WINNERS NEED WITNESSES: CERTIFIED REPLAY SELECTION AND REPORTING WITH PAIRS

Anonymous authors
Paper under double-blind review

ABSTRACT

Selecting a good configuration and estimating its performance are distinct statistical tasks. We formulate joint selection and reporting: under a compute budget, a procedure returns an ε-optimal configuration and an η-accurate performance estimate with a joint certificate, or abstains. Shared randomness can make comparisons nearly noiseless while absolute-performance estimation remains costly. We establish a finite-pool minimax characterization and, for two candidates with known Gaussian covariance and complete-pair queries, instance-dependent bounds matching up to logarithmic factors. PAIRS combines paired sequential selection with protected reporting. A report-aware controller has a variance- and gap-dependent completion bound; paid variance learning can change the cheapest acceptable winner. Under the stated information-use assumptions, reports are selection-conditionally unbiased and incorrect joint certification has probability at most δ. Four replay benchmarks yield reference-relative $P_\varepsilon = 93.2\%$, versus 81.1% for uniform allocation at equal mean selection cost. In a 20-run Split-CIFAR-10 task-order study, fresh same-winner reporting reduces reference-relative macro absolute signed bias by 92%. A 1,000-pair application evaluation gives a 2.30 ratio between the smallest evaluated budgets meeting prespecified completion and correct-joint-output targets, with reporting and audit charged and abstentions counted as failures. A heterogeneous-variance study shows that additional comparisons can reduce total cost by shifting selection toward an easier-to-report acceptable candidate. PAIRS makes reporting precision and completion explicit objectives of replay selection.

1 MOTIVATION: A CORRECT WINNER CAN HAVE AN INCORRECT REPORT

Suppose twenty configurations have identical expected performance. Every selected configuration is optimal, yet reporting the largest observed score overstates its value. In 100,000 controlled trials with fifty independent Gaussian worlds per candidate and marginal standard deviation 0.1, winner optimism is 0.02639 native units. A nominal 95% known-variance interval covers the common mean only 60.49% of the time, despite 100% selection success. Appendix A derives the population quantities. Figure 1 separates selection success, winner optimism, and reporting coverage: choosing a correct winner and reporting its value reliably are different tasks.

Replay configurations change buffer admission, retrieval, replay ratios, latent layers, and schedules (Isele & Cosgun, 2018; Aljundi et al., 2019a;b; Buzzega et al., 2020; Shim et al., 2021; Pellegrini et al., 2019). Shared task orders and initialization reduce comparison noise, but reusing the winning score couples selection and stopping bias (Shin et al., 2021; Russo & Zou, 2016). Best-arm identification, common random numbers, and confidence sequences address selection, dependence, and repeated inspection (Gabillon et al., 2012; Kalyanakrishnan et al., 2012; Garivier & Kaufmann, 2016; Nelson & Matejcik, 1995; Kim & Nelson, 2001; Saad et al., 2023; Howard et al., 2021). Protected information is additionally needed for conditionally unbiased reports (Leiner et al., 2025; García Rasines & Young, 2023).

PAIRS makes completion, selection quality, and reporting precision joint objectives. Contrast information ranks candidates; level information anchors their absolute performance. We establish protected reporting, two-candidate Gaussian cost bounds, and bounded-pool minimax complexity.

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.

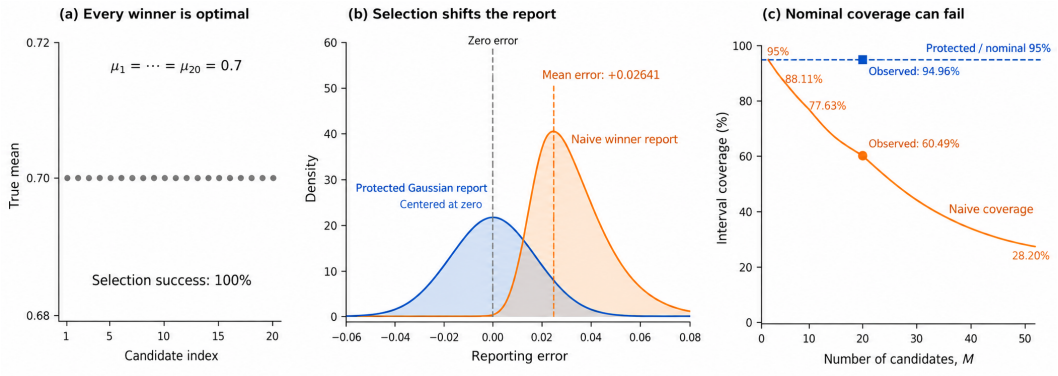


Figure 1: An optimal winner can have an unreliable report. (a) All twenty candidates are optimal. (b) Analytical reporting-error distributions for $n = 50$, $\sigma = 0.1$, and independent candidate outcomes. (c) Analytical coverage versus pool size; observed markers use 100,000 trials. Protected reporting remains centered and calibrated (Appendices A and G).

A report-aware controller and paid variance learning capture how reporting requirements change the cheapest acceptable winner. Experiments evaluate selection, same-winner bias, and total cost.

2 The Joint Task and Its Observable Output

Fix a benchmark, a target world distribution, and a finite pool $\mathcal{Q} = \{1, \ldots, M\}$ of executable configurations. Utility is $Y_t(i) = g(H_t(i))$ for taskwise record $H_t(i)$ and a map g fixed before selection; $\mu_i = \mathbb{E}Y_t(i)$ and $\mu_* = \max_i \mu_i$. Bounded accuracy uses percentage points, $Y_t(i) \in [0, 100]$; the separate Gaussian model uses its stated native units. The estimand is performance under reruns from the target world law.

For total budget B, a procedure returns

$$(\hat{i}, \widehat{\mu}_{\hat{i}}, I_{\hat{i}}, C_B, \mathcal{L}), \qquad C_B \in \{0, 1\}, \tag{1}$$

where $\mathcal{L}$ records computation and evidence use. Write $A_B = \{C_B = 1\}$ and define

$$G = \{\mu_* - \mu_{\hat{i}} \le \varepsilon,\ |\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}}| \le \eta\}. \tag{2}$$

The joint task has two requirements, error control and completion:

$$\Pr(A_B \cap G^c) \le \delta, \qquad \Pr(A_B) \ge \gamma_B. \tag{3}$$

Write $Q_B = \Pr(A_B)$, $J_B = \Pr(A_B \cap G)$, and $W_B = \Pr(A_B \cap G^c)$ for completion, correct joint output, and incorrect certification. Then $Q_B = J_B + W_B$, and Eq. (3) implies $J_B \ge \gamma_B - \delta$. Conditional error is $W_B/Q_B \le \delta/Q_B$ when $Q_B > 0$, not automatically δ. All three rates use every run; abstentions contribute no correct joint output. A cap-returned empirical leader without a certificate remains a diagnostic output.

Let $\mathcal{S}$ be the selection transcript: candidate generation, revealed selection observations, query and stopping decisions, permitted selector randomization, failures, and precommitted reporting design. For completed outputs, conditional reporting unbiasedness means

$$\mathbb{E}[\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}} \mid \mathcal{S}, A_B] = 0 \tag{4}$$

where $\Pr(A_B \mid \mathcal{S}) > 0$. This is distinct from interval coverage and the precision requirement in Eq. (2). On completion, the reporting radius is at most η. Completion must not reveal the protected report's value through either direct inspection or outcome-dependent cost signals.

Each action fixes candidates, a fresh world, coupling, and information release before observing outcomes. The ledger separates accepted evaluations, attempts, accelerator-hours, and wall-clock makespan across proposal, selection, reporting, gate, audit, and reference. Total-cost comparisons state their included stages.

3 Information for Ranking and Information for Reporting

3.1 Shared Noise Cancels in Contrasts, Not in Levels

Consider independent worlds with two candidates,

$$Y_{1,t} = \theta + \Delta/2 + \sigma_c Z_t + \sigma_e E_{1,t},$$
$$Y_{2,t} = \theta - \Delta/2 + \sigma_c Z_t + \sigma_e E_{2,t}, \tag{5}$$

where all base noises are independent standard normals. Throughout, σ_c denotes shared noise and σ_e candidate-specific noise; the marginal variance is $v = \sigma_c^2 + \sigma_e^2$. The difference $D_t = Y_{1,t} - Y_{2,t}$ and average $A_t = (Y_{1,t} + Y_{2,t})/2$ obey

$$D_t \sim N(\Delta, v_D), \quad A_t \sim N(\theta, v_A), \quad v_D = 2\sigma_e^2, \quad v_A = \sigma_c^2 + \sigma_e^2/2, \tag{6}$$

and are independent. As $\sigma_e \to 0$, comparisons become arbitrarily precise while level estimation retains noise σ_c^2. This elementary decomposition supplies the two directions used in the joint lower bound. Pairwise variance reduction alone therefore cannot predict absolute-reporting cost.

3.2 A Protected Contrast-Based Gaussian Witness

There is a structured alternative to assigning every observation to a separate bank. Let $X \sim N_M(\mu, V)$ with known positive-definite V, and let $C \in \mathbb{R}^{(M-1)\times M}$ have full row rank and span the contrast space. Put $H = CVC^\top$, draw $W \sim N(0, H)$ independently, and fix $\lambda > 0$. Define

$$F = CX + \sqrt{\lambda}W, \qquad G = X - \lambda^{-1/2} VC^\top H^{-1} W. \tag{7}$$

The selector receives F; a reporting service seals G. The raw X and randomization W are not available to the selector. This is a contrast-specific use of Gaussian information splitting (Leiner et al., 2025; García Rasines & Young, 2023).

Proposition 1 (Protected witness)**.** *Under Eq.* (7)*, $F \perp G$, $\mathbb{E}G = \mu$, and*

$$V_G = V + \lambda^{-1} VC^\top (CVC^\top)^{-1} CV. \tag{8}$$

For any $\hat{i} = \phi(F)$, $\mathbb{E}[G_{\hat{i}} - \mu_{\hat{i}} \mid F] = 0$. Relative to full-vector fission, the reduction in reporting covariance is the positive-semidefinite matrix

$$\lambda^{-1}\{V - VC^\top (CVC^\top)^{-1} CV\}. \tag{9}$$

For $V = \sigma_c^2 \mathbf{1}\mathbf{1}^\top + \sigma_e^2 I$,

$$(V_G)_{ii} = \sigma_c^2 + \sigma_e^2\{1 + (1 - 1/M)/\lambda\}. \tag{10}$$

Only contrast noise incurs the splitting penalty.

With independent complete worlds and fixed C, V, λ, apply Eq. (7) at every time. If the winner, stopping time $\tau \geq 1$, and all budgets are measurable using only the F transcript, then

$$\overline{G}_{\hat{i},\tau} \mid \mathcal{S} \sim N\left(\mu_{\hat{i}}, \frac{(V_G)_{\hat{i}\hat{i}}}{\tau}\right). \tag{11}$$

This conditional law gives a valid report after selection and stopping. Unknown covariance requires a separate construction; independent plug-in estimation is not exact Gaussian protection (Dharamshi et al., 2026).

4 Fundamental Limits of the Joint Task

We state the complexity result for Eq. (5), with known $\sigma_e > 0, \sigma_c \geq 0$, $|\Delta| > \varepsilon$, and the action set consisting of complete two-candidate worlds, each costing two evaluations. This fixes the observation model and cost on both sides of the comparison. Let N be the number of worlds requested, $g = |\Delta| + \varepsilon$, and $k_{\gamma,\delta} = \mathrm{kl}(\gamma - \delta, \delta)$ for $\gamma > 2\delta$.

Theorem 2 (Two sources of necessary computation)**.** *Any procedure satisfying Eq.* (3) *with completion at least γ throughout this model class obeys*

$$\mathbb{E}N \;\geq\; k_{\gamma,\delta} \max\left\{\frac{2v_D}{g^2}, \frac{v}{2\eta^2}\right\}, \qquad v = \sigma_c^2 + \sigma_e^2. \tag{12}$$

The evaluation-cost bound is twice Eq. (12)*.*

Two alternatives change the acceptable winner or its absolute mean. Their maximum avoids double-counting reusable observations; Appendix F gives the change-of-measure proof.

Theorem 3 (Achievable cost in the same action model)**.** *Gaussian confidence-sequence selection followed by a fresh reporting sample whose size is fixed before its outcomes are observed has joint error at most δ and, without a hard cap, completes almost surely. Its expected number of full worlds is*

$$O\left(1 + \frac{v_D}{g^2}\left[\log\frac{1}{\delta} + \log\left(e + \frac{v_D}{g^2}\right)\right] + \frac{\sigma_c^2 + \sigma_e^2}{\eta^2}\log\frac{1}{\delta}\right). \tag{13}$$

For fixed nonvanishing completion targets, this matches the instance-dependent scales in Eq. (12)*, up to logarithmic factors and the unavoidable cost of an initial observation.*

Appendix F gives the proof and budget-dependent completion bound. For iid $[0,1]^M$ worlds, unrestricted within-world dependence, and unit coordinate costs, the finite-pool joint minimax rate is

$$\mathcal{C}^* = \Theta\left[\left(M/\varepsilon^2 + 1/\eta^2\right)\log(1/\delta)\right]. \tag{14}$$

Theorem 17 gives its parameter range and uniform completion requirement; median elimination followed by fresh reporting attains it.

5 PAIRS: Allocate Comparisons and Protect the Report

5.1 Admissible Replay Queries

PAIRS freezes the proposal-generated pool, utility, coupling, and cost rules before racing (Appendix Q). Shared exogenous state includes split, task order, initialization primitives, augmentation, loader schedule, and evaluation sampling. Candidate-induced buffer, retrieval, and optimizer states remain separate, preserving each marginal target law.

Queries depend only on completed selection information; blocks complete before simultaneous elimination. Failed blocks use immutable same-world retries, score-independent whole-block deletion, or a prespecified failure score included in the estimand. Appendix C treats outcome-dependent partial queries.

5.2 Two Reporting Routes

For the general route, simultaneous pairwise confidence sequences satisfy

$$\Pr\{\exists n,\; i < j : \mu_i - \mu_j \notin [L_n(i,j), U_n(i,j)]\} \leq \delta_{\mathrm{sel}}. \tag{15}$$

Candidate i is removed when $U_n(i,j) < -\varepsilon$ for an active challenger j. A surviving empirical leader is selection-certified when

$$\max_{j \in A \setminus \{\hat{i}\}} U_n(j, \hat{i}) \leq \varepsilon. \tag{16}$$

All surviving pairs have legal observation prefixes. On the joint coverage event an optimum is never eliminated, so Eq. (16) certifies an ε-optimal candidate.

Before opening a fresh reporting bank, freeze its sample count and interval rule with conditional coverage at least $1 - \delta_{\mathrm{rep}}$ and radius at most η. Appendix M gives Gaussian and bounded-score counts, reserve, and audit. Independent variance pilots additionally charge acquisition cost and error budget (Proposition 12).

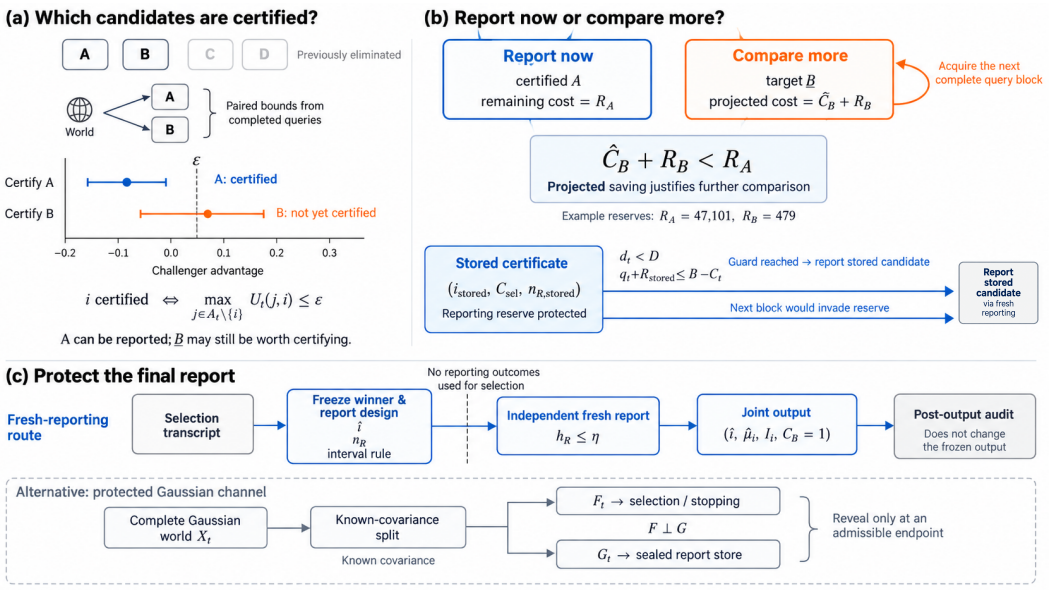


Figure 2: Joint decision and evidence flow. (a) Paired bounds identify certified candidates. (b) Further comparison protects a stored certificate and reserve, including eight audit observations; reaching the guard triggers fresh reporting. (c) Fresh reporting freezes the winner and report design before acquisition. Known-covariance Gaussian protection instead seals a reporting channel. Audit leaves the output unchanged.

The structured route uses complete Gaussian worlds and Eq. (7). Candidate elimination is logical: all coordinates continue to be evaluated because the protected-witness proof assumes complete vectors. Selection operates on F alone. For fixed λ, the reporting threshold is

$$n_R(i,\lambda) = \left\lceil z^2_{1-\delta_{\mathrm{rep}}/2}(V_G)_{ii}/\eta^2 \right\rceil . \tag{17}$$

Stop only when both Eq. (16) and $n \geq n_R(\hat{\imath},\lambda)$ hold. The report is then revealed once. This makes reporting precision part of the stopping rule without exposing reporting values.

Choose the route and fixed splitting parameter before production from declared costs, known parameters, or paid independent pilot evidence. Appendix L gives the comparison–reporting balancing objective. Certification still requires valid selection bounds and the reporting-precision reserve.

For bounded pools, periodic full-active-set queries give a report-aware completion bound with thresholds $\widetilde{O}(s^2_{ij}/g^2 + w/g)$ for paired variance s^2_{ij}, certification margin g, and range w (Theorem 21). A fresh variance pilot attains expected reporting cost $O((v_i/\eta^2 + w/\eta)\log(1/\delta))$ (Theorem 27). After a certificate is available, an economic guard can cap additional comparison cost at a fraction β of its reporting reserve, giving a pathwise $(1+\beta)$ cost bound relative to immediate reporting from the same history (Proposition 24). Figure 2 connects certification, report-aware continuation, and protected reporting.

Here C_t is charged cost, q_t the next whole-block cost, and d_t counts blocks since the first stored certificate. Its frozen report and audit reserve R_{stored} remains protected; the forecast $\widehat{C}_j$ guides continuation, never certification. The structured route stops on F alone; incomplete report acquisition returns $C_B = 0$.

Algorithm 1 PAIRS with guarded continuation and protected reporting

Require: Frozen pool, target law, budget B, ε, η, $\delta_{\mathrm{sel}} + \delta_{\mathrm{rep}} \le \delta$

Precommit route, query/report/cost/failure rules and continuation limit D
Initialize comparisons, sealed store and ledger; stored answer $\leftarrow \varnothing$
while true **do**
  if fresh route and an affordable certified answer exists **then**
    Choose the certified i with smallest reporting reserve, including the stored answer
    If none is stored, store i, its certificate and frozen design; protect R_{stored}, set $d_t = 0$
    if no uncertified target j has $\widehat{C}_j + R_j < R_i$ **then**
      Freeze i's report design; **return** Algorithm 2 for i
    end if
  end if
  Plan the next whole query and cost q_t from completed selection information
  if stored answer exists and ($d_t \ge D$, no query, or $q_t + R_{\mathrm{stored}} > B - C_t$) **then**
    return Algorithm 2 with the stored answer and frozen design
  end if
  If no complete admissible query is affordable, **return** $C_B = 0$ and the ledger
  Execute the whole block; charge all attempts and apply the failure rule
  Reveal selection information, seal protected information, update and eliminate simultaneously
  If an answer was stored before this block, increment d_t
  if structured route, selection certified, $n \ge n_R(\hat{i}, \lambda)$, and reserves fit **then**
    Freeze output rule; reveal the protected report and **return** $C_B = 1$
  end if
end while

Theorem 4 (Joint certificate and conditional report)**.** *Let $\mathcal{T}$ include selection, the reporting design, and all failure and cost signals determining completion, with $A_B \in \mathcal{T}$. Assume conditional reporting bias zero and error probability at most δ_{rep} on every completing transcript. Under Eq.* (15)*, Algorithm 1 satisfies Eq.* (4) *and*

$$J_B \ge (1 - \delta_{\mathrm{rep}})(Q_B - \delta_{\mathrm{sel}})_+, \quad W_B \le Q_B - (1 - \delta_{\mathrm{rep}})(Q_B - \delta_{\mathrm{sel}})_+. \tag{18}$$

In particular, $W_B \le \delta_{\mathrm{sel}} + \delta_{\mathrm{rep}}$, including at a hard cap with abstention.

Conditioning on the completing transcript yields the envelope. Under iid complete worlds, available observations, and unit acquisition costs, Proposition 11 gives a sufficient fixed-allocation budget with certain completion; the experiments measure adaptive completion at smaller caps.

6 Evidence: Selection, Reporting, and Total Cost

6.1 Replay Selection Evidence

The replay study contains 190 outer replicates: 60 for five-task Split-CIFAR-10, 50 for ten-task Split-CIFAR-100, 40 for CORe50 NICv2-style latent replay, and 40 for a frozen RS-MCTS schedule pool. Candidate counts are 16, 24, 12, and 32. Learner families are ER/DER++, ER/ASER, latent replay, and the released schedule selector (Krizhevsky, 2009; Lomonaco & Maltoni, 2017; Chaudhry et al., 2019; Buzzega et al., 2020; Shim et al., 2021; Pellegrini et al., 2019). Selection uses $\varepsilon = 0.20$ pp and $\delta_{\mathrm{sel}} = 0.05$. Appendix Q gives candidate coordinates and final-task-average utilities; the complete schedule matrices appear in Appendix Q.2.

PAIRS improves reference-relative P_ε in every setting at the same mean selection count (Table 1); the pooled difference is 12.1 pp. Uniform budgets $B_u = 31, 46, 24, 62$ are fixed before the evaluated race: each candidate receives $\lfloor B_u/M \rfloor$ observations, with the remainder assigned by an outcome-independent permutation (Appendix N.4). PAIRS uses 35.4% fewer evaluations than fixed $k = 3$; fixed $k = 5$ reaches 95.8% at 103.2 evaluations. Appendix N presents all baselines, tail regret, selection GPU-hours, and the joint-count analysis of certified outputs and paired contrasts.

Table 1: Replay selection over all returned incumbents, including uncertified cap outputs. P_ε and exact-best agreement (EB) are finite-reference percentages; regret is in pp. Evals count accepted selection observations. Uncert. is the percentage of cap outputs without a selection certificate ($13/190$ for PAIRS); uniform allocation returns a candidate without a certification endpoint (denoted –).

Setting	Method	P_ε	EB	Regret	Evals	Uncert.
CIFAR-10	Uniform	86.7	71.7	.11	31	–
	PAIRS	96.7	85.0	.05	31	3.3
CIFAR-100	Uniform	80.0	54.0	.19	46	–
	PAIRS	92.0	70.0	.09	46	8.0
CORe50	Uniform	85.0	65.0	.14	24	–
	PAIRS	95.0	80.0	.07	24	5.0
RS-MCTS	Uniform	70.0	42.5	.27	62	–
	PAIRS	87.5	60.0	.13	62	12.5
Pooled	Uniform	81.1	59.5	.171	40.0	–
	PAIRS	93.2	74.7	.082	40.0	6.8
Macro	Uniform	80.4	58.3	.178	40.75	–
	PAIRS	92.8	73.8	.085	40.75	7.2

6.2 Mechanism and Calibration

Full state-complete pairing reduces the observed difference-variance ratio to 0.33 [0.27,0.40] and the controlled racing-cost ratio to 0.61 [0.54,0.69]. Sharing only splits gives a variance ratio of 0.81; adding task order and initialization gives 0.62 and 0.45. Pairing therefore improves comparison efficiency. Absolute-reporting cost also contains the shared-level variance in Eq. (6).

The calibration experiment uses 10,000 repetitions per row. Practical and robust CS coverage is 95.1% [94.7,95.5] and 94.8% [94.3,95.2], respectively; repeatedly inspected fixed-time intervals cover at 84.2% and 78.1%. These calibration endpoints assess repeated inspection. The bounded and Gaussian experiments below evaluate complete selection-and-reporting procedures. In the correlation study, zero correlation gives 44 versus 45 evaluations, negative correlation gives 51 versus 45, and a near tie requires 99 evaluations. The efficiency benefit depends on positive shared variation and resolvable candidate gaps.

6.3 Application-Level Reporting and Joint Cost

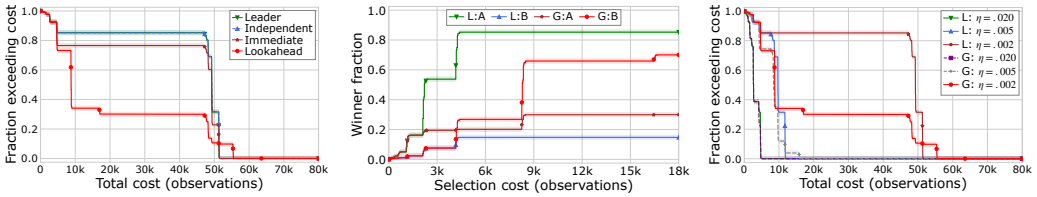
We evaluate the joint application protocol in a Split-CIFAR-10 task-order study with four task orders, four ER configurations, and five outer runs per order (Appendix V). The protocol uses $\varepsilon = 0.20$ pp, $\eta = 1.00$ pp, and $(\delta_{\mathrm{sel}}, \delta_{\mathrm{rep}}) = (0.02, 0.03)$. All 20 paired-selection outputs have both a naive selection-reused report and a fresh same-winner report. Equal-setting macro absolute signed bias decreases from 0.875 to 0.070 pp, a 92% reduction; macro MAE decreases by 69.6% (Table 2). Appendix E defines the protocol and estimands.

Paired selection (P) and independent selection (I) both complete all 20 runs, with reference-v2 joint correctness $19/20$ and $18/20$. Their selection, reporting, and audit ledgers contain 2,601,620 and 5,983,726 attempted observations, giving an I/P ratio of 2.30 (paired-resampling 95% range $[2.271, 2.329]$). The joint-correctness difference is $+5$ pp, with range $[-10, 20]$ pp. The additional 190-run record supplies reporting and acquisition diagnostics (Appendix S).

The application budget evaluation compares PEBR-FR-Paired and ε-LUCB-FR-Independent over 1,000 paired runs, with identical fresh reporting. At its recorded native-unit targets $(\varepsilon, \eta) = (.05, .10)$, the smallest qualifying evaluated budgets are 10,000 and 23,000, a ratio of 2.30. Qualifying means completion at least .95 and correct joint output at least .90, counting abstentions as failures. Simultaneous inference over all budget comparisons gives a joint-success difference lower bound of -1.249 pp, above the -2 pp noninferiority margin. The mean actual-attempt ratio is 1.99821. Appendix W gives the error allocation and distinguishes this budget endpoint from the 20-run total-attempt endpoint.

Table 2: Joint application results from the 20-run Split-CIFAR-10 task-order study (Appendix V). Bias and MAE compare fresh versus naive reports on the same P-selected winners. Completion, reference-v2 joint correctness, and cost compare paired (P) versus independent (I) selection, both with fresh reporting. Attempts include selection failures, reporting, and audit.

Same-winner reporting endpoint	Fresh	Naive	Contrast
Macro absolute signed bias (pp)	0.070	0.875	-92.0%
Macro MAE (pp)	0.266	0.875	-69.6%
Joint-procedure endpoint	**P**	**I**	**Contrast**
Completion	20/20	20/20	–
Reference-scored joint correctness	19/20	18/20	$+5$ pp
Total attempted observations	2,601,620	5,983,726	I/P $= 2.300$



(a) Cost comparison (b) Acceptable-winner choice (c) Reporting-precision check

Figure 3: Cost distributions and acceptable-winner choice. (a) Total-cost survival curves at $\eta = .002$. (b) Cumulative fractions by selection cost and winner. (c) Paired leader (L) and guarded lookahead (G) across reporting tolerances. Each curve uses 10,000 runs at cap 200,000, $\varepsilon = .020$, and $(\delta_{\text{sel}}, \delta_{\text{rep}}) = (.02, .03)$, including eight audit observations. Shading: simultaneous 95% DKW bands over 12 distinct distributions.

6.4 More Comparisons Can Reduce Joint Cost

In the four-candidate CIFAR allocation study, A and B satisfy $\varepsilon = .020$ in native units, with marginal standard deviations .200 and .020. At $\eta = .002$, their fresh reports require 47,093 and 471 observations. Guarded lookahead selects B in 7,003 of 10,000 runs, versus 1,474 for paired leader, buying 3,217.00 additional selection observations to save 25,777.30 report observations (Table 3). The total-cost ratio is 2.1065 (paired-bootstrap 95% interval $[2.0681, 2.1464]$); the joint-success difference is $+0.08$ pp with one-sided 95% lower bound -0.433 pp. Relative to paired immediate, continuation reduces cost by a further factor of 1.9075. Leader and lookahead retain 100% observed ε-success while exact-best frequency falls from 85.26% to 29.97%: the joint objective rewards the cheaper acceptable answer. Appendix Z.9 gives the integer ledger and budget frontier, and Propositions 29–30 establish a separation with paid variance learning. Figure 3 shows the cost distributions and the shift toward the less expensive acceptable winner.

At separately executed budgets (Figure 4), simultaneous lower bounds first meet $Q \geq .95$ and $J \geq .90$ at 20,000 for guarded lookahead and 60,000 for paired leader: a 3.0 evaluated-grid budget ratio.

6.5 Controlled Selection and Reporting Studies

At correlation 0.9, contrast-specific protection reduces interval width by 26.0% for the same selection information. In 50,000 sequential trials at $\rho = .67$ and a 500-world cap, $(Q, J, W) = (99.74, 98.13, 1.61)\%$. Figure 5 and Appendices I–K give the assumption tests, bounded sequential paths, and finite-reference analysis. Execution records and disjoint reporting, audit, and gate banks are specified in Appendix M.

Table 3: Report-aware allocation at $B = 200{,}000$: 10,000 runs per policy, $\varepsilon = .020$, $\eta = .002$, $(\delta_s, \delta_r) = (.02, .03)$. Every policy completes; total includes eight audit observations per run. Leader and lookahead use the same paired comparison and fresh-report rules.

Policy	Selection	Reporting	Total	J (%)
Independent leader	2,738.02	40,206.93	42,952.95	96.89
Paired leader	2,720.73	40,220.92	42,949.65	96.88
Paired immediate	2,720.73	36,164.80	38,893.53	96.87
Paired guarded lookahead	5,937.73	14,443.61	20,389.34	96.96

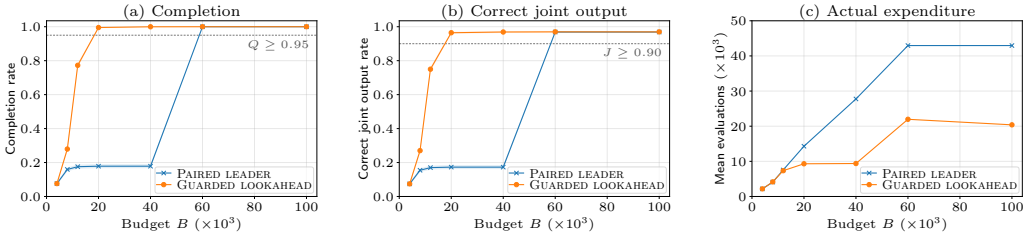


Figure 4: Executed budget frontier at $\eta = .002$, with other targets as in Figure 3. Each method–budget setting uses 10,000 runs, including abstentions in all denominators and costs. Shading: simultaneous 95% Clopper–Pearson intervals over 28 probabilities. Dashed lines mark $Q = .95$ and $J = .90$; connecting segments guide the eye. Audit is charged on completed runs.

7 Related Work

Pure exploration identifies acceptable arms (Gabillon et al., 2012; Kalyanakrishnan et al., 2012; Garivier & Kaufmann, 2016; Degenne & Koolen, 2019); covariance-adaptive identification, sequential ranking, and common random numbers exploit dependent comparisons (Saad et al., 2023; Nelson & Matejcik, 1995; Kim & Nelson, 2001; Görder & Kolonko, 2019). PAIRS additionally targets absolute reporting and budget-dependent completion. Confidence sequences support adaptive stopping and off-policy inference (Howard et al., 2021; Karampatziakis et al., 2021; Waudby-Smith et al., 2024). Our checkpoint rule uses finite-sample error spending rather than asymptotic monitoring (Maharaj et al., 2023); variance-adaptive and robust boundaries retain their own assumptions (Maurer & Pontil, 2009; Audibert et al., 2009; Wang & Ramdas, 2023).

Data fission protects inference after selection (Leiner et al., 2025; García Rasines & Young, 2023; Dharamshi et al., 2026). Related targets include winner inference, directional claims, inference-aware policy choice, and tune-then-deploy inference (Andrews et al., 2024; Heller & Solari, 2024; Bastani et al., 2025; Xu et al., 2026). Predictive-model-assisted sequences, proxy-assisted identification, and selective-conformal certificates address different information and acceptance rules (Kilian et al., 2026; Ao et al., 2026; Yu & Liu, 2026); PAIRS protects a frozen winner's report, with audit occurring after the output.

Multi-fidelity optimization allocates training effort (Jamieson & Talwalkar, 2016; Li et al., 2018; 2020; Falkner et al., 2018; Wu et al., 2020); replay methods optimize stored examples, retrieval, or representations (Hao et al., 2023; Ostapenko et al., 2022). PAIRS couples replay comparisons with certified absolute reporting.

8 Conclusion

PAIRS makes selection, reporting precision, and completion one budgeted objective. Its bounds separate comparison and absolute-performance costs; its application results connect protected reporting to lower bias and joint cost. Additional comparisons can save computation by finding a cheaper-to-report acceptable winner.

AI USE STATEMENT

Generative AI tools assisted with manuscript reconstruction, language editing, literature retrieval, mathematical derivations and proof drafts, and source compilation. The authors directed the research and provided the experimental records. They take responsibility for the derivations, citations, implementations, analyses, and conclusions.

REFERENCES

Rahaf Aljundi, Eugene Belilovsky, Tinne Tuytelaars, Laurent Charlin, Massimo Caccia, Min Lin, and Lucas Page-Caccia. Online continual learning with maximal interfered retrieval. In *Advances in Neural Information Processing Systems*, volume 32, 2019a. URL https://proceedings.neurips.cc/paper/2019/hash/15825aee15eb335cc13f9b559f166ee8-Abstract.html.

Rahaf Aljundi, Min Lin, Baptiste Goujaud, and Yoshua Bengio. Gradient based sample selection for online continual learning. In *Advances in Neural Information Processing Systems*, volume 32, 2019b. URL https://proceedings.neurips.cc/paper/2019/hash/e562cd9c0768d5464b64cf61da7fc6bb-Abstract.html.

Isaiah Andrews, Toru Kitagawa, and Adam McCloskey. Inference on winners. *The Quarterly Journal of Economics*, 139(1):305–358, 2024. doi: 10.1093/qje/qjad043. URL https://academic.oup.com/qje/article/139/1/305/7276491.

Ruicheng Ao, Hongyu Chen, Siyang Gao, Hanwei Li, and David Simchi-Levi. Best arm identification with LLM judges and limited human audits. arXiv preprint arXiv:2601.21471, 2026. URL https://arxiv.org/abs/2601.21471.

Jean-Yves Audibert, Rémi Munos, and Csaba Szepesvári. Exploration–exploitation tradeoff using variance estimates in multi-armed bandits. *Theoretical Computer Science*, 410(19):1876–1902, 2009. doi: 10.1016/j.tcs.2009.01.016. URL https://doi.org/10.1016/j.tcs.2009.01.016.

Hamsa Bastani, Osbert Bastani, and Bryce McLaughlin. Beating the winner's curse via inference-aware policy optimization. arXiv preprint arXiv:2510.18161, 2025. URL https://arxiv.org/abs/2510.18161.

Pietro Buzzega, Matteo Boschini, Angelo Porrello, Davide Abati, and Simone Calderara. Dark experience for general continual learning: A strong, simple baseline. In Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin (eds.), *Advances in Neural Information Processing Systems*, volume 33, pp. 15920–15930. Curran Associates, Inc., 2020. URL https://proceedings.neurips.cc/paper/2020/hash/b704ea2c39778f07c617f6b7ce480e9e-Abstract.html.

Arslan Chaudhry, Marcus Rohrbach, Mohamed Elhoseiny, Thalaiyasingam Ajanthan, Puneet K. Dokania, Philip H. S. Torr, and Marc'Aurelio Ranzato. On tiny episodic memories in continual learning. arXiv preprint arXiv:1902.10486, 2019. URL https://arxiv.org/abs/1902.10486.

Rémy Degenne and Wouter M. Koolen. Pure exploration with multiple correct answers. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://papers.nips.cc/paper_files/paper/2019/hash/60cb558c40e4f18479664069d9642d5a-Abstract.html.

Ameer Dharamshi, Anna Neufeld, Lucy L. Gao, Jacob Bien, and Daniela Witten. Decomposing Gaussians with unknown covariance. *Biometrika*, 113(1):asaf057, 2026. doi: 10.1093/biomet/asaf057. URL https://academic.oup.com/biomet/article/113/1/asaf057/8209918.

Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. *Journal of Machine Learning Research*, 7(39):1079–1105, 2006. URL https://www.jmlr.org/papers/v7/evendar06a.html.

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. PMLR, 2018. URL https://proceedings.mlr.press/v80/falkner18a.html.

Victor Gabillon, Mohammad Ghavamzadeh, and Alessandro Lazaric. Best arm identification: A unified approach to fixed budget and fixed confidence. In *Advances in Neural Information Processing Systems*, volume 25, pp. 3221–3229, 2012. URL https://proceedings.neurips.cc/paper_files/paper/2012/hash/8b0d268963dd0cfb808aac48a549829f-Abstract.html.

Daniel García Rasines and G. Alastair Young. Splitting strategies for post-selection inference. *Biometrika*, 110(3):597–614, 2023. doi: 10.1093/biomet/asac070. URL https://academic.oup.com/biomet/article/110/3/597/6955603.

Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Proceedings of the 29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pp. 998–1027. PMLR, 2016. URL https://proceedings.mlr.press/v49/garivier16a.html.

Björn Görder and Michael Kolonko. Ranking and selection: A new sequential Bayesian procedure for use with common random numbers. *ACM Transactions on Modeling and Computer Simulation*, 29(1):2:1–2:24, 2019. doi: 10.1145/3241042. URL https://doi.org/10.1145/3241042.

Jie Hao, Kaiyi Ji, and Mingrui Liu. Bilevel coreset selection in continual learning: A new formulation and algorithm. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/a0251e494a7e75d59e06d37e646f46b7-Abstract-Conference.html.

Ruth Heller and Aldo Solari. Simultaneous directional inference. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(3):650–670, 2024. doi: 10.1093/jrsssb/qkad137. URL https://academic.oup.com/jrsssb/article/86/3/650/7499129.

Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021. doi: 10.1214/20-AOS1991. URL https://doi.org/10.1214/20-AOS1991.

David Isele and Akansel Cosgun. Selective experience replay for lifelong learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), 2018. doi: 10.1609/aaai.v32i1.11595. URL https://ojs.aaai.org/index.php/AAAI/article/view/11595.

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

Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone. PAC subset selection in stochastic multi-armed bandits. In *Proceedings of the 29th International Conference on Machine Learning*, pp. 655–662. Omnipress, 2012. URL https://icml.cc/2012/papers/359.pdf.

Nikos Karampatziakis, Paul Mineiro, and Aaditya Ramdas. Off-policy confidence sequences. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 5301–5310. PMLR, 2021. URL https://proceedings.mlr.press/v139/karampatziakis21a.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 https://www.jmlr.org/papers/v17/kaufman16a.html.

