The Universal Price of Best-Arm Identification

Pahan Dewasurendra
Johns Hopkins University

Abstract

The static oracle for fixed-budget best-arm identification knows the arm means and selects the allocation with the best error exponent. No adaptive algorithm can match this oracle on every instance when there are at least three arms. How large must the worst-case loss be?

For unit-variance Gaussian arms, we determine this price to within an additive constant below 0.2 for every number K of arms. If P_K is the smallest possible worst-case ratio between the oracle and algorithmic error exponents, then

$$C_K \leq P_K \leq U_K, \qquad 0 < U_K - C_K < 0.2,$$

where both endpoints are explicit sequences and equal $\log K + O(1)$. Thus the exact leading constant in the price of universality is one.

The lower bound strengthens a recent impossibility result using a translated equal-gap hard family. The upper bound is constructive. We optimize the batch lengths of a successive-elimination design and prove that its exponent is at least $\Gamma^\star_{\text{so}}/U_K$ on every instance. The schedule is exactly minimax within the full class of fixed-weight, one-at-a-time batched elimination designs. The key is a sharp information inequality relating the static oracle to the large-deviation cost of incorrectly eliminating the best arm. Equal-gap active sets are the unique extremizers. The optimized design strictly improves the classical Successive Rejects worst-case ratio for every $K \geq 3$.

1 Introduction

Fixed-budget best-arm identification allocates T observations among K unknown arms and recommends the arm with largest mean. Its central performance quantity is the exponential decay rate of the error probability. If the means were known before sampling, the best static allocation would attain an instance-dependent exponent $\Gamma^\star_{\text{so}}(\mu)$ (Glynn and Juneja, 2004). This *static oracle* is a natural benchmark in bandit learning, ranking and selection, and adaptive experimentation.

The benchmark cannot be achieved uniformly. Qin (2022) posed the fixed-budget complexity question, and Degenne (2023) formalized the obstruction through the difficulty ratio. Very recent work closed the qualitative question: for every $K \geq 3$ and every regular one-parameter natural exponential family, each consistent algorithm loses a strict factor on some instance (Goldberger, 2026). For Gaussian arms, that work proves the explicit lower bound

$$L_K := 1 + \sum_{j=3}^{K} \frac{1}{(1+\sqrt{j-1})^2}. \tag{1}$$

The exact sum is $\log K + O(1)$, although its stated simpler corollary is $1 + \log(K)/8$.

This leaves a quantitative problem. Is the lower bound an artifact of a hard family, or is a logarithmic loss sufficient? Existing elimination algorithms have order-$\log K$ gap-based guarantees (Audibert et al., 2010; Carpentier and Locatelli, 2016; Degenne, 2023), but these do not identify the leading constant relative to $\Gamma^\star_{\text{so}}$. Recent large-deviation analyses give refined exponents for Successive Rejects and Continuous Rejects (Wang et al., 2023, 2024). Batched arm elimination can also strictly dominate uniform allocation (Imbens et al., 2025). None of these results optimizes the worst-case static-oracle ratio.

Contributions. We nearly determine the minimax price of a universal fixed-budget algorithm. Let

$$U_K := 1 + \sum_{j=3}^{K} \frac{j}{(j-1)(1+\sqrt{j-1})^2}. \tag{2}$$

Put $c_j = (1+\sqrt{j-1})^2$ and define the second explicit sequence by

$$C_2 = 1, \qquad C_j = C_{j-1} + \frac{C_{j-1}}{c_j C_{j-1} - 1}. \tag{3}$$

For unit-variance Gaussian arms, we prove

$$C_K \le P_K \le U_K, \qquad U_K - C_K < 0.2, \tag{4}$$

where P_K is the infimum, over all consistent algorithms, of the worst-case ratio $\Gamma^\star_{\mathrm{so}}/\Gamma_{\mathsf{A}}$. In particular, $\lim_{K\to\infty} P_K/\log K = 1$.

The upper bound is attained by a simple optimized batched elimination design. When j arms remain, every surviving arm has received a cumulative fraction proportional to

$$a_j := \frac{j}{(j-1)(1+\sqrt{j-1})^2}. \tag{5}$$

The resulting batch weights are explicit, positive, and independent of the instance. We prove that this schedule is exactly optimal among all fixed-weight one-at-a-time batched elimination designs. This is a minimax theorem for the algorithm class studied by Imbens et al. (2025), not only a sufficient condition for improving uniform allocation.

The lower endpoint improves (1). It follows from a continuum of translated equal-gap alternatives around one hierarchical baseline and an exact allocation program. Our main upper-bound ingredient is a sharp comparison between two information geometries. The static oracle is controlled by pairwise best-versus-challenger deviations. Elimination fails when the best empirical mean falls below every arm in an active set. We prove that the ratio of these costs is at most a_j for every instance with j active arms. Equality holds exactly when all $j-1$ gaps are equal. The proof requires a reciprocal-truncated-variance inequality because, for unequal gaps, the information projection need not move every challenger.

Finally, classical Successive Rejects has exact worst-case oracle ratio $\ell_K = \frac{1}{2} + \sum_{i=2}^{K} i^{-1}$. Our optimized schedule replaces this by $U_K < \ell_K$ for every $K \ge 3$. The asymptotic additive improvement is about 1.28.

2 Problem and Price

Arm $i \in [K]$ returns independent observations from $\mathcal{N}(\mu_i, 1)$. Let $b = b(\mu)$ be the unique maximizer of μ_i and write $\mathcal{M}_K$ for all such mean vectors. An algorithm family $\mathsf{A} = (\mathsf{A}_T)_{T\ge 1}$ chooses arms sequentially, then recommends $\widehat{b}_T$. Its error probability and lower error exponent are

$$p_{\mu,T}(\mathsf{A}) = \mathbb{P}_\mu(\widehat{b}_T \ne b),$$
$$\Gamma_{\mathsf{A}}(\mu) = \liminf_{T\to\infty} \frac{1}{T} \log \frac{1}{p_{\mu,T}(\mathsf{A})}.$$

The family is consistent if $p_{\mu,T}(\mathsf{A}) \to 0$ for every $\mu \in \mathcal{M}_K$.

