RETHINKING SUCCESSIVE REJECTS ALGORITHMS

Anonymous authors
Paper under double-blind review

ABSTRACT

In the best-arm identification problem, an algorithm adaptively selects arms to find the arm of the largest mean reward. Successive Rejects (SR) is a simple fixed-budget best-arm identification algorithm that eliminates one candidate arm at each phase. At first glance, SR seems to have limited adaptivity, since it follows predetermined phase lengths. However, for Gaussian arms with common variance, SR's worst-instance error exponent normalized by the popular complexity measure $H_1(P)$ is optimal up to a constant factor, which partly explains why SR's performance is competitive compared with state-of-the-art adaptive sampling algorithms. In this paper, we extend the framework of SR by optimizing its predefined schedule of phase lengths. We formulate phase-schedule design as an optimization problem for the normalized error exponent. Then, we introduce a version of SR with the optimized schedule, which we call H_1-balanced Successive Rejects (H_1-SR). This schedule is optimal within the class of SR-type algorithms with deterministic equal-count schedules. For Gaussian arms with common variance, H_1-SR achieves at least $11/12$ of the optimal normalized rate over all adaptive algorithms for every $K \geq 2$, and this approximation ratio converges to one as $1 - O(1/\log K)$ as $K \to \infty$. This is a significant improvement over original SR, whose corresponding fraction of the optimal normalized rate converges to $1/4$. Finally, experiments on ten synthetic instances and two dataset-derived Gaussian benchmarks compare the optimized schedule with original SR at finite horizons. They show substantial gains on many instances while also exposing a regime in which more aggressive early elimination is costly.

1 INTRODUCTION

Best-arm identification asks how to use noisy measurements to select the alternative with the largest mean. The forecaster, who follows an algorithm, adaptively chooses which arm to measure and, after a prescribed budget of T observations, outputs a final recommendation: its estimate of the best arm.

Successive Rejects (SR; Audibert et al., 2010) is a particularly simple elimination algorithm that involves $K - 1$ phases. Each phase gives the same number of additional samples to every remaining arm, then eliminates the arm with the smallest cumulative empirical mean. The phase lengths are fixed in advance. Sequential Halving (SH; Karnin et al., 2013) uses a similar design but eliminates approximately half the remaining arms in each phase. These methods show that simple elimination can achieve an exponentially small probability of error. The multiplicative constant in the exponential rate, however, determines how much budget is needed when small error probabilities matter.

The two methods also differ in how they use observations. Original SH ranks arms using samples collected in the current phase, whereas SR retains earlier samples in each surviving arm's empirical mean. This distinction can affect finite-budget performance. SR remains a competitive benchmark alongside more recent algorithms such as Continuous Rejects (Wang et al., 2023) and Almost Tracking (Komiyama et al., 2026). Its continued relevance raises two questions: how close is its error rate to the best attainable rate, and can a small change to its schedule close the remaining gap?

The limited adaptivity of SR makes its performance gap particularly intriguing. Once the rejection rule is fixed—sample all surviving arms equally and eliminate an arm with the smallest cumulative empirical mean—the only remaining design choice is the *clock*: how many observations should each surviving arm have when each elimination takes place. Original SR makes one particular choice.

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.

It is not immediate whether that choice is well matched to the statistical difficulty of avoiding an erroneous elimination.

We therefore study the phase schedule itself as an optimization variable. Our argument proceeds from the outside in. We first place an information bound on the H_1-normalized rate of an arbitrary adaptive algorithm, where the H_1 is a popularly used complexity measure of the instance. We then restrict attention to the class of SR-type algorithms with deterministic equal-count schedules, which is a subset of all adaptive algorithms that includes the original SR, and ask for the best schedule in that class. A sharp analysis of the event that the true best arm is simultaneously below all surviving competitors turns the statistical problem into a small deterministic balancing problem. Solving it produces the phase schedule; the algorithm is a consequence of that optimization.

1.1 How should we compare identification algorithms?

An algorithm can work well when one arm is clearly better and struggle when several arms have nearly equal means. Comparing raw error probabilities across such instances confounds algorithm quality with problem difficulty. We therefore account for the gaps between arm means before comparing how quickly errors decrease with the budget. Following Komiyama et al. (2022; 2026), we compare the worst-instance exponential rate after normalization by an instance complexity. We use H_1, the sum of inverse squared gaps to the best arm, as a complexity measure that we normalize the error rate. A larger normalized rate means a stronger guarantee relative to this measure of difficulty. Appendix **??** explains how the guarantees transfer to uniformly comparable complexities.

Our benchmark is the best such rate over *all* adaptive algorithms, not just elimination methods. The algorithm does not know the true gaps; H_1 is used to evaluate it, not as an input. This criterion does not require matching a gap-informed oracle on every instance (Degenne, 2023). We give the formal definition in Section 2.1.

Our algorithm can be used to the best-arm identification with sub-Gaussian rewards, including rewards supported in $[0, 1]$ (or any finite region). Our optimality is for Gaussian rewards as we provides the error rate lower bound to that case.

1.2 Contributions

The contributions follow the same progression as the analysis.

- **An explicit characterization of the best rate for any adaptive algorithm.** For Gaussian arms we derive an explicit upper bound $R^\star_K \le 1/(\sigma^2 B_K)$ on the best H_1-normalized rate attainable by arbitrary adaptive algorithms.
- **Optimization within the class of SR-type algorithms.** We characterize the worst-instance rate of deterministic equal-count successive-rejection schedules. Its unique solution gives us H_1-SR algorithm, that achieves the matching rate.
- **Comparison with all algorithms and with original SR.** The optimized schedule satisfies $R_{H_1\text{-SR}}/R^\star_K \ge B_K/D_K \ge 11/12$ for every K, with $B_K/D_K = 1 - O(1/\log K)$, whereas original SR satisfies $R_{\mathrm{SR}}/R^\star_K \to 1/4$ as $K \to \infty$. Figure 1 illustrates the difference between SR and H_1-SR in terms of budget allocation across phases.
- **Finite-budget experiments.** On ten synthetic and two dataset-derived Gaussian benchmarks we compare H_1-SR with original SR. The experiments show that H_1-SR often improves terminal error, while a deliberately highlighted failure case shows why the asymptotic worst-instance guarantee should not be read as pointwise finite-budget dominance.

1.3 Related work

Elimination algorithms Beyond SR (Audibert et al., 2010) and Sequential Halving (Karnin et al., 2013; Zhao et al., 2023), Shahrampour et al. (2017) study general sequential-elimination schedules, including nonlinear cumulative phase lengths. However, they have not characterized the optimal scheduling of the phase lengths. Wang et al. (2023) develop a large-deviation analysis of adaptive identification and introduce Continuous Rejects, which dynamically chooses the phase lengths.

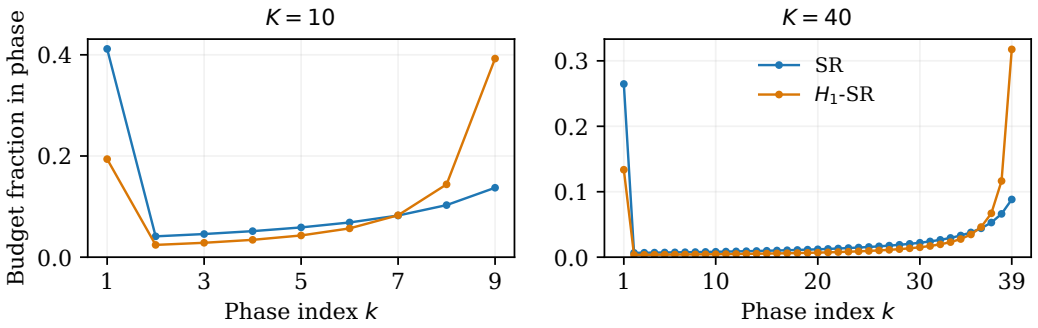


Figure 1: Fraction of the total budget spent in each phase, for the original SR and our H_1-SR. Each phase includes all pulls of its surviving arms, not a cumulative per-arm count. Both methods spend the same total budget and eliminate one arm per phase; H_1-SR spends more of the budget in the final phases.

Barrier et al. (2023) consider fixed-budget identification problems for a class of reward distributions bounded on $[0, 1]$ and consider the SR algorithm therein.

Minimax rates and implementable algorithms. Audibert et al. (2010); Carpentier & Locatelli (2016) show that there is an $O(\log K)$ gap in the complexity of identifying the best arm between the fixed-budget setting (considered in this paper) and fixed-confidence settings, in which the algorithm can choose when to stop sampling. A notable difference that subsequent research disclosed is that optimization in these two settings is fundamentally different. Namely, the fixed-confidence setting allows a uniformly good algorithm. Namely, there exist several algorithms such as Track and Stop (Garivier & Kaufmann, 2016) and Top-two Thompson Sampling (Russo, 2020; Jourdan et al., 2023; Bandyopadhyay et al., 2024) that achieve this property, meaning they perform optimally on all problem instances, whereas in the fixed-budget setting, no such uniformly good algorithm exists (Komiyama et al., 2022; Wang et al., 2023; Degenne, 2023; Wang et al., 2024). Therefore, the latter setting is closer to traditional statistical decision theory[1] (Jennison et al., 1982), where the optimality depends on the choice of the complexity measure $H(P)$, such as admissibility (Imbens et al., 2025) and minimax optimality (Komiyama et al., 2022). In particular, Komiyama et al. (2022) formulate minimax rates given a complexity measure $H(P)$, and introduce the delayed optimal tracking (DOT) algorithm. However, DOT is almost impossible to implement in practice as it needs to solve a dynamic program with a long horizon. Later on, Komiyama et al. (2026) proposed a practical algorithm called simple tracking that does not require the information on the horizon T, and its version called almost tracking, which has a provable constant-factor rate guarantee. However, its constant on top of the rate guarantee is impractically large.

Our main results start from a more practical viewpoint. Starting from the popularly used SR algorithm, we improve its performance without introducing additional complexity into the algorithm. Our main algorithm, H_1-SR, significantly improves the practical performance of SR, and we derive nearly minimax rate guarantees for the most popularly used complexity measure $H_1(P)$.