Valentin Kilian, Stefano Cortinovis, and François Caron. Asymptotically log-optimal Bayes-assisted confidence sequences for bounded means. arXiv preprint arXiv:2605.07964, 2026. URL https://arxiv.org/abs/2605.07964.

Seong-Hee Kim and Barry L. Nelson. A fully sequential procedure for indifference-zone selection in simulation. *ACM Transactions on Modeling and Computer Simulation*, 11(3):251–273, 2001. doi: 10.1145/502109.502111. URL https://doi.org/10.1145/502109.502111.

Marcus Klasson, Hedvig Kjellström, and Cheng Zhang. Learn the time to learn: Replay scheduling in continual learning. *Transactions on Machine Learning Research*, 2023. URL https://openreview.net/forum?id=Q4aAITDgdP.

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

James Leiner, Boyan Duan, Larry Wasserman, and Aaditya Ramdas. Data fission: Splitting a single data point. *Journal of the American Statistical Association*, 120(549):135–146, 2025. doi: 10.1080/01621459.2023.2270748. URL https://www.tandfonline.com/doi/full/10.1080/01621459.2023.2270748.

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/hash/a06f20b349c6cf09a6b171c71b88bbfc-Abstract.html.

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://www.jmlr.org/papers/v18/16-558.html.

Vincenzo Lomonaco and Davide Maltoni. CORe50: A new dataset and benchmark for continuous object recognition. In Sergey Levine, Vincent Vanhoucke, and Ken Goldberg (eds.), *Proceedings of the 1st Annual Conference on Robot Learning*, volume 78 of *Proceedings of Machine Learning Research*, pp. 17–26. PMLR, 2017. URL https://proceedings.mlr.press/v78/lomonaco17a.html.

Akash Maharaj, Ritwik Sinha, David Arbour, Ian Waudby-Smith, Simon Z. Liu, Moumita Sinha, Raghavendra Addanki, Aaditya Ramdas, Manas Garg, and Viswanathan Swaminathan. Anytime-valid confidence sequences in an enterprise A/B testing platform. In *Companion Proceedings of the ACM Web Conference 2023*, pp. 396–400. Association for Computing Machinery, 2023. doi: 10.1145/3543873.3584635. URL https://doi.org/10.1145/3543873.3584635.

Andreas Maurer and Massimiliano Pontil. Empirical Bernstein bounds and sample-variance penalization. In *Proceedings of the 22nd Annual Conference on Learning Theory*, 2009. URL https://www.learningtheory.org/colt2009/papers/012.pdf.

Barry L. Nelson and Frank J. Matejcik. Using common random numbers for indifference-zone selection and multiple comparisons in simulation. *Management Science*, 41(12):1935–1945, 1995. doi: 10.1287/mnsc.41.12.1935. URL https://doi.org/10.1287/mnsc.41.12.1935.

Anna Neufeld, Ameer Dharamshi, Lucy L. Gao, and Daniela Witten. Data thinning for convolution-closed distributions. *Journal of Machine Learning Research*, 25(57):1–35, 2024. URL https://www.jmlr.org/papers/v25/23-0446.html.

Oleksiy Ostapenko, Timothee Lesort, Pau Rodriguez, Md Rifat Arefin, Arthur Douillard, Irina Rish, and Laurent Charlin. Continual learning with foundation models: An empirical study of latent replay. In *Proceedings of the 1st Conference on Lifelong Learning Agents*, volume 199 of *Proceedings of Machine Learning Research*, pp. 60–91. PMLR, 2022. URL https://proceedings.mlr.press/v199/ostapenko22a.html.

Lorenzo Pellegrini, Gabriele Graffieti, Vincenzo Lomonaco, and Davide Maltoni. Latent replay for real-time continual learning. arXiv preprint arXiv:1912.01100, 2019. URL https://arxiv.org/abs/1912.01100.

Daniel Russo and James Zou. Controlling bias in adaptive data analysis using information theory. In *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pp. 1232–1240. PMLR, 2016. URL https://proceedings.mlr.press/v51/russo16.html.

El Mehdi Saad, Gilles Blanchard, and Nicolas Verzelen. Covariance-adaptive best arm identification. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/e82ef7865f29b40640f486bbbe7959a7-Abstract-Conference.html.

Dongsub Shim, Zheda Mai, Jihwan Jeong, Scott Sanner, Hyunwoo Kim, and Jongseong Jang. Online class-incremental continual learning with adversarial Shapley value. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(11):9630–9638, 2021. doi: 10.1609/aaai.v35i11.17159. URL https://ojs.aaai.org/index.php/AAAI/article/view/17159.

Jaehyeok Shin, Aaditya Ramdas, and Alessandro Rinaldo. On the bias, risk, and consistency of sample means in multi-armed bandits. *SIAM Journal on Mathematics of Data Science*, 3(4): 1278–1300, 2021. doi: 10.1137/20M1361249. URL https://epubs.siam.org/doi/abs/10.1137/20M1361249.

Hongjian Wang and Aaditya Ramdas. Huber-robust confidence sequences. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206 of *Proceedings of Machine Learning Research*, pp. 9662–9679. PMLR, 2023. URL https://proceedings.mlr.press/v206/wang23p.html.

Ian Waudby-Smith, Lili Wu, Aaditya Ramdas, Nikos Karampatziakis, and Paul Mineiro. Anytime-valid off-policy inference for contextual bandits. *ACM / IMS Journal of Data Science*, 1(3):1–42, 2024. doi: 10.1145/3643693. URL https://doi.org/10.1145/3643693.

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 Conference on Uncertainty in Artificial Intelligence*, volume 115 of *Proceedings of Machine Learning Research*, pp. 788–798. PMLR, 2020. URL https://proceedings.mlr.press/v115/wu20a.html.

Yang Xu, Jiefu Zhang, Haixiang Sun, Zihan Zhou, Tianyu Cao, and Vaneet Aggarwal. Towards reliable LLM evaluation: Correcting the winner's curse in adaptive benchmarking. arXiv preprint arXiv:2605.05973, 2026. URL https://arxiv.org/abs/2605.05973.

Xiaoli Yu and Jiamiao Liu. A joint finite-sample certificate for adaptive selective conformal risk control. arXiv preprint arXiv:2606.08517, 2026. URL https://arxiv.org/abs/2606.08517.

A Why Perfect Selection Does Not Validate a Winner's Report

Let $\overline{X}_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2/n)$ for $i = 1, \ldots, M$, and select $\hat{i} = \arg\max_i \overline{X}_i$ with a deterministic tie rule. Every candidate is optimal, so selection success is one for any nonnegative tolerance. Writing $Z_{(M)} = \max_i Z_i$ for independent standard normals,

$$\mathbb{E}[\overline{X}_{\hat{i}} - \mu] = \frac{\sigma}{\sqrt{n}} \int_{-\infty}^{\infty} zM\phi(z)\Phi(z)^{M-1}\, dz. \tag{A.1}$$

A usual known-variance interval centered at the selected sample mean is $[\overline{X}_{\hat{i}} \pm z_{.975}\sigma/\sqrt{n}]$. Its exact coverage is

$$\Pr\{-z_{.975} \le Z_{(M)} \le z_{.975}\} = .975^M - .025^M. \tag{A.2}$$

For $M = 20, n = 50, \sigma = .1$, numerical integration gives expected optimism 0.02641008557 and the coverage formula gives 60.268768%. The 100,000-repetition experimental row reports 0.02639 and 60.49%; these are empirical estimates rather than the exact constants. The score scale in this example is a native Gaussian scale, not an accuracy percentage-point scale. With equicorrelated candidate averages, the common noise does not affect the winning index, and expected optimism is multiplied by $\sqrt{1-\rho}$ at fixed marginal variance.

B Proofs for Protected Reporting

Proof of Proposition 1. Let $P = VC^\top(CVC^\top)^{-1}CV$ and $H = CVC^\top$. Independence of X and W implies

$$\mathrm{Cov}(F, G) = CV - CVC^\top H^{-1}CV = 0.$$

Since (F, G) is jointly Gaussian, F and G are independent. Direct expansion gives $\mathbb{E}G = \mu$ and $\mathrm{Var}\, G = V + \lambda^{-1}P$. Conditional on F, the selected coordinate is fixed, so $\mathbb{E}[G_{\hat{i}} - \mu_{\hat{i}} \mid F] = 0$. For the positive-semidefinite comparison, write $P = V^{1/2}\Pi V^{1/2}$, where $\Pi = V^{1/2}C^\top(CVC^\top)^{-1}CV^{1/2}$ is an orthogonal projection. Then $V - P = V^{1/2}(I - \Pi)V^{1/2} \succeq 0$.

If $V = \sigma_c^2 \mathbf{1}\mathbf{1}^\top + \sigma_e^2 I$ and $C\mathbf{1} = 0$ with rank $M - 1$, then $P = \sigma_e^2(I - \mathbf{1}\mathbf{1}^\top/M)$. Taking diagonal entries proves Eq. (10). Full-vector fission instead gives reporting covariance $(1 + 1/\lambda)V$. Its contrast projection can be coupled to the same F selection channel, so the comparison need not change the winner. □

Proposition 5 (Contrast witnesses at a selection stopping time)**.** *Let the complete vectors X_t be independent $N_M(\mu, V)$, and construct (F_t, G_t) by Eq. (7) with fixed C, V, λ. Let $\mathcal{S}$ be generated by the F stream and independent selector randomness. If $\hat{i}$, an almost-surely finite positive integer R, and R distinct retained indices are $\mathcal{S}$-measurable, then the mean of the retained G values at coordinate $\hat{i}$, conditional on $\mathcal{S}$, is normal with mean $\mu_{\hat{i}}$ and variance $(V_G)_{\hat{i}\hat{i}}/R$.*

Proof. Proposition 1 and independence across worlds imply that the entire F stream is independent of the entire G stream. Conditional on $\mathcal{S}$, the coordinate, count, and distinct indices are fixed. The selected G vectors retain their independent Gaussian laws. Their average has the asserted distribution. A completion event determined by $\mathcal{S}$ does not change this conditional statement. □

This proposition includes $R = \tau$ and indices $1, \ldots, \tau$ for any finite F-only stopping time. The condition is on the complete transcript, including cost and failure signals used by the selector. If such signals depend on raw X or protected G beyond F, they require a separate independence argument.

B.1 An explicit selection boundary for the Gaussian route

For each unordered pair allocate $\delta_{ij} > 0$ with $\sum_{i<j} \delta_{ij} \le \delta_{\text{sel}}$. Because the row space of C is the contrast space, a vector a_{ij} satisfies $a_{ij}^\top C = e_i^\top - e_j^\top$. The released observations $Z_{ij,t} = a_{ij}^\top F_t$ have

mean $\mu_i - \mu_j$ and known variance $s_{ij}^2 = (1+\lambda)(e_i - e_j)^\top V(e_i - e_j)$. The explicit time-uniform boundary

$$r_{ij,n} = \sqrt{\frac{2s_{ij}^2}{n} \log \frac{\pi^2 n^2}{3\delta_{ij}}}, \qquad [L_n, U_n] = [\overline{Z}_{ij,n} - r_{ij,n}, \overline{Z}_{ij,n} + r_{ij,n}] \tag{B.1}$$

satisfies Eq. (15). Indeed the Gaussian two-sided tail probability is at most $6\delta_{ij}/(\pi^2 n^2)$; summing over time and pairs gives δ_{sel}. This newly specified Gaussian construction is distinct from the practical or robust boundaries in the earlier replay tables.

Proof of Theorem 4. Let E_s denote simultaneous pairwise coverage. An optimal candidate i^* has $\mu_{i^*} - \mu_j \geq 0$ for every challenger, so it is not eliminated on E_s. If selection stops with Eq. (16), then $\mu_* - \mu_{\hat{i}} \leq U_n(i^*, \hat{i}) \leq \varepsilon$; the singleton case is immediate. Hence a selection-certified error lies in E_s^c.

Let $\mathcal{T}$ contain $\mathcal{S}$, the frozen reporting count, and all failure, retention, and cost signals used to determine completion. The completed, selection-correct event $H = A_B \cap \{\mu_* - \mu_{\hat{i}} \leq \varepsilon\}$ is $\mathcal{T}$-measurable. By the first part, $\Pr(H) \geq (Q_B - \delta_{\mathrm{sel}})_+$. Under the assumed reporting law,

$$J_B = \mathbb{E}[\mathbf{1}_H \Pr\{|\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}}| \leq \eta \mid \mathcal{T}\}] \geq (1 - \delta_{\mathrm{rep}}) \Pr(H).$$

This proves the first bound in Eq. (18); $W_B = Q_B - J_B$ proves the second. For $Q_B \geq \delta_{\mathrm{sel}}$, it is at most $\delta_{\mathrm{sel}} + \delta_{\mathrm{rep}} Q_B - \delta_{\mathrm{sel}}\delta_{\mathrm{rep}} \leq \delta_{\mathrm{sel}} + \delta_{\mathrm{rep}}$; for $Q_B < \delta_{\mathrm{sel}}$, use $W_B \leq Q_B$. Conditional unbiasedness given $\mathcal{T}$ on A_B implies $\mathbb{E}[\mathbf{1}_{A_B}(\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}}) \mid \mathcal{S}] = 0$ and hence Eq. (4) by iterated expectation. The product is defined as zero on noncompleting outputs, where a report need not exist. Fresh reporting and the protected Gaussian route have this law when completion-related signals reveal no additional reporting information. A pilot variance bound with failure probability δ_P instead requires its separate allocation, as in Proposition 12. □

C Admissible Queries, Missing Blocks, and Selective Completion

Let $W_t = (Y_t(i) : i \in \mathcal{Q})$ denote a complete potential-score vector with law P, conditional on the frozen proposal transcript. Before world t, the active set and request decision are measurable using completed selection information. For a whole-block acceptance indicator R_t, suppose $R_t \perp W_t$ conditional on this past. Whenever acceptance has positive probability, the accepted vector retains law P, since

$$\Pr(W_t \in H, R_t = 1 \mid \text{past}) = P(H) \Pr(R_t = 1 \mid \text{past}).$$

Division and iteration prove the target law for chronological accepted blocks. Surviving pairs observe prefixes of that sequence. A deterministic same-world retry preserves the original random world; an outcome-dependent redraw generally does not. A fixed failure-outcome convention is part of a different target utility law and must apply to the whole inferential block.

Selective completion has a different contract. Suppose H is revealed anchor information, D is the potential pairwise difference, and the completion indicator satisfies $\Pr(R = 1 \mid H, D) = q(H) > 0$. Then

$$\widetilde{D} = \frac{RD}{q(H)}, \qquad \mathbb{E}\widetilde{D} = \mathbb{E}D, \qquad \mathbb{E}\widetilde{D}^2 = \mathbb{E}\frac{D^2}{q(H)}. \tag{C.1}$$

These identities follow by conditioning on (H, D). They preserve a mean, not the original variance or bounded range. If $|D| \leq c$ and $q(H) \geq q_{\min} > 0$, the corrected range is at most $c/q_{\min}$. Any confidence sequence must use that corrected observation model. When completion is rare exactly where D^2 is large, the second moment can make a cheap selective query less efficient than a complete pair.

D Reference Evidence and Reporting Estimands

Suppose an independent reference procedure provides simultaneous intervals $[\ell_i, u_i]$ for all candidate means with error at most δ_{ref}. Then the selected regret belongs to

$$[\max\{0, \max_i \ell_i - u_{\hat{i}}\},\ \max\{0, \max_i u_i - \ell_{\hat{i}}\}]. \tag{D.1}$$

The report's absolute error belongs to

$$[\max\{0, \ell_{\hat{i}} - \widehat{\mu}_{\hat{i}}, \widehat{\mu}_{\hat{i}} - u_{\hat{i}}\},\ \max\{|\widehat{\mu}_{\hat{i}} - \ell_{\hat{i}}|, |\widehat{\mu}_{\hat{i}} - u_{\hat{i}}|\}]. \tag{D.2}$$

A certified run is a definite joint success when both upper endpoints meet their respective tolerances; it is a definite failure if either lower endpoint exceeds its tolerance; otherwise it is inconclusive. Reference uncertainty is separate from the algorithm's certificate error budget. Across replicates these classifications require simultaneous or otherwise explicitly allocated reference error control, rather than treating every per-replicate interval as jointly certain.

Table 4 applies these rules to four reported records at $\varepsilon = 0.20$ pp and $\eta = 1.00$ pp. The maximum competitor lower and upper endpoints are 80.15 and 80.32 pp. All four outputs carry a frozen joint certificate; the reference evidence can confirm it, contradict one component, or leave it unresolved.

Table 4: Joint scoring from simultaneous reference intervals. All numerical entries are in pp. The two error columns give lower and upper bounds.

Winner interval	Report	Regret bounds	Report-error bounds	Classification
$[80.20, 80.30]$	80.70	$[0, .12]$	$[.40, .50]$	Definite success
$[79.70, 79.80]$	80.10	$[.35, .62]$	$[.30, .40]$	Selection failure
$[80.00, 80.20]$	80.60	$[0, .32]$	$[.40, .60]$	Inconclusive
$[80.20, 80.30]$	81.45	$[0, .12]$	$[1.15, 1.25]$	Reporting failure

For a method m, write $e_{b,m} = \widehat{\mu}_{b,m} - \mu_{\hat{i}_{b,m}}$. An absolute pooled bias is $|B^{-1}\sum_b e_{b,m}|$; an average of absolute setting-level biases is $S^{-1}\sum_s |B_s^{-1}\sum_{b\in s} e_{b,m}|$. They are different from mean absolute error $B^{-1}\sum_b |e_{b,m}|$ and can yield different percentage reductions. A bias-reduction claim must state which averaging convention is used and whether reports include all outputs or a prespecified certified subset. Stratified diagnostics condition on declared summaries of selection history, with stratum counts and uncertainty; they do not identify bias conditional on every possible complete transcript from a finite sample.

E Joint Replay Reporting Protocol and Aggregate Endpoints

This appendix defines the joint replay reporting protocol, estimands, reliability targets, and common accounting scope used by the application comparisons. The empirical reporting and cost results are reported in Appendix V. The protocol uses fresh winner-only reporting, with selection and reporting on disjoint world banks, iid winner utilities in $[0, 100]$ pp, and a reporting count fixed before its outcomes are revealed. The cost scope includes reporting and audit, and joint-success rates count abstentions as failures. Table 5 summarizes the protocol and endpoint definitions. Gaussian reporting experiments are specified in Appendix G. The selection-only study fixes $\varepsilon = .20$ pp and $\delta_{\mathrm{sel}} = .05$; the joint study allocates the total error budget across selection and reporting. Equation (H.1) defines bias on a declared set of paired same-winner reports. Completion and joint success use all outer runs, including abstentions.

The bounded Hoeffding reporting rule at the joint target prescribes

$$n_R = \left\lceil \frac{100^2}{2(1.00)^2} \log \frac{2}{0.03} \right\rceil = 20{,}999. \tag{E.1}$$

This sufficient sample size follows from the bounded reporting rule. The reporting ledger links each acquired sample count to its interval rule. For an observed count K_{cert} of certified outputs using this rule, the reporting acquisitions alone require $K_{\mathrm{cert}} n_R$ evaluations, before selection, audit, or failed attempts are charged. If a cohort of N runs actually attains empirical completion $\widehat{Q} = K_{\mathrm{cert}}/N \geq \gamma_\star$, then

$$T_{\mathrm{report}} \geq \lceil \gamma_\star N \rceil\, 20{,}999 \tag{E.2}$$

whenever every certified run uses this bounded-reporting rule. The realized reporting cost uses the observed K_{cert}; the population requirement $Q \geq \gamma_\star$ defines the reliability target. A variance-aware design uses its corresponding sample requirement. The observation unit is a fresh configuration–world outcome, including the stochastic execution specified by the target world law. Appendix L

Table 5: Joint replay reporting protocol and metric definitions. Selection-only and joint reporting allocate their error budgets to different endpoints.

Field	Definition or setting
Bias averaging	Equal-setting absolute signed bias, Eq. (H.2)
Estimand denominators	Same-winner paired reports for bias; all outer runs for completion and joint success
Joint $(\varepsilon, \eta, \delta)$	$(0.20\text{ pp}, 1.00\text{ pp}, 0.05)$
Error split	$\delta_{\text{sel}} = 0.02$, $\delta_{\text{rep}} = 0.03$
Completion and joint-success targets	$Q \geq 95\%$ and $J \geq 90\%$
Joint-success noninferiority margin	$J_{\text{PAIRS}} - J_{\text{comparator}} \geq -2$ pp
Route used on each replay setting	Fresh winner-only; count frozen before reporting
Application studies	Reporting and attempts: Appendix V; qualifying budgets: Appendix W; additional diagnostics: Appendix S.
Cost scope	Selection, reporting, and audit; all-in costs separate

Table 6: Conditional accounting bounds for the fixed Hoeffding reporting rule, $n_R = 20{,}999$. T_P and T_C denote total PAIRS and comparator evaluations under one common cost scope, with $R = T_C/T_P$. Other stages can only increase the required total.

Fixed assumptions	Necessary implication
$K_{\text{cert}} \geq 181$	$T_P \geq 3{,}800{,}819$
$K_{\text{cert}} \geq 181$, $T_C = 6{,}900{,}000$	$R \leq 1.8154$
$T_P = 3{,}000{,}000$, $N = 190$	$K_{\text{cert}} \leq 142$; $\widehat{Q} \leq 74.74\%$

gives the corresponding counts under bounded and known-Gaussian assumptions. The prespecified reliability requirements are $Q \geq 95\%$ and $J \geq 90\%$, with a -2 pp margin for joint-success noninferiority. Attainment is evaluated with paired uncertainty, separately from the point cost ratio.

The rule also gives conditional accounting bounds for comparing totals. For $N = 190$ and attained $\widehat{Q} \geq 0.95$, at least 181 certificates require $181 \cdot 20{,}999 = 3{,}800{,}819$ reporting observations. Table 6 propagates that requirement. Each row derives a necessary cost or completion bound from its stated assumptions.

F A RESTRICTED CHARACTERIZATION OF JOINT SELECTION AND REPORTING

One alternative reverses the acceptable winner; a second moves its mean in the least-cost covariance direction while preserving that winner. A complete world's KL divergence in level–contrast coordinates is

$$\mathrm{KL}((\theta, \Delta), (\theta + h, \Delta + d)) = h^2/(2v_A) + d^2/(2v_D). \tag{F.1}$$

Changing measure for each alternative yields two necessary inequalities and hence their maximum. Adding the bounds would count reusable observations twice. Multiple acceptable answers require an answer-aware change of measure; $|\Delta| > \varepsilon$ makes the acceptable candidate unique (Degenne & Koolen, 2019).

We characterize joint selection and reporting for two Gaussian candidates with known covariance and complete-pair queries. Every admissible action observes a complete pair and costs two candidate evaluations. In world t,

$$Y_{t1} = \theta + \Delta/2 + \sigma_c Z_t + \sigma_e E_{t1}, \tag{F.2}$$

$$Y_{t2} = \theta - \Delta/2 + \sigma_c Z_t + \sigma_e E_{t2}, \tag{F.3}$$

where all Z_t, E_{t1}, E_{t2} are independent standard normal variables, $\sigma_e > 0$ and $\sigma_c \geq 0$ are known, and $\theta \in \mathbb{R}$ is unknown. Let $\mu_1 = \theta + \Delta/2$, $\mu_2 = \theta - \Delta/2$, $\varepsilon \geq 0$, $\eta > 0$, and $|\Delta| > \varepsilon$. Thus the best candidate is the unique ε-optimal candidate. The pair average $A_t = (Y_{t1} + Y_{t2})/2$ and difference $D_t = Y_{t1} - Y_{t2}$ are independent, with

$$A_t \sim N(\theta, a), \qquad D_t \sim N(\Delta, 2\sigma_e^2), \qquad a = \sigma_c^2 + \sigma_e^2/2. \tag{F.4}$$

Write $v = \sigma_c^2 + \sigma_e^2$ for the marginal candidate variance; then $a \le v \le 2a$.

A procedure purchases a stopping-time number N of complete pairs, then either abstains or returns $(\widehat{q}, \widehat{V}, C = 1)$. Its certificate is incorrect when

$$C = 1 \quad \text{and} \quad \left\{ \max_j \mu_j - \mu_{\widehat{q}} > \varepsilon \text{ or } |\widehat{V} - \mu_{\widehat{q}}| > \eta \right\}. \tag{F.5}$$

Uniform (δ, γ) validity means that, on every instance of this model, the probability of (F.5) is at most δ and $\Pr(C = 1) \ge \gamma$. Internal randomization is independent of the worlds. All stages, including reporting or any audit that purchases observations, count in N. There are no uncharged observations.

Theorem 6 (Complete-pair lower bound)**.** *Suppose $0 < \delta < \gamma/2$ and $0 < \gamma \le 1$. For any uniformly (δ, γ)-valid procedure and any $|\Delta| > \varepsilon$, define*

$$g = |\Delta| + \varepsilon, \qquad K_{\delta,\gamma} = \mathrm{kl}(\gamma - \delta, \delta), \tag{F.6}$$

where $\mathrm{kl}(p, q) = p \log(p/q) + (1 - p) \log((1 - p)/(1 - q))$. Then

$$\mathbb{E}_{\theta,\Delta}[N] \;\ge\; K_{\delta,\gamma} \max\left\{ \frac{4\sigma_e^2}{g^2}, \frac{v}{2\eta^2} \right\}. \tag{F.7}$$

The expected candidate-evaluation cost is $2\mathbb{E}[N]$ in this action model. If $N \le B$ almost surely, the right side of (F.7) *is also a necessary lower bound on B.*

Proof. The claim is immediate if $\mathbb{E}[N] = \infty$, so assume finite expectation. Let $P = P_{\theta,\Delta}$ and $P' = P_{\theta',\Delta'}$ denote one-world laws. The Gaussian factorization in (F.4) gives

$$\mathrm{KL}(P \| P') = \frac{(\theta - \theta')^2}{2a} + \frac{(\Delta - \Delta')^2}{4\sigma_e^2}. \tag{F.8}$$

For the stopped transcript, the log likelihood ratio is the sum of the one-world log likelihood ratios up to N. The event $\{N \ge t\}$ is measurable before world t is opened. Independence of the next world therefore yields

$$\mathrm{KL}(P^{\mathrm{tr}} \| (P')^{\mathrm{tr}}) = \mathbb{E}_P[N] \, \mathrm{KL}(P \| P'). \tag{F.9}$$

For completeness, the exchange of expectation and summation follows from $\sum_t \mathbb{E}_P[\mathbf{1}\{N \ge t\}|L_t|] = \mathbb{E}_P[N]\mathbb{E}_P[|L_1|] < \infty$, where L_t is a Gaussian one-world log likelihood ratio. Independent internal randomness contributes zero additional likelihood ratio. Mapping the transcript to any binary event H can only decrease relative entropy, so the left side is at least $\mathrm{kl}(P(H), P'(H))$.

Let $s = \mathrm{sign}(\Delta)$ and let q^* be its best candidate. For the comparison alternative take $\theta' = \theta$ and $\Delta' = -s(\varepsilon + h)$, with $h > 0$. Set $H = \{C = 1, \widehat{q} = q^*\}$. Under P, completion and joint validity give $P(H) \ge \gamma - \delta$. Under P', q^* is more than ε suboptimal, hence $P'(H) \le \delta$. Since $\gamma - \delta > \delta$, monotonicity of binary relative entropy for $p > q$ implies $\mathrm{kl}(P(H), P'(H)) \ge K_{\delta,\gamma}$. Equations (F.8)–(F.9) now imply

$$\mathbb{E}_P[N] \ge \frac{4\sigma_e^2 K_{\delta,\gamma}}{(|\Delta| + \varepsilon + h)^2}.$$

Letting $h \downarrow 0$ gives the first term.

For reporting, write $V = \sigma_c^2 \mathbf{1}\mathbf{1}^\top + \sigma_e^2 I_2$ for the one-world covariance and take

$$\mu' = \mu + h \frac{V e_{q^*}}{V_{q^* q^*}}, \qquad h > 2\eta.$$

The winner's mean increases by h and the other mean by $h\sigma_c^2/v$. Since $\sigma_e > 0$, its winning gap increases, so q^* remains the unique acceptable candidate. This alternative belongs to the same unrestricted mean model. Its one-world divergence is $\frac{1}{2}(\mu' - \mu)^\top V^{-1}(\mu' - \mu) = h^2/(2v)$. Use the event

$$H = \{C = 1, \widehat{q} = q^*, |\widehat{V} - \mu_{q^*}| \le \eta\}.$$

Again $P(H) \ge \gamma - \delta$. The two closed accuracy intervals are disjoint because their centers differ by more than 2η. Thus H is an incorrect report under P' and $P'(H) \le \delta$. The same argument gives

$$\mathbb{E}_P[N] \ge \frac{2v K_{\delta,\gamma}}{h^2}.$$

Letting $h \downarrow 2\eta$ gives the second term. Both inequalities constrain the same expected sample count, so they combine through a maximum. □

Why the effective gap is $|\Delta| + \varepsilon$. The closest parameter at which the original winner is not ε-optimal lies beyond the opposite ε boundary. Consequently the change-of-measure distance is $|\Delta| + \varepsilon$. Using $|\Delta| - \varepsilon$ would analyze a different certification task. The completion requirement is equally material: $\gamma - \delta$, rather than γ or $1 - \delta$, is the available lower bound on a correct completed output. If $\gamma \leq 2\delta$, the binary-event argument above does not give the displayed positive lower bound.

F.1 A CONSTRUCTIVE SEQUENTIAL COMPARISON AND PROTECTED REPORT

Choose $\delta_s, \delta_r \in (0, 1)$ with $\delta_s + \delta_r \leq \delta$. From complete racing pairs, maintain

$$\overline{D}_n = \frac{1}{n}\sum_{t=1}^{n} D_t, \qquad w_n = \sqrt{\frac{4\sigma_e^2}{n}\log\left(\frac{\pi^2 n^2}{3\delta_s}\right)}. \tag{F.10}$$

Stop at

$$\tau = \inf\{n \geq 1 : w_n \leq |\overline{D}_n| + \varepsilon\}, \tag{F.11}$$

and select candidate 1 if $\overline{D}_\tau \geq 0$, otherwise candidate 2. The tie convention is fixed before sampling. Reserve the deterministic number of fresh complete reporting pairs

$$r = \max\left\{1, \left\lceil \frac{2v}{\eta^2}\log\left(\frac{2}{\delta_r}\right)\right\rceil\right\}. \tag{F.12}$$

Return $\widehat{V}$ as the mean of the selected component in those fresh pairs. The unselected components still count toward the cost because the action set permits only complete pairs. Report outcomes do not change r, the selected candidate, or the reporting endpoint. The algorithm uses neither θ nor Δ; instance-dependent quantities below occur only in the analysis.

Theorem 7 (Achievable validity, expected cost, and budgeted completion)**.** *For every instance of* (F.3)*, the above procedure completes almost surely and issues an incorrect joint certificate with probability at most $\delta_s + \delta_r$. Conditional on the complete selection history $\mathcal{S}_\tau$,*

$$\mathbb{E}[\widehat{V} \mid \mathcal{S}_\tau] = \mu_{\widehat{q}}, \qquad \Pr(|\widehat{V} - \mu_{\widehat{q}}| > \eta \mid \mathcal{S}_\tau) \leq \delta_r. \tag{F.13}$$

For $g = |\Delta| + \varepsilon$, define

$$H = \frac{16\sigma_e^2}{g^2}, \qquad b_0 = \frac{\pi^2}{3\delta_s}, \qquad n_0 = \left\lceil \max\{1, 4H\log(e + 2H\sqrt{b_0})\}\right\rceil. \tag{F.14}$$

Then

$$\mathbb{E}[N] \leq n_0 + H + r, \qquad \Pr(\tau > m) \leq e^{-m/H} \quad (m \geq n_0). \tag{F.15}$$

For an integer budget $B \geq r + 1$, run the race for at most $m = B - r$ pairs and abstain if $\tau > m$. This capped procedure uses at most B pairs, retains the same certificate-error bound, and satisfies

$$\Pr(C = 1) \geq 1 - e^{-(B-r)/H} \qquad (B - r \geq n_0). \tag{F.16}$$

In particular, for $0 < \gamma < 1$ it suffices that $B - r \geq \max\{n_0, H\log(1/(1-\gamma))\}$ to attain completion at least γ on this instance. Its probability of a correct completed output is at least $\Pr(C = 1) - (\delta_s + \delta_r)$; completion and correctness are distinct probabilities.

Proof. Since $D_t \sim N(\Delta, 2\sigma_e^2)$, the Gaussian tail inequality gives

$$\Pr(|\overline{D}_n - \Delta| > w_n) \leq 2\exp\{-nw_n^2/(4\sigma_e^2)\} = \frac{6\delta_s}{\pi^2 n^2}.$$

Summing over n proves simultaneous coverage of Δ with probability at least $1 - \delta_s$. On this event, if candidate 1 is selected, then $-\Delta \leq -\overline{D}_\tau + w_\tau \leq \varepsilon$. If candidate 2 is selected, the symmetric inequality yields $\Delta \leq \varepsilon$. Thus every stopping output is ε-optimal on the simultaneous coverage event.

All reporting worlds are fresh and independent of $\mathcal{S}_\tau$. After conditioning on that history, $\widehat{q}$ is fixed and its report is normal with mean $\mu_{\widehat{q}}$ and variance v/r. The same tail inequality and (F.12) yield (F.13). A union bound gives the joint error claim.

We next verify the explicit threshold in (F.14). The inequality $w_n \le g/2$ is equivalent to $n \ge H\log(b_0 n^2)$. Let $L = \log(e + 2H\sqrt{b_0}) \ge 1$ and $t = n/(2H)$. For $n \ge n_0$, $t \ge 2L$; the function $t - \log t$ is increasing for $t \ge 1$, and

$$t - \log t \ge 2L - \log(2L) \ge L \ge \log(2H\sqrt{b_0}).$$

The middle inequality uses $L \ge \log(2L)$ for $L \ge 1$. Rearranging proves $w_n \le g/2$ for every $n \ge n_0$.

Assume $\Delta > \varepsilon$; the other sign is symmetric. If $\tau > n$ and $n \ge n_0$, then

$$|\overline{D}_n| < w_n - \varepsilon \le (\Delta - \varepsilon)/2,$$

and hence $\overline{D}_n - \Delta < -g/2$. The one-sided Gaussian tail bound yields

$$\Pr(\tau > n) \le \exp\{-ng^2/(16\sigma_e^2)\} = e^{-n/H}.$$

This bound tends to zero, proving almost-sure completion. Summing the tail probabilities gives

$$\mathbb{E}[\tau] = \sum_{n=0}^{\infty} \Pr(\tau > n) \le n_0 + \sum_{n=n_0}^{\infty} e^{-n/H} \le n_0 + H,$$

where $n_0 \ge 1$ and $\sum_{n=1}^{\infty} e^{-n/H} = 1/(e^{1/H} - 1) \le H$. Adding r proves the expected-cost claim. The capped procedure returns a certificate precisely when $\tau \le B - r$. Abstention cannot introduce an incorrect certificate, and completion is bounded by the same tail argument. Conditioning additionally on completion preserves (F.13) because completion is determined before opening reporting outcomes. □

Comparison of the two bounds. With $\delta_s = \delta_r = \delta/2$, the construction uses

$$\mathbb{E}[N] = O\left(1 + \frac{\sigma_e^2}{g^2}\left[\log\frac{1}{\delta} + \log\left(e + \frac{\sigma_e^2}{g^2}\right)\right] + \frac{v}{\eta^2}\log\frac{1}{\delta}\right). \tag{F.17}$$

Since a sum of two nonnegative terms is at most twice their maximum, (F.17) has the variance and tolerance scales of Theorem 6, up to logarithmic factors, universal constants, and the additive cost of buying at least one pair per stage. For fixed positive completion target γ and sufficiently small δ, $K_{\delta,\gamma}$ has order $\log(1/\delta)$. The comparison term tends to zero as $\sigma_e^2 \to 0$ at fixed gap, whereas the reporting lower bound retains its σ_c^2/η^2 dependence. This is the comparison–reporting asymmetry.

Fresh reporting attains the constructive upper bound in this complete-pair model. The expected-cost bound characterizes acquisition efficiency, while Eq. (F.16) gives completion under a hard cap. When $\varepsilon = 0$ and gaps can approach zero, no fixed finite budget can satisfy a nontrivial uniform completion target over the whole model, as is also apparent from the lower bound.

F.2 A REPORTING LENGTH CHOSEN BEFORE OPENING REPORT OUTCOMES

Proposition 8 (Selection-measurable reporting length)**.** *Let $\mathcal{S}$ contain all selection information, including the selected candidate $\widehat{q}$ and an integer $1 \le R < \infty$ chosen before opening the reporting bank. Suppose conditional on $\mathcal{S}$, the fresh reporting observations for candidate $\widehat{q}$ are iid normal with mean $\mu_{\widehat{q}}$ and known variance $v_{\widehat{q}}$. Then the R-sample mean satisfies*

$$\mathbb{E}[\widehat{V} - \mu_{\widehat{q}} \mid \mathcal{S}] = 0.$$

If $R \ge (2v_{\widehat{q}}/\eta^2)\log(2/\delta_r)$ almost surely, its conditional error probability is at most δ_r. The assertion remains true after restricting to any completion event in $\mathcal{S}$, with the reporting length fixed by that pre-report transcript.

Proof. Conditional on $\mathcal{S}$, both R and $\widehat{q}$ are fixed. The report is therefore normal with the asserted mean and variance $v_{\widehat{q}}/R$. Its mean and the Gaussian tail bound prove the result. Conditioning on a selection-measurable completion event changes none of these conditional distributions. □

F.3 When stopping on a protected fission stream is legitimate

Proposition 9 (Gaussian fission with protected reporting outcomes). *Let $Y_t \overset{\text{iid}}{\sim} N(\mu, \Sigma)$ and independently $W_t \overset{\text{iid}}{\sim} N(0, \Sigma)$, with known covariance Σ. For fixed $\kappa > 0$, define*

$$F_t = Y_t + \kappa W_t, \qquad G_t = Y_t - \kappa^{-1} W_t.$$

The entire streams $(F_t)_{t\geq 1}$ and $(G_t)_{t\geq 1}$ are independent, with respective covariances $(1+\kappa^2)\Sigma$ and $(1+\kappa^{-2})\Sigma$ per world, and both have mean μ. Let $\mathcal{F}$ be generated by the full F stream and independent algorithmic randomness. Suppose $\widehat{q}$, a finite positive count R, and R distinct retained world indices $I_1, \ldots, I_R$ are $\mathcal{F}$-measurable. None of the protected G outcomes is used to construct these objects. Then

$$\widehat{V}_G = \frac{1}{R}\sum_{j=1}^{R} G_{I_j,\widehat{q}} \;\Big|\; \mathcal{F} \sim N\left(\mu_{\widehat{q}}, \frac{(1+\kappa^{-2})\Sigma_{\widehat{q}\widehat{q}}}{R}\right). \tag{F.18}$$

Thus the report is selection-conditionally unbiased and its conditional Gaussian interval has nominal coverage. In particular, R may be an almost-surely finite stopping time based exclusively on F, with retained indices $1, \ldots, R$.

Proof. The cross covariance within each world is $\mathrm{Cov}(F_t, G_t) = \Sigma - \Sigma = 0$. Joint Gaussianity gives independence within each world; independence of the worlds gives independence of the full streams. After conditioning on $\mathcal{F}$, the index set, coordinate, and sample count are fixed, while the corresponding G vectors remain iid with their unconditional Gaussian law. Taking their mean proves (F.18). □

Information flow. Query decisions, retries, failures, cost-based stopping, retention, and selection are measurable with respect to the released F stream and independent randomness. The unfissioned Y remains protected because its correlation with G would transmit reporting information into those decisions. This measurability condition applies to precommitted rules as well as adaptive ones. The resulting guarantee uses the stated Gaussian covariance model and the fission-adjusted variances.