For fixed proportions $w \in \Delta_K$, sampling each arm $w_i T + o(T)$ times and recommending the largest empirical mean has exponent

$$\Gamma(\mu, w) = \min_{i\ne b} \frac{w_b w_i}{w_b + w_i} \frac{(\mu_b - \mu_i)^2}{2}. \tag{6}$$

The static-oracle exponent is

$$\Gamma^\star_{\mathrm{so}}(\mu) = \max_{w\in\Delta_K} \Gamma(\mu, w). \tag{7}$$

Zero coordinates are allowed by continuity. This is the reciprocal of the static-oracle difficulty used by Degenne (2023) and Goldberger (2026).

Definition 1 (Price of universality)**.** *The minimax static-oracle ratio is*

$$P_K = \inf_{\mathsf{A}\ \textit{consistent}} \sup_{\mu\in\mathcal{M}_K} \frac{\Gamma^\star_{\mathrm{so}}(\mu)}{\Gamma_{\mathsf{A}}(\mu)}, \tag{8}$$

with the ratio infinite when $\Gamma_{\mathsf{A}}(\mu) = 0$.

This criterion differs from minimax analysis over a prescribed gap hardness (Komiyama et al., 2022). It asks for the smallest simultaneous loss relative to the instance-specific static oracle. It also differs from large-deviation admissibility, which asks whether one fixed design dominates another (Wang et al., 2024; Imbens et al., 2025).

3 Batched Elimination

We use the fixed-weight batched arm elimination class. Start with all arms. Within each batch, sample every surviving arm equally. At the end of a batch, discard the arm with smallest empirical mean. Continue until one arm remains.

It is convenient to parametrize a design by cumulative per-arm fractions. Let q_j be the fraction of the total budget received by every arm that survives through the batch with j active arms. A schedule is feasible when

$$q_2 \ge q_3 \ge \cdots \ge q_K \ge 0, \qquad 2q_2 + \sum_{j=3}^{K} q_j = 1. \tag{9}$$

For budget T, use cumulative counts $N_j = \lfloor q_j T \rfloor$ and give any unused observations to the last two arms. Equivalently, the idealized batch fractions are

$$\beta_K = K q_K, \qquad \beta_j = j(q_j - q_{j+1}), \quad 2 \le j < K. \tag{10}$$

They are nonnegative and telescope to one.

For $J \subseteq [K]$ containing b, define the information needed to make the best arm empirically worst within J:

$$I_J(\mu) = \inf_{\lambda_b \le \min_{i \in J \setminus \{b\}} \lambda_i} \frac{1}{2} \sum_{i \in J} (\lambda_i - \mu_i)^2. \quad (11)$$

Set

$$I_j(\mu) = \min_{J \ni b,\ |J|=j} I_J(\mu). \quad (12)$$

Large-deviation analyses of elimination give the following bound (Wang et al., 2023; Imbens et al., 2025). We include a direct Gaussian argument in the supplement.

Proposition 2 (Finite-budget elimination bound)**.** *For every feasible schedule q and every T,*

$$p_{\mu,T}(\mathrm{BAE}(q)) \le \sum_{j=2}^{K} \binom{K-1}{j-1} \exp\{-N_j I_j(\mu)\}. \quad (13)$$

If $N_j = 0$, its summand is interpreted as one. Consequently,

$$\Gamma_{\mathrm{BAE}(q)}(\mu) \ge \min_{2 \le j \le K} q_j I_j(\mu). \quad (14)$$

Proof. If the algorithm errs, the best arm is empirically smallest in some active set J when it is discarded. At the j-arm stage, every member of J has N_j observations. A Gaussian Chernoff bound for this convex event is $\exp\{-N_j I_J(\mu)\}$. Union-bound over the $\binom{K-1}{j-1}$ possible sets containing the best arm, then over stages. The supplement gives the one-line separating-hyperplane derivation of the Gaussian bound. Taking exponents proves (14). $\square$

Expected simple regret and error probability have the same exponent on every fixed unique-best instance. Thus Proposition 2 is also the error-exponent form of the bound in Imbens et al. (2025).

4 Sharp Information Comparison

The key question is how $I_j(\mu)$ compares with $\Gamma^\star_{\mathrm{so}}(\mu)$. A generic pairwise bound loses the sharp constant because the event defining I_j couples all active empirical means.

Theorem 3 (Oracle-to-elimination inequality)**.** *For every $\mu \in \mathcal{M}_K$ and $j \in \{2, \dots, K\}$,*

$$\Gamma^\star_{\mathrm{so}}(\mu) \le a_j I_j(\mu), \qquad a_j = \frac{j}{(j-1)(1+\sqrt{j-1})^2}. \quad (15)$$

The constant is sharp. Equality is approached by instances whose closest $j-1$ challengers have one common gap and whose remaining gaps diverge.

We give the proof because it is the source of the optimized schedule. Order the gaps as $0 < d_2 \le \cdots \le d_K$. Restricting the oracle comparison to the closest $j-1$ challengers can only increase its value. Put $n = j-1$ and $A_i = d_i^2/2$ for $i = 2, \dots, j$. A standard one-dimensional representation of the j-arm static oracle is

$$\Gamma^\star_{\mathrm{so},j} = \max_{0 < x < \min_i A_i} \frac{x}{1 + \sum_{i=2}^{j} x/(A_i - x)}. \quad (16)$$

The projection in (11) gives

$$I_j(\mu) = \min_{r \ge 0} \frac{1}{2} \left\{ r^2 + \sum_{i=2}^{j} (d_i - r)_+^2 \right\}. \quad (17)$$

The closest $j-1$ arms minimize I_J. Details for both reductions are in the supplement.

Fix x in (16) and let $z_i = \sqrt{A_i/x} > 1$. The ratio between its objective and (17) is the reciprocal of AB, where

$$A = 1 + \sum_{i=2}^{j} \frac{1}{z_i^2 - 1}. \quad (18)$$

$$B = \min_{t \ge 0} \left\{ t^2 + \sum_{i=2}^{j} (z_i - t)_+^2 \right\}. \quad (19)$$

Lemma 4 (Reciprocal-truncated variance)**.** *For every $z_1, \dots, z_n > 1$,*

$$\left(1 + \sum_{i=1}^{n} \frac{1}{z_i^2 - 1}\right) \min_{t \ge 0} \left\{ t^2 + \sum_{i=1}^{n} (z_i - t)_+^2 \right\} \ge \frac{n(1+\sqrt{n})^2}{n+1}. \quad (20)$$

Equality holds exactly when $z_i^2 = 1 + \sqrt{n}$ for every i.