2 From Arbitrary Adaptive Algorithms to an Optimized Rejection Schedule

2.1 Model and Performance Criterion

We follow the notation of Komiyama et al. (2026). There are $K \geq 2$ arms with unknown means $\boldsymbol{P} = (P_1, \ldots, P_K)$. At round t, algorithm $\mathcal{A}$ chooses $I(t) \in [K]$ from the history up to round $t-1$,

[1]Note that, in the statistical decision theory framework, it is usually called the risk measure.

then observes a fresh reward $X(t)$ from that arm. Write

$$N_i(t) = \sum_{u=1}^{t} \mathbf{1}\{I(u) = i\}, \qquad Q_i(t) = \frac{\sum_{u=1}^{t} X(u)\mathbf{1}\{I(u) = i\}}{N_i(t)} \quad (N_i(t) > 0). \tag{1}$$

Thus P_i is a true mean and $Q_i(t)$ is an empirical mean; write $\boldsymbol{Q}(t) = (Q_1(t), \ldots, Q_K(t))$ once all arms have been sampled. After the known horizon T, the algorithm recommends $J(T)$.

Assume a unique best arm and define

$$i^\star = i^\star(\boldsymbol{P}) = \arg\max_{i \in [K]} P_i, \qquad P^\star = \max_i P_i, \qquad \Delta_i = P^\star - P_i > 0 \quad (i \neq i^\star).$$

Parenthesized subscripts denote true ranks: $P_{(1)} > P_{(2)} \geq \cdots \geq P_{(K)}$. The unique-best assumption makes the target recommendation unambiguous and ensures that every suboptimal gap is strictly positive. For the exact-rate and minimax analysis, observations from arm i are independent draws from $N(P_i, \sigma^2)$, with a common variance $\sigma^2 > 0$, and the reward streams are independent across arms.

We evaluate an algorithm through four quantities:

$$\mathrm{PoE}_{\mathcal{A}}(\boldsymbol{P}) = \mathbb{P}^{\mathcal{A}}_{\boldsymbol{P}}(J(T) \neq i^\star(\boldsymbol{P})), \qquad H_1(\boldsymbol{P}) = \sum_{i \neq i^\star(\boldsymbol{P})} \Delta_i^{-2},$$

$$E_{\mathcal{A}}(\boldsymbol{P}) = \liminf_{T \to \infty} \frac{1}{T} \log \frac{1}{\mathrm{PoE}_{\mathcal{A}}(\boldsymbol{P})}, \qquad R_{\mathcal{A}} = \inf_{\boldsymbol{P}:\ \text{unique best}} H_1(\boldsymbol{P}) E_{\mathcal{A}}(\boldsymbol{P}). \tag{2}$$

Here $\mathrm{PoE}_{\mathcal{A}}(\boldsymbol{P})$ is the probability of recommending a suboptimal arm after T observations; as in Komiyama et al. (2022), its dependence on T is implicit. The complexity $H_1(\boldsymbol{P})$ increases when competitors approach the best arm, reflecting the greater difficulty of distinguishing their means. The exponent $E_{\mathcal{A}}(\boldsymbol{P})$ measures the asymptotic decrease of error with budget: when a rate limit exists, the error behaves as $\exp\{-TE_{\mathcal{A}}(\boldsymbol{P}) + o(T)\}$. Finally, $R_{\mathcal{A}}$ is the smallest H_1-normalized exponent over all unique-best instances. It measures the algorithm's worst-instance guarantee after accounting for this complexity.

Let $\mathfrak{A}_K$ denote the class of all adaptive algorithm families in this model, including randomized and horizon-dependent ones. For each horizon T, arm selections may depend on the past observations and internal randomness, and the final recommendation may depend on the full observed history and internal randomness. These rules may depend on K and T, but not on the unknown means. Our ideal objective is to attain the H_1-minimax rate $R^\star_K = \sup_{\mathcal{A} \in \mathfrak{A}_K} R_{\mathcal{A}}$, or to guarantee a large fraction of it with a practical algorithm. We determine SR's normalized rate and improve it by changing only its phase lengths.

We exclusively consider H_1 here. Appendix I relates this convention to the other complexities studied in the literature (Audibert et al., 2010; Garivier & Kaufmann, 2016; Komiyama et al., 2022).

For the finite-budget upper bound, we also allow independent rewards with a common sub-Gaussian variance proxy σ^2: every centered reward has moment generating function at most $e^{\lambda^2\sigma^2/2}$ for all $\lambda \in \mathbb{R}$. In particular, our algorithm also has error rate guarantees under any rewards supported in $[0, 1]$ by setting $\sigma^2 = 1/4$ in the analysis. For these extensions, the error probability depends on the full reward laws, and could be the case that we can further improve the bounds when we have additional knowledge about the reward distributions.

2.2 An information bound for every adaptive algorithm

We first describe the possible achievable rates for any adaptive algorithm, including not only SR-type algorithms, but also Continuous Rejects (Wang et al., 2023) and Simple Tracking (Komiyama et al., 2026). Namely, let

$$h_m = \sum_{j=1}^{m} j^{-1}, \qquad B_K = \begin{cases} 8, & K = 2, \\ 43/7 + 2h_{K-1}, & K \geq 3. \end{cases}$$

The constant B_K is an inverse-rate information constraint.

Theorem 1 (Rate upper bound)**.** *For Gaussian arms and the criterion in* (2)*,* $R_K^\star \le 1/(\sigma^2 B_K)$.

The proof is deferred to Appendix E, which gives the full change-of-measure argument and derives the explicit constant B_K from a single limiting allocation constrained by a hierarchy of Gaussian alternatives. The theorem is a worst-instance asymptotic restriction; it is not a finite-budget lower bound for every individual instance.

2.3 The class of SR-type algorithms

We now focus on the class of SR-type algorithms. We consider all deterministic schedules (ie., the phase lengths are fixed in advance) and we preserve SR's sampling and rejection rules. Let $\mathfrak{A}_K^{\mathrm{SR}} \subseteq \mathfrak{A}_K$ be the class of SR-type algorithms that sample every surviving arm to the same predetermined cumulative count, reject one arm with a smallest empirical mean, and recommend the last survivor. With s arms remaining, denote the cumulative count by $m_s(T)$. The counts are positive integers satisfying

$$m_K(T) \le \cdots \le m_2(T), \qquad 2m_2(T) + \sum_{s=3}^{K} m_s(T) \le T.$$

The budget counts each rejected arm at its final sample count and both finalists at $m_2(T)$. Tie breaking in elimination is fixed.

In summary, an algorithm in $\mathfrak{A}_K^{\mathrm{SR}}$ is characterized by the choice of the deterministic schedule $(m_K(T), \dots, m_2(T))$. For example, as we discussed in Section 2.3, the original version of SR (Audibert et al., 2010) corresponds to $m_s/T \propto 1/s$. We next relate the deterministic schedule to the algorithm's performance (i.e., error rate $R_{\mathcal{A}}$).

The rest of this section shows how the deterministic schedule $(m_K(T), \dots, m_2(T))$ can be optimized.

2.4 How the deterministic schedule controls performance

Consider a phase with s surviving arms, while the true best arm is still active. Rejecting it requires its empirical mean to be no larger than those of all $s-1$ competitors at once. To analyze this event without conditioning on survival, fix a deterministic candidate set $A = \{i^\star, i_1, \dots, i_{s-1}\}$ containing the true best. Define its true gaps and complexity by

$$\delta_j = P_{i^\star} - P_{i_j} > 0, \qquad H_A = \sum_{j=1}^{s-1} \delta_j^{-2},$$

that is, $\delta_j = \Delta_{i_j}$. For the proof, let $\overline{Q}_{i,m}$ be the mean of the first m rewards in arm i's potential independent sequence. This is an equivalent way to construct the sequential observations, not an extra observation available to the algorithm. At any deterministic common count m, the empirical means of the arms in A are represented by $\overline{Q}_{i,m}$.

We now relate the deterministic schedule to the SR-type algorithm's error exponent through concentration inequalities. Namely, for Gaussian rewards with variance σ^2, the probability that an arm's empirical mean falls below a fixed threshold $q < P_i$ satisfies

$$\mathbb{P}\big(\overline{Q}_{i,m} \le q\big) \le \exp\big(-m\,\mathrm{KL}\big(N(q,\sigma^2) \,\|\, N(P_i,\sigma^2)\big)\big) = \exp\left(-\frac{m(q-P_i)^2}{2\sigma^2}\right). \tag{3}$$

After subtracting $P^\star$, the true mean vector on A is $p_A = (0, -\delta_1, \dots, -\delta_{s-1})$. A necessary condition for rejecting the best arm is that the shifted empirical vector belongs to $\mathcal{C}_A = \{q \in \mathbb{R}^s : q_0 \le q_j,\ 1 \le j < s\}$. This is a closed convex polyhedral cone, the intersection of the half-spaces $q_0 - q_j \le 0$. It contains exactly the vectors in which the true best arm has a smallest empirical mean, allowing ties. The quantity d_A is the smallest Euclidean displacement of the *whole mean vector* that makes the true best arm empirically worst:

$$d_A^2 = \inf_{q \in \mathcal{C}_A} \|q - p_A\|^2 = \min_{u \ge 0} \left\{ u^2 + \sum_{j=1}^{s-1} (\delta_j - u)_+^2 \right\} \ge \frac{(s-1)^2}{sH_A}. \tag{4}$$

Here $(x)_+ = \max\{x, 0\}$ and u is the downward displacement of the best-arm empirical mean from 0 to $-u$. For this fixed coordinate, the nearest feasible competitor mean is $\max\{-\delta_j, -u\}$, requiring an upward displacement $(\delta_j - u)_+$. Summing these squared displacements and minimizing over u gives the middle expression. The inequality is tight when the gaps in A are equal. Projection onto this cone and a sub-Gaussian Chernoff bound, which gives a finite-sample guarantee with the same quadratic structure in the exponent as (3), give

$$\mathbb{P}\big(\overline{Q}_{i^\star,m} \le \overline{Q}_{i,m} \text{ for every } i \in A \setminus \{i^\star\}\big) \le \exp\left(-\frac{m(s-1)^2}{2s\sigma^2 H_A}\right). \tag{5}$$

Appendix B proves both the projection step and the sharp inverse-gap inequality. The large-deviation analysis of SR in (Wang et al., 2023; Komiyama et al., 2026) identifies the cost of such misleading empirical means. Here we bound that cost, at the end of a phase $K+1-s$ where s arms remains, sharply in terms of H_A. Since $H_A \le H_1$, (5) also yields an upper bound in terms of H_1 that is uniform over all such sets A. The analysis is enabled by the fact that an SR-type algorithm conducts uniform exploration at each phase, which is much more simpler than the design of algorithms appearing in (Wang et al., 2023; Komiyama et al., 2026).

