WHEN LLMS ANSWER WITH RANGES: SAMPLE-OPTIMAL SELECTION UNDER INTERVAL-VALUED RESPONSES

Anonymous authors
Paper under double-blind review

ABSTRACT

Large language models (LLMs) are increasingly used as stochastic simulators to support selection from large sets of candidate designs and products. In many applications with high uncertainty, LLMs are more likely to return a numerical interval rather than a scalar. Motivated by the potentially vast number of candidates, we study large-scale selection under interval-valued responses. The objective is lexicographic: maximize the expected midpoint and, among alternatives that tie on midpoint, minimize the expected radius. We introduce **I**nterval-**R**esponse **I**nterlaced **S**election (IRIS) with two greedy algorithms. F-IRIS uses a **F**ixed midpoint tolerance as an oracle benchmark; A-IRIS replaces it with a data-dependent candidate set. Theoretically, we prove that F-IRIS and A-IRIS are both sample optimal, that is, a budget linear in the number of alternatives achieves a positive probability of correct selection. By choosing larger finite initialization and budget coefficients, both methods can achieve any pre-specified accuracy. Experiments with Gaussian, uniform, and skewed midpoint noise support the asymptotic results. A study based on repeated Llama queries shows when interval information improves over midpoint-only selection.

1 INTRODUCTION

Large language models (LLMs) have advanced rapidly and are now used across a wide range of domains. Many LLM applications ultimately involve a selection problem: a practitioner repeatedly queries a model to compare prompts, product configurations, generated designs, or virtual-screening candidates, and then selects one alternative with maximum population mean. Because model responses are stochastic, reliable selection may require repeated queries, especially when the candidate pool is large. Yet the associated token and computational costs can become substantial at scale. The central objective is therefore to maximize the quality of the final selection under a finite query budget. This is the pure-exploration objective studied in best-arm identification (BAI) and ranking and selection (R&S) (Audibert et al., 2010; Karnin et al., 2013; Hong et al., 2021).

LLM responses need not be scalar. When asked for willingness to pay (WTP), completion time, or a plausible range, a model often return an interval to reflect uncertainty. Collapsing that interval to its midpoint discards the information about its width, whereas optimizing width alone can favor a precise but poor alternative. A possible solution is to construct a weighted score such as midpoint minus a multiple of width, thereby reducing the problem to a scalar selection problem. However, this approach requires choosing a weight that determines the trade-off between midpoint and radius performance. We instead consider a lexicographic target. The primary objective is expected midpoint; expected radius resolves ties among alternatives with the best midpoint. This target is useful when response level has priority and width matters only after that primary requirement has been met. The resulting problem is therefore a lexicographic two-objective selection problem.

The statistical challenge is that neither the set of midpoint-optimal alternatives nor its size is known in advance. Radius samples spent on a midpoint-inferior alternative are wasted for the final decision, but eliminating alternatives too quickly can permanently remove the target. The difficulty becomes sharper as the number of alternatives K grows while the total budget remains proportional to K. Classical BAI methods allocate samples to a scalar mean, and most analyses assume a unique best

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.

arm. Large-scale R&S shows that greedy allocation can still be sample optimal under positive separation (Li et al., 2025), but the interval problem must first locate a set under midpoint signal and then select within it using width information.

We propose **I**nterval-**R**esponse **I**nterlaced **S**election (IRIS), a fixed-budget framework for this setting. Following an initial exploration phase, both IRIS algorithms interlace midpoint and radius updates. At each midpoint epoch, they greedily sample the alternative with the largest empirical midpoint, allowing the procedure to correct a spurious leader caused by limited initial sampling. At each radius epoch, they sample the alternative with the smallest empirical radius within a dynamically recomputed midpoint candidate set. Once the budget is exhausted, the procedure returns the alternative selected most frequently during the radius epochs. This modal recommendation reduces unrobustness near the end of the sampling process. F-IRIS uses a fixed tolerance smaller than the unknown population midpoint gap. It can therefore be used when prior information or expert knowledge provides a credible lower bound on that gap; otherwise, it serves as an oracle benchmark. A-IRIS instead constructs the candidate set from confidence bounds and requires no externally specified tolerance, making it the practical variant of the framework.

Our contributions are as follows.

- We formulate selection from interval-valued stochastic responses as a lexicographic two-objective Top-1 selection problem. The formulation retains the midpoint and radius as separate estimands and includes ordinary scalar BAI as the zero-radius special case.
- We prove sample optimality and consistency for F-IRIS and A-IRIS in a large-scale regime. A linear budget in K is sufficient to achieve a positive probability of correct selection (PCS). For any pre-specified error level, finite choices of the initial sample size and linear budget coefficient independent of K ensure that the asymptotic PCS reaches the error level. The required conditions are mild: midpoint observations are sub-Gaussian, while the radii of midpoint-optimal alternatives need only be integrable, covering many commonly used distributions.
- We provide extensive numerical evidence across midpoint gaps, problem scales, and three noise distributions, demonstrating the outperformance of IRIS and the practical value of interval information. On repeated `Llama-3.2-3B-Instruct` WTP responses for laptop screening, IRIS can separate alternatives that midpoint-only selection cannot. Our sensitivity analysis further demonstrates how IRIS supports different practically meaningful definitions of near-optimal midpoint performance.

2 Related Work

Best-arm identification and ranking and selection. BAI studies pure exploration: samples are collected to identify an arm with the largest mean, rather than to maximize reward while learning. Fixed-budget methods include successive rejects and sequential halving (Audibert et al., 2010; Karnin et al., 2013); fixed-confidence methods include LUCB and Track-and-Stop (Kalyanakrishnan et al., 2012; Garivier & Kaufmann, 2016). R&S develops closely related guarantees for simulated systems, including indifference-zone procedures, OCBA, and parallel or large-scale selection (Chen et al., 2000; Kim & Nelson, 2001; Ni et al., 2017). Of particular relevance, Li et al. (2025) prove that greedy and explore-first greedy rules can achieve nonvanishing probability of correct selection with a budget linear in K. IRIS shares the large-K fixed-budget goal, but its observations have two ordered components and its primary optimum can be nonunique. Lexicographic BAI also prioritizes vector rewards (Xue et al., 2026); here the radius resolves ties in an unknown midpoint-optimal set. Appendix A gives a fuller comparison.

LLM selection and uncertainty. LLMs are now used both to generate candidates and to evaluate them. Fixed-budget selection has been applied to LLM-enabled virtual screening (Li et al., 2024b), and prompt optimization can be cast directly as BAI over a candidate prompt pool (Shi et al., 2024). Other work studies whether LLM judgments reproduce human preferences or experimental behavior (Li et al., 2024a; Goli & Singh, 2024; Aher et al., 2023). Semantic entropy and related methods quantify uncertainty over free-form text generations (Kuhn et al., 2023; Farquhar et al., 2024). IRIS addresses a different estimand: a numerical interval explicitly returned by the model. Its radius

is an observed response feature, which may represent human uncertainty of WTP, not a calibrated confidence interval.

Interval-valued data and multi-criteria selection. Interval-valued data arise naturally when a variable is summarized by its observed range over a period, as in daily high–low stock prices and minimum–maximum temperatures. Statistical methods for interval-valued data commonly represent each interval by its midpoint and radius (Sun et al., 2022; González-Rivera & Lin, 2013). Related work considers several alternative decision rules for interval-valued or vector-valued outcomes, including directly ordering intervals, identifying the set of Pareto-optimal alternatives, and finding the best arm separately for each objective (Sengupta & Pal, 2000; Andradóttir & Lee, 2021; Chen et al., 2025). IRIS instead returns one alternative under a strict midpoint-then-radius preference, without scalarization or estimation of a full Pareto set.

3 Problem Formulation

We consider a fixed-budget ranking-and-selection problem with K alternatives. Querying alternative k returns an interval $[X^l_{k,\ell}, X^u_{k,\ell}]$ at replication ℓ. We retain its midpoint and radius,

$$X^m_{k,\ell} := \frac{X^l_{k,\ell} + X^u_{k,\ell}}{2}, \qquad X^r_{k,\ell} := \frac{X^u_{k,\ell} - X^l_{k,\ell}}{2},$$

with means $\mu^m_k := \mathbb{E}[X^m_{k,1}]$ and $\mu^r_k := \mathbb{E}[X^r_{k,1}]$. The midpoint measures the level of the response, while the radius records its reported width, representing fluctuation or uncertainty of prices. When $X^r_{k,\ell} = 0$, the interval reduces to a scalar. We therefore adopt a lexicographic preference. We first maximize the expected midpoint, thereby identifying the alternatives with the highest expected response. Because multiple alternatives may attain the same optimal midpoint, we then minimize the expected radius within this set, favoring more precise responses without sacrificing expected midpoint performance. Writing

$$G_m := \arg \max_{1 \le k \le K} \mu^m_k,$$

the target is

$$k^* := \arg \min_{k \in G_m} \mu^r_k.$$

After sampling n observations of alternative k via LLMs, we denote the empirical means as

$$\bar{X}^m_k(n) := \frac{1}{n} \sum_{\ell=1}^{n} X^m_{k,\ell}, \qquad \bar{X}^r_k(n) := \frac{1}{n} \sum_{\ell=1}^{n} X^r_{k,\ell}.$$

A policy adaptively chooses which alternative to query, subject to a total budget B_K, and returns $\widehat{k}$. Its probability of correct selection is $\mathrm{PCS}_K := \Pr(\widehat{k} = k^*)$.

We begin by stating the assumptions used in our analysis. Without loss of generality, we index the alternatives so that $G_m = \{1, \dots, K_1\}$ and $k^* = 1$.

Assumption 1 (Fixed optimal set and positive gaps)**.** K_1 is unknown and fixed as $K \to \infty$. There are constants $\mu^m_* \in \mathbb{R}$ and $\gamma_1, \gamma_2 > 0$, independent of K, such that

$$\mu^m_1 = \cdots = \mu^m_{K_1} = \mu^m_*, \quad \sup_{k > K_1} \mu^m_k \le \mu^m_* - \gamma_1, \quad \mu^r_j - \mu^r_1 \ge \gamma_2 \quad (2 \le j \le K_1).$$

Assumption 1 considers the more practical scenario where K_1 is unknown. If K_1 were known, the problem could be decomposed into two existing ranking and selection tasks: the top-K_1 selection algorithm in Li et al. (2024b) could be directly applied to identify the midpoint candidates, followed by the top-1 selection algorithm in Li et al. (2025) to determine the candidate with the optimal radius. However, in practice, the size of the optimal midpoint set is rarely known a priori. Moreover, Assumption 1 does not impose any structural condition on the population radius means of candidates with inferior midpoint means. This prevents the direct use of simple weighted scores, such as $\bar{X}^m_k(n) - \eta \bar{X}^r_k(n)$, because a candidate with a poor midpoint mean may still achieve a large score due to its sufficiently small radius. Therefore, although Assumption 1 is relatively mild and allows for broad applicability, it introduces a fundamental algorithmic challenge: how to reliably identify the target candidate without prior knowledge of the optimal midpoint set.

Assumption 2 (Sampling paths)**.** For each alternative, the pair-valued observations $\{(X^m_{k,\ell}, X^r_{k,\ell}) : \ell \geq 1\}$ are independent and identically distributed over replications, and the complete pair-valued sampling paths are independent across alternatives. For $k \in G_m$, the distribution of the pair-valued path does not depend on K.

Assumption 2 imposes standard conditions in simulation optimization. Each LLM response is generated by a separate query under the same prompt and decoding configuration, making it natural to model repeated responses as independent draws. Because the midpoint and radius are computed from the same interval response, the assumption allows them to be dependent.

Assumption 3 (Observation distributions)**.** Radii in G_m are nonnegative and integrable. For a known finite $\bar{\sigma}_m$, midpoint observations satisfy, uniformly in k and K,

$$\mathbb{E}\big[\exp\{u(X^m_{k,\ell} - \mu^m_k)\}\big] \leq \exp(u^2\bar{\sigma}^2_m/2), \qquad u \in \mathbb{R}.$$

Assumption 3 requires midpoint observations to be uniformly sub-Gaussian, a class that includes a wide range of distributions like Gaussian, uniform and skewed beta distribution. For radii, we only require the integrability condition for alternatives in G_m and no condition is imposed outside G_m. In contrast to Sun et al. (2022), who model interval ranges using a log-Gaussian distribution, our analysis does not restrict any distributional assumption for radius.

