THE STATISTICAL PRICE OF UNKNOWN TRANSIENT AMPLITUDE IN BEST-ASYMPTOTE IDENTIFICATION

Anonymous authors
Paper under double-blind review

ABSTRACT

When several learning or optimization runs share a budget of updates, the run with the smallest current loss need not be the run with the smallest limiting loss because the runs are still improving. We study how to identify the run with the best limit (best-asymptote identification) when the sth update of run i produces $Y_{i,s} = \mu_i + \beta_i s^{-r} + \varepsilon_{i,s}$ with Gaussian noise: each pull advances only the arm it is spent on, the decay exponent r and the noise variance σ^2 are known, and the limits μ_i and the transient amplitudes β_i are not. The question is how much not knowing the amplitudes costs compared with an oracle that reveals them. We give three exact answers. For a single arm, the sharp asymptotic variance for estimating μ_i is $\sigma^2\kappa(r)/n$ instead of the oracle's σ^2/n, where $\kappa(r) = (1-r)^2/r^2$ for $0 < r < 1/2$ and $\kappa(r) = 1$ for $r \geq 1/2$; unknown amplitude, therefore, has a first-order cost only when the transient decays slowly. For two arms whose limits differ by $c/\sqrt{T}$, with amplitudes in any fixed compact subinterval of $(0, \infty)$ and a budget cap $\tau_T \leq CT$ almost surely, the exact local asymptotic minimax misidentification error over all parameter-free policies with predictable, history-dependent allocation and adapted stopping is $\Phi(-c\sqrt{C}/[2\sigma\sqrt{\kappa(r)}])$; it is attained by deterministic balanced sampling with full-prefix least squares, and the oracle value is the same formula with $\kappa(r) = 1$. Hence, no adaptive allocation or stopping rule can remove the nuisance penalty: matching the oracle requires $\kappa(r)$ times as much first-order budget. For K arms, when at most one challenger may approach the best arm at the $T^{-1/2}$ scale and every other arm keeps a fixed margin, a sublinear screening pilot followed by an independent-suffix comparison attains the two-arm value, and the optimized deterministic all-arm-prefix budget is $K/2$ times larger. Exact finite-prefix risk calculations and Gaussian simulations illustrate the boundary at $r = 1/2$ and the screening mechanism.

1 INTRODUCTION

Consider several learning or optimization runs that share a finite update budget. Each run improves along its own trajectory, and updating one run does not advance the others. The run with the smallest loss today therefore need not have the smallest loss after convergence. We model the sth update of arm i as

$$Y_{i,s} = \mu_i + \beta_i s^{-r} + \varepsilon_{i,s}, \qquad \varepsilon_{i,s} \overset{\text{iid}}{\sim} N(0, \sigma^2), \tag{1}$$

where μ_i is the limiting loss, $\beta_i \geq 0$ is the transient amplitude, $r > 0$ is the decay exponent, and s is the arm's local pull count. The exponent r and the variance σ^2 are known, the amplitudes are not, and the goal is to identify an arm in $\arg\min_i \mu_i$.

What is the statistical price of not knowing the transient amplitudes? As a benchmark, suppose an oracle reveals every β_i. Subtracting $\beta_i s^{-r}$ then leaves independent observations of μ_i with variance σ^2. Without that side information, the limiting level and the transient amplitude must be learned from the same prefix of the curve. We show that the resulting sharp variance ratio is

$$\kappa(r) := \begin{cases} (1-r)^2/r^2, & 0 < r < 1/2, \\ 1, & r \geq 1/2. \end{cases} \tag{2}$$

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.

Below $r = 1/2$, the empirical correlation between the constant regressor and the decaying regressor s^{-r} has a nonzero limit, so estimating the amplitude removes first-order information about the intercept. At and above the boundary the correlation vanishes, and intercept estimation with unknown amplitude becomes first-order as efficient as the oracle.

This fixed-arm geometry does not by itself settle the bandit problem: a policy might concentrate updates on a promising arm, or stop early once the observed curves are easy to separate. We prove that neither option improves the local two-arm minimax constant. Under an almost-sure budget cap, the same factor $\kappa(r)$ is the exact asymptotic unknown-to-oracle budget ratio, and the lower bound allows endogenous, data-dependent allocation and stopping without assuming a deterministic limiting allocation.

This objective differs from the finite-horizon curve values and the regret studied by existing curve-bandit methods (Cella et al., 2021; Takemori et al., 2024; Mussi et al., 2024). For $r < 1/2$, the coefficient in (2) is already implicit in the inverse-Gram analysis of Takemori et al. (2024). Relative to that calculation, our results add the revealed-amplitude efficiency interpretation, the saturation of the ratio at one for $r \geq 1/2$, and the exact transfer of the constant to adaptive best-asymptote identification.

Contributions. The paper makes three contributions.

1. **A sharp fixed-arm boundary.** We compute the intercept Schur information and prove that $\sigma^2\kappa(r)/n$ is the sharp asymptotic and local minimax mean-squared-error constant, against the oracle constant σ^2/n (Theorem 1).
2. **An exact adaptive two-arm value.** For gaps of at least $c/\sqrt{T}$, an almost-sure cap CT, and amplitudes in any fixed compact subinterval of $(0, \infty)$, the local minimax error over parameter-free policies with predictable allocation and adapted stopping is $\Phi(-c\sqrt{C}/[2\sigma\sqrt{\kappa(r)}])$. Balanced full-prefix least squares attains it, and the oracle value is the same formula with $\kappa(r) = 1$ (Theorem 2).
3. **A fixed-K one-challenger extension.** When only one challenger may approach the best arm at the root-T scale and all remaining arms have a fixed margin, a sublinear screen followed by an independent-suffix comparison attains the two-arm value. Over this class, the optimized deterministic all-arm-prefix budget constant is $K/2$ times the adaptive screening constant (Theorem 3).

Simultaneous K-way local ties remain open. Experiments verify the exact finite-prefix risk law and show the screen–then–compare error approaching its asymptotic prediction; they do not claim practical superiority.

Notation. Φ is the standard normal distribution function, $z_\alpha := \Phi^{-1}(1-\alpha)$, e_k is the kth standard basis vector, and I_d is the $d \times d$ identity. Throughout, s indexes the local pulls of one arm, n is a prefix length, t is a global round, T is the budget scale, and $N_i(t)$ is the number of pulls of arm i after t global rounds. A hat marks an estimate of μ with unknown amplitude and a tilde its revealed-amplitude counterpart; the labels un and or refer to the unknown-amplitude problem and to the oracle, and $\bullet \in \{\mathrm{un}, \mathrm{or}\}$ stands for either.

2 Related works

Cella et al. (2021) use the same inverse-polynomial model, after an exchange of their transient-amplitude and asymptote notation. Their Theorem 1 gives a two-arm lower bound and Theorem 3 an explore–then–commit (ETC) upper bound; Theorem 2 builds confidence intervals for future losses from local pull counts, and Theorem 4 gives a K-arm REST-SURE upper bound. Their work already covers the model, joint coefficient estimation, future-loss confidence intervals, and Δ^{-2}-scale exploration. The stateful local-pull setting and the best-function-identification framing follow the arm-as-function framework of Dorn et al. (2026), which studies the regret of F-LCB rather than asymptote identification; none of its results are used here.

The inverse-power basis of Takemori et al. (2024) contains our two-regressor curve as the case $(\rho_1, \rho_2) = (0, r)$. Their Propositions 4.2–4.3 develop the corresponding Gram geometry, Theorems 4.4 and 5.6 give K-arm sequential-halving upper bounds, and Theorem 6.3 gives a lower bound

for an allocation-regular adaptive class. For $r < 1/2$, their asymptotic inverse-Gram coefficient for the intercept already yields $(1-r)^2/r^2$. They neither interpret this coefficient relative to a revealed-amplitude oracle nor establish its saturation at one for $r \geq 1/2$.

Mussi et al. (2024) analyze adaptive K-arm R-UCBE and R-SR, with lower bounds that apply to all algorithms, for nonparametric concave rising curves (their Theorems 4.1–6.2). Their bias–noise transition occurs at the corresponding exponent $r = 1/2$. All three papers rank a finite-horizon curve value; we rank its limiting intercept. Table 1 summarizes the comparison.

Statistical positioning. The fixed-arm result is a Gaussian LAN and local minimax calculation (Le Cam, 1960; Hájek, 1972; Le Cam, 1986); its Schur complement projects the intercept score off the nuisance-amplitude score (Cox & Reid, 1987). Under predictable allocation, each arm's classical regression-design information (Kiefer & Wolfowitz, 1959) becomes random and data-dependent. Recent work characterizes attainable batched adaptive limits (Hirano & Porter, 2025); continuously updated contextual bandits can have random-information LAQ/SDE limits (van den Akker et al., 2025a), while standard LAN can reappear at a fixed unique best arm (van den Akker et al., 2025b). Appendix J derives the dominated limit experiment for the joint score and information in our Gaussian prefix model, whereas Theorem 2 optimizes over all predictable policies under an almost-sure budget cap without imposing a limiting allocation.

The closest stationary-design comparison is generalized Neyman allocation. For known-variance Gaussian arms, Kato (2024b, Theorem 4.1 and Theorems 4.2–4.4) derive a fixed-budget exponent bound and a matching static GNA strategy, and Kato (2025, Theorem 2.3 and Theorems 4.2–4.3) give adaptive K-arm GNA–A2IPW and matching locally minimax constants for general mean-parameterized distributions. These criteria send $T \to \infty$ before shrinking the gap neighborhood, whereas we set $\Delta_T = c/\sqrt{T}$ from the outset. Our predictable-allocation result under a deterministic interior limiting allocation is Theorem 15; the fixed-K theorem instead optimizes over all predictable policies but restricts the parameter class, so that at most one challenger lies at root-T distance while every other arm keeps a fixed margin. It therefore neither extends the Neyman-allocation principle nor characterizes simultaneous K-way near ties; it isolates the exact variance inflation caused by estimating a transient mean amplitude from a stateful prefix. Unlike Kato (2024a, Proposition 2.1 and Theorem 4.1), who estimate unknown Gaussian reward variances in a stationary problem, we assume a known innovation variance σ^2 and estimate an unknown amplitude in the mean.

3 The Single-Arm Price of Unknown Amplitude

Fix one arm and drop its index. With $z_s = (1, s^{-r})^\top$ and $\theta = (\mu, \beta)^\top$, the observations $Y_s = z_s^\top \theta + \varepsilon_s$ form a two-regressor linear model with Gram matrix

$$Q_n := \sum_{s=1}^{n} z_s z_s^\top = \begin{pmatrix} n & a_n \\ a_n & b_n \end{pmatrix}, \quad a_n := \sum_{s=1}^{n} s^{-r}, \quad b_n := \sum_{s=1}^{n} s^{-2r}. \tag{3}$$

For a fixed ridge parameter $\lambda > 0$, let $Q_{n,\lambda} = Q_n + \lambda I_2$ and $q_{n,\lambda} = e_1^\top Q_{n,\lambda}^{-1} e_1$. The ridge intercept is

$$\widehat{\mu}_{n,\lambda} := e_1^\top Q_{n,\lambda}^{-1} \sum_{s=1}^{n} z_s Y_s. \tag{4}$$

Table 1: Theorem-level comparison with the closest curve-bandit work. The targets differ, so the guarantees are not normalized to a common loss.