The bound (5) is the key design input. Its coefficient depends only on the number s of surviving arms. The next subsection packages this coefficient into a convenient quantity and asks how to choose all checkpoint counts so that no active-set size becomes the bottleneck.

2.5 A SMALL BALANCING PROBLEM WITH AN EXPLICIT SOLUTION

The simultaneous-rejection bound in Section 2.4 identifies

$$a_s = \frac{2s}{(s-1)^2}, \qquad 2 \le s \le K,$$

as the information requirement associated with the checkpoint having s surviving arms. Recall that $m_s(T)$ is the cumulative count per surviving arm at the checkpoint with s arms. An SR-type algorithm is characterized by $(m_s(T))_{2\le s\le K}$. Since this is a vector sequence of T, we take the limit as $T \to \infty$. When the limits exist,

$$q_s = \lim_{T\to\infty} \frac{m_s(T)}{T}, \qquad 2 \le s \le K.$$

Thus q_s is a cumulative per-arm fraction, not the fraction spent in that phase. For schedules without full limits,[2] Appendix D uses the same notation for limits along a common subsequence. Since the error exponent is determined by the worst phase, the guaranteed normalized exponent is $\min_s q_s/(\sigma^2 a_s)$. Maximizing this among the class of limit fractions $(q_2, \dots, q_K)$ gives

$$\max_{\substack{0\le q_K \le \cdots \le q_2 \\ 2q_2 + \sum_{s=3}^K q_s \le 1}} \min_{2\le s\le K} \frac{q_s}{a_s}. \tag{6}$$

Define the normalizing constant

$$D_K := 2a_2 + \sum_{s=3}^{K} a_s = 4 + 2h_{K-1} + 2h^{(2)}_{K-1}, \qquad h^{(2)}_m = \sum_{j=1}^{m} j^{-2}. \tag{7}$$

The optimal limit fractions are given by q_s such that (6) holds, is achieved by the following:

Proposition 2 (Information balancing)**.** *The value of* (6) *is* $1/D_K$*, and its unique optimizer is* $q_s = a_s/D_K$ *for all* s*.*

Proof. If the minimum is at least v, every $q_s \ge v a_s$. The budget gives $1 \ge v(2a_2 + \sum_{s=3}^K a_s) = vD_K$. Equality is feasible because a_s decreases with s, and equality in the budget forces equality in every coordinate lower bound. □

[2]Of course, such cases, for example, $m_s(T)$ quickly switches for different T, do not improve the rate!

Algorithm 1 H_1-balanced Successive Rejects (H_1-SR)

Require: Number of arms $K \geq 2$; horizon $T \geq K$; fixed tie-breaking order
1: $A_1 \leftarrow [K]$, $n_0 \leftarrow 0$; compute D_K from (7)
2: **for** $k = 1, \ldots, K-1$ **do**
3: $s \leftarrow K + 1 - k$; $n_k \leftarrow 1 + \lfloor (T-K)a_s/D_K \rfloor$
4: Pull each $i \in A_k$ exactly $n_k - n_{k-1}$ additional times
5: Let t_k be the current round; $j_k \leftarrow \arg\min_{i \in A_k} Q_i(t_k)$
6: $A_{k+1} \leftarrow A_k \setminus \{j_k\}$
7: **end for**
8: Set $J(T)$ to the unique element of A_K
9: If fewer than T pulls were used, pad the budget and ignore those observations
10: **return** $J(T)$

2.6 The resulting algorithm and its guarantee

Proposition 2 determines the asymptotically balanced checkpoint fractions. We implement those fractions at finite horizon by integer rounding. For $T \geq K$, let $A_1 = [K]$ and $n_0 = 0$. In phase $k \in \{1, \ldots, K-1\}$, the active set has size $s = K + 1 - k$. Set

$$n_k = 1 + \left\lfloor \frac{T-K}{D_K} a_{K+1-k} \right\rfloor. \tag{8}$$

The checkpoint n_k is the cumulative sample count *per surviving arm*, not the global time. Each arm in A_k receives $n_k - n_{k-1}$ additional samples. At the resulting global round t_k, all active arms satisfy $N_i(t_k) = n_k$, and an arm minimizing the empirical mean $Q_i(t_k)$ is removed. The phase therefore uses $(K+1-k)(n_k - n_{k-1})$ pulls. Algorithm 1 summarizes the procedure.

The following algorithm bounds the probability of error for Algorithm 1.

Theorem 3 (Finite-budget error upper bound)**.** *For every $T \geq K$ and every unique-best instance with independent σ^2-sub-Gaussian arm rewards,*

$$\mathrm{PoE}_{H_1\text{-SR}}(\boldsymbol{P}) \leq \min\left\{1, (2^{K-1}-1)\exp\left(-\frac{T-K}{\sigma^2 D_K H_1(\boldsymbol{P})}\right)\right\}. \tag{9}$$

Consequently, at every fixed instance the error exponent is at least $1/(\sigma^2 D_K H_1(\boldsymbol{P}))$. This gives $R_{H_1\text{-SR}} \geq 1/(\sigma^2 D_K)$.

In the Gaussian model, the asymptotic rate guarantee is tight: Theorem 4 below provides the matching upper bound, yielding $R_{H_1\text{-SR}} = 1/(\sigma^2 D_K)$.

2.7 Class optimality and comparison with any algorithm

Since H_1-SR is designed to optimize the allocation of pulls across arms in a successive rejects framework, it achieves the best possible error exponent within this class of algorithms.

Theorem 4 (Optimal clock within the class of SR-type algorithms)**.** *For Gaussian arms,* $\sup_{\mathcal{A} \in \mathfrak{A}_K^{\mathrm{SR}}} R_{\mathcal{A}} = 1/(\sigma^2 D_K)$*, attained by* H_1-SR.

The upper bound is enforced by a different hard instance at each checkpoint. Consider s arms with small gaps and $K - s$ arms with much lower means. With high probability the latter are removed first, leaving the difficult comparison to the s-arm checkpoint. Too few samples at that checkpoint make best-arm rejection too likely. Applying this argument to every s yields simultaneous lower requirements on the counts; their sum exhausts the budget at the proposed schedule. Appendix D makes this argument precise.

This theorem identifies the best rate within $\mathfrak{A}_K^{\mathrm{SR}}$. Combining it with the bound for arbitrary adaptive algorithms gives a separate comparison with the unrestricted optimum.

Corollary 5 (Uniform approximation and leading constant)**.** *For Gaussian arms and every fixed* $K \geq 2$*,*

$$\frac{1}{\sigma^2 D_K} \leq R_K^\star \leq \frac{1}{\sigma^2 B_K}, \qquad \frac{R_{H_1\text{-SR}}}{R_K^\star} \geq \frac{B_K}{D_K} \geq \frac{11}{12}. \tag{10}$$

Moreover $B_K/D_K = 1-O(1/\log K)$. As $K \to \infty$ after the fixed-K horizon limit, $R_{H_1\text{-SR}}/R^\star_K \to 1$ and $R^\star_K \sim 1/(2\sigma^2 \log K)$.

Corollary 5 shows that H_1-SR is very close to the optimal rate attainable by any algorithm, which is somewhat surprising given the fact that it determines the length of the schedule beforehand and does not revisit the arms once they are eliminated.

3 H_1-SR VERSUS ORIGINAL SUCCESSIVE REJECTS

3.1 SHARP NORMALIZED-RATE COMPARISON

For $T > K$, original SR uses

$$n_k^{\mathrm{SR}} = \left\lceil \frac{T-K}{\overline{\log}(K)(K+1-k)} \right\rceil, \qquad \overline{\log}(K) = \frac{1}{2} + \sum_{i=2}^{K} \frac{1}{i}. \tag{11}$$

Theorem 6 (Sharp normalized original-SR comparison)**.** *For Gaussian arms under* (2)*, $R_{\mathrm{SR}} = 1/(8\sigma^2\overline{\log}(K))$. For $T > K$,*

$$\mathrm{PoE}_{\mathrm{SR}}(\boldsymbol{P}) \le \min\left\{1, \frac{K(K-1)}{4} \exp\left(-\frac{T-K}{8\sigma^2\overline{\log}(K) H_1(\boldsymbol{P})}\right)\right\}.$$

Consequently,

$$\frac{R_{\mathrm{SR}}}{R^\star_K} \ge \frac{B_K}{8\overline{\log}(K)} \ge \frac{1}{4}.$$

$$\frac{R_{H_1\text{-SR}}}{R_{\mathrm{SR}}} = \frac{8\overline{\log}(K)}{D_K} \longrightarrow 4, \qquad \frac{R_{\mathrm{SR}}}{R^\star_K} \longrightarrow \frac{1}{4}. \tag{12}$$

Theorem 6 shows that, for $K \ge 3$, H_1-SR achieves a strictly better worst-instance normalized exponent compared to the original SR algorithm, and it is by 4 for large K.

For general σ^2-sub-Gaussian arms, the displayed finite-budget bound remains valid with prefactor $K(K-1)/2$ instead of $K(K-1)/4$. The smaller Gaussian prefactor uses an additional one-half tail factor; the exact-rate and minimax comparisons in the theorem are Gaussian.

The exact error exponent of original SR is established in prior work (Wang et al., 2024; Komiyama et al., 2026). Appendix F gives the short self-contained argument needed for the H_1 normalization used here.