We end this section with two definitions.

Definition 1 (Sample optimality)**.** We call a policy *sample optimal* if some finite c yields $\liminf_{K\to\infty} \mathrm{PCS}_K > 0$ under $B_K = \lfloor cK \rfloor$.

Definition 2 (Consistency)**.** For every $\alpha \in (0, 1)$, finite constant $c(\alpha)$ can be chosen independent of K such that $\liminf_{K\to\infty} \mathrm{PCS}_K > 1 - \alpha$ under $B_K = \lfloor c(\alpha)K \rfloor$.

4 FIXED-TOLERANCE IRIS

Since the size of the midpoint-optimal set is unknown, the algorithm must learn which alternatives remain competitive in midpoint performance. Given the positive separation γ_1, a natural approach is to construct a dynamic candidate set using a tolerance parameter. We first study F-IRIS with a fixed tolerance $\epsilon \in (0, \gamma_1)$, which serves as an oracle benchmark. The next section develops an adaptive version that does not require a prespecified tolerance.

Let $B_K = \lfloor cK \rfloor$. We begin by allocating n_e observations to every alternative. As shown by Li et al. (2025), this explore-first step can improve PCS over a purely greedy allocation by reducing the probability that early sampling noise causes the procedure to commit to an inferior alternative. Let $T_K := B_K - Kn_e$ denote the remaining budget after initialization. We allocate these T_K observations through interlaced midpoint and radius epochs, allowing the midpoint candidate set and empirical radius estimates to be updated throughout the procedure. In particular, for some constant $\lambda \in (0, 1)$ that determines the fraction of midpoint-channel sampling, define

$$a_s := \lfloor \lambda s \rfloor - \lfloor \lambda(s-1) \rfloor, \qquad s = 1, \ldots, T_K.$$

Then $a_s = 1$ marks a midpoint epoch and $a_s = 0$ marks a radius epoch. For example, if $\lambda = 1/3$, the sequence is $(a_s)_{s\geq 1} = (0, 0, 1, 0, 0, 1, \ldots)$, so every third epoch is a midpoint epoch and the remaining epochs are radius epochs. At time $t \in \{Kn_e, Kn_e + 1, \ldots, B_K\}$, the dynamic fixed-tolerance set is

$$\mathcal{C}^\epsilon_t := \left\{ k : \bar{X}^m_k(n_k(t)) \geq \max_i \bar{X}^m_i(n_i(t)) - \epsilon \right\}, \tag{1}$$

where $n_i(t)$ denotes the sample size of alternative i. Such a candidate set in equation 1 is influenced by the empirical midpoint leader. Thus the midpoint epochs revisit the current empirical leader, allowing an initially exaggerated alternative to be corrected by adding samples. At each radius epoch, the algorithm restricts attention to the current candidate set and greedily samples the alternative with the smallest empirical mean radius, allowing an inferior radius leader caused by sampling noise to be corrected as additional observations are collected. The algorithm records the recommendation at each radius epoch and returns the most frequently recommended alternative. Using the modal recommendation rather than the final one prevents a brief fluctuation near the end of the sampling horizon from determining the output. The F-IRIS algorithm is shown in Algorithm 1.

Algorithm 1: Fixed-Tolerance IRIS (F-IRIS)

Input : K, B_K, n_e, ϵ, and $\lambda \in (0,1)$
Output: Selected alternative $\widehat{k}_{\mathrm{F}}$
Sample each alternative n_e times; set $n_k = n_e$, $t = Kn_e$, and $V_k = 0$ for all k;
for $s = 1, \ldots, T_K$ **do**
  if $a_s = 1$ **then**
    $q_t \leftarrow \min \arg\max_k \bar{X}_k^m(n_k)$;
  else
    form $\mathcal{C}_t^\epsilon$ using equation 1;
    $q_t \leftarrow \min \arg\min_{k \in \mathcal{C}_t^\epsilon} \bar{X}_k^r(n_k)$; $V_{q_t} \leftarrow V_{q_t} + 1$;
  end
  Sample q_t once, update both empirical means, and increment n_{q_t} and t;
end
$\widehat{k}_{\mathrm{F}} \leftarrow \min \arg\max_k V_k$;

To state the guarantee, fix $0 < \delta < \frac{1}{2}\min\{\epsilon, \gamma_1 - \epsilon\}$ and define

$$a_{\epsilon,\delta} := \frac{(\gamma_1 - \epsilon - \delta)^2}{2\bar{\sigma}_m^2}, \qquad C_{\epsilon,\delta}(n_e) := \frac{e^{-a_{\epsilon,\delta} n_e}}{(1 - e^{-a_{\epsilon,\delta}})^2},$$

$$p_m(n_e, \delta) := 1 - 2K_1 \sum_{n=n_e}^{\infty} e^{-n\delta^2/(2\bar{\sigma}_m^2)}, \qquad L_\lambda := \lceil \lambda^{-1} \rceil.$$

For $K_1 = 1$, set $p_r(n_e) = 1$. Otherwise, fix $\delta_1 \in (0, \gamma_2)$ and define

$$p_r(n_e) := \Pr\left(\sup_{n \geq n_e} \bar{X}_1^r(n) \leq \mu_1^r + \delta_1, \min_{2 \leq j \leq K_1} \inf_{n \geq n_e} \bar{X}_j^r(n) > \mu_1^r + \delta_1 \right).$$

For every $\eta \in (0, 1/2)$, define

$$n_r(\eta) := \min\{n \geq 1 : p_r(n) \geq 1 - \eta\}. \tag{2}$$

Theorem 1 (Sample optimality of F-IRIS)**.** *Define the Fréchet lower bound* $\pi_{\mathrm{F}}(n_e, \delta) := p_m(n_e, \delta) + p_r(n_e) - 1$. *Suppose Assumptions 1–3 hold. If the integer* n_e *and* c *satisfy*

$$n_e \geq \max\left\{ \left\lceil \frac{2\bar{\sigma}_m^2}{\delta^2} \log\left(\frac{2K_1}{\eta\left[1 - \exp\left\{-\delta^2/(2\bar{\sigma}_m^2)\right\}\right]} \right) \right\rceil, n_r(\eta) \right\}$$
$$\textit{and} \quad (1 - \lambda)(c - n_e) > 2L_\lambda C_{\epsilon,\delta}(n_e), \tag{3}$$

then Algorithm 1 satisfies

$$\liminf_{K \to \infty} \mathrm{PCS}_K \geq \pi_{\mathrm{F}}(n_e, \delta)\left(1 - \frac{2L_\lambda C_{\epsilon,\delta}(n_e)}{(1 - \lambda)(c - n_e)}\right) > 0. \tag{4}$$

Theorem 2 (Consistency of F-IRIS)**.** *Under the same assumptions, for* $\alpha \in (0,1)$, *set*

$$n_e(\alpha) \geq \max\left\{ \left\lceil \frac{2\bar{\sigma}_m^2}{\delta^2} \log\left[\frac{6K_1}{\alpha\left(1 - \exp\left\{-\delta^2/(2\bar{\sigma}_m^2)\right\}\right)} \right] \right\rceil, \; n_r\left(\frac{\alpha}{3}\right) \right\},$$
$$c(\alpha) \geq n_e(\alpha) + \frac{6L_\lambda C_{\epsilon,\delta}(n_e(\alpha))}{(1 - \lambda)\alpha}.$$

Then $n_e(\alpha)$ *and* $c(\alpha)$ *are finite and independent of* K, *and Algorithm 1 with* $B_K = \lfloor c(\alpha) K \rfloor$ *satisfies*

$$\liminf_{K \to \infty} \mathrm{PCS}_K \geq 1 - \alpha.$$

Theorem 1 provides a lower bound on PCS and establishes the sample optimality of F-IRIS, while Theorem 2 establishes its consistency. The implicit definition of $n_r(\alpha)$ in equation 2 reflects our assumption that radii need only be integrable. If, in addition, $\mathrm{Var}(X_{j,\ell}^r) \leq \sigma_r^2 < \infty$ for all $j \in G_m$, an explicit sufficient choice is $n_r(\alpha) = 1$ when $K_1 = 1$, and $n_r(\alpha/3) = \left\lceil \frac{12\sigma_r^2}{\alpha}\left[\frac{1}{\delta_1^2} + \frac{K_1 - 1}{(\gamma_2 - \delta_1)^2}\right] \right\rceil$ when $K_1 > 1$. The supporting lemmas and proofs are provided in Appendix E.

5 ADAPTIVE-TOLERANCE IRIS

F-IRIS is applicable when prior information or domain expertise provides a credible lower bound on the midpoint gap γ_1. In most applications, however, this gap is unknown, so a valid fixed tolerance cannot be specified in advance. Estimating γ_1 directly is also impractical in large-scale problems, because doing so requires enough samples to distinguish the midpoint-optimal set from its closest inferior competitor before the selection procedure is run. Instead, A-IRIS replaces the fixed tolerance with confidence bounds for the midpoint means. For $\beta > 0$, define

$$w(n) := \bar{\sigma}_m \sqrt{\frac{2(\beta + 2\log n)}{n}}, \quad L_k(n) := \bar{X}_k^m(n) - w(n), \quad U_k(n) := \bar{X}_k^m(n) + w(n).$$

The parameter $\beta > 0$ controls the confidence level of the time-uniform bounds. Increasing β makes the bounds wider and reduces the probability of excluding a midpoint-optimal alternative. We take $\beta = 1$ in our experiments. The adaptive candidate set is

$$\mathcal{C}_t^{\mathrm{ad}} := \left\{ k : U_k(n_k(t)) \geq \max_i L_i(n_i(t)) \right\}. \tag{5}$$

In equation 5, the lower and upper confidence bounds account for the different levels of uncertainty caused by unequal sample sizes. An alternative is excluded only when its upper confidence bound falls below the largest lower confidence bound, providing sufficient evidence that its expected midpoint is suboptimal. As in F-IRIS, A-IRIS alternates between midpoint and radius epochs until the budget is exhausted. It samples the empirical midpoint leader at each midpoint epoch and recomputes the candidate set in equation 5 before sampling at each radius epoch. The complete procedure is given in Algorithm 2.

Note that we use known sub-Gaussian upper bound $\bar{\sigma}_m$ in A-IRIS and its theories for the theoretical analysis. In practice, $\bar{\sigma}_m$ can be estimated via a high-probability upper bound during the exploration phase or updated throughout the sampling process. In our LLM experiment, we compute the midpoint sample standard deviation for each alternative from its n_e initial observations and set $\bar{\sigma}_m$ to their 0.95 quantile.