Lemma 4 implies that the objective in (16) divided by I_j is at most $(n+1)/(n(1+\sqrt{n})^2) = a_j$. Maximizing over x proves Theorem 3. The equality conditions yield equal gaps.

The truncation in Lemma 4 matters. If gaps are unequal, close challengers can remain fixed in the projection while only a farther suffix and the best arm move. The proof conditions on this active suffix, applies Jensen to reduce it to one scalar, then shows that every inactive coordinate makes the inequality strict. The full argument is in the supplement.

5 An Optimized Universal Design

The constants a_j are decreasing in j. Define U_K by (2) and choose

$$q_j^\star = \frac{a_j}{U_K}, \qquad 2 \le j \le K. \quad (21)$$

Since $a_2 = 1/2$, these cumulative fractions satisfy (9). Equations (10) and (21) give all batch lengths explicitly.

Theorem 5 (Universal upper bound)**.** *The optimized batched elimination design is consistent and satisfies*

$$\Gamma_{\mathrm{BAE}(q^\star)}(\mu) \geq \frac{\Gamma^\star_{\mathrm{so}}(\mu)}{U_K} \quad \text{for every } \mu \in \mathcal{M}_K. \tag{22}$$

Consequently, $P_K \leq U_K$.

Proof. Theorem 3 gives $q_j^\star I_j(\mu) \geq \Gamma^\star_{\mathrm{so}}(\mu)/U_K$ for every j. Apply Proposition 2 and take the minimum over j. The right side is positive on every unique-best instance, which also proves consistency. □

The schedule is not just sufficient. It solves the minimax design problem for the whole fixed-weight batched elimination class.

Theorem 6 (Exact batched-elimination optimum)**.** *For every feasible cumulative schedule q,*

$$\sup_{\mu\in\mathcal{M}_K} \frac{\Gamma^\star_{\mathrm{so}}(\mu)}{\Gamma_{\mathrm{BAE}(q)}(\mu)} = \max_{2\leq j\leq K} \frac{a_j}{q_j}, \tag{23}$$

with either side infinite if some $q_j = 0$. Therefore

$$\inf_{q \text{ feasible}} \sup_{\mu} \frac{\Gamma^\star_{\mathrm{so}}(\mu)}{\Gamma_{\mathrm{BAE}(q)}(\mu)} = U_K, \tag{24}$$

and $q^\star$ is the unique optimal cumulative schedule.

Proof sketch. Theorem 3 and Proposition 2 give

$$\Gamma_{\mathrm{BAE}(q)}(\mu) \geq \frac{\Gamma^\star_{\mathrm{so}}(\mu)}{\max_j a_j/q_j},$$

which proves the upper half of (23).

Fix j. Give the top $j-1$ challengers a common gap and send all remaining gaps to infinity. The oracle-to-elimination ratio converges to a_j. With probability tending to one on the exponential scale, the distant arms are eliminated first. The event that the best arm is then eliminated at the j-arm batch has exponent at most $q_j I_j$. This proves the reverse inequality in (23). The supplement gives the full limiting argument.

If the left side of (23) is at most R, then $q_j \geq a_j/R$ for all j. Feasibility yields

$$1 = 2q_2 + \sum_{j=3}^{K} q_j \geq \frac{2a_2 + \sum_{j=3}^K a_j}{R} = \frac{U_K}{R}.$$

Thus $R \geq U_K$. Equality is attained by (21) and Theorem 5. □

6 A Stronger Universal Lower Bound

We first strengthen the recent lower endpoint L_K. The construction adds a translation parameter to each equal-gap alternative. This forces every algorithm to balance cumulative and arm-specific information.

Theorem 7 (Translated hard family)**.** *For every $K \geq 3$, each consistent algorithm satisfies*

$$\sup_{\mu\in\mathcal{M}_K} \frac{\Gamma^\star_{\mathrm{so}}(\mu)}{\Gamma_{\mathsf{A}}(\mu)} \geq C_K > L_K. \tag{25}$$

Consequently, $P_K \geq C_K$.

Proof sketch. Fix large M and use the baseline $x_1 = 0$, $x_2 = -1$, and $x_i = -M^i$ for $i \geq 3$. For $j \in \{2,\dots,K\}$, put $d_j = M^{j+1/2}$. For each $\alpha \in [0,1]$, define an alternative with means

$$\lambda_i^{j,\alpha} = \begin{cases} -\alpha d_j, & i < j, \\ (1-\alpha)d_j, & i = j, \\ x_i, & i > j. \end{cases}$$

Arm j is uniquely best. Its first $j-1$ gaps equal d_j, while all later gaps are asymptotically larger. If $S_j(w) = \sum_{i=1}^j w_i$, then, uniformly over w and α,

$$\begin{aligned}\frac{1}{\Gamma^\star_{\mathrm{so}}(\lambda^{j,\alpha})} \sum_i w_i \frac{(x_i - \lambda_i^{j,\alpha})^2}{2} \\ = c_j\{\alpha^2 S_{j-1} + (1-\alpha)^2 w_j\} + o_M(1).\end{aligned} \tag{26}$$

Minimizing over α gives $c_j S_{j-1} w_j / S_j$, interpreted as zero when $S_j = 0$. The change-of-measure inequality of Degenne (2023), applied to all these alternatives, therefore gives

$$P_K \geq \left\{ \sup_{w\in\Delta_K} \min_{2\leq j\leq K} \frac{c_j S_{j-1}(w) w_j}{S_j(w)} \right\}^{-1}. \tag{27}$$

The program is exact. By homogeneity, its reciprocal is the least total mass of a nonnegative vector for which every displayed fraction is at least one. The $j = 2$ constraint requires cumulative mass at least one, with equality at $w_1 = w_2 = 1/2$. Given cumulative mass s through arm $j-1$, the least next mass is $s/(c_j s - 1)$. Iteration gives exactly (3). Since this increment is strictly larger than $1/c_j$, we obtain $C_K > L_K$. The supplement gives the uniform limits and allocation argument in full. □

7 The Universal Price

Combining Theorems 7 and 5 gives the main characterization.