3.2 INTUITION ON THE RATIO

Recall that $m_s(T)$ is the cumulative number of samples per surviving arm at the phase with s arms remaining. Ignoring integer rounding, the two schedules satisfy

$$\text{Original SR:} \quad m_s(T) \approx \frac{T}{s\overline{\log}(K)}, \qquad H_1\text{-SR:} \quad m_s(T) \approx \frac{T}{D_K} \frac{2s}{(s-1)^2}.$$

Their ratio is therefore

$$\frac{m_s^{H_1\text{-SR}}(T)}{m_s^{\mathrm{SR}}(T)} \longrightarrow \frac{2\overline{\log}(K)}{D_K} \left(\frac{s}{s-1}\right)^2 \qquad (T \to \infty).$$

The ratio increases as fewer arms remain: H_1-SR spends less before early eliminations and more on the final comparisons (Figure 1). Since $2\overline{\log}(K)/D_K \to 1$, the final-two ratio tends to 4, matching the large-K rate improvement in Theorem 6.

4 NUMERICAL EXPERIMENTS

The theory compares worst-instance normalized asymptotic exponents and does not imply pointwise finite-budget dominance. We therefore compare H_1-SR with original SR on twelve fixed-horizon Gaussian problems: ten synthetic instances with $K = 40$, a MovieLens 1M-derived instance with $K = 31$, and an Open Bandit-derived instance with $K = 80$. The synthetic configurations follow Komiyama et al. (2026) and are listed in Appendix A; the other two instances use means derived from MovieLens 1M (Harper & Konstan, 2015) and Open Bandit (Saito et al., 2021). Rewards are unit-variance Gaussian. Each cell uses $10{,}000$ repetitions at the configured horizon T.

Our primary empirical quantity is the ratio $\widehat{\mathrm{PoE}}_{H_1\text{-SR}}/\widehat{\mathrm{PoE}}_{\mathrm{SR}}$: values below one favor H_1-SR and values above one favor SR. Paired terminal comparisons use two-sided exact McNemar tests with Holm correction over the twelve predeclared comparisons at family-wise level 0.05.

Table 1: Terminal error (%) over $10{,}000$ repetitions. Ratio $= \widehat{\mathrm{PoE}}_{H_1\text{-SR}}/\widehat{\mathrm{PoE}}_{\mathrm{SR}}$; boldface marks a Holm-corrected supported improvement.

			Terminal $\widehat{\mathrm{PoE}}$ (%)			
Instance	K	T	H_1-SR	SR	Ratio	Supported
Synthetic 1	40	6,000	**9.92**	18.05	0.550	H_1-SR
Synthetic 2	40	1,000	**1.37**	4.61	0.297	H_1-SR
Synthetic 3	40	4,000	7.72	7.45	1.036	–
Synthetic 4	40	4,500	**9.88**	14.56	0.679	H_1-SR
Synthetic 5	40	25,000	**9.47**	13.73	0.690	H_1-SR
Synthetic 6	40	5,000	**5.54**	11.30	0.490	H_1-SR
Synthetic 7	40	8,000	15.88	**4.98**	3.189	SR
Synthetic 8	40	2,000	**2.01**	4.17	0.482	H_1-SR
Synthetic 9	40	1,500	**1.76**	5.08	0.346	H_1-SR
Synthetic 10	40	2,500	**7.03**	12.36	0.569	H_1-SR
MovieLens 1M	31	10,000	**18.27**	24.87	0.735	H_1-SR
Open Bandit	80	3,000	9.37	10.27	0.912	–

H_1-SR has lower empirical terminal error on ten of twelve instances; after Holm correction, nine comparisons support H_1-SR, one supports SR, and two are inconclusive. MovieLens has ratio 0.735, whereas synthetic instance 7 has ratio 3.189. In that instance, H_1-SR uses 27 observations per arm before the first elimination versus 53 for SR, but 2023 versus 1054 at the final-two checkpoint. This illustrates the finite-budget early/late allocation tradeoff. Appendix A visualizes all trajectories. Regarding the synthetic instance 7, Appendix H defines dominance and discusses this tradeoff in relation to the worst-instance rate guarantee.

5 CONCLUSION AND LIMITATIONS

Optimizing the deterministic SR checkpoints yields H_1-SR, which is optimal within this class of SR-type algorithms and achieves at least $11/12$ of the optimal Gaussian H_1-normalized rate for every fixed K, with the certified fraction tending to one as K grows; original SR approaches $1/4$. Experiments show substantial finite-budget gains on many instances but also a clear non-dominance case. The finite-budget bound extends to independent sub-Gaussian rewards; matching optimality beyond the Gaussian setting and the early/late elimination tradeoff remain open.

AI USE STATEMENT

OpenAI Codex assisted with research-note development, mathematical claims and proof arguments, literature searches, experimental planning, proof exposition, and manuscript and artifact preparation. The authors reviewed the AI-assisted material and take responsibility for the final theoretical, experimental, and expository content.

Reproducibility statement

Algorithm 1 gives the complete H_1-SR procedure, including the integer checkpoints, tie handling, and unused-budget convention. The implementation is a direct translation of this pseudocode and the SR baseline uses the schedule in (11). Appendix A specifies the experimental instances, horizons, repetition count, statistical comparison, and trajectory interpretation. Appendices B–G contain the full proofs and hard-instance constructions needed to verify the theoretical claims.

References

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. URL https://www.learningtheory.org/colt2010/papers/59Audibert.pdf.

Agniv Bandyopadhyay, Sandeep Juneja, and Shubhada Agrawal. Optimal top-two method for best arm identification and fluid analysis. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

Antoine Barrier, Aurélien Garivier, and Gilles Stoltz. On best-arm identification with a fixed budget in non-parametric multi-armed bandits. In *Proceedings of the 34th International Conference on Algorithmic Learning Theory*, volume 201 of *Proceedings of Machine Learning Research*, pp. 136–181. PMLR, 2023. URL https://proceedings.mlr.press/v201/barrier23a.html.

Alexandra Carpentier and Andrea Locatelli. Tight (lower) bounds for the fixed budget best arm identification bandit problem. In *29th Annual Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pp. 590–604. PMLR, 2016. URL https://proceedings.mlr.press/v49/carpentier16.html.

Rémy Degenne. On the existence of a complexity in fixed budget bandit identification. In *Proceedings of the 36th Conference on Learning Theory*, volume 195 of *Proceedings of Machine Learning Research*, pp. 1131–1154. PMLR, 2023. URL https://proceedings.mlr.press/v195/degenne23a.html.

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

F. Maxwell Harper and Joseph A. Konstan. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, 5(4):19:1–19:19, 2015. doi: 10.1145/2827872.

Guido Imbens, Chao Qin, and Stefan Wager. Admissibility of completely randomized trials: A large-deviation approach, 2025. URL https://arxiv.org/abs/2506.05329. Preprint.

Christopher Jennison, Iain M. Johnstone, and Bruce W. Turnbull. Asymptotically optimal procedures for sequential adaptive selection of the best of several normal means. In *Statistical Decision Theory and Related Topics III*, volume 2, pp. 55–86. Academic Press, 1982.

Marc Jourdan, Rémy Degenne, and Emilie Kaufmann. An $varepsilon$-best-arm identification algorithm for fixed-confidence and beyond. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

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. PMLR, 2013. URL https://proceedings.mlr.press/v28/karnin13.html.

Junpei Komiyama, Taira Tsuchiya, and Junya Honda. Minimax optimal algorithms for fixed-budget best arm identification. In *Advances in Neural Information Processing Systems*, volume 35, pp. 10393–10404, 2022. URL https://papers.neurips.cc/paper_files/paper/2022/hash/43c18853329c7504996b255252b6cb1f-Abstract-Conference.html.

Junpei Komiyama, Kyoungseok Jang, and Junya Honda. Rate-optimal design for anytime best arm identification. In *Proceedings of the 29th International Conference on Artificial Intelligence and Statistics*, volume 300 of *Proceedings of Machine Learning Research*, pp. 748–756. PMLR, 2026. URL https://proceedings.mlr.press/v300/komiyama26a.html.

Daniel Russo. Simple bayesian algorithms for best-arm identification. *Operations Research*, 68(6): 1625–1647, 2020. doi: 10.1287/opre.2019.1911.

Yuta Saito, Shunsuke Aihara, Megumi Matsutani, and Yusuke Narita. Open bandit dataset and pipeline: Towards realistic and reproducible off-policy evaluation. In *Advances in Neural Information Processing Systems Datasets and Benchmarks Track*, 2021.

Shahin Shahrampour, Mohammad Noshad, and Vahid Tarokh. On sequential elimination algorithms for best-arm identification in multi-armed bandits. *IEEE Transactions on Signal Processing*, 65 (16):4281–4292, 2017. doi: 10.1109/TSP.2017.2706192. URL https://arxiv.org/abs/1609.02606.

Po-An Wang, Ruo-Chun Tzeng, and Alexandre Proutiere. Best arm identification with fixed budget: A large deviation perspective. In *Advances in Neural Information Processing Systems*, volume 36, pp. 16804–16815, 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/364d565b4b726c607aa40e1632045873-Abstract-Conference.html.

Po-An Wang, Kaito Ariu, and Alexandre Proutiere. On universally optimal algorithms for A/B testing. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 50065–50091. PMLR, 2024.

Yao Zhao, Connor Stephens, Csaba Szepesvári, and Kwang-Sung Jun. Revisiting simple regret: Fast rates for returning a good arm. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 42110–42158. PMLR, 2023.

A EXPERIMENTAL INSTANCES AND TRAJECTORY PLOTS

The synthetic configurations follow the benchmark family used by Komiyama et al. (2026). All ten use $K = 40$ independent Gaussian arms with unit variance. With arms indexed by $i \in \{1, \dots, 40\}$, the means used in our experiments are

$$\begin{aligned}
&\text{Instance 1:} && P_i = 1 - 0.05(i-1),\\
&\text{Instance 2:} && P_i = 10\left(\frac{i-1}{39}\right)^{0.8},\\
&\text{Instance 3:} && P_i = 1 - \frac{\sqrt{i-1}}{10},\\
&\text{Instance 4:} && (P_1, \dots, P_{40}) = (1, 0.9, 0.9, 0.9, 0.9, \underbrace{0, \dots, 0}_{35}),\\
&\text{Instance 5:} && P_1 = \sin\left(\frac{39\pi}{80}\right), \quad P_i = \sin\left(\frac{9\pi(40-i-1)}{800}\right),\ 2 \le i \le 40,\\
&\text{Instance 6:} && P_i = 0.75\, 3^{-i/10},\\
&\text{Instance 7:} && (P_1, \dots, P_{40}) = (1, \underbrace{0.8, \dots, 0.8}_{39}),\\
&\text{Instance 8:} && (P_1, \dots, P_{40}) = (1, \underbrace{0.8, 0.8, 0.8}_{3}, \underbrace{0.5, \dots, 0.5}_{6}, \underbrace{0.2, \dots, 0.2}_{10}, \underbrace{0, \dots, 0}_{20}),\\
&\text{Instance 9:} && (P_1, \dots, P_{40}) = (1, 0.8, 0.8, \underbrace{0, \dots, 0}_{37}),\\
&\text{Instance 10:} && (P_1, \dots, P_{40}) = (1, 0.9, 0.85, 0.8, \underbrace{0, \dots, 0}_{36}).
\end{aligned}$$

For Instances 5 and 8, the displayed definitions state explicitly the configurations used in our experiments. The two additional problems are Gaussian instances whose means are derived from MovieLens 1M and Open Bandit, as described in Section 4.