	Cella et al.	Takemori et al.	Mussi et al.	This paper
Target	terminal loss or regret	post-exploration loss	reward at T	limiting intercept
Allocation	ETC; REST-SURE	K-arm elimination	K-arm rejection	capped two-arm; screened K
Guarantee	lower and upper bounds	error bounds	error and regret bounds	exact asymptotic local minimax risk
Amplitude	jointly estimated; no oracle ratio	factor implicit for $r < 1/2$	no separate amplitude parameter	oracle ratio; saturation

The unregularized Schur information of the intercept and its inverse are

$$D_n := n - \frac{a_n^2}{b_n}, \qquad \bar{q}_n := D_n^{-1} = e_1^\top Q_n^{-1} e_1. \tag{5}$$

Power-sum asymptotics give

$$\frac{D_n}{n} \longrightarrow \frac{1}{\kappa(r)}, \tag{6}$$

and the corresponding finite-prefix law is already exact: for $n \geq 2$, the unregularized intercept $\widehat{\mu}_{n,0} = e_1^\top Q_n^{-1} \sum_{s=1}^n z_s Y_s$ satisfies

$$\widehat{\mu}_{n,0} - \mu \sim N\left(0, \frac{\sigma^2}{D_n}\right), \qquad \frac{n}{\sigma^2}\mathbb{E}(\widehat{\mu}_{n,0} - \mu)^2 = \frac{n}{D_n} \longrightarrow \kappa(r). \tag{7}$$

Below $r = 1/2$ the nuisance score retains first-order alignment with the intercept score; at and above $r = 1/2$ the alignment vanishes. For the minimax statements we work on the parameter space

$$\Theta_S := \{(\mu, \beta) : \mu^2 + \beta^2 \leq S^2,\ \beta \geq 0\} \tag{8}$$

and write $B(\theta_0, \rho)$ for the closed Euclidean ball of radius ρ around θ_0.

Theorem 1 (Sharp statistical price of unknown transient amplitude)**.** *For every fixed $\lambda > 0$ and $\theta = (\mu, \beta) \in \Theta_S$, the ridge intercept in* (4) *satisfies*

$$\sqrt{n}(\widehat{\mu}_{n,\lambda} - \mu) \Rightarrow N(0, \sigma^2\kappa(r)), \qquad \sup_{\theta \in \Theta_S} \left| n\mathbb{E}_\theta(\widehat{\mu}_{n,\lambda} - \mu)^2 - \sigma^2\kappa(r) \right| \to 0. \tag{9}$$

If β is known, the corrected mean $\widetilde{\mu}_n = n^{-1}\sum_{s=1}^n (Y_s - \beta s^{-r})$ obeys $\sqrt{n}(\widetilde{\mu}_n - \mu) \sim N(0, \sigma^2)$ exactly. Moreover, for every $\theta_0 \in \mathrm{int}(\Theta_S)$ and every sufficiently small $\rho > 0$ with $B(\theta_0, \rho) \subset \mathrm{int}(\Theta_S)$,

$$\lim_{n\to\infty} \frac{n}{\sigma^2\kappa(r)} \inf_{\widehat{\mu}_n} \sup_{\theta \in B(\theta_0,\rho)} \mathbb{E}_\theta(\widehat{\mu}_n - \mu)^2 = 1, \tag{10}$$

where the infimum is over all estimators based on the first n observations.

Proof. See Appendix A.

The exact two-regressor risk and a local least-favorable direction show that $\kappa(r)$ is the asymptotic variance-inflation factor relative to the revealed-amplitude oracle, not merely a rate constant. Appendix B extends the Schur criterion to arbitrary transient shapes, including logarithmic and arbitrarily slow monotone decay.

4 Deterministic Two-Arm Benchmark

For deterministic prefixes n_1, n_2 define the pooled intercept information

$$J(n_1, n_2) := \left(D_{n_1}^{-1} + D_{n_2}^{-1}\right)^{-1}. \tag{11}$$

Fix $\theta_0 = (\mu_0, \beta_0) \in \mathrm{int}(\Theta_S)$ and $\rho > 0$ with $B(\theta_0, \rho) \subset \mathrm{int}(\Theta_S)$, and consider the two local classes

$$\mathcal{P}_+(\Delta) := \{(\theta_1, \theta_2) \in B(\theta_0, \rho)^2 : \mu_2 - \mu_1 \geq \Delta\},$$
$$\mathcal{P}_-(\Delta) := \{(\theta_1, \theta_2) \in B(\theta_0, \rho)^2 : \mu_1 - \mu_2 \geq \Delta\},$$

on which arm 1 and arm 2, respectively, are best by at least Δ, and let $\mathcal{R}_T(\psi; n_1, n_2, \Delta)$ be the larger of the two worst-case misidentification probabilities of a recommendation rule ψ over them.

Proposition 1 (Exact two-arm deterministic benchmark)**.** *For fixed $c > 0$, put $\Delta_T = c/\sqrt{T}$. Every deterministic $n_1, n_2 \geq 2$ summing to T satisfies, for all sufficiently large T,*

$$\inf_\psi \mathcal{R}_T(\psi; n_1, n_2, \Delta_T) = \Phi\left(-\frac{\Delta_T\sqrt{J(n_1, n_2)}}{\sigma}\right), \tag{12}$$

and the rule that recommends the arm with the smaller OLS intercept attains the infimum. If $J_T^* := \max_{n_1+n_2=T,\, n_1,n_2\geq 2} J(n_1, n_2)$, then

$$J_T^*/T \to \{4\kappa(r)\}^{-1}, \qquad \min_{\substack{n_1+n_2=T \\ n_1,n_2\geq 2}} \inf_\psi \mathcal{R}_T(\psi; n_1, n_2, \Delta_T) \to \Phi\left(-\frac{c}{2\sigma\sqrt{\kappa(r)}}\right). \tag{13}$$

Proof. See Appendix C.

The proposition transfers the fixed-arm Schur information to two-arm identification and makes balanced sampling the deterministic benchmark. We next ask whether data-dependent allocation or stopping can recover the lost information.

5 Adaptive Two-Arm Identification

Fix a compact interval $\mathcal{B} \subset (0,\infty)$ with $\beta_0 \in \mathrm{int}(\mathcal{B})$. An allocation-and-stopping procedure is *parameter-free* if, conditional on the known problem inputs and the observed history, its kernels are the same for every unknown parameter vector. Its allocation is *predictable* if it chooses each arm before observing the next outcome, and its stopping time τ_T is adapted to the same filtration. For a parameter-free policy sequence $\boldsymbol{\pi} = (\pi_T)_{T\ge 1}$ and a local-scale bound M, define

$$\mathcal{R}_{T,M}^{\pi_T,\mathcal{B}}(c;r) := \inf_{\psi} \max\Big\{ \sup_{\substack{|u_i|\le M,\ u_2-u_1\ge c \\ \beta\in\mathcal{B}^2}} \mathbb{P}_{u,\beta}^{\pi_T}(\psi \neq 1), \sup_{\substack{|u_i|\le M,\ u_1-u_2\ge c \\ \beta\in\mathcal{B}^2}} \mathbb{P}_{u,\beta}^{\pi_T}(\psi \neq 2)\Big\}, \qquad \mu_i = \mu_0 + u_i/\sqrt{T}, \tag{14}$$

where the infimum is over parameter-free, possibly randomized recommendations ψ. Let $\mathfrak{P}_C$ be the set of such policy sequences with $\tau_T \le CT$ almost surely, and define the lower and upper minimax envelopes

$$\underline{\mathfrak{R}}_{\mathrm{end,un}}^{\star,\mathcal{B}}(c,C;r) := \inf_{\boldsymbol{\pi}\in\mathfrak{P}_C} \lim_{M\to\infty} \liminf_{T\to\infty} \mathcal{R}_{T,M}^{\pi_T,\mathcal{B}}(c;r),$$

$$\overline{\mathfrak{R}}_{\mathrm{end,un}}^{\star,\mathcal{B}}(c,C;r) := \inf_{\boldsymbol{\pi}\in\mathfrak{P}_C} \lim_{M\to\infty} \limsup_{T\to\infty} \mathcal{R}_{T,M}^{\pi_T,\mathcal{B}}(c;r).$$

The oracle envelopes $\underline{\mathfrak{R}}_{\mathrm{end,or}}^{\star,\mathcal{B}}$ and $\overline{\mathfrak{R}}_{\mathrm{end,or}}^{\star,\mathcal{B}}$ are defined in the same way, except that allocation, stopping, and recommendation may depend on the revealed amplitude vector β.

Theorem 2 (Exact asymptotic two-arm value under adaptive sampling)**.** *For every $r, c, C > 0$,*

$$\underline{\mathfrak{R}}_{\mathrm{end,un}}^{\star,\mathcal{B}}(c,C;r) = \overline{\mathfrak{R}}_{\mathrm{end,un}}^{\star,\mathcal{B}}(c,C;r) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma\sqrt{\kappa(r)}}\right). \tag{15}$$

The value does not depend on the particular $\mathcal{B}$ satisfying the interior condition, and both oracle envelopes equal $\Phi(-c\sqrt{C}/[2\sigma])$. At a fixed local gap and budget, the unknown-amplitude experiment thus retains $1/\kappa(r)$ of the oracle's squared Gaussian separation; equivalently, it needs $\kappa(r)$ times the first-order budget to reach the same error.

Proof. See Appendix E.

Proof roadmap. The lower bound for $0 < r < 1/2$ mixes over the nuisance parameters instead of fixing a single least-favorable pair. Consider the local amplitude paths $\beta_{i,T} = \beta_0 + v_i T^{r-1/2}$ and the local coordinates $\gamma = (u_1, v_1, u_2, v_2)^\top$. Write $d = (c/2)(-1,0,1,0)^\top$ for the signed gap and $G\xi = (\xi_0, \xi_1, \xi_0, \xi_2)^\top$ for the common intercept and the two amplitude nuisances. We mix $\gamma = \pm d + G\xi$ under the proper prior $\xi \sim N(0, a^2 I_3)$ and denote the resulting laws by $\mathbb{P}_{\pm,a,T}$; the zero-gap mixture $\gamma = G\xi$ is $\mathbb{P}_{0,a,T}$. When the policy makes the sth local pull of arm 1, its design row is $T^{-1/2}(1,(s/T)^{-r},0,0)$, and analogously for arm 2. Let H collect the rows selected through τ_T, and set $h = Hd$, $\mathsf{Z} = HG$, and $Q_T = H^\top H$. Because the action and stopping kernels are parameter-free, they cancel from every likelihood ratio, and one obtains the exact sequential identities

$$\frac{d\mathbb{P}_{\pm,a,T}}{d\mathbb{P}_{0,a,T}} = \exp\left\{\pm\frac{\mathcal{M}_T}{2} - \frac{V_T}{8}\right\}, \tag{16}$$

$$\mathbb{E}_{0,a,T} e^{t\mathcal{M}_T - t^2 V_T/2} = 1 \qquad (t\in\mathbb{R}), \tag{17}$$

$$\frac{V_T}{4} = h^\top(\sigma^2 I + a^2 \mathsf{Z}\mathsf{Z}^\top)^{-1} h = \frac{1}{\sigma^2}\min_{\zeta\in\mathbb{R}^3}\left\{(d-G\zeta)^\top Q_T (d-G\zeta) + \sigma^2 a^{-2}\|\zeta\|^2\right\}, \tag{18}$$

in which $\mathcal{M}_T$ is a score martingale and V_T its bracket. The bracket formula holds on every realized design path, and no step conditions on the adaptively selected design (Lemma 8). With $p_i = N_i(\tau_T)/T$, uniform prefix Riemann sums identify the bracket limit for fixed a as a continuum ridge Schur complement, and as $a \to \infty$ these complements converge uniformly in p to

$$J_0(p) := \frac{c^2}{\sigma^2} \frac{p_1 p_2}{\kappa(r)(p_1 + p_2)} \le \frac{c^2 C}{4\sigma^2 \kappa(r)}, \qquad p_1 + p_2 \le C, \tag{19}$$

where the ratio is read as zero at $p_1 = p_2 = 0$. Hence, for every $\epsilon > 0$, one can choose first a finite a and then T so large that every admissible policy and every sample path satisfy

$$V_T \le \frac{c^2 C}{\sigma^2 \kappa(r)} + \epsilon \tag{20}$$

(Lemmas 9 and 10). Since $\mathcal{M}_T$ need not be conditionally Gaussian given V_T, set $\overline{V} = c^2 C/\{\sigma^2 \kappa(r)\} + \epsilon$ and enlarge the experiment by drawing, conditionally on the stopped history, $W \sim N(0, \overline{V} - V_T)$ (Lemma 7). The completed score $\mathcal{M}_T + W$ is exactly $N(0, \overline{V})$, so data processing bounds the original equal-prior error from below by $\Phi(-\sqrt{\overline{V}}/2)$. To embed the proper but unbounded prior in (14), we restrict it, for $M > c/2$, to

$$\mathcal{K}_M = \{|\xi_0| \le M - c/2,\ |\xi_1| \le M,\ |\xi_2| \le M\}.$$

For fixed M the corresponding amplitude paths lie in $\mathcal{B}^2$ for all large T, and truncation changes either mixture law in total variation by at most $\mathbb{P}(\xi \notin \mathcal{K}_M)$. Taking $\liminf_T$, then $M \to \infty$, and finally $\epsilon \downarrow 0$ gives the lower bound.

