one end on each branch. The condition $\mathcal{R} = \emptyset$ says that w_{d_1} is a maximal residual coordinate, the same event that makes λ_{hi} infinite on the other branch. The negative-S branch additionally requires the stated condition on tied residuals.

B.11 Full-ordering and capture certificates

On a row with $S > 0$ let $A_{(2)} = u_{d_1} - u_{d_{(2)}}$ and $g = \min_{1 \leq i < n}(u_{d_{(i)}} - u_{d_{(i+1)}})$ be the design's smallest adjacent gap, and with $\lambda^\star(A) = A / \sqrt{2n - A^2}$ as in Section 4.1. If g is beyond the band then the composite orders the arms exactly as the design orders them, for every $w \in \mathcal{W}$, with critical

fidelity $\lambda^\star(g)$ against $\lambda^\star(A_{(2)})$ for the capture. Since $g \le A_{(2)}$ the ordering is the more expensive claim; on $S < 0$ the same argument reverses every pair. The distinction is substantial: the ratio $\lambda^\star(A_{(2)})/\lambda^\star(g)$ has median 29.27 on walker and 38.36 on cheetah, only 1.4% of walker's adjacent gaps and 3.3% of cheetah's are beyond the band, and *no* row in either domain has its whole ordering certified — against 19 and 51 rows whose capture is.

B.12 Composition-chain coherence

Let $\mathcal{P}^{(L)} = \sum_{i\ge 2} g_i^2 \|\varepsilon_i\|^2$ denote the Pythagorean sum of the weighted residuals. The ratio $\|E^{(L)}\|^2/\mathcal{P}^{(L)}$ is above 1 at every depth in both domains with a pooled z of $+156$, $+512$ and $+759$ on walker through $m = L-1 = 2, 3, 4$ and $+172$, $+575$, $+851$ on cheetah, and medians rising from 1.504 to 1.911 and from 1.600 to 2.156. Writing τ for the total horizon of the legs strictly between a pair, the mean cosine decays monotonically in τ from $+0.675$ to $+0.073$ on walker and from $+0.823$ to $+0.221$ on cheetah, and leave-one-catalog-out predicts the held-out catalog's ratio with Pearson 0.656 over $62\,976$ walker cells and 0.870 over $47\,232$ cheetah cells — measurement rather than theory.

B.13 Reachable-state separation

The first-leg axis of the archive's tensor reads sixteen distinct trajectories rather than sixteen copies of one, and they are separated: at $t = 10$ the state reachable by one first leg differs from another's by 0.93 of what a different rollout root costs on walker, 0.63 on cheetah and 0.26 on pointmass. What the archive does not carry is a mixed closed-loop trace, since its t axis holds one first leg fixed throughout.

B.14 State-restoration checks

Every root here is a recorded state the pipeline restores and reads back, and the design is matched-root, so state and observation must be the same object. On walker and cheetah they are, the restored observation equalling the recorded one in $200/200$ roots and times at exactly $0.000\mathrm{e}{+00}$ in every coordinate. On quadruped they disagree in $200/200$ cases — a median per-root maximum of 6.28, a worst case of 130 — with the same twenty-six coordinates failing every time, all in blocks that read solver-dependent contact quantities the recorded state does not carry, and no choice of roots repairing it.