A.1 TRAJECTORY PLOTS

Figures 2 and 3 report the empirical probability of error along the trajectories configured for their terminal horizons. The blue curve is H_1-SR and the orange curve is original SR; shading gives pointwise exact 95% binomial intervals and the vertical axis is logarithmic. A point at $t < T$ is an intermediate recommendation from the same T-configured run, not a separate run retuned to horizon t, so these curves do not constitute anytime guarantees.

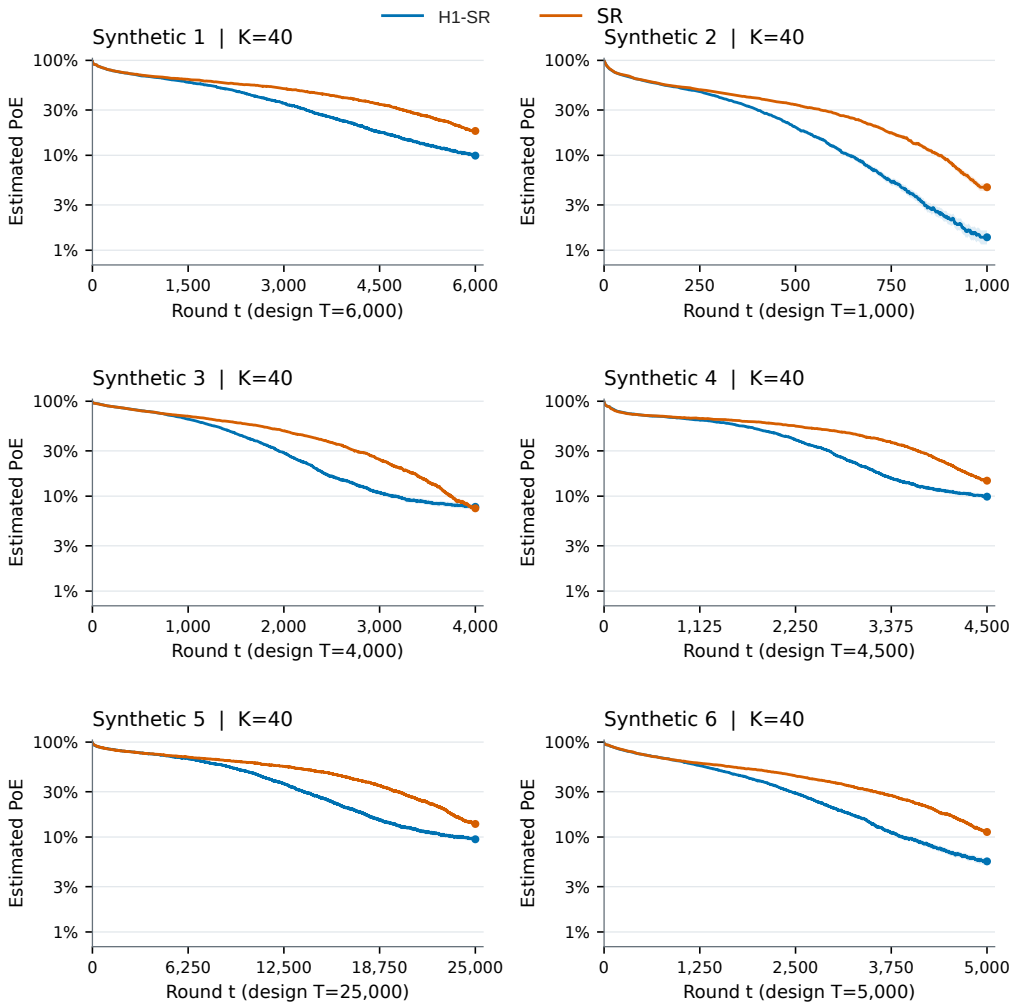


Figure 2: Trajectory comparisons for synthetic instances 1–6. The terminal points correspond to the fixed-horizon results in Table 1.

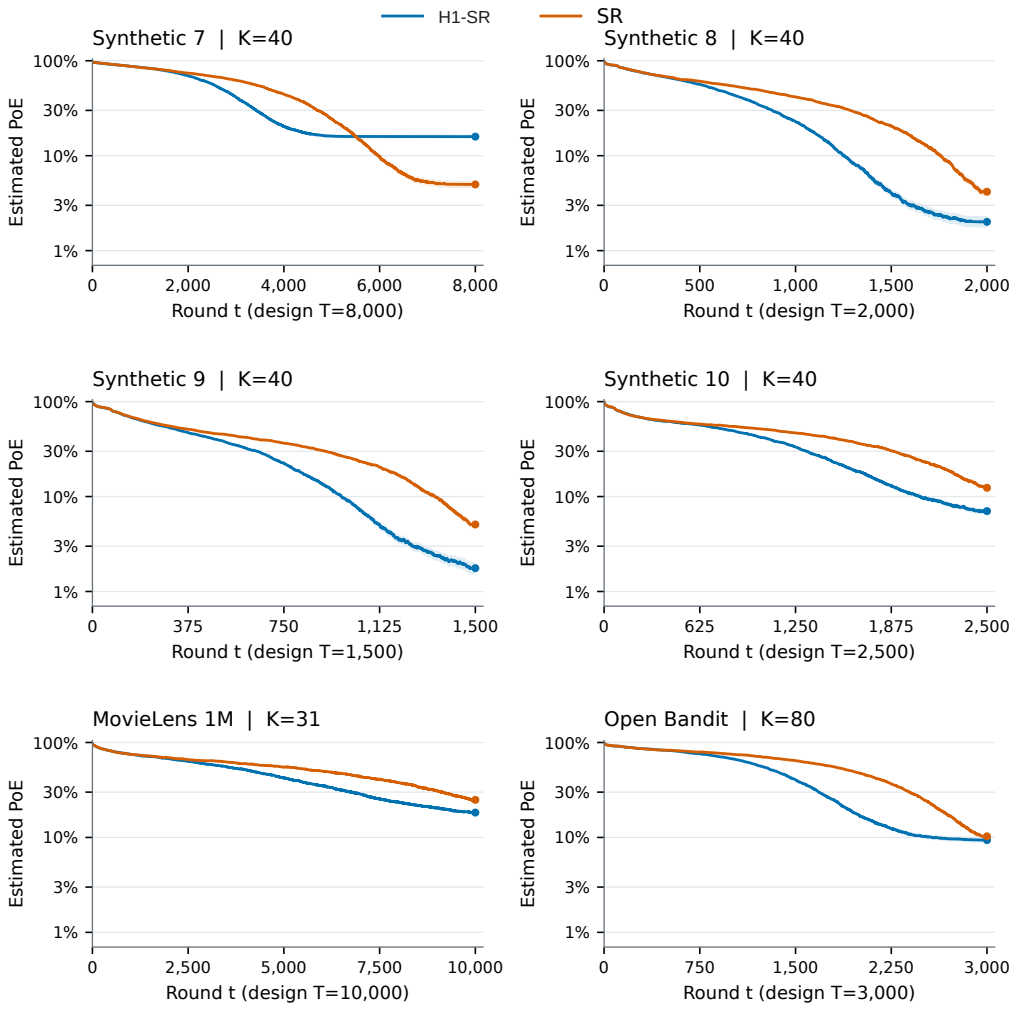


Figure 3: Trajectory comparisons for synthetic instances 7–10, MovieLens 1M, and Open Bandit. Synthetic instance 7 illustrates the finite-budget cost of allocating less sampling effort before early irreversible eliminations.

B JOINT REJECTION BOUNDS AND GAUSSIAN SHARPNESS

For the proofs, we use the equivalent construction in which each arm has an independent stream of potential rewards. Its first m rewards define the sample mean $\overline{Q}_{i,m}$, even on paths where the algorithm does not actually collect all of them. At a deterministic count these sample means have their unconditional laws; we never condition those laws on survival. This construction does not change the sequential observation model in the main text.