For $r \ge 1/2$, the equal-amplitude symmetric endpoints already have the exact bracket $V_T = c^2 \tau_T/(\sigma^2 T) \le c^2 C/\sigma^2$ (Lemma 11), which gives the same bound with $\kappa(r) = 1$. For the upper bound, sampling each arm $n_T = \lfloor CT/2 \rfloor$ times gives

$$\widehat{\mu}_2 - \widehat{\mu}_1 \sim N\left(\mu_2 - \mu_1, \frac{2\sigma^2}{D_{n_T}}\right), \qquad \frac{D_{n_T}}{n_T} \longrightarrow \frac{1}{\kappa(r)}, \tag{21}$$

so the OLS sign rule matches the lower bound uniformly over $\mathcal{B}^2$. For the oracle lower bound in every regime, the two symmetric endpoints share the same amplitude vector, which therefore stays identical under both signs even after revelation; their bracket is at most $c^2 C/\sigma^2$, and balanced corrected means attain the resulting value. Appendix E contains the truncation and uniformity arguments in full.

The minimax identity does not make terminal counts sufficient for evaluating an individual adaptive policy: for $0 < r < 1/2$, Appendix J derives the joint score–information limit, Theorem 4 identifies the subclass in which the allocation is ancillary, and Appendix D exhibits policies with identical terminal-count laws but different local two-point risks.

6 From Two Arms to K Arms

The two-arm value extends to fixed K whenever a sublinear pilot can locate the only local decision boundary. For $\eta \in (0, \rho/4)$, a best arm b, and a single potentially local challenger $j \ne b$, let

$$\mathcal{C}_{b,j}(\Delta, \eta) := \left\{(\theta_i)_{i=1}^K \in B(\theta_0, \rho)^K :\ \mu_j - \mu_b \ge \Delta, \quad \mu_k - \max(\mu_b, \mu_j) \ge \eta\ (k \notin \{b, j\})\right\}. \tag{22}$$

Screen–then–compare. Choose $\omega \in (0, 1)$ and set $m_T = \lceil T^\omega \rceil$. First pull every arm m_T times and select the two smallest pilot OLS intercepts. Allocate the remaining budget equally to this pair. Finally, discard the pilot observations, fit OLS to the two independent suffixes, and recommend the smaller intercept. The pilot costs $Km_T = o(T)$ pulls, and suffix independence permits an exact two-arm comparison. In the revealed-amplitude experiment, the pilot and suffix comparisons use amplitude-corrected sample means instead.

Let $\mathfrak{R}^{K,\mathrm{pair}}_{T,\bullet}(c, C, \eta)$ denote the minimax error over parameter-free policies with predictable allocation, randomized recommendation, and $\tau_T \le CT$ almost surely, where the maximum is taken over all ordered pairs (b, j), which are unknown to the policy, and over $\mathcal{C}_{b,j}(c/\sqrt{T}, \eta)$. For $\bullet = \mathrm{un}$ the amplitudes are unknown; for $\bullet = \mathrm{or}$ all kernels may depend on the revealed amplitude vector. For fixed η, let $B^{K,\mathrm{pair}}_{\bullet}(\Delta, \alpha)$ be the smallest integer B for which the corresponding worst-case error under $\tau \le B$ almost surely is at most α.

Theorem 3 (Exact asymptotic screened-K value)**.** *For every fixed $K \geq 2$, $r > 0$, $c, C > 0$, and admissible η,*

$$\lim_{T\to\infty} \mathfrak{R}^{K,\mathrm{pair}}_{T,\mathrm{un}}(c, C, \eta) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma\sqrt{\kappa(r)}}\right), \qquad \lim_{T\to\infty} \mathfrak{R}^{K,\mathrm{pair}}_{T,\mathrm{or}}(c, C, \eta) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma}\right). \quad (23)$$

For every target error $\alpha \in (0, 1/2)$, the minimal almost-sure budget caps satisfy, as $\Delta \downarrow 0$,

$$B^{K,\mathrm{pair}}_{\mathrm{un}}(\Delta, \alpha) \sim \frac{4\sigma^2\kappa(r)z^2_\alpha}{\Delta^2}, \qquad B^{K,\mathrm{pair}}_{\mathrm{or}}(\Delta, \alpha) \sim \frac{4\sigma^2 z^2_\alpha}{\Delta^2}. \quad (24)$$

Thus the unknown-amplitude first-order budget inflation remains $\kappa(r)$, while adaptive screening reduces the optimized deterministic all-arm-prefix budget by the factor $K/2$.

Proof. See Appendix G.

With probability $1 - o(1)$ the pilot finds the competitive pair, all but $o(T)$ pulls then go to that pair, and suffix OLS retains $1/\kappa(r)$ of the oracle intercept information per pull in every decay regime. The $K/2$ reduction relies on the fixed margin and does not cover simultaneous K-way root-T ties.

Secondary consequences. Appendix F gives the exact deterministic-prefix Gaussian K-arm reduction, Theorem 15 its stable-proportion predictable extension, and Theorem 9 the deterministic one-challenger budget behind the $K/2$ comparison. At a fixed nonzero gap, Theorem 16 gives an unknown-to-oracle budget ratio of one under a known amplitude-difference envelope. These results delimit the scope of the theory but lie outside the main minimax argument.

7 Finite-Sample Checks

Does the sharp factor appear at finite prefixes? We evaluate the fixed-λ ridge intercept in (4) at $(\mu, \beta, \sigma, \lambda) = (.25, .75, .20, 1)$ on a grid of 16 decay exponents, nine of them in $[.45, .55]$, and eight prefix lengths from 2^6 to 2^{20}. In every cell we compute the exact Gaussian risk and draw 200,000 independent Gaussian sufficient statistics. The revealed-amplitude corrected mean has normalized risk exactly one. As the prefix grows, the unknown-amplitude risk moves pointwise toward $\kappa(r)$, but the convergence is strongly nonuniform near the boundary: at $n = 2^{20}$, the exact normalized risks are 1.83 at $r = .49$ and 1.75 at $r = .50$, against the limits 1.083 and 1, whereas at $r = 1$ the risk is already 1.0004. Across all 128 cells, the simulated unknown-amplitude risks track the exact values, with a maximum absolute standardized deviation of 3.08. This checks the finite-prefix calculation behind Theorem 1, not its local minimax lower bound.

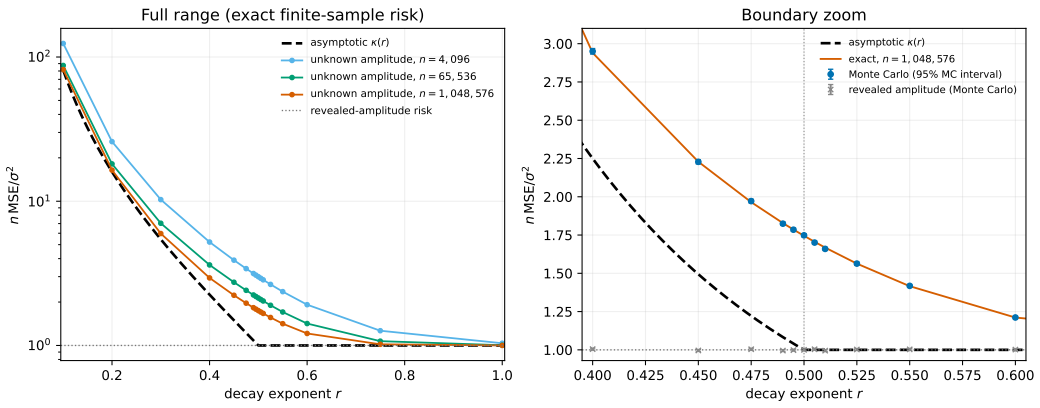

Figure 1: **Direct finite-prefix check of the sharp factor.** Shown is $n\,\mathbb{E}(\widehat{\mu} - \mu)^2/\sigma^2$ versus r. Left: exact unknown-amplitude risks at three prefix lengths, the pointwise limit $\kappa(r)$ (dashed), and the revealed-amplitude value one (dotted). Right: boundary zoom at $n = 2^{20}$; the solid curve is the exact risk, and the points show 200,000 Monte Carlo replications with 95% Monte Carlo intervals.

Does screening approach the fixed-K prediction? We test the constructive policy of Theorem 3 on the boundary instance $\mu_1 = 0$, $\mu_2 = \Delta$, and $\mu_k = \Delta + \eta$ for $k \geq 3$, with $\eta = 1.5$, $\Delta = c/\sqrt{B}$, unit noise variance, and a fixed amplitude profile equally spaced from .5 to 1.5. Across $K \in \{2, 4, 8\}$ and exponents below, at, and above $r = 1/2$, the pair-screening error decreases with the budget B and the squared critical-signal gain moves toward $K/2$. Figure 2 shows the representative case $K = 8$ and the slower convergence near the boundary. Gaussian integration gives the exact quantities for the declared policies, and Monte Carlo checks them independently. This tests the construction, not finite-sample minimaxity; Appendix K.1 gives the protocol and additional cases.

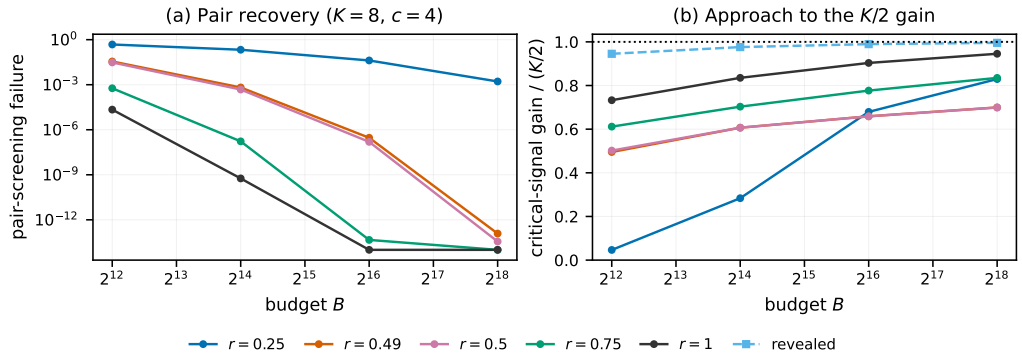

Figure 2: **Finite-budget behavior of screen–then–compare for $K = 8$.** (a) Exact pair-screening failure probability for unknown amplitudes and $c = 4$. (b) The squared critical-signal ratio $[c_{\mathrm{bal}}(B, r)/c_{\mathrm{stc}}(B, r)]^2$, normalized by $K/2$, at error .10. Solid curves use unknown amplitudes, the dashed curve uses revealed amplitudes, and the dotted line marks the asymptotic prediction. For the logarithmic display in (a), values below 10^{-14} are floored at 10^{-14}.