Table 1: Static-oracle ratios. Gold. is the lower bound in Goldberger (2026), $[C_K, U_K]$ is our global interval, and SR is Successive Rejects.

K	Gold. L_K	C_K	U_K	SR ℓ_K
3	1.172	1.207	1.257	1.333
5	1.417	1.479	1.575	1.783
10	1.802	1.884	2.020	2.429
100	3.472	3.572	3.753	4.687

Theorem 8 (Price sandwich)**.** *For every $K \geq 3$,*

$$C_K \leq P_K \leq U_K, \qquad 0 < U_K - C_K < 0.2. \tag{28}$$

Both bounds are $\log K + O(1)$*. Hence*

$$\lim_{K\to\infty} \frac{P_K}{\log K} = 1. \tag{29}$$

Proof. For $j \geq 3$, $c_j > j$ and $C_{j-1} \geq 1$, so

$$\frac{1}{c_j} < \frac{C_{j-1}}{c_j C_{j-1} - 1} < \frac{j}{(j-1)c_j} = a_j.$$

Hence $L_K < C_K < U_K$. The increment of $U_K - C_K$ at stage j is less than $1/\{(j-1)c_j\}$. Direct outward-rounded recursion through $j = 20$ gives a gap below 0.159. The remaining decreasing series is bounded by its first term plus its integral and is below 0.040. Thus the total gap is below 0.199. Finally, $L_K < C_K < U_K$ and both outer sequences are $\log K + O(1)$. ☐

Table 1 shows that the interval is already narrow for small K. It also compares the exact worst-case ratio of classical Successive Rejects. The supplement proves that this ratio is

$$\ell_K = \frac{1}{2} + \sum_{i=2}^{K} \frac{1}{i}. \tag{30}$$

The optimized design is strictly better for all $K \geq 3$.

8 Related Work and Scope

Fixed-budget optimality. Classical fixed-budget methods include Successive Rejects (Audibert et al., 2010), Sequential Halving (Karnin et al., 2013), and confidence-bound designs. Gap-hardness lower bounds show a logarithmic minimax overhead (Carpentier and Locatelli, 2016), while Barrier et al. (2023) sharpen nonparametric guarantees. Komiyama et al. (2022) characterize a minimax exponent through a global optimization and propose a conceptual tracking algorithm. Their objective fixes a hardness class rather than comparing to the static oracle at every instance. For that criterion, Komiyama et al. (2026) give an anytime rate-optimal design and an exact Gaussian Successive Rejects exponent. Kato (2026) instead obtain exact minimax and Bayes constants for polynomial-scale simple regret under shrinking gaps.

Large deviations and admissibility. Wang et al. (2023) develop large-deviation tools for adaptive allocations, sharpen the Successive Rejects exponent, and introduce Continuous Rejects. Wang et al. (2024) prove uniform sampling admissible for two Bernoulli arms and derive the exact Successive Rejects exponent. For at least three Gaussian arms, Imbens et al. (2025) construct batched designs that strictly dominate uniform allocation. Their general exponent bound is an input to our analysis. We instead optimize the worst-case ratio to the instance-specific static oracle and prove exact minimaxity within their fixed-weight elimination class. Fei and Branke (2026) smooth the static maximin allocation with saddlepoint-weighted entropic regularization. Their adaptive procedure has no proved error exponent and does not address a single design's worst-case ratio to the instance-specific oracle.

Complexity and reductions. Degenne (2023) show that any fixed-budget complexity would have to equal the static-oracle difficulty, then rule this out in several cases. Goldberger (2026) prove nonexistence for all $K \geq 3$ and all regular one-parameter natural exponential families. Our translated family strengthens their Gaussian lower bound, while optimized elimination gives the nearly matching upper endpoint. Balagopalan et al. (2026) give a general reduction from fixed-confidence to fixed-budget identification up to logarithmic factors. That nonasymptotic reduction targets broad structured problems and does not optimize the sharp static-oracle exponent considered here.

Scope. Our sharp upper theorem assumes known common Gaussian variance. Goldberger's lower bound applies much more broadly by local quadratic approximation. The reciprocal-truncated-variance lemma is Euclidean, and extending the additive 0.2 characterization to general exponential families requires controlling nonquadratic information projections away from local neighborhoods. We also optimize fixed-weight one-at-a-time elimination, not Continuous Rejects or all adaptive algorithms. The global interval in Theorem 8 quantifies exactly what remains outside that class.

9 Conclusion

The impossibility of matching the fixed-budget static oracle does not leave an unquantified gap. For Gaussian best-arm identification, the unavoidable worst-case loss is $\log K$ with leading constant one, and its finite-K value lies in an explicit interval narrower than 0.2. An optimized batched elimination schedule attains the upper endpoint and is exactly minimax in its natural algorithm class. The remaining global gap is additive, not multiplicative.

A Static-Oracle Reductions

We first justify the scalar representation of the Gaussian static oracle. By translation and permutation, take arm 1 to be best, set $\mu_1 = 0$, and write $\mu_i = -d_i$ with $0 < d_2 \le \cdots \le d_K$. Let $A_i = d_i^2/2$.

Lemma 9 (Scalar oracle representation)**.** *For a problem with j arms,*

$$\Gamma^\star_{\mathrm{so},j} = \max_{0<x<A_2} \frac{x}{1 + \sum_{i=2}^{j} x/(A_i - x)}. \tag{31}$$

For the original K-arm instance, $\Gamma^\star_{\mathrm{so}}(\mu) \le \Gamma^\star_{\mathrm{so},j}$ when the latter contains the best arm and its closest $j-1$ challengers.

Proof. For proportions w, the pairwise exponent against arm i is

$$A_i \frac{w_1 w_i}{w_1 + w_i}.$$

At an interior optimum all pairwise exponents are equal. Otherwise, weight can be moved from a slack challenger to a binding challenger. Let the common value be γ and put $x = \gamma/w_1$. Solving

$$A_i \frac{w_1 w_i}{w_1 + w_i} = \gamma$$

gives

$$w_i = w_1 \frac{x}{A_i - x}, \qquad w_1 = \left(1 + \sum_{i=2}^{j} \frac{x}{A_i - x}\right)^{-1}.$$

Therefore γ equals the objective in (31). Conversely, every $x \in (0, A_2)$ defines valid positive weights through these formulas and equalizes all challengers. This proves the representation.