Lemma 7 (A fixed-subset rejection bound)**.** *Suppose independent rewards from every arm are σ^2-sub-Gaussian. Fix a deterministic set $A = \{i^\star, i_1, \dots, i_{s-1}\}$ of $s \geq 2$ arms containing the true best arm. Define its true gaps by $\delta_j = P_{i^\star} - P_{i_j} > 0$, and put $H_A = \sum_{j=1}^{s-1} \delta_j^{-2}$. At a common deterministic count $m \geq 1$, the probability that the best arm has the smallest sample mean in this set is at most*

$$\exp\left\{-\frac{m(s-1)^2}{2s\sigma^2 H_A}\right\}. \tag{13}$$

Proof. Translate the best mean to zero and write $\nu = (0, -\delta_1, \dots, -\delta_{s-1})$. Let X be the vector of shifted sample means, so that $\mathbb{E}X = \nu$. The joint rejection event is the cone

$$\mathcal{C} = \{x \in \mathbb{R}^s : x_0 \leq x_j \text{ for } 1 \leq j < s\}.$$

Let x° be the Euclidean projection of ν onto $\mathcal{C}$, and set $v = x^\circ - \nu$. The projection inequality gives $v^\top(x - \nu) \geq \|v\|^2$ for every $x \in \mathcal{C}$. Independence and the sub-Gaussian assumption give, for every $\lambda \in \mathbb{R}$,

$$\mathbb{E}\exp\{\lambda v^\top (X - \nu)\} \leq \exp\left\{\frac{\lambda^2 \sigma^2 \|v\|^2}{2m}\right\}.$$

Applying the Chernoff bound at $\lambda = m/\sigma^2$ therefore implies

$$\mathbb{P}(X \in \mathcal{C}) \leq \exp\left\{-\frac{m \operatorname{dist}(\nu, \mathcal{C})^2}{2\sigma^2}\right\}. \tag{14}$$

For a proposed first coordinate $-u$, the closest feasible competitor coordinate is $\max\{-\delta_j, -u\}$. Values $u < 0$ cannot improve on $u = 0$. Consequently,

$$\operatorname{dist}(\nu, \mathcal{C})^2 = \min_{u \geq 0}\left\{u^2 + \sum_{j=1}^{s-1}(\delta_j - u)_+^2\right\}.$$

For any $u \geq 0$, put $d_j = (\delta_j - u)_+$ and $V = u^2 + \sum_j d_j^2$. Since $\delta_j \leq u + d_j$, Cauchy–Schwarz gives

$$\begin{aligned} s - 1 &\leq u\sum_j \delta_j^{-1} + \sum_j d_j \delta_j^{-1} \\ &\leq \sqrt{V}\sqrt{\left(\sum_j \delta_j^{-1}\right)^2 + H_A} \leq \sqrt{sH_A V}. \end{aligned}$$

For the second inequality, apply Cauchy–Schwarz to the vectors

$$a = (u, d_1, \dots, d_{s-1}), \qquad b = \left(\sum_j \delta_j^{-1}, \delta_1^{-1}, \dots, \delta_{s-1}^{-1}\right).$$

Their inner product is $u\sum_j \delta_j^{-1} + \sum_j d_j\delta_j^{-1}$, while $\|a\|^2 = V$ and $\|b\|^2 = (\sum_j \delta_j^{-1})^2 + H_A$; hence $a^\top b \leq \|a\|\,\|b\|$ gives the stated bound. The last inequality uses $(\sum_j \delta_j^{-1})^2 \leq (s-1)H_A$, again by Cauchy–Schwarz. Thus $\operatorname{dist}(\nu, \mathcal{C})^2 \geq (s-1)^2/(sH_A)$. Together with (14), this proves (13), and hence (5).

□

The error upper bound alone does not establish an exact exponent. The following Gaussian lower bound is used in the hard instances of Appendices C and D.

Lemma 8 (Exact Gaussian rejection exponent)**.** *With Gaussian rewards of common variance σ^2, the rejection probability for a fixed set satisfies*

$$\lim_{m\to\infty} -\frac{1}{m}\log \mathbb{P}(X \in \mathcal{C}) = \frac{\operatorname{dist}(\nu,\mathcal{C})^2}{2\sigma^2}.$$

For equal gaps Δ, this limit is $(s-1)\Delta^2/(2s\sigma^2)$.

Proof. Here $X \sim N(\nu, \sigma^2 I_s/m)$. Choose interior points $y_\varepsilon \in \mathcal{C}$ converging to x° and a ball $B(y_\varepsilon, r)$ contained in $\mathcal{C}$. Integrating the Gaussian density over that ball yields

$$\mathbb{P}(X \in \mathcal{C}) \geq \operatorname{vol}(B(y_\varepsilon, r)) \left(\frac{m}{2\pi\sigma^2}\right)^{s/2} \times \exp\left\{-\frac{m}{2\sigma^2} \sup_{x\in B(y_\varepsilon,r)} \|x-\nu\|^2\right\}.$$

Take the upper limit of $-m^{-1}\log$ of the probability, then let $r \downarrow 0$ and $\varepsilon \downarrow 0$. Combining the result with (14) proves

$$\lim_{m\to\infty} -\frac{1}{m}\log \mathbb{P}(X \in \mathcal{C}) = \frac{\operatorname{dist}(\nu,\mathcal{C})^2}{2\sigma^2}. \tag{15}$$

For equal gaps, the minimizer has every coordinate equal to $-(s-1)\Delta/s$, and its squared distance is $(s-1)\Delta^2/s$. Interior points can be obtained by moving its first coordinate down by ε and all other coordinates up by ε. This proves the claimed equal-gap rate. Ties have probability zero. □

C THE BALANCED CLOCK AND ITS GAUSSIAN WORST-INSTANCE RATE

Budget feasibility and phase balancing. The coefficients $a_s = 2s/(s-1)^2$ strictly decrease with $s \geq 2$. For example, their continuous extension has derivative $-2(s+1)/(s-1)^3 < 0$. Thus the checkpoints in (8) are nondecreasing across successive rejection phases. Summing each rejected arm's final cumulative count and the counts of both finalists gives

$$\begin{aligned}\text{number of observations used} &= 2n_{K-1} + \sum_{k=1}^{K-2} n_k \\ &\leq K + \frac{T-K}{D_K}\left(2a_2 + \sum_{s=3}^{K} a_s\right) = T.\end{aligned}$$

This calculation includes $T = K$. Any remaining observations are taken after the recommendation is stored and do not affect it.

As in Section 2.5, $m_s(T)$ is the cumulative per-arm count at the s-arm checkpoint and $q_s = \lim_{T\to\infty} m_s(T)/T$, whenever the limit exists. For H_1-SR, $m_s(T) = n_{K+1-s}$ and $q_s = a_s/D_K$. For any such limiting schedule the budget constraint is

$$2q_2 + \sum_{s=3}^{K} q_s \leq 1. \tag{16}$$

If $\gamma = \min_{2\leq s\leq K} q_s/a_s$, then $q_s \geq \gamma a_s$ for every s. Hence $1 \geq \gamma(2a_2 + \sum_{s=3}^{K} a_s) = \gamma D_K$. Equality is attained by $q_s = a_s/D_K$, which satisfies the required ordering. Equality also forces every coordinate to have this value. This proves the balancing assertion of Proposition 2. The statistical upper bound showing optimality beyond this max–min calculation is proved in Appendix D.

Proof of Theorem 3. If the best arm is first rejected in phase k, its active set has size $s = K+1-k$. For each possible deterministic active subset A of this size, the rejection event is contained in the

event of Lemma 7, evaluated at the deterministic count n_k. The subset complexity satisfies $H_A \leq H_1(\boldsymbol{P})$, while the rounding in (8) ensures

$$\frac{n_k}{a_s} \geq \frac{T-K}{D_K}.$$

All these subset events share the exponential bound in (9), before multiplication by its prefactor. There are

$$\sum_{s=2}^{K} \binom{K-1}{s-1} = 2^{K-1} - 1$$

subsets containing the best arm and at least one competitor. A union bound, and the trivial bound by one, prove the finite-budget statement. Taking the horizon limit at fixed K and a fixed instance gives only the lower bound

$$\liminf_{T\to\infty} \frac{1}{T} \log \frac{1}{\mathrm{PoE}_{H_1\text{-SR}}(\boldsymbol{P})} \geq \frac{1}{\sigma^2 D_K H_1(\boldsymbol{P})}.$$

For the Gaussian model defining the normalized rate, it follows that

$$R_{H_1\text{-SR}} \geq \frac{1}{\sigma^2 D_K}.$$

□

D Optimality within the class of SR-type algorithms

The following proof goes as following: For each s, we construct an instance that is difficult at the checkpoint with s arms remaining, and then take the worst case over s.

Proof of Theorem 4. Work in the Gaussian model and consider an algorithm $\mathcal{A}$ in the class of SR-type algorithms. In particular, each rejection removes an empirical worst arm; only its deterministic common checkpoints $m_K(T) \leq \cdots \leq m_2(T)$ are allowed to change. By compactness, along a horizon subsequence all fractions $m_s(T)/T$ converge to limits q_s. They satisfy $0 \leq q_K \leq \cdots \leq q_2$ and (16). If $R_{\mathcal{A}} = 0$, the result is immediate. Otherwise fix any finite $0 < r < R_{\mathcal{A}}$.

A positive rate requires $q_K > 0$. Indeed, on a fixed equal-gap Gaussian instance the first-checkpoint rejection probability is at least $c_0 e^{-c_1 m_K(T)}$ for constants $c_0, c_1 > 0$. This follows from the Gaussian ball bound above, uniformly over integer counts $m_K(T) \geq 1$. If $q_K = 0$, then $m_K(T) = o(T)$ along the subsequence, so the full-sequence lower error exponent is zero, contradicting $R_{\mathcal{A}} > 0$. This boundary case is needed before choosing the far-arm gap below.

Fix $s \in \{2, \ldots, K\}$ and $\Delta > 0$. For a fixed $M > \Delta$, consider the instance with one best mean 0, $s-1$ near means $-\Delta$, and $K-s$ far means $-M$. Let E_T be the unconditioned event that all near competitors have sample means larger than the best at count $m_s(T)$. Lemma 8 gives, along the selected subsequence,

$$\mathbb{P}(E_T) = \exp\{-TI_s + o(T)\}, \qquad I_s = \frac{q_s(s-1)\Delta^2}{2s\sigma^2}. \tag{17}$$

Let S_T denote the event that all s near arms, including the best, survive to the s-arm checkpoint. If S_T fails, some near arm is rejected at an earlier checkpoint while a far arm remains active. Because the rejected arm is empirically worst, at that checkpoint a far sample mean is at least as large as a near sample mean. Every relevant true gap is at least $M - \Delta$, and every checkpoint count is at least $m_K(T)$. The Gaussian two-arm tail and a finite union over arms and earlier checkpoints yield a constant $L_K < \infty$ such that

$$\mathbb{P}(S_T^c) \leq L_K \exp\left\{-\frac{m_K(T)(M-\Delta)^2}{4\sigma^2}\right\}. \tag{18}$$

For $s = K$, the event S_T is certain and this bound is unnecessary. Choose M large enough that $q_K(M-\Delta)^2/(4\sigma^2) > I_s$. On $E_T \cap S_T$ the best arm is rejected at the s-arm checkpoint. Thus, without any independence assumption between these events,

$$\mathrm{PoE}_{\mathcal{A}}(\boldsymbol{P}) \geq \mathbb{P}(E_T) - \mathbb{P}(S_T^c) = \exp\{-TI_s + o(T)\}.$$