Algorithm 2: Adaptive-Tolerance IRIS (A-IRIS)

Input : $K, B_K, n_e, \beta, \bar{\sigma}_m$, and $\lambda \in (0, 1)$
Output: Selected alternative $\widehat{k}_{\mathrm{A}}$
Sample each alternative n_e times; set $n_k = n_e$, $t = Kn_e$, and $V_k = 0$ for all k;
for $s = 1, \ldots, T_K$ **do**
  if $a_s = 1$ **then**
    $q_t \leftarrow \min \arg\max_k \bar{X}_k^m(n_k)$;
  else
    form $\mathcal{C}_t^{\mathrm{ad}}$ using equation 5;
    $q_t \leftarrow \min \arg\min_{k \in \mathcal{C}_t^{\mathrm{ad}}} \bar{X}_k^r(n_k)$; $V_{q_t} \leftarrow V_{q_t} + 1$;
  end
  Sample q_t once, update both empirical means, and increment n_{q_t} and t;
end
$\widehat{k}_{\mathrm{A}} \leftarrow \min \arg\max_k V_k$;

Fix $\delta_0 \in (0, \gamma_1)$ and put $\Delta_m := \gamma_1 - \delta_0$, denote

$$n_{\mathrm{sep}} := \min\{n \geq n_e : w(n) \leq \Delta_m/2\}, \qquad a_m := \frac{\Delta_m^2}{8\bar{\sigma}_m^2},$$

$$C_{\mathrm{ad}}(n_e) := n_{\mathrm{sep}} - n_e + \frac{e^{-a_m n_{\mathrm{sep}}}}{1 - e^{-a_m}}, \qquad p_A(n_e) := 1 - 2K_1 e^{-\beta} \sum_{n=n_e}^{\infty} n^{-2}.$$

Theorem 3 (Sample optimality of A-IRIS)**.** *Under Assumptions 1–3, if the integer n_e and c satisfy*

$$n_e \geq \max\left\{ \left\lceil \frac{16\bar{\sigma}_m^2\beta}{\delta_0^2} + \frac{32\bar{\sigma}_m^2}{\delta_0^2} \log\max\left\{1, \frac{32\bar{\sigma}_m^2}{\delta_0^2}\right\} \right\rceil, \; 1 + \left\lceil \frac{2K_1 e^{-\beta}}{\eta} \right\rceil, \; n_r(\eta) \right\},$$

and $(1 - \lambda)(c - n_e) > 2L_\lambda C_{\mathrm{ad}}(n_e)$,

then Algorithm 2 satisfies

$$\liminf_{K\to\infty} \mathrm{PCS}_K > 0. \tag{6}$$

Theorem 4 (Consistency of A-IRIS)**.** *Under the same assumptions, fix $\beta > 0$, $\lambda \in (0,1)$, and $\delta_0 \in (0, \gamma_1)$. For every $\alpha \in (0,1)$, choose an integer $n_e(\alpha)$ satisfying*

$$n_e(\alpha) \geq \max\left\{ \left\lceil \frac{16\bar{\sigma}_m^2\beta}{\delta_0^2} + \frac{32\bar{\sigma}_m^2}{\delta_0^2} \log \max\left\{1, \frac{32\bar{\sigma}_m^2}{\delta_0^2}\right\} \right\rceil, 1 + \left\lceil \frac{6K_1 e^{-\beta}}{\alpha} \right\rceil,\ n_r\left(\frac{\alpha}{3}\right) \right\},$$

and choose $c(\alpha)$ such that

$$c(\alpha) \geq n_e(\alpha) + \frac{6L_\lambda C_{\mathrm{ad}}(n_e(\alpha))}{(1-\lambda)\alpha}.$$

Then, Algorithm 2 with $B_K = \lfloor c(\alpha)K \rfloor$ satisfies

$$\liminf_{K\to\infty} \mathrm{PCS}_K \geq 1 - \alpha.$$

Appendix F gives the lemmas and complete proofs.

6 Numerical Experiments

We conduct two experiments to illustrate the asymptotic guarantees: Experiment 1 examines sample optimality, and Experiment 2 examines consistency. Both share the following settings. Each sample returns an interval, which is decomposed into a midpoint and a radius. Unless stated otherwise, the midpoint variance is one, and the radius follows a log-normal distribution with log-variance 0.01. We compare F-IRIS and A-IRIS with EFG, the benchmark procedure proposed by Li et al. (2025). EFG greedily samples the alternative with the largest empirical midpoint and returns the alternative with the largest final empirical midpoint, without using radius information. The experimental designs are adapted from Li et al. (2025) to interval-valued responses. PCS is estimated from 200 independent macro-replications. Finally in Section 6.3, we apply the proposed methods to a case study with LLM-enabled laptop screening.

6.1 Experiment 1: midpoint separation and distributional robustness

This experiment varies both problem size and the population midpoint gap. We use $K \in \{2^2, \ldots, 2^{14}\}$ and $\gamma_1 \in \{0.1, 0.3, 0.5, 0.7, 1\}$. The total budget is $B_K = 100K$, every alternative receives $n_e = 80$ initial observations, and the midpoint-channel fraction is $\lambda = 0.375$. F-IRIS uses the oracle tolerance $\epsilon = \gamma_1/2$. A-IRIS uses $\beta = 1$ and $\bar{\sigma}_m = 1$.

We consider two configurations: the unique midpoint-optimum setting (UMOS), with $K_1 = 1$, and the tied midpoint-optima setting (TMOS), with $K_1 = 4$. In both configurations, $\mu_k^m = \gamma_1$ for $1 \leq k \leq K_1$ and $\mu_k^m = 0$ for $k > K_1$. In UMOS, all alternatives share the same radius distribution $\log X_{k,\ell}^r \sim \mathcal{N}(0.1, 0.01)$. For TMOS, the radius observations satisfy $\log X_{k,\ell}^r \mid \nu_k \sim \mathcal{N}(\nu_k, 0.01)$ with $\nu_1 = 0$ and $\nu_2 = \nu_3 = \nu_4 = 0.1$. For each $k > 4$, the log-radius location is randomly 0 with probability 1/4 and 0.1 with probability 3/4. Thus, midpoint-inferior alternatives share the same radius distributions as midpoint-optimal alternatives, so radius information alone cannot identify G_m. We repeat the design under three midpoint noise distributions: Gaussian, $\mathcal{N}(\mu_k^m, 1)$; uniform, $\mu_k^m + \mathrm{Unif}(-\sqrt{3}, \sqrt{3})$; and a centered, variance-one transformation of $\mathrm{Beta}(2,5)$. All three noise distributions have mean zero and variance one.

Figure 1 shows the selection performance of the two IRIS variants under Gaussian midpoint noise; the uniform and beta results are reported in Appendix B. When $\gamma_1 \geq 0.5$, both IRIS variants attain PCS close to one at large K. A-IRIS exhibits a nonmonotone pattern: its PCS initially declines before rising rapidly toward one. This pattern is consistent with the limited post-exploration budget at small K, which restricts midpoint refinement and radius-based selection. In TMOS, EFG remains near $1/K_1$ because midpoint information alone cannot distinguish the four midpoint-optimal alternatives, whereas IRIS uses radius information to resolve the tie.

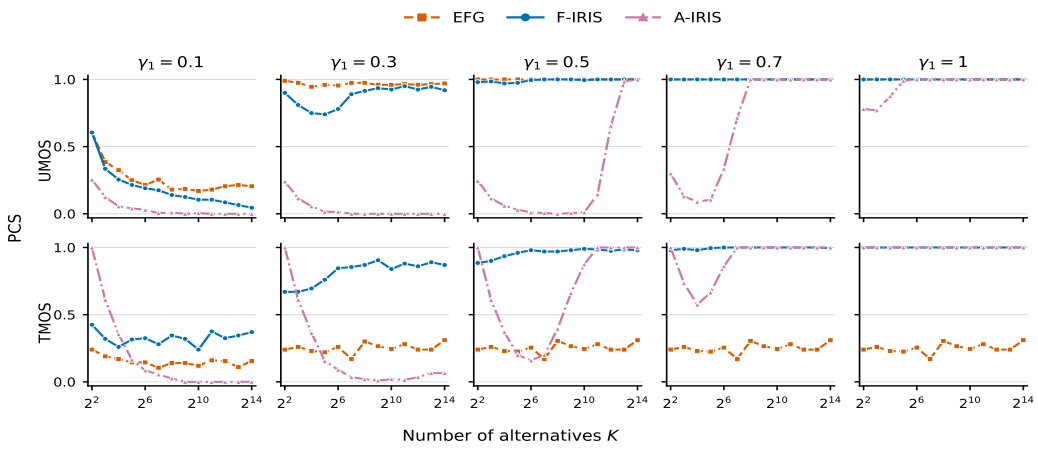

Figure 1: Experiment 1 with Gaussian midpoint observations.

Smaller midpoint gaps are more challenging. For $\gamma_1 = 0.1$ and 0.3, the PCS of A-IRIS declines toward zero over the tested range of K. At the initialization size $n_e = 80$, its confidence width is approximately 0.494, exceeding both gaps and leaving many inferior alternatives in the candidate set. F-IRIS shows a similar qualitative pattern to EFG in UMOS, while radius-based selection improves its performance in TMOS. The next experiment shows that increasing the per-alternative budget n with $\gamma_1 = 0.1$ substantially improves the PCS of both IRIS variants. The similar results under Gaussian, uniform, and skewed beta noise support robustness across the three tested distributions.

6.2 Experiment 2: Increasing the Budget per Alternative

Experiment 2 examines the consistency under the three midpoint distributions. We use TMOS with $K_1 \in \{4, 8\}$, $K \in \{2^4, 2^8, 2^{12}\}$, and $\gamma_1 = 0.1$. We vary n between 100 and 50,000, set $n_e = 0.8n$, and keep $\lambda = 0.375$. F-IRIS uses $\epsilon = 0.05$; the other algorithmic parameters and the radius model match Experiment 1.

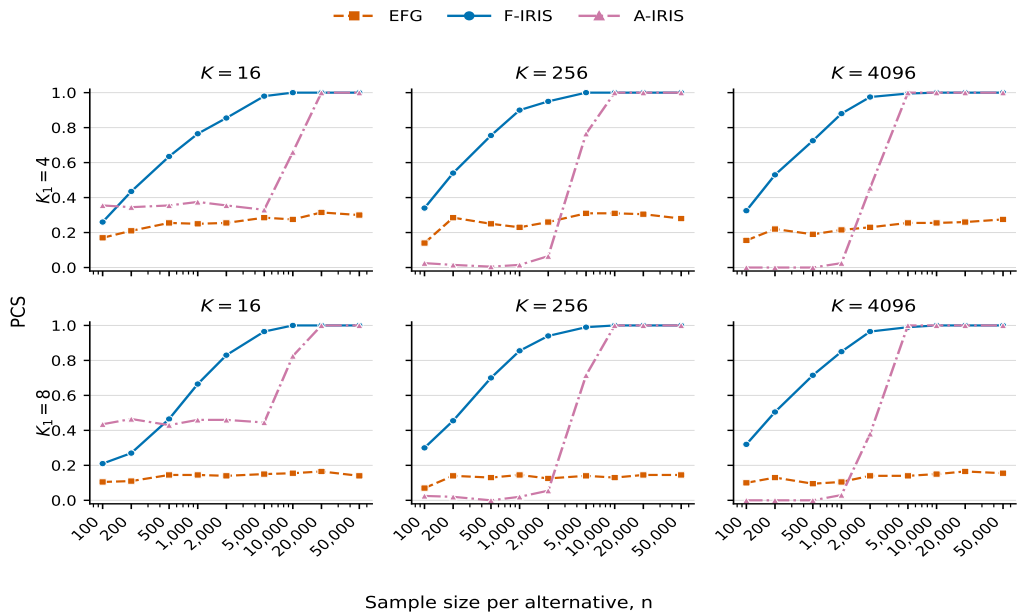

Figure 2: Experiment 2 under Gaussian midpoint observations.

Figure 2 shows that both IRIS procedures approach PCS one as the budget per alternative increases, for every displayed combination of K and K_1. F-IRIS improves earlier because it is supplied with a valid fixed tolerance. A-IRIS requires a larger budget before its confidence-based candidate set becomes selective, after which its PCS rises rapidly. EFG stays near $1/K_1$: additional midpoint samples cannot identify the smallest-radius alternative when the first K_1 population midpoints are equal. These results illustrate the quantifier order in Theorems 2 and 4: increasing the finite initialization and linear budget coefficients raises the asymptotic PCS, even though a small fixed coefficient can perform poorly at large K. Appendix C repeats the experiment with centered uniform and standardized Beta midpoint noise, respectively.

6.3 Case study: LLM-enabled laptop screening

We next test the algorithms on interval-valued responses from an LLM. Following the design of Li et al. (2024b), each alternative represents a laptop configuration, and the model is queried to simulate consumers' willingness to pay (WTP), offering a less costly alternative to survey-based elicitation. We ask `Llama-3.2-3B-Instruct` to return either a price range or a single price, which is treated as a degenerate interval. The four design spaces contain $K \in \{36, 360, 1080, 3240\}$ alternatives. We generated 4,000 valid responses per alternative on NVIDIA A100 GPUs and use these responses as a finite empirical population. The algorithms only observe the responses drawn within their assigned budget; the full population is used after each run to determine whether the selection is correct. Appendix D gives the prompts, hardware, candidate attributes, and evaluation protocol.