For the comparison, fix any K-arm proportions w and let $s = \sum_{i=1}^{j} w_i \le 1$. Normalize the first j weights as $v_i = w_i/s$. The minimum exponent over all challengers is at most the minimum over the first $j-1$ challengers. By homogeneity, that restricted minimum is s times its value under v, hence is at most $s\Gamma^\star_{\mathrm{so},j} \le \Gamma^\star_{\mathrm{so},j}$. Maximizing over w proves the claim. □

We next justify the information projection used for elimination.

Lemma 10 (Best-is-worst projection)**.** *For a set containing the best arm and challengers with gaps $d_1, \ldots, d_n > 0$, the information needed to make the best mean no larger than every challenger is*

$$I(d_1, \ldots, d_n) = \min_{r \ge 0} \frac{1}{2} \left\{ r^2 + \sum_{i=1}^{n} (d_i - r)_+^2 \right\}. \tag{32}$$

Among all sets of n challengers, the closest n minimize this value.

Proof. Again put the best mean at zero and challenger i at $-d_i$. At a feasible information projection, write the new best mean as $-r$. An optimizer has $r \ge 0$. For a fixed r, challenger i is moved only if its original mean is below $-r$, and its closest feasible value is $-r$. Its squared movement is therefore $(d_i - r)_+^2$. This proves (32).

The objective is nondecreasing in each gap before minimization over r. Hence replacing a selected challenger by a closer one cannot increase the information. The closest n challengers minimize it. □

B Reciprocal-Truncated Variance

We prove the analytic lemma that supplies the sharp constant.

Lemma 11 (Reciprocal-truncated variance)**.** *For $n \ge 1$ and $z_1, \ldots, z_n > 1$,*

$$\left(1 + \sum_{i=1}^{n} \frac{1}{z_i^2 - 1}\right) \min_{t \ge 0} \left\{ t^2 + \sum_{i=1}^{n} (z_i - t)_+^2 \right\} \ge F_n, \tag{33}$$

where

$$F_n = \frac{n(1 + \sqrt{n})^2}{n + 1}.$$

Equality holds if and only if $z_i^2 = 1 + \sqrt{n}$ for all i.

Proof. Let t minimize the truncated quadratic and let $I = \{i : z_i > t\}$, $m = |I|$, and $l = n - m$. At least one coordinate is active. The first-order condition gives

$$t = \frac{\sum_{i \in I} z_i}{m + 1}. \tag{34}$$

If $l > 0$, every inactive coordinate is larger than one and no larger than t, so $t > 1$.

Write $\bar{z} = m^{-1} \sum_{i \in I} z_i$ and $y = \bar{z}^2$. The map $z \mapsto (z^2 - 1)^{-1}$ is strictly convex on $(1, \infty)$. Jensen's inequality, the inactive bounds $z_i \le t$, and (34) give

$$1 + \sum_{i=1}^{n} \frac{1}{z_i^2 - 1} \ge 1 + \frac{m}{y - 1} + \frac{l}{m^2 y/(m+1)^2 - 1}, \tag{35}$$

$$t^2 + \sum_{i=1}^{n} (z_i - t)_+^2 = \sum_{i \in I} z_i^2 - \frac{(\sum_{i \in I} z_i)^2}{m + 1} \ge \frac{my}{m + 1}. \tag{36}$$

It remains to prove

$$P_{m,n}(y) := \frac{my}{m+1} \left(1 + \frac{m}{y - 1} + \frac{n - m}{m^2 y/(m+1)^2 - 1}\right) \ge F_n. \tag{37}$$

The last term is omitted if $m = n$. When $m = n$, put $u = y - 1$. Direct algebra gives

$$P_{n,n}(y) = F_n + \frac{n}{n+1}\frac{(u-\sqrt{n})^2}{u}. \tag{38}$$

Thus the desired bound holds, with equality only at $y = 1 + \sqrt{n}$.

Suppose now that $m < n$, put $u = y-1$ and $l = n-m$. The inactive-coordinate contribution in (37) is

$$\begin{aligned} D_m(u) &= \frac{m(m+1)(u+1)}{m^2u-(2m+1)} \\ &= \frac{m+1}{m} + \frac{(m+1)^3}{m\{m^2u-(2m+1)\}}. \end{aligned} \tag{39}$$

Its denominator is positive on the inactive domain. Since it is smaller than m^2u, equations (38) and (39), followed by AM-GM, yield

$$P_{m,n}(y) > F_m + l\frac{m+1}{m} + 2\sqrt{Q} - \frac{2m^{3/2}}{m+1}, \tag{40}$$

where

$$Q = \frac{m^3}{(m+1)^2} + l\frac{(m+1)^2}{m^2}.$$

Let $g(x) = x^3/(x+1)^2$. Its derivative is

$$g'(x) = \frac{x^2(x+3)}{(x+1)^3} < 1 \qquad (x > 0).$$

Consequently,

$$Q > g(m) + l > g(m+l) = g(n).$$

Since $F_s = s + 2s^{3/2}/(s+1)$, substitution into (40) gives

$$P_{m,n}(y) > F_n + \frac{l}{m} > F_n.$$

An inactive coordinate therefore makes the inequality strict. In the all-active case, equality in Jensen and (38) requires every z_i to be equal with $z_i^2 = 1 + \sqrt{n}$. The converse follows by direct substitution. □

Proof of the oracle-to-elimination inequality. Use Lemmas 9 and 10. Fix x in (31) and set $z_i = \sqrt{A_i/x} > 1$. Scaling (32) by x gives

$$I_j = x \min_{t\geq 0}\left\{t^2 + \sum_{i=2}^{j}(z_i - t)_+^2\right\}.$$

The objective in (31) is

$$\frac{x}{1+\sum_{i=2}^{j}(z_i^2-1)^{-1}}.$$

Lemma 11, with $n = j-1$, bounds their ratio by

$$\frac{j}{(j-1)(1+\sqrt{j-1})^2} = a_j.$$

Maximize over x.