8 Limitations and Conclusion

The fixed-arm variance inflation $\kappa(r)$ is also the exact local unknown-to-oracle first-order budget ratio for two-arm best-asymptote identification, even after optimizing over predictable allocation and almost-surely capped stopping; when at most one challenger may be local, sublinear screening attains the same value for every fixed K and reduces the optimized deterministic all-arm-prefix budget by the factor $K/2$. The two constants describe different bottlenecks: $\kappa(r)$ is a within-arm efficiency loss caused by estimating the transient amplitude, and it survives optimization over allocation and stopping, whereas $K/2$ is an across-arm allocation gain from screening out arms separated from the only local pair, and it depends on the one-challenger margin.

The theory assumes known r and σ^2, Gaussian innovations, and an almost-sure rather than expected budget cap. The fixed-K result requires a fixed margin; simultaneous local ties and random active sets remain open. Policy–specific score–information limits are subsequential and restricted to $0 < r < 1/2$; the fixed-confidence result assumes a known amplitude-difference envelope, and our Gaussian checks do not establish practical superiority for hyperparameter optimization.

AI USE STATEMENT

Generative AI tools assisted in the literature search, proof verification, proof refinements, software implementation, and editing of the manuscript. The authors reviewed all AI-assisted work, checked literature claims against the primary sources, and independently verified every theorem statement and proof before inclusion. They also checked the simulation code against the exact formulas and preserved frozen protocols, deviations, and hashes for the reported experiments. The authors take full responsibility for the final text, claims, proofs, code, and artifacts.

REPRODUCIBILITY STATEMENT

Appendices A–K contain complete proofs and a summary of the experimental checks. The supplementary material includes everything needed for result reproduction.

REFERENCES

Leonardo Cella, Massimiliano Pontil, and Claudio Gentile. Best model identification: A rested bandit formulation. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 1362–1372. PMLR, 2021. URL https://proceedings.mlr.press/v139/cella21a.html.

David R. Cox and Nancy Reid. Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society: Series B (Methodological)*, 49(1):1–39, 1987. doi: 10.1111/j.2517-6161.1987.tb01422.x.

Yuriy Dorn, Aleksandr Katrutsa, Ilgam Latypov, and Anastasia Soboleva. Functional multi-armed bandit and the best function identification problems. In *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems*, pp. 2913–2921. IFAAMAS, 2026. doi: 10.65109/SFMN9947. URL https://ifaamas.csc.liv.ac.uk/Proceedings/aamas2026/pdfs/SFMN9947.pdf.

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

Jaroslav Hájek. Local asymptotic minimax and admissibility in estimation. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pp. 175–194. University of California Press, Berkeley, 1972. URL https://digicoll.lib.berkeley.edu/record/113183.

Keisuke Hirano and Jack R. Porter. Asymptotic representations for sequential decisions, adaptive experiments, and batched bandits. *CoRR*, abs/2302.03117, 2025. URL https://arxiv.org/abs/2302.03117.

Masahiro Kato. Locally optimal fixed-budget best arm identification in two-armed gaussian bandits with unknown variances. *CoRR*, abs/2312.12741, 2024a. URL https://arxiv.org/abs/2312.12741.

Masahiro Kato. Worst-case optimal multi-armed gaussian best arm identification with a fixed budget. *CoRR*, abs/2310.19788, 2024b. URL https://arxiv.org/abs/2310.19788.

Masahiro Kato. Generalized neyman allocation for locally minimax optimal best-arm identification. *CoRR*, abs/2405.19317, 2025. URL https://arxiv.org/abs/2405.19317.

Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of best-arm identification in multi-armed bandit models. *Journal of Machine Learning Research*, 17(1):1–42, 2016. URL https://www.jmlr.org/papers/v17/kaufman16a.html.

Jack Kiefer and Jacob Wolfowitz. Optimum designs in regression problems. *The Annals of Mathematical Statistics*, 30(2):271–294, 1959. doi: 10.1214/aoms/1177706252.

Lucien Le Cam. *Asymptotic Methods in Statistical Decision Theory*. Springer Series in Statistics. Springer, New York, 1986. doi: 10.1007/978-1-4612-4946-7. URL https://doi.org/10.1007/978-1-4612-4946-7.

Lucien M. Le Cam. Locally asymptotically normal families of distributions: Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. *University of California Publications in Statistics*, 3(2):37–98, 1960.

Marco Mussi, Alessandro Montenegro, Francesco Trovò, Marcello Restelli, and Alberto Maria Metelli. Best arm identification for stochastic rising bandits. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 36953–36989. PMLR, 2024. URL https://proceedings.mlr.press/v235/mussi24b.html.

Sho Takemori, Yuhei Umeda, and Aditya Gopalan. Model-based best arm identification for decreasing bandits. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 1567–1575. PMLR, 2024. URL https://proceedings.mlr.press/v238/takemori24a.html.

Ramon van den Akker, Bas J. M. Werker, and Bo Zhou. Valid post-contextual bandit inference. *CoRR*, abs/2505.13897, 2025a. URL https://arxiv.org/abs/2505.13897.

Ramon van den Akker, Bas J. M. Werker, and Bo Zhou. Local asymptotic normality for multi-armed bandits. *CoRR*, abs/2512.12192, 2025b. URL https://arxiv.org/abs/2512.12192.

A Proof of the Single-Arm Price Theorem

This appendix proves Theorem 1. Fix one arm and suppress its index, so that the model (1) reads, for a fixed known exponent $r > 0$,

$$Y_s = \mu + \beta s^{-r} + \varepsilon_s, \qquad \varepsilon_s \overset{\text{iid}}{\sim} N(0, \sigma^2), \qquad s = 1, \dots, n. \tag{25}$$

As in Section 3, write $\theta = (\mu, \beta)^\top$, $z_s = (1, s^{-r})^\top$, and

$$Q_n := \sum_{s=1}^{n} z_s z_s^\top = \begin{pmatrix} n & a_n \\ a_n & b_n \end{pmatrix}, \qquad a_n := \sum_{s=1}^{n} s^{-r}, \qquad b_n := \sum_{s=1}^{n} s^{-2r}, \tag{26}$$

$$Q_{n,\lambda} := Q_n + \lambda I_2, \qquad \widehat{\theta}_{n,\lambda} := Q_{n,\lambda}^{-1} \sum_{s=1}^{n} z_s Y_s, \qquad \widehat{\mu}_{n,\lambda} := e_1^\top \widehat{\theta}_{n,\lambda}, \tag{27}$$

where the ridge parameter $\lambda > 0$ stays fixed as $n \to \infty$. Recall that

$$\kappa(r) := \begin{cases} (1-r)^2/r^2, & 0 < r < 1/2, \\ 1, & r \geq 1/2. \end{cases} \tag{28}$$

The lower bounds are stated on the one-arm parameter space

$$\Theta_S := \{(\mu, \beta) \in \mathbb{R}^2 : \mu^2 + \beta^2 \leq S^2, \beta \geq 0\}. \tag{29}$$

so that an interior point $\theta_0 = (\mu_0, \beta_0)$ satisfies $\|\theta_0\| < S$ and $\beta_0 > 0$. The argument has four steps: the Schur-complement geometry of the design, the exact ridge risk, an exact finite-sample information bound, and the local minimax lower bound.

A.1 Schur-complement geometry

We use the ridge and the unregularized intercept leverages

$$q_{n,\lambda} := e_1^\top Q_{n,\lambda}^{-1} e_1 = \frac{1}{n + \lambda - a_n^2/(b_n + \lambda)}, \tag{30}$$

$$\bar{q}_n := e_1^\top Q_n^{-1} e_1 = \frac{1}{n - a_n^2/b_n}, \tag{31}$$

$$\varsigma_{n,\lambda} := \frac{a_n}{b_n + \lambda}, \qquad \bar{\varsigma}_n := \frac{a_n}{b_n}. \tag{32}$$

For every $r > 0$ and $n \geq 2$, Q_n is positive definite because the vectors $(1, \ldots, 1)$ and $(1^{-r}, \ldots, n^{-r})$ are not collinear.

Lemma 1 (Ridge and information geometry)**.** *For each fixed $r > 0$ and fixed $\lambda > 0$,*

$$nq_{n,\lambda} \longrightarrow \kappa(r), \qquad n\bar{q}_n \longrightarrow \kappa(r), \tag{33}$$

and

$$\frac{\varsigma_{n,\lambda}}{\sqrt{n}} \longrightarrow 0, \qquad \frac{\bar{\varsigma}_n}{\sqrt{n}} \longrightarrow 0. \tag{34}$$

In particular, $nq_{n,\lambda}$ and $n\bar{q}_n$ are bounded away from zero and infinity for all sufficiently large n.

Proof. For $\upsilon > 0$, the standard power-sum estimates are

$$\sum_{s=1}^{n} s^{-\upsilon} = \begin{cases} \dfrac{n^{1-\upsilon}}{1-\upsilon}(1 + o(1)), & 0 < \upsilon < 1, \\ \log n + O(1), & \upsilon = 1, \\ \zeta(\upsilon) + o(1), & \upsilon > 1. \end{cases} \tag{35}$$

If $0 < r < 1/2$, then

$$a_n = \frac{n^{1-r}}{1-r}(1 + o(1)), \qquad b_n = \frac{n^{1-2r}}{1-2r}(1 + o(1)), \tag{36}$$

so, both with and without the fixed ridge term,

$$\frac{a_n^2}{n(b_n + \lambda)} \longrightarrow \frac{1-2r}{(1-r)^2}, \qquad \frac{a_n^2}{nb_n} \longrightarrow \frac{1-2r}{(1-r)^2}. \tag{37}$$

The residual-information fraction is $1 - (1-2r)/(1-r)^2 = r^2/(1-r)^2$, which proves (33) in this regime. Moreover, $\varsigma_{n,\lambda} \asymp \bar{\varsigma}_n \asymp n^r = o(\sqrt{n})$.

At $r = 1/2$, $a_n = 2\sqrt{n} + O(1)$ and $b_n = \log n + O(1)$. Hence

$$\frac{a_n^2}{n(b_n + \lambda)} = O((\log n)^{-1}) \to 0, \qquad \frac{a_n^2}{nb_n} = O((\log n)^{-1}) \to 0, \tag{38}$$

while $\varsigma_{n,\lambda}/\sqrt{n}$ and $\bar{\varsigma}_n/\sqrt{n}$ are $O((\log n)^{-1})$. Thus the critical constant is exactly one.

If $r > 1/2$, then $b_n \to \zeta(2r) \in (0, \infty)$, whereas

$$a_n = \begin{cases} O(n^{1-r}), & 1/2 < r < 1, \\ O(\log n), & r = 1, \\ O(1), & r > 1. \end{cases} \tag{39}$$

Thus $a_n/\sqrt{n} \to 0$, and both explained fractions $a_n^2/[n(b_n + \lambda)]$ and $a_n^2/(nb_n)$ vanish. This gives (33) with $\kappa(r) = 1$ and also (34). □

For later feasibility and bias checks, the preceding calculation also gives

$$\frac{1 + \varsigma_{n,\lambda}}{\sqrt{n}}, \quad \frac{1 + \bar{\varsigma}_n}{\sqrt{n}} = \begin{cases} O(n^{r-1/2}), & 0 < r < 1/2, \\ O((\log n)^{-1}), & r = 1/2, \\ O(n^{1/2-r}), & 1/2 < r < 1, \\ O((\log n)/\sqrt{n}), & r = 1, \\ O(n^{-1/2}), & r > 1, \end{cases} \tag{40}$$

and every displayed remainder tends to zero for fixed $r > 0$.