G Controlled Joint-Selection and Reporting Studies

We conduct five controlled Gaussian studies to examine information allocation, reporting precision, adaptive stopping, and selective completion. Experiments A–C use fully evaluated world banks; D studies a two-candidate sequential rule; E studies outcome-dependent completion of paired observations. Scores, biases, and root mean squared errors (RMSEs) use the Gaussian model's native units. Selection success, reporting coverage, correct joint output, and computation are reported separately.

All studies use master seed 2026091907, with four NumPy SeedSequence streams for fixed fission (A and B), misspecified fission (C), selective completion (E), and adaptive protected reporting (D). The execution environment is Python 3.13.5, NumPy 2.3.5, and SciPy 1.17.0. A and B use $100{,}000$ independent repetitions per setting, C uses $200{,}000$, D uses $50{,}000$ per correlation setting, and E uses $1{,}000{,}000$ worlds. Parameter settings within a study share primitive noise. Mean intervals use a normal approximation with Monte Carlo standard errors; proportion intervals use Wilson 95% intervals. These are pointwise intervals across settings.

G.1 A: Protecting Only the Information Used for Selection

There are $M = 20$ candidates with equal true means 0.7, $n = 50$ worlds, marginal standard deviation $\sigma = 0.1$ per world, and information-allocation parameter $\lambda = 1$. Set $v = \sigma^2/n$ and $P = I - \mathbf{1}\mathbf{1}^{\mathsf{T}}/M$. Generate

$$X \sim \mathcal{N}\big(\mu, v[\rho \mathbf{1}\mathbf{1}^{\mathsf{T}} + (1-\rho)I]\big), \tag{G.1}$$

$$W_c \sim \mathcal{N}(0, v(1-\rho)P), \qquad W_c \perp X, \tag{G.2}$$

$$F = PX + \sqrt{\lambda} W_c, \qquad G = X - W_c/\sqrt{\lambda}. \tag{G.3}$$

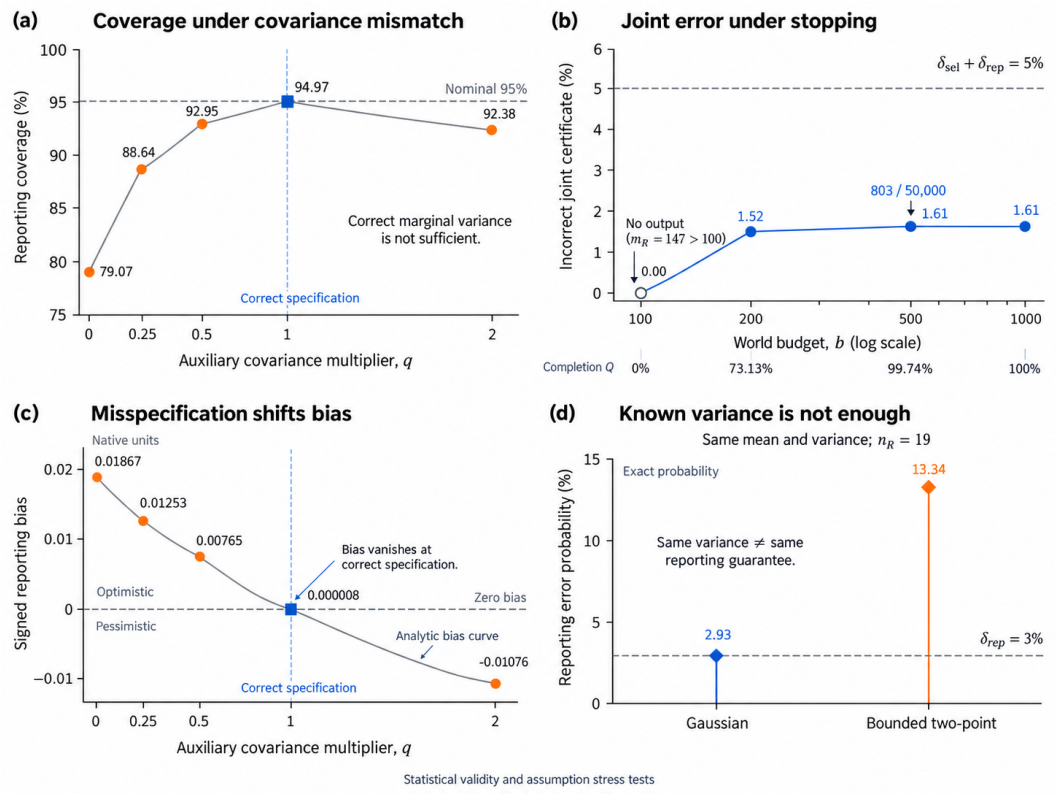


Figure 5: Statistical validity and assumption stress tests. (a,c) Coverage and signed bias under covariance misspecification: 200,000 trials per setting; the bias curve is analytical. (b) Joint error and completion: 50,000 trials at $\rho = .67$, with all runs in both denominators (Appendix G). (d) Exact probabilities of error exceeding 1 pp with 19 fresh observations. The laws share mean 50 and variance 4 (Eq. (R.8)).

The structured report is $G_{\hat{\imath}}$, where $\hat{\imath} = \arg\max_i F_i$. For independent $Z_{\text{level}} \sim \mathcal{N}(0,1)$, the full-vector comparator reports the same winner using

$$G_{\text{full}} = G - \sqrt{v[\rho + (1-\rho)/M]/\lambda}\, Z_{\text{level}} \mathbf{1}. \tag{G.4}$$

The two protected procedures therefore have identical selection decisions. The naive columns in Table 7 instead use the original winner $\arg\max_i X_i$, its unprotected score, and an ordinary 95% interval. Gaussian interval half-widths are $z_{0.975}$ times the relevant standard deviation. Each bank costs $Mn = 1{,}000$ candidate evaluations.

Table 7: Experiment A: reporting under equal true means. Every selected candidate is optimal. Coverage concerns the selected candidate's true mean at the fixed reporting endpoint. Each row uses 100,000 repetitions.

	Naive winner report		Full-vector	Structured report	
ρ	Mean optimism	Coverage (%)	RMSE	RMSE	Coverage (%)
0.00	0.02639	60.49	0.02000	0.01975	94.96
0.50	0.01865	79.02	0.02002	0.01717	95.04
0.67	0.01515	84.58	0.02003	0.01620	95.08
0.90	0.00833	91.92	0.02005	0.01480	94.99

For equicorrelated observations, write $s^2 = 0.1^2/n$ for a candidate mean's marginal variance. The structured reporting variance is

$$v_{\text{struct}} = s^2 \left[1 + \frac{(1-\rho)(1-1/M)}{\lambda}\right], \qquad v_{\text{full}} = s^2(1 + 1/\lambda). \tag{G.5}$$

Thus, at $\rho = 0.9$ and $\lambda = 1$, the ratio of Gaussian interval widths is $\sqrt{1.095/2} \simeq 0.740$, a theoretical reduction of about 26.0%. The 26% precision gain holds at the same fully evaluated bank and selection information. Its conversion to end-to-end acquisition savings depends on the reporting route, including the option to evaluate only the selected winner afresh.

The first row gives the equal-means counterexample: selection is always correct, but the ordinary winner interval covers 60.49% of the time. The corresponding population coverage is $0.975^{20} - 0.025^{20} \simeq 60.27\%$, and expected optimism is approximately 0.02641 under the exact known-variance Gaussian law.

G.2 B: The Information-Allocation Tradeoff

The setting again has $M = 20$ candidates, $n = 50$ worlds, and $\sigma = 0.1$, now with $\mu_1 = 0.72$ and $\mu_i = 0.70$ for $i > 1$, selection tolerance $\varepsilon = 0.01$, and correlation $\rho = 0.67$. Each setting uses $100{,}000$ repetitions. Table 8 varies λ. Since the gap exceeds the tolerance, selection success requires selecting the uniquely better candidate. The success counts are $62{,}559$, $46{,}438$, and $25{,}545$ for $\lambda = 0.25, 1, 4$, respectively.

Table 8: Experiment B: selection and reporting respond in opposite directions to information allocation. Each setting uses $100{,}000$ repetitions.

λ	Selection success (%)	Structured reporting RMSE
0.25	62.56	0.02119
1.00	46.44	0.01619
4.00	25.55	0.01471

Increasing λ protects reporting more strongly but adds noise to the selection channel. The decrease in RMSE therefore accompanies a decrease in selection success. The fixed-λ comparison exposes the two tolerances and costs that a joint allocation policy must balance.

G.3 C: Covariance Misspecification Restores Dependence

This experiment uses $M = 20$ candidates with mean 0.7, $n = 50$ worlds, $\sigma = 0.1$, $\rho = 0.5$, $\lambda = 1$, and $200{,}000$ repetitions per setting. Keeping the observation law from A, we replace the auxiliary covariance by $v(1-\rho)qP$ for $q \in \{0, 0.25, 0.5, 1, 2\}$ and compute F and G using the same formulas. Intervals use the true unconditional marginal reporting variance

$$v_G(q) = v\left[1 + q(1-\rho)(1-1/M)/\lambda\right]. \tag{G.6}$$

This design isolates residual selection dependence from marginal variance miscalibration in Table 9.

Table 9: Experiment C: noise-variance misspecification. Correct marginal report variance does not remove dependence between selection and reporting. Each row uses $200{,}000$ repetitions.

Noise-variance multiplier q	Signed report bias	95% interval coverage (%)
0.00	$+0.01867$	79.07
0.25	$+0.01253$	88.64
0.50	$+0.00765$	92.95
1.00	$+0.000008$	94.97
2.00	-0.01076	92.38

The correctly specified row is close to nominal coverage and has negligible observed signed bias. Under- and over-specification restore selection–report dependence, with over-specification reversing

the bias. Exact covariance determines the cancellation in the Gaussian independence proof; using a covariance estimate requires accounting for its estimation error even when the pilot is independent.

The direction and magnitude have an exact expression in this equal-mean model. With $\kappa_M = \mathbb{E}\max_{i\le M} Z_i$, Gaussian regression of G onto F gives

$$b(q) = \kappa_M \frac{\sigma\sqrt{1-\rho}}{\sqrt{n}} \frac{1-q}{\sqrt{1+\lambda q}}. \tag{G.7}$$

Indeed, on the contrast space the cross-covariance multiplier is $1-q$ and the selection-variance multiplier is $1+\lambda q$. Taking the expected winning contrast yields Eq. (G.7). At $q = 0.9$ and 1.1 the population signed biases are $+0.001355$ and -0.001289, respectively; $q = 1$ restores zero bias.

G.4 D: Completion under Adaptive Stopping

This study uses two Gaussian candidates with marginal variance $\sigma^2 = 0.01$, gap $\Delta = 0.02$, $\varepsilon = 0.01$, reporting tolerance $\eta = 0.02$, $\lambda = 1$, and error allocations $\delta_{\text{sel}} = \delta_{\text{rep}} = 0.025$. We run $50{,}000$ repetitions at each $\rho \in \{0, 0.67, 0.9\}$; Table 10 reports $\rho = 0.67$. Each trajectory has at most $1{,}000$ worlds, generated in batches of 500 trajectories. The selection differences have mean Δ and variance $v_D = 2\sigma^2(1-\rho)(1+\lambda)$, and the protected report has per-world variance $v_R = \sigma^2[1+(1-\rho)/(2\lambda)]$.

For the mean difference $\bar{D}_t$, the known-variance Gaussian mixture confidence sequence has half-width

$$r_t = \frac{\sqrt{v_D(t+\kappa)\log[(t+\kappa)/(\delta_{\text{sel}}^2\kappa)]}}{t}, \qquad \kappa = 20. \tag{G.8}$$

Define the minimum report size

$$m_R = \left\lceil z^2_{1-\delta_{\text{rep}}/2} v_R/\eta^2 \right\rceil. \tag{G.9}$$

Certification is allowed only at $t \ge \max(5, m_R)$. Candidate 1 is eligible when $\bar{D}_t - r_t \ge -\varepsilon$; candidate 2 is eligible when $\bar{D}_t + r_t \le \varepsilon$. When both are eligible, the sign of $\bar{D}_t$ breaks the tie. Let τ be the first eligible time. At world budget b, completion is $\mathbf{1}\{\tau \le b\}$ and logical candidate cost is $2\min(\tau, b)$. We evaluate $b \in \{50, 100, 200, 500, 1000\}$. For $\rho = 0.67$, $v_D = 0.0132$, $v_R = 0.01165$, and $m_R = 147$.

After generating the selection path, we generate reporting error from its exact conditional law, $Z\sqrt{v_R/\min(\tau, b)}$ with independent $Z \sim \mathcal{N}(0,1)$. The reporting interval uses half-width $z_{1-\delta_{\text{rep}}/2}\sqrt{v_R/\min(\tau, b)}$. This Gaussian-law evaluation records logical candidate acquisitions, with sealed-channel execution time and audit acquisition outside its cost scope. All three outcome columns in Table 10 use all runs as their denominator.

Table 10: Experiment D: sequential two-candidate selection with protected reporting. Completion is $\Pr(A_B)$; correct joint output is $\Pr(A_B \cap G)$; incorrect joint certificate is $\Pr(A_B \cap G^c)$. Abstentions are failures for correct joint output. All rates use $50{,}000$ runs.

World budget	Completion (%)	Correct joint output (%)	Incorrect certificate (%)
100	0.00	0.00	0.00
200	73.13	71.61	1.52
500	99.74	98.13	1.61
1,000	100.00	98.39	1.61

At budget 500, $49{,}868$ runs complete: $49{,}065$ yield a correct joint output and 803 yield an incorrect joint certificate. The remaining 132 runs abstain. The reporting intervals cover in $48{,}627$ of the $49{,}868$ completed runs, giving 97.51% conditional coverage against the nominal 97.5% level. Mean logical candidate cost is 373.85292. An incorrect joint certificate means either a non-ε-optimal choice or reporting error exceeding η; interval noncoverage is a distinct event because the interval radius can be smaller than η. Budgets 50 and 100 cannot reach $m_R = 147$, explaining their zero completion. These outcomes demonstrate the difference between controlling incorrect certificates and returning a correct joint answer within budget.

G.5 E: Selective Completion and Its Variance Cost

Generate $Y_1, Z \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$ and $Y_2 = 0.5Y_1 + \sqrt{0.75}Z$, so both outcomes have unit variance and their correlation is 0.5. Let $D = Y_1 - Y_2$. After observing Y_1, draw a completion indicator $R_{\text{seen}} \sim \text{Bernoulli}(p(Y_1))$, where $p(Y_1) = 0.9$ for $Y_1 > 0$ and 0.1 otherwise. Conditional on Y_1, this draw is independent of the unrevealed outcome. Among $N = 1{,}000{,}000$ worlds, $500{,}779$ are completed. The naive estimate averages D over completed worlds. The inverse-probability estimate is the Horvitz–Thompson mean

$$\widehat{\Delta}_{\text{IPW}} = \frac{1}{N}\sum_{t=1}^{N} \frac{R_{\text{seen},t} D_t}{p(Y_{1,t})}. \tag{G.10}$$

The denominator includes every generated world, as required by the Horvitz–Thompson construction.

Table 11: Experiment E: mean-difference estimates under outcome-dependent completion across 1,000,000 worlds. The population mean difference is zero.

Estimator	Difference mean estimate
Full observation of all worlds	-0.00077
Average over completed comparisons only	$+0.31711$
Inverse-probability correction	$+0.00150$

The inverse-probability estimate has a 95% Monte Carlo interval of $[-0.00313, 0.00612]$, which includes the true mean difference zero. The analytic naive mean is 0.3191538, and the analytic IPW per-world variance is 5.5555556. Correcting selective completion removes the large directional bias but raises the measured per-world variance from 0.99966 to 5.56668. The result quantifies the precision cost of correction under this known-propensity completion rule.

Expected acquisitions are 1.5 per selectively completed world, compared with two for a complete pair. At the same variance of the final mean, IPW requires 5.5556 times as many worlds in this model, giving expected acquisition cost ratio $1.5(5.5556)/2 = 4.1667$. Lower per-world acquisition cost can therefore accompany higher cost at a fixed estimation precision.

G.6 Information allocation and reporting quality

The five studies isolate the reporting-precision benefit of contrast protection, its allocation trade-off, the role of correct covariance, completion under sequential inference, and the variance cost of selective-completion correction. Their Gaussian model endpoints complement the four replay benchmarks, whose bias and total-cost aggregates concern the application-level joint task.

H Joint Reporting Estimands and Cost Accounting

To separate the effect of reporting from the effect of choosing a different winner, fix the selected configuration within each reporting comparison. For setting s and outer replicate r, let $\hat{i}_{sr}$ denote this frozen winner. With $m \in \{N, P\}$ indexing the naïve and PAIRS reports, respectively, define

$$e^m_{sr} = \widehat{\mu}^m_{sr} - \mu_{s,\hat{i}_{sr}}, \qquad \widehat{b}^m_s = \frac{1}{N_s}\sum_{r=1}^{N_s} e^m_{sr}. \tag{H.1}$$

This same-winner comparison holds the target configuration fixed within each pair. The all-report estimand includes diagnostic reports for uncertified incumbents. An available-report analysis averages over the produced reports and states its corresponding denominator; completion and joint success use all outer runs.

An equal-setting absolute-bias summary first averages signed errors within each setting and then takes an average of their absolute values:

$$B_m = \frac{1}{S}\sum_{s=1}^{S} |\widehat{b}^m_s|, \qquad R_{\text{bias}} = 1 - \frac{B_P}{B_N}, \quad B_N > 0. \tag{H.2}$$

Equal weighting gives each setting the same contribution regardless of its number of outer replicates. This statistic differs from mean absolute error, which averages $|e_{sr}^m|$, and from pooled signed bias, which averages all signed errors with replicate-count weights. Averaging setting-specific percentage reductions also produces a different estimand. Signed setting-level biases accompany the relative reduction because taking absolute values obscures direction and becomes non-smooth near zero. With a finite reference mean in place of $\mu_{s,\hat{i}_{sr}}$, these errors are reference-relative. A shared reference cancels in the paired difference between reports, but not in their individual absolute biases.

For a full selection-and-reporting method m, let C_{sr}^m indicate issuance of a joint certificate and let G_{sr}^m denote satisfaction of both the selection and reporting targets in Eq. (2). Setting-specific targets are fixed before outcomes are inspected. With $N = \sum_s N_s$, the pooled completion, correct-joint-output, and incorrect-certificate rates are

$$\begin{aligned} \widehat{Q}_m &= N^{-1} \sum_{s,r} C_{sr}^m, & \widehat{J}_m &= N^{-1} \sum_{s,r} C_{sr}^m \mathbf{1}\{G_{sr}^m\}, \\ \widehat{W}_m &= N^{-1} \sum_{s,r} C_{sr}^m \mathbf{1}\{(G_{sr}^m)^c\}, & \widehat{A}_m &= 1 - \widehat{Q}_m. \end{aligned} \tag{H.3}$$

All four rates use every outer replicate. Thus $\widehat{Q}_m = \widehat{J}_m + \widehat{W}_m$, and abstentions count as failures for correct joint output. Completion is defined by issuance of a joint certificate. Conditional correctness among certified outputs, $\widehat{J}_m / \widehat{Q}_m$, is reported separately when completion is nonzero. Population comparisons use paired uncertainty intervals and a prespecified equivalence or noninferiority margin.

To distinguish the sources of incorrect certification, let $E_{\text{sel}} = \{\mu_* - \mu_{\hat{i}} > \varepsilon\}$ and $E_{\text{rep}} = \{|\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}}| > \eta\}$. Define

$$\begin{aligned} W_{\text{sel}} &= \Pr(A_B \cap E_{\text{sel}}), & W_{\text{rep}} &= \Pr(A_B \cap E_{\text{rep}}), \\ W_{\text{both}} &= \Pr(A_B \cap E_{\text{sel}} \cap E_{\text{rep}}), & W_B &= W_{\text{sel}} + W_{\text{rep}} - W_{\text{both}}. \end{aligned} \tag{H.4}$$

Empirical versions count these events over the same all-run denominator. The allocated ratio $\delta_{\text{sel}} : \delta_{\text{rep}}$ limits errors; it does not prescribe the observed proportions of selection and reporting failures. Report-interval noncoverage, error exceeding η, and a negative post-decision audit are also different events.

Let T_{sr}^m be the number of attempted candidate–world evaluations charged to a run under a common accounting scope. The aggregate cost ratio is

$$R_{\text{cost}} = \frac{\sum_{s,r} T_{sr}^{\text{baseline}}}{\sum_{s,r} T_{sr}^{\text{PAIRS}}}. \tag{H.5}$$

This ratio includes unsuccessful and abstaining runs and is not the arithmetic mean of setting-specific ratios. Proposal, pilot, racing, reporting, gate, audit, and reference computations are recorded separately. For each stage, attempts form the disjoint partition

$$N_{\text{attempt}} = N_{\text{accepted}} + N_{\text{completed,unused}} + N_{\text{failed}}. \tag{H.6}$$

Completed but unused evaluations include successful computations discarded with an incomplete whole block. A successful retry is already in the accepted or completed-unused count; it is not a fourth additive category. Each physical attempt is charged once, including failures. A protected observation serving two inferential roles incurs one acquisition cost, and shared setup costs are allocated once under a stated amortization rule. Appendix P applies these rules to a stage-resolved ledger.

A cost-at-target comparison additionally fixes completion and correct-output requirements $(\gamma_\star, J_\star)$ before inspecting outcomes. If $Q_m(C)$ and $J_m(C)$ denote those probabilities along a declared cost curve, its target cost is

$$C_m^\star = \inf\{C : Q_m(C) \geq \gamma_\star,\ J_m(C) \geq J_\star\}. \tag{H.7}$$

A fixed-log cost ratio describes expenditure at the evaluated operating points. A cost-at-target comparison additionally assesses whether both procedures meet the targets, using a prespecified noninferiority comparison when appropriate. In the finite-reference study, the correct-joint-output difference interval $[-4.21, 6.32]$ pp crosses the -2 pp margin, leaving the noninferiority comparison inconclusive. The cost-ratio interval describes the more precisely estimated expenditure contrast.

We report accepted logical evaluations, attempted evaluations, measured accelerator-hours, and wall-clock makespan as distinct quantities. Accelerator-hours accumulate device allocation times; wall-clock measures the full execution interval, including concurrency. Acquisition ratios and elapsed-time ratios are computed from their respective measurements.

Outer-replicate uncertainty calculations resample within settings while retaining each replicate's complete vector of method outcomes. Shared splits or task orders define higher-level clusters when applicable. To incorporate finite-reference uncertainty as well, whole reference blocks must be resampled jointly across candidates, with each reference draw reused for every dependent method and replicate. The complete nonlinear statistic is then recomputed, including absolute values and aggregate ratios. This reference-aware calculation is distinct from holding the reference bank fixed during outer-replicate resampling. These intervals describe the sampled settings; generalization to unseen settings requires a separate sampling model.

For paired binary success outcomes A, B, put $p_{10} = \Pr(A = 1, B = 0)$ and $p_{01} = \Pr(A = 0, B = 1)$. Their difference has variance $p_{10}+p_{01}-(p_{10}-p_{01})^2$. Marginal success rates determine $p_{10}-p_{01}$ but not the discordance sum, so the paired uncertainty is not identified by marginal rates alone. For a non-smooth absolute-bias ratio, simultaneous signed-bias intervals can instead be propagated conservatively: if they imply $|b_N| \in [l_N, u_N]$ and $|b_P| \in [l_P, u_P]$ with $l_N > 0$, then

$$1 - \frac{|b_P|}{|b_N|} \in \left[1 - \frac{u_P}{l_N},\ 1 - \frac{l_P}{u_N}\right]. \tag{H.8}$$

This interval inherits the simultaneous coverage event of its inputs. Shared reference uncertainty must already be included in those input intervals.

H.1 Reporting subsets and completion denominators

The reporting subset determines which errors enter the same-winner comparison in Equation (H.1). Let $\mathcal{R}_s$ contain the outer replicates in setting s retained for both reports. The subset-specific mean is

$$\widehat{b}_s^m(\mathcal{R}_s) = \frac{1}{|\mathcal{R}_s|} \sum_{r \in \mathcal{R}_s} e_{sr}^m, \qquad |\mathcal{R}_s| > 0, \quad m \in \{N, P\}. \tag{H.9}$$

Equation (H.2) then aggregates these setting-level means. The two reports use the same retained outer IDs and the same winner within each ID. The denominator counts available report pairs.

Three subsets answer different questions. The all-available subset contains every outer ID with both same-winner reports, including uncertified diagnostic reports. The PAIRS-certified subset intersects those available pairs with the PAIRS joint-certificate flag. The both-method-certified subset additionally requires the complete comparator program to issue a certificate, yielding the intersection of the two certification events. Each subset retains its own denominator when report availability or certification differs.

Table 12: Twelve-run reporting cohort. Both error columns refer to the same winner within each row and are measured in pp. The certificate flag is the PAIRS joint-certificate flag. Rows with paired reports enter the bias calculation.

Outer ID	Paired reports	Joint certified	Naïve error	PAIRS error
01	Yes	Yes	0.30	−0.08
02	Yes	Yes	0.35	−0.04
03	Yes	Yes	0.40	0.00
04	Yes	Yes	0.45	0.02
05	Yes	Yes	0.50	0.04
06	Yes	Yes	0.55	0.06
07	Yes	Yes	0.60	0.08
08	Yes	Yes	0.65	0.08
09	Yes	No	0.60	0.12
10	Yes	No	0.60	0.12
11	No	No	Missing	Missing
12	No	No	Missing	Missing

Table 13: Absolute signed bias and mean absolute error (MAE) on common reporting subsets of Table 12. Errors are in pp; reductions are percentages. Both-method-certified membership additionally requires the comparator certificate flags.

		Absolute signed bias			MAE		
Subset	n	Naïve	PAIRS	Reduction	Naïve	PAIRS	Reduction
Available report pairs	10	0.500	0.040	92.00	0.500	0.064	87.20
PAIRS-certified pairs	8	0.475	0.020	95.79	0.475	0.050	89.47

Table 12 contains ten available report pairs. Their empirical absolute biases are $B_N = 0.50$ pp and $B_P = 0.04$ pp, giving $1 - 0.04/0.50 = 92\%$. Restricting both columns to the eight certified outer IDs yields $B_N = 0.475$ pp and $B_P = 0.020$ pp, giving $1 - 0.020/0.475 = 95.7895\%$. Both analyses take the absolute value after averaging signed errors within the corresponding subset.

The distinction is numerically consequential in this cohort (Table 13). Across the ten available PAIRS reports, the signed errors sum to 0.40 pp while their absolute values sum to 0.64 pp. Absolute signed bias is therefore 0.040 pp and mean absolute error is 0.064 pp. The respective reductions relative to the same naïve reports are 92.00% and 87.20%. Restricting to the eight certified reports changes both summaries, while preserving their distinct definitions.

Completion retains all twelve outer runs in its denominator: $\widehat{Q} = 8/12 = 66.6667\%$, whereas paired-report availability is $10/12 = 83.3333\%$. The four joint abstentions comprise two runs with diagnostic reports and two without reports. Changing the reporting subset changes the bias estimand, while completion and correct joint output retain the all-run denominator.

The conditional unbiasedness in Equation (4) extends to a subset whose membership is fixed by the permitted pre-report information. Let $\mathcal{G}$ denote that information across the outer runs and suppose $\mathcal{R}_s$ is $\mathcal{G}$-measurable, with $\mathbb{E}[e^P_{sr} \mid \mathcal{G}] = 0$. On $|\mathcal{R}_s| > 0$, linearity gives

$$\mathbb{E}\left[\widehat{b}^P_s(\mathcal{R}_s) \mid \mathcal{G}\right] = \frac{1}{|\mathcal{R}_s|} \sum_{r \in \mathcal{R}_s} \mathbb{E}[e^P_{sr} \mid \mathcal{G}] = 0. \tag{H.10}$$

Thus the selection transcript and precommitted reporting requirements can define a conditionally unbiased reporting subset. Filtering on the report value, report error, or reference-scored correctness would introduce outcome dependence. The empirical absolute bias remains a nonlinear finite-sample statistic even when signed reporting error has expectation zero.

I Executed Bounded-Noise Racing with Fresh Reporting

I.1 Executed Bounded-Outcome Sequential Study

The executed PEBR-FR-SYN procedure uses independent bounded shared-noise worlds in $[0,1]^K$, with $K = 16, 24, 12, 32$ and 190 outer runs. At $n_\ell = 32\,2^\ell$, $\ell = 0, \ldots, 8$, allocate $\alpha_{ij,\ell} = \delta_{\mathrm{sel}}/[\binom{K}{2}(\ell+1)(\ell+2)]$. With unbiased difference sample variance $s^2_{ij,n}$, its implemented radius is

$$r_{ij,n} = \sqrt{\frac{2s^2_{ij,n}\log(4/\alpha_{ij,\ell})}{n}} + \frac{14\log(4/\alpha_{ij,\ell})}{3(n-1)}. \tag{I.1}$$

All active candidates complete each requested block before simultaneous zero-gap elimination ($L_{j,i} > 0$). With surviving set A and $U_{i,i} = 0$, its leader stops when $\max_{j \in A} U_{j,\hat{i}} \le \varepsilon$. The next block leaves the fixed fresh-report and 128-observation audit reserve affordable. Pairwise prefix coverage follows from summing $\alpha_{ij,\ell}$ (Maurer & Pontil, 2009). Appendix I specifies the bounded observation law and complete execution traces for this rule.

PEBR-FR-SYN produces 190 correct joint certificates: empirical $(Q, J, W) = (100\%, 100\%, 0\%)$. Its 9,783,367 evaluations comprise 9,459,712 selection, 299,335 reporting, and 24,320 audit acquisitions. The HFR-SYN boundary control uses 9,488,455 evaluations and has $(Q, J, W) = (0, 0, 0)$ at the same caps. The comparison shows how boundary width affects completion within a budget.

The saved trajectory continues at $n = 4096$ because $U = 0.010270 > 0.01$, and certifies at 8192 with $U = 0.001542$ (Table 17).

Two additional paths isolate contamination resolution: clean EB certifies after $90{,}112$ acquisitions; a persistent 0.5% whole-world contamination correction yields abstention at $524{,}288$ (Appendix U).

We evaluate Paired Empirical-Bernstein Racing with Fresh Reporting (PEBR-FR-SYN, version 0.1.0) on four bounded-noise selection tasks. Each evaluation acquires one candidate–world observation. The study tests sequential elimination, joint certification, and stage-resolved acquisition costs under an explicitly specified observation model.

I.2 Observation model and reporting protocol

For candidate i and independent world t, generate

$$X_{i,t} = \mu_i + 0.09Z_t + 0.025E_{i,t}, \qquad Z_t, E_{i,t} \overset{\text{iid}}{\sim} \text{Uniform}[-1, 1]. \tag{I.2}$$

The K means are equally spaced over $[0.66, 0.78]$. All scores lie in $[0, 1]$. The common component Z_t cancels in paired differences, but the algorithm estimates their variance from observations rather than using its population value. Table 14 gives the four configurations; tolerances are in native score units.

Table 14: Bounded-noise task configurations. R is the number of outer runs, B the per-run total evaluation cap, and m the fixed fresh-reporting sample size.

Task	K	R	ϵ	η	B	m
SYN-U1	16	50	0.010	0.030	60,000	2,334
SYN-U2	24	50	0.015	0.035	80,000	1,715
SYN-U3	12	45	0.020	0.040	65,000	1,313
SYN-U4	32	45	0.025	0.050	100,000	840

We set $\delta_{\text{sel}} = 0.02$ and $\delta_{\text{rep}} = 0.03$. Every frozen incumbent, including an uncertified one, receives a report from independent fresh worlds, using

$$m = \left\lceil \frac{\log(2/\delta_{\text{rep}})}{2\eta^2} \right\rceil, \qquad \widehat{\mu}_{\text{rep}} = m^{-1} \sum_{t=1}^{m} X^{\text{rep}}_{\hat{i},t}. \tag{I.3}$$

Each run also generates 128 independent audit observations. Audit outcomes do not influence selection or reporting. Before requesting a complete query block, the procedure must be able to pay for that block and the reporting and audit stages without exceeding B. Reporting an uncertified incumbent does not create a joint certificate.

I.3 Scheduled confidence bounds and decisions

Intervals are updated at predetermined pair counts $n_\ell = 32\,2^\ell$, for $\ell = 0, \ldots, 8$. Queries are fixed from the completed history before revealing the next block. Candidates start together, eliminated candidates are not reintroduced, and each observed pair follows a prefix of its independent-world sequence. Elimination is a simultaneous update after a whole block completes.

For each unordered pair, allocate

$$\alpha_{ij} = \frac{\delta_{\text{sel}}}{\binom{K}{2}}, \qquad \alpha_{ij,\ell} = \frac{\alpha_{ij}}{(\ell+1)(\ell+2)}. \tag{I.4}$$

With $D_t = X_{i,t} - X_{j,t} \in [-1, 1]$, compute its unbiased sample variance from the online blockwise mean and covariance summaries:

$$s^2_{ij,n} = \frac{1}{n-1} \sum_{t=1}^{n} (D_t - \overline{D}_{ij,n})^2 = s^2_{i,n} + s^2_{j,n} - 2s^{\text{cov}}_{ij,n}. \tag{I.5}$$

The empirical-Bernstein fixed-time inequality, scaled to a range of width two and applied to both tails (Maurer & Pontil, 2009), gives the implemented radius

$$r_{ij,n}^{\mathrm{EB}} = \sqrt{\frac{2s_{ij,n}^2 \log(4/\alpha_{ij,\ell})}{n}} + \frac{14 \log(4/\alpha_{ij,\ell})}{3(n-1)}. \tag{I.6}$$

The HFR-SYN control uses the same query and reporting protocol with

$$r_{ij,n}^{\mathrm{H}} = \sqrt{\frac{2 \log(2/\alpha_{ij,\ell})}{n}}. \tag{I.7}$$

Both use $[L_{ij,n}, U_{ij,n}] = [\overline{D}_{ij,n} - r_{ij,n}, \overline{D}_{ij,n} + r_{ij,n}]$. Previously computed intervals may remain displayed between checkpoints, but no new fixed-time interval is inspected without an allocated error budget.

Candidate i is eliminated when some currently active j satisfies $L_{j,i}(n) > 0$. For the empirical leader $\hat{i}$ of the remaining set S_n, selection is certified when

$$\max_{j \in S_n} U_{j,\hat{i}}(n) \le \epsilon, \tag{I.8}$$

where the self-comparison is zero. If the next complete block is unaffordable or the final checkpoint is reached without certification, the procedure abstains from issuing a joint certificate.

This implementation uses zero-gap elimination, equivalently $U_{i,j} < 0$. Both this rule and the $U_{i,j} < -\varepsilon$ rule in the general presentation preserve a population-best candidate on simultaneous coverage. The zero-gap rule determines the active sets and acquisition costs reported here.

All M candidates receive 32 observations before the first inspection, so this schedule charges at least $32M$ selection evaluations before any decision: 512, 768, 384, and 1,024 for its four pool sizes. The replay-selection table reports means of 31, 46, 24, and 62 accepted selection observations under its study-specific acquisition schedule; the present study uses the checkpoint schedule above.

Proposition 10 (Validity at the scheduled checkpoints)**.** *Under independent worlds, the prefix-preserving query protocol above, and the bounded observation model, either boundary satisfies*

$$\Pr(\exists i < j, \ell : \mu_i - \mu_j \notin [L_{ij,n_\ell}, U_{ij,n_\ell}]) \le \delta_{\mathrm{sel}},$$

where the union includes intervals actually inspected. The probability of issuing a joint certificate with either regret exceeding ϵ or reporting error exceeding η is at most $\delta_{\mathrm{sel}} + \delta_{\mathrm{rep}} = 0.05$.

Proof. Pre-generate the independent potential world vectors. For every fixed pair and deterministic checkpoint, the two-sided empirical-Bernstein or Hoeffding bound fails with probability at most $\alpha_{ij,\ell}$. Adaptively observing fewer of these prefixes cannot enlarge the union of failures. Summing over pairs and checkpoints gives the claim because $\sum_{\ell \ge 0}[(\ell+1)(\ell+2)]^{-1} = 1$. On simultaneous coverage, no maximizer can be eliminated by a strictly positive lower bound. Thus a population maximizer remains in S_n, and Equation (I.8) certifies global ϵ-optimality. Conditional on the full selection transcript, the report uses a fixed number of independent bounded observations of the frozen incumbent. Hoeffding's inequality and Equation (I.3) bound its error probability by δ_{rep}. A union bound proves the joint guarantee. This argument permits adaptive stopping at allocated checkpoints; it does not require independence between candidate-pair intervals. □

I.4 Completed runs and an executed stopping trajectory

PEBR-FR-SYN issued 190 correct joint certificates in 190 runs, at a total cost of 9,783,367 evaluations (Table 15). HFR-SYN used 9,488,455 evaluations and issued no joint certificates within the same task-specific caps and checkpoint schedule. Thus the boundary comparison separates expenditure from completion: the empirical-Bernstein procedure completes every run, whereas the Hoeffding control exhausts its stopping opportunities without certification.

The fixed report and audit counts identify the stage decomposition in Table 16. Selection counts equal the recorded total minus those two stages. Every column is measured in acquired candidate–world observations. The bounded tasks use $50, 50, 45, 45$ outer runs; the replay-selection benchmarks use $60, 50, 40, 40$.

Table 15: Executed PEBR-FR-SYN results. Total evaluations include selection, fresh reporting, and audit. Correctness is assessed against the generating means.

Task	Joint certificates / runs	Correct joint certificates	Evaluations
SYN-U1	50/50	50	2,394,332
SYN-U2	50/50	50	2,811,894
SYN-U3	45/45	45	1,249,613
SYN-U4	45/45	45	3,327,528
Total	190/190	190	9,783,367

Table 16: PEBR-FR-SYN acquisition ledger over all runs. Reporting is $R_s m_s$; audit is $128R_s$; selection is the remainder of the recorded total.

Task	Selection	Fresh report	Audit	Total
SYN-U1	2,271,232	116,700	6,400	2,394,332
SYN-U2	2,719,744	85,750	6,400	2,811,894
SYN-U3	1,184,768	59,085	5,760	1,249,613
SYN-U4	3,283,968	37,800	5,760	3,327,528
Total	9,459,712	299,335	24,320	9,783,367

Table 17 shows the critical comparison between candidates 1 and 0 in SYN-U1, replicate 0. At $n = 4,096$, the competitor's advantage upper bound is $0.010270 > \epsilon = 0.010$, requiring continuation. At $n = 8,192$, it is 0.001542, and selection is certified. The run uses 48,128 selection evaluations and $2,334 + 128$ reporting and audit evaluations, totaling 50,590.

Table 17: Executed SYN-U1 trajectory for $D = X_1 - X_0$. The active count is measured on entry to each checkpoint. Values are rounded for display.

n	Active	$\overline{D}$	s^2	r^{EB}	$U_{1,0}$	Decision
1,024	16	−0.006717	0.00043009	0.066467	0.059750	Continue
2,048	9	−0.007719	0.00041579	0.034564	0.026845	Continue
4,096	5	−0.007812	0.00041882	0.018081	0.010270	Continue
8,192	3	−0.007988	0.00041844	0.009531	0.001542	Certify

The implementation entry point is `python src/selector.py --out selector`, with core functions `run_one(cfg, replicate, method="EB")` and `radius(n, sample_var, alpha_look, method)`. The run identifier is `NEW-SYN-20260919-95615365a334`, with master seed 2026091970.

Recomputation reproduces the boundaries in all 2,934 saved checkpoint records and the decisions in eight complete paths, one per task and boundary. The run identifier links these checks to the bounded-noise execution above.

The checkpoint statistics also reconcile the boundary and acquisition ledger. At $n = 4096$, $\ell = 7$, $K = 16$, and $\delta_{\mathrm{sel}} = 0.02$ give $\alpha_{ij,7} = 0.02/(120 \cdot 8 \cdot 9) = 2.3148148 \times 10^{-6}$. Substitution of the displayed variance into Eq. (I.6) separates the variance and bounded-range terms:

$$r^{\mathrm{EB}}_{ij,4096} \simeq 0.00171381 + 0.01636749 = 0.01808131. \tag{I.9}$$

This agrees with the recorded radius up to display rounding and retains the continuation decision. At $n = 8192$, the displayed statistics give $r^{\mathrm{EB}} \simeq 0.0095305514$ and $U_{1,0} \simeq 0.0015425514 < 0.010$. These calculations check the displayed critical pair; the global stopping condition in Eq. (I.8) also checks every other active competitor. The active-set sizes and observation increments give the exact selection acquisition count

$$16(1024) + 9(1024) + 5(2048) + 3(4096) = 48{,}128. \tag{I.10}$$

Adding 2,334 fresh reporting and 128 audit observations recovers the recorded total of 50,590. Thus the run identity, boundary, checkpoint statistics, decisions, and cost ledger describe the same executed path.