For evaluation, we regard alternatives within $\epsilon_{\text{eval}} = 30$ of the largest population mean midpoint as practically tied in WTP and define the ground truth as the member of this set with the smallest population mean radius. The tolerance of 30 corresponds to approximately 2.0%–2.6% of the largest mean midpoint across the four design spaces. We use this tolerance to treat alternatives with similar mean WTP as practically tied and then select the one with the smallest mean radius. To show the sensitivity to $\epsilon_{\text{eval}} = 30$, Appendix D.3 reports results over a grid of evaluation tolerances ϵ_{eval}. We compare EFG, F-IRIS, and A-IRIS at per-alternative budgets $n \in \{100, 200, 400, 800\}$, with $n_e = 0.8n$, $\lambda = 0.375$, and 500 paired macro-replications. F-IRIS uses the fixed tolerance $\epsilon = 30$. A-IRIS uses $\beta = 1$ and estimates $\bar{\sigma}_m$ via exploration samples.

A-IRIS improves on midpoint-only EFG without requiring a fixed tolerance matched to the evaluation target. At $n = 800$ and $K = 360$, its PCS increases from EFG's 0.008 to 0.768, a gain of 76.0 percentage points. At $K = 3240$, PCS increases from 0.028 to 0.144, a gain of 11.6 percentage points, or approximately 414% relative to EFG. A-IRIS also achieves PCS 0.076 at $K = 36$, where EFG records no correct selections, and performs similarly to EFG at $K = 1080$, where both attain PCS above 0.98. F-IRIS serves as a target-matched benchmark here: its fixed tolerance is set to the same value used to define the empirical ground truth. Appendix D reports the complete results.

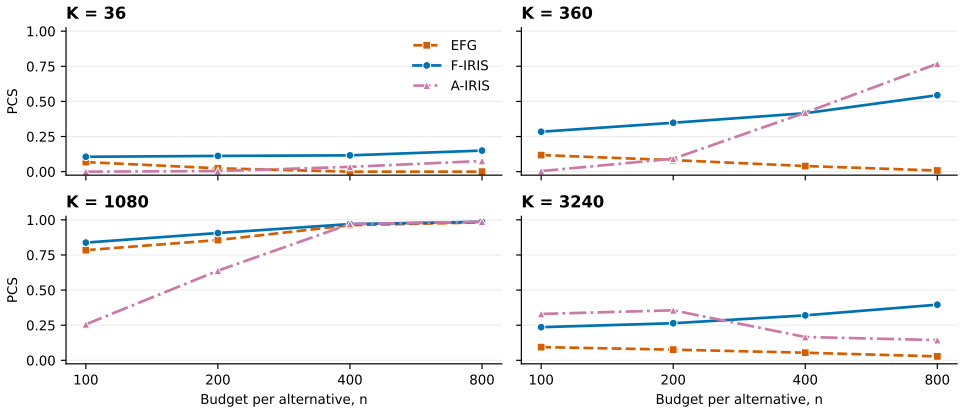

Figure 3: Empirical PCS for LLM-enabled laptop screening.

AI USE STATEMENT

Generative AI was used in two roles. First, `Llama-3.2-3B-Instruct` generated the interval-valued responses studied in the laptop-screening experiment; this use is part of the research method and is documented in Appendix D. Second, `GPT-5.6` supported parts of the related work literature search, code development, verification of mathematical derivations, and language polishing. The authors inspected the resulting code, reviewed every mathematical and empirical claim. The authors take responsibility for the final content of the paper.

REPRODUCIBILITY STATEMENT

The complete codebase for all experiments are provided as supplementary files. Appendix D documents the model, prompt, parsing rules, candidate spaces, and evaluation target for the LLM study. Appendices E and F provide complete proofs for the theorems.

REFERENCES

Gustaf Ahdritz, Tian Qin, Nikhil Vyas, Boaz Barak, and Benjamin L. Edelman. Distinguishing the knowable from the unknowable with language models. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 503–549. PMLR, 2024.

Gati V. Aher, Rosa I. Arriaga, and Adam Tauman Kalai. Using large language models to simulate multiple humans and replicate human subject studies. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 337–371. PMLR, 2023.

Sigrún Andradóttir and Judy S. Lee. Pareto set estimation with guaranteed probability of correct selection. *European Journal of Operational Research*, 292(1):286–298, 2021.

Javier Arroyo, Rosa Espínola, and Carlos Maté. Different approaches to forecast interval time series: A comparison in finance. *Computational Economics*, 37(2):169–191, 2011.

Jean-Yves Audibert, Sébastien Bubeck, and Rémi Munos. Best arm identification in multi-armed bandits. In *Proceedings of the 23rd Conference on Learning Theory*, pp. 41–53, 2010.

Chun-Hung Chen, Jianwu Lin, Enver Yücesan, and Stephen E. Chick. Simulation budget allocation for further enhancing the efficiency of ordinal optimization. *Discrete Event Dynamic Systems*, 10 (3):251–270, 2000.

Zhirui Chen, P. N. Karthik, Yeow Meng Chee, and Vincent Y. F. Tan. Optimal multi-objective best arm identification with fixed confidence. In *Proceedings of the 28th International Conference on Artificial Intelligence and Statistics*, volume 258 of *Proceedings of Machine Learning Research*, pp. 1162–1170, 2025.

Stephen E. Chick and Koichiro Inoue. New two-stage and sequential procedures for selecting the best simulated system. *Operations Research*, 49(5):732–743, 2001.

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.

Weiwei Fan, L. Jeff Hong, and Barry L. Nelson. Indifference-zone-free selection of the best. *Operations Research*, 64(6):1499–1514, 2016.

Weiwei Fan, L. Jeff Hong, and Xiaowei Zhang. Distributionally robust selection of the best. *Management Science*, 66(1):190–208, 2020.

Sebastian Farquhar, Jannik Kossen, Lorenz Kuhn, and Yarin Gal. Detecting hallucinations in large language models using semantic entropy. *Nature*, 630:625–630, 2024.

Peter I. Frazier. A fully sequential elimination procedure for indifference-zone ranking and selection with tight bounds on probability of correct selection. *Operations Research*, 62(4):926–942, 2014.

Aurélien Garivier and Emilie Kaufmann. Optimal best arm identification with fixed confidence. In *Proceedings of the 29th Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pp. 998–1027, 2016.

Ali Goli and Amandeep Singh. Frontiers: Can large language models capture human preferences? *Marketing Science*, 43(4):709–722, 2024.

Gloria González-Rivera and Wei Lin. Constrained regression for interval-valued data. *Journal of Business & Economic Statistics*, 31(4):473–490, 2013.

L. Jeff Hong, Weiwei Fan, and Jun Luo. Review on ranking and selection: A new perspective. *Frontiers of Engineering Management*, 8(3):321–343, 2021.

L. Jeff Hong, Guangxin Jiang, and Ying Zhong. Solving large-scale fixed-budget ranking and selection problems. *INFORMS Journal on Computing*, 34(6):2930–2949, 2022.

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, 2012.

Zohar Karnin, Tomer Koren, and Oren Somekh. Almost optimal exploration in multi-armed bandits. In *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pp. 1238–1246, 2013.

Gregory Keslin, Barry L. Nelson, Bernardo Pagnoncelli, Matthew Plumlee, and Hamed Rahimian. Ranking and contextual selection. *Operations Research*, 73(5):2695–2707, 2024.

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.

Taeho Kim, Kyoung-Kuk Kim, and Eunhye Song. Selection of the most probable best. *Operations Research*, 73(6):3199–3218, 2025.

Lorenz Kuhn, Yarin Gal, and Sebastian Farquhar. Semantic uncertainty: Linguistic invariances for uncertainty estimation in natural language generation. In *International Conference on Learning Representations*, 2023.

Yan Leng and Yuan Yuan. SUVA: A probabilistic framework for auditing LLMs with an application to social preferences. *Information Systems Research*, 2026. Articles in Advance.

Peiyao Li, Noah Castelo, Zsolt Katona, and Miklos Sarvary. Frontiers: Determining the validity of large language models for automated perceptual analysis. *Marketing Science*, 43(2):254–266, 2024a.

Zaile Li, Weiwei Fan, and L. Jeff Hong. Efficient budget allocation for large-scale LLM-enabled virtual screening. *arXiv preprint*, 2024b. URL `https://arxiv.org/abs/2408.09537`.

Zaile Li, Weiwei Fan, and L. Jeff Hong. The (surprising) sample optimality of greedy procedures for large-scale ranking and selection. *Management Science*, 71(2):1238–1259, 2025.

Lei Liu, Ya-Chen Tina Shih, Robert L. Strawderman, Daowen Zhang, Bankole A. Johnson, and Haitao Chai. Statistical analysis of zero-inflated nonnegative continuous data: A review. *Statistical Science*, 34(2):253–279, 2019.

Eric C. Ni, Dragos F. Ciocan, Shane G. Henderson, and Susan R. Hunter. Efficient ranking and selection in parallel computing environments. *Operations Research*, 65(3):821–836, 2017.

Edward Paulson. A sequential procedure for selecting the population with the largest mean from k normal populations. *The Annals of Mathematical Statistics*, 35(1):174–180, 1964.

Xin Qiu and Risto Miikkulainen. Semantic density: Uncertainty quantification for large language models through confidence measurement in semantic space. In A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang (eds.), *Advances in Neural Information Processing Systems*, volume 37, pp. 134507–134533. Curran Associates, Inc., 2024.

Yosef Rinott. On two-stage selection procedures and related probability-inequalities. *Communications in Statistics - Theory and Methods*, 7(8):799–811, 1978.

Sebastian Rojas-Gonzalez, Juergen Branke, and Inneke Van Nieuwenhuyse. Bi-objective ranking and selection using stochastic kriging. *European Journal of Operational Research*, 322(2):599–614, 2025.

Atanu Sengupta and Tapan Kumar Pal. On comparing interval numbers. *European Journal of Operational Research*, 127(1):28–43, 2000.

Chengshuai Shi, Kun Yang, Zihan Chen, Jundong Li, Jing Yang, and Cong Shen. Efficient prompt optimization through the lens of best arm identification. In *Advances in Neural Information Processing Systems*, volume 37, pp. 99646–99685, 2024.

Yuying Sun, Xinyu Zhang, Alan T.K. Wan, and Shouyang Wang. Model averaging for interval-valued data. *European Journal of Operational Research*, 301(2):772–784, 2022.

Mengxin Wang, Dennis J. Zhang, and Heng Zhang. Large language models for market research: A data-augmentation approach. *Marketing Science*, 45(4):728–751, 2026a.

Wen Wang, Siqi Pei, and Tianshu Sun. Unraveling generative AI from a human intelligence perspective: A battery of experiments. *Information Systems Research*, 2026b. Articles in Advance.

Wenyu Wang, Hong Wan, and Xi Chen. Bonferroni-free and indifference-zone-flexible sequential elimination procedures for ranking and selection. *Operations Research*, 72(5):2119–2134, 2024.

Di Wu, Yuhao Wang, and Enlu Zhou. Data-driven ranking and selection under input uncertainty. *Operations Research*, 72(2):781–795, 2024.

Bo Xue, Yuanyu Wan, Zhichao Lu, and Qingfu Zhang. Beyond the lower bound: Bridging regret minimization and best arm identification in lexicographic bandits. *Proceedings of the AAAI Conference on Artificial Intelligence*, 40(32):27414–27422, 2026.

Wei Zhong, Chen Qian, Wanjun Liu, Liping Zhu, and Runze Li. Feature screening for interval-valued response with application to study association between posted salary and required skills. *Journal of the American Statistical Association*, 118(542):805–817, 2023.

Ying Zhong and L. Jeff Hong. Knockout-tournament procedures for large-scale ranking and selection in parallel computing environments. *Operations Research*, 70(1):432–453, 2022.

A Additional Related Work

A.1 Classical and large-scale ranking and selection

R&S allocates simulation effort across a finite set of stochastic alternatives and returns a final selection. Fixed-precision procedures seek a prescribed PCS, whereas fixed-budget procedures seek the best decision attainable with a given number of samples (Hong et al., 2021). The fixed-precision literature includes early sequential and two-stage procedures (Paulson, 1964; Rinott, 1978), sequential elimination methods such as KN and BIZ (Kim & Nelson, 2001; Frazier, 2014), and procedures that relax a fixed indifference-zone specification (Fan et al., 2016; Wang et al., 2024). In fixed-budget R&S, OCBA uses estimated means and variances to focus simulation effort on alternatives that matter for PCS (Chen et al., 2000); Bayesian procedures can also account for opportunity cost and unequal sampling cost (Chick & Inoue, 2001).