A.2 EXACT RIDGE RISK AND THE REVEALED-AMPLITUDE LAW

Let $\boldsymbol{Z}_n \in \mathbb{R}^{n\times 2}$ have rows $z_s^\top$, and put $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^\top$. Since $Q_n = \boldsymbol{Z}_n^\top \boldsymbol{Z}_n$,

$$\widehat{\theta}_{n,\lambda} - \theta = -\lambda Q_{n,\lambda}^{-1}\theta + Q_{n,\lambda}^{-1}\boldsymbol{Z}_n^\top \varepsilon. \tag{41}$$

Direct inversion gives

$$e_1^\top Q_{n,\lambda}^{-1} = q_{n,\lambda}(1, -\varsigma_{n,\lambda}), \tag{42}$$

so the exact intercept bias is

$$\mathrm{Bias}_n(\theta) := \mathbb{E}_\theta[\widehat{\mu}_{n,\lambda} - \mu] = -\lambda q_{n,\lambda}(\mu - \varsigma_{n,\lambda}\beta). \tag{43}$$

If $\theta \in \Theta_S$, then

$$|\mathrm{Bias}_n(\theta)| \le \lambda q_{n,\lambda} S\sqrt{1+\varsigma_{n,\lambda}^2}. \tag{44}$$

Lemma 1 implies

$$\sup_{\theta\in\Theta_S} \sqrt{n}|\mathrm{Bias}_n(\theta)| \to 0, \qquad \sup_{\theta\in\Theta_S} n\mathrm{Bias}_n(\theta)^2 \to 0. \tag{45}$$

The coefficient of ε_s in the ridge intercept is $q_{n,\lambda}(1 - \varsigma_{n,\lambda}s^{-r})$. Hence, with

$$R_n := \sum_{s=1}^{n}(1 - \varsigma_{n,\lambda}s^{-r})^2, \tag{46}$$

the exact finite-sample distribution and mean-squared error are

$$\widehat{\mu}_{n,\lambda} - \mu \sim N\big(\mathrm{Bias}_n(\theta),\, \sigma^2 q_{n,\lambda}^2 R_n\big)\,, \tag{47}$$

$$\mathbb{E}_\theta[(\widehat{\mu}_{n,\lambda} - \mu)^2] = \mathrm{Bias}_n(\theta)^2 + \sigma^2 q_{n,\lambda}^2 R_n. \tag{48}$$

The variance term has the exact simplification

$$R_n = q_{n,\lambda}^{-1} - \lambda(1 + \varsigma_{n,\lambda}^2). \tag{49}$$

Indeed, substitution of $\varsigma_{n,\lambda} = a_n/(b_n + \lambda)$ yields

$$R_n = n - \frac{a_n^2(b_n + 2\lambda)}{(b_n+\lambda)^2} = n + \lambda - \frac{a_n^2}{b_n+\lambda} - \lambda - \frac{\lambda a_n^2}{(b_n+\lambda)^2}.$$

It follows that

$$nq_{n,\lambda}^2 R_n = nq_{n,\lambda} - \lambda n q_{n,\lambda}^2(1+\varsigma_{n,\lambda}^2) \longrightarrow \kappa(r), \tag{50}$$

because

$$nq_{n,\lambda}^2(1+\varsigma_{n,\lambda}^2) = (nq_{n,\lambda})^2\frac{1+\varsigma_{n,\lambda}^2}{n} \longrightarrow 0. \tag{51}$$

The variable in (47) is Gaussian for every n; thus (45) and (50) prove

$$\sqrt{n}(\widehat{\mu}_{n,\lambda} - \mu) \Longrightarrow N(0, \sigma^2\kappa(r)), \tag{52}$$

as well as the uniform risk limit

$$\sup_{\theta\in\Theta_S} \big|n\mathbb{E}_\theta[(\widehat{\mu}_{n,\lambda} - \mu)^2] - \sigma^2\kappa(r)\big| \longrightarrow 0. \tag{53}$$

If the transient amplitude β is revealed, the corrected mean

$$\widetilde{\mu}_n := \frac{1}{n}\sum_{s=1}^{n}(Y_s - \beta s^{-r}) \tag{54}$$

satisfies the exact finite-sample law

$$\sqrt{n}(\widetilde{\mu}_n - \mu) \sim N(0, \sigma^2). \tag{55}$$

Therefore the first-order variance inflation due solely to the unknown amplitude is $\kappa(r)$: it exceeds one for $0 < r < 1/2$, equals one for $r \ge 1/2$, and diverges as r^{-2} when $r \downarrow 0$.

A.3 EXACT FINITE-SAMPLE INFORMATION BOUND

For $n \geq 2$, the score and Fisher information in the full two-parameter Gaussian model are

$$\mathcal{S}_\theta(y) = \frac{1}{\sigma^2} \boldsymbol{Z}_n^\top (y - \boldsymbol{Z}_n \theta), \qquad \mathcal{I}_n(\theta) = \frac{1}{\sigma^2} Q_n. \tag{56}$$

Let $\widehat{\mu}_n$ have finite variance and be unbiased for μ at every $(\mu, \beta) \in \mathbb{R}^2$. Differentiating the Gaussian density gives

$$\mathbb{E}_\theta[(\widehat{\mu}_n - \mu)\mathcal{S}_\theta] = e_1. \tag{57}$$

Taking $w = \mathcal{I}_n(\theta)^{-1} e_1$ and applying Cauchy–Schwarz to the covariance of $\widehat{\mu}_n$ and $w^\top \mathcal{S}_\theta$ yields

$$\mathrm{Var}_\theta(\widehat{\mu}_n) \geq e_1^\top \mathcal{I}_n(\theta)^{-1} e_1 = \sigma^2 e_1^\top Q_n^{-1} e_1 = \sigma^2 \bar{q}_n. \tag{58}$$

The ordinary least-squares estimator $\widehat{\theta}_{n,0} = Q_n^{-1} \boldsymbol{Z}_n^\top y$ is unbiased and satisfies

$$\mathrm{Cov}_\theta(\widehat{\theta}_{n,0}) = \sigma^2 Q_n^{-1}. \tag{59}$$

Its intercept attains equality in (58); hence $\sigma^2 \bar{q}_n$ is the exact finite-sample information price among globally unbiased estimators.

A.4 ARBITRARY-ESTIMATOR LOWER BOUNDS

Define the least-favorable direction

$$\mathbf{q}_n := \frac{Q_n^{-1} e_1}{\sqrt{\bar{q}_n}} = \sqrt{\bar{q}_n}(1, -\bar{\varsigma}_n)^\top. \tag{60}$$

It obeys

$$\mathbf{q}_n^\top Q_n \mathbf{q}_n = 1, \qquad e_1^\top \mathbf{q}_n = \sqrt{\bar{q}_n}, \qquad \|\mathbf{q}_n\| \longrightarrow 0. \tag{61}$$

The identities are immediate from the definition, while

$$\|\mathbf{q}_n\|^2 = \bar{q}_n(1 + \bar{\varsigma}_n^2) = (n\bar{q}_n)\frac{1 + \bar{\varsigma}_n^2}{n} \longrightarrow 0 \tag{62}$$

follows from Lemma 1. In particular, the nuisance-coordinate perturbation has the explicit orders

$$|e_2^\top \mathbf{q}_n| = \sqrt{\bar{q}_n}\,\bar{\varsigma}_n = \begin{cases} O(n^{r-1/2}), & 0 < r < 1/2, \\ O((\log n)^{-1}), & r = 1/2, \\ O(n^{1/2-r}), & 1/2 < r < 1, \\ O((\log n)/\sqrt{n}), & r = 1, \\ O(n^{-1/2}), & r > 1. \end{cases} \tag{63}$$

Thus no hidden feasibility failure occurs at the critical exponent. If θ_0 is interior, choose

$$0 < \rho < \min\{\beta_0, S - \|\theta_0\|\}. \tag{64}$$

Then $B(\theta_0, \rho) \subset \Theta_S$, and $\theta_0 + \sigma \mathbf{q}_n \in B(\theta_0, \rho)$ for all sufficiently large n.

For such n, set

$$\theta_{0,n} := \theta_0, \qquad \theta_{1,n} := \theta_0 + \sigma \mathbf{q}_n. \tag{65}$$

Let $P_{j,n}$ be the joint law under $\theta_{j,n}$. Since their covariance is $\sigma^2 I_n$,

$$\mathrm{KL}(P_{0,n} \| P_{1,n}) = \frac{1}{2\sigma^2} \|\boldsymbol{Z}_n(\theta_{1,n} - \theta_{0,n})\|^2 = \frac{1}{2} \mathbf{q}_n^\top Q_n \mathbf{q}_n = \frac{1}{2}. \tag{66}$$

Their intercepts are separated by

$$\Delta_n := \mu_{1,n} - \mu_{0,n} = \sigma e_1^\top \mathbf{q}_n = \sigma \sqrt{\bar{q}_n}. \tag{67}$$

For an arbitrary estimator $\widehat{\mu}_n$, use the test $\psi_n = \mathbf{1}\{\widehat{\mu}_n \geq (\mu_{0,n} + \mu_{1,n})/2\}$. On its error event under either hypothesis the squared estimation error is at least $\Delta_n^2/4$, so

$$\max_{j \in \{0,1\}} \mathbb{E}_{\theta_{j,n}}[(\widehat{\mu}_n - \mu_{j,n})^2] \geq \frac{\Delta_n^2}{8}\{P_{0,n}(\psi_n = 1) + P_{1,n}(\psi_n = 0)\} \tag{68}$$

$$\geq \frac{\Delta_n^2}{8}\{1 - \mathrm{TV}(P_{0,n}, P_{1,n})\}. \tag{69}$$

Pinsker's inequality and (66) give $\mathrm{TV}(P_{0,n}, P_{1,n}) \le 1/2$. Therefore

$$\max_{j\in\{0,1\}} \mathbb{E}_{\theta_{j,n}}[(\widehat{\mu}_n - \mu_{j,n})^2] \ge \frac{\sigma^2 \bar{q}_n}{16}, \qquad n\bar{q}_n \longrightarrow \kappa(r). \tag{70}$$

This is a finite-sample two-point bound for every estimator, with the two parameters converging to the same interior point.

It remains to establish the exact constant-one local asymptotic minimax bound. Fix any sufficiently small ball $B(\theta_0, \rho) \subset \Theta_S$, and consider the one-dimensional submodel

$$\theta_n(t) := \theta_0 + \sigma t \mathbf{q}_n, \qquad |t| \le L_n, \qquad L_n := \frac{\rho}{2\sigma \|\mathbf{q}_n\|}. \tag{71}$$

Then $L_n \to \infty$, the submodel lies in the ball, its Fisher information for t is exactly

$$(\sigma \boldsymbol{Z}_n \mathbf{q}_n)^\top (\sigma^2 I_n)^{-1} (\sigma \boldsymbol{Z}_n \mathbf{q}_n) = \mathbf{q}_n^\top Q_n \mathbf{q}_n = 1, \tag{72}$$

and its target is

$$\mu_n(t) = \mu_0 + \sigma\sqrt{\bar{q}_n}\, t. \tag{73}$$

For completeness, we prove the required van Trees inequality. On $[-1, 1]$, let

$$\varphi(z) := \cos^2(\pi z/2)\mathbf{1}\{|z| \le 1\}, \qquad \varphi_L(t) := L^{-1}\varphi(t/L). \tag{74}$$

This density integrates to one, vanishes at the endpoints, and has prior Fisher information

$$I(\varphi_L) := \int_{-L}^{L} \frac{\varphi_L'(t)^2}{\varphi_L(t)}\, dt = \frac{\pi^2}{L^2}. \tag{75}$$

If $f_t(y)$ is the likelihood in the submodel, define the joint score

$$\Psi(t, y) := \partial_t \log f_t(y) + \frac{\varphi_L'(t)}{\varphi_L(t)}. \tag{76}$$

For any estimator $\delta_n(y)$, integration by parts in t gives

$$\mathbb{E}[(\delta_n(Y) - t)\Psi(t, Y)] = 1. \tag{77}$$

The boundary term vanishes because $\varphi_L(\pm L) = 0$. The likelihood score has conditional mean zero, so its cross term with the prior score vanishes. By (72) and (75),

$$\mathbb{E}[\Psi(t, Y)^2] = 1 + \frac{\pi^2}{L^2}. \tag{78}$$

Cauchy–Schwarz now yields

$$\int_{-L}^{L} \mathbb{E}_t[(\delta_n - t)^2]\varphi_L(t)\, dt \ge \frac{1}{1 + \pi^2/L^2}. \tag{79}$$

For rigor, if the integrated risk is infinite the inequality is immediate. Otherwise apply the integration-by-parts calculation first to the clipped estimator $\delta_{n,m} = (-m) \vee (\delta_n \wedge m)$, then send $m \to \infty$. Joint L^2 convergence and $\Psi \in L^2$ justify the limit in (77).