Table 18: Finite-reference means and simultaneous intervals for DEMO-C000 in the controlled Gaussian study, with known generating means.

Task	Oracle mean	Reference mean	Reference interval
DEMO-B1	0.860000	0.859752	$[0.857276, 0.862228]$
DEMO-B2	0.835000	0.833928	$[0.830214, 0.837642]$
DEMO-B3	0.810000	0.809646	$[0.804694, 0.814598]$
DEMO-B4	0.785000	0.784026	$[0.776598, 0.791454]$

The displayed path also gives a local sensitivity calculation. At $n = 4096$, this comparison would pass if a valid alternative half-width were at most $0.010 - (-0.007812) = 0.017812$, about 1.49% below the displayed EB radius. If all other certificate conditions also passed and the earlier query path were unchanged, omitting the final 3×4096 acquisitions would reduce total cost from 50,590 to 38,302, or about 24.29%. This conditional calculation quantifies the cost sensitivity to the boundary at this checkpoint.

J Executed Finite-Reference Uncertainty Study

J.1 Executed Reference-Uncertainty Study

Four Gaussian reference banks contain 84 candidates and 1024 worlds each. Methods and outer runs share their task's bank. Finite references change this controlled study's macro bias reduction from 92% to 90.7582%; outer-plus-reference resampling gives $[84.64, 95.44]\%$. Simultaneous intervals classify 180 PAIRS-DEMO certificates as 141 definitely correct, three incorrect, and 36 inconclusive (Appendix J).

The separate 2,000-run Gaussian comparison retains complete arrays. Fixed-64 paired and fixed-149 independent selection have matched contrast precision, followed by 19 fresh report and eight audit observations. Both complete 99.10% of runs; reference-relative joint correctness is 94.75% versus 94.85%, with paired outer-plus-reference difference interval $[-1.85, 1.15]$ pp. Their 2.30220 acquisition-cost ratio isolates fixed-design covariance effects (Appendix T).

We study finite-reference evaluation of frozen selection and reporting outputs. The four controlled tasks, DEMO-B1–B4, combine calibrated joint-output records with independent Gaussian reference banks. Their oracle bias reduction is 92%; the analysis below measures how finite reference information changes its point estimate and uncertainty.

Each task contains 84 candidates, giving 336 candidate means across four tasks. For task s, a reference bank contains 1,024 independent world rows, each preserving all 84 candidate outcomes:

$$X_{swi} = \mu_{si} + \sigma_s\left(\sqrt{0.67}Z_{sw} + \sqrt{0.33}E_{swi}\right), \qquad \sigma_s \in \{0.02, 0.03, 0.04, 0.06\}, \tag{J.1}$$

where the Gaussian innovations are independent except for the shared Z_{sw} within each row. The master seed is 2026091971. This generates 344,064 reference outcomes. Banks are independent across tasks; within a task, all methods and outer replicates share one bank. The archive includes complete world matrices, candidate order, mean covariance, and 570 output-to-reference links. One bank identifier is `NEW-SIM-REF-DEMO-B1-N1024-SEED2026091971-0`.

The offline reference error budget assigns 0.025 to simultaneous candidate-mean intervals and 0.025 to simultaneous pairwise-difference intervals, independently of the selection and reporting budgets. Known Gaussian standard errors and simultaneous correction give the intervals in Table 18 for candidate DEMO-C000.

Sharing a reference bank creates uncertainty that does not average away with the number of outer replicates. Let $p_i = R^{-1}\sum_{r=1}^{R}\mathbf{1}\{\hat{i}_r = i\}$ be the observed selection frequency. The reference-

Table 19: Signed bias by task and macro absolute bias in the controlled study. Finite reference changes point estimates as well as uncertainty.

	Oracle		Finite reference	
Task	Naïve	PAIRS-DEMO	Naïve	PAIRS-DEMO
DEMO-B1	0.020000	0.001000	0.020191	0.001191
DEMO-B2	0.025000	0.002000	0.026079	0.003079
DEMO-B3	0.030000	0.003000	0.029567	0.002567
DEMO-B4	0.025000	0.002000	0.025531	0.002531
Macro absolute bias	0.025000	0.002000	0.025342	0.002342

Table 20: Finite-reference sensitivity of the controlled study. Intervals are approximate 95% percentile bootstrap intervals. Rate differences are percentage points (pp). The comparator is PR-CS-FR-DEMO.

Metric	Point estimate	Outer only	Outer + reference
Absolute bias reduction	90.76%	[85.70, 95.07]%	[84.64, 95.44]%
Selection success vs. uniform	+12.11 pp	[+6.84, +17.89]	[+6.32, +17.89]
PAIRS-DEMO joint success	92.11%	[88.42, 95.79]%	[86.84, 95.26]%
Comparator joint success	91.58%	[87.37, 95.26]%	[85.79, 95.26]%
Joint-success difference	+0.53 pp	[−4.21, +5.26]	[−4.21, +6.32]
Fixed-log cost ratio	2.300	[2.264, 2.337]	[2.264, 2.337]

relative signed bias and its conditional reference variance are

$$\widehat{b} = \frac{1}{R}\sum_{r}\widehat{\mu}_r - \sum_{i} p_i \widehat{\mu}_i^{\mathrm{ref}}, \tag{J.2}$$

$$\mathrm{Var}_{\mathrm{ref}}(\widehat{b} \mid \text{observed selections}) = p^{\top}\,\mathrm{Cov}(\widehat{\mu}^{\mathrm{ref}})p = \frac{\sigma_s^2}{1024}\left(0.67 + 0.33\sum_{i} p_i^2\right). \tag{J.3}$$

The corresponding standard errors are approximately 0.000521, 0.000780, 0.001042, and 0.001563. In particular, repeatedly referencing the same bank leaves the common-noise contribution intact. Table 19 shows the resulting point-estimate change: macro absolute bias reduction becomes $1 - 0.0023420667/0.0253420667 = 90.7582\%$, versus 92% under oracle means. The order of aggregation matters even without reference error: the ratio of oracle macro absolute biases is $1 - 0.002/0.025 = 92\%$, whereas averaging the four task-specific reductions gives $(95 + 92 + 90 + 92)/4 = 92.25\%$. We use the former, as defined in Eq. (H.2).

We propagate outer-run and reference uncertainty using 10,000 paired resamples with seed 2026091972. Within each task, all methods use the same resampled outer indices. Independently, we resample entire reference-world rows, preserving candidate dependence, and compute one reference mean vector shared by all resampled methods and outputs in that task. Candidate choices, certificate flags, and costs remain frozen; reference resampling never reruns selection. Table 20 compares this procedure with resampling outer runs alone.

The two analyses yield identical cost ratios in all 10,000 paired draws because reference resampling leaves the logged decisions and costs fixed. Quality-at-target comparisons depend on the success probabilities after reference propagation. Nonlinearity, skewness, and nonsmoothness near zero bias, tied optima, or success thresholds affect percentile endpoints; the intervals summarize resampling uncertainty, while simultaneous reference intervals provide the output-level classifications below.

The point-scored joint-success rates correspond to 175 correct outputs for PAIRS-DEMO and 174 for PR-CS-FR-DEMO, each out of 190 runs. Both methods issue 180 joint certificates, giving completion $180/190 = 94.7368\%$ and ten abstentions each. Among their certificates, five and six, respectively, are incorrect under reference-point scoring. These point-scored counts differ from classification using simultaneous reference intervals in Table 22.

To identify what these margins say about a paired comparison, let n_{ab} count the runs whose two correct-joint-output indicators are (a, b). All compatible paired tables have the form

$$(n_{11}, n_{10}, n_{01}, n_{00}) = (k, 175 - k, 174 - k, k - 159), \quad k = 159, \ldots, 174. \tag{J.4}$$

Here n_{00} includes incorrect certificates and abstentions. Every table has difference $100/190 = 0.5263$ pp, but discordance ranges from one to 31 runs. Table 21 gives the exhaustive enumeration and the corresponding normal-approximation intervals, holding the reference point scores fixed and ignoring setting strata. The standard error used for this enumeration is

$$\widehat{\mathrm{SE}} = \sqrt{\frac{(n_{10} + n_{01})/190 - (1/190)^2}{189}}. \tag{J.5}$$

The paired bootstrap uses the observed joint records and their setting and reference links. This marginal-count enumeration displays the range of compatible pairing structures; sparse discordance makes its normal approximation particularly sensitive.

Table 21: All paired contingency tables compatible with 175 and 174 correct joint outputs out of 190 runs. Intervals are $100\{1/190 \pm z_{0.975}\widehat{\mathrm{SE}}\}$ in pp for each compatible structure, using fixed reference point scores and unstratified margins.

n_{11}	n_{10}	n_{01}	n_{00}	Discordance	Interval (pp)
159	16	15	0	31	$[-5.232, 6.284]$
160	15	14	1	29	$[-5.043, 6.096]$
161	14	13	2	27	$[-4.847, 5.900]$
162	13	12	3	25	$[-4.645, 5.697]$
163	12	11	4	23	$[-4.433, 5.486]$
164	11	10	5	21	$[-4.213, 5.265]$
165	10	9	6	19	$[-3.981, 5.034]$
166	9	8	7	17	$[-3.737, 4.790]$
167	8	7	8	15	$[-3.479, 4.531]$
168	7	6	9	13	$[-3.202, 4.255]$
169	6	5	10	11	$[-2.903, 3.956]$
170	5	4	11	9	$[-2.576, 3.628]$
171	4	3	12	7	$[-2.209, 3.262]$
172	3	2	13	5	$[-1.785, 2.838]$
173	2	1	14	3	$[-1.264, 2.316]$
174	1	0	15	1	$[-0.505, 1.558]$

For a conservative output-level analysis, simultaneous reference intervals induce regret bounds $R_i^L \le \max_j \mu_j - \mu_i \le R_i^U$. Selection is definitely correct when $R_i^U \le \epsilon$, and definitely incorrect when $R_i^L > \epsilon$. Given a report a and a candidate-mean interval $[\ell_i, u_i]$, reporting is definitely accurate when $\max\{|a - \ell_i|, |a - u_i|\} \le \eta$, and definitely inaccurate when $\mathrm{dist}(a, [\ell_i, u_i]) > \eta$. A joint output is definitely correct if both components are correct, definitely incorrect if either component is incorrect, and otherwise inconclusive.

Table 22 gives the PAIRS-DEMO observed-set identification range $141/190 \le J_{\mathrm{observed}} \le (180 - 3)/190$, or 74.21–93.16%. This range propagates reference uncertainty conditional on the observed outputs; population-success inference additionally involves outer-sampling uncertainty. Inconclusive outputs retain their unresolved status. With preserved reports, complete reference-world matrices support row resampling, means and joint covariance support propagation under distributional or asymptotic assumptions, and simultaneous marginal intervals support conservative output-level bounds.

K Finite-Reference Precision

To examine reference uncertainty separately from selection, we generate a bank of 4,096 Gaussian worlds for 84 candidates using seed 2026091961:

$$X_{i,w} = \mu_i + 0.04\left(\sqrt{0.67}Z_w + \sqrt{0.33}E_{i,w}\right), \qquad \mu_i = 0.8 - 0.001i, \quad i = 0, \ldots, 83, \tag{K.1}$$

Table 22: Conservative reference-based classification of frozen certificates. Each method has 190 observed runs; ten abstentions remain outside the 180 issued joint certificates.

Method	Certificates	Correct	Incorrect	Inconclusive
PAIRS-DEMO	180	141	3	36
PR-CS-FR-DEMO	180	151	0	29

where the common and candidate-specific noise variables are independent standard normals across their respective indices. This controlled study has known means and marginal standard deviation 0.04.

At each fixed bank size n, the pointwise normal half-width is $1.96(0.04)/\sqrt{n}$. A Bonferroni interval simultaneous over all 84 candidates uses critical value 3.433775. Table 23 reports these fixed-n widths and the realized error of candidate 0; the simultaneous guarantee covers candidates within each preselected bank size.

Table 23: Reference-bank precision in the controlled Gaussian study. A larger bank reduces uncertainty, but the realized error along one sample path need not decrease monotonically.

Worlds	Pointwise 95% half-width	Simultaneous half-width	Candidate 0 error
64	.009800	.017169	-.007976
256	.004900	.008584	+.001036
1,024	.002450	.004292	+.001460
4,096	.001225	.002146	+.001080

At $n = 1{,}024$, candidate 0 has reference mean 0.801459820 and simultaneous interval $[0.797167601, 0.805752038]$, while its true mean is 0.8. Given simultaneous intervals $[L_i, U_i]$, a selected candidate's regret lies in

$$\left[\max\{0, \max_j L_j - U_i\},\ \max\{0, \max_j U_j - L_i\}\right]. \tag{K.2}$$

The upper endpoint being at most ε establishes success; the lower endpoint exceeding ε establishes failure; otherwise the result remains inconclusive. At $\varepsilon = 0.01$, Table 24 illustrates all three outcomes, including unresolved candidates on both sides of the true regret threshold.

Table 24: Reference-relative classification at $n = 1{,}024$ and $\varepsilon = 0.01$.

Candidate	True regret	Reference regret interval	Classification
0	.000	[0,.008584]	Definite success
5	.005	[0,.015111]	Inconclusive
12	.012	[.003861,.021030]	Inconclusive
20	.020	[.011710,.028879]	Definite failure

A shared reference also contributes common uncertainty to estimated reporting bias. As an analytic illustration, suppose 190 independent reports of the same candidate have error standard deviation 0.01, and a shared independent reference averages 256 worlds of standard deviation 0.04. The standard error of the report-minus-reference average is

$$\sqrt{0.01^2/190 + 0.04^2/256} = 0.0026031, \tag{K.3}$$

compared with the report-only standard error $0.01/\sqrt{190} = 0.0007255$. The shared reference cancels in the difference between two reports of the same winner, while its uncertainty remains in each method's absolute bias and the resulting bias-reduction ratio.

L Explicit Joint Budgets and Reporting Extensions

L.1 A sufficient budget for a bounded finite pool

A fixed allocation makes the joint guarantee explicit for any finite pool. Assume $M \geq 2$ candidates and independent, identically distributed whole-world vectors $(Y_t(1), \ldots, Y_t(M))$, with every coordinate in $[a, a+R]$ for $R > 0$. Dependence within a world is unrestricted. Every requested observation is available, and a candidate–world evaluation costs one unit. The action set permits complete M-candidate selection worlds followed by fresh winner-only reporting worlds. This fixed-allocation construction gives a distribution-free sufficient budget. The complete two-candidate Gaussian action model of Theorems 2 and 3 instead supports the instance-dependent complexity characterization.

Fix $\varepsilon, \eta > 0$, $\delta_{\mathrm{sel}}, \delta_{\mathrm{rep}} \in (0,1)$, and $K = \binom{M}{2}$. Evaluate all candidates on exactly n_S selection worlds, where

$$n_S = \left\lceil \frac{2R^2}{\varepsilon^2} \log \frac{2K}{\delta_{\mathrm{sel}}} \right\rceil, \qquad r_{n_S} = R\sqrt{\frac{2}{n_S} \log \frac{2K}{\delta_{\mathrm{sel}}}}. \tag{L.1}$$

Choose $\hat{i} \in \arg\max_i \bar{Y}_i$ using a fixed deterministic tie rule, and set $U(j,i) = \bar{Y}_j - \bar{Y}_i + r_{n_S}$. Since $\bar{Y}_j - \bar{Y}_{\hat{i}} \leq 0$ and $r_{n_S} \leq \varepsilon$, the selection certificate $\max_{j \neq \hat{i}} U(j, \hat{i}) \leq \varepsilon$ is available for every realized selection sample. Freeze the winner and then collect exactly

$$n_R = \left\lceil \frac{R^2}{2\eta^2} \log \frac{2}{\delta_{\mathrm{rep}}} \right\rceil \quad \text{fresh winner observations}, \qquad B_0^{\mathrm{joint}} = Mn_S + n_R. \tag{L.2}$$

Let $\widehat{\mu}_{\hat{i}}$ be their mean and return $I_{\hat{i}} = [\widehat{\mu}_{\hat{i}} - h_R, \widehat{\mu}_{\hat{i}} + h_R]$, where $h_R = R\sqrt{\log(2/\delta_{\mathrm{rep}})/(2n_R)} \leq \eta$.

Proposition 11 (Bounded finite-pool joint budget)**.** *For $B \geq B_0^{\mathrm{joint}}$, the construction above completes with probability one, satisfies*

$$\Pr\{\mu_* - \mu_{\hat{i}} > \varepsilon \text{ or } |\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}}| > \eta\} \leq \delta_{\mathrm{sel}} + \delta_{\mathrm{rep}}, \tag{L.3}$$

and has $\mathbb{E}[\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}} \mid \mathcal{S}] = 0$. Its reporting interval has conditional coverage at least $1 - \delta_{\mathrm{rep}}$ given $\mathcal{S}$.

Proof. For each unordered pair, $Y_t(i) - Y_t(j) \in [-R, R]$ is independent across worlds. The two-sided Hoeffding inequality (Maurer & Pontil, 2009) gives

$$\Pr\left\{\left|\bar{Y}_i - \bar{Y}_j - (\mu_i - \mu_j)\right| > r_{n_S}\right\} \leq 2\exp\left(-\frac{n_S r_{n_S}^2}{2R^2}\right) = \frac{\delta_{\mathrm{sel}}}{K}.$$

A union bound covers all pairs, in either orientation, with probability at least $1 - \delta_{\mathrm{sel}}$. On this event, $\mu_j - \mu_{\hat{i}} \leq U(j, \hat{i}) \leq \varepsilon$ for every j. Conditional on $\mathcal{S}$, the reporting observations are independent draws from the frozen winner's target law. Their mean is unbiased and Hoeffding's inequality bounds its two-sided error beyond h_R by δ_{rep}. Integrating this conditional bound and taking a union bound proves Eq. (L.3). Both sample counts are fixed and the selection certificate is always available, so the declared budget suffices for completion on every sample path. □

A marginal-mean union bound gives the same joint guarantee with an alternative selection allocation:

$$n'_S = \left\lceil \frac{2R^2}{\varepsilon^2} \log \frac{2M}{\delta_{\mathrm{sel}}} \right\rceil, \qquad h_S = R\sqrt{\frac{\log(2M/\delta_{\mathrm{sel}})}{2n'_S}} \leq \frac{\varepsilon}{2}. \tag{L.4}$$

Indeed, evaluating all candidates on n'_S complete worlds gives $\Pr\{\max_i |\bar{Y}_i - \mu_i| > h_S\} \leq \delta_{\mathrm{sel}}$ by marginal Hoeffding inequalities and a union bound. For a true maximizer i_* and the deterministically tie-broken empirical leader $\hat{i}$, the complementary simultaneous-coverage event implies

$$\mu_* - \mu_{\hat{i}} \leq (\mu_{i_*} - \bar{Y}_{i_*}) + (\bar{Y}_{\hat{i}} - \mu_{\hat{i}}) \leq 2h_S \leq \varepsilon.$$

The observable bound $U'(j, \hat{i}) = \bar{Y}_j - \bar{Y}_{\hat{i}} + 2h_S \leq \varepsilon$ provides a certificate on every sample path. Attaching the same fresh reporting count n_R therefore preserves completion probability one, conditional reporting unbiasedness, and joint error at most $\delta_{\mathrm{sel}} + \delta_{\mathrm{rep}}$, with total budget $Mn'_S + n_R$.

Table 25: Deterministic sufficient joint budgets in candidate–world evaluations for $R = 100$, $\varepsilon = 0.2$ pp, $\eta = 1$ pp, $\delta_{\text{sel}} = 0.02$, and $\delta_{\text{rep}} = 0.03$. Both columns include the same $n_R = 20{,}999$ fresh winner observations.

M	Pairwise allocation $Mn_S + n_R$	Marginal allocation $Mn'_S + n_R$
12	52,789,955	42,561,467
16	75,162,295	59,043,079
24	122,727,863	93,419,711
32	173,008,967	129,155,527

Table 26: Deterministic reporting design counts at $\delta_{\text{rep}} = 0.03$. The Hoeffding row assumes $Y \in [0, 100]$ and uses Eq. (L.2). The Gaussian row instead assumes an exactly Gaussian score law with known $\sigma = 2$ pp. Selection cost is additional to both fresh winner sample counts.

η (pp)	0.2	0.5	1	2	3	3.5	4	5
Hoeffding	524,964	83,995	20,999	5,250	2,334	1,715	1,313	840
Gaussian	471	76	19	5	3	2	2	1

The pairwise and marginal constructions can be compared before sampling: choose the smaller count using only the declared $M, R, \varepsilon, \delta_{\text{sel}}$ and budget, and use its corresponding certificate. This deterministic choice yields the sufficient budget $M \min\{n_S, n'_S\} + n_R$ without an additional error allocation because the certificate choice is fixed before outcomes are observed.

Reporting counts depend on the observation model. For independent Gaussian winner observations with known standard deviation σ, the fresh-mean normal interval requires

$$n_R^G = \left\lceil \frac{z^2_{1-\delta_{\text{rep}}/2}\sigma^2}{\eta^2} \right\rceil,$$

where z_p is the standard-normal p-quantile. Table 26 compares this rule at $\sigma = 2$ pp with the bounded Hoeffding rule. Both use the same unit cost per fresh winner observation; their distributional assumptions determine the required counts.

L.2 A PILOT-ASSISTED FRESH-REPORTING EXTENSION

A separate pilot can adapt reporting cost to the winner's variance while keeping the final mean unbiased. We analyze the following pilot-assisted reporting design. After freezing $\hat{i}$, draw $2m$ independent pilot observations $Y_1^P, \ldots, Y_{2m}^P$ of that winner, with a prespecified integer $m \geq 1$. The pilot is independent of selection and is excluded from the final reporting mean. For $\delta_P \in (0, 1)$, define

$$Z_k = \frac{(Y_{2k}^P - Y_{2k-1}^P)^2}{2}, \qquad v_U = \min\left\{ \frac{R^2}{4},\ \bar{Z} + \frac{R^2}{2}\sqrt{\frac{\log(1/\delta_P)}{2m}} \right\}. \tag{L.5}$$

Conditional on the pre-pilot selection transcript, $\mathbb{E}Z_k = v_{\hat{i}} := \text{Var}(Y(\hat{i}))$ and $0 \leq Z_k \leq R^2/2$. One-sided Hoeffding concentration, together with $v_{\hat{i}} \leq R^2/4$, therefore gives $\Pr\{v_{\hat{i}} > v_U \mid \mathcal{S}\} \leq \delta_P$. Freeze the pilot-dependent count

$$n_R^P = \left\lceil \frac{2v_U + (2/3)R\eta}{\eta^2} \log\frac{2}{\delta_{\text{rep}}} \right\rceil, \tag{L.6}$$

then draw n_R^P new winner observations independently of selection and pilot. Report their mean with interval $[\widehat{\mu}_{\hat{i}} - \eta, \widehat{\mu}_{\hat{i}} + \eta]$.

Proposition 12 (Pilot-assisted reporting). *Let $\mathcal{T} = \sigma(\mathcal{S}, Y_1^P, \ldots, Y_{2m}^P)$. The fresh reporting mean satisfies $\mathbb{E}[\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}} \mid \mathcal{T}] = 0$, and hence the same identity conditional on $\mathcal{S}$. On each $\mathcal{T}$-transcript for which $v_{\hat{i}} \leq v_U$, reporting coverage is at least $1 - \delta_{\text{rep}}$. Unconditionally, reporting error is at most $\delta_P + \delta_{\text{rep}}$. Combining this report with Proposition 11's selection stage gives joint error at most $\delta_{\text{sel}} + \delta_P + \delta_{\text{rep}}$.*

Proof. Given $\mathcal{T}$, the winner and reporting count are fixed, while the fresh observations remain independent with mean $\mu_{\hat{i}}$ and variance $v_{\hat{i}}$. This proves unbiasedness on every such transcript. The two-sided Bernstein bound is

$$\Pr\{|\widehat{\mu}_{\hat{i}} - \mu_{\hat{i}}| > \eta \mid \mathcal{T}\} \le 2\exp\left\{-\frac{n_R^P \eta^2}{2v_{\hat{i}} + (2/3)R\eta}\right\}.$$

When $v_{\hat{i}} \le v_U$, Eq. (L.6) makes this at most δ_{rep}. Splitting over the variance-bound event gives the unconditional bound; adjoining the selection failure event gives the joint bound. The $1 - \delta_{\mathrm{rep}}$ conditional coverage statement applies on the variance-bound event, whereas conditional unbiasedness holds throughout. □

The pilot costs $2m$ winner evaluations. With the fixed selection stage, the realized evaluation cost is $Mn_S + 2m + n_R^P$. A deterministic reserve of $Mn_S + 2m + \lceil [R^2/2 + (2/3)R\eta] \log(2/\delta_{\mathrm{rep}})/\eta^2 \rceil$ suffices for completion with probability one because $v_U \le R^2/4$. The total balances the tighter variance bound against the pilot acquisitions.

L.3 Closed-form planning for the protected Gaussian route

In the shared-noise model, an a priori contrast design scale $d_0 > 0$ gives the planning objective

$$\min_{\lambda \in \Lambda} c_{\mathrm{world}} \max\left\{\frac{a(1+\lambda)\sigma_e^2}{d_0^2}, \frac{b[\sigma_c^2 + \sigma_e^2\{1 + (1 - 1/M)/\lambda\}]}{\eta^2}\right\}, \tag{L.7}$$

where a, b are confidence-bound factors and Λ is a fixed positive grid. The comparison term increases with λ while the reporting term decreases.

The allocation objective in Eq. (L.7) has a closed-form continuous solution. Write

$$A = \frac{a\sigma_e^2}{d_0^2}, \qquad B_0 = \frac{b(\sigma_c^2 + \sigma_e^2)}{\eta^2}, \qquad C_0 = \frac{b\sigma_e^2(1 - 1/M)}{\eta^2},$$

and assume $A, C_0 > 0$ and $B_0 \ge 0$. After removing the positive common cost factor, the objective is $f(\lambda) = \max\{A(1+\lambda), B_0 + C_0/\lambda\}$ for $\lambda > 0$. The first branch increases strictly and the second decreases strictly. They cross once, so the minimizer is the positive root of $A\lambda^2 + (A - B_0)\lambda - C_0 = 0$:

$$\lambda_* = \frac{B_0 - A + \sqrt{(A - B_0)^2 + 4AC_0}}{2A}. \tag{L.8}$$

For a fixed finite positive grid Λ, evaluate f at the nearest grid point below and above λ_* and choose the smaller value; if λ_* lies outside the grid's range, use the nearest endpoint. This rule follows from the decrease of f before the crossing and its increase afterward, giving the optimum of the stated planning surrogate.

M Reporting Execution and Comparison Protocols

The experimental records link frozen pools and utilities to paired outcomes, correctness and certification indicators, and stage-resolved costs. Run configurations retain covariance specifications, confidence boundaries and error allocation, controller parameters, and execution versions. Budget-specific traces record queries, stopping decisions, accepted whole-block prefixes, and costs, including abstentions. Fresh reporting, post-output audit, and deployment gating use distinct banks (Algorithm 2). With known Gaussian marginal variance v_i, use $n_R = \lceil z^2_{1-\delta_{\mathrm{rep}}/2} v_{\hat{i}}/\eta^2 \rceil$; for bounded scores of range R, use $n_R = \lceil R^2 \log(2/\delta_{\mathrm{rep}})/(2\eta^2) \rceil$. The count and interval rule are frozen before opening reporting outcomes. A fresh report estimates the selected candidate under the same target world law used to define its utility. Its observations are independent of proposal and selection information. The reporting count and radius are fixed before the first reporting outcome is opened. Algorithm 2 specializes the reporting step of Algorithm 1; the pilot-assisted construction in Proposition 12 has an additional pilot error allocation.

The fixed-sample proof uses complete observations. Retries retain the same immutable world; a prespecified failure score becomes part of the utility; and whole-world deletion follows the score-independent missingness condition of Appendix C. These rules preserve the target observation law and the precommitted reporting count.

Algorithm 2 Precommitted fresh winner reporting

Require: Frozen selection transcript $\mathcal{S}$, winner $\hat{i}$, selection certificate C_{sel}, remaining budget, $\eta, \delta_{\mathrm{rep}}$, valid reporting model

From $\mathcal{S}$, fix count $n_R \geq 1$, radius $h_R \leq \eta$, and acquisition cost
Reserve any required post-decision audit cost
if $C_{\mathrm{sel}} = 0$ or the complete report and required reserve do not fit **then**
    Return joint abstention; flag any separately produced report as diagnostic
end if
Draw exactly n_R fresh winner observations under the target world law
Charge every attempt under the declared failure rule
if the prescribed complete sample cannot be obtained within budget **then**
    Return abstention without issuing a joint certificate
end if
Set $\widehat{\mu}_{\hat{i}} = n_R^{-1} \sum_{t=1}^{n_R} Y_t^{\mathrm{rep}}(\hat{i})$
Freeze $(\hat{i}, \widehat{\mu}_{\hat{i}}, [\widehat{\mu}_{\hat{i}} - h_R, \widehat{\mu}_{\hat{i}} + h_R], C_B = 1)$
Open the independent audit without changing the frozen joint output
Return the frozen joint output and the separate audit record

An optional deployment gate uses an additional bank $\mathcal{B}_{\mathrm{gate}}$ and its own error allocation. Its bank is disjoint from the held-out audit $\mathcal{B}_{\mathrm{audit}}$. Gate evidence determines deploy versus fallback; audit evidence is opened after deployment is fixed. Gate Type-I error concerns the true gain of a deployed candidate, whereas negative audit frequency concerns observed held-out outcomes. Reporting, gate, and audit therefore retain separate records and costs.

M.1 What each comparison isolates

Comparison protocols hold the target utility, tolerances, error allocations, and required completion or correct-output targets fixed. Their acquisition costs include unused acquired observations and the declared reporting and audit reserves.

Paired selection followed by a fresh winner report controls for the reporting route while preserving the benefit of paired differences. Ordinary independent splitting is equivalent when it uses the same selection observations, winner-only reporting reserve, boundaries, and stopping decisions. Coupling the random inputs then gives identical queries, outputs, and pathwise evaluation cost. Computational gains arise from a change in allocation, observation structure, or reporting design.

A full-vector-fission comparison holds the released contrast channel and stopping decisions fixed while changing the protected reporting channel. Equations (8)–(9) then isolate the reporting-variance reduction. Both methods pay for every acquired vector coordinate. When selection determines the stopping time, a narrower reporting interval improves precision at the same acquisition count. Covariance-aware best-arm identification followed by a fresh report provides another selection-and-reporting composition (Saad et al., 2023), whose allocation rule and confidence boundary determine its cost.

These comparison protocols isolate the contributions of allocation and reporting. Equation (H.7) defines their empirical cost-at-target endpoints. Each experiment retains its own method configuration, execution cohort, and acquisition ledger.

N Full Replay Selection Results

The selection study uses $\varepsilon = 0.20$ pp and $\delta_{\mathrm{sel}} = 0.05$ across four settings. EB is exact-best agreement with the finite reference; P_ε is reference-relative selection success. Mean regret and its 90th percentile, r_{90}, are in percentage points. Evals and GPU-h are per-replicate selection-stage averages; end-to-end computation additionally includes the other stages. The four settings use 60, 50, 40, and 40 outer replicates, respectively.

The uncertified rate U records return at the selection cap without a selection certificate. For PAIRS, the counts are $2/60$, $4/50$, $2/40$, and $5/40$, totaling $13/190$. All returned incumbents enter P_ε, EB,

and regret, including these cap outputs. Selection certification, reference-relative correctness, and joint-reporting completion are separate endpoints. The U column applies to the two anytime-valid racing methods; the other methods are marked with a dash. Matched-uniform costs and finite-reference scoring follow Section 6.1.

N.1 SPLIT-CIFAR SELECTION

Table 27: Split-CIFAR-10: selection results over 60 outer replicates and 16 frozen candidates. EB, P_ε, and U are percentages; regret is in pp.

Method	EB	P_ε	Mean r	r_{90}	Evals	GPU-h	U
Single realization	45.0	66.7	.27	.72	16	2.9	–
Fixed $k = 2$, paired	68.3	85.0	.13	.32	32	5.8	–
Fixed $k = 3$, paired	81.7	93.3	.07	.18	48	8.6	–
Fixed $k = 5$, paired	91.7	98.3	.03	.08	80	14.4	–
Uniform, matched budget	71.7	86.7	.11	.27	31	5.6	–
Successive halving	73.3	88.3	.10	.24	30	5.4	–
Unpaired AV racing	81.7	93.3	.07	.18	42	7.6	5.0
PAIRS	85.0	96.7	.05	.12	31	5.6	3.3

Table 28: Split-CIFAR-100: selection results over 50 outer replicates and 24 frozen candidates. EB, P_ε, and U are percentages; regret is in pp.

Method	EB	P_ε	Mean r	r_{90}	Evals	GPU-h	U
Single realization	28.0	52.0	.45	1.12	24	15.6	–
Fixed $k = 2$, paired	50.0	76.0	.23	.58	48	31.2	–
Fixed $k = 3$, paired	66.0	88.0	.12	.30	72	46.8	–
Fixed $k = 5$, paired	80.0	94.0	.07	.18	120	78.0	–
Uniform, matched budget	54.0	80.0	.19	.47	46	29.9	–
Successive halving	56.0	82.0	.18	.43	45	29.3	–
Unpaired AV racing	66.0	88.0	.12	.29	65	42.3	10.0
PAIRS	70.0	92.0	.09	.22	46	29.9	8.0

N.2 STREAM AND SCHEDULE SELECTION

Table 29: CORe50: selection results over 40 outer replicates and 12 frozen candidates. EB, P_ε, and U are percentages; regret is in pp.

Method	EB	P_ε	Mean r	r_{90}	Evals	GPU-h	U
Single realization	40.0	62.5	.32	.80	12	10.8	–
Fixed $k = 2$, paired	62.5	82.5	.16	.40	24	21.6	–
Fixed $k = 3$, paired	75.0	92.5	.09	.23	36	32.4	–
Fixed $k = 5$, paired	87.5	97.5	.04	.10	60	54.0	–
Uniform, matched budget	65.0	85.0	.14	.35	24	21.6	–
Successive halving	67.5	87.5	.13	.31	23	20.7	–
Unpaired AV racing	75.0	92.5	.09	.23	34	30.6	7.5
PAIRS	80.0	95.0	.07	.17	24	21.6	5.0

Table 30: Frozen RS-MCTS pool: selection results over 40 outer replicates and 32 frozen candidates. EB, P_ε, and U are percentages; regret is in pp.

Method	EB	P_ε	Mean r	r_{90}	Evals	GPU-h	U
Single realization	17.5	37.5	.62	1.48	32	11.2	–
Fixed $k = 2$, paired	35.0	62.5	.33	.82	64	22.4	–
Fixed $k = 3$, paired	55.0	82.5	.18	.46	96	33.6	–
Fixed $k = 5$, paired	67.5	92.5	.10	.26	160	56.0	–
Uniform, matched budget	42.5	70.0	.27	.67	62	21.7	–
Successive halving	45.0	72.5	.25	.62	60	21.0	–
Unpaired AV racing	55.0	82.5	.18	.45	90	31.5	15.0
PAIRS	60.0	87.5	.13	.33	62	21.7	12.5

N.3 PAIRED SUCCESSES AND IDENTIFIABLE UNCERTAINTY

Reference-relative correctness and selection certification are separate run-level events. Let their marginal counts in N runs be K_{correct} and K_{cert}. The number satisfying both obeys the sharp count bounds

$$\max\{0, K_{\text{correct}} + K_{\text{cert}} - N\} \le K_{\text{correct}\cap\text{cert}} \le \min\{K_{\text{correct}}, K_{\text{cert}}\}. \tag{N.1}$$

The lower bound follows by counting the union and the upper bound by inclusion in each event. Both endpoints are attainable by arranging the overlap of the two sets. Table 31 applies this identity to the setting-level counts. The overlap can vary even when the two marginal counts agree.

Table 31: Reference-relative selection success and certificate counts. The final column contains deterministic identification bounds conditional on the observed margins.

Setting	Runs	PAIRS correct	Uniform correct	PAIRS certified	Correct ∩ certified
CIFAR-10	60	58	52	58	56–58
CIFAR-100	50	46	40	46	42–46
CORe50	40	38	34	38	36–38
RS-MCTS	40	35	28	35	30–35
Total	190	177	154	177	164–177

The overlap is therefore between $164/190 = 86.32\%$ and $177/190 = 93.16\%$ of all runs. Among the 177 selection-certified outputs, reference-relative correctness lies between $164/177 = 92.66\%$ and 100% at this level of aggregation. The bounds range over all correctness–certification overlaps consistent with the margins. Finite-reference uncertainty and reporting error are separate components of the joint analysis.

The reference-relative success counts for PAIRS versus matched uniform are $58/60$ versus $52/60$, $46/50$ versus $40/50$, $38/40$ versus $34/40$, and $35/40$ versus $28/40$. They sum to $177/190$ versus $154/190$, a pooled difference of 12.1053 pp. Their equal-setting mean difference is 12.375 pp. Paired uncertainty also depends on the discordances between the two methods.

For n paired runs with success counts x and y, let d be the number of discordant pairs and $\widehat{\Delta} = (x - y)/n$. The possible values satisfy

$$|x - y| \le d \le \min\{x + y, 2n - x - y\}, \qquad d \equiv x - y \pmod 2. \tag{N.2}$$

The plug-in standard error of the paired difference is $\{(d/n - \widehat{\Delta}^2)/(n-1)\}^{1/2}$. The observed CIFAR-10 comparisons against fixed $k = 3$ and unpaired AV have different discordance counts despite identical comparator margins (Table 32). Both differences are $2/60 = 3.3333$ pp, but their paired standard errors differ.

Table 32: Observed CIFAR-10 paired selection outcomes over the same 60 outer runs. The four cells count both correct, PAIRS alone correct, comparator alone correct, and neither correct. Intervals are pointwise 95% paired Wald intervals in pp with reference scoring fixed.

Comparator	n_{11}	n_{10}	n_{01}	n_{00}	Difference	Interval
Fixed $k = 3$	56	2	0	2	3.3333	$[-1.2470, 7.9137]$
Unpaired AV	54	4	2	0	3.3333	$[-4.6908, 11.3574]$

Table 33 shows the Wald intervals obtained at the two endpoints of the feasible discordance range. The observed discordance count selects an interval within this family. The Wald approximation is sensitive to small discordance counts. Joint resampling in Appendix H retains outer IDs, method outcomes, and reference-bank links.

For CIFAR-10, the full compatible set of $(n_{11}, n_{10}, n_{01}, n_{00})$ is $(50, 8, 2, 0)$, $(51, 7, 1, 1)$, or $(52, 6, 0, 2)$. Across the four settings, the discordance sets are respectively $\{6, 8, 10\}$, $\{6, 8, 10, 12, 14\}$, $\{4, 6, 8\}$, and $\{7, 9, 11, 13, 15, 17\}$. Their Cartesian product contains $3 \cdot 5 \cdot 3 \cdot 6 = 270$ stratified contingency-table structures. Enumerating them gives the complete sensitivity analysis over overlaps compatible with the marginal counts.

Table 33: Paired uncertainty compatible with the marginal selection counts. Differences and interval endpoints are in pp. The intervals use $z_{0.975}$ and the minimum or maximum possible discordance, holding the reference scores fixed.