Large candidate sets motivate algorithms whose sample requirement grows linearly in K. Median elimination gives a linear-order PAC guarantee for approximate selection (Even-Dar et al., 2006). Parallel good-selection and knockout procedures address large-scale simulation settings (Ni et al., 2017; Zhong & Hong, 2022; Hong et al., 2022). Li et al. (2025) provide a boundary-crossing analysis of greedy and explore-first greedy procedures and show sample optimality under a positive mean gap. Their target is the largest scalar mean. IRIS instead has an exact K_1-way maximum in the primary component and a secondary ordering within that set. The proof must control both the

workload generated by midpoint-inferior alternatives and the radius competition among midpoint-optimal alternatives.

Other R&S extensions address robust objectives, input uncertainty, observed covariates, or multiple objectives (Fan et al., 2020; Wu et al., 2024; Kim et al., 2025; Keslin et al., 2024; Andradóttir & Lee, 2021; Rojas-Gonzalez et al., 2025). These problems alter either the definition of an alternative's value or the information structure. None uses the strict midpoint-then-radius preference studied here.

A.2 Pure exploration in machine learning

BAI is the bandit counterpart of fixed-budget R&S. Successive rejects and sequential halving discard arms in rounds and obtain near-optimal fixed-budget rates (Audibert et al., 2010; Karnin et al., 2013). In the fixed-confidence setting, LUCB samples empirical boundary arms through confidence bounds (Kalyanakrishnan et al., 2012), while Track-and-Stop tracks information-theoretically optimal sampling proportions (Garivier & Kaufmann, 2016). Top-m and structured variants change the decision class, but still derive a scalar value for each arm. IRIS is closer to constrained pure exploration: the radius criterion is relevant only on the unknown population argmax set of the midpoint criterion.

Lexicographic bandits also impose an order on vector-valued objectives. Xue et al. (2026) connect regret minimization with fixed-confidence BAI and analyze elimination procedures that use this priority structure. IRIS differs in three respects: it uses a fixed budget, allows several arms to share the optimal primary mean, and asks whether a budget linear in K can retain a nonvanishing PCS. In multi-objective BAI, Chen et al. (2025) identify the best arm for every objective under a fixed-confidence criterion. Their output is an objective-indexed collection of arms; IRIS seeks one arm under a strict midpoint-then-radius preference.

This connection is also relevant to LLM systems. Shi et al. (2024) formulate prompt selection from a generated pool as fixed-budget BAI, and Li et al. (2024b) use large-scale R&S for LLM-enabled virtual screening. Both optimize scalar evaluations. Our formulation covers responses for which the model supplies a numerical range.

A.3 Interval responses and LLM-generated judgments

Interval-valued data appear in posted salaries and daily price ranges; reducing these observations to one endpoint or their center can remove information relevant to estimation and forecasting (Zhong et al., 2023; Arroyo et al., 2011). Center-and-range models treat the two components separately and enforce valid endpoint ordering (Sun et al., 2022; González-Rivera & Lin, 2013). In our LLM data, exact scalar responses produce a mass at radius zero, while interval responses produce a right-skewed positive component. This is compatible with semicontinuous or zero-inflated modeling (Liu et al., 2019), although IRIS itself does not fit a parametric radius model.

Uncertainty measures for LLM outputs answer a different question. Semantic uncertainty, semantic entropy, and semantic density aggregate variation over meanings or estimate confidence in a generated answer (Kuhn et al., 2023; Farquhar et al., 2024; Qiu & Miikkulainen, 2024). Work separating epistemic and aleatoric uncertainty likewise concerns the source of predictive uncertainty (Ahdritz et al., 2024). IRIS uses interval width as a measure of response imprecision and incorporates it into the selection objective: we first maximize expected midpoint and then, among midpoint-optimal alternatives, minimize expected radius.

Studies comparing LLM and human judgments suggest that LLMs can provide a low-cost supplement to human preference elicitation, although their agreement with human responses varies across tasks, models, and elicitation protocols (Li et al., 2024a; Goli & Singh, 2024; Wang et al., 2026a;b; Aher et al., 2023; Leng & Yuan, 2026). Despite possible biases, model-generated responses allow researchers to screen large candidate pools before conducting more costly human surveys. Motivated by this use, our empirical study examines laptop selection using repeated LLM-generated WTP responses, treating the collected outputs as a finite empirical population for evaluating the selection procedures.

B EXPERIMENT 1 WITH UNIFORM AND BETA MIDPOINT NOISE

Figures 4-5 show the PCS results of Experiment 1 under uniform and beta distribution.

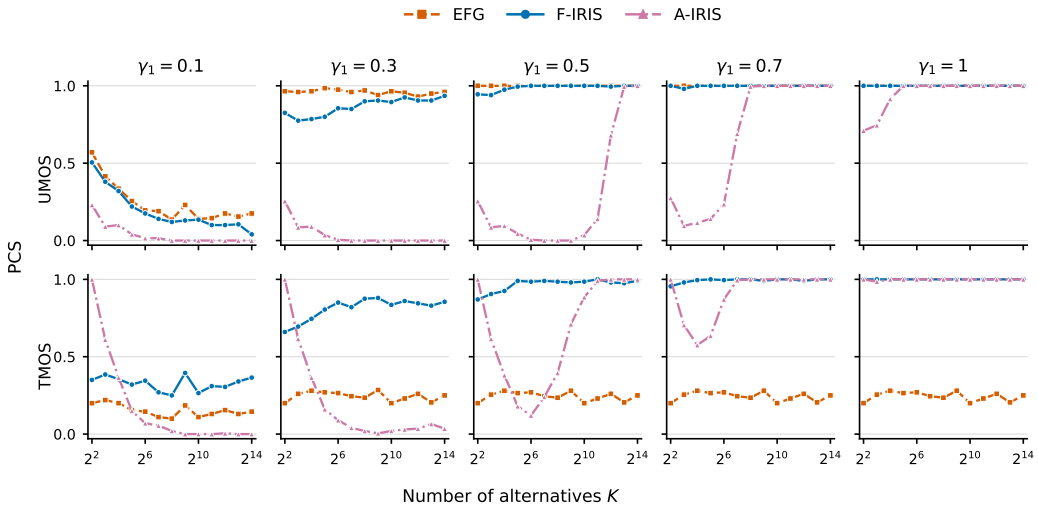

Figure 4: Experiment 1 with centered uniform midpoint observations of variance one.

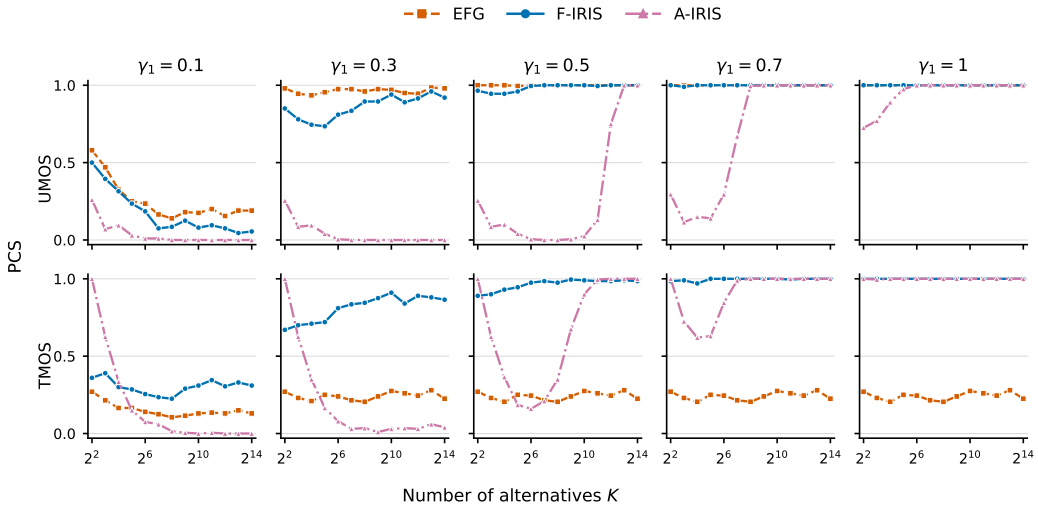

Figure 5: Experiment 1 with centered, standardized $\mathrm{Beta}(2, 5)$ midpoint observations.

C EXPERIMENT 2 WITH UNIFORM AND BETA MIDPOINT NOISE

We repeat Experiment 2 after replacing the Gaussian midpoint noise with $\mathrm{Unif}(-\sqrt{3}, \sqrt{3})$. The noise is centered and has variance one, so this change preserves the midpoint means, variances, budget, and algorithmic parameters used in the main experiment.

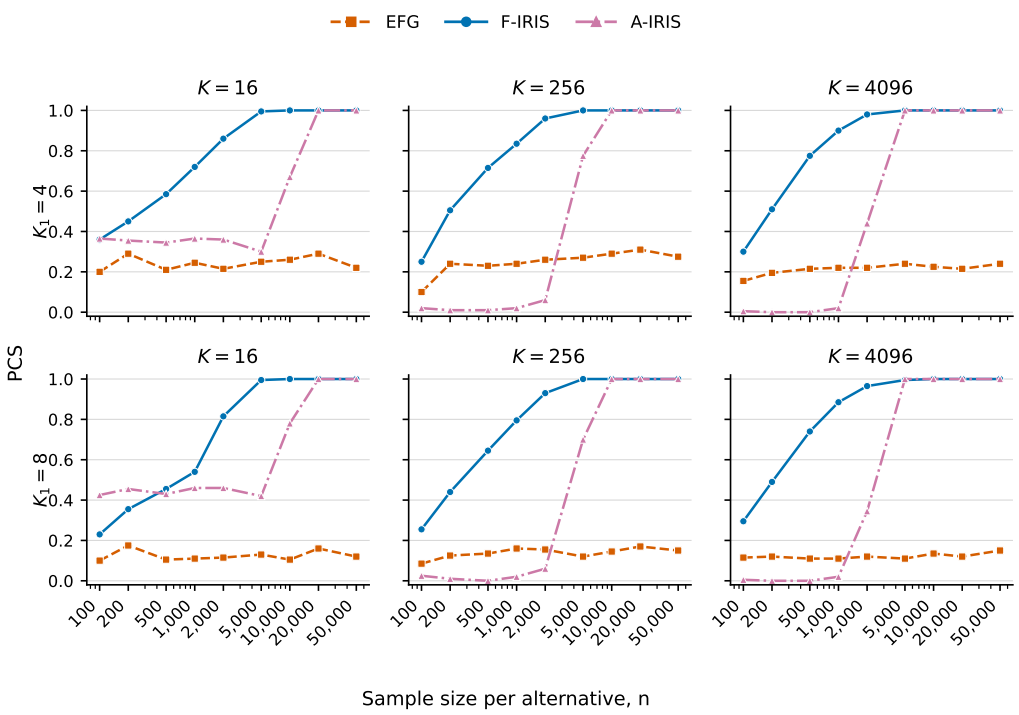

Figure 6: Experiment 2 with centered uniform midpoint noise of variance one.

Figure 6 closely matches the Gaussian result in Figure 2. F-IRIS improves steadily with n, whereas A-IRIS begins to improve only after the confidence-based candidate set becomes sufficiently selective. Both procedures attain PCS one at the largest budgets for every displayed (K, K_1) pair. EFG remains near $1/K_1$ because midpoint samples alone cannot distinguish the K_1 midpoint-optimal alternatives. Thus, the consistency pattern in Experiment 2 is not specific to Gaussian midpoint observations.

We also repeat Experiment 2 using a centered, variance-one transformation of $\mathrm{Beta}(2, 5)$ midpoint noise. This distribution is bounded and asymmetric, unlike the Gaussian and centered uniform distributions.