Given an arbitrary intercept estimator $\widehat{\mu}_n$, take $L = L_n$ and

$$\delta_n(Y) := \frac{\widehat{\mu}_n(Y) - \mu_0}{\sigma\sqrt{\bar{q}_n}}. \tag{80}$$

Equations (71)–(79) imply the finite-n bound

$$\sup_{\theta \in B(\theta_0,\rho)} \mathbb{E}_\theta[(\widehat{\mu}_n - \mu)^2] \ge \frac{\sigma^2 \bar{q}_n}{1 + \pi^2/L_n^2}. \tag{81}$$

Taking the infimum over $\widehat{\mu}_n$, using $L_n \to \infty$ and $n\bar{q}_n \to \kappa(r)$, proves the lower half of

$$\lim_{n\to\infty} \frac{n}{\sigma^2 \kappa(r)} \inf_{\widehat{\mu}_n} \sup_{\theta \in B(\theta_0,\rho)} \mathbb{E}_\theta[(\widehat{\mu}_n - \mu)^2] = 1. \tag{82}$$

For the reverse inequality, use the ridge intercept itself in the infimum; the uniform risk limit (53) gives a limsup at most one. This completes the proof of Theorem 1.

Scope. The result is pointwise in the fixed known exponent $r > 0$. It is not uniform as $r \downarrow 0$, where $\kappa(r) \asymp r^{-2}$, nor across shrinking neighborhoods of the critical exponent $r = 1/2$, where the convergence is logarithmically slow. The theorem quantifies the asymptote-estimation risk of a fixed arm with known shape and unknown amplitude; it is neither an adaptation theorem for unknown r nor a regret lower bound for an adaptive multi-arm policy. At a boundary point with $\beta_0 = 0$, reversing the sign of $\mathbf{q}_n$ still gives a one-sided two-point bound, but the symmetric constant-one local minimax statement is made only at interior points.

B General-Shape Schur Geometry

This appendix replaces the power transient s^{-r} of Section 3 by an arbitrary deterministic shape $\phi : \{1, 2, \ldots\} \to \mathbb{R}$, so that a single arm produces $Y_s = \mu + \beta\phi(s) + \varepsilon_s$. With $z_s = (1, \phi(s))^\top$, the quantities $a_n = \sum_{s \le n} \phi(s)$, $b_n = \sum_{s \le n} \phi(s)^2$, Q_n, $Q_{n,\lambda} = Q_n + \lambda I_2$, and the ridge leverage $q_{n,\lambda} = e_1^\top Q_{n,\lambda}^{-1} e_1$ are defined exactly as in (3)–(5). The proposition below characterizes when the ridge intercept keeps the parametric rate $q_{n,\lambda} = \Theta(n^{-1})$; the two examples that follow exhibit a logarithmic shape with a $(\log n)^2$ inflation and a monotone shape whose leverage decays arbitrarily slowly.

Proposition 2 (General-shape Schur geometry)**.** *Let $b_n^\circ = b_n - a_n^2/n$ be the centered energy of the shape. Then*

$$q_{n,\lambda}^{-1} = \lambda + n\frac{\lambda + b_n^\circ}{\lambda + b_n} = \lambda + n - \frac{a_n^2}{\lambda + b_n}. \tag{83}$$

If $\phi(n) \to 0$, then $q_{n,\lambda} \to 0$. Moreover, $q_{n,\lambda} = \Theta(n^{-1})$ if and only if

$$\liminf_{n\to\infty} \frac{\lambda + b_n^\circ}{\lambda + b_n} > 0. \tag{84}$$

Proof. Write $\bar{\phi}_n := a_n/n$. Expanding around the empirical mean of the shape values gives

$$b_n^\circ = \sum_{s=1}^{n} \big(\phi(s) - \bar{\phi}_n\big)^2 = b_n - \frac{a_n^2}{n} \ge 0. \tag{85}$$

Since

$$Q_{n,\lambda} = \begin{pmatrix} \lambda + n & a_n \\ a_n & \lambda + b_n \end{pmatrix},$$

direct inversion yields

$$\begin{aligned} q_{n,\lambda} &= \frac{\lambda + b_n}{(\lambda + n)(\lambda + b_n) - a_n^2} \\ &= \frac{1}{\lambda + n - a_n^2/(\lambda + b_n)} = \frac{1}{\lambda + n(\lambda + b_n^\circ)/(\lambda + b_n)}. \end{aligned} \tag{86}$$

This proves the identities in (83). Because $0 \le b_n^\circ \le b_n$, it also shows

$$\frac{1}{\lambda + n} \le q_{n,\lambda} \le \frac{1}{\lambda}. \tag{87}$$

Suppose that $\phi(n) \to 0$. Cesàro convergence applied to $\phi(s)^2$ gives $b_n/n \to 0$. From (86),

$$q_{n,\lambda}^{-1} \ge \lambda + \frac{\lambda n}{\lambda + b_n} \longrightarrow \infty, \tag{88}$$

so $q_{n,\lambda} \to 0$. Neither monotonicity nor a sign restriction on ϕ is needed for this implication.

This leverage limit also gives pointwise consistency of the ridge intercept for every fixed θ. Since $0 \preceq Q_n \preceq Q_{n,\lambda}$,

$$\mathrm{Var}_\theta(\widehat{\mu}_{n,\lambda}) = \sigma^2 e_1^\top Q_{n,\lambda}^{-1} Q_n Q_{n,\lambda}^{-1} e_1 \le \sigma^2 q_{n,\lambda} \longrightarrow 0.$$

Moreover,

$$\mathbb{E}_\theta \widehat{\mu}_{n,\lambda} - \mu = -\lambda e_1^\top Q_{n,\lambda}^{-1} \theta.$$

The first coordinate of $e_1^\top Q_{n,\lambda}^{-1}$ is $q_{n,\lambda}$, and positive definiteness of $Q_{n,\lambda}^{-1} \preceq \lambda^{-1} I$ gives

$$|(Q_{n,\lambda}^{-1})_{12}|^2 \le (Q_{n,\lambda}^{-1})_{11}(Q_{n,\lambda}^{-1})_{22} \le q_{n,\lambda}/\lambda.$$

Thus the bias also vanishes for every fixed θ.

Finally, divide the inverse leverage in (86) by n:

$$\frac{q_{n,\lambda}^{-1}}{n} = \frac{\lambda}{n} + \frac{\lambda + b_n^\circ}{\lambda + b_n}. \tag{89}$$

The lower bound in (87) already implies $q_{n,\lambda} = \Omega(n^{-1})$. Equation (89) shows that $q_{n,\lambda} = O(n^{-1})$ if and only if $(\lambda + b_n^\circ)/(\lambda + b_n)$ is eventually bounded away from zero, which is precisely (84). This proves the claimed equivalence. □

The logarithmic shape. We show that a logarithmic transient inflates the leverage by a $(\log n)^2$ factor. Fix $\upsilon > 0$ and put $L := \log n$. Monotone sum–integral comparison, followed by repeated integration by parts, gives

$$\sum_{s=1}^{n} [\log(s+1)]^{-\upsilon} = nL^{-\upsilon}\left(1 + \frac{\upsilon}{L} + \frac{\upsilon(\upsilon+1)}{L^2} + O(L^{-3})\right). \tag{90}$$

Indeed, the sum differs by $O(1)$ from $\int_2^{n+1} (\log x)^{-\upsilon}\, dx$, and iterating

$$\int_2^x (\log t)^{-\upsilon}\, dt = x(\log x)^{-\upsilon} + \upsilon \int_2^x (\log t)^{-\upsilon-1}\, dt + O(1)$$

proves (90); replacing $n+1$ by n does not change the displayed remainder.

Now take $\phi(s) = [\log(s+1)]^{-\nu}$ with fixed $\nu > 0$, and apply (90) with $\upsilon = \nu$ and $\upsilon = 2\nu$. Expanding the ratio gives

$$\frac{a_n^2}{n b_n} = 1 - \frac{\nu^2}{L^2} + O(L^{-3}). \tag{91}$$

Moreover, $b_n \sim nL^{-2\nu}$, and hence

$$\frac{b_n}{\lambda + b_n} = 1 + o(L^{-3}).$$

Combining this fact with (91) yields

$$\frac{q_{n,\lambda}^{-1}}{n} = \frac{\lambda}{n} + 1 - \frac{a_n^2}{n(\lambda + b_n)} = \frac{\nu^2}{L^2}\big(1 + o(1)\big). \tag{92}$$

Therefore

$$q_{n,\lambda} \sim \frac{(\log n)^2}{\nu^2 n}. \tag{93}$$

Arbitrarily slow positive monotone decay. Consistency alone carries no rate: the leverage can decay arbitrarily slowly even for a positive nonincreasing shape. For every positive deterministic sequence $t_n \to 0$, there is a positive nonincreasing shape $\phi(n) \to 0$ for which $q_{n,\lambda} \to 0$, but

$$\limsup_{n\to\infty} \frac{q_{n,\lambda}}{t_n} = \infty. \tag{94}$$

To prove this, first note the variational form of the Schur complement:

$$q_{n,\lambda}^{-1} = \lambda + \min_{\zeta \in \mathbb{R}} \left\{ \lambda\zeta^2 + \sum_{s=1}^{n} \big(1 - \zeta\phi(s)\big)^2 \right\}. \tag{95}$$

It follows by expanding the quadratic in ζ, whose minimizer is $a_n/(\lambda + b_n)$.

Construct consecutive finite blocks. On block k, set $\phi(s) = 1/k$. Suppose blocks $1, \dots, k-1$ have already been chosen and their last index is n_{k-1}, with $\mathsf{n}_0 := 0$. Define the finite constant

$$\Xi_k := \lambda + \lambda k^2 + \sum_{s=1}^{\mathsf{n}_{k-1}} \big(1 - k\phi(s)\big)^2. \tag{96}$$

For any n in block k, choosing $\zeta = k$ in (95) makes every contribution from the new block equal to zero. Consequently,

$$q_{n,\lambda}^{-1} \le \Xi_k, \qquad q_{n,\lambda} \ge \Xi_k^{-1}. \tag{97}$$

Because $t_n \to 0$, we may choose a finite endpoint $\mathsf{n}_k > \mathsf{n}_{k-1}$ so large that

$$t_{\mathsf{n}_k} \le \frac{1}{k\Xi_k}. \tag{98}$$

Then $q_{\mathsf{n}_k,\lambda}/t_{\mathsf{n}_k} \ge k$, proving (94). The resulting block sequence is positive, nonincreasing, and tends to zero. Proposition 2 therefore still ensures $q_{n,\lambda} \to 0$, completing the construction.