For sharpness, give all $j-1$ challengers one gap d. The information projection pools all j means and gives

$$I_j = \frac{j-1}{2j}d^2. \tag{41}$$

The static oracle puts weight $1/(1+\sqrt{j-1})$ on the best arm and equal remaining weight on the challengers, giving

$$\Gamma^\star_{\mathrm{so},j} = \frac{d^2}{2(1+\sqrt{j-1})^2}. \tag{42}$$

Their ratio is a_j. In a K-arm problem, send all other gaps to infinity. The static-oracle exponent converges to (42), while I_j is unchanged. □

C Batched-Elimination Exponent

We give a direct proof specialized to Gaussian rewards. Let N_j be the cumulative number of observations per surviving arm at the end of the batch with j arms.

Proposition 12. *For every feasible schedule and every budget T,*

$$p_{\mu,T}(\mathrm{BAE}(q)) \leq \sum_{j=2}^{K}\binom{K-1}{j-1}\exp\{-N_jI_j(\mu)\}. \tag{43}$$

If $N_j = 0$, its summand is interpreted as one. In particular, when $N_j = q_jT + O(1)$,

$$\Gamma_{\mathrm{BAE}(q)}(\mu) \geq \min_{2\leq j\leq K} q_jI_j(\mu). \tag{44}$$

Proof. If any N_j is zero, the corresponding summand makes the bound trivial. Suppose therefore that all cumulative counts are positive. An error occurs only if the best arm is discarded at some stage. Fix a stage with j active arms and a candidate set J of size j containing the best arm. If this is the candidate set and the best arm is discarded, then its empirical mean after N_j observations is no larger than every empirical mean in $J \setminus \{b\}$ after the same number of observations.

Let

$$C_J = \{x \in \mathbb{R}^J : x_b \leq \min_{i\in J\setminus\{b\}} x_i\}$$

and let $\lambda^\star$ be the Euclidean projection of μ_J onto C_J. The projection inequality gives

$$\langle \mu_J - \lambda^\star, x - \lambda^\star\rangle \leq 0 \qquad (x \in C_J).$$

Thus, if the vector $\widehat{\mu}_J$ of sample means belongs to C_J, then

$$\langle \lambda^\star - \mu_J, \widehat{\mu}_J - \mu_J\rangle \geq \|\lambda^\star - \mu_J\|_2^2.$$

The left side is centered Gaussian with variance $\|\lambda^\star - \mu_J\|_2^2/N_j$. The standard Gaussian tail bound therefore gives

$$\mathbb{P}_\mu(\widehat{\mu}_J \in C_J) \le \exp\left\{-\frac{N_j}{2}\|\lambda^\star - \mu_J\|_2^2\right\} = \exp\{-N_j I_J(\mu)\}.$$

There are $\binom{K-1}{j-1}$ sets of size j containing b. Union-bounding over these sets and all stages proves (43). Substituting $N_j = q_j T + O(1)$ and taking lower error exponents proves (44). □

The same result follows from the general large-deviation analysis of Wang et al. (2023). Imbens et al. (2025) state it for the expected simple-regret exponent. On a fixed unique-best instance,

$$\Delta_{\min}\mathbb{P}(\widehat{b}_T \ne b) \le \mathbb{E}[\mu_b - \mu_{\widehat{b}_T}] \le \Delta_{\max}\mathbb{P}(\widehat{b}_T \ne b),$$

so the two exponents coincide.

D Exact Optimality in the Elimination Class

We supply the lower half of the batched-elimination minimax theorem. Fix a stage j and a gap $d > 0$. Let arm 1 have mean zero, arms $2, \dots, j$ have mean $-d$, and arms $j+1, \dots, K$ have mean $-M$.

Suppose first that every q_s is positive. Let $G_{M,T}$ be the event that no top-j arm is discarded at a stage with more than j active arms. If $G_{M,T}$ fails at stage $s > j$, some top arm has an empirical mean no larger than that of a surviving distant arm. Their mean separation is at least $M - d$. A Gaussian tail bound and a union bound over arms and stages give

$$\liminf_{T\to\infty} \frac{1}{T} \log \frac{1}{\mathbb{P}(G_{M,T}^{\mathsf{c}})} \ge \min_{s>j} \frac{q_s(M-d)^2}{4}. \tag{45}$$

The right side diverges as $M \to \infty$. When $j = K$, define $G_{M,T}$ to be the sure event and omit (45).

Let $E_{j,T}$ be the event that arm 1 has strictly smallest empirical mean among $[j]$ at the j-arm stage. Every count is $N_j + O(1)$, where the $O(1)$ term only accounts for unused rounded observations. The Gaussian large-deviation upper and lower bounds for the open cone give

$$\lim_{T\to\infty} \frac{1}{T} \log \frac{1}{\mathbb{P}(E_{j,T})} = q_j I_{[j]}(\mu^{(M,j)}). \tag{46}$$

Indeed, the closure of this cone has Euclidean projection cost $I_{[j]}$, and an arbitrarily small displacement of its projection lies in the open cone and has cost arbitrarily close to $I_{[j]}$.

On $E_{j,T} \cap G_{M,T}$, the algorithm reaches candidate set $[j]$ and then discards the best arm. Therefore

$$p_{\mu^{(M,j)},T}(\mathrm{BAE}(q)) \ge \mathbb{P}(E_{j,T}) - \mathbb{P}(G_{M,T}^{\mathsf{c}}).$$

Choose M so that the exponent in (45) is strictly larger than the exponent in (46). The elementary subtraction rule for two exponentially decaying sequences then shows

$$\Gamma_{\mathrm{BAE}(q)}(\mu^{(M,j)}) \le q_j I_{[j]}(\mu^{(M,j)}). \tag{47}$$

The distant-arm gaps diverge with M, so the scalar oracle formula and the projection formula converge to their j-arm equal-gap values. Equations (41) and (42) then give

$$\liminf_{M\to\infty} \frac{\Gamma^\star_{\mathrm{so}}(\mu^{(M,j)})}{\Gamma_{\mathrm{BAE}(q)}(\mu^{(M,j)})} \ge \frac{a_j}{q_j}.$$

If some $q_s = 0$, consistency already fails on an instance where a discarded arm at that stage is best, and the worst-case ratio is infinite.