The full-sequence lower exponent is no larger than the corresponding upper exponent along this subsequence. Since $H_1(\boldsymbol{P}) = (s-1)/\Delta^2 + (K-s)/M^2$, we obtain

$$r \le \left(\frac{s-1}{\Delta^2} + \frac{K-s}{M^2} \right) I_s.$$

Here the horizon limit has been taken with M fixed. Now let $M \to \infty$, giving

$$r \le \frac{q_s(s-1)^2}{2s\sigma^2} = \frac{q_s}{\sigma^2 a_s}.$$

This holds for every $s \in \{2, \dots, K\}$, so

$$r \le \min_{2 \le s \le K} \frac{q_s}{\sigma^2 a_s}.$$

Thus the smallest of these bounds over all checkpoints limits the rate. Equivalently, $q_s \ge \sigma^2 r a_s$ for every s, and (16) implies $1 \ge \sigma^2 r D_K$. Since this holds for every finite $0 < r < R_{\mathcal{A}}$, it gives $R_{\mathcal{A}} \le 1/(\sigma^2 D_K)$. Theorem 3 and Appendix C show that H_1-SR attains this value. The proof does not assume that a schedule's normalized checkpoints converge along the full horizon sequence. □

E A CONVERSE FOR ALL ADAPTIVE ALGORITHMS

The one-shot *game-optimal* bound R^{go} in Komiyama et al. (2022; 2026) is an upper bound on attainable normalized rates for any algorithm. Our argument uses essentially the same change-of-measure principle. For Gaussian rewards with common variance σ^2 and the complexity H_1 used here, our construction gives B_K such that

$$R^\star_K \le R^{\mathrm{go}} \le \frac{1}{\sigma^2 B_K}.$$

Thus $1/(\sigma^2 B_K)$ is an explicit upper bound on the one-shot game value, but not necessarily its exact value, and it is a useful simple formula. We first prove the allocation statement needed to impose several constraints on one common vector. An algorithm may choose which arm to sample using past observations, internal randomness, and the known budget T. The true arm means are unknown and are not inputs to the algorithm. Throughout this section rewards are Gaussian with the common variance σ^2. We use $\boldsymbol{\theta}$ for the true mean vector of a reference instance whose unique best arm is arm 1, and $\boldsymbol{P}$ for the true mean vector of an alternative instance with a different best arm.

The following allocation statement follows directly from the definition of R^{go} in Komiyama et al. (2022; 2026) and holds for every finite $0 < r < R^{\mathrm{go}}$. We derive it here for completeness.

Lemma 9 (One allocation for every fixed alternative)**.** *Fix a finite r with $0 < r < R^{\mathrm{go}}$. For any fixed unique-best instance $\boldsymbol{\theta}$ with best arm 1, there exists $u(\boldsymbol{\theta})$ in the probability simplex such that, for every fixed unique-best instance $\boldsymbol{P}$ whose best arm is not 1,*

$$r \le \frac{H_1(\boldsymbol{P})}{2\sigma^2} \sum_{i=1}^{K} u_i(\boldsymbol{\theta})(\theta_i - P_i)^2. \tag{19}$$

Proof. Since $r < R^{\mathrm{go}}$, the definition of the one-shot game provides an allocation map $w(\cdot)$ into the probability simplex such that

$$r \le H_1(\boldsymbol{P}) \sum_{i=1}^{K} w_i(\boldsymbol{\theta}) \, \mathrm{KL}\big(N(\theta_i, \sigma^2) \,\|\, N(P_i, \sigma^2)\big)$$

for every unique-best reference instance $\boldsymbol{\theta}$ and every unique-best alternative $\boldsymbol{P}$ with a different best arm. For Gaussian distributions with common variance, the KL divergence is $(\theta_i - P_i)^2/(2\sigma^2)$. Setting $u(\boldsymbol{\theta}) = w(\boldsymbol{\theta})$ proves (19). The same allocation map works for all these pairs; in particular, $u(\boldsymbol{\theta})$ does not depend on the alternative $\boldsymbol{P}$. The strict inequality $r < R^{\mathrm{go}}$ avoids assuming that the supremum defining the game value is attained. □

We now use this allocation statement to prove $R^{\mathrm{go}} \le 1/(\sigma^2 B_K)$. We first derive constraints on a common allocation vector using Gaussian instances with gaps at different scales, then show that these constraints imply $1 \ge \sigma^2 r B_K$ for every finite $0 < r < R^{\mathrm{go}}$. For a nonnegative vector u and $2 \le s \le K$, define

$$\Psi_s(u) = \inf_{\substack{z_i>0,\ i<s \\ t\in\mathbb{R}}} \left(\sum_{i<s} z_i^{-2}\right)\left[u_s t^2 + \sum_{i<s} u_i (z_i - t)^2\right]. \tag{20}$$

Lemma 10 (Simultaneous prefix information constraints)**.** *For every finite r with $0 < r < R^{\mathrm{go}}$, there exists one probability vector u satisfying*

$$\Psi_s(u) \ge 2\sigma^2 r, \qquad 2 \le s \le K. \tag{21}$$

Proof. For $0 < \varepsilon < 1$, choose the strictly ranked reference instance

$$\theta_1^\varepsilon = 0, \qquad \theta_j^\varepsilon = -\sum_{\ell=2}^{j} \varepsilon^{2(K-\ell+1)}, \quad j \ge 2.$$

Apply Lemma 9 for each fixed $\boldsymbol{\theta}^\varepsilon$. Along a sequence $\varepsilon \downarrow 0$, extract a limit u of the resulting allocation vectors. For each $s \in \{2, \dots, K\}$, we call the arms $\{1, \dots, s\}$ the *prefix* and the remaining arms $\{s+1, \dots, K\}$ the *suffix*, which is empty when $s = K$. Put $L_s = \varepsilon^{2(K-s)+1}$. Then

$$\theta_i^\varepsilon / L_s \to 0 \quad (i \le s), \qquad \theta_i^\varepsilon / L_s \to -\infty \quad (i > s).$$

For fixed positive $z_1, \dots, z_{s-1}$ and $t \ge 0$, define

$$P_s^\varepsilon = L_s t, \qquad P_i^\varepsilon = L_s(t - z_i)\ (i < s), \qquad P_i^\varepsilon = \theta_i^\varepsilon\ (i > s).$$

For sufficiently small ε, arm s is the unique best arm, and the suffix contributes asymptotically nothing to the scaled complexity:

$$L_s^2 H_1(\boldsymbol{P}^\varepsilon) \longrightarrow \sum_{i<s} z_i^{-2}.$$

The suffix means are unchanged, so its contribution to the quadratic information cost is exactly zero. Applying (19) and then taking $\varepsilon \downarrow 0$ yields

$$2\sigma^2 r \le \left(\sum_{i<s} z_i^{-2}\right)\left[u_s t^2 + \sum_{i<s} u_i (z_i - t)^2\right].$$

For fixed positive z, the minimizing translation is $t = (\sum_{i<s} u_i z_i)/(\sum_{i\le s} u_i) \ge 0$ whenever the denominator is positive. If all prefix coordinates vanish, the expression is zero, contradicting $r > 0$. Thus the restriction $t \ge 0$ loses no minimizing value, and taking the infimum proves (21).

The allocation map is fixed before taking $\varepsilon \downarrow 0$. The convergent subsequence of allocation vectors is common to all the finitely many prefixes and all fixed choices of their parameters. No horizon limit is needed in this argument. □

Proof of Theorem 1. If $R^{\mathrm{go}} = 0$, the claimed bound is immediate. Otherwise fix any finite $0 < r < R^{\mathrm{go}}$ and use the vector in Lemma 10. Write $c = \sigma^2 r$. At $s = 2$, direct minimization gives

$$\Psi_2(u) = \frac{u_1 u_2}{u_1 + u_2} \le \frac{u_1 + u_2}{4}.$$

The denominator is positive by the prefix constraint. Hence $u_1 + u_2 \ge 8c$. For $K = 2$, this already gives $1 \ge cB_2$ with $B_2 = 8$. Now suppose $K \ge 3$. For $s \ge 3$, choose $z_i = 1$ in (20) and minimize over t. With $C_s = \sum_{i<s} u_i$, this gives

$$\Psi_s(u) \le (s-1)\frac{C_s u_s}{C_s + u_s} \le (s-1)u_s.$$

Thus $u_s \ge 2c/(s-1)$. To retain the coupling between the first three arms, put $C = u_1 + u_2 \ge 8c$. The same inequality at $s = 3$, before discarding its denominator, gives

$$\frac{C u_3}{C + u_3} \ge c, \qquad u_3 \ge \frac{cC}{C - c}.$$

The function $C \mapsto C^2/(C-c)$ is increasing for $C \geq 8c$, since its derivative is $C(C-2c)/(C-c)^2 > 0$. Consequently,

$$u_1 + u_2 + u_3 \geq \frac{C^2}{C-c} \geq \frac{64c}{7}.$$

Combining this with the individual bounds for $s \geq 4$ yields

$$1 = \sum_i u_i \geq c\left(\frac{64}{7} + 2\sum_{s=4}^{K} \frac{1}{s-1}\right) = c\left(\frac{43}{7} + 2h_{K-1}\right) = \sigma^2 r B_K.$$

An empty sum is zero, so this includes $K = 3$. Because the inequality holds for every finite $0 < r < R^{\mathrm{go}}$, it implies $R^{\mathrm{go}} \leq 1/(\sigma^2 B_K)$, including ruling out an infinite game value. Combining this with the bound $R_K^\star \leq R^{\mathrm{go}}$ from Komiyama et al. (2022; 2026) proves the theorem. □

F THE EXACT EXPONENT OF ORIGINAL SUCCESSIVE REJECTS

The exact error exponent of original SR is established in prior work by Wang et al. (2024) and was rediscovered by Komiyama et al. (2026). Here, we will re-derive it for completeness.