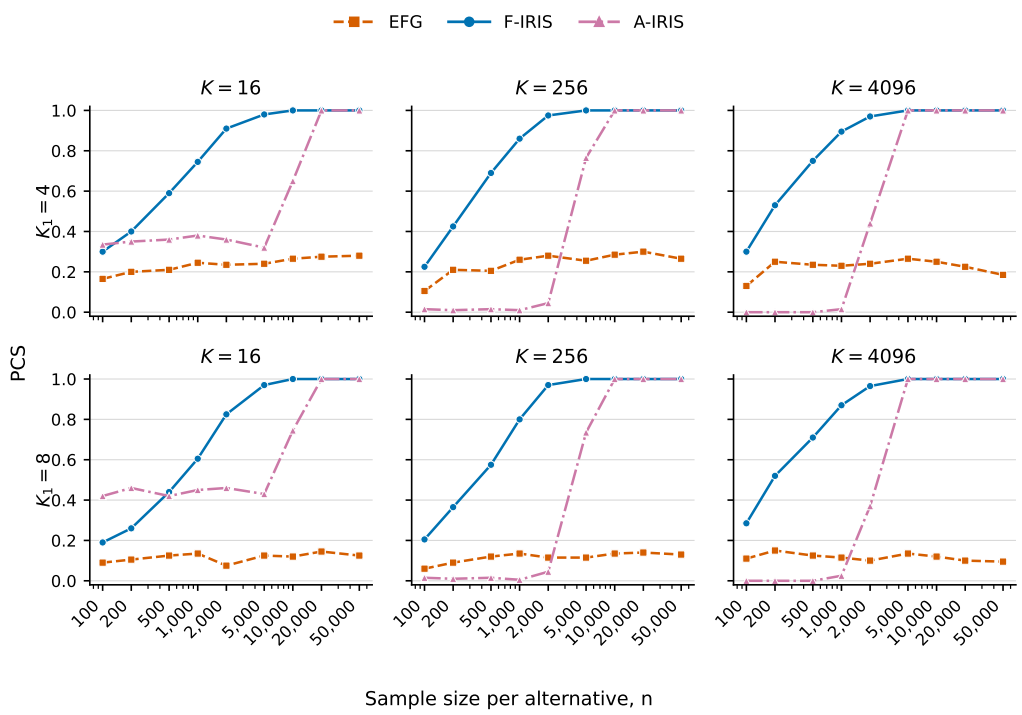

Figure 7: Experiment 2 with centered, standardized $\mathrm{Beta}(2, 5)$ midpoint noise.

Figure 7 again exhibits the consistency pattern seen under Gaussian and uniform midpoint noise. F-IRIS reaches high PCS at smaller budgets, while A-IRIS has a longer transient region before rising rapidly to one. EFG remains near $1/K_1$. The agreement across all three distributions indicates that the observed convergence is not tied to symmetry or Gaussian tails.

D Details for the LLM-Enabled Screening Experiment

D.1 Response generation and candidate spaces

We adapt the laptop WTP application of Li et al. (2024b), while changing the elicited response from a scalar to an interval. The system message is

> You are a customer. You are selected at random while shopping for laptops to participate in a survey.

For each configuration, the complete user message is:

> The customer is asked: What is the maximum price you would be willing to pay for a Lenovo laptop with {`candidate attributes`}?
>
> Please respond with either a plausible price range or a single price. When more than one price seems reasonable, prefer a price range. Use a single price only when one specific price seems clearly appropriate. Give numbers only, using one of these formats: 800-1000 or 900 (no descriptions).

The placeholder {`candidate attributes`} is replaced by the configuration's attribute values and labels: {`cpu`} `CPU`, {`ram`} `RAM`, and {`storage`} `Storage Drive`, followed by {`gpu`} `GPU` for $K \geq 360$, {`resolution`} `Screen Display Resolution` for $K \geq 1080$, and {`screen size`} `Screen Size` for $K = 3240$. The attributes are separated by commas, with "and" before the final attribute. Each query describes only one candidate configuration. For a parsed response $[L, U]$, we store the midpoint $(L + U)/2$ and radius $(U - L)/2$. A scalar p is stored as

$[p, p]$. Ambiguous or invalid generations are rejected and queried again, up to 20 attempts, rather than imputed.

The candidate spaces are nested Cartesian products. For $K = 36$, the attributes are six CPUs (Intel Core i5/i7/i9 and AMD-R5/R7/R9), three RAM levels (16, 32, and 64 GB), and two storage levels (256 and 512 GB). For $K = 360$, we use the same six CPUs, four RAM levels (8–64 GB), three storage levels (256–1024 GB), and five GPU labels. The $K = 1080$ space adds three display resolutions, and the $K = 3240$ space adds three screen sizes. Attribute labels and Cartesian-product structure follow Li et al. (2024b).

We use the Hugging Face implementation of `Llama-3.2-3B-Instruct` in bfloat16. Generation uses temperature 1, a maximum of 32 new tokens, and batches of 32 prompts. Each collection shard runs on one NVIDIA A100 GPU with four CPU cores and 32 GB of host memory. We retain 4,000 valid responses per alternative, giving 144,000, 1,440,000, 4,320,000, and 12,960,000 interval observations for $K = 36$, 360, 1080, and 3240, respectively.

D.2 Evaluation and algorithm settings

The 4,000 responses for alternative i define a finite empirical distribution with population means $\widetilde{\mu}_i^m$ and $\widetilde{\rho}_i$. For an evaluation tolerance ϵ_{eval}, let

$$\widetilde{G}(\epsilon_{\text{eval}}) = \left\{ i : \widetilde{\mu}_i^m \geq \max_j \widetilde{\mu}_j^m - \epsilon_{\text{eval}} \right\}, \tag{7}$$

$$\widetilde{i}^*(\epsilon_{\text{eval}}) \in \arg \min_{i \in \widetilde{G}(\epsilon_{\text{eval}})} \widetilde{\rho}_i. \tag{8}$$

PCS is the fraction of runs that select $\widetilde{i}^*(\epsilon_{\text{eval}})$. Thus, ϵ_{eval} defines a single empirical target and does not change the 0–1 PCS metric. At the primary value $\epsilon_{\text{eval}} = 30$, the competitive sets contain 3, 3, 1, and 5 alternatives for $K = 36$, 360, 1080, and 3240. That is, we treat the top 3, 3, 1, and 5 designs as candidates with highest and close WTP population mean and the target candidate has smallest radius. This setting includes both UMOS ($K_1 = 1$) and TMOS ($K_1 > 1$).

Each method receives total budget $B = nK$, where $n \in \{100, 200, 400, 800\}$, and first draws $n_e = 0.8n$ observations from every alternative. EFG allocates the remaining budget to the largest current midpoint. F-IRIS uses $\epsilon = 30$ and A-IRIS uses $\beta = 1$. For both IRIS methods, $\lambda = 0.375$ of the post-initialization epochs update the midpoint leader; the other epochs sample by radius within the method's current candidate set. A-IRIS estimates $\bar{\sigma}_m$ separately in each macro-replication: it computes one midpoint standard deviation per alternative from the n_e current observations and uses their 0.95 quantile. This estimate is then fixed for the remainder of that replication. No full-population observation enters an algorithmic decision. We use 500 macro-replications to evaluate PCS. Table 1 gives the numerical values underlying Figure 3.

D.3 Sensitivity to the empirical target

The threshold controls which near-maximal mean midpoints are treated as practically equivalent, so the target can change discontinuously with ϵ_{eval}. Figure 8 reports the full sensitivity calculation. The dashed line marks the primary value 30, and dotted lines mark tolerances at which the radius-optimal empirical target changes. The results are quite robust to the selection of truth epsilon. When ϵ_{eval} is small or 0, the midpoint optimal candidate set is singular, the problem reduces to Top-1 selection. Hence EFG performs better as budget is not enough. When truth epsilon is modest (i.e., 20-60), IRIS outperforms EFG because IRIS uses extra width information. When the true epsilon is sufficiently large, both IRIS and EFG fail to select the target when $K = 3,240$. A possible reason is that a larger tolerance admits alternatives with lower mean midpoints, which may become the evaluation target because of their smaller mean radii.

Table 1: PCS in the LLM-enabled screening experiment at $\epsilon_{\text{eval}} = 30$.

K	n	n_e	EFG	F-IRIS	A-IRIS
36	100	80	0.068	**0.106**	0.000
36	200	160	0.024	**0.112**	0.004
36	400	320	0.000	**0.116**	0.034
36	800	640	0.000	**0.150**	0.076
360	100	80	0.118	**0.284**	0.004
360	200	160	0.082	**0.348**	0.092
360	400	320	0.040	0.416	**0.420**
360	800	640	0.008	0.544	**0.768**
1080	100	80	0.784	**0.838**	0.256
1080	200	160	0.856	**0.906**	0.638
1080	400	320	0.962	**0.970**	**0.970**
1080	800	640	0.982	**0.986**	**0.986**
3240	100	80	0.094	0.236	**0.330**
3240	200	160	0.076	0.264	**0.356**
3240	400	320	0.054	**0.320**	0.166
3240	800	640	0.028	**0.396**	0.144

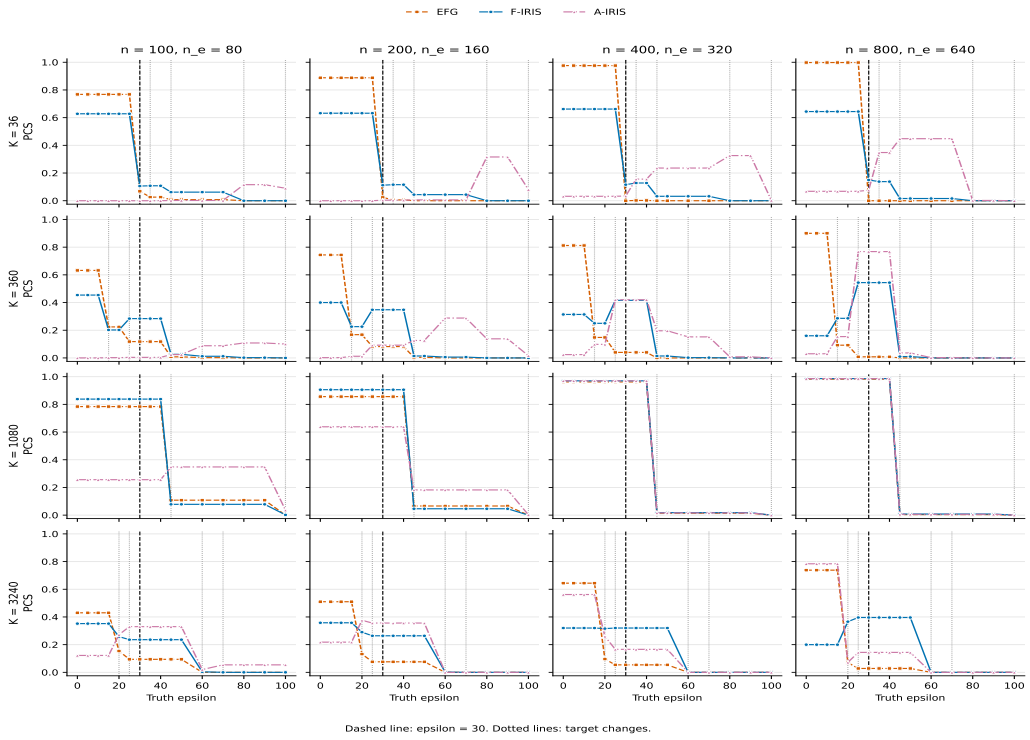

Figure 8: PCS sensitivity to the evaluation tolerance. Dotted vertical lines mark values at which the radius-optimal target changes; the dashed line marks $\epsilon_{\text{eval}} = 30$.

E PROOFS FOR FIXED-TOLERANCE IRIS

We first isolate two events involving only the fixed set G_m. Define

$$A_\delta := \bigcap_{j=1}^{K_1} \bigcap_{n \geq n_e} \{|\bar{X}_j^m(n) - \mu_*^m| \leq \delta\},$$

$$A_r(n_e) := \left\{ \sup_{n \geq n_e} \bar{X}_1^r(n) \leq \mu_1^r + \delta_1 \right\} \cap \bigcap_{j=2}^{K_1} \left\{ \inf_{n \geq n_e} \bar{X}_j^r(n) > \mu_1^r + \delta_1 \right\},$$

where $A_r(n_e) = \Omega$ when $K_1 = 1$.

Lemma 1. *Under Assumptions 1–3,* $\Pr(A_\delta) \geq p_m(n_e, \delta)$ *and* $\Pr(A_r(n_e)) = p_r(n_e) > 0$. *Moreover,* $p_m(n_e, \delta) \to 1$ *and* $p_r(n_e) \to 1$ *as* $n_e \to \infty$.