C Proof of the Deterministic Two-Arm Benchmark

This appendix proves Proposition 1 and derives the corresponding local inverse-budget constants. Throughout, fix $r > 0$, abbreviate $x_s = s^{-r}$, and recall from (3)–(5) that

$$a_n = \sum_{s=1}^{n} x_s, \qquad b_n = \sum_{s=1}^{n} x_s^2, \qquad D_n = n - \frac{a_n^2}{b_n}.$$

Because $x_1, \dots, x_n$ are not constant when $n \ge 2$, the unregularized Gram matrix

$$Q_n = \sum_{s=1}^{n} \begin{pmatrix} 1 \\ x_s \end{pmatrix} (1 \quad x_s) = \begin{pmatrix} n & a_n \\ a_n & b_n \end{pmatrix}$$

is positive definite and $D_n > 0$. Direct inversion gives

$$e_1^\top Q_n^{-1} e_1 = \frac{b_n}{n b_n - a_n^2} = \frac{1}{D_n}. \tag{99}$$

C.1 Design limits and uniform feasibility

Lemma 2 (Design limits)**.** *For every fixed $r > 0$,*

$$\frac{D_n}{n} \longrightarrow \frac{1}{\kappa(r)} \quad \textit{and} \quad \frac{a_n}{b_n} = o(\sqrt{n}). \tag{100}$$

Moreover,

$$\max_{2 \le n \le T} \frac{a_n/b_n}{\sqrt{T}} \longrightarrow 0. \tag{101}$$

Proof. For $0 < r < 1/2$, the power-sum asymptotics are

$$a_n \sim \frac{n^{1-r}}{1-r}, \qquad b_n \sim \frac{n^{1-2r}}{1-2r}.$$

It follows that

$$\frac{D_n}{n} = 1 - \frac{a_n^2}{n b_n} \longrightarrow 1 - \frac{1-2r}{(1-r)^2} = \frac{r^2}{(1-r)^2} = \frac{1}{\kappa(r)}.$$

At $r = 1/2$, $a_n = 2\sqrt{n} + O(1)$ and $b_n = \log n + O(1)$, so $a_n^2/(n b_n) \to 0$. If $r > 1/2$, then b_n converges to a finite positive limit and $a_n^2/n \to 0$: use $a_n = O(n^{1-r})$ for $1/2 < r < 1$, $a_n = O(\log n)$ for $r = 1$, and $a_n = O(1)$ for $r > 1$. This proves the first limit in (100).

The same estimates give

$$\frac{a_n}{b_n} = \begin{cases} \dfrac{1-2r}{1-r} n^r (1+o(1)), & 0 < r < 1/2, \\ \dfrac{2\sqrt{n}}{\log n}(1+o(1)), & r = 1/2, \\ O(n^{1-r}), & 1/2 < r < 1, \\ O(\log n), & r = 1, \\ O(1), & r > 1. \end{cases}$$

Every line is $o(\sqrt{n})$. For the triangular-array statement, fix $\epsilon > 0$ and choose n_0 such that $a_n/b_n \le \epsilon\sqrt{n}$ for $n \ge n_0$. The contribution from $n_0 \le n \le T$ to the maximum in (101) is at most ϵ, while the maximum over the finitely many $2 \le n < n_0$ is $O(T^{-1/2})$. Letting $T \to \infty$ and then $\epsilon \downarrow 0$ proves (101). □

C.2 Exact minimax risk for a deterministic allocation

For a recommendation rule ψ, the composite risk of Section 4 is

$$\mathcal{R}_T(\psi; n_1, n_2, \Delta) := \max\Big\{ \sup_{(\theta_1,\theta_2)\in\mathcal{P}_+(\Delta)} \mathbb{P}_\theta(\psi = 2), \\ \sup_{(\theta_1,\theta_2)\in\mathcal{P}_-(\Delta)} \mathbb{P}_\theta(\psi = 1)\Big\}. \tag{102}$$

where $\mathcal{P}_+(\Delta)$ and $\mathcal{P}_-(\Delta)$ are the product-ball classes defined before Proposition 1.

Lemma 3 (Exact deterministic-allocation identity)**.** *Fix $n_1, n_2 \ge 2$ and let*

$$J = J(n_1, n_2) = \left(D_{n_1}^{-1} + D_{n_2}^{-1}\right)^{-1}.$$

Whenever the two least-favorable parameter points constructed in the proof belong to the local product ball,

$$\inf_{\psi} \mathcal{R}_T(\psi; n_1, n_2, \Delta) = \Phi\left(-\frac{\Delta\sqrt{J}}{\sigma}\right). \tag{103}$$

For every fixed $c > 0$ and $\Delta_T = c/\sqrt{T}$, this identity holds for all sufficiently large T, uniformly over all deterministic $n_1, n_2 \ge 2$ with $n_1 + n_2 = T$.

Proof. Let

$$\widehat{\theta}_i = Q_{n_i}^{-1} \sum_{s=1}^{n_i} \begin{pmatrix} 1 \\ x_s \end{pmatrix} Y_{i,s}, \qquad \widehat{\mu}_i = e_1^\top \widehat{\theta}_i.$$

The two arms are independent, and (99) gives

$$\widehat{\mu}_i \sim N\left(\mu_i, \frac{\sigma^2}{D_{n_i}}\right).$$

Thus, with $\Delta_{12} = \mu_2 - \mu_1$ the true signed gap,

$$\widehat{\Delta} := \widehat{\mu}_2 - \widehat{\mu}_1 \sim N\left(\Delta_{12}, \frac{\sigma^2}{J}\right). \tag{104}$$

The rule that recommends arm 1 when $\widehat{\Delta} > 0$ and arm 2 otherwise has worst error $\Phi(-\Delta\sqrt{J}/\sigma)$ over each separated class. This proves the upper bound in (103).

For the matching lower bound, abbreviate

$$D_i = D_{n_i}, \qquad a_i = a_{n_i}, \qquad b_i = b_{n_i},$$

and set

$$\delta_1^\mu = \frac{2\Delta D_2}{D_1 + D_2}, \qquad \delta_2^\mu = -\frac{2\Delta D_1}{D_1 + D_2}, \qquad \delta_i^\beta = -\delta_i^\mu \frac{a_i}{b_i}. \tag{105}$$

Center two parameter pairs at the interior point $\theta_0 = (\mu_0, \beta_0)$:

$$\mu_i^+ = \mu_0 - \frac{\delta_i^\mu}{2}, \quad \beta_i^+ = \beta_0 - \frac{\delta_i^\beta}{2}, \qquad \mu_i^- = \mu_0 + \frac{\delta_i^\mu}{2}, \quad \beta_i^- = \beta_0 + \frac{\delta_i^\beta}{2}. \tag{106}$$

Since $\delta_1^\mu - \delta_2^\mu = 2\Delta$, the plus point has $\mu_2^+ - \mu_1^+ = \Delta$, while the minus point has $\mu_1^- - \mu_2^- = \Delta$.

Let P_+ and P_- be the joint observation laws at these two points. On arm i, the difference between the minus and plus mean vectors at local time s is

$$\delta_i^\mu + \delta_i^\beta x_s = \delta_i^\mu \left(1 - \frac{a_i}{b_i} x_s\right).$$

Its squared norm is

$$\sum_{s=1}^{n_i} (\delta_i^\mu + \delta_i^\beta x_s)^2 = (\delta_i^\mu)^2 \left(n_i - \frac{a_i^2}{b_i}\right) = (\delta_i^\mu)^2 D_i.$$

Therefore the squared Mahalanobis distance between P_+ and P_- is

$$\begin{aligned} d_{\mathrm{Mah}}^2 &= \frac{(\delta_1^\mu)^2 D_1 + (\delta_2^\mu)^2 D_2}{\sigma^2} \\ &= \frac{4\Delta^2}{\sigma^2} \frac{D_1 D_2}{D_1 + D_2} = \frac{4\Delta^2 J}{\sigma^2}. \end{aligned} \tag{107}$$

The equal-prior Bayes error for two Gaussian laws with common covariance and Mahalanobis distance d_{Mah} is $\Phi(-d_{\mathrm{Mah}}/2)$. Hence every recommendation obeys

$$\begin{aligned} \mathcal{R}_T(\psi; n_1, n_2, \Delta) &\geq \max\{P_+(\psi = 2), P_-(\psi = 1)\} \\ &\geq \frac{1}{2}\{P_+(\psi = 2) + P_-(\psi = 1)\} \\ &\geq \Phi(-d_{\mathrm{Mah}}/2) = \Phi\left(-\frac{\Delta\sqrt{J}}{\sigma}\right). \end{aligned}$$

This matches the OLS upper bound whenever the two points in (106) are feasible.

It remains to check feasibility uniformly at $\Delta_T = c/\sqrt{T}$. From (105),

$$\frac{|\delta_i^\mu|}{2} \leq \Delta_T, \qquad \frac{|\delta_i^\beta|}{2} \leq \Delta_T \max_{2 \leq n \leq T} \frac{a_n}{b_n}.$$

The first quantity tends to zero, and the second is $o(1)$ by Lemma 2. Thus both per-arm perturbation vectors in (106) converge uniformly to zero. Because the closed ball $B(\theta_0, \rho)$ lies in the interior of Θ_S, both least-favorable pairs belong to the required product ball for all sufficiently large T, uniformly over the allocation. □

C.3 Optimizing the deterministic allocation

Lemma 4 (Optimal value and asymptotic balance)**.** *Let*

$$J_T^* := \max_{\substack{n_1 + n_2 = T \\ n_1, n_2 \geq 2}} J(n_1, n_2).$$

Then

$$\frac{J_T^*}{T} \longrightarrow \frac{1}{4\kappa(r)}. \tag{108}$$

Moreover, if an allocation sequence satisfies

$$\frac{J(n_{1,T}, n_{2,T})}{J_T^*} \longrightarrow 1, \tag{109}$$

then

$$\frac{n_{1,T}}{T} \longrightarrow \frac{1}{2}, \qquad \frac{n_{2,T}}{T} \longrightarrow \frac{1}{2}.$$

Proof. For $n_1 = \lfloor T/2 \rfloor$ and $n_2 = \lceil T/2 \rceil$, Lemma 2 gives

$$\frac{D_{n_1}}{T} \longrightarrow \frac{1}{2\kappa(r)}, \qquad \frac{D_{n_2}}{T} \longrightarrow \frac{1}{2\kappa(r)},$$

and hence $J(n_1, n_2)/T \to 1/(4\kappa(r))$.

Conversely, take any sequence of allocations and pass to a subsequence along which $n_{1,T}/T \to p \in [0, 1]$. Since $0 < D_n \le n$, if $p \in \{0, 1\}$, then

$$\frac{J(n_{1,T}, n_{2,T})}{T} \le \frac{\min(D_{n_{1,T}}, D_{n_{2,T}})}{T} \le \frac{\min(n_{1,T}, n_{2,T})}{T} \longrightarrow 0.$$

If $p \in (0, 1)$, both counts diverge, and Lemma 2 yields

$$\frac{J(n_{1,T}, n_{2,T})}{T} \longrightarrow \frac{p(1-p)}{\kappa(r)} \le \frac{1}{4\kappa(r)}.$$

Together with the balanced lower bound, this proves (108).

If (109) holds, then

$$\frac{J(n_{1,T}, n_{2,T})}{T} \longrightarrow \frac{1}{4\kappa(r)}.$$

Every convergent subsequence of $n_{1,T}/T$ must therefore be interior and satisfy $p(1-p) = 1/4$, whose unique solution in $[0, 1]$ is $p = 1/2$. Hence the full sequence converges to $1/2$, as does its complementary allocation fraction. □

Lemma 3 and Lemma 4, with $\Delta_T = c/\sqrt{T}$, give

$$\min_{\substack{n_1+n_2=T \\ n_i \ge 2}} \inf_{\psi} \mathcal{R}_T(\psi; n_1, n_2, \Delta_T) \longrightarrow \Phi\left(-\frac{c}{2\sigma\sqrt{\kappa(r)}}\right).$$

This proves the unknown-amplitude statements of Proposition 1. Note that OLS attains the exact minimax identity for every deterministic allocation, and balanced sampling attains the optimized first-order limit.

C.4 Revealed-amplitude class

In the revealed-amplitude experiment the parameter classes are the same product-ball classes $\mathcal{P}_{\pm}(\Delta)$, but the realized pair (β_1, β_2) is revealed before sampling. Formally, an oracle rule is a measurable family of rules indexed by the revealed pair, and $\mathcal{R}_T^{\mathrm{or}}$ is the worst error of that family over the same classes.