Proof of Theorem 6. We first work with Gaussian rewards and prove the finite-budget bound for $T > K$. At phase k, put $s = K + 1 - k$. If the true best is rejected, at least one true rank $r \in \{s, \ldots, K\}$ remains active and has a sample mean no smaller than the best. Indeed, among s active arms there must be one of those $K - s + 1$ ranks, since only $s - 1$ arms have better true ranks. For a deterministic common count m,

$$\mathbb{P}\{\overline{Q}_{(r),m} \geq \overline{Q}_{i^\star,m}\} = \Phi\left(-\Delta_{(r)}\sqrt{\frac{m}{2\sigma^2}}\right) \leq \frac{1}{2}\exp\left\{-\frac{m\Delta_{(r)}^2}{4\sigma^2}\right\}.$$

Here Φ is the standard normal distribution function; the displayed tail inequality follows, for example, by applying the Gaussian density shift to its integral over the positive half-line. The SR checkpoint satisfies $n_k^{\mathrm{SR}} \geq (T-K)/(s\overline{\log}(K))$. Also $H_1(\boldsymbol{P}) \geq (s-1)/\Delta_{(s)}^2$ and $\Delta_{(r)} \geq \Delta_{(s)}$. Therefore

$$n_k^{\mathrm{SR}}\Delta_{(r)}^2 \geq \frac{T-K}{\overline{\log}(K)H_1(\boldsymbol{P})}\frac{s-1}{s} \geq \frac{T-K}{2\overline{\log}(K)H_1(\boldsymbol{P})}.$$

At phase k, there are k possible ranks $r \in \{K+1-k, \ldots, K\}$ to consider. Summing over all phases gives $\sum_{k=1}^{K-1} k = K(K-1)/2$ events in the union bound. The union bound proves the asserted prefactor $K(K-1)/4$ and exponent, with a further minimum with one if necessary. The explicit one-sample-per-arm convention at $T = K$ removes the zero-count endpoint of the original checkpoint formula and does not affect any asymptotic claim.

For independent σ^2-sub-Gaussian rewards, the difference of the two sample means instead has variance proxy $2\sigma^2/m$. The Chernoff bound gives

$$\mathbb{P}\{\overline{Q}_{(r),m} \geq \overline{Q}_{i^\star,m}\} \leq \exp\left\{-\frac{m\Delta_{(r)}^2}{4\sigma^2}\right\}.$$

The same union bound over phases and ranks therefore proves the same finite-budget exponential term with prefactor $K(K-1)/2$, rather than $K(K-1)/4$. The additional factor $1/2$ per pairwise event above is Gaussian-specific.

To prove Gaussian sharpness, first let $K \geq 3$ and consider the fixed instance

$$\boldsymbol{P}^{(M)} = (0, -1, -M, \ldots, -M), \qquad M > 1.$$

Write $m = n_{K-1}^{\mathrm{SR}}$ and define the unconditioned final reversal event $E_T = \{\overline{Q}_{2,m} \geq \overline{Q}_{1,m}\}$. Since $m/T \to 1/(2\overline{\log}(K))$, the normal-tail exponent gives

$$\mathbb{P}(E_T) = \exp\left\{-\frac{T}{8\sigma^2\overline{\log}(K)} + o(T)\right\}. \tag{22}$$

Let S_T be the event that arms 1 and 2 both reach the final phase. If either is removed earlier, a far arm must then have a sample mean at least as large. The smallest relevant gap is $M-1$ and the smallest checkpoint fraction tends to $1/(K\overline{\log}(K))$. A finite union bound consequently gives

$$\mathbb{P}(S_T^c) \le \exp\left\{-\frac{T(M-1)^2}{4\sigma^2 K\overline{\log}(K)} + o(T)\right\}. \tag{23}$$

Choose fixed M such that $(M-1)^2 > K/2$. The exponent in (23) is strictly larger than the exponent in (22). On $S_T \cap E_T$, SR rejects arm 1 in its final phase. Hence

$$\mathrm{PoE}_{\mathrm{SR}}(\boldsymbol{P}^{(M)}) \ge \mathbb{P}(E_T) - \mathbb{P}(S_T^c) = \exp\left\{-\frac{T}{8\sigma^2\overline{\log}(K)} + o(T)\right\}.$$

Since $H_1(\boldsymbol{P}^{(M)}) = 1 + (K-2)/M^2$, first taking $T \to \infty$ for this fixed instance and then $M \to \infty$ proves

$$R_{\mathrm{SR}} \le \frac{1}{8\sigma^2\overline{\log}(K)}.$$

For $K = 2$, the same upper bound follows directly from the final two-arm Gaussian comparison. The finite-budget guarantee supplies the reverse inequality for every K, proving exact equality. □

G Uniform factors and the large-arm limit

Proof of Corollary 5. Theorems 3 and 1 immediately imply

$$\frac{R_{H_1\text{-SR}}}{R_K^\star} \ge \frac{B_K}{D_K}.$$

For $K = 2$, $B_2 = D_2 = 8$ and the algorithm is optimal. We use $11/12$ as a convenient rational lower bound, rather than as the exact minimum of B_K/D_K that is a bit involved. Since $D_K > 0$, it suffices to show that $12B_K - 11D_K \ge 0$ for every K. For $K \ge 3$, put $m = K-1$ and define

$$G_m = 12B_K - 11D_K = \frac{208}{7} + 2h_m - 22h_m^{(2)}.$$

For $m \ge 3$, its successive differences are

$$G_m - G_{m-1} = \frac{2(m-11)}{m^2}.$$

Thus, over the integers $m \ge 2$, the minimum is attained at $m = 10$ and $m = 11$. Exact arithmetic gives

$$G_{10} = G_{11} = \frac{187633}{127008} > 0.$$

It follows that $B_K/D_K > 11/12$ for every $K \ge 3$, proving the uniform factor.

The exact SR rate and the all-algorithm sandwich give

$$\frac{B_K}{8\overline{\log}(K)} \le \frac{R_{\mathrm{SR}}}{R_K^\star} \le \frac{D_K}{8\overline{\log}(K)}.$$

Moreover,

$$\frac{B_K}{8\overline{\log}(K)} \ge \frac{1}{4} + \frac{7 - 2/K}{8\overline{\log}(K)} > \frac{1}{4},$$

because $B_K \ge 6 + 2h_{K-1}$ and $\overline{\log}(K) = h_{K-1} + 1/K - 1/2$. Since $h_{K-1} = \log K + O(1)$ and $h_{K-1}^{(2)} = O(1)$, both B_K and D_K equal $2\log K + O(1)$, whereas $8\overline{\log}(K) = 8\log K + O(1)$. More precisely, for $K \ge 3$,

$$D_K - B_K = 2h_{K-1}^{(2)} - \frac{15}{7} = O(1).$$

Using $h_{K-1} = \log K + \gamma + O(K^{-1})$ and $h_{K-1}^{(2)} = \pi^2/6 + O(K^{-1})$, where γ is Euler's constant, also gives the first-order expansion

$$\frac{B_K}{D_K} = 1 - \frac{\pi^2/6 - 15/14}{\log K} + O\big((\log K)^{-2}\big).$$

Since the achieved rate cannot exceed the all-algorithm optimum,

$$0 \le 1 - \frac{R_{H_1\text{-SR}}}{R_K^\star} \le \frac{D_K - B_K}{D_K} = O\left(\frac{1}{\log K}\right).$$

The preceding bounds also give

$$\frac{R_{H_1\text{-SR}}}{R_K^\star} \longrightarrow 1, \qquad \frac{R_{\mathrm{SR}}}{R_K^\star} \longrightarrow \frac{1}{4}.$$

The exact improvement of the balanced rate over SR is $8\overline{\log}(K)/D_K \to 4$. These are limits of fixed-K asymptotic error exponents: the horizon limit is taken first. In particular, no claim about discarding a K-dependent finite-budget prefactor when K grows with T is needed. □

H Non-dominance of H_1-SR over original SR

At a fixed budget T, we say that an algorithm $\mathcal{A}$ dominates another algorithm $\mathcal{A}'$ if

$$\mathrm{PoE}_{\mathcal{A}}(\boldsymbol{P}) \le \mathrm{PoE}_{\mathcal{A}'}(\boldsymbol{P}) \quad \text{for every unique-best instance } \boldsymbol{P},$$

with strict inequality for at least one instance. Both probabilities are evaluated at the same budget and under the same true reward distributions.

The optimality of H_1-SR concerns the worst-instance normalized asymptotic rate $R_{\mathcal{A}}$. Maximizing this criterion can involve tradeoffs between performance on different instances $\boldsymbol{P}$ and $\boldsymbol{P}'$. It does not imply dominance over original SR at a fixed budget.

Synthetic instance 7 illustrates this distinction. It has one best arm with mean 1 and 39 competitors with mean 0.8, all with unit-variance Gaussian rewards. In this case, the first elimination matters most to the probability of correctly identifying the best arm. At $T = 8000$, H_1-SR takes 27 observations per arm before the first elimination, compared with 53 for original SR. In other words, H_1-SR allocates fewer samples before early eliminations. As a result, SR outperforms H_1-SR in this instance.

This provides empirical evidence against finite-budget dominance by H_1-SR, despite its better worst-instance normalized asymptotic rate. Such tradeoffs across underlying distributions are common in statistical decision theory (Jennison et al., 1982): a procedure can perform better on some distributions and worse on others, so neither procedure dominates the other.

This differs from fixed-confidence best-arm identification, where algorithms such as Track-and-Stop can attain the instance-dependent lower bound on expected sample complexity asymptotically as $\delta \to 0$ (Garivier & Kaufmann, 2016). Such algorithms are asymptotically optimal on each instance, so attaining the optimal leading-order sample complexity on one instance does not require sacrificing it on another. In contrast, our optimality guarantee concerns the worst-instance normalized asymptotic rate. Within the class of SR-type algorithms, improving performance on some instances can come at the cost of worse performance on others.

I The inverse-gap Complexity Measures

Our complexity is the sum over suboptimal arms only:

$$H_1(\boldsymbol{P}) = \sum_{i \ne i^\star} \Delta_i^{-2}.$$

In the original SR paper, Audibert et al. (2010) assign the best arm the artificial gap $\Delta_{(2)}$ and include its inverse square in their complexity. Writing their complexity as H_1^{SR}, we have

$$H_1^{\mathrm{SR}} = H_1 + \Delta_{(2)}^{-2}, \qquad \frac{K}{K-1} H_1 \le H_1^{\mathrm{SR}} \le 2H_1.$$

Indeed, $\Delta_{(2)}^{-2}$ is the largest of the $K-1$ terms in H_1, so it lies between $H_1/(K-1)$ and H_1.