Proof. For $j \leq K_1$, the sub-Gaussian bound gives $\Pr(|\bar{X}_j^m(n) - \mu_*^m| > \delta) \leq 2e^{-n\delta^2/(2\bar{\sigma}_m^2)}$. Hence, by a union bound, $\Pr(A_\delta^c) \leq 2K_1 \sum_{n=n_e}^{\infty} e^{-n\delta^2/(2\bar{\sigma}_m^2)} = 1 - p_m(n_e, \delta)$. The geometric series also implies $p_m(n_e, \delta) \to 1$.

To prove the radius claim, first consider i.i.d. integrable increments $(Y_\ell)_{\ell \geq 1}$ with $\mathbb{E}Y_1 < 0$, and write $S_n = \sum_{\ell=1}^n Y_\ell$. Choose $\eta > 0$ such that $p := \Pr(Y_1 \leq -\eta) > 0$. For $m \geq 1$, define $M_m := \sup_{q \geq 0} \sum_{\ell=m+1}^{m+q} Y_\ell$, where the sum is zero when $q = 0$. The strong law gives $M_m < \infty$ almost surely. Since the distribution of M_m does not depend on m, there exists m such that $\Pr(M_m < m\eta) > 0$.

Let $B_m := \bigcap_{\ell=1}^m \{Y_\ell \leq -\eta\} \cap \{M_m < m\eta\}$. Independence yields $\Pr(B_m) = p^m \Pr(M_m < m\eta) > 0$. On B_m, we have $S_n \leq -n\eta < 0$ for $n \leq m$, and $S_n \leq -m\eta + M_m < 0$ for $n > m$. Together with $S_n/n \to \mathbb{E}Y_1 < 0$, this implies $\sup_{n \geq 1} S_n/n < 0$ almost surely on B_m. Therefore, $\Pr(\sup_{n \geq 1} S_n/n < 0) > 0$.

Set $a := \mu_1^r + \delta_1$. For alternative 1, take $Y_\ell = X_{1,\ell}^r - a$, so that $\mathbb{E}Y_\ell = -\delta_1 < 0$. For $2 \leq j \leq K_1$, take $Y_\ell = a - X_{j,\ell}^r$, so that $\mathbb{E}Y_\ell \leq \delta_1 - \gamma_2 < 0$. The preceding argument gives $\Pr(\sup_{n \geq 1} \bar{X}_1^r(n) < a) > 0$ and $\Pr(\inf_{n \geq 1} \bar{X}_j^r(n) > a) > 0$. Independence across alternatives then implies $p_r(n_e) = \Pr(A_r(n_e)) > 0$ for every $n_e \geq 1$. On $A_r(n_e)$, $\bar{X}_1^r(n) \leq a < \bar{X}_j^r(n')$ for all $j \geq 2$ and $n, n' \geq n_e$.

Finally, the strong law gives $\bar{X}_1^r(n) \to \mu_1^r < a$ and $\bar{X}_j^r(n) \to \mu_j^r > a$. Since K_1 is fixed, $A_r(n_e) \uparrow \Omega$ up to a null set as $n_e \to \infty$, and thus $p_r(n_e) \to 1$. When $K_1 = 1$, the radius claims follow directly from $A_r(n_e) = \Omega$. □

For $k > K_1$, define the permanent midpoint exit time

$$\tau_k^\epsilon := \inf\{n \geq n_e : \bar{X}_k^m(\ell) < \mu_*^m - \delta - \epsilon \text{ for every } \ell \geq n\}.$$

Lemma 2 (Inferior midpoint workload)**.** *For every* $k > K_1$,

$$\Pr(\tau_k^\epsilon > n) \leq \frac{e^{-a_{\epsilon,\delta} n}}{1 - e^{-a_{\epsilon,\delta}}}, \qquad \mathbb{E}[(\tau_k^\epsilon - n_e)_+] \leq C_{\epsilon,\delta}(n_e).$$

On A_δ, *neither channel of Algorithm 1 samples* k *after its count reaches* τ_k^ϵ. *Consequently, with* $Z_K^\epsilon := \sum_{k > K_1} (\tau_k^\epsilon - n_e)_+$,

$$\mathbb{E}[Z_K^\epsilon \mid A_\delta \cap A_r(n_e)] \leq (K - K_1) C_{\epsilon,\delta}(n_e).$$

Proof. Write $a := a_{\epsilon,\delta}$. For $k > K_1$ and $n \geq n_e$, $\{\tau_k^\epsilon > n\} \subseteq \bigcup_{\ell \geq n} \{\bar{X}_k^m(\ell) \geq \mu_*^m - \delta - \epsilon\}$. Since $\mu_k^m \leq \mu_*^m - \gamma_1$, the sub-Gaussian bound and a union bound give

$$\Pr(\tau_k^\epsilon > n) \leq \sum_{\ell=n}^{\infty} \Pr\big(\bar{X}_k^m(\ell) - \mu_k^m \geq \gamma_1 - \delta - \epsilon\big) \leq \sum_{\ell=n}^{\infty} e^{-a\ell} = \frac{e^{-an}}{1 - e^{-a}}.$$

Consequently,

$$\mathbb{E}[(\tau_k^\epsilon - n_e)_+] = \sum_{n=n_e}^{\infty} \Pr(\tau_k^\epsilon > n) \le \frac{e^{-an_e}}{(1-e^{-a})^2} = C_{\epsilon,\delta}(n_e).$$

On A_δ, if $n_k(t) \ge \tau_k^\epsilon$, then $\bar{X}_k^m(n_k(t)) < \mu_*^m - \delta - \epsilon$, whereas $\min_{j \le K_1} \bar{X}_j^m(n_j(t)) \ge \mu_*^m - \delta$. Thus, $\bar{X}_k^m(n_k(t)) < \min_{j \le K_1} \bar{X}_j^m(n_j(t)) - \epsilon$, which implies $k \notin \mathcal{C}_t^\epsilon$ and $k \notin \arg\max_i \bar{X}_i^m(n_i(t))$. Neither channel therefore samples k once its count reaches τ_k^ϵ.

Finally, Z_K^ϵ depends only on the midpoint paths of alternatives outside G_m, while $A_\delta \cap A_r(n_e)$ depends only on paths within G_m. Assumption 2 implies $Z_K^\epsilon \perp\!\!\!\perp (A_\delta \cap A_r(n_e))$. Hence, whenever $\Pr(A_\delta \cap A_r(n_e)) > 0$,

$$\mathbb{E}[Z_K^\epsilon \mid A_\delta \cap A_r(n_e)] = \mathbb{E}[Z_K^\epsilon] = \sum_{k > K_1} \mathbb{E}[(\tau_k^\epsilon - n_e)_+] \le (K - K_1) C_{\epsilon,\delta}(n_e).$$

□

Let $H_K := T_K - \lfloor \lambda T_K \rfloor$ be the number of radius epochs, and let D_K count the radius epochs at which the recorded recommendation is not alternative 1.

Lemma 3 (Incorrect-vote occupation)**.** *There are at most $L_\lambda - 1$ consecutive radius epochs and $H_K/K \to (1-\lambda)(c - n_e)$. On $A_\delta \cap A_r(n_e)$,*

$$D_K \le L_\lambda (Z_K^\epsilon + 1).$$

If $D_K < H_K/2$, Algorithm 1 selects alternative 1.

Proof. Let $M_K := \lfloor \lambda T_K \rfloor$, and write $m_1 < \cdots < m_{M_K}$ for the midpoint epochs, with $m_0 := 0$. Define $I_j := \{m_{j-1} + 1, \ldots, m_j\}$ for $j = 1, \ldots, M_K$ and the terminal block $I_{M_K+1} := \{m_{M_K} + 1, \ldots, T_K\}$. Let $R_j := \{s \in I_j : a_s = 0\}$. The deterministic schedule gives $|R_j| \le L_\lambda - 1$. Moreover, $H_K = T_K - \lfloor \lambda T_K \rfloor$ and $H_K/K \longrightarrow (1-\lambda)(c - n_e)$.

For each block, define $d_j := \sum_{s \in R_j} \mathbf{1}\{q_s \ne 1\}$ and $z_j := \sum_{s \in I_j} \mathbf{1}\{q_s \notin G_m\}$. We next show that, on $A_\delta \cap A_r(n_e)$, for $j \le M_K$,

$$d_j > 0 \quad \Longrightarrow \quad z_j \ge 1.$$

The reason is as follows. Fix a nonterminal block with $d_j > 0$, and choose an incorrect radius epoch $s \in R_j$. If $1 \in \mathcal{C}_s^\epsilon$, then $A_r(n_e)$ implies $\bar{X}_1^r(n_1(s)) < \bar{X}_k^r(n_k(s))$ for any $k \in G_m \setminus \{1\}$. Hence the radius-greedy rule gives that $q_s \ne 1$ results in $q_s \notin G_m$, and hence $z_j \ge 1$. Otherwise, $1 \notin \mathcal{C}_s^\epsilon$. On A_δ, $\mu_*^m - \delta \le \bar{X}_k^m(n_k(t)) \le \mu_*^m + \delta$ for $k \in G_m$. Thus

$$\max_i \bar{X}_i^m(n_i(s)) > \bar{X}_1^m(n_1(s)) + \epsilon \ge \mu_*^m - \delta + \epsilon > \mu_*^m + \delta,$$

where the last inequality uses $2\delta < \epsilon$. Choose a leader $b \notin G_m$ attaining this maximum. If any inferior alternative is sampled between epochs s and $m_j - 1$, then $z_j \ge 1$. If none is sampled, the empirical midpoint of b remains unchanged, and therefore

$$\bar{X}_b^m(n_b(m_j)) > \mu_*^m + \delta \ge \max_{k \in G_m} \bar{X}_k^m(n_k(m_j)).$$

Consequently, the midpoint epoch m_j samples an inferior alternative, again giving $z_j \ge 1$.

It follows that $d_j \le (L_\lambda - 1)\mathbf{1}\{d_j > 0\} \le L_\lambda z_j$ for $j \le M_K$ and $d_{M_K+1} \le L_\lambda - 1 \le L_\lambda$ for the terminal block. Using the inferior-sampling bound from the preceding lemma,

$$D_K = \sum_{j=1}^{M_K+1} d_j \le L_\lambda \left(\sum_{j=1}^{M_K} z_j + 1 \right) \le L_\lambda (Z_K^\epsilon + 1).$$

Finally, the vote counts satisfy $V_1 = H_K - D_K$ and $\max_{k \ne 1} V_k \le \sum_{k \ne 1} V_k = D_K$. Thus, if $D_K < H_K/2$,

$$V_1 = H_K - D_K > D_K \ge \max_{k \ne 1} V_k,$$

and Algorithm 1 returns alternative 1. □

Proof of Theorem 1. By conditional Markov's inequality and Lemmas 2 and 3,

$$\limsup_{K\to\infty} \Pr\left(D_K \geq H_K/2 \mid A_\delta \cap A_r(n_e)\right) \leq \frac{2L_\lambda C_{\epsilon,\delta}(n_e)}{(1-\lambda)(c-n_e)}.$$

Write $q_\delta := \delta^2/(2\bar{\sigma}_m^2)$. The lower bound on n_e gives

$$1 - p_m(n_e,\delta) = \frac{2K_1 e^{-q_\delta n_e}}{1-e^{-q_\delta}} \leq \eta, \text{ and } p_r(n_e) \geq 1-\eta.$$

The second inequality follows from $n_e \geq n_r(\eta)$. Hence the Fréchet bound yields

$$\begin{aligned}\Pr(A_\delta \cap A_r(n_e)) &\geq \Pr(A_\delta) + \Pr(A_r(n_e)) - 1 \\ &\geq p_m(n_e,\delta) + p_r(n_e) - 1 \\ &= \pi_{\mathrm{F}}(n_e,\delta) \geq 1 - 2\eta > 0.\end{aligned}$$

Combining this inequality with the conditional success bound proves equation 4. The budget condition makes its second factor strictly positive. □

Proof of Theorem 2. Fix an admissible δ and apply Theorem 1 with $\eta = \alpha/3$. Write $n_e := n_e(\alpha)$ and $c := c(\alpha)$. The lower bound on n_e and the definition of $n_r(\alpha/3)$ give

$$p_m(n_e,\delta) \geq 1 - \frac{\alpha}{3}, \text{ and } p_r(n_e) \geq 1 - \frac{\alpha}{3}.$$