Thus every feasible schedule has worst-case ratio at least $\max_j a_j/q_j$. If this maximum is R, then $q_j \ge a_j/R$. The budget identity gives

$$1 = 2q_2 + \sum_{j=3}^K q_j \ge \frac{2a_2 + \sum_{j=3}^K a_j}{R} = \frac{U_K}{R}.$$

Hence $R \ge U_K$. The schedule $q_j = a_j/U_K$ attains equality by the upper theorem. Equality in the budget display forces every coordinate to equal a_j/U_K, which also proves uniqueness.

For completeness, the matching upper bound holds schedule by schedule. The oracle-to-elimination inequality and the BAE exponent bound give

$$\Gamma_{\mathrm{BAE}(q)}(\mu) \ge \min_j q_j I_j(\mu) \ge \frac{\Gamma^\star_{\mathrm{so}}(\mu)}{\max_j a_j/q_j}.$$

Thus the exact worst-case ratio of every positive feasible schedule is $\max_j a_j/q_j$. If some $q_j = 0$, both this expression and the worst-case ratio are infinite.

E Translated Equal-Gap Lower Bound

We prove the new all-algorithm lower endpoint. Fix K and let M tend to infinity. Define a unique-best baseline by

$$x_1 = 0, \qquad x_2 = -1, \qquad x_i = -M^i \quad (3 \le i \le K).$$

For $j \in \{2, \dots, K\}$, put $d_j = M^{j+1/2}$. For every $\alpha \in [0,1]$, define

$$\lambda_i^{j,\alpha} = \begin{cases} -\alpha d_j, & i < j, \\ (1-\alpha)d_j, & i = j, \\ x_i, & i > j. \end{cases} \tag{48}$$

Arm j is uniquely best under $\lambda^{j,\alpha}$, while arm 1 is best under x. Thus every member of this alternative family changes the best arm.

Lemma 13 (Uniform translated cost)**.** *Let $c_j = (1+\sqrt{j-1})^2$ and $S_j(w) = \sum_{i=1}^{j} w_i$. Uniformly over $w \in \Delta_K$ and $\alpha \in [0,1]$,*

$$\frac{1}{\Gamma^\star_{\mathrm{so}}(\lambda^{j,\alpha})} \sum_{i=1}^{K} w_i \frac{(x_i - \lambda_i^{j,\alpha})^2}{2} = c_j \left\{\alpha^2 S_{j-1}(w) + (1-\alpha)^2 w_j\right\} + o_M(1). \tag{49}$$

Consequently, the infimum over α converges uniformly to

$$\frac{c_j S_{j-1}(w) w_j}{S_j(w)}. \tag{50}$$

The ratio is interpreted as zero when $S_j(w) = 0$.

Proof. The first $j-1$ gaps of $\lambda^{j,\alpha}$ are exactly d_j. Every later gap divided by d_j diverges uniformly in α. The scalar oracle representation in Lemma 9, after dividing all squared gaps by d_j^2, therefore gives

$$\Gamma^\star_{\mathrm{so}}(\lambda^{j,\alpha}) = \frac{d_j^2}{2c_j}\{1 + o_M(1)\}$$

uniformly in α. This is also immediate from the equal-gap oracle in (42), since the normalized later gaps diverge.

For $i < j$, $x_i/d_j = o_M(1)$ uniformly over the finite index set, and

$$\frac{x_i - \lambda_i^{j,\alpha}}{d_j} = \alpha + o_M(1).$$

Likewise,

$$\frac{x_j - \lambda_j^{j,\alpha}}{d_j} = -(1-\alpha) + o_M(1).$$

The remaining coordinates do not move. Squaring and summing proves (49), with errors uniform over the two compact simplices. Finally,

$$\min_{0 \le \alpha \le 1} \{\alpha^2 S_{j-1} + (1-\alpha)^2 w_j\} = \frac{S_{j-1} w_j}{S_j},$$

including the zero-mass cases by continuity. $\square$

We now invoke the standard fixed-budget change-of-measure inequality in the difficulty-ratio form of Degenne (2023). For a consistent algorithm A, a baseline x, and any family D of unique-best alternatives to x, it states

$$\left\{\sup_{\lambda \in D} \frac{\Gamma^\star_{\mathrm{so}}(\lambda)}{\Gamma_{\mathsf{A}}(\lambda)}\right\}^{-1} \le \sup_{w \in \Delta_K} \inf_{\lambda \in D} \frac{1}{\Gamma^\star_{\mathrm{so}}(\lambda)} \sum_{i=1}^{K} w_i \frac{(x_i - \lambda_i)^2}{2}. \tag{51}$$

Indeed, Degenne's ratio is

$$\limsup_T \frac{T}{\log(1/p_{\lambda,T})} \Big/ \frac{1}{\Gamma^\star_{\mathrm{so}}(\lambda)} = \frac{\Gamma^\star_{\mathrm{so}}(\lambda)}{\Gamma_{\mathsf{A}}(\lambda)},$$

so (51) uses exactly the convention of the main paper. Apply this with all alternatives in (48). Uniformity in Lemma 13 and finiteness of the index set give

$$P_K \ge \rho_K^{-1}, \qquad \rho_K = \sup_{w \in \Delta_K} \min_{2 \le j \le K} \frac{c_j S_{j-1}(w) w_j}{S_j(w)}. \tag{52}$$

Lemma 14 (Exact allocation program)**.** *The value in (52) is $\rho_K = 1/C_K$, where*

$$C_2 = 1, \qquad C_j = C_{j-1} + \frac{C_{j-1}}{c_j C_{j-1} - 1}. \tag{53}$$

Moreover, $L_K < C_K < U_K$ for every $K \ge 3$.

Proof. For a nonnegative vector z, define

$$\phi_j(z) = \frac{c_j S_{j-1}(z) z_j}{S_j(z)}.$$

These functions are homogeneous of degree one. Hence ρ_K^{-1} is the least total mass $S_K(z)$ subject to $\phi_j(z) \ge 1$ for every j.

Since $c_2 = 4$,

$$\phi_2(z) = \frac{4 z_1 z_2}{z_1 + z_2} \le z_1 + z_2.$$

Thus $S_2 \ge 1$, with equality attainable at $z_1 = z_2 = 1/2$. Given $S_{j-1} = s \ge 1$, the next constraint is equivalent to

$$z_j \ge \frac{s}{c_j s - 1}.$$

The map $s \mapsto s + s/(c_j s - 1)$ has derivative $1 - (c_j s - 1)^{-2} > 0$ on this domain. The least final mass therefore starts at $S_2 = 1$ and takes equality at every later constraint, which is precisely (53).