Setting	PAIRS / uniform	Difference	d range	Min-d interval	Max-d interval
CIFAR-10	$58/52$	10.0	6–10	$[2.35, 17.65]$	$[-.10, 20.10]$
CIFAR-100	$46/40$	12.0	6–14	$[2.90, 21.10]$	$[-2.43, 26.43]$
CORe50	$38/34$	10.0	4–8	$[.58, 19.42]$	$[-3.68, 23.68]$
RS-MCTS	$35/28$	17.5	7–17	$[5.57, 29.43]$	$[-2.21, 37.21]$

Table 34: Pooled paired uncertainty after incorporating the observed CIFAR-10 comparisons. Uniform ranges over all 270 margin-compatible structures; fixed $k = 3$ and unpaired AV range over 90 structures each with their CIFAR-10 discordances fixed. Differences and 95% Wald intervals are in pp; reference uncertainty is held fixed.

Comparator	Structures	Difference	Narrowest interval	Widest interval
Uniform	270	12.1053	$[7.4353, 16.7752]$	$[5.0269, 19.1837]$
Fixed $k = 3$	90	3.6842	$[0.9796, 6.3888]$	$[-1.9087, 9.2772]$
Unpaired AV	90	3.6842	$[0.2720, 7.0964]$	$[-2.2832, 9.6516]$

Holding the four setting sample sizes fixed, let $w_s = N_s/190$ and $\widehat{\Delta}_s = (x_s - y_s)/N_s$. For each comparator, the stratified plug-in variance under independent outer replicates and settings is

$$\widehat{\Delta} = \sum_s w_s \widehat{\Delta}_s, \qquad \widehat{V} = \sum_s w_s^2 \frac{d_s/N_s - \widehat{\Delta}_s^2}{N_s - 1}. \tag{N.3}$$

The pooled point differences are $23/190 = 12.1053$ pp against uniform and $7/190 = 3.6842$ pp against each of fixed $k = 3$ and unpaired AV. Their equal-setting macro differences are 12.3750 and 3.7083 pp, respectively. Table 34 gives the narrowest and widest $\widehat{\Delta} \pm z_{0.975}\sqrt{\widehat{V}}$ intervals compatible with these counts. For fixed $k = 3$ and unpaired AV, the CIFAR-10 cells in Table 32 fix $d_1 = 2$ and $d_1 = 6$, leaving $5 \cdot 3 \cdot 6 = 90$ possible structures over the remaining settings for each comparison. Uniform retains 270 structures. These are conditional normal-approximation sensitivity intervals with reference scores fixed. The actual CIFAR-10 information narrows the fixed-$k = 3$ envelope and distinguishes it from unpaired AV.

N.4 Outcome-independent uniform allocation

The fixed-budget allocation rule writes $B_u = Mq + r$, where $q = \lfloor B_u/M \rfloor$ and $0 \le r < M$. Each candidate receives q observations; the first r candidates in an outcome-independent random permutation receive one extra observation. For $(M, B_u) = (16, 31), (24, 46), (12, 24), (32, 62)$, the respective allocations are 15 candidates twice and one once; 22 twice and two once; all 12 twice; and 30 twice and two once. Table 35 gives the fixed per-run budgets and permutation seeds. The budgets were frozen on 2026-09-19 at 10:00:00 UTC for all four settings. Remainder allocation uses the permutation independently of outcomes. The preassigned B_u remains fixed throughout the evaluated race. Matching therefore concerns the selection budget, with reporting, audit, and reference charged as separate stages.

O Quantitative Analysis of Reporting and Computation

O.1 Population bias and finite-outer absolute bias

Conditional unbiasedness describes the expected signed reporting error. Its finite-outer absolute estimate has a different scale. To quantify the distinction, consider equal-mean Gaussian candidates with marginal world standard deviation $\sigma = 2$ pp, equicorrelation $\rho = 0.67$, $n_S = 10$ selection worlds per candidate, and $n_R = 20$ fresh winner observations. The selected candidate is optimal regardless of the selection outcome. If $\kappa_M = \mathbb{E} \max_{i \le M} Z_i$ for independent standard normals,

Table 35: Fixed uniform-allocation configuration. Each candidate receives q observations and the first r candidates in an outcome-independent permutation receive one more. The budget column gives the fixed per-run allocation.

Setting	M	B_u	q	r	Permutation seed
CIFAR-10	16	31	1	15	2026091901
CIFAR-100	24	46	1	22	2026091902
CORe50	12	24	2	0	2026091903
RS-MCTS	32	62	1	30	2026091904

Appendix A gives

$$b_N = \frac{\sigma\sqrt{1-\rho}}{\sqrt{n_S}}\kappa_M, \qquad b_P = 0, \qquad s_R = \frac{\sigma}{\sqrt{n_R}}. \tag{O.1}$$

Here b_N and b_P are population signed biases. For N_s independent outer reports scored against the population mean, $\widehat{b}_s^P \sim N(0, s_R^2/N_s)$. Integrating the folded-normal density yields

$$\mathbb{E}|\widehat{b}_s^P| = \sqrt{\frac{2}{\pi}}\frac{s_R}{\sqrt{N_s}}. \tag{O.2}$$

The settings in Table 36 vary the candidate and outer counts while retaining this Gaussian observation law. All four have fresh-report $\mathrm{RMSE} = s_R = 0.44721$ pp and $\mathrm{MAE} = \sqrt{2/\pi}s_R = 0.35682$ pp, even though their population reporting bias is zero.

Table 36: Finite-outer absolute-bias scales under Eq. (O.1). Entries are population quantities and expected absolute estimates in pp; M and N_s specify the Gaussian setting.

M	N_s	Naive bias b_N	Fresh bias b_P	$\mathbb{E}\lvert\widehat{b}_s^P\rvert$
16	60	0.64162	0.00000	0.04607
24	50	0.70763	0.00000	0.05046
12	40	0.59193	0.00000	0.05642
32	40	0.75195	0.00000	0.05642
Equal-setting mean		0.67328	0.00000	0.05234

The ratio $1 - 0.0523416/0.6732794 = 92.2259\%$ compares the expected absolute fresh-bias estimate with the population naive bias. It explains how a nonzero finite-sample absolute bias can coexist with a zero-bias reporting mechanism. The expectation of an empirical ratio also depends on its random denominator and is a different quantity. With a shared finite reference, Eq. (O.2) acquires an additional reference-error component, as studied in Appendix K.

O.2 Gaussian reporting precision and selection dependence

The population counterparts of Experiment A follow from the maximum distribution and Eq. (G.5). For $M = 20$, $n = 50$, $\sigma = 0.1$, and $\lambda = 1$, Table 37 gives the mean optimism, ordinary winner-interval coverage, structured RMSE, and structured-versus-full interval-width reduction. These quantities use the Gaussian model's native score scale and complement the experimental estimates in Table 7.

Table 37: Population reporting quantities for the Gaussian model of Experiment A. Coverage and interval-width reduction are percentages; optimism and RMSE use native score units.

ρ	Naive optimism	Naive coverage	Structured RMSE	Width reduction
0.00	0.026410	60.269	0.019748	1.258
0.50	0.018675	78.962	0.017176	14.122
0.67	0.015171	84.530	0.016208	18.960
0.90	0.008352	91.839	0.014799	26.007

For $0 < \rho < 1$, the naive coverage column is obtained by conditioning on $Z_{(M)} = z$. Let $z_c = z_{0.975}$ and $f_M(z) = M\phi(z)\Phi(z)^{M-1}$. The coverage is

$$\int_{-\infty}^{\infty} f_M(z) \left[\Phi\left(\frac{z_c - \sqrt{1-\rho}\, z}{\sqrt{\rho}} \right) - \Phi\left(\frac{-z_c - \sqrt{1-\rho}\, z}{\sqrt{\rho}} \right) \right] dz. \tag{O.3}$$

At $\rho = 0$, it reduces to $.975^M - .025^M$. Under this Gaussian model, structured reporting has population bias zero and interval coverage 95%. Joint success additionally requires selection correctness.

For Experiment B's uniquely better candidate, set $d = (0.72 - 0.70)/[\sigma\sqrt{(1-\rho)(1+\lambda)/n}]$. Selection success is $\int \phi(z)\Phi(z+d)^{M-1}dz$. At $\lambda = 0.25, 1, 4$, its values are 62.706%, 46.666%, and 25.663%, while structured RMSE is 0.021232, 0.016208, and 0.014686. Increasing protection therefore improves absolute reporting precision while reducing the probability of identifying the better candidate at this fixed bank size.

O.3 From difference variance to total cost

For a paired comparison with marginal standard deviations σ_i, σ_j and correlation ρ_{ij},

$$v_{ij} = \sigma_i^2 + \sigma_j^2 - 2\rho_{ij}\sigma_i\sigma_j. \tag{O.4}$$

Equal marginal variances give a paired-to-independent difference-variance ratio $1 - \rho$. Under unequal variances, the ratio also depends on their relative scales. At $\rho = 0.8$, $(\sigma_i, \sigma_j) = (1,1), (1,3), (1,10)$ yield independent-to-paired variance ratios 5.00, 1.92, and 1.19. Thus variance balance determines how strongly positive correlation improves comparison precision. For a single pair with $\rho = -0.2$, pairing increases difference variance by 20%.

Now let a comparator's cost be $C_S + C_F$, where C_S is selection and C_F contains equally charged reporting, audit, and other unchanged stages. If PAIRS changes selection cost to sC_S, define $f = C_F/(C_S + C_F)$. Then

$$R_{\text{total}} = \frac{C_S + C_F}{sC_S + C_F} = \frac{1}{f + s(1-f)}. \tag{O.5}$$

When selection cost is variance-dominated and the remaining factors agree, a variance ratio can serve as s. Table 38 shows how unchanged stages attenuate the selection-stage gain.

Table 38: Total cost ratios from Eq. (O.5). The two columns assume the indicated selection-cost multiplier s, with all other stage costs equal between procedures.

Unchanged fraction f	$s = 0.33$	$s = 0.61$
0.10	2.519	1.541
0.30	1.883	1.376
0.50	1.504	1.242
0.80	1.155	1.085

With $s = 0.33$, a total ratio of at least 2.3 requires $f \leq 0.1564$. With $s = 0.61$, this same-stage model permits at most $1/0.61 = 1.639$ even at $f = 0$. Thus the observed difference-variance ratio, a controlled selection-cost ratio, and the application-level total-cost ratio describe different comparisons. A comparator-to-PAIRS cost ratio of 2.3 means PAIRS uses 43.48% of the comparator cost, a 56.52% reduction.

O.4 Charging the variance pilot

The pilot in Proposition 12 trades its own acquisition cost against the remaining report size. For $R = 100$, $\eta = 1$ pp, $\delta_P = 0.005$, and $\delta_{\text{rep}} = 0.025$, condition on the pilot summary $\bar{Z} = 4$ squared pp. Equations (L.5)–(L.6) give the counts in Table 39. The comparison uses the 20,999-observation Hoeffding report at total reporting error budget $0.03 = \delta_P + \delta_{\text{rep}}$.

A small pilot leaves a broad variance bound, whereas a large pilot can consume its precision benefit. The 4.65 ratio at $m = 500$ describes the report subproblem conditional on the stated pilot summary. The unconditional expected total requires averaging the pilot-dependent count over the pilot distribution.

Table 39: Pilot-dependent reporting acquisitions given $\bar{Z} = 4$ squared pp. The cost ratio is $20{,}999/(2m + n_R^P)$; selection and audit costs are outside this reporting subproblem.

Pilot pairs m	Pilot observations	v_U	Fresh report	Total	Cost ratio
10	20	2,500.00	22,203	22,223	0.94
100	200	817.81	7,460	7,660	2.74
500	1,000	367.95	3,517	4,517	4.65
1,000	2,000	261.35	2,583	4,583	4.58
10,000	20,000	85.38	1,041	21,041	1.00

An exactly Gaussian observation law with unknown variance $v > 0$ permits a different variance bound. After the winner is fixed, take $k \geq 2$ independent pilot observations and their unbiased sample variance s_P^2. Since $(k-1)s_P^2/v \sim \chi^2_{k-1}$, define

$$v_U^G = \frac{(k-1)s_P^2}{\chi^2_{\delta_P, k-1}}, \qquad n_R^G = \left\lceil \frac{z^2_{1-\delta_{\mathrm{rep}}/2} v_U^G}{\eta^2} \right\rceil. \tag{O.6}$$

The variance bound fails with probability δ_P. On its complementary event, a fresh independent mean based on the frozen count n_R^G has error beyond η with probability at most δ_{rep}. Its conditional bias given selection and pilot is zero on every transcript. For $k = 30$, $s_P^2 = 4$ squared pp, $\delta_P = 0.005$, $\delta_{\mathrm{rep}} = 0.025$, and $\eta = 1$ pp, the variance bound is 8.84069 squared pp and the fresh count is 45, costing 75 observations including the pilot. This construction uses exact Gaussian sampling rather than boundedness alone.

P Stage-Resolved Physical Compute and Cost Scope

An evaluation count measures acquired candidate–world observations; accelerator-hours measure device allocation; makespan also depends on concurrency and orchestration. This experiment records all three quantities by stage. Its execution cohort is separate from the four-setting selection study in Appendix N.

P.1 Attempt-level reconciliation

Table 40 separates accepted, completed-unused, and failed attempts according to Eq. (H.6). For PAIRS selection, $576 + 4 + 20 = 600$; for comparator selection, $1870 + 5 + 25 = 1900$. The 18 and 22 successful selection retries are already included in accepted observations. Reporting and audit are acquired separately by each method. Proposal and reference are shared study costs; the pilot and deployment gate are not used in this record.

Table 40: Stage-resolved accounting record. Unused means completed but not accepted for inference. Retry is a subset of accepted observations, not an additional charge. Reporting and audit rows apply to each method.

Method / stage	Accepted	Unused	Failed	Attempts	Retry	GPU-h
PAIRS / selection	576	4	20	600	18	120.0
Comparator / selection	1,870	5	25	1,900	22	342.0
Each / reporting	300	0	0	300	0	60.0
Each / audit	100	0	0	100	0	20.0
Shared proposal	160	0	4	164	4	32.8
Shared reference	1,024	0	8	1,032	8	206.4
Pilot / gate	0	0	0	0	0	0.0

Reporting precision depends on both sample count and observation law. With 300 iid observations in $[0, 100]$, the 97% Hoeffding half-width is

$$h_H = 100\sqrt{\frac{\log(2/0.03)}{2 \cdot 300}} = 8.3663 \text{ pp}. \tag{P.1}$$

For comparison, under a strictly Gaussian model with known marginal standard deviation 2 pp, the 97% half-width for the same count is

$$h_G = z_{0.985}\frac{2}{\sqrt{300}} = 0.25058 \text{ pp.} \tag{P.2}$$

At $\eta = 1$ pp, the bounded-score interval is too wide, whereas the known-Gaussian interval meets the precision target. The two calculations quantify the role of the observation assumption at a fixed reporting cost.

P.2 Online and all-in ratios

Counting selection, reporting, and audit gives 1,000 attempts and 200 GPU-h for PAIRS, versus 2,300 attempts and 422 GPU-h for the comparator. The attempt ratio is 2.300, but the accelerator-hour ratio is 2.110. For an all-in comparison, split the shared proposal and reference equally between the two methods. Each method is then charged 598 additional attempts and 119.6 additional GPU-h. Table 41 reports the resulting totals. The equal split is an explicit amortization rule; a different number of users of the shared banks changes it.

Table 41: Cost ratios for the same stage-resolved record. Ratios are comparator / PAIRS. All-in totals split shared proposal and reference equally. Wall-clock excludes those shared stages and uses the recorded schedule described below.

Accounting scope / unit	PAIRS	Comparator	Ratio
Selection + reporting + audit / attempts	1,000	2,300	2.300
Selection + reporting + audit / GPU-h	200.0	422.0	2.110
All-in / attempts	1,598	2,898	1.814
All-in / GPU-h	319.6	541.6	1.695
Method execution / wall-clock h	82.0	137.5	1.677

The all-in attempts sum to 4496 across both methods, and all-in GPU-h sum to 861.2, charging each shared acquisition once. These ratios compare expenditure within each accounting scope. Cost at matched reliability, defined in Eq. (H.7), additionally conditions the comparison on completion and joint-success requirements.

P.3 Concurrency and makespan

The recorded execution schedule uses four GPUs for selection, two for reporting, and one for audit, with these stages executed in sequence within a method. Including two hours of orchestration gives

$$\begin{aligned} T_P &= 120/4 + 60/2 + 20 + 2 = 82 \text{ h}, \\ T_C &= 342/4 + 60/2 + 20 + 2 = 137.5 \text{ h}. \end{aligned} \tag{P.3}$$

With sufficient devices for the methods to execute concurrently, a shared proposal stage taking 8.2 h followed by the two methods and then a 25.8 h reference stage gives study makespan

$$T_{\text{study}} = 8.2 + \max\{82, 137.5\} + 25.8 = 171.5 \text{ h}. \tag{P.4}$$

The critical path determines study makespan. Equation (P.3) combines the stage-specific device allocations and orchestration time.

P.4 Relation to the four-setting selection ledger

Multiplying the displayed mean PAIRS selection GPU-hours by each setting's outer count gives approximately 336, 1495, 864, and 868 GPU-h, respectively, or 3563 GPU-h in total. The analogous displayed evaluation means give 7600 accepted selection evaluations. This aggregation inherits the rounding of the setting means. The resulting totals cover selection; proposal, reporting, audit, reference, and failed attempts belong to their respective stage accounts.

P.5 HOMOGENEOUS-COST EXECUTION

The homogeneous-cost experiment uses 1,000 and 2,300 evaluations at 0.2 GPU-h per evaluation, with four GPUs fully occupied, no failures, and no orchestration overhead in the timed execution. Its accelerator totals are 200 and 460 GPU-h and its makespans are 50 and 115 h, giving a ratio of 2.3 in both units. With 20 h of serial overhead charged to each method, the corresponding end-to-end makespans are 70 and 135 h and their ratio is 1.928571. The homogeneous execution and the 200/422 GPU-h experiment have different per-evaluation selection costs and scheduling conditions. Each cost comparison uses its own execution cohort; cross-cohort analysis matches outputs and stage costs by study and run identifiers.

Q CANDIDATE SPECIFICATIONS, UTILITY, AND EVIDENCE ACCESS

The configuration space contains four candidate families with 16, 24, 12, and 32 members. Table 42 lists their varying coordinates. Each candidate combines these coordinates with the fixed training settings to define an executable configuration.

Table 42: Candidate-family coordinates. Memory denotes replay-buffer capacity; draw and batch sizes count replay examples. The last family is indexed by the schedule parameters in Equation (Q.2).

Setting	Varying coordinates	M
CIFAR-10	Learner $\{\mathrm{ER}, \mathrm{DER{+}{+}}\}$; memory $\{200, 500\}$; replay draw $\{16, 32\}$; learning rate $\{0.03, 0.10\}$.	16
CIFAR-100	Learner ER or ASER$_\mu$-PyTorch; memory $\{1000, 2000, 5000\}$; replay batch $\{10, 20\}$; learning rate $\{0.05, 0.10\}$.	24
CORe50	Latent layer `conv5_3/dw`, `conv5_4/dw`, or `pool6`; memory $\{500, 1000, 1500, 3000\}$.	12
Replay schedule	Eight exponents a and four uniform-mixing coefficients λ; each pair specifies a 20×20 schedule matrix.	32

CIFAR-10 fixes the current batch size at 32 and uses 50 epochs. DER++ distinguishes the stored-logit and replay-label retrievals and their loss coefficients. CIFAR-100 uses an online single pass with current batch size 10; the ASER branch uses `asvm`, $k = 3$, and `n_smp_cls`$= 1.5$. CORe50 uses AR1F with MobileNetV1 under NICv2-391. These fixed settings and the varying coordinates jointly identify the families.

The candidate identifiers use the coordinate ordering in Table 43. Let $i_1, \ldots, i_d$ be zero-based indices into the listed coordinate values, with the last coordinate varying fastest. The one-based candidate number is

$$k = 1 + \sum_{h=1}^{d} i_h \prod_{u=h+1}^{d} m_u, \tag{Q.1}$$

where m_u is the size of coordinate u and an empty product is one. Appending the two-digit number k to the setting prefix enumerates all $16 + 24 + 12 + 32 = 84$ configurations. For example, `C10-07` denotes ER with memory 500, replay draw 32, and learning rate 0.03; `C100-12` denotes ER with memory 5000, replay batch 20, and learning rate 0.10; and `CORE-03` denotes memory 500 at `pool6`. The identifier fixes the varying coordinates, while the training configuration includes the fixed settings above.

The schedule family changes how replay is distributed over previously seen tasks. Its parameters are

$$a \in \{-3, -2, -1, -0.5, 0.5, 1, 2, 3\}, \qquad \lambda \in \{0, 0.25, 0.50, 0.75\}.$$

For training task $t \geq 2$, the replay probability assigned to task $j < t$ is

$$p_{t,j}(a, \lambda) = (1 - \lambda) \frac{j^a}{\sum_{k=1}^{t-1} k^a} + \frac{\lambda}{t - 1}. \tag{Q.2}$$

The first row has no replay; entries with $j \geq t$ are zero. Every later row sums to one. Positive exponents favor recent tasks, negative exponents favor earlier tasks, and increasing λ moves the al-

Table 43: Candidate indexing in the 84-row family manifest. Coordinates run from slowest to fastest; values follow the ascending numerical order in Table 42. Learner and layer orders are specified explicitly.

IDs	Coordinate sizes	Ordered coordinates
`C10-01--16`	$(2, 2, 2, 2)$	Learner (ER, DER++), memory, draw, learning rate
`C100-01--24`	$(2, 3, 2, 2)$	Learner (ER, ASER), memory, batch, learning rate
`CORE-01--12`	$(4, 3)$	Memory, layer (`conv5_3/dw`, `conv5_4/dw`, `pool6`)
`RS-01--32`	$(8, 4)$	Exponent a, uniform mixing λ

Table 44: Task-three replay probabilities $p_{3,1}$ computed from the 32-member parametric schedule family. These values describe the probability matrices before any integer replay-count allocation.

a	$\lambda = 0$	$\lambda = 0.25$	$\lambda = 0.50$	$\lambda = 0.75$
-3.0	0.888889	0.791667	0.694444	0.597222
-2.0	0.800000	0.725000	0.650000	0.575000
-1.0	0.666667	0.625000	0.583333	0.541667
-0.5	0.585786	0.564340	0.542893	0.521447
0.5	0.414214	0.435660	0.457107	0.478553
1.0	0.333333	0.375000	0.416667	0.458333
2.0	0.200000	0.275000	0.350000	0.425000
3.0	0.111111	0.208333	0.305556	0.402778

location toward uniform replay. Parameter pairs index this family. An MCTS-exported pool instead uses its exported schedules and candidate-to-run mapping.

The task-three allocation makes the geometry of this parametric family explicit. It satisfies

$$p_{3,1} = \frac{1-\lambda}{1+2^a} + \frac{\lambda}{2}, \qquad p_{3,1} - \frac{1}{2} = (1-\lambda)\left(\frac{1}{1+2^a} - \frac{1}{2}\right). \tag{Q.3}$$

Because the grid has $a \neq 0$ and $\lambda < 1$, every allocation is strictly between zero and one and differs from $1/2$. Table 44 lists all 32 values; the second task's coordinate is $1 - p_{3,1}$. In this indexing, `RS-17` has $a = 0.5$ and $\lambda = 0$, hence its task-three replay probabilities are $(0.414214, 0.585786, 0)$ to six decimals. For a replay draw of 20, the task-three integer allocation is $(8, 12, 0)$.

For a task-based record H containing final accuracies $a_{T,t}$ in percentage points, the primary utility is the final task average

$$Y(q) = g(H(q)) = \frac{1}{T}\sum_{t=1}^{T} a_{T,t}(q), \qquad g(H) \in [0, 100]. \tag{Q.4}$$

This endpoint is higher-is-better and weights tasks equally. An $\varepsilon = 0.20$ pp selection tolerance therefore concerns a difference in final average accuracy. Training-curve averages define a separate utility. On a fraction-scale accuracy record, the corresponding tolerances are $\varepsilon = 0.002$ and $\eta = 0.01$; conversion precedes inference and reference scoring.

Table 45: Evidence access under fresh winner reporting. Each bank has a distinct role; audit and reference outcomes do not feed back into selection or deployment.

Bank	Permitted use
Proposal	Generate candidates, freeze the pool, and perform independent budget planning.
Selection	Choose admissible queries, eliminate candidates, select the winner, and stop.
Pilot, if used	Determine the report count through the stated valid pilot rule.
Reporting	Estimate the frozen winner's performance and form its reporting interval.
Gate, if used	Apply the precommitted deployment or fallback decision.
Audit	Evaluate outcomes after the selection, report, and deployment decisions are frozen.
Reference	Score bias, regret, and joint correctness with finite-reference uncertainty.

Bank manifests link observations to setting, outer replicate, world, and candidate. The target world law specifies whether the data split is fixed or randomized alongside training; resampling test images from one checkpoint estimates a different target from rerunning the configuration in a fresh world. The complete run configuration additionally records the optimizer, learning-rate and normalization schedules, preprocessing, tie rules, loss coefficients, and candidate-specific random-number namespace.

Each bank record specifies its parent study, stage role, target world law, generation procedure, world IDs, candidate order, and sample count. It separately identifies shared randomness and candidate-specific randomness, the reporting route, the information-release rule, and the observed access history. Reference records also specify the uncertainty model used in scoring. Executable candidate identity comprises the configuration, utility implementation, code revision, and checkpoint or schedule artifact.

Fresh reporting relies on independent world banks. Protected Gaussian reporting instead releases F and seals G from the same complete acquisition in Eq. (7). The two roles may therefore carry the same world ID while their information channels remain independent under the stated Gaussian model. For that route, the relevant access record identifies the released channel, the covariance specification, and whether raw observations or protected values entered selection. The shared acquisition incurs one physical charge.

Q.1 Native actions and executable schedule identity

The native replay-scheduling action space uses $t-1$ interchangeable slots at training task t, each assigned to one historical task. Normalized slot counts determine the replay proportions (Klasson et al., 2023). Thus, for $t \geq 2$, its probability actions can be written as

$$\mathcal{A}_t^{\mathrm{native}} = \left\{ p \in \mathbb{R}_{\geq 0}^{t-1} : \sum_{j<t} p_j = 1,\ (t-1)p_j \in \mathbb{Z}_{\geq 0} \right\}. \tag{Q.5}$$

In particular, $\mathcal{A}_3^{\mathrm{native}} = \{(1,0),(1/2,1/2),(0,1)\}$. Let $\mathcal{Q}_{\mathrm{native}}$ contain schedule probability matrices whose rows belong to these native action sets, and let $\mathcal{Q}_{\mathrm{param}}$ be the 32-member family in Eq. (Q.2). For a common task horizon $T \geq 3$,

$$\mathcal{Q}_{\mathrm{param}} \cap \mathcal{Q}_{\mathrm{native}} = \varnothing. \tag{Q.6}$$

Indeed, Eq. (Q.3) gives $0 < p_{3,1} < 1$ and $p_{3,1} \neq 1/2$ for every grid point. The third-task row therefore excludes membership in the native family. Parametric schedules therefore use an extended probability action space.

Integer replay counts form a separate representation. The native method rounds task allocations while preserving total memory capacity (Klasson et al., 2023). For memory capacity m, let $\mathcal{I}_m(P)$ denote the executed integer allocation matrix produced by a specified rounding or projection rule. Distinct probability matrices can map to the same integer schedule, so equivalence at execution is determined by $\mathcal{I}_m(P)$. An executable pool records the proposed matrix, any action-space extension or projection, the actual integer allocation, and the associated candidate and run IDs. If candidates are deduplicated after a transformation, the relevant pool size is the number of distinct resulting executable configurations.

Q.2 Complete probability matrices and coordinate correspondence

The schedule archive contains all 32 probability matrices for $T = 20$, indexed by the eight exponents and four uniform-mixing coefficients in Eq. (Q.2). The machine-readable supplement stores each matrix separately, preserving all $32 \times 20 \times 20 = 12{,}800$ probability entries. Its matrix index follows the `RS-01--32` coordinate order in Table 43. The matrices are distinct, nonnegative, and zero on and above the diagonal. The first row is zero; each later row sums to one to the stored decimal precision. Across all entries, the maximum absolute difference from Eq. (Q.2) is below 5.0×10^{-13}, and the maximum deviation of a later row sum from one is 1.5×10^{-12}. These residuals quantify the numerical precision of the stored matrices.

Table 46: Early rows of full-horizon probability matrix 17 ($a = 0.5$, $\lambda = 0$). Values are displayed to nine decimals; coordinates $j \geq 5$ are zero in these rows. The machine-readable matrix contains the complete 20×20 record at its stored precision.

Task t	$p_{t,1}$	$p_{t,2}$	$p_{t,3}$	$p_{t,4}$
1	0.000000000	0.000000000	0.000000000	0.000000000
2	1.000000000	0.000000000	0.000000000	0.000000000
3	0.414213562	0.585786438	0.000000000	0.000000000
4	0.241180955	0.341081377	0.417737668	0.000000000
5	0.162700453	0.230093188	0.281805452	0.325400907

Table 47: Integer replay allocation for `SCHED-P17`. Each row gives the numbers of replay examples drawn from the five tasks; the first task has no replay and each subsequent row sums to 20.

Current task	Task 1	Task 2	Task 3	Task 4	Task 5	Total
1	0	0	0	0	0	0
2	20	0	0	0	0	20
3	8	12	0	0	0	20
4	5	7	8	0	0	20
5	3	5	6	6	0	20

Table 46 shows the early rows of matrix 17, whose parameters are $a = 0.5$ and $\lambda = 0$. The full archive retains all 20 rows for every candidate. The probability matrices and the integer execution mapping $\mathcal{I}_m$ have the distinct roles defined in Appendix Q.1.

Coordinate tuples also provide a stable join between differently ordered candidate inventories. The additional CORe50 inventory varies memory fastest, whereas Table 43 varies layer fastest. With zero-based memory index $b \in \{0, 1, 2, 3\}$ and layer index $l \in \{0, 1, 2\}$, their respective one-based row numbers are

$$k_{\text{inventory}} = 1 + 4l + b, \qquad k_{\text{family}} = 1 + 3b + l.$$

Thus `AUDIT-CORE-03` and `CORE-07` share memory 1500 and layer `conv5_3/dw`, while `AUDIT-CORE-09` and `CORE-03` share memory 500 and layer `pool6`. The supplement preserves both labels and their coordinate correspondence; execution records use their own candidate and run keys.

Q.3 A FIVE-TASK SCHEDULE WITH AN INTEGER EXECUTION MAP

The parametric candidate `SCHED-P17` uses $a = 0.5$, $\lambda = 0$, a five-task horizon, and a replay draw of 20. Its probability rows are the five-task prefix in Table 46. The first task has no replay. For $t \geq 2$, the execution map applies largest-remainder allocation: take $b_{t,j} = \lfloor 20 p_{t,j} \rfloor$, then assign the remaining $20 - \sum_{j<t} b_{t,j}$ positions to the largest fractional remainders, breaking ties by smaller task ID. Table 47 gives the complete integer allocation read by the training procedure.

At task five, the unrounded allocation is $(3.254009, 4.601864, 5.636109, 6.508018)$. Flooring leaves two positions; the remainders allocate them to tasks three and two, yielding $(3, 5, 6, 6)$. The candidate identity comprises the probability parameters, horizon, draw size, and rounding rule. It belongs to the parametric schedule family; native MCTS actions use the distinct action set in Appendix Q.1.

R COMPLETION-AWARE GUARANTEES AND QUANTITATIVE DESIGN ANALYSIS

We derive completion and reporting guarantees and evaluate their numerical scales under explicit observation models, query sets, and acquisition costs.

R.1 Completion, conditional reporting, and cost information

Theorem 4 conditions on all information that determines whether a joint output is completed. At $\delta_{\mathrm{sel}} = 0.02$ and $\delta_{\mathrm{rep}} = 0.03$, Eq. (18) gives the following population envelope.

Table 48: Completion-aware bounds for $\delta_{\mathrm{sel}} = 0.02$, $\delta_{\mathrm{rep}} = 0.03$. The input Q is the population completion probability.

Q (%)	Lower bound on J (%)	Upper bound on W (%)
90	85.36	4.64
95	90.21	4.79
98	93.12	4.88
100	95.06	4.94

Cost information can violate the conditional reporting assumption even when the report value is not explicitly inspected. Let a single report observation be $Y \sim \mathrm{Bernoulli}(1/2)$, with acquisition cost one when $Y = 1$ and two when $Y = 0$. A budget-one rule that completes only affordable outcomes has $A = \{Y = 1\}$, $Q = 1/2$, and $\mathbb{E}[Y \mid A] = 1$. Its completion-conditional bias is $1/2$. Reporting protection therefore covers failures, retention, and budget signals as well as direct access to outcomes. The outcome-independent failure model below preserves that protection.

R.2 The affine witness for a fixed selection channel

Proposition 13 (Fixed-channel affine protection)**.** *Let $V \succ 0$, $\Omega \succ 0$, $X \sim N_M(\mu, V)$, and $U \sim N_{M-1}(0, \Omega)$ independently. Fix $F = CX + U$. Among reports $G = AX + BU + b$ with fixed coefficients, unbiasedness for every $\mu \in \mathbb{R}^M$ and independence from F uniquely give*

$$G = X - VC^\top \Omega^{-1} U, \qquad V_G = V + VC^\top \Omega^{-1} CV. \tag{R.1}$$

Allowing additional independent zero-mean noise only increases covariance in positive-semidefinite order. If C has rank $M - 1$ and $\ker C = \mathrm{span}\{\mathbf{1}\}$, then

$$V - VC^\top (CVC^\top)^{-1} CV = \frac{\mathbf{1}\mathbf{1}^\top}{\mathbf{1}^\top V^{-1} \mathbf{1}}. \tag{R.2}$$

Proof. The identity $A\mu + b = \mu$ for all μ forces $A = I$ and $b = 0$. Gaussian independence requires $\mathrm{Cov}(G, F) = VC^\top + B\Omega = 0$, determining B uniquely. Expansion gives the covariance in Eq. (R.1). An independent noise term contributes its nonnegative covariance. For Eq. (R.2), the matrix on the left is positive semidefinite of rank one and is annihilated by C, so it equals $c\mathbf{1}\mathbf{1}^\top$. Multiplication by $V^{-1}\mathbf{1}$ gives $\mathbf{1}$ and identifies $c = (\mathbf{1}^\top V^{-1}\mathbf{1})^{-1}$. □

For $\Omega = \lambda CVC^\top$, full-vector fission adds, relative to contrast-only protection, exactly $\lambda^{-1}\mathbf{1}\mathbf{1}^\top/(\mathbf{1}^\top V^{-1}\mathbf{1})$. This rank-one covariance difference isolates the additional level-noise penalty of full-vector fission. The optimum above is within the fixed-channel affine class.

R.3 Predictable Gaussian protection and weighted reporting

Let $\mathcal{H}_{t-1}$ contain past released F values and independent selector randomization. Before world t, choose an $\mathcal{H}_{t-1}$-measurable $\Omega_t \succ 0$. Obtain a fresh $X_t \sim N_M(\mu, V)$ and fresh auxiliary Gaussian innovation, conditionally independent of each other and of the past, and construct

$$F_t = CX_t + U_t, \qquad G_t = X_t - VC^\top \Omega_t^{-1} U_t, \quad U_t \mid \mathcal{H}_{t-1} \sim N(0, \Omega_t). \tag{R.3}$$

The covariance V is known and fixed. The raw X_t, auxiliary innovations or reconstructing seeds, and G_t are unavailable to the selector. Failures and cost signals used by the selector may use released information or additional independent randomness, but no further information about these sealed quantities.

Proposition 14 (Predictable-channel reporting)**.** *Suppose the winner, an almost-surely finite nonempty retained index set I, stopping, and completion are measurable with respect to the allowed transcript $\mathcal{T}$. Conditional on that transcript,*

$$\mathcal{L}((G_t)_{t\in I} \mid \mathcal{T}) = \bigotimes_{t\in I} N_M(\mu, V_{G,t}), \qquad V_{G,t} = V + VC^\top \Omega_t^{-1} CV.$$

For $v_{t,i} = (V_{G,t})_{ii}$, define

$$P_i = \sum_{t\in I} v_{t,i}^{-1}, \qquad \widehat{\mu}_i^{\mathrm{pw}} = P_i^{-1} \sum_{t\in I} \frac{G_{t,i}}{v_{t,i}}. \tag{R.4}$$

Then $\widehat{\mu}_{\hat{\imath}}^{\mathrm{pw}} \mid \mathcal{T} \sim N(\mu_{\hat{\imath}}, P_{\hat{\imath}}^{-1})$. These weights minimize conditional variance among transcript-measurable linear unbiased combinations of the selected coordinate's observations.

Proof. Given the past, Eq. (R.1) makes F_t and G_t independent. Their conditional joint density factors into an F_t density and a G_t density with covariance $V_{G,t}$. Multiplying these factors over time and conditioning on the released transcript fixes all channel choices without conditioning on a protected value. Retention and stopping depend only on that transcript, giving the asserted product law. The weighted sum is therefore Gaussian with the displayed mean and variance. For weights w_t summing to one, Cauchy–Schwarz gives $1 \le (\sum_t w_t^2 v_{t,i})(\sum_t v_{t,i}^{-1})$, with equality for weights proportional to $v_{t,i}^{-1}$. □

The precision condition $z_{1-\delta_{\mathrm{rep}}/2}/\sqrt{P_{\hat{\imath}}} \le \eta$ is transcript-measurable and can determine completion without opening G. Here the protected variance may depend on earlier F values, so conditional factorization supplies the guarantee. The variance optimum concerns scalar weights on the selected coordinate with the acquired channels fixed. For a fixed $M = 20$, $\sigma = .1$, $\rho = .67$ design with 25 worlds at $\lambda = .25$ and 25 at $\lambda = 4$, the ordinary mean has RMSE 0.018255, while precision weighting gives 0.017081 in native units.

Selection inference must also use the changing channel variance. For $a_{ij}^\top C = e_i^\top - e_j^\top$, put

$$Z_{ij,t} = a_{ij}^\top F_t, \quad s_{ij,t}^2 = a_{ij}^\top (CVC^\top + \Omega_t) a_{ij}, \quad V_{ij,n} = \sum_{t=1}^{n} s_{ij,t}^2.$$

Choose $\delta_{ij} > 0$ with $\sum_{i<j} \delta_{ij} \le \delta_{\mathrm{sel}}$ and fix $\zeta_{ij} > 0$ before observing data. A Gaussian mixture construction (Howard et al., 2021) gives the pairwise half-width

$$r_{ij,n} = \frac{1}{n}\sqrt{(V_{ij,n} + \zeta_{ij}) \log \frac{V_{ij,n} + \zeta_{ij}}{\zeta_{ij}\delta_{ij}^2}}. \tag{R.5}$$

Indeed, with $S_n = \sum_{t\le n}(Z_{ij,t} - \Delta_{ij})$, $\exp\{uS_n - u^2 V_{ij,n}/2\}$ is a nonnegative martingale for each fixed u. Mixing over $u \sim N(0, 1/\zeta_{ij})$ gives

$$M_n = \sqrt{\frac{\zeta_{ij}}{\zeta_{ij} + V_{ij,n}}} \exp\left\{ \frac{S_n^2}{2(\zeta_{ij} + V_{ij,n})} \right\}.$$

Ville's inequality bounds $\Pr(\sup_n M_n \ge 1/\delta_{ij})$ by δ_{ij}; inversion gives Eq. (R.5). A union bound over pairs proves simultaneous coverage. The accumulated conditional variance $V_{ij,n}$ adapts the boundary to the varying Ω_t; acquisition efficiency then depends on the channel policy.

R.4 A VARIANCE-SENSITIVE BOUNDED PILOT

Retain the independent pilot of Proposition 12 and write $L = R^2/2$, so $0 \le Z_k \le L$ and $\mathbb{E}Z_k = v$. With $b = L\log(1/\delta_P)/m$, an alternative upper bound is

$$v_U^C = \min\left\{ \frac{R^2}{4},\ \bar{Z} + b + \sqrt{b^2 + 2b\bar{Z}} \right\}. \tag{R.6}$$

To prove it, convexity gives $\mathbb{E}e^{-uZ_k} \le \exp\{(v/L)(e^{-uL}-1)\}$ for $u \ge 0$. Chernoff optimization and the multiplicative lower-tail bound yield

$$\Pr(\bar{Z} \le v - x \mid \mathcal{S}) \le \exp\{-mx^2/(2Lv)\}, \qquad 0 \le x \le v.$$

Consequently $v - \bar{Z} \le \sqrt{2bv}$ except on an event of probability at most δ_P. Solving for $\sqrt{v}$ and using $v \le R^2/4$ proves $\Pr(v > v_U^C \mid \mathcal{S}) \le \delta_P$. Using v_U^C in Eq. (L.6), then collecting a new independent final report, preserves Proposition 12's unbiasedness and $\delta_P + \delta_{\mathrm{rep}}$ reporting-error bound.