Consequently,

$$\Pr(A_\delta \cap A_r(n_e)) \geq p_m(n_e,\delta) + p_r(n_e) - 1 \geq 1 - \frac{2\alpha}{3}.$$

The condition on c yields

$$\frac{2L_\lambda C_{\epsilon,\delta}(n_e)}{(1-\lambda)(c-n_e)} \leq \frac{\alpha}{3}.$$

Therefore, Theorem 1 gives

$$\liminf_{K\to\infty} \mathrm{PCS}_K \geq \left(1 - \frac{2\alpha}{3}\right)\left(1 - \frac{\alpha}{3}\right) \geq 1 - \alpha.$$

The finiteness of n_e follows from Lemma 1, and both constants are independent of K. □

F Proofs for Adaptive-Tolerance IRIS

Define the all-time confidence event for the optimal midpoint paths,

$$A := \bigcap_{j=1}^{K_1} \bigcap_{n\geq n_e} \{|\bar{X}_j^m(n) - \mu_*^m| \leq w(n)\}, \qquad X_*^L := \max_{j\leq K_1} \inf_{n\geq n_e} L_j(n).$$

Lemma 4 (Adaptive confidence event)**.** *If $w(n_e) \leq \delta_0/2$, then $\Pr(A) \geq p_A(n_e)$ and, on A, $X_*^L \geq \mu_*^m - \delta_0$.*

Proof. For each $j \leq K_1$ and $n \geq n_e$, sub-Gaussianity gives

$$\Pr\{|\bar{X}_j^m(n) - \mu_*^m| > w(n)\} \leq 2e^{-\beta} n^{-2}.$$

By a union bound,

$$\Pr(A^c) \leq \sum_{j=1}^{K_1} \sum_{n=n_e}^{\infty} \Pr\big(|\bar{X}_j^m(n) - \mu_*^m| > w(n)\big) \leq 2K_1 e^{-\beta} \sum_{n=n_e}^{\infty} n^{-2} = 1 - p_A(n_e).$$

Hence $\Pr(A) \geq p_A(n_e)$. Since w is decreasing, on A we have $L_j(n) \geq \mu_*^m - 2w(n) \geq \mu_*^m - \delta_0$ for all $j \leq K_1$ and $n \geq n_e$. □

For $k > K_1$, define

$$\tau_k := \inf\{n \geq n_e : U_k(n) < X_*^L\}, \qquad M_K := \max_{k>K_1} \tau_k,$$

with $M_K = n_e$ when $K = K_1$, and

$$Q_K := 1 + \sum_{k>K_1} (\tau_k - n_e)_+ + K_1(M_K - n_e)_+. \tag{9}$$

Lemma 5 (Adaptive exit workload)**.** *On A, an inferior alternative is never sampled after its count reaches τ_k. Moreover,*

$$\limsup_{K\to\infty} \frac{\mathbb{E}[Q_K \mid A \cap A_r(n_e)]}{K} \leq C_{\mathrm{ad}}(n_e). \tag{10}$$

Proof. Define $\widetilde{\tau}_k := \inf\{n \geq n_{\mathrm{sep}} : \bar{X}_k^m(n) - \mu_k^m < \Delta_m/2\}$ and $\widetilde{M}_K := \max_{k>K_1} \widetilde{\tau}_k$. On A, at $n = \widetilde{\tau}_k$,

$$U_k(n) < \mu_k^m + \Delta_m/2 + w(n) \leq \mu_*^m - \gamma_1 + \Delta_m = \mu_*^m - \delta_0 \leq X_*^L.$$

Thus $\tau_k \leq \widetilde{\tau}_k$ and $M_K \leq \widetilde{M}_K$ on A.

For $n \geq n_{\mathrm{sep}}$, $\{\widetilde{\tau}_k > n\} \subseteq \{\bar{X}_k^m(n) - \mu_k^m \geq \Delta_m/2\}$. Hence sub-Gaussianity gives

$$\Pr(\widetilde{\tau}_k > n) \leq e^{-n\Delta_m^2/(8\bar{\sigma}_m^2)} = e^{-a_m n}.$$

Since $\widetilde{\tau}_k \geq n_{\mathrm{sep}} \geq n_e$,

$$\mathbb{E}[(\widetilde{\tau}_k - n_e)_+] = n_{\mathrm{sep}} - n_e + \sum_{n=n_{\mathrm{sep}}}^{\infty} \Pr(\widetilde{\tau}_k > n) \leq n_{\mathrm{sep}} - n_e + \frac{e^{-a_m n_{\mathrm{sep}}}}{1 - e^{-a_m}} = C_{\mathrm{ad}}(n_e).$$

Choose $j_* \in \arg\max_{j\leq K_1} \inf_{n\geq n_e} L_j(n)$. If $n_k(t) = \tau_k$, then

$$U_k(\tau_k) < X_*^L \leq L_{j_*}(n_{j_*}(t)) \leq \max_i L_i(n_i(t)),$$

and $\bar{X}_k^m(\tau_k) \leq U_k(\tau_k) < L_{j_*}(n_{j_*}(t)) \leq \bar{X}_{j_*}^m(n_{j_*}(t))$. Thus neither channel samples k. Its count and empirical means remain unchanged, while $L_{j_*}(n_{j_*}(t')) \geq X_*^L$ at every later time t'. Consequently, k is never sampled again.

The variables $(\widetilde{\tau}_k)_{k>K_1}$ depend only on inferior midpoint paths and are independent of $A \cap A_r(n_e)$. Therefore,

$$\mathbb{E}\left[\sum_{k>K_1} (\tau_k - n_e)_+ \mid A \cap A_r(n_e)\right] \leq \sum_{k>K_1} \mathbb{E}[(\widetilde{\tau}_k - n_e)_+] \leq (K - K_1)C_{\mathrm{ad}}(n_e).$$

For $K - K_1 \geq 1$, a union bound yields $\Pr(\widetilde{M}_K > n) \leq \min\{1, (K-K_1)e^{-a_m n}\}$ for $n \geq n_{\mathrm{sep}}$. Set $h_K := \max\{n_{\mathrm{sep}}, \lceil a_m^{-1} \log(K - K_1)\rceil\}$. Then

$$\begin{aligned}
\mathbb{E}[(M_K - n_e)_+ \mid A \cap A_r(n_e)] &\leq \mathbb{E}[(\widetilde{M}_K - n_e)_+] \\
&\leq h_K - n_e + \sum_{n=h_K}^{\infty} (K - K_1)e^{-a_m n} \\
&\leq h_K - n_e + \frac{1}{1 - e^{-a_m}} = O(\log K).
\end{aligned}$$

Using equation 9, we obtain

$$\frac{\mathbb{E}[Q_K \mid A \cap A_r(n_e)]}{K} \leq \frac{1}{K} + \frac{K - K_1}{K} C_{\mathrm{ad}}(n_e) + \frac{K_1}{K}\left(h_K - n_e + \frac{1}{1 - e^{-a_m}}\right).$$

Since K_1 is fixed and $h_K = O(\log K)$, taking $\limsup_{K\to\infty}$ proves equation 10. □

Lemma 6 (Adaptive vote charging)**.** *Let D_K^{ad} be the number of incorrect radius-channel votes. On $A \cap A_r(n_e)$,*

$$D_K^{\mathrm{ad}} \le L_\lambda Q_K.$$

Fewer than $H_K/2$ incorrect votes imply correct selection.

Proof. Use the block decomposition and the quantities d_j, z_j from the proof of Lemma 3. For each nonterminal block ending at midpoint epoch m_j, define $b_j = \mathbf{1}\{q_{m_j} \in G_m, n_{q_{m_j}}(m_j) < M_K\}$. We claim that, on $A \cap A_r(n_e)$,

$$d_j > 0 \quad \Longrightarrow \quad z_j + b_j \ge 1.$$

If $z_j \ge 1$, the claim is immediate. Otherwise, an incorrect radius vote in the block requires $1 \notin \mathcal{C}_s^{\mathrm{ad}}$, because $A_r(n_e)$ gives

$$1 \in \mathcal{C}_s^{\mathrm{ad}},\ q_s \ne 1 \quad \Longrightarrow \quad q_s \notin G_m.$$

On A, there is therefore an inferior blocker b satisfying

$$L_b(n_b(s)) > U_1(n_1(s)) \ge \mu_*^m.$$

Since $z_j = 0$, the blocker is not sampled and keeps alternative 1 excluded through the remaining radius epochs. At epoch m_j, write $q := q_{m_j} \in G_m$. Then

$$\mu_*^m + w(n_q(m_j)) \ge \bar{X}_q^m(n_q(m_j)) \ge \bar{X}_b^m(n_b(m_j)) > \mu_*^m + w(n_b(m_j)).$$

The strict decrease of w and Lemma 5 imply

$$n_q(m_j) < n_b(m_j) < \tau_b \le M_K,$$

so $b_j = 1$. The same spacing argument as in Lemma 3 now gives

$$D_K^{\mathrm{ad}} \le L_\lambda \left(1 + \sum_j z_j + \sum_j b_j\right).$$

Moreover,

$$\sum_j z_j \le \sum_{k > K_1} (\tau_k - n_e)_+, \qquad \sum_j b_j \le K_1 (M_K - n_e)_+,$$

and hence $D_K^{\mathrm{ad}} \le L_\lambda Q_K$. The majority argument is identical to that in Lemma 3. □

Proof of Theorem 3. Conditional Markov's inequality and Lemmas 5 and 6 give

$$\limsup_{K\to\infty} \Pr(D_K^{\mathrm{ad}} \ge H_K/2 \mid A \cap A_r(n_e)) \le \frac{2L_\lambda C_{\mathrm{ad}}(n_e)}{(1-\lambda)(c-n_e)}.$$

The first lower bound on n_e ensures $w(n_e) \le \delta_0/2$. Moreover,

$$1 - p_A(n_e) = 2K_1 e^{-\beta} \sum_{n=n_e}^{\infty} n^{-2} \le \frac{2K_1 e^{-\beta}}{n_e - 1} \le \eta, \text{ and } p_r(n_e) \ge 1 - \eta,$$

where the last inequality follows from $n_e \ge n_r(\eta)$. Lemmas 4 and 1, together with the Fréchet bound, therefore give

$$\begin{aligned} \Pr(A \cap A_r(n_e)) &\ge \Pr(A) + \Pr(A_r(n_e)) - 1 \\ &\ge p_A(n_e) + p_r(n_e) - 1 \ge 1 - 2\eta > 0. \end{aligned}$$

Combining the two bounds yields

$$\liminf_{K\to\infty} \mathrm{PCS}_K \ge (1-2\eta)\left(1 - \frac{2L_\lambda C_{\mathrm{ad}}(n_e)}{(1-\lambda)(c-n_e)}\right) > 0,$$

which proves equation 6. □

Proof of Theorem 4. Fix $\alpha \in (0,1)$ and write $n_e := n_e(\alpha)$ and $c := c(\alpha)$. Apply Theorem 3 with $\eta = \alpha/3$. The first lower bound on n_e ensures $w(n_e) \le \delta_0/2$. Moreover,

$$1 - p_A(n_e) = 2K_1 e^{-\beta} \sum_{n=n_e}^{\infty} n^{-2} \le \frac{2K_1 e^{-\beta}}{n_e - 1} \le \frac{\alpha}{3}.$$

Since $n_e \ge n_r(\alpha/3)$, the definition of $n_r(\alpha/3)$ gives $p_r(n_e) \ge 1 - \alpha/3$. Thus

$$\Pr(A \cap A_r(n_e)) \ge p_A(n_e) + p_r(n_e) - 1 \ge 1 - \frac{2\alpha}{3}.$$

The condition on c also yields

$$\frac{2L_\lambda C_{\mathrm{ad}}(n_e)}{(1-\lambda)(c - n_e)} \le \frac{\alpha}{3}.$$

Combining the Fréchet bound with the conditional success bound gives

$$\liminf_{K\to\infty} \mathrm{PCS}_K \ge \left(1 - \frac{2\alpha}{3}\right)\left(1 - \frac{\alpha}{3}\right) \ge 1 - \alpha.$$

The quantity $n_r(\alpha/3)$ is finite by Lemma 1, and all parameter choices are finite and independent of K. □