For $j \ge 3$, $c_j > j$ and $C_{j-1} \ge 1$. Therefore

$$\frac{1}{c_j} < \frac{C_{j-1}}{c_j C_{j-1} - 1} < \frac{j}{(j-1)c_j} = a_j. \tag{54}$$

Summing these strict inequalities from the common initial value one proves $L_K < C_K < U_K$. $\square$

F Sequence Calculations

F.1 Feasibility of the optimized schedule

Put $m = j - 1$. Then

$$a_j = \frac{m+1}{m(1+\sqrt{m})^2}.$$

This is strictly decreasing in $m \geq 1$, as is seen by differentiating its continuous extension. Also $a_2 = 1/2$. Therefore $q_j = a_j/U_K$ is decreasing and

$$2q_2 + \sum_{j=3}^{K} q_j = \frac{2a_2 + \sum_{j=3}^{K} a_j}{U_K} = 1.$$

It defines nonnegative batch fractions through $\beta_K = Kq_K$ and $\beta_j = j(q_j - q_{j+1})$.

F.2 Width below 0.2

By (54), the stage-j increment of $U_K - C_K$ is positive and smaller than

$$f(j-1), \qquad f(x) = \frac{1}{x(1+\sqrt{x})^2}.$$

Direct evaluation of the recursion with outward rounding gives

$$C_{20} > 2.3388734, \qquad U_{20} < 2.4973514,$$

so $U_{20} - C_{20} < 0.159$. The function f is positive and decreasing, and

$$\begin{aligned}\sum_{m=20}^{\infty} f(m) &\leq f(20) + \int_{20}^{\infty} f(x)\,dx \\ &= f(20) + 2\log\frac{1+\sqrt{20}}{\sqrt{20}} - \frac{2}{1+\sqrt{20}} \\ &< 0.040.\end{aligned}$$

Thus $U_K - C_K < 0.199 < 0.2$ for every K.

Since
$$\frac{1}{(1+\sqrt{m})^2} = \frac{1}{m} + O(m^{-3/2}),$$
both L_K and U_K equal $\log K + O(1)$. Since $L_K < C_K < U_K$, the same holds for C_K.

G Classical Successive Rejects

Classical Successive Rejects has cumulative fractions

$$q_j^{\mathrm{SR}} = \frac{1}{j\ell_K}, \qquad \ell_K = \frac{1}{2} + \sum_{i=2}^{K} \frac{1}{i}.$$

The oracle-to-elimination inequality gives worst-case ratio at most

$$\max_j \frac{a_j}{q_j^{\mathrm{SR}}} = \ell_K \max_j j a_j.$$

With $m = j-1$,

$$ja_j = \frac{(m+1)^2}{m(1+\sqrt{m})^2} \leq 1,$$

because $m + 1 \leq m + \sqrt{m}$. Equality holds at $m = 1$, or $j = 2$. An instance with two close arms and all remaining gaps diverging attains this ratio by the lower construction in the previous section. Thus the exact worst-case ratio is ℓ_K.

For $K \geq 3$, the optimized value U_K is strictly below ℓ_K. Indeed, $2a_2 = 1$ and $a_j < 1/j$ for every $j \geq 3$, so

$$U_K = 1 + \sum_{j=3}^{K} a_j < 1 + \sum_{j=3}^{K} \frac{1}{j} = \ell_K.$$

Numerically, $\ell_K - U_K$ converges to approximately 1.28.

References

Audibert, J.-Y., Bubeck, S., and Munos, R. (2010). Best arm identification in multi-armed bandits. In *Proceedings of the 23rd Conference on Learning Theory*, pages 41–53.

Balagopalan, K., Li, Y., Zhao, Y., Nguyen, T., Daitche, A., Nassif, H., and Jun, K.-S. (2026). Fixed budget is no harder than fixed confidence in best-arm identification up to logarithmic factors. In *Proceedings of the 43rd International Conference on Machine Learning*, Proceedings of Machine Learning Research.

Barrier, A., Garivier, A., and Stoltz, G. (2023). 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*, pages 136–181.

Carpentier, A. and Locatelli, A. (2016). Tight (lower) bounds for the fixed budget best arm identification bandit problem. In *Proceedings of the 29th Conference on Learning Theory*, volume 49 of *Proceedings of Machine Learning Research*, pages 590–604.

Degenne, R. (2023). 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*, pages 1131–1154.

Fei, X. and Branke, J. (2026). Annealed entropic allocation for ranking and selection. *arXiv preprint arXiv:2606.11347*.

Glynn, P. W. and Juneja, S. (2004). A large deviations perspective on ordinal optimization. In *Proceedings of the 2004 Winter Simulation Conference*, pages 577–585. IEEE.

Goldberger, M. (2026). Fundamental limitations of fixed-budget best-arm identification. *arXiv preprint arXiv:2607.11635*.

Imbens, G., Qin, C., and Wager, S. (2025). Admissibility of completely randomized trials: A large-deviation approach. *arXiv preprint arXiv:2506.05329*.

Karnin, Z., Koren, T., and Somekh, O. (2013). 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*, pages 1238–1246.

Kato, M. (2026). Minimax and Bayes optimal best-arm identification. *arXiv preprint arXiv:2506.24007*.

Komiyama, J., Jang, K., and Honda, J. (2026). Rate-optimal design for anytime best arm identification. *arXiv preprint arXiv:2510.23199*.

Komiyama, J., Tsuchiya, T., and Honda, J. (2022). Minimax optimal algorithms for fixed-budget best arm identification. In *Advances in Neural Information Processing Systems*, volume 35, pages 10393–10404.

Qin, C. (2022). Open problem: Optimal best arm identification with fixed-budget. In *Proceedings of the 35th Conference on Learning Theory*, volume 178 of *Proceedings of Machine Learning Research*, pages 5650–5654.

Wang, P.-A., Ariu, K., and Proutiere, A. (2024). 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*, pages 50065–50091.

Wang, P.-A., Tzeng, R.-C., and Proutiere, A. (2023). Best arm identification with fixed budget: A large deviation perspective. In *Advances in Neural Information Processing Systems*, volume 36, pages 16804–16815.