Table 49: Alternative bounded pilot calculations at $R = 100$, $\eta = 1$ pp, $\delta_P = .005$, $\delta_{\mathrm{rep}} = .025$, conditional on $\bar{Z} = 4$ squared pp. Each total includes $2m$ pilot observations and the fresh report.

m	Hoeffding v_U	Chernoff v_U^C	Old total	New total
100	817.8118	537.8020	7,660	5,206
250	518.6998	219.8599	5,339	2,719
500	367.9477	113.8258	4,517	2,290
1,000	261.3499	60.7197	4,583	2,825
2,000	185.9739	34.0213	5,923	4,591

At $m = 500$, conditional on the displayed pilot statistic, the reporting-only ratio to the fixed Hoeffding report is $20{,}999/2{,}290 = 9.1699$; selection and audit are additional acquisitions. The relative tightness depends on the pilot statistic: with $b/L = .01$ and $\bar{Z}/L = .30$, the old and new bounds are $0.370711L$ and $0.388102L$, both below the cap $L/2$. The table compares separately chosen rules. A generic data-dependent minimum of two bounds requires joint error control; allocating $\delta_H + \delta_C = \delta_P$ before taking that minimum is sufficient.

R.5 The cost ceiling for reuse alone

Proposition 15 (Same-selector reuse ceiling)**.** *Consider the same complete-vector Gaussian selection process and winner under two reporting routes. Let selection take n_S full worlds, fresh winner-only reporting require n_R observations, and structured reporting require n_G full worlds. Use the same known-variance Gaussian precision rule in both routes. Then $V_G \succeq V$ implies $n_G \ge n_R$. Without a hard cap, or when both routes complete their prescribed reports,*

$$C_F = Mn_S + n_R, \quad C_G = M\max(n_S, n_G), \qquad \frac{C_F}{C_G} \le 1 + \frac{1}{M}. \tag{R.7}$$

Proof. $C_G \ge M\max(n_S, n_R)$, while $C_F \le (M+1)\max(n_S, n_R)$. Division proves the claim. □

The ceiling is 1.08333, 1.06250, 1.04167, and 1.03125 for $M = 12, 16, 24, 32$. A 2.3 ratio therefore requires differences beyond reuse of the same complete selection vectors, such as allocation, active-set queries, or completion.

For $M = 20$, $\sigma = 2$ pp, $\rho = .9$, $\lambda = 1$, $\eta = 1$ pp, and $\delta_{\mathrm{rep}} = .03$, the counts are $n_R = 19$ and $n_G = 21$. If the common selection path uses $n_S = 10, 21, 50, 100$, the respective (C_F, C_G) pairs are $(219, 420)$, $(439, 420)$, $(1019, 1000)$, and $(2019, 2000)$. Reuse can therefore cost more when too few full vectors have been acquired. Separately, selection costs of 1,000 versus 3,000 with a common 20,999-observation report and 128-observation audit give totals 22,127 versus 24,127, a ratio of 1.090387. This is Eq. (O.5) with the unchanged stages charged explicitly.

R.6 The distributional requirement for Gaussian reporting

Under an exactly Gaussian winner law, the count $n_R = \lceil z_{.985}^2 \sigma^2 \rceil$ targets $\eta = 1$ pp and $\delta_{\mathrm{rep}} = .03$. For $\sigma = 1, 2, 3, 5, 10$ pp the counts are $5, 19, 43, 118, 471$, and the respective probabilities of absolute error at most 1 pp are 97.4653%, 97.0702%, 97.1171%, 97.0186%, and 97.0012%. The Gaussian interval itself has exactly 97% coverage; integer rounding can make its radius smaller than 1 pp.

The role of Gaussianity is exposed by the bounded distribution

$$Y = \begin{cases} 73, & \text{with probability } 4/533, \\ 50 - 4/23, & \text{with probability } 529/533. \end{cases} \tag{R.8}$$

It has $\mathbb{E}Y = 50$, $\mathrm{Var}(Y) = 4$, and $Y \in [0, 100]$. For 19 independent observations, zero high values gives error $4/23 < 1$; one high value already gives error $1.045767 > 1$, and additional high values increase the error. Thus

$$\Pr(|\bar{Y} - 50| > 1) = 1 - (529/533)^{19} = 13.3356\%.$$

This exact probability shows how a bounded distribution with the same known variance exceeds the Gaussian reporting error target.

R.7 A FIXED-DESIGN COMPLETION–CORRECTNESS CURVE

An exactly calculable Gaussian design provides a population comparison for completion and joint correctness. Set $M = 20$, marginal $\sigma = 2$ pp, one winner with gap $\Delta = .5$ pp over 19 tied alternatives, $\varepsilon = .2$ pp, $\eta = 1$ pp, $\delta_{\text{sel}} = .02$, and $\delta_{\text{rep}} = .03$. Each row below fixes its selection length n before observing outcomes. All candidates are evaluated on n complete worlds with equicorrelation ρ. The winner is reported using 19 independent fresh observations only when the fixed-time selection certificate passes. There are no failures or audits in this design.

Write $a = \sigma\sqrt{(1-\rho)/n}$, $z_S = z_{1-\delta_{\text{sel}}/\binom{M}{2}}$, and $h = \max\{\sqrt{2}z_S a - \varepsilon, 0\}$. Conditioning on the residual noise of a candidate gives

$$\begin{aligned} q_* &= \int \phi(z)\Phi\left(z + \frac{\Delta - h}{a}\right)^{M-1} dz, \\ q_- &= (M-1)\int \phi(z)\Phi\left(z - \frac{h}{a}\right)^{M-2} \Phi\left(z - \frac{h+\Delta}{a}\right) dz. \end{aligned} \tag{R.9}$$

Here q_* is a correct-winner certificate probability and q_- includes certificates for all incorrect winners. Since $\Delta > \varepsilon$,

$$Q = q_* + q_-, \quad J = q_*\{2\Phi(\sqrt{19}/2) - 1\}, \quad W = Q - J, \quad \mathbb{E}C = Mn + 19Q. \tag{R.10}$$

The small but nonzero q_- term is included in the calculations.

Table 50: Population outcomes of fixed-n Gaussian designs at $\rho = .67$. Each row is a single-inspection design with its own precommitted length. Costs include fresh reporting on completion.

n	Q (%)	J (%)	W (%)	Expected evaluations
50	0.3310	0.3213	0.0097	1,000.06
100	13.6567	13.2566	0.4001	2,002.59
150	55.0136	53.4018	1.6118	3,010.45
200	87.2278	84.6722	2.5556	4,016.57
250	97.7792	94.9145	2.8648	5,018.58
300	99.7302	96.8082	2.9219	6,018.95
400	99.9980	97.0683	2.9298	8,019.00

The smallest fixed lengths meeting $Q \geq .95$ and $J \geq .90$ are $n = 692, 346, 229, 70$ for $\rho = 0, .5, .67, .9$, respectively, with expected costs 13,858.05, 6,938.05, 4,598.07, and 1,418.14. These design comparisons use deterministic Gaussian integration. An adaptive choice to stop across these lengths needs its own time-uniform boundary.

R.8 FINITE INSPECTION SCHEDULES AND SAMPLING RESOLUTION

The bounded selector in Appendix I has nine predeclared checkpoints. If these are the only possible inspections, $\alpha_{ij,\ell} = \delta_{\text{sel}}/[9\binom{M}{2}]$ supplies a finite-schedule alternative to its infinite summable spending. The fixed-time empirical Bernstein inequality (Maurer & Pontil, 2009), followed by a

union bound over all pairs and nine looks, remains valid. At the displayed $n = 4096$ pair with $M = 16$, $\delta_{\text{sel}} = .02$, $\bar{D} = -.007812$ and $s^2 = .00041882$, the half-width becomes 0.015582658 and the advantage upper bound becomes 0.007770658. The original spending gives 0.018081306 and 0.010269306 instead. Only the former upper bound is below $\varepsilon = .01$.

This calculation changes one pair's decision at fixed sufficient statistics. If every other required comparison also passes and the earlier path is unchanged, avoiding the final $3 \cdot 4096$ evaluations would reduce 50,590 to 38,302, or 24.2894%. A different spending rule can change earlier eliminations; its actual total requires rerunning or replaying the full alternative path.

Repeated fixed-time intervals require different accounting. For independent new batches with exact 95% intervals, stop at the first interval excluding the true zero mean, or after L batches. The returned interval covers only when all L intervals cover, with probability $.95^L$. For $L = 5, 10, 20$, this is 77.3781%, 59.8737%, and 35.8486%. The calculation concerns independent batches; nested-sample inspection induces a different dependence structure.

Table 51: Sampling resolution in a single iid Bernoulli model. The first panel gives predictive count ranges when the population probability is fixed. The second gives one-sided 95% Clopper–Pearson lower bounds conditional on observed counts.

Repetitions	Population probability	Central 95% count range
10,000	.95	9,457–9,542
100,000	.95	94,864–95,135
190	.95	174–186
190	.98	182–189
190	.93	169–183
Observed count	**Empirical rate (%)**	**Lower bound (%)**
181/190	95.2632	91.8798
185/190	97.3684	94.5468
186/190	97.8947	95.2474
190/190	100.0000	98.4357

Completion targets concern population probabilities, with uncertainty quantified from observed counts. Table 51 uses one iid Bernoulli population; fixed benchmark strata call for setting-specific inference. For paired noninferiority at margin $m_0 = .02$, true difference zero, one-sided size .05, power .80, and discordance probability d, normal planning gives $N \simeq (z_{.95} + z_{.80})^2 d/m_0^2$. For $d = .01, .02, .05, .10$, rounding upward gives $155, 310, 773, 1546$ pairs. These are normal-approximation planning counts; the observed paired contingencies determine the subsequent comparison.

R.9 Reference, acquisition failures, and transformed pools

For 84 Gaussian candidates with known marginal standard deviation 4 pp, the fixed-bank Bonferroni mean-interval half-width is $h_N = 4z_{1-.05/(2\cdot 84)}/\sqrt{N}$. At $N = 256, 1024, 4096, 16384, 65536$ worlds it is $0.858444, 0.429222, 0.214611, 0.107305, 0.053653$ pp. Half-widths .2, .1, .05 pp require 4,717, 18,866, and 75,462 worlds under this model. Classification also depends on distance from the regret and reporting thresholds: without a positive margin, no fixed finite bank can generally resolve every output. If 190 independent report errors have standard deviation 1 pp and share an independent 256-observation reference of standard deviation 4 pp, the bias-estimate standard error is $\sqrt{1/190 + 16/256} = 0.260314$ pp, compared with 0.072548 pp from reports alone. Increasing only the outer count leaves the reference contribution at 0.25 pp.

When each attempt independently succeeds with probability p, independently of every potential score, the mean attempts for 19 accepted reports is $19/p$. For $p = .99, .95, .90$, these means are $19.1919, 20, 21.1111$; the smallest caps satisfying $\Pr\{\text{Bin}(B, p) \geq 19\} \geq .99$ are $21, 23, 25$ under the stated score-independent failure law. An independent post-decision audit remains additional evidence, with Gaussian 97% half-width $2z_{.985}/\sqrt{n}$ if $\sigma = 2$ pp is known, compared with $100\sqrt{\log(2/.03)/(2n)}$ under bounded Hoeffding reporting. At $n = 128$ these are 0.383621 and 12.808239 pp; the observation model therefore determines audit precision at a fixed count.

Proposition 16 (Transferring guarantees through a pool transformation)**.** *Let π map the original pool onto the transformed pool and suppose $\sup_i |\mu_i - \mu'_{\pi(i)}| \le \xi$. If a transformed winner is ε-optimal, any corresponding original candidate is $(\varepsilon + 2\xi)$-optimal. A report within η of the transformed winner's mean is within $\eta + \xi$ of its corresponding original mean.*

Proof. For the selected original candidate i, every original challenger j satisfies $\mu_j - \mu_i \le \mu'_{\pi(j)} - \mu'_{\pi(i)} + 2\xi \le \varepsilon + 2\xi$. The reporting claim follows by the triangle inequality. □

At $\xi = .03$ pp, $(\varepsilon, \eta) = (.20, 1.00)$ transfers to $(.26, 1.03)$ pp. The transfer uses a uniform mean bound for the complete executable schedules, including all their task-dependent allocations.

For uniform allocation, an integer budget $B = Mq + r$, $0 \le r < M$, can be implemented by giving each candidate q evaluations and assigning one extra to the first r candidates in a precommitted random permutation. Fixing B before the race specifies this outcome-independent allocation. Elimination thresholds also determine the active-set path: $\bar{D}_{ij} = -.10$, radius .03, and $\varepsilon = .20$ give $U_{ij} = -.07$, which passes a zero-threshold elimination test but not an $U_{ij} < -\varepsilon$ test.

R.10 Numerical scales of the strengthened lower bound

At $\sigma = 2$ pp, $|\Delta| = .5$ pp, $\varepsilon = .2$ pp, $\eta = 1$ pp, $\delta = .05$, and $\gamma = .95$, the binary divergence is $\mathrm{kl}(.90, .05) = 2.376205$. Table 52 evaluates Theorem 2. Its reporting term uses marginal variance $v = 4$ rather than pair-average variance v_A; the strengthening factor is $v/v_A = 2/(1+\rho) \in [1, 2]$ in this shared-noise model.

Table 52: Necessary information scales in the complete-pair Gaussian action model. The two middle columns bound the same expected world count; the last is twice their maximum.

ρ	Ranking worlds	Reporting worlds	Joint evaluation lower bound
0	77.5904	4.7524	155.1808
.5	38.7952	4.7524	77.5904
.67	25.6048	4.7524	51.2097
.9	7.7590	4.7524	15.5181
.99	.7759	4.7524	9.5048

The necessary acquisition scales vary with the gap. For $\varepsilon = .2$, the ranking scales at $|\Delta| = 1, .5, .21$, relative to $|\Delta| = .5$, are .3403, 1, and 2.9149. The unique-acceptable-winner argument requires $|\Delta| > \varepsilon$; the regime $|\Delta| \le \varepsilon$ has multiple acceptable candidates.

S Application Run-Level Reporting and Acquisition Record

This appendix presents completion, acquisition-cost, and reporting-error diagnostics for an additional PAIRS cohort. It contains 190 outer runs, with 60, 50, 40, and 40 runs in the four application settings, and produces 177 completed outputs and 177 pairs of same-winner reports. Table 1 describes the selection study; Appendix E defines the joint reporting protocol. Primary application evidence comprises the task-order reporting and cost study in Appendix V and the 1,000-pair budget evaluation in Appendix W.

S.1 Completion and actual reporting acquisitions

Each completed run acquires 20,999 report observations and 128 audit observations; the 13 abstaining runs incur selection cost only. The total reporting count is $177 \cdot 20{,}999 = 3{,}716{,}823$. Table 54 charges every acquisition to its stage, including selection on abstaining runs. The reporting count follows the fixed bounded-score rule in Eq. (E.1). The acquisition unit is the configuration–world observation defined by the experiment's target law.

The completion intervals in Table 53 treat outer completion indicators as iid Bernoulli observations within each setting. A pointwise 95% Clopper–Pearson interval is obtained by inverting the two

Table 53: PAIRS run-level reporting study. Completion is K/N over all runs; intervals are pointwise exact binomial 95% intervals. Pooled bounds combine simultaneous setting intervals with fixed replicate weights. Acquisition totals include abstaining runs.

Setting	K/N	Completion % [95% CI]	Report obs.	All attempts
CIFAR-10	58/60	96.67 [88.47,99.59]	1,217,942	1,256,086
CIFAR-100	46/50	92.00 [80.77,97.78]	965,954	1,010,242
CORe50	38/40	95.00 [83.08,99.39]	797,962	818,186
RS-MCTS	35/40	87.50 [73.20,95.81]	734,965	780,405
Pooled	177/190	93.16 [78.74,98.87]	3,716,823	3,864,919

Table 54: PAIRS reporting-study stage totals. Counts are acquisition attempts. The total covers selection, reporting, and audit; proposal, gate, and reference have separate accounting scopes.

Setting	Selection	Reporting	Audit	Total
CIFAR-10	30,720	1,217,942	7,424	1,256,086
CIFAR-100	38,400	965,954	5,888	1,010,242
CORe50	15,360	797,962	4,864	818,186
RS-MCTS	40,960	734,965	4,480	780,405
Total	125,440	3,716,823	22,656	3,864,919

binomial tails at level 0.025. For the fixed mixture $Q_{\mathrm{mix}} = \sum_s (N_s/190) Q_s$, we first construct setting intervals $[l_s, u_s]$ at level $1 - 0.05/4$ and then report $[\sum_s N_s l_s/190, \sum_s N_s u_s/190]$. A union bound gives at least 95% coverage for this conservative pooled interval without assuming equal completion probabilities across settings. The target population is the fixed mixture of these four settings.

S.2 RECORDED CORRECTNESS FLAGS AND REPORT DIAGNOSTICS

The joint-correct field contains counts 56, 46, 38, and 35, totaling 175; the incorrect-certificate field contains 2, 0, 0, and 0. In CIFAR-10 rows R009 and R010, completion and both component-correctness flags equal one, whereas the joint flag equals zero. Applying $C\,\mathbf{1}\{G_{\mathrm{sel}}\}\mathbf{1}\{G_{\mathrm{rep}}\}$ to the component fields gives 177 joint-correct and zero incorrect outputs. Thus the stored joint field and the component-derived calculation differ on these two rows. We preserve both field mappings. Because the stored joint field and the component-derived flags disagree on two records, this cohort is used only for completion, acquisition, and reporting diagnostics; it is not used for headline joint-correctness claims. Appendices V and W provide the primary joint comparisons.

The reporting subset contains all 177 rows with both reports, with setting denominators 58, 46, 38, and 35. Errors subtract each row's scalar reference value. Table 55 reports signed bias and MAE as point diagnostics conditional on these reference values. Completion retains all 190 runs in its denominator.

Table 55: Same-winner report diagnostics for the additional run record. All values are in pp. Macro absolute signed bias takes absolute values after averaging signed errors within settings; MAE averages absolute row errors. The summary row therefore changes the first pair of columns from signed to absolute bias, as labeled.

		Signed bias		MAE	
Setting	Reports	Naive	PAIRS	Naive	PAIRS
CIFAR-10	58	+0.5828	+0.0234	0.5828	0.0276
CIFAR-100	46	+0.0370	−0.3922	0.1152	0.3922
CORe50	38	+0.3000	−0.0800	0.3000	0.0800
RS-MCTS	35	+0.3000	−0.0800	0.3000	0.0800
Macro absolute bias / MAE	177	0.3049	0.1439	0.3245	0.1449

The unrounded macro absolute-bias reduction is 52.806834% and the macro MAE reduction is 55.333526% within this available-report cohort. These are diagnostics for the available-report cohort; the primary reporting-bias and total-attempt results use the task-order cohort in Appendix V, while qualifying-budget results use Appendix W.

T Complete-Array Studies of Reporting, Covariance, and Replay Decisions

We study paired reporting in a fixed-design Gaussian experiment and examine individual sequential decisions in six bounded-noise executions. The Gaussian settings use candidate counts $16, 24, 12, 32$ and observations in the model's native utility units. Their fixed allocation isolates the effect of covariance on comparison cost. The bounded executions record adaptive elimination and stopping, and a calculation-validation dataset tests the associated accounting and export operations.

T.1 Gaussian comparison with fully retained observations

There are 500 independent outer runs per setting, hence $2{,}000$ paired runs. The four mean vectors have leading values $84, 66, 77, 71$, respectively. In each setting, candidate zero is best and the remaining candidates have gaps $1.5, 1.6, \ldots, 1.5 + 0.1(M-2)$. The observation standard deviation is $\sigma = 2$. The paired procedure, denoted P64-FR, observes $n_P = 64$ complete worlds with within-world correlation

$$\rho = 1 - \frac{64}{149} = 0.5704697987.$$

The independent procedure, I149-FR, observes $n_I = 149$ worlds with independent candidate coordinates. The resulting contrast variances agree:

$$\mathrm{Var}(\bar{Y}_i^P - \bar{Y}_j^P) = \frac{2\sigma^2(1-\rho)}{64} = \frac{2\sigma^2}{149} = \mathrm{Var}(\bar{Y}_i^I - \bar{Y}_j^I). \tag{T.1}$$

Across methods, the execution couples the sample means through identical residual contrast means and a different common level. Consequently the two methods select the same candidate and make the same selection-certification decision in each paired run. Complete observations are reconstructed from independent Gaussian sample means and centered Gaussian residuals; the saved world arrays contain the resulting observations, not just their means.

Both methods inspect only their own fixed endpoint. For $\delta_{\mathrm{sel}} = 0.02$, the simultaneous normal halfwidth is

$$h_M = z_{1-\delta_{\mathrm{sel}}/(2\binom{M}{2})}\sqrt{8/149}. \tag{T.2}$$

The empirical leader is certified if every competitor's advantage upper bound is at most $\varepsilon = 0.2$. An independent, shared availability variable with probability 0.99 determines whether the precommitted report can be completed. On completion, each method acquires 19 fresh winner observations and eight further audit observations. Fresh reports are independent of selection, with marginal variance $4/19$ and cross-method correlation 0.7. Their known-variance normal intervals use $\delta_{\mathrm{rep}} = 0.03$ and halfwidth $z_{0.985}2/\sqrt{19} < \eta = 1$. The audit has no effect on the output.

The execution identifiers are `SIM_FIXED64_PAIRED_GAUSSIAN_FR19_A8_v1` and `SIM_FIXED149_INDEPENDENT_GAUSSIAN_FR19_A8_v1`. The sample sizes were fixed before the recorded draws to study a near-2.3 comparison at equal contrast precision. Equation (T.1) identifies the covariance mechanism responsible for this fixed-allocation cost ratio.

Table 56 separates completion from joint correctness. All $2{,}000$ paired runs select candidate zero and pass selection certification; the 18 abstentions arise from the independent availability variable. Against generating means, the incorrect-certificate frequencies are 3.10% and 2.90% of all runs. Evaluation against the finite reference gives disagreement frequencies of 4.35% and 4.25%. Equal allocation makes the macro and pooled rates identical. The saved identifiers `SIM_C10`, `SIM_C100`, `SIM_CORE`, and `SIM_RS` map in order to G16, G24, G12, and G32 in these tables.

The accounting identity is $T_m = n_m M + 27C_m$ for each run. Table 57 includes every selection acquisition and the report/audit acquisitions of completing runs. Offline reference scoring acquires an additional shared $256(16 + 24 + 12 + 32) = 21{,}504$ observations, recorded separately from the two procedures' selection-and-reporting totals. All costs in this table are candidate–world acquisitions. The raw observation archive has $9{,}056{,}000$ entries, including masks and unacquired potential report/audit outcomes.

Table 56: Complete-array Gaussian comparison. Each row has 500 paired runs. P/I denotes P64-FR/I149-FR. Q is completion; J_μ scores against generating means and J_{ref} against the finite reference. All rates include abstentions in the denominator.

Setting	M	Completed	Q (%)	J_μ P/I (%)	J_{ref} P/I (%)
G16	16	495/500	99.00	96.20 / 96.00	96.20 / 95.40
G24	24	496/500	99.20	95.20 / 96.80	94.40 / 94.80
G12	12	497/500	99.40	96.40 / 96.40	94.60 / 93.80
G32	32	494/500	98.80	96.20 / 95.60	93.80 / 95.40
Pooled / macro	–	1,982/2,000	99.10	96.00 / 96.20	94.75 / 94.85

Table 57: Acquired candidate–world observations in the Gaussian comparison. Reporting and audit columns are per method. Potential observations for noncompleted outputs remain in the arrays but are not charged as acquired. Finite-reference generation is reported separately.

Setting	Selection P/I	Report	Audit	Total P	Total I	I/P
G16	512,000 / 1,192,000	9,405	3,960	525,365	1,205,365	2.29434
G24	768,000 / 1,788,000	9,424	3,968	781,392	1,801,392	2.30536
G12	384,000 / 894,000	9,443	3,976	397,419	907,419	2.28328
G32	1,024,000 / 2,384,000	9,386	3,952	1,037,338	2,397,338	2.31105
Total	2,688,000 / 6,258,000	37,658	15,856	2,741,514	6,311,514	2.30220

T.2 Paired uncertainty and model-specific completion

Each setting has an independent reference bank of 256 complete Gaussian worlds, marginal standard deviation 2, and within-world correlation 0.75. All methods and outputs in that setting use the same bank. Ten thousand stored bootstrap draws resample outer runs within settings using identical indices for the two methods. The extended analysis independently resamples complete reference-world rows, then shares the new reference vector across all sampled outputs of that setting. Winners, certificates, costs, and completion masks remain frozen. The two-source procedure retains both between-method pairing and common reference uncertainty.

The finite-reference paired correctness counts $(n_{11}, n_{10}, n_{01}, n_{00})$ are $(1841, 54, 56, 49)$; against generating means they are $(1883, 37, 41, 39)$. Here n_{10} means P64-FR alone is correct. The shared-reference analysis broadens the joint-quality comparison (Table 58), whereas all 10,000 cost-ratio draws remain identical because reference resampling leaves the acquisitions fixed. The cost interval therefore reflects the precision of expenditure under this fixed design.

For this Gaussian model, completion also admits a direct bound. Put $s = \sqrt{8/149}$ and let $\Delta_j = 1.5 + 0.1(j-1)$, $j = 1, \ldots, M-1$. On the event $E = \bigcap_j \{\bar{Y}_0 - \bar{Y}_j \geq h_M - \varepsilon\}$, candidate zero is the empirical leader because $h_M > \varepsilon$ in all four settings. Every competitor upper bound is then at most ε, so selection is certified. A union bound gives

$$f_M = \sum_{j=1}^{M-1} \Phi\left(\frac{h_M - \varepsilon - \Delta_j}{s}\right), \tag{T.3}$$

$$Q_M \geq 0.99(1 - f_M), \qquad J_M \geq 0.99(1 - f_M)\big[2\Phi(\sqrt{19}/2) - 1\big]. \tag{T.4}$$

Independent availability and fresh reporting supply the two factors. Averaging over the four settings yields $Q \geq 98.9614\%$ and $J \geq 96.0620\%$ for both procedures under the stated Gaussian model. The coupled contrasts and identical fresh-report marginals also give equal population Q and J across methods; the realized joint-correctness rates are 96.00%/96.20%. A simultaneous analysis of the observed oracle counts gives lower bounds 97.1128% for both completion rates and 92.8165%/93.0733% for joint correctness. It uses 24 Bonferroni-adjusted one-sided binomial statements, yielding more conservative bounds than the model-specific calculation.

Table 58: Gaussian study: point estimates and approximate 95% percentile intervals from 10,000 paired resamples. Rate and difference entries are percentages and percentage points, respectively. The cost ratio is unitless.

Metric	Estimate	Outer only	Outer + reference
$Q_P = Q_I$	99.10	[98.65,99.50]	[98.65,99.50]
$J_{\mathrm{ref},P}$	94.75	[93.75,95.70]	[91.00,96.05]
$J_{\mathrm{ref},I}$	94.85	[93.90,95.80]	[91.55,96.15]
$J_{\mathrm{ref},P} - J_{\mathrm{ref},I}$	-0.10	[-1.10,0.90]	[-1.85,1.15]
T_I/T_P	2.30220	[2.30210,2.30232]	[2.30210,2.30232]

Table 59: All six bounded sequential executions. P: practical EB; R: contamination-aware EB with $c = 0.01$. $C = 1$ indicates a joint output. Dashes mean no report was acquired. Every issued joint certificate is correct against the generating means.

Variant	Gap	Last n	Selection	Report	Audit	Total	C
P	1.20	1,024	2,816	1,890	16	4,722	1
R	1.20	2,048	5,376	1,890	16	7,282	1
P	0.45	2,048	4,864	1,890	16	6,770	1
R	0.45	8,192	17,152	–	–	17,152	0
P	0.08	4,096	8,960	1,890	16	10,866	1
R	0.08	8,192	17,152	–	–	17,152	0
Total	–	–	56,320	7,560	64	63,944	4

T.3 Bounded practical and contamination-aware sequential records

Six additional executions use four candidates with observations supported on $[70, 85]$, paired-difference width $W = 30$, $\varepsilon = 0.2$, $\eta = 0.5$, and $(\delta_{\mathrm{sel}}, \delta_{\mathrm{rep}}) = (0.02, 0.03)$. The checkpoints are $n_\ell = 64\,2^\ell$, $\ell = 0, \ldots, 7$. For each of the six candidate pairs, set

$$\alpha_{ij,\ell} = \frac{0.02}{6(\ell+1)(\ell+2)}.$$

The implemented practical radius and its contamination-aware extension are

$$r_{ij,n}(c) = \sqrt{\frac{2s_{ij,n}^2 \log(4/\alpha_{ij,\ell})}{n}} + \frac{7W \log(4/\alpha_{ij,\ell})}{3(n-1)} + Wc, \quad c \in \{0, 0.01\}. \tag{T.5}$$

The first two terms are the scaled empirical-Bernstein boundary (Maurer & Pontil, 2009). For iid whole-world contamination of probability at most c, the mixture mean of a bounded difference can differ from its clean mean by at most Wc. The added term therefore targets the clean difference under bounded-support contamination. All generated values in these executions satisfy the support restriction.

Whole active blocks precede interval updates. Candidates with $L_{j,i} > 0$ are removed simultaneously, and the empirical leader stops when all surviving competitor upper bounds are at most ε. These executions use the terminal checkpoint 8192 as their sample cap. Completed selections receive 1,890 independent clean fresh report observations and 16 audit observations, in addition to the selection acquisitions. The bounded report halfwidth is $15\sqrt{\log(2/0.03)/(2 \cdot 1890)} = 0.4999824$. Contamination affects selection worlds only; the reporting bank is clean.

The contamination-aware radius retains a 0.30-unit correction even as sampling variance shrinks. Two of its three runs reach the terminal checkpoint without certification, despite returning ε-optimal empirical leaders. This separates output quality from certificate availability. The archive retains 63,944 acquired observations, 144 pair intervals, and 40 checkpoint records, including active sets, elimination witnesses, sample variances, and cost increments. Potential worlds beyond the acquired active prefixes are stored separately. Each decision record links to the radius in Eq. (T.5) and its observed prefix.

T.4 COMPLETE ARCHIVE AND CALCULATION-LEVEL VALIDATION

The accompanying archive contains the source files, CSVs, arrays, manifests, and resampling indices. For the Gaussian study, each of the 2,000 paired records links to its row in the complete observation arrays and to a setting-specific reference bank. There are 3,964 output–reference links, one per completed method output. Each setting's bootstrap indices have shapes $10{,}000 \times 500$ for outer runs and $10{,}000 \times 256$ for reference worlds. For the bounded study, observed-prefix CSVs identify acquired data separately from unused potential worlds. Acquisition masks define the charged observations, and shared reference IDs identify repeated use of a single bank.

The archive also contains a synthetic calculation-validation fixture, `SIM_FIX40`, used to test denominator handling, joins, exports, and resampling. Its assigned residuals and acquisition fields are test constants rather than executed selector outputs and are not used in any empirical claim.

A six-world reference calculation checks the evaluation rule directly. With a frozen winner and report 80.97, whole-world resampling changes that winner's reference mean from 80.10 to 79.9167. Its absolute reporting discrepancy therefore changes from 0.87 to 1.0533, crossing the $\eta = 1$ threshold while the selected candidate, output, and cost remain fixed. This calculation illustrates how reference propagation changes evaluation flags with deployment and expenditure held fixed; its unit of analysis is the six-world bank.

U Bounded Sequential Paths and Contamination Resolution

Two four-candidate paths examine how a broad score range and a contamination allowance affect certification near the selection tolerance. The clean observation model is

$$Y_{w,i} = \mu_i + Z_w + E_{w,i}, \qquad \mu = (82.00, 81.88, 81.55, 81.20), \tag{U.1}$$

where $Z_w \sim U[-9, 9]$, $E_{w,i} \sim U[-0.025, 0.025]$, and all primitive variables are independent except for the shared Z_w within each world. Scores are bounded by $[0, 100]$. The procedure uses the corresponding pairwise range width $W = 200$, estimates difference variance from observed prefixes, and sets $\varepsilon = 0.20$ pp and $\delta_{\mathrm{sel}} = 0.02$.

Intervals are inspected at $n_\ell = 32\,2^\ell$, $\ell = 0, \ldots, 12$, using $\alpha_\ell = 0.02/[6(\ell+1)(\ell+2)]$ for each of the six pairs. With the unbiased difference sample variance $s^2_{ij,n}$, the clean empirical-Bernstein radius is

$$r^{\mathrm{clean}}_{ij,n} = \sqrt{\frac{2s^2_{ij,n}\log(4/\alpha_\ell)}{n}} + \frac{7W\log(4/\alpha_\ell)}{3(n-1)}. \tag{U.2}$$

The scaled fixed-time bound and pair–checkpoint error allocation give simultaneous prefix coverage as in Proposition 10 (Maurer & Pontil, 2009). Each requested active block completes before all eliminations with $L_{j,i} > 0$ are applied simultaneously. For the remaining empirical leader $\hat{i}$, the stopping statistic is $U_{\max} = \max_{j \neq \hat{i}} U_{j,\hat{i}}$ over active competitors. Selection is certified when $U_{\max} \leq 0.20$; the final checkpoint is 131,072.

The contamination-aware path replaces an entire world by $(70, 85, 85, 85)$ with probability $c = 0.005$, independently across worlds. For any bounded pairwise difference, the contaminated and clean means differ by at most $Wc = 1.00$. This path therefore uses $r^{\mathrm{cont}}_{ij,n} = r^{\mathrm{clean}}_{ij,n} + 1.00$, with the empirical variance computed from its contaminated observations. Its intervals target the clean means in Eq. (U.1).

Table 60 records all 24 inspected checkpoints. The clean path removes D at 8,192 and C at 16,384. At the latter checkpoint, the A–B mean difference is approximately 0.119914 and the radius is 0.336622, leaving $U_{B,A} = 0.216707 > 0.20$. At 32,768, the difference is 0.119936 and the radius is 0.171069, giving $U_{B,A} = 0.0511327826$ and selection certification. Its acquisition count is

$$4(8192) + 3(8192) + 2(16384) = 90{,}112. \tag{U.3}$$

The contamination-aware path keeps A as its empirical leader and retains all four candidates through the cap. Its final upper bound is 1.0085108473, so it abstains after $4(131{,}072) = 524{,}288$ acquisitions. Together the paths charge $614{,}400$ observations and inspect 136 pair intervals: 58 on the clean path and 78 on the contaminated path.

The contaminated path reveals a resolution limit of this radius. For the closest competitor B, its population advantage over A under the mixture is

$$\Delta^{\mathrm{mix}}_{B,A} = (1 - 0.005)(-0.12) + 0.005(85 - 70) = -0.0444. \tag{U.4}$$

Along an indefinitely observed prefix, the sampling terms in Eq. (U.2) vanish, while the contamination correction remains 1.00. The corresponding upper bound converges to $-0.0444 + 1.00 = 0.9556$, above $\varepsilon = 0.20$. Thus the conservative contamination allowance remains decisive even with arbitrarily precise estimation of the mixture mean. This separates sampling uncertainty from the ambiguity introduced by the contamination model.

Table 60: Complete checkpoint paths under the clean and contaminated observation laws. A is the clean path's active set on entry to the checkpoint; the contaminated path retains all four candidates. Counts are cumulative selection acquisitions. A dash marks a stopped path.

	Clean EB				Contamination-aware EB	
n	$U_{\max}$	Acquisitions	A	Update	$U_{\max}$	Acquisitions
32	117.056527	128	ABCD	Continue	118.056527	128
64	65.678254	256	ABCD	Continue	66.678254	256
128	35.070044	512	ABCD	Continue	36.070044	512
256	18.342365	1,024	ABCD	Continue	19.342365	1,024
512	9.465323	2,048	ABCD	Continue	10.711437	2,048
1,024	4.823003	4,096	ABCD	Continue	6.047319	4,096
2,048	2.416099	8,192	ABCD	Continue	3.557702	8,192
4,096	1.176870	16,384	ABCD	Continue	2.310568	16,384
8,192	0.541382	32,768	ABCD	Remove D	1.657408	32,768
16,384	0.216707	57,344	ABC	Remove C	1.325081	65,536
32,768	0.051133	90,112	AB	Certify	1.139633	131,072
65,536	–	–	–	–	1.049690	262,144
131,072	–	–	–	–	1.008511	524,288

V SPLIT-CIFAR-10 JOINT SELECTION AND REPORTING STUDY

We evaluate joint selection and reporting across four task orders on CIFAR-10, with five outer runs per order and four ER configurations. All 20 runs enter the analysis. The setting is the task order, and each setting receives weight $1/4$. This study underlies the 92% reporting-bias result in the abstract and the total-attempt ratio in Table 2. The qualifying-budget comparison appears in Appendix W. The four-benchmark selection study and the 190-run reporting diagnostics use separate cohorts.

V.1 EXECUTABLE CONFIGURATIONS AND UTILITY

The official 50,000-image training and 10,000-image test sets are partitioned into five two-class tasks. Table 61 fixes their order. The model is a CIFAR-adapted ResNet-18 with a 3×3, stride-one first convolution, no initial max-pooling, and one ten-class output head. Class-incremental evaluation uses all ten logits without a task identifier. After training task five, the utility is

$$Y_w(i) = \frac{1}{5}\sum_{k=1}^{5} 100\,\frac{\#\{\text{correct test predictions on task } k\}}{\#\{\text{test examples on task } k\}}. \tag{V.1}$$

The unit is percentage points, and larger values are better.

Each task uses 50 epochs of SGD with momentum 0.9, weight decay 0.0005, and no Nesterov acceleration. The current-data and replay batch sizes are both 32; a smaller memory supplies all available examples. The loss is the sum of mean current-batch and replay-batch cross-entropy, with zero replay loss for an empty memory. Training applies reflect padding of four pixels, random 32×32 cropping, and horizontal flipping with probability 0.5. Pixels are scaled to $[0, 1]$ without mean–standard-deviation normalization. The learning rate follows a within-task cosine decay to zero. Momentum is retained between tasks. Reservoir admission occurs only on a stream sample's first encounter, and replay samples uniformly without replacement. A new world resets the model, optimizer, and memory.

Method P uses paired selection and fresh winner-only reporting. Within a selection world, candidates share initialization, data order, and augmentation primitives; their buffer contents and replay-sampling trajectories remain candidate-specific. Method I uses independent candidate worlds with the same pool, boundary form, tolerances, and reporting rule. The naive estimate N reuses P's selection estimate for P's frozen winner. Thus N versus P compares reporting on the same configuration, whereas P versus I compares complete selection-and-reporting procedures.

Both procedures use $\varepsilon = 0.20$ pp, $\eta = 1.00$ pp, $\delta_{\text{sel}} = 0.02$, and $\delta_{\text{rep}} = 0.03$. Each completed output has a fixed fresh report of

$$m = \left\lceil \frac{100^2}{2\eta^2} \log \frac{2}{\delta_{\text{rep}}} \right\rceil = 20{,}999 \tag{V.2}$$

observations and a separate 128-observation post-output audit. Table 65 contains every frozen winner, report, joint score, and total attempted-observation count.

V.2 SAME-WINNER BIAS, ACCURACY, AND JOINT OUTPUT

For method $m \in \{N, P, I\}$, let $e_{m,sr} = \widehat{\mu}_{m,sr} - \widehat{\mu}^{\text{ref}}_{s,\hat{i}_{m,sr}}$ using reference version 2. The reporting comparison fixes $\hat{i}_{N,sr} = \hat{i}_{P,sr}$. Define

$$b_{m,s} = \frac{1}{5}\sum_{r=1}^{5} e_{m,sr}, \qquad B_m = \frac{1}{4}\sum_{s=1}^{4} |b_{m,s}|, \qquad \text{MAE}_m = \frac{1}{20}\sum_{s,r} |e_{m,sr}|. \tag{V.3}$$

The order of averaging and taking absolute values distinguishes B_m from MAE. Table 62 gives the setting-level results. In aggregate,

$$1 - \frac{B_P}{B_N} = 1 - \frac{0.070}{0.875} = 92\%, \qquad 1 - \frac{\text{MAE}_P}{\text{MAE}_N} = 1 - \frac{0.266}{0.875} = 69.6\%. \tag{V.4}$$

Fresh reporting reduces both endpoints, with the larger reduction in the setting-level signed-bias measure.

Table 61: Task orders and candidate configurations for the 20-run CIFAR-10 study. Each ordered pair lists official CIFAR-10 class IDs. Candidates share the training protocol described in the text.

Setting	Task 1	Task 2	Task 3	Task 4	Task 5
S1	(0,1)	(2,3)	(4,5)	(6,7)	(8,9)
S2	(8,9)	(6,7)	(4,5)	(2,3)	(0,1)
S3	(4,5)	(0,1)	(8,9)	(2,3)	(6,7)
S4	(2,3)	(8,9)	(6,7)	(0,1)	(4,5)

Candidate	A	B	C	D
ER memory capacity	200	200	500	500
Initial learning rate	0.03	0.10	0.03	0.10

Table 62: Task-order results. Bias and MAE are in pp and use reference version 2. Joint correctness applies both the 0.20 pp selection and 1.00 pp reporting criteria. All five outputs complete in every setting for both methods. Costs are attempted observations, including reporting, audit, and failed selection attempts.

	b_N	b_P	MAE_P	Joint P	Joint I	Attempts P	Attempts I
S1	0.50	$+0.04$	0.156	5/5	5/5	580,774	1,318,254
S2	0.75	-0.06	0.176	5/5	5/5	670,888	1,539,467
S3	1.00	$+0.08$	0.184	5/5	4/5	613,539	1,432,950
S4	1.25	-0.10	0.548	4/5	4/5	736,419	1,693,055
All	–	–	0.266	19/20	18/20	2,601,620	5,983,726

All selected configurations are reference-relative ε-optimal. Reporting error exceeds 1.00 pp for P on R018 (1.12 pp), and for I on R013 and R019 (1.08 and 1.14 pp). Accordingly, $(Q, J_{\mathrm{ref}}, W_{\mathrm{ref}})$ is $(100\%, 95\%, 5\%)$ for P and $(100\%, 90\%, 10\%)$ for I. The paired contingency counts $(n_{11}, n_{10}, n_{01}, n_{00}) = (17, 2, 1, 0)$ give a joint-success difference of 5 pp. These are reference-relative empirical endpoints.

Two exported joint flags in this cohort (S2/R009 and S2/R010) were generated using a stale 0.10-pp reporting-tolerance cache. Recomputing the flags from frozen outputs with the declared 1.00-pp tolerance changes both to correct. No winner, report, reference value, decision, or cost changes.

V.3 Attempted observations and physical execution

The cost ratio is $5{,}983{,}726/2{,}601{,}620 = 2.3$ under the same nominal precision and error targets. Table 63 includes 8 and 866 failed selection attempts, respectively. The paired uncertainty analysis in Table 64 reports this ratio alongside the joint-success difference. They measure distinct features of the procedures: resource use and attained output quality.

Each counted reporting observation corresponds to a distinct, independently executed configuration–world realization: $419{,}980$ per method across the 20 reports. Cached reads replay stored outcomes during analysis and are never counted as additional independent observations.

A physical timing pair records a physical generation cost of 650 seconds for a complete five-task configuration execution and an analysis-time retrieval cost of 0.72 seconds for reading its cached utility. Both operations retain one allocated GPU, corresponding to 0.180556 and 0.000200 allocation-hours. Generation produces a new configuration–world utility; retrieval accesses the stored observation. At a constant 650-second execution time, a $20{,}999$-observation report would require $3{,}791.4861$ GPU-hours. This is a budget extrapolation from the measured execution, while the study's cost ratio uses the attempted observation ledger. Repeated retrievals do not increase the number of independent reporting worlds.

Table 63: Complete P/I attempted-observation ledger for the task-order study. Reference-bank construction is an offline scoring stage; the ratio here covers selection, reporting, and audit.

Stage	P	I
Accepted selection	2,179,072	5,560,320
Failed selection attempts	8	866
Completed-unused selection	0	0
Fresh reporting	419,980	419,980
Post-output audit	2,560	2,560
Total attempts	2,601,620	5,983,726

Table 64: Paired resampling in the 20-run task-order study (20,000 draws). Outer-only analysis fixes the reference means; two-source analysis also resamples the 12 complete reference-world rows in each setting.

Endpoint	Estimate	Outer-only 95% range	Outer + reference
Absolute signed-bias reduction (%)	92.00	[74.2123,94.8630]	[69.7524,94.1765]
$J_{\mathrm{ref},P}$ (%)	95.00	[85,100]	[85,100]
$J_{\mathrm{ref},I}$ (%)	90.00	[75,100]	[75,100]
Joint-success difference (pp)	$+5$	$[-10,20]$	$[-10,20]$
Attempted-cost ratio I/P	2.300000	[2.271188,2.329128]	[2.271188,2.329128]

V.4 Shared finite-reference uncertainty

Each setting has one independent reference bank of 12 complete four-candidate worlds. Table 66 gives all 192 scores. All runs and methods in a setting use the same candidate-mean vector. The primary version-2 means are $(a_s, a_s - 0.12, a_s - 0.45, a_s - 0.80)$, where $(a_1, a_2, a_3, a_4) = (82, 68, 76, 71)$.

Let p_s be the vector of observed P winner frequencies in setting s, and let $\widehat{\Sigma}_s$ be the unbiased covariance matrix of its 12 reference-world vectors. Conditional on the frozen winners, the estimated reference contribution to the variance of the setting-level signed bias is

$$\widehat{V}_{\mathrm{ref},s} = p_s^{\top} \frac{\widehat{\Sigma}_s}{12} p_s. \tag{V.5}$$

Its standard errors are 0.09740, 0.10058, 0.04715, and 0.07361 pp for S1–S4. The shared bank contributes once through its joint covariance; reusing it in five outer runs does not create five independent references.

The two-source resampling analysis draws outer indices within each setting, using the same indices for P, I, and N, and independently resamples complete reference-world rows. Each resampled reference vector is shared across that setting's selected outputs. Winners, reports, decisions, and costs remain frozen. Table 64 reports the central 95% percentile ranges from 20,000 draws. The bias interval changes when reference uncertainty is propagated. The cost interval remains unchanged because reference resampling does not alter recorded computation. With only five outer runs per setting and 12 reference worlds, these ranges describe resampling uncertainty; the pathwise guarantees derive from the sequential and reporting bounds.

The joint-success range spans zero and the -2 pp noninferiority margin. This cohort therefore estimates the cost contrast precisely while leaving the joint-success contrast less resolved. A boundary case shows why reference uncertainty matters: resampling S4 rows $(10, 10, 11, 11, 7, 7, 8, 8, 9, 9, 6, 12)$ yields a mean for candidate A of 71.178333 pp. R018's frozen report of 72.12 then has reference-relative error 0.941667 pp rather than 1.12 pp, crossing the reporting tolerance without changing the selected configuration or its cost.

For version sensitivity, shifting the four reference banks by $(+0.08, -0.05, +0.03, -0.04)$ pp gives the earlier version-1 scores. The corresponding macro absolute signed biases are 0.870 and 0.040 pp, or a 95.4023% reduction. Version 2 is used consistently for the primary 92% result, all joint scores, and the tables in this section.

Table 65: Run-level training results for the local 20-run cohort. P uses paired selection and I uses independent selection; both use fresh reporting. The naive report estimates the same winner selected by P. All report values are in percentage points. Total attempts include accepted and failed selection evaluations, reporting, and audit. The final two columns give reference-scored joint-correctness flags recomputed at $\varepsilon = 0.20$ pp and $\eta = 1.00$ pp; all runs completed. The final row sums attempts and joint-correct outputs.

ID	P winner	Naive report	P fresh report	I winner	I fresh report	P total attempts	I total attempts	J_P^{ref}	J_I^{ref}
S1/R001	A	82.22	81.80	A	82.06	111,239	250,538	1	1
S1/R002	A	82.43	82.03	B	81.76	111,239	266,929	1	1
S1/R003	B	82.44	82.06	A	82.20	119,434	258,733	1	1
S1/R004	A	82.61	81.91	A	82.08	127,623	275,126	1	1
S1/R005	A	82.68	82.28	B	81.58	111,239	266,928	1	1
S2/R006	A	68.45	67.76	A	68.10	127,623	291,513	1	1
S2/R007	B	68.48	67.99	A	67.75	135,820	316,085	1	1
S2/R008	A	68.75	67.95	B	67.92	127,623	299,708	1	1
S2/R009	B	68.76	68.06	A	68.14	144,007	324,269	1	1
S2/R010	A	69.07	67.70	A	67.82	135,815	307,892	1	1
S3/R011	A	76.60	75.74	B	76.00	119,431	275,118	1	1
S3/R012	A	76.82	76.14	A	75.95	111,239	266,930	1	1
S3/R013	A	76.97	76.30	A	77.08	127,623	299,699	1	0
S3/R014	B	77.07	75.89	A	76.20	119,431	283,308	1	1
S3/R015	A	77.42	76.21	B	75.66	135,815	307,895	1	1
S4/R016	A	71.80	70.38	A	71.11	144,007	332,464	1	1
S4/R017	B	71.90	70.53	A	71.04	152,199	357,049	1	1
S4/R018	A	72.23	72.12	B	70.63	135,815	316,081	0	1
S4/R019	A	72.47	70.58	A	69.86	160,391	373,423	1	0
S4/R020	A	72.73	70.77	A	71.06	144,007	314,038	1	1
Total attempts / joint-correct outputs						2,601,620	5,983,726	19	18

Table 66: Reference-world observations for the local 20-run training cohort. Each setting contains 12 shared world rows for candidates A–D. The same bank scores the outputs of P (paired selection) and I (independent selection) across all five outer runs in that setting. Values and column means are in percentage points. Panels pair settings for presentation; world indices are local to each setting.

	(a) S1				S2			
World	A	B	C	D	A	B	C	D
01	81.89	81.71	81.33	80.97	67.48	67.43	67.06	66.65
02	82.19	82.12	81.69	81.40	68.05	67.76	67.65	67.22
03	81.49	81.22	80.94	80.68	67.73	67.79	67.32	66.99
04	82.30	82.12	81.83	81.51	68.38	68.19	67.90	67.54
05	81.83	81.74	81.29	80.96	67.99	67.91	67.51	67.09
06	81.59	81.46	81.18	80.89	68.50	68.41	68.22	67.77
07	82.11	82.05	81.77	81.43	68.52	68.33	68.04	67.75
08	81.81	81.64	81.41	81.00	67.95	68.00	67.45	67.18
09	82.51	82.54	82.16	81.72	68.27	67.97	67.78	67.41
10	81.70	81.64	81.27	80.89	67.62	67.57	67.20	66.86
11	82.17	82.02	81.81	81.44	68.01	67.85	67.59	67.31
12	82.41	82.30	81.92	81.51	67.50	67.35	66.88	66.63
Mean	82.00	81.88	81.55	81.20	68.00	67.88	67.55	67.20

	(b) S3				S4			
World	A	B	C	D	A	B	C	D
01	76.12	75.99	75.54	75.15	70.89	70.80	70.42	70.17
02	76.05	75.85	75.59	75.23	70.94	70.88	70.51	70.09
03	76.17	76.03	75.76	75.31	70.87	70.72	70.35	70.14
04	76.07	75.94	75.61	75.25	70.52	70.35	70.04	69.82
05	76.31	76.09	75.85	75.43	70.71	70.58	70.23	70.04
06	75.87	75.78	75.58	75.16	71.04	70.85	70.48	70.27
07	75.88	75.77	75.56	75.25	71.11	70.96	70.68	70.23
08	75.95	75.91	75.51	75.17	71.06	70.88	70.59	70.31
09	75.83	75.73	75.34	75.09	71.13	71.04	70.75	70.26
10	75.93	75.82	75.49	75.15	71.48	71.41	71.06	70.58
11	75.69	75.67	75.25	74.97	71.29	71.18	70.87	70.36
12	76.13	75.98	75.52	75.24	70.96	70.91	70.62	70.13
Mean	76.00	75.88	75.55	75.20	71.00	70.88	70.55	70.20

W Application Budget Evaluation

The original application evidence includes a 1,000-pair budget evaluation of PEBR-FR-Paired (J-new) and ε-LUCB-FR-Independent (J-base). PEBR-FR-Paired uses same-world paired observations, empirical-Bernstein racing and adaptive elimination. The comparator samples arms independently using LUCB leader/challenger queries and ε-PAC stopping. Both use the same independent fresh winner-reporting protocol. This budget evaluation uses native-unit targets $(\varepsilon, \eta) = (.05, .10)$, a selection/report split $(.02, .03)$, 210 fresh report observations and 32 audit observations. On an accuracy scale normalized to $[0, 1]$, these tolerances are 5 and 10 percentage points. Define $Q(B)$ as completion probability and $J(B)$ as the probability of a correct joint output, with abstentions counted as zero. The fixed budget grid is

$$\mathcal{B} = \{6000, 8000, 10000, 15000, 20000, 23000, 30000\}.$$

The evaluation targets are $Q \geq .95$, $J \geq .90$, and a paired noninferiority margin of $-.02$ for $J_{\text{new}} - J_{\text{base}}$.

Reference simultaneous mean intervals receive error budget .005. On their coverage event, definite-correct and not-definitely-wrong indicators obey $H^- \leq H^{\text{true}} \leq H^+$. With 922,220 worlds and four bounded candidate utilities, the Hoeffding half-width is 0.001999999852. All completed outputs in the supplied table are definite: the two regular report values are $99/210$ and $111/210$, while six method-specific rows have $128/210$ and exceed the reporting tolerance.

For independent identically distributed outer pairs, allocate the remaining .045 to 154 one-sided Clopper–Pearson bounds: 56 cover both ends of Q, J at 14 method-budget points, and 98 cover two discordance probabilities for all 49 cross-method budget pairs. Thus $\alpha = .045/154$. The union bound allows dependence between budget points and protects selecting a pair of points after inspecting the entire grid.

Table 67: Selected budget points in the application paired records. Intervals use the simultaneous grid error allocation; denominator is 1,000 throughout.

Method	Budget	Completed	Joint correct	Mean attempts
J-new	10,000	987	985	6,375.454
J-base	23,000	988	984	12,739.496

Method	Q interval (%)	J interval (%)
J-new	[96.97,99.60]	[96.69,99.49]
J-base	[97.12,99.65]	[96.55,99.43]

At lower grid points the completion upper bound is below .95: this holds through 8,000 for J-new and 20,000 for J-base. The smallest supported grid budgets are therefore 10,000 and 23,000, giving a threshold-budget ratio of 2.30. This ratio describes the evaluated grid, not the unknown optimum between its points.

At the selected points the paired table is $(n_{11}, n_{10}, n_{01}, n_{00}) = (982, 3, 2, 13)$, giving a difference of .10 percentage points. The simultaneous one-sided bound is

$$J_{\text{new}} - J_{\text{base}} \geq L(p_{10}) - U(p_{01}) = 0.000124494598 - 0.012614923324 = -0.012490428726.$$

It exceeds the $-.02$ noninferiority margin. With reference-inconclusive outputs, the analogous construction uses the conservative indicators H^-_{new} and H^+_{base}.

Completed runs contribute their recorded completion cost T_r, and abstaining runs contribute $B - 242$ under the recorded cost rule. Mean attempts are 6,375.454 and 12,739.496, giving a ratio of 1.998210. This acquisition ratio and the 2.30 threshold-budget ratio measure expenditure and qualifying budget, respectively. Budget-indexed accounting follows complete query trajectories: a budget-invariant policy uses a truncated trajectory, whereas a budget-dependent policy uses its trajectory at that budget. Whole-block cost is the last affordable complete prefix; it equals $B - 242$ when that amount is attainable.

W.1 A THREE-STAGE PHYSICAL COST ACCOUNT

The R0901 stage ledger includes all acquisitions and failed attempts. J-new's incorrect report on this run is included in its total cost. Allocated GPU-hours sum devices times allocation duration; makespan spans the first start to the last finish.

Table 68: R0901: selection, reporting and audit cost. Proposal and reference generation are outside this three-stage account.

Quantity	J-new	J-base	Base/new
Accepted observations	9,792	22,782	—
Attempted evaluations	9,800	22,800	2.326531
Allocated GPU-hours	164.833	341.500	2.071790
Makespan (hours)	42.750	86.750	2.029240

These are single-pair totals. The separate CPU cache-analysis times of 2.4 and 4.1 seconds are not substituted for world-generation costs.

X FINITE-POOL JOINT COMPLEXITY

Let $\mathcal{P}_M$ contain all distributions on $[0,1]^M$, with $M \geq 2$. World vectors are iid; dependence among candidates within a world is unrestricted. Before seeing a world, a procedure requests a nonempty subset $A \subseteq [M]$, observes its coordinates and pays $|A|$. All requested observations are available. Reporting observations are charged under the same rule, so a winner-only report costs one unit per observation.

Write $\mu_i = \mathbb{E}Y(i)$ and $\mu_* = \max_i \mu_i$. A completed output $(\hat{\imath}, \hat{\mu})$ is correct when

$$\mu_* - \mu_{\hat{\imath}} \leq \varepsilon, \qquad |\hat{\mu} - \mu_{\hat{\imath}}| \leq \eta.$$

For every $P \in \mathcal{P}_M$, require completion probability at least $\gamma = .95$ and probability of a completed incorrect output at most δ. Let $\mathcal{C}^*$ be the infimum, over all procedures with these uniform guarantees, of their worst-case expected paid acquisition cost.

Theorem 17 (Finite-pool joint minimax rate)**.** *For $0 < \varepsilon \leq 1/16$, $0 < \eta \leq 1/8$ and $0 < \delta \leq .1$,*

$$\mathcal{C}^*(M, \varepsilon, \eta, \delta; .95) = \Theta\left[\left(\frac{M}{\varepsilon^2} + \frac{1}{\eta^2}\right)\log\frac{1}{\delta}\right].$$

The upper-bound procedure completes with probability one at a deterministic finite budget and has an exactly selection-conditionally unbiased report. The lower bound permits adaptive subsets, stopping and abstention.

The upper bound uses median elimination followed by a frozen independent report. Median elimination supplies the finite-pool selection rate (Even-Dar et al., 2006); making its reporting cost explicit supplies a benchmark for the joint task. The proof below allows correlated candidate coordinates and uses the same paid action space on both sides.

X.1 CONSTRUCTIVE UPPER BOUND

Set $\delta_s = .4\delta$ and $\delta_r = .6\delta$. At round ℓ, use fresh worlds to evaluate every active candidate n_ℓ times, where

$$a_\ell = \frac{\varepsilon}{4}(3/4)^{\ell-1}, \qquad d_\ell = \delta_s 2^{-\ell}, \qquad n_\ell = \left\lceil \frac{2}{a_\ell^2}\log\frac{3}{d_\ell}\right\rceil.$$

Retain the $\lceil N_\ell/2\rceil$ largest round-specific sample means, using a fixed candidate-ID tie rule. Stop when one candidate remains. Freeze this winner and average

$$n_R = \left\lceil \frac{\log(2/\delta_r)}{2\eta^2}\right\rceil$$

new independent observations for its report.

Proof of the upper bound. Condition on the history before a round, so its active set is fixed. If its best mean is m, the best candidate is underestimated below $m - a_\ell/2$ with probability at most $u = \exp(-n_\ell a_\ell^2/2)$. Any candidate with mean below $m - a_\ell$ exceeds this threshold with probability at most u. The expected number of such overestimated bad candidates is at most $N_\ell u$. Markov's inequality bounds the probability that at least $\lceil N_\ell/2 \rceil$ are overestimated by $2u$. Therefore the best remaining mean drops by more than a_ℓ with probability at most $3u \le d_\ell$. This argument uses marginal tails and linearity of expectation, not independence across candidates.

Since $\sum_\ell a_\ell = \varepsilon$ and $\sum_\ell d_\ell = \delta_s$, the final candidate is acceptable with probability at least $1 - \delta_s$. The independent fixed-length report has conditional mean $\mu_{\hat{i}}$ and conditional error probability at most δ_r. A union bound establishes the joint guarantee.

For odd pool sizes, the exact entering size is $N_\ell = \lceil M/2^{\ell-1} \rceil$ for $1 \le \ell \le L = \lceil \log_2 M \rceil$, and $N_\ell \le 2M/2^{\ell-1}$. Summing the geometrically decreasing costs gives

$$C_s \le \frac{64M}{\varepsilon^2}\left[9 \log \frac{3}{\delta_s} + 81 \log 2\right] + 4M.$$

Adding n_R proves the asserted rate. All round sizes are predetermined, so this procedure completes at budget $C_s + n_R$. □

For a small pool, one can compare the two deterministic costs before acquisition and instead choose uniform sampling with $n_S = \lceil 2\varepsilon^{-2} \log(2M/\delta_s) \rceil$ observations per candidate. A simultaneous error at most $\varepsilon/2$ then certifies the empirical leader. Choosing the smaller predetermined cost preserves both the rate and the error allocation.

X.2 Lower bound under the same paid queries

Let $K = \mathrm{kl}(\gamma - \delta, \delta)$. The change-of-measure inequality relates distinguishable outputs to accumulated query information (Kaufmann et al., 2016). We give the two required constructions.

For selection, take independent Bernoulli coordinates with means $(1/2{+}2\varepsilon, 1/2, \ldots, 1/2)$. For each $j \ge 2$, change only its mean to $1/2 + 4\varepsilon$. The event of completing and returning candidate 1 has probability at least $\gamma - \delta$ under the original model and at most δ under the alternative. Consequently

$$\mathbb{E}N_j \ \mathrm{kl}(1/2, 1/2 + 4\varepsilon) \ge K.$$

Here N_j counts paid observations in every stage. Independent coordinates make the information of any queried subset the sum of its purchased coordinate contributions. Summing over j at the same base instance gives

$$L_s = \frac{(M-1)K}{\mathrm{kl}(1/2, 1/2 + 4\varepsilon)}.$$

For reporting, use independent coordinates with all means $1/2$, versus all means $1/2 + 2\eta + h$ for $h > 0$. The two correct-report intervals are disjoint. The event of completing with a report within η of $1/2$ again has probabilities at least $\gamma - \delta$ and at most δ. Each purchased coordinate contributes the same information, so letting $h \downarrow 0$ gives

$$L_r = \frac{K}{\mathrm{kl}(1/2, 1/2 + 2\eta)}.$$

These two hard instances yield $\sup_P \mathbb{E}_P C \ge \max\{L_s, L_r\}$. Using $\max(x, y) \ge (x + y)/2$, $M - 1 \ge M/2$ and $\mathrm{kl}(1/2, 1/2 + x) = -\frac{1}{2}\log(1 - 4x^2)$ proves the lower rate in Theorem 17. In the stated parameter range, $\frac{1}{2}\log(1/\delta) \le K \le \log(1/\delta)$.

The two terms characterize worst-case task complexity. Instance-dependent optimal acquisition additionally depends on the acceptable answer, gap geometry, dependence structure and action costs. That allocation question is distinct from the minimax benchmark.

X.3 Four-candidate execution counts

For $M = 4$, $\varepsilon = .05$, $\eta = .02$ and $(\delta_s, \delta_r) = (.02, .03)$, the median schedule costs 583,168 selection observations. The smaller uniform schedule takes 4,794 observations per candidate and

5,250 for reporting, totaling 24,426. The supplied world tables give selection success counts $(2871, 2371, 2387, 2366)$ and 3,156 report successes for winner A. Its report is 0.601142857 with Hoeffding half-width 0.019999298. The respective selection and reporting hard-instance lower bounds are 81.772042 and 740.185440. These finite-parameter values illustrate the conservative constants of the guaranteed schedule; the minimax theorem concerns matching rates over the model class.

Y Protected Reporting with Unknown Covariance and Shared Events

Information separation must protect the selected coordinate from the *entire* selection transcript. Independent hidden worlds provide this property for arbitrary within-world dependence. If the winner is frozen before generation, R winner-only observations cost R; generating an entire M-candidate hidden bank first costs MR.

Y.1 Two Gaussian worlds with unknown fixed covariance

Independent Gaussian observations permit decompositions even when their common covariance is unknown (Dharamshi et al., 2026). For iid $X_1, X_2 \sim N_M(\mu, V)$, define

$$F = (2X_1 + X_2)/3, \qquad G = -X_1 + 2X_2.$$

Then $\mathbb{E}F = \mathbb{E}G = \mu$, $\mathrm{Var}(F) = 5V/9$, $\mathrm{Var}(G) = 5V$ and $\mathrm{Cov}(F, G) = 0$. Joint Gaussianity gives $F \perp G$ as full vectors, for every fixed V. Thus an F-measurable winner satisfies $\mathbb{E}[G_{\hat{i}} \mid F] = \mu_{\hat{i}}$.

After freezing the winner, use 30 independent G pilot observations. For the supplied variance $s_P^2 = .020$ and pilot error .005, the Gaussian variance upper bound is

$$v_U = \frac{29(.020)}{\chi^2_{.005,29}} = 0.044203446280.$$

With report error .025 and $\eta = .10$, freeze

$$R = \left\lceil \frac{z^2_{.9875} v_U}{.10^2} \right\rceil = 23$$

and use new independent G blocks for the final mean. The reporting error is at most $.005 + .025 = .03$ by the variance event and conditional normal tail bound. The final report is conditionally unbiased given selection and pilot information. If every pilot and report block acquires two complete four-candidate worlds, its cost is $2 \cdot 4(30 + 23) = 424$. Selection cost is additional. Choosing the winner using candidate-specific pilot bounds would instead require simultaneous variance coverage.

Y.2 Why decorrelation is insufficient

Let $X \sim \mathrm{Bernoulli}(1/2)$ and independent $W \sim N(0, 1/4)$. The channels $F = X + W$ and $G = X - W$ have zero covariance, but

$$\mathbb{E}[G \mid F = f] = \frac{2}{1 + e^{-(4f-2)}} - f.$$

At $f = 1$ this expectation is 0.761594156, not the target $1/2$.

Proposition 18 (Single-Bernoulli protection)**.** *Let a parameter-independent randomized mechanism map one $X \sim \mathrm{Bernoulli}(p)$ to (F, G), with integrable G. If $\mathbb{E}_p[G \mid F] = p$ for every $p \in (0, 1)$, the law of F does not depend on p.*

Proof. For any measurable set A, write $f_x(A) = \mathbb{P}(F \in A \mid X = x)$ and $g_x(A) = \mathbb{E}[G\mathbf{1}_{\{F \in A\}} \mid X = x]$. Conditional unbiasedness requires

$$(1 - p)g_0(A) + pg_1(A) = p\{(1 - p)f_0(A) + pf_1(A)\}.$$

The left side is affine in p; equality for all p forces the quadratic coefficient $f_1(A) - f_0(A)$ to vanish. Hence F has the same law for $X = 0$ and $X = 1$ and contains no information about p. □

Y.3 POISSON PROTECTION AND SHARED-EVENT THINNING

For $X \sim \mathrm{Poisson}(\lambda)$, binomial thinning $U \mid X \sim \mathrm{Binomial}(X, 1/2)$ and $V = X - U$ gives independent Poisson halves (Neufeld et al., 2024). The report $2V$ is unbiased for λ. With independent candidate counts, applying the construction to each coordinate gives full-vector protection. The supplied five-world table yields selection estimates $(3.6, 4.4)$ and a selected-B report of 3.6, with 97% exact interval $[1.503293, 7.218524]$.

Candidatewise thinning need not protect a correlated pool: with independent thinning coins and $i \neq j$, $\mathrm{Cov}(U_i, V_j) = \mathrm{Cov}(X_i, X_j)/4$. The shared-count model $X_A = X_B = Z$ with $Z \sim \mathrm{Poisson}(\lambda)$ therefore has $\mathrm{Cov}(U_A, V_B) = \lambda/4$. Marginal $U_i \perp V_i$ is not the full-vector independence required by winner selection.

A stronger construction is available when shared latent events are observed. The next proposition is an analytic extension of the supplied material; it does not add experimental observations.

Proposition 19 (Protection by thinning shared events once)**.** *Let $Z_1, \dots, Z_K$ be observed independent Poisson counts with means θ_k, and let $X = AZ$ for a known fixed matrix A. Independently thin each latent count once: $U_k \mid Z_k \sim \mathrm{Binomial}(Z_k, 1/2)$, $V_k = Z_k - U_k$. Then $F = 2AU$ and $G = 2AV$ are independent full vectors, with $\mathbb{E}F = \mathbb{E}G = A\theta = \mathbb{E}X$. Their coordinates may be correlated within each vector. Every F-measurable winner has a conditionally unbiased G report.*

Proof. Poisson splitting makes the entire vectors U and V independent because the latent counts and their randomizations are independent across k. Deterministic linear maps preserve independence, and $\mathbb{E}U = \mathbb{E}V = \theta/2$. Conditioning on F fixes the selected index but leaves the law of G unchanged. □

The construction assigns each shared event consistently to a channel, rather than independently splitting its appearances in different candidates. It requires the latent event representation or an appropriate joint thinning kernel. It is not inferred from a covariance matrix alone. Accuracy utilities with arbitrary shared noise remain covered by the independent-world reporting route.

Z INSTANCE-DEPENDENT COMPLETION OF THE FINITE-POOL CONTROLLER

The forced full-active-set queries of the report-aware controller provide a quantitative completion guarantee. The bound depends on the variance of each acquired difference and on the distance of a candidate from the acceptance boundary. It applies to the implemented acquisition policy, including its continue-versus-report decisions.

Let fresh potential world vectors be iid, with every coordinate in the common interval $[a_0, a_0 + w]$, where $w > 0$. Within-world dependence is unrestricted. Query subsets are chosen before their worlds are opened. All requested observations are available, and observing candidate i costs a deterministic $c_i > 0$. Write

$$\mu_* = \max_i \mu_i, \qquad \Delta_i = \mu_* - \mu_i, \qquad s_{ij}^2 = \mathrm{Var}(Y(i) - Y(j)), \qquad K = \binom{M}{2}.$$

The controller uses blocks of $b \geq 1$ worlds, a full-active-set query at least every $f \geq 1$ query rounds, and the empirical-Bernstein radius

$$r_{ij,n} = \sqrt{\frac{2\widehat{s}_{ij,n}^2 z_n}{n}} + \frac{14 w z_n}{3(n-1)}, \qquad z_n = \log \frac{4Kn(n+1)}{\delta_s}, \qquad n \geq 2. \tag{Z.1}$$

Intervals are intersected with $[-w, w]$. Elimination removes i when an active challenger satisfies $L(j, i) > \varepsilon$, and certification of i requires $U(j, i) \leq \varepsilon$ for every active j. Updates are simultaneous. The final report uses the implemented frozen count

$$m_H = \left\lceil \frac{w^2 \log(2/\delta_r)}{2\eta^2} \right\rceil, \qquad m_R(i) = \min \left\{ m_H, \left\lceil \frac{[2\overline{v}_i + (2/3)w\eta] \log(2/\delta_r)}{\eta^2} \right\rceil \right\}, \tag{Z.2}$$

with a minimum of one observation, where the simultaneous marginal-variance upper bounds have error at most δ_v. A fixed independent audit uses $a \geq 0$ additional winner observations. Its values do not affect the output.

For completeness, the marginal variance bound uses disjoint pairs from candidate i's predictably acquired iid stream. At $k \geq 1$ pairs put

$$W_{i,t} = (Y_{i,2t} - Y_{i,2t-1})^2/2, \qquad \overline{v}_{i,k} = \min\left\{\frac{w^2}{4}, \overline{W}_{i,k} + \frac{w^2}{2}\sqrt{\frac{\log(Mk(k+1)/\delta_v)}{2k}}\right\}.$$

Before the first pair, use $w^2/4$. Hoeffding's inequality, summed over candidate and prefix, gives the stated simultaneous variance guarantee.

The report-aware policy assigns each active i the reserve $R_i = c_i(m_R(i) + a)$ and a nonnegative plug-in estimate $\widehat{C}_i$ of the remaining comparison cost. For an uncertified pair, this estimate holds the observed difference mean and variance fixed and projects the additional pair count needed for Eq. (Z.1) to place its upper bound below ε. It sums these counts times $c_i + c_j$; a nonpositive projected margin gives infinite cost. This projection chooses queries, whereas certification always uses the actual interval. The target minimizes $R_i + \widehat{C}_i$, with candidate-ID ties. Among affordable certified candidates, the immediate choice minimizes R_i. The `joint_lookahead` policy defers that report only when an uncertified target has a strictly smaller projected total cost. Outside forced full-set rounds it queries the target with the challenger of largest advantage upper bound, in a fresh whole block. With at most two active candidates it queries the full set. Every query preserves at least the smallest active reserve. If a proposed continuation would violate the cap while a certificate is affordable, it reports that certificate. The `report_aware_immediate` policy omits deferral; `paired_leader` reports when the empirical leader is certified and affordable. These rules specify the controller analyzed below.

Z.1 A VARIANCE-SENSITIVE THRESHOLD FOR THE IMPLEMENTED BOUNDARY

Fix an analysis failure probability $\kappa \in (0, 1)$; this parameter does not alter the controller or its certificate error allocations. Put

$$d = \min\{\delta_s, \kappa\}, \qquad \ell_n = \log\frac{4Kn(n+1)}{d}, \qquad A(v, g) = \max\left\{\frac{128v}{g^2}, \frac{144w}{g}\right\}, \tag{Z.3}$$

and, for $g > 0$, define the integer threshold

$$N(v, g) = \left\lceil \max\left\{2, 4A(v, g)\log\left(e + 2A(v, g)\sqrt{\frac{8K}{d}}\right)\right\}\right\rceil. \tag{Z.4}$$

The constants give an explicit sufficient threshold; the exact implemented radius can yield substantially earlier stopping.

Lemma 20 (Pairwise radius envelope)**.** *With probability at least $1 - \kappa$, simultaneously for every pair and every acquired prefix of length $n \geq 2$,*

$$r_{ij,n} \leq \sqrt{\frac{2s_{ij}^2\ell_n}{n}} + \frac{9w\ell_n}{n-1}. \tag{Z.5}$$

For every $n \geq N(s_{ij}^2, g)$ the right side is at most $g/4$.

Proof. Predictable acquisition produces an iid extended stream for each pair: after the last actual query, append unused independent draws from its target law. Every observed pair prefix is a prefix of this stream. The sample-standard-deviation bound for variables in an interval of width $2w$ gives

$$\Pr\left\{\widehat{s}_{ij,n} > s_{ij} + 2w\sqrt{\frac{2\log(1/\beta_n)}{n-1}}\right\} \leq \beta_n, \qquad \beta_n = \frac{\kappa}{Kn(n+1)}.$$

This is the scaled upper-tail variance inequality of Maurer & Pontil (2009), Theorem 10. Summing over pairs and all $n \geq 2$ costs at most κ. On the resulting event, use $z_n \leq \ell_n$ and $\log(1/\beta_n) \leq \ell_n$ in Eq. (Z.1). The added square-root term is at most $4w\ell_n/\sqrt{n(n-1)}$. Since $4 + 14/3 = 26/3 < 9$, this proves Eq. (Z.5).

For the explicit threshold, abbreviate $A = A(v, g)$, $C = 8K/d$, and $L = \log(e + 2A\sqrt{C}) \geq 1$. For $n \geq N(v, g)$ let $t = n/(2A)$, so $t \geq 2L$. The function $t - \log t$ is increasing for $t \geq 1$, and

$$t - \log t \geq 2L - \log(2L) \geq L \geq \log(2A\sqrt{C}).$$

It follows that $n \geq A \log(Cn^2) \geq A\ell_n$. Thus the square-root term is at most $g/8$. Because $n - 1 \geq n/2$ and $A \geq 144w/g$, the linear term is also at most $g/8$. □

Z.2 Completion under the actual continue-versus-report policy

Assume $0 < \varepsilon < w$ and no candidate lies exactly on the acceptance boundary. Partition the pool into

$$\mathcal{G} = \{i : \Delta_i < \varepsilon\}, \qquad \mathcal{B} = \{i : \Delta_i > \varepsilon\}.$$

Fix a maximizer i_* for analysis. Each unacceptable candidate has a sufficient elimination horizon

$$h_i = \left\lceil \frac{N(s^2_{i_* i}, \Delta_i - \varepsilon)}{b} \right\rceil \quad (i \in \mathcal{B}). \tag{Z.6}$$

For candidates in $\mathcal{G}$, every pair has strictly positive mutual certification margin $\varepsilon - |\mu_i - \mu_j|$. Let

$$H = \max \left\{ 1,\ \max_{i \in \mathcal{B}} h_i,\ \max_{\substack{i<j \\ i,j \in \mathcal{G}}} \left\lceil \frac{N(s^2_{ij}, \varepsilon - |\mu_i - \mu_j|)}{b} \right\rceil \right\}, \tag{Z.7}$$

where maxima over empty collections are zero. Define

$$C^*_{\mathrm{sel}} = fb \left(\sum_{i \in \mathcal{B}} c_i h_i + H \sum_{i \in \mathcal{G}} c_i \right),$$
$$R_{\max} = (m_H + a) \max_i c_i, \qquad B_* = C^*_{\mathrm{sel}} + R_{\max}. \tag{Z.8}$$

Theorem 21 (Finite-pool instance-dependent completion)**.** *Run the implemented* `joint_lookahead` *policy with budget $B \geq B_*$, at least $fH + 1$ decision rounds, and sufficient selection, reporting, and audit banks. Under the observation and cost assumptions above,*

$$\Pr\{\textit{joint output completed with cost at most } B_*\} \geq 1 - \delta_s - \kappa, \tag{Z.9}$$
$$\Pr\{\textit{incorrect joint certificate}\} \leq \delta_s + \delta_v + \delta_r, \tag{Z.10}$$
$$\Pr\{\textit{correct joint output}\} \geq 1 - \delta_s - \kappa - \delta_v - \delta_r. \tag{Z.11}$$

The report is unbiased conditional on its pre-report transcript. The same completion result holds for `paired_leader` *and* `report_aware_immediate` *under their shared inference rules.*

Proof. Intersect the simultaneous comparison-coverage event with the radius envelope event of Lemma 20. Their joint probability is at least $1 - \delta_s - \kappa$. Every $i \in \mathcal{G}$ survives: a valid lower bound on any challenger's advantage cannot exceed $\Delta_i < \varepsilon$. In particular i_* survives all simultaneous updates.

If an unacceptable i remains through h_i forced full-set queries, its pair with i_* has at least bh_i observations. Hence

$$L(i_*, i) \geq \Delta_i - 2r_{i_* i} \geq \Delta_i - \frac{\Delta_i - \varepsilon}{2} > \varepsilon,$$

so it is removed at the next decision. The factor two accounts for both the confidence half-width and the sample-mean error on the coverage event. It can be queried in at most fh_i blocks, since the first forced full-set query occurs in round zero and subsequent ones occur at most f rounds apart.

By the decision following H full-set queries all unacceptable candidates are removed. For any surviving $i, j \in \mathcal{G}$,

$$U(j, i) \leq \mu_j - \mu_i + 2r_{ij} \leq \mu_j - \mu_i + \frac{\varepsilon - |\mu_i - \mu_j|}{2} < \varepsilon.$$

Thus every surviving candidate is certified. The lookahead rule can defer only toward an uncertified target; at this decision no such target remains. It therefore reports. The leader-based and immediate policies also report.

The weighted query count is at most C^*_{sel}, and every report-plus-audit reserve is at most $R_{\max}$. Before this terminal decision, the cost of the next whole query is included in the same bound. Therefore $B \geq B_*$ prevents a budget interruption on the stated event. The decision-round and bank assumptions prevent the remaining non-statistical interruptions. This proves Eq. (Z.9).

On the comparison event, every certified output is ε-optimal because an optimum survives. On the simultaneous marginal-variance event, the selected fresh report has conditional error at most δ_r by Hoeffding or Bernstein according to the frozen choice in Eq. (Z.2). Union bounds give Eqs. (Z.10)–(Z.11). Conditioning on the pre-report transcript fixes the winner and positive report count; fresh reporting draws retain their target mean. This proves conditional unbiasedness. □

The dependence

$$N(s^2, g) = O\left[\left(\frac{s^2}{g^2} + \frac{w}{g}\right) \log\left(e + \frac{K}{d}\left(\frac{s^2}{g^2} + \frac{w}{g}\right)^2\right)\right]$$

makes the pairing effect explicit. Shared noise enters through the variance of a difference, whereas the report reserve uses marginal variance or the bounded-range fallback. The certificate-error budget and the analysis probability κ play different roles: κ controls the cost envelope, not whether a displayed interval is valid.

Z.3 A STORED-CERTIFICATE GUARD WITH BOUNDARY-INDEPENDENT COMPLETION

A bounded-deferral guard can retain report-aware exploration while removing the acceptance-boundary condition. When any candidate is first certified, store that candidate, its certificate, and its frozen fresh-report count. Reserve their reporting and audit cost. Permit at most $D \geq 0$ additional whole query blocks while preserving this reserve. During those blocks, allow an earlier terminal choice of another certified candidate; otherwise report the stored candidate when the block allowance is exhausted or the next query would invade its reserve. This guard specifies an additional policy rule.

Let

$$H_* = \max\left\{1, \max_{j \neq i_*}\left\lceil \frac{N(s^2_{i_* j}, \varepsilon + \Delta_j)}{b} \right\rceil\right\}. \tag{Z.12}$$

Proposition 22 (Bounded-deferral completion)**.** *For $0 < \varepsilon < w$, the guarded controller needs no exclusion of $\Delta_i = \varepsilon$. If all banks and decision rounds are sufficient, then a budget*

$$B \geq b(fH_* + D)\sum_i c_i + R_{\max} \tag{Z.13}$$

gives completion with probability at least $1 - \delta_s - \kappa$ and the same joint-certificate error and conditional-unbiasedness guarantees as Theorem 21.

Proof. On the same coverage and radius events, an optimum i_* is never eliminated. After H_* full-set queries, each remaining challenger has

$$U(j, i_*) \leq -\Delta_j + 2r_{i_* j} \leq -\Delta_j + \frac{\varepsilon + \Delta_j}{2} \leq \varepsilon.$$

Consequently some candidate is certified by this decision, irrespective of acceptance-boundary ties. The pre-certificate query cost is at most $bfH_* \sum_i c_i$. At most D further blocks cost at most $bD\sum_i c_i$, and the stored final reserve is at most $R_{\max}$. Equation (Z.13) therefore permits all required stages. The stored certificate remains a valid statement about fixed means on the simultaneous coverage event, even if subsequent intervals change. The stored report count is fixed before any final reporting value is opened. The preceding validity and conditional-expectation argument applies to either terminal candidate. □

Z.4 Conditional cost of continuing a certified race

An available certificate creates a choice between reporting now and acquiring information about a cheaper acceptable answer. Freeze the complete history $\mathcal{T}_0$ at such a choice, with paid prefix cost C_0 and certified candidate i_0. Let R_0 be its precommitted report and audit cost. Both branches start from this same history. The immediate branch pays $C_{\mathrm{I}} = C_0 + R_0$. A continuation branch acquires additional information at cost K, then reports a certified candidate I at cost R_I, retaining i_0 as its fallback. Costs use the same nonnegative acquisition units and stage definitions in both branches.

Proposition 23 (Exact continuation value)**.** *With integrable costs, the pathwise and conditional-mean differences are*

$$C_{\mathrm{G}} - C_{\mathrm{I}} = K - (R_0 - R_I), \tag{Z.14}$$

$$\mathbb{E}[C_{\mathrm{G}} - C_{\mathrm{I}} \mid \mathcal{T}_0] = \mathbb{E}[K \mid \mathcal{T}_0] - \mathbb{E}[R_0 - R_I \mid \mathcal{T}_0]. \tag{Z.15}$$

If the branch either retains $A = i_0$ or switches to B, the two report costs are fixed by $\mathcal{T}_0$, and $p = \mathbb{P}(I = B \mid \mathcal{T}_0)$, it improves conditional expected cost exactly when $\mathbb{E}[K \mid \mathcal{T}_0] < p(R_A - R_B)$.

Proof. Subtract $C_0 + R_0$ from $C_0 + K + R_I$ and take conditional expectations. For the two-answer branch, $R_0 - R_I = (R_A - R_B)\mathbf{1}\{I = B\}$. □

For a specified Gaussian state, the comparison is explicit. Consider a retained pair with gap $\Delta = \mu_A - \mu_B = .005$, marginal standard deviations .200 and .020, and correlation .5, giving $v_{AB} = .0364$. Condition on $n_0 = 1024$ paired observations with $\overline{D}_{n_0} = .005$ and a stored certificate for A. Retain B through $m = 3072$ additional pairs, so $N = 4096$; assume its other certificate comparisons are already discharged. For $\varepsilon = .020$ and radius $r_N = .0121435811$, switch to B when $\overline{D}_N + r_N \le \varepsilon$. Independence of future worlds gives

$$p = \Phi\left(\frac{N(\varepsilon - r_N) - n_0\overline{D}_{n_0} - m\Delta}{\sqrt{mv_{AB}}}\right) = .8657279. \tag{Z.16}$$

Here the known-variance checkpoint radius uses $\alpha_N = .02/(6 \cdot 8 \cdot 9)$ and $r_N = \Phi^{-1}(1 - \alpha_N/2)\sqrt{v_{AB}/N}$. At $\eta = .002$ and reporting error .03, the report counts are 47093 and 471; equal audit charges cancel. The 6144 added acquisitions therefore yield conditional expected saving $p(47093 - 471) - 6144 = 34217.97$. The gap enters this evaluation of the continuation rule, not its information set.

The stored-certificate controller also admits an economic guard. Fix $\beta \ge 0$ at $\mathcal{T}_0$. Authorize a new whole-block query only if its cost keeps cumulative continuation spending below βR_0 and leaves the fallback reserve R_0 payable under the hard cap. Switch only to a certified candidate with $R_I \le R_0$; otherwise report i_0. Every added pilot, query, or other charged acquisition belongs to K. The guard uses a finite continuation schedule and freezes the final report count before opening its fresh observations.

Proposition 24 (Economic-guard overhead and validity)**.** *For deterministic acquisition charges and available scheduled observations, the economic guard completes whenever the immediate branch is payable and satisfies*

$$C_{\mathrm{G}} \le C_0 + (1 + \beta)R_0 \le (1 + \beta)C_{\mathrm{I}} \tag{Z.17}$$

on every path. Over the full execution law, if all stored and new selection certificates share a simultaneous error budget δ_s, and the chosen fresh-report rule has conditional error at most δ_r, its joint incorrect-certificate probability is at most $\delta_s + \delta_r$. The final fresh mean remains unbiased conditional on the complete pre-report transcript.

Proof. The restrictions give $K \le \beta R_0$ and $R_I \le R_0$; substitute them in $C_0 + K + R_I$, using $C_0 \ge 0$. The protected reserve and finite schedule make the fallback payable at termination. On the simultaneous selection event, every retained certificate is valid for the fixed candidate means, including the fallback. Conditioning on the pre-report transcript fixes the reported candidate and sample count; independent fresh observations then give the stated mean and error probability. A union bound combines the selection and reporting errors. □

For example, $\beta = .15$ with $R_0 = 47093 + 8 = 47101$ permits up to 7065.15 units of additional acquisition, covering the 6144-unit continuation above while bounding relative overhead by 15%.

Z.5 AN ANSWER-AWARE INFORMATION LOWER BOUND

The paid action space also gives an instance-dependent information bound. Let P be an iid law on bounded candidate-world vectors. A predictable fresh subset action S has law P_S and cost $c(S) = \sum_{i \in S} c_i$. Singleton pilot and report observations use the same paid actions. A procedure is uniformly (γ, δ)-valid if, for every model in the declared class, it completes with probability at least γ and issues an incorrect joint certificate with probability at most δ. Let $I(P) = \{i : \mu_*(P) - \mu_i(P) \le \varepsilon\}$. For $i \in I(P)$, define the event and answer-specific alternatives

$$\begin{aligned} E_i(P) &= \{\text{completed output } i, |\widehat{\mu} - \mu_i(P)| \le \eta\}, \\ \mathcal{A}_i(P) &= \{Q : i \notin I(Q) \text{ or } |\mu_i(Q) - \mu_i(P)| > 2\eta\}. \end{aligned} \tag{Z.18}$$

Write $d_\delta(p) = \mathrm{kl}(p, \delta)$ for $p > \delta$ and zero otherwise.

Proposition 25 (Paid-action information relaxation)**.** *Under a finite hard cap, or the likelihood integrability conditions for the stopped change-of-measure identity, every uniformly valid procedure has expected cost at least*

$$\begin{aligned} &\inf_{\substack{n_S \ge 0,\ p_i \ge 0 \\ \gamma - \delta \le \sum_{i \in I(P)} p_i \le 1}} \sum_{S \ne \varnothing} c(S) n_S, \\ &\textit{subject to} \quad \sum_{S \ne \varnothing} n_S \, \mathrm{KL}(P_S \| Q_S) \ge d_\delta(p_i) \quad \forall i \in I(P),\ Q \in \mathcal{A}_i(P). \end{aligned} \tag{Z.19}$$

Proof. Set $p_i = P(E_i(P))$ and let N_S count all acquired action-S worlds. Correct completed outputs have probability at least $\gamma - \delta$, so these disjoint events satisfy the probability constraints. For each alternative $Q \in \mathcal{A}_i(P)$ the event $E_i(P)$ implies an incorrect joint certificate under Q: either the answer is unacceptable or its accurate-report interval is disjoint from that under P. Therefore $Q(E_i(P)) \le \delta$. The stopped likelihood-ratio chain rule and binary data processing give

$$\sum_S \mathbb{E}_P[N_S] \, \mathrm{KL}(P_S \| Q_S) \ge \mathrm{kl}(P(E_i(P)), Q(E_i(P))) \ge d_\delta(p_i).$$

The last inequality uses monotonicity in the second argument below p_i when $p_i > \delta$, and nonnegativity otherwise. Actual expected counts and event probabilities are consequently a feasible point of Eq. (Z.19); its objective is their paid cost. □

For a unique acceptable answer, its probability is at least $\gamma - \delta$, recovering the usual single-answer KL constraint. With several acceptable answers, the probabilities must remain answer-specific. Equation (Z.19) is a lower-bound relaxation, not a claim that every multiple-answer instance attains this value. In a Gaussian raw-subset model with fixed known positive-definite covariance V, its action information is $\frac{1}{2}(\mu_S - \nu_S)^\top V_{SS}^{-1}(\mu_S - \nu_S)$. Protected-channel observations instead use their own observation kernel.

Corollary 26 (Gaussian selection and reporting information)**.** *Consider the full Gaussian location family $\{\mathcal{N}(\mu, \Sigma) : \mu \in \mathbb{R}^M\}$ with fixed known $\Sigma \succ 0$ and fresh raw-subset actions. For expected paid action counts $n_S = \mathbb{E}_\mu N_S$, let E_S extract the coordinates in S and set*

$$J(n) = \sum_{S \ne \varnothing} n_S E_S^\top \Sigma_{SS}^{-1} E_S. \tag{Z.20}$$

Assume $J(n) \succ 0$. A boundary answer with $\mu_* - \mu_i = \varepsilon$ must satisfy $p_i \le \delta$. For a strictly ε-acceptable answer i, write $g_{ij} = \varepsilon + \mu_i - \mu_j > 0$ and $a_{ji} = e_j - e_i$. The constraints indexed by i in the stopped-information argument of Proposition 25 reduce, when $d_\delta(p_i) > 0$, to

$$\begin{aligned} a_{ji}^\top J(n)^{-1} a_{ji} &\le \frac{g_{ij}^2}{2 d_\delta(p_i)} \quad (j \ne i), \\ e_i^\top J(n)^{-1} e_i &\le \frac{2\eta^2}{d_\delta(p_i)}. \end{aligned} \tag{Z.21}$$

Here $p_i = \mathbb{P}_\mu(E_i(\mu))$ retains its answer-specific meaning, with $p_i \ge 0$ and $\gamma - \delta \le \sum_{i \in I(\mu)} p_i \le 1$. Answers with $d_\delta(p_i) = 0$ impose no corresponding information constraint. Minimizing $\sum_S c(S) n_S$ under these conditions gives the positive-definite specialization of the paid-action relaxation.

Proof. For $h = \nu - \mu$, the accumulated Gaussian KL divergence is $\frac{1}{2}h^\top J(n)h$. Weighted Cauchy–Schwarz gives $h^\top Jh \geq (a^\top h)^2/(a^\top J^{-1}a)$, approached at the half-space boundary along $J^{-1}a$. Thus the infimum over $\nu_j - \nu_i > \varepsilon$ is $g_{ij}^2/(2a_{ji}^\top J^{-1}a_{ji})$; over $|\nu_i - \mu_i| > 2\eta$ it is $2\eta^2/(e_i^\top J^{-1}e_i)$. The full location family permits these perturbations. Each alternative must supply at least $d_\delta(p_i)$ information, yielding Eq. (Z.21); the probability constraints follow from the same disjoint correct-output events as in Proposition 25. For a boundary answer, an arbitrarily small mean perturbation makes it unacceptable, giving zero infimum KL and hence $d_\delta(p_i) = 0$. □

The contrast directions a_{ji} quantify evidence for acceptability, whereas e_i quantifies evidence for the selected mean. Their shared information matrix connects the two requirements without identifying comparison precision with reporting precision.

Z.6 Variance-adaptive fresh reporting

After the selection transcript $\mathcal{T}_0$ freezes candidate i, its independent fresh outcomes lie in $[a, a+R]$, with mean μ_i and unknown variance v_i. Write $x = \log(1/\delta_v)$, $z = \log(2/\delta_r)$, and $B_Z = R^2/2$. Purchase

$$m = \left\lceil \frac{Rx}{2\eta} \right\rceil, \qquad Z_k = \frac{(Y_{2k}^P - Y_{2k-1}^P)^2}{2}, \qquad \overline{v} = \min\left\{ \frac{R^2}{4}, 2\overline{Z}_m + \frac{2B_Z x}{m} \right\}. \tag{Z.22}$$

All $2m$ pilot acquisitions are charged. Freeze a separate fresh-report count

$$n_R = \min\left\{ \left\lceil \frac{R^2 z}{2\eta^2} \right\rceil, \left\lceil \frac{[2\overline{v} + (2/3)R\eta]z}{\eta^2} \right\rceil \right\}. \tag{Z.23}$$

The final report is the mean of exactly n_R independent new winner observations, with interval radius η.

Theorem 27 (Unknown-variance reporting with instance-dependent cost)**.** *Conditional on $\mathcal{T}_0$, the variance bound in Eq.* (Z.22) *fails with probability at most δ_v. The final report is unbiased conditional on the complete pre-report transcript, and its error beyond η has probability at most $\delta_v + \delta_r$. If selection error is at most δ_s, joint incorrect certification is at most $\delta_s + \delta_v + \delta_r$. Its expected additional winner observations satisfy*

$$\mathbb{E}[2m + n_R \mid \mathcal{T}_0] \leq \frac{4v_i z}{\eta^2} + \frac{Rx + (14/3)Rz}{\eta} + 3. \tag{Z.24}$$

For deterministic winner cost c_i, multiply the right side by c_i. Selection and independent audit costs are additional.

Proof. The independent pilot variables satisfy $0 \leq Z_k \leq B_Z$ and $\mathbb{E}[Z_k \mid \mathcal{T}_0] = v_i$. For a nonnegative $Z \leq B_Z$, $\mathbb{E}Z^2 \leq B_Z v_i$ and $e^{-u} \leq 1 - u + u^2/2$ imply

$$\log \mathbb{E}e^{-\lambda(Z - v_i)} \leq \lambda^2 B_Z v_i/2, \qquad \lambda \geq 0.$$

Chernoff optimization gives

$$\Pr\left\{ v_i - \overline{Z}_m > \sqrt{2B_Z v_i x/m} \mid \mathcal{T}_0 \right\} \leq e^{-x}.$$

The zero-variance case is immediate. On the complementary event, $\sqrt{2B_Z v_i x/m} \leq v_i/2 + B_Z x/m$, so $v_i \leq 2\overline{Z}_m + 2B_Z x/m$. The deterministic variance bound $v_i \leq R^2/4$ proves the stated upper-confidence property.

The pre-report transcript fixes the candidate, pilot, count, completion decision, and chosen branch. Fresh final observations retain their mean, giving conditional unbiasedness. On the variance-coverage event, Hoeffding or bounded Bernstein gives error at most δ_r. A union bound proves the certificate assertions.

For cost, $\mathbb{E}[\overline{v} \mid \mathcal{T}_0] \leq 2v_i + 2B_Z x/m$. The Bernstein branch bounds the minimum in Eq. (Z.23), and $\lceil u \rceil \leq u + 1$, giving

$$\mathbb{E}[2m + n_R \mid \mathcal{T}_0] \leq 2m + 1 + \frac{[4v_i + 4B_Z x/m + (2/3)R\eta]z}{\eta^2}.$$

Use $B_Z x/m \leq R\eta$ and $2m \leq Rx/\eta + 2$ to obtain Eq. (Z.24). □

A hard cap reserves $2m + n_H$, where $n_H = \lceil R^2 z/(2\eta^2) \rceil$; completion is decided without final-report values. A frozen winner needs no candidate-wise pilot correction. This variance-learning extension charges pilot and report acquisitions separately. Application tables retain their recorded fixed reporting protocols.

Corollary 28 (A matching deterministic-instance reporting scale)**.** *In the iid $[0,1]^2$ fresh-subset model with unit candidate costs, consider two deterministic candidates with means $(0.6, 0.4)$ and $\varepsilon = 0.1$. Fix $\gamma > 2\delta > 0$. For $0 < \eta \le 1/12$, any uniformly (γ, δ)-valid joint procedure has expected cost at least $\mathrm{kl}(\gamma - \delta, \delta)/(12\eta)$ on this instance. A fixed valid selection stage followed by Theorem 27 has cost $O(\eta^{-1})$ on the same instance, with fixed error allocations.*

Proof. Keep candidate 2 at 0.4; let candidate 1 equal 0.6 with probability $1-6\eta$ and 1 with probability 6η. It remains uniquely acceptable, but its mean shifts by $2.4\eta > 2\eta$, separating accurate-report intervals. A completed accurate candidate-1 report therefore has probabilities at least $\gamma - \delta$ at the original instance and at most δ at the alternative. Fresh actions containing candidate 1 contribute KL $-\log(1-6\eta)$; other actions contribute zero. Stopped change of measure gives

$$\mathbb{E}[N_1] \ge \frac{\mathrm{kl}(\gamma - \delta, \delta)}{-\log(1-6\eta)} \ge \frac{\mathrm{kl}(\gamma - \delta, \delta)}{12\eta}.$$

Here N_1 counts all purchased candidate-1 observations and $-\log(1-u) \le 2u$ for $u \le 1/2$. A uniformly valid fixed-budget bounded selection stage has cost independent of η at fixed $\varepsilon = 0.1$. Its selected candidate has zero variance at the deterministic instance, so Eq. (Z.24) supplies the matching $O(\eta^{-1})$ upper bound. □

Z.7 Report-aware choice beyond contrast information

Consider the reporting subproblem after two candidates have been certified to satisfy the selection tolerance. One observation costs one candidate–world evaluation. In two Gaussian instances the means are $(82, 81.95)$, the selection tolerance is $\varepsilon = 0.20$, and the covariance matrices are

$$V_1 = \mathrm{diag}(25, 1), \qquad V_2 = \mathrm{diag}(1, 25).$$

The means specify evaluation instances, not values revealed to the procedure. In both instances each candidate is acceptable, and every fresh contrast has the identical distribution $Y_A - Y_B \sim N(0.05, 26)$. The following comparison isolates the additional reporting cost; the earlier cost and possible error of establishing acceptability are separate.

Fix $\eta = 0.20$ and $\delta_r = 0.03$. The common Gaussian reporting service, invoked only after the winner is frozen, purchases

$$n(v) = \left\lceil z^2_{1-\delta_r/2} v/\eta^2 \right\rceil$$

independent new observations and reports their mean. Its known variance is not exposed to the winner decision. Here $n_H = n(25) = 2944$ and $n_L = n(1) = 118$. Both winner rules may know the unordered variance set $\{1, 25\}$ but not its assignment. A *contrast-blind winner rule* always returns a candidate and may use arbitrarily many independent paired contrasts, predictable contrast-based stopping, and independent internal randomization, but no raw marginal values, variance labels, or variance-dependent side channel before freezing its winner.

Proposition 29 (Contrast-blind reporting separation)**.** *For every contrast-blind winner rule using this fixed reporting service,*

$$\max_{k=1,2} \mathbb{E}_{P_k}[C_{\mathrm{report}}] \ge \frac{n_H + n_L}{2} = 1531.$$

The bound is attained by choosing each candidate with probability $1/2$. A variance-assignment oracle pays 118, giving a report-only ratio $1531/118 = 12.9746$.

Proof. The full contrast transcript, including contrast-based stopping and internal randomization, has the same distribution under P_1 and P_2. Consequently the probability p of choosing A is common to both instances. The respective expected reporting costs are $pn_H + (1-p)n_L$ and $pn_L + (1-p)n_H$. Their average is $(n_H + n_L)/2$, so at least one is no smaller. Choosing $p = 1/2$ attains equality. The oracle selects the variance-one candidate in each instance. □

The comparison class is defined by the winner's available information. Retaining marginal observations permits variance learning, as the following construction shows.

Proposition 30 (A fully charged variance-aware choice)**.** *Suppose the unordered variance set* $\{1, 25\}$ *is known but its assignment is not. Buy 30 independent raw pilot observations from each candidate, using worlds disjoint from the earlier selection data and the final reporting data. Choose the candidate with smaller unbiased sample variance, breaking a tie deterministically. With the same fixed reporting service, this rule has expected pilot-plus-report cost*

$$60 + 118 + (2944 - 118)q = 178.000000000207,$$
$$q = \Pr\{F_{29,29} < 1/25\} = 7.3131004450 \times 10^{-14},$$

in either instance, and report error at most 0.03*. Moreover, without a known-variance service, the rule may instead buy exactly 118 new observations of its chosen candidate. This procedure has deterministic pilot-plus-report cost* 178*, a conditionally unbiased report, and unconditional report error strictly below* 0.03 *in both instances.*

Proof. Gaussian sample variances satisfy $29S_i^2/v_i \sim \chi^2_{29}$. Independent pilot streams make these two chi-square variables independent. The high-variance candidate is selected exactly when $25U/V < 1$, where U, V are independent $\chi^2_{29}/29$ variables, so the error probability is q as displayed. The same-service expected cost follows by conditioning on this event; the service has error at most 0.03 for either selected variance.

For the version without the service, set

$$\beta_L = 2\Phi(-0.20\sqrt{118}) = 0.0298137431406,$$
$$\beta_H = 2\Phi(-0.20\sqrt{118/25}) = 0.663917249216.$$

The new reporting stream is independent of all pilot values. Hence the exact report-error probability is

$$(1 - q)\beta_L + q\beta_H = 0.029813743140655 < 0.03.$$

The pre-report transcript fixes the candidate and sample count; conditional expectation of the final sample mean is therefore the selected candidate's mean. All 60 pilot observations and all 118 report observations are charged. □

The paid comparison changes the worst-instance ratio from the oracle-only 12.9746 to $1531/(178 + 2826q) = 8.6011$ against the same-service class. This is an analytic achievable construction for the specified Gaussian reporting subproblem. It does not assume that the earlier selection data, the pilot data, and the final report are three independent views of the same raw draws: they are separately purchased streams. If acceptability was itself certified with error at most δ_s, a union bound gives joint error at most $\delta_s + 0.03$, with that selection cost added to both procedures. An independent audit is additional. For the fixed-118 version, unbiasedness holds conditional on the pre-report transcript, whereas the error guarantee averages over pilot outcomes. Conditional error on the rare high-variance selection event is β_H, not 0.03. The same-service version instead retains conditional error control on every pilot transcript.

Z.8 Exact cost of the bounded variance pilot

The variance-adaptive reporting rule of Theorem 27 can be evaluated without Monte Carlo error on a two-point outcome law. Let $Y = 50 \pm s$ with equal probabilities, $0 < s \le 50$. The algorithm knows the common range $[0, 100]$, not the smaller support. Set $\eta = 1$, $\delta_v = 0.005$, $\delta_r = 0.025$, $x = \log 200$, and $z = \log 80$. There are $m = 265$ pilot pairs, costing 530 observations. If K denotes the number of unequal pilot pairs, then

$$K \sim \mathrm{Binomial}(265, 1/2), \quad \overline{v}(K) = \min\left\{2500, \frac{4s^2K}{265} + \frac{10000x}{265}\right\},$$

and the fresh report count is

$$n_s(K) = \min\left\{21911, \lceil [2\overline{v}(K) + 200/3]z \rceil\right\}.$$

Thus the exact expected additional observation count is the finite sum

$$\mathbb{E}C_s = 530 + 2^{-265} \sum_{k=0}^{265} \binom{265}{k} n_s(k). \tag{Z.25}$$

An independent fixed Hoeffding report assigning its entire error budget 0.03 to the report costs 20999 observations. Table 69 evaluates the finite sum and the theorem's bound. These distributional calculations have no Monte Carlo error.

Table 69: Exact bounded-pilot calculation at a total report error budget of 0.03. Costs include all 530 pilot observations. Ratio is the fixed Hoeffding cost 20999 divided by the exact expected pilot-plus-report cost.

s	Variance	Expected cost	Theorem bound	Ratio	Report error
0.2	0.04	2575.98	2578.48	8.1518	0
2	4	2645.00	2647.89	7.9391	2.60×10^{-122}
10	100	4327.69	4330.59	4.8522	8.09×10^{-10}
25	625	13529.95	13532.84	1.5520	5.96×10^{-6}
50	2500	22441.00	46398.04	0.9357	0.003086

At $s = 50$ the Hoeffding branch wins whenever $K \geq 61$. The complementary event has probability 5.0526×10^{-20}; its expected saving from the cap is 7.8682×10^{-18} observations. Thus the expected cost rounds to 22441.00. Its approximately 1442-observation overhead over the fixed 20999 report consists of 530 pilot observations and 912 additional report observations from reserving 0.005 of the total error budget for variance learning. The general theorem's bound of 46398.04 remains valid but loose at this endpoint; the deterministic capped cost is 22441.

Z.9 Report-aware allocation records and statistical targets

The four-candidate CIFAR allocation study compares four policies using the same covariance information, inspection times, and fresh-reporting rule. Reporting precision is calibrated under the Gaussian score model with known marginal variances. Scores and tolerances use native units. The heterogeneous setting has

$$\mu = (.600, .595, .550, .500), \qquad \sigma = (.200, .020, .050, .100), \qquad \varepsilon = .020.$$

Candidates A and B satisfy the tolerance. A has the largest mean; B has one hundredth of A's marginal variance. The study uses $\delta_s = .02$, $\delta_r = .03$, and checkpoints $n_\ell = 32\,2^\ell$, $\ell = 0, \ldots, 9$. The winner is frozen before independent fresh reporting. Its fixed report count is

$$n_R(i) = \left\lceil z^2_{.985}\sigma_i^2/\eta^2 \right\rceil, \qquad z_{.985} = 2.1700903776. \tag{Z.26}$$

At $\eta = .002$, the candidate-specific counts are $(47093, 471, 2944, 11774)$. Eight independent audit observations are charged on completion and do not change the frozen output. Total counts include selection, fresh reporting, and audit; Q, J, W use all runs, including abstentions.

Table 70: Candidate-specific fresh reporting counts in the heterogeneous study. Adding eight audit observations gives the protected report-plus-audit reserve. Counts use the known-variance normal rule with $\delta_r = .03$.

η	A	B	C	D
.020	471	5	30	118
.005	7,535	76	471	1,884
.002	47,093	471	2,944	11,774

The independent-leader policy uses separate candidate observations. The paired-leader policy uses same-world comparisons and selects the empirical leader. Paired immediate uses reporting cost when choosing among currently certifiable candidates. Guarded lookahead additionally permits further comparisons to make a cheaper-reporting candidate certifiable, subject to the remaining budget

and the stored-certificate guard. This allocation-policy ablation isolates reporting-aware choice beyond pairing: all policies use the same covariance information and fresh-reporting rule. The intervals here use known-covariance normal calibration; Theorem 21 analyzes the bounded-score controller.

Each policy–precision setting has 10,000 runs at cap 200000. Two variance settings, three reporting precisions, and four policies give 24 groups and 240,000 method-run records. Comparisons pair policies by aligned replicate indices and shared worlds; uncertainty is evaluated at the paired-replicate level.

Z.9.1 Cost decomposition and acceptable-winner choice

At $\eta = .002$, guarded lookahead increases mean selection cost from 2720.7296 to 5937.7280, a difference of 3216.9984. Mean reporting cost falls from 40220.9172 to 14443.6134, saving 25777.3038 observations. Including audit, total cost falls from 42949.6468 to 20389.3414. Thus the net saving is 22560.3054 and the leader-to-guarded cost ratio is 2.1064754. A paired 5,000-draw percentile bootstrap gives $[2.0680607, 2.1463518]$ (seed 2026092503).

The probability of selecting B rises from 14.74% to 70.03%. Consequently exact-best frequency decreases from 85.26% to 29.97%, and mean regret rises from $.000737$ to $.0035015$ native units. Both policies select an ε-acceptable candidate in every recorded run. The reporting saving comes from choosing the lower-variance acceptable answer, with the exact-best frequencies quantifying the corresponding change in winner identity.

The corresponding winner counts are $(N_A, N_B) = (8526, 1474)$ for paired leader and $(2997, 7003)$ for guarded lookahead. They account exactly for the integer reporting ledger:

$$\begin{aligned} 8526(47093) + 1474(471) &= 402{,}209{,}172, \\ 2997(47093) + 7003(471) &= 144{,}436{,}134. \end{aligned} \tag{Z.27}$$

Table 71 gives the complete selection, reporting, and audit totals. The observed 2.1064754 mean-cost ratio also equals $429{,}496{,}468/203{,}893{,}414$ because the policies have equal numbers of runs.

Table 71: Integer acquisition ledger at cap 200,000 and $\eta = .002$, aggregated over 10,000 runs per policy. Every run completes. Winner-dependent reporting accounts for the reduction in total cost.

Stage	Paired leader	Guarded lookahead
Selection	27,207,296	59,377,280
Fresh reporting	402,209,172	144,436,134
Audit	80,000	80,000
Total	429,496,468	203,893,414

The four-policy comparison separates three mechanisms. Pairing changes total cost from 42952.9530 to 42949.6468. Choosing the cheapest currently certifiable candidate reduces it to 38893.5328, saving 4056.1140 observations. Continuing comparisons to certify a cheaper answer reduces it by another 18504.1914 to 20389.3414. Thus the immediate-to-guarded ratio is 1.9075424. This last contrast isolates the gain from continuation beyond report-aware terminal choice.

Let n_{10} count runs correct only for guarded lookahead and n_{01} runs correct only for paired leader. Their joint-success table is

$$(n_{11}, n_{10}, n_{01}, n_{00}) = (9525, 171, 163, 141).$$

The correct-output difference is $(171 - 163)/10000 = .0008$, or $+.08$ pp. Taking a one-sided Clopper–Pearson lower bound for p_{10} and upper bound for p_{01}, each with error $.025$, gives a simultaneous difference lower bound $-.004327134$, or $-.4327134$ pp. This is an exploratory paired comparison at the specified cap, separate from the prespecified 1,000-pair application comparison in Appendix W.

Z.9.2 The budget frontier

Table 73 reports a separate budget sweep at $\eta = .002$ with 10,000 runs per policy–budget point. Simultaneous evaluation covers 14 points, two metrics (Q, J), and two tails: each of 56 Clopper–Pearson statements receives error $.05/56$. For guarded lookahead at $B = 20000$, the lower bounds

Table 72: Heterogeneous-variance study: all 12 policy–precision groups at cap 200,000, with 10,000 runs per row. $Q = 1$ throughout, so $W = 1 - J$; mean audit cost is eight. All counts are candidate–world observations.

η	Policy	Selection	Reporting	Total	J (%)
.020	Independent leader	2738.0224	402.1718	3148.1942	96.96
	Paired leader	2720.7296	402.3116	3131.0412	96.96
	Paired immediate	2720.7296	361.7696	3090.4992	97.01
	Guarded lookahead	2720.7296	361.7696	3090.4992	97.01
.005	Independent leader	2738.0224	6433.3057	9179.3281	96.90
	Paired leader	2720.7296	6435.5434	9164.2730	96.89
	Paired immediate	2720.7296	5786.6104	8515.3400	96.88
	Guarded lookahead	4101.6960	3927.0817	8036.7777	96.94
.002	Independent leader	2738.0224	40206.9306	42952.9530	96.89
	Paired leader	2720.7296	40220.9172	42949.6468	96.88
	Paired immediate	2720.7296	36164.8032	38893.5328	96.87
	Guarded lookahead	5937.7280	14443.6134	20389.3414	96.96

are $Q \geq .9926284$ and $J \geq .9584531$; for paired leader at $B = 60000$, they are $Q \geq .9992981$ and $J \geq .9629909$. These are the smallest evaluated budgets whose simultaneous lower bounds exceed the targets $Q \geq .95$, $J \geq .90$, giving a grid-budget ratio of 3.0. The smaller evaluated caps have completion upper bounds below $.95$. This ratio concerns this study's grid, not the original application's 2.30 budget ratio or the 2.1065 mean-cost ratio at cap 200,000.

Table 73: Budget-dependent completion and correct output in the heterogeneous setting, $\eta = .002$. Rates are percentages of all 10,000 runs per row and policy; wrong-certificate rate is $W = Q - J$. Mean cost includes unsuccessful runs. L: paired leader; G: guarded lookahead.

Budget	Q_L	J_L	Mean cost L	Q_G	J_G	Mean cost G
4,000	7.60	7.43	2164.9608	7.60	7.43	2164.9608
8,000	15.92	15.44	4096.5512	27.95	27.02	4154.1749
12,000	17.55	17.01	7547.0469	77.27	74.99	7340.3569
20,000	17.84	17.29	14300.2824	99.52	96.46	9306.5184
40,000	17.84	17.29	27756.4616	99.97	96.89	9382.4019
60,000	100.00	96.88	42949.6468	100.00	96.98	21967.7422
100,000	100.00	96.88	42949.6468	100.00	96.96	20389.3414

The policy depends on the cap. More available budget can permit different deferral decisions, explaining why the guarded policy's mean consumption at 60,000 exceeds that at 40,000. Each cap is evaluated as its own policy configuration; the frontier is not a truncation of one high-budget trajectory.

Table 74 expresses the frontier as integer output and acquisition counts. Completion decomposes exactly as correct joint output plus incorrect certification; every remaining run is an abstention. The charged totals include all runs at each cap.

At $\eta = .002$, the report-plus-audit reserves are $(R_A, R_B, R_C, R_D) = (47101, 479, 2952, 11782)$. For a completed output i, selection acquisition therefore obeys $C_{\text{sel}} \leq B - R_i$. At the five evaluated caps below 47,101, reporting A is unaffordable, and C and D fail the selection tolerance. Consequently every correct joint output at these caps selects B. For guarded lookahead the resulting counts are 743, 2702, 7499, 9646, and 9689; for paired leader they are 743, 1544, 1701, 1729, and 1729, in increasing budget order. These identities connect the completion frontier to the choice of an affordable acceptable winner.

The two qualifying grid points have joint-output rates 96.46% (guarded, 20,000) and 96.88% (leader, 60,000), a difference of $-.42$ pp. Their marginal simultaneous lower bounds establish the stated completion and correctness targets. The paired $+.08$ pp comparison above concerns the common cap 200,000.

Table 74: Integer budget-frontier ledger, 10,000 runs per row. Q, J, W are completed, correct-joint, and incorrect-certificate counts; A counts abstentions. Total is the sum of charged candidate–world observations. L: paired leader; G: guarded lookahead.

Budget	Policy	Q	J	W	A	Total
4,000	L	760	743	17	9,240	21,649,608
	G	760	743	17	9,240	21,649,608
8,000	L	1,592	1,544	48	8,408	40,965,512
	G	2,795	2,702	93	7,205	41,541,749
12,000	L	1,755	1,701	54	8,245	75,470,469
	G	7,727	7,499	228	2,273	73,403,569
20,000	L	1,784	1,729	55	8,216	143,002,824
	G	9,952	9,646	306	48	93,065,184
40,000	L	1,784	1,729	55	8,216	277,564,616
	G	9,997	9,689	308	3	93,824,019
60,000	L	10,000	9,688	312	0	429,496,468
	G	10,000	9,698	302	0	219,677,422
100,000	L	10,000	9,688	312	0	429,496,468
	G	10,000	9,696	304	0	203,893,414

Z.9.3 A PRESSURE TEST WITH NO ALTERNATIVE ACCEPTABLE WINNER

The second setting has $\mu = (.60, .57, .54, .51)$, common marginal standard deviation .10, equicorrelation .90, and $\varepsilon = .02$. Only A is acceptable. Mean selection cost is 525.4176 for independent leader and 128.0512 for paired leader. Paired immediate and guarded lookahead return the same outputs and costs as paired leader at each precision. There is no lower-report-cost acceptable alternative to pursue. The covariance matrix in this pressure setting is $\Sigma_{\text{pressure}} = .001I_4 + .009\mathbf{1}\mathbf{1}^\top$, so every paired difference has variance .002, compared with .020 under independent observations.

Table 75: High-shared-noise pressure test at cap 200,000; 10,000 runs per policy and precision. All policies complete, and every paired policy has the same records. I and P denote independent and paired leader.

η	Report count	Total I	Total P	I/P	J_I (%)	J_P (%)
.020	118	651.4176	254.0512	2.5641	97.01	97.02
.005	1884	2417.4176	2020.0512	1.1967	96.99	97.00
.002	11774	12307.4176	11910.0512	1.0334	96.99	97.00

Pairing removes the same 397.3664 mean selection observations at all three precisions. Its relative end-to-end gain shrinks as reporting becomes dominant (Table 75). Together with the heterogeneous result, this locates the report-aware mechanism: it benefits instances with multiple acceptable candidates and unequal reporting costs.

Z.10 ADDITIONAL APPLICATION ACCOUNTING AND PAIRED REPORT ANALYSIS

For the 20 paired-selection winners in Appendix V, define $D_r = \widehat{\mu}_{N,r} - \widehat{\mu}_{F,r}$ using the same winner's naive and fresh reports. Reanalysis of the displayed two-decimal values gives setting means $.46, .81, .92, 1.35$ pp and an equal-setting macro mean $.885$ pp. A setting-stratified paired bootstrap with 20,000 draws (seed 2026092502) gives a percentile 95% range $[.7035, 1.0415]$ pp. This measures a signed report difference, whereas the 92% result measures reference-relative macro absolute signed bias. The same reference cancels from D_r; rounding of the displayed reports remains part of this reanalysis.

The execution accounting treats each counted candidate–world observation as a distinct, independently executed complete configuration–world run. The measured mean duration on the specified single-GPU execution hardware is 650 seconds per run. Table 76 reports generation accelerator-hours calibrated by that measured mean, $H(N) = 650N/3600$. These are generation-time estimates from counted executions and mean timing, not cache retrieval or a sum of individual job-duration

records. The calibration applies to the listed application counts and does not replace other studies' timing protocols.

Table 76: Application generation cost calibrated by the measured mean of 650 seconds per independently executed configuration–world on one GPU. The 20-run P and I counts retain their original selection, reporting, and audit definitions. Reference/proposal generation outside these counts is additional. GPU-hours and wall-clock makespan are distinct.

Counted workload	Executions N	Calibrated generation GPU-h
One fixed fresh report	20,999	3,791.49
Twenty fresh reports (each method)	419,980	75,829.72
Paired selection/report/audit	2,601,620	469,736.94
Independent selection/report/audit	5,983,726	1,080,394.97