The corrected sample means satisfy

$$\widetilde{\mu}_i := \frac{1}{n_i} \sum_{s=1}^{n_i} (Y_{i,s} - \beta_i x_s) \sim N\left(\mu_i, \frac{\sigma^2}{n_i}\right).$$

Thus, with

$$J^{\mathrm{or}}(n_1, n_2) := (n_1^{-1} + n_2^{-1})^{-1} = \frac{n_1 n_2}{T},$$

the sign of $\widetilde{\mu}_2 - \widetilde{\mu}_1$ has worst error $\Phi(-\Delta\sqrt{J^{\mathrm{or}}}/\sigma)$. For the matching lower bound, take

$$\delta_1^{\mu} = \frac{2\Delta n_2}{T}, \qquad \delta_2^{\mu} = -\frac{2\Delta n_1}{T},$$

set $\mu_i^{\pm} = \mu_0 \mp \delta_i^{\mu}/2$, and leave both amplitudes equal to β_0. The two points have opposite gaps of magnitude Δ, and their squared Mahalanobis distance is

$$\frac{(\delta_1^{\mu})^2 n_1 + (\delta_2^{\mu})^2 n_2}{\sigma^2} = \frac{4\Delta^2 J^{\mathrm{or}}(n_1, n_2)}{\sigma^2}.$$

For all sufficiently small Δ, these points belong to the revealed-amplitude classes. The same two-point Bayes argument therefore gives the exact identity

$$\inf_{\psi} \mathcal{R}_T^{\mathrm{or}}(\psi; n_1, n_2, \Delta) = \Phi\left(-\frac{\Delta\sqrt{J^{\mathrm{or}}(n_1, n_2)}}{\sigma}\right). \tag{110}$$

The product $n_1 n_2$ is maximized by the two nearest integers to $T/2$, so

$$\max_{n_1+n_2=T} J^{\mathrm{or}}(n_1, n_2) = \frac{T}{4} + O(T^{-1}).$$

Taking $\Delta_T = c/\sqrt{T}$ in (110) gives the oracle limit $\Phi(-c/(2\sigma))$; the same value reappears as the oracle statement of Theorem 11 in Appendix H.

C.5 Formal inversion to local budgets

Fix $\alpha \in (0, 1/2)$ and write $z_\alpha = \Phi^{-1}(1-\alpha) > 0$. Define

$$T_{\mathrm{un}}(\Delta, \alpha) := \min\Big\{T \geq 4 : \min_{\substack{n_1+n_2=T \\ n_i \geq 2}} \inf_{\psi} \mathcal{R}_T(\psi; n_1, n_2, \Delta) \leq \alpha\Big\}, \tag{111}$$

$$T_{\mathrm{or}}(\Delta, \alpha) := \min\Big\{T \geq 4 : \min_{\substack{n_1+n_2=T \\ n_i \geq 2}} \inf_{\psi} \mathcal{R}_T^{\mathrm{or}}(\psi; n_1, n_2, \Delta) \leq \alpha\Big\}. \tag{112}$$

These are inverse fixed-budget complexities, the least horizons at which an optimized deterministic design reaches error α; they are not stopping times.

Corollary 1 (Local inverse-budget constants)**.** *For every fixed $\alpha \in (0, 1/2)$,*

$$T_{\mathrm{un}}(\Delta, \alpha) \sim \frac{4\sigma^2 \kappa(r) z_\alpha^2}{\Delta^2}, \tag{113}$$

$$T_{\mathrm{or}}(\Delta, \alpha) \sim \frac{4\sigma^2 z_\alpha^2}{\Delta^2}, \tag{114}$$

as $\Delta \downarrow 0$. Consequently,

$$\frac{T_{\mathrm{un}}(\Delta, \alpha)}{T_{\mathrm{or}}(\Delta, \alpha)} \longrightarrow \kappa(r).$$

Moreover, in the iterated local-then-high-confidence limit,

$$\lim_{\alpha\downarrow 0} \lim_{\Delta\downarrow 0} \frac{T_{\mathrm{un}}(\Delta, \alpha)\Delta^2}{8\sigma^2 \kappa(r) \log(1/\alpha)} = 1, \tag{115}$$

$$\lim_{\alpha\downarrow 0} \lim_{\Delta\downarrow 0} \frac{T_{\mathrm{or}}(\Delta, \alpha)\Delta^2}{8\sigma^2 \log(1/\alpha)} = 1. \tag{116}$$

Proof. First, both minimizing budgets are of order Δ^{-2}. For a lower bound, restrict the unknown-amplitude class to fixed equal amplitudes. For any allocation, the resulting ordinary two-sample Gaussian subproblem has effective information at most

$$\frac{n_1 n_2}{T} \leq \frac{T}{4}.$$

Thus achieving worst error at most α requires

$$T \geq \frac{4\sigma^2 z_\alpha^2}{\Delta^2}.$$

The same lower bound applies in the revealed-amplitude model. Conversely, balanced OLS has $J(\lfloor T/2 \rfloor, \lceil T/2 \rceil)/T \to 1/(4\kappa(r))$, so taking $T = \mathsf{C}/\Delta^2$ with a sufficiently large constant C gives error below α. Balanced corrected means give the analogous oracle upper bound. Hence, for constants $0 < \mathsf{C}_1 < \mathsf{C}_2 < \infty$ independent of small Δ, both minimizing budgets lie in $[\mathsf{C}_1/\Delta^2, \mathsf{C}_2/\Delta^2]$.

The least-favorable pairs used in Lemma 3 are uniformly feasible throughout this budget range. Indeed,

$$\Delta \max_{2\le n\le \mathsf{C}_2/\Delta^2} \frac{a_n}{b_n} = \sqrt{\mathsf{C}_2}\left[\max_{2\le n\le \mathsf{C}_2/\Delta^2} \frac{a_n/b_n}{\sqrt{\mathsf{C}_2/\Delta^2}}\right] \longrightarrow 0$$

by (101). Thus the exact identity

$$\min_{n_1+n_2=T} \inf_{\psi} \mathcal{R}_T(\psi; n_1, n_2, \Delta) = \Phi\left(-\frac{\Delta\sqrt{J_T^*}}{\sigma}\right)$$

holds uniformly over all relevant T.

Now fix $\epsilon \in (0,1)$ and compare the two budgets

$$T_\pm(\Delta) = (1\pm\epsilon)\frac{4\sigma^2\kappa(r)z_\alpha^2}{\Delta^2},$$

rounded to admissible integers. Since $J_T^*/T \to 1/(4\kappa(r))$, uniformly once $T \ge \mathsf{C}_1/\Delta^2$, the risk at $T_+(\Delta)$ is eventually below α, while no $T \le T_-(\Delta)$ in the relevant range can reach α. The coarser Gaussian-submodel lower bound already excludes smaller T. Letting $\epsilon \downarrow 0$ proves (113). Equation (110) and $\max J^{\mathrm{or}} = T/4 + O(T^{-1})$ give (114) by the same, simpler argument. Their ratio is the constant $\kappa(r)$ that also appears in (250) of Appendix H. Finally, $z_\alpha^2 \sim 2\log(1/\alpha)$ as $\alpha \downarrow 0$, which gives the two iterated limits. The order of limits is essential. □

D Adaptive Two-Arm Problem: Ancillary Optimum, Completion Lemma, and Fixed-Pair Envelope

This appendix collects the tools behind Theorem 2 and two results that delimit its scope. It first solves the two-arm problem exactly within the class of policies whose limiting allocation is ancillary (Theorem 4). It then proves the Gaussian-bracket completion lemma (Lemma 7), the testing device on which the unrestricted lower bound rests, and shows by an explicit fixed-pair envelope (Theorem 5) why a single least-favorable pair cannot close the unrestricted value; the diffuse-prior argument that does close it appears in Appendix E. Finally, it records why the marginal law of the terminal allocation does not determine the risk of a policy. Throughout, $0 < r < 1/2$ and $\tau_T \le CT$ pathwise.

D.1 Model and local minimax problem

There are two arms, and the sth local observation from arm i is

$$Y_{i,s} = \mu_i + \beta_i s^{-r} + \varepsilon_{i,s}, \qquad \varepsilon_{i,s} \overset{\text{iid}}{\sim} N(0,\sigma^2), \qquad 0 < r < \frac{1}{2}. \tag{117}$$

At each global round a parameter-free predictable kernel chooses an arm and reveals that arm's next unused observation; independent internal randomization is included in the initial σ-field $\mathcal{F}_0$. The stopping time τ_T obeys

$$\tau_T \le CT \quad \text{almost surely} \tag{118}$$

for a fixed $C > 0$. Write $N_i = N_i(\tau_T)$ and $P_{i,T} = N_i/T$.

Fix an interior equal-arm point (μ_0, β_0) and use local coordinates

$$\mu_{i,T} = \mu_0 + \frac{u_i}{\sqrt{T}}, \qquad \beta_{i,T} = \beta_0 + v_i T^{r-1/2}, \qquad \gamma_i = (u_i, v_i)^\top. \tag{119}$$

For $M, c > 0$, let

$$\begin{aligned}\Gamma_{+,M}(c) &= \{\gamma \in [-M,M]^4 : u_2 - u_1 \ge c\},\\ \Gamma_{-,M}(c) &= \{\gamma \in [-M,M]^4 : u_1 - u_2 \ge c\}.\end{aligned} \tag{120}$$

Arm 1 is correct on $\Gamma_{+,M}(c)$ and arm 2 on $\Gamma_{-,M}(c)$. For a policy and terminal recommendation ψ_T, define

$$\mathcal{R}_{T,M}(\pi,\psi;c) = \max\left\{\sup_{\gamma\in\Gamma_{+,M}(c)} \mathbb{P}_\gamma(\psi_T = 2), \sup_{\gamma\in\Gamma_{-,M}(c)} \mathbb{P}_\gamma(\psi_T = 1)\right\}. \tag{121}$$

For every fixed cube, all of these local points lie in the original interior parameter set once T is large enough.

For $p \geq 0$, put

$$g_r(x) = (1, x^{-r})^\top,$$
$$Q_r(p) = \int_0^p g_r(x) g_r(x)^\top dx = \begin{pmatrix} p & A_r(p) \\ A_r(p) & B_r(p) \end{pmatrix}, \tag{122}$$

where

$$A_r(p) = \frac{p^{1-r}}{1-r}, \qquad B_r(p) = \frac{p^{1-2r}}{1-2r}. \tag{123}$$

The intercept Schur complement is

$$D_r(p) = p - \frac{A_r(p)^2}{B_r(p)} = \frac{p}{\kappa(r)}, \qquad \kappa(r) = \frac{(1-r)^2}{r^2}. \tag{124}$$

We use $D_r(0) = 0$.

D.2 Exact optimization over the allocation-ancillary class

Under the central law define the normalized terminal score

$$X_{i,T} = \frac{1}{\sqrt{T}} \sum_{s=1}^{N_i} \varepsilon_{i,s} g_r(s/T), \qquad \mathcal{Q}_r(P) = \mathrm{diag}\{Q_r(P_1), Q_r(P_2)\}. \tag{125}$$

Every subsequential central limit $(P_T, X_T) \Rightarrow (P, X)$ has local density

$$L_\gamma(P, X) = \exp\left\{ \frac{\gamma^\top X}{\sigma^2} - \frac{\gamma^\top \mathcal{Q}_r(P)\gamma}{2\sigma^2} \right\}. \tag{126}$$

This follows from the exact Gaussian likelihood, the cancellation of the predictable kernels, the uniform prefix Riemann sums, and the exponential-martingale moments available under bounded stopping (Theorem 14). The following definition isolates the class of policies whose allocation carries no local parameter signal.

Definition 1 (Full local allocation ancillarity)**.** A joint limit (P, X) is *full-local-allocation-ancillary* if the marginal law of P under the experiment (126) is the same as its central law for every $\gamma \in \mathbb{R}^4$. A policy sequence is full-local-allocation-ancillary if every one of its joint subsequential limits is so.

The word "full" means that ancillarity is required in the target intercept direction as well as in every nuisance direction; the allocation and the horizon may nevertheless be data-dependent and random. This is exactly the class in which the limiting terminal allocation carries no signal under any fixed local intercept-or-amplitude direction. We do not claim that it is the largest regular class under other notions of regularity.

Lemma 5 (Ancillarity is exactly conditional Gaussianity)**.** *Let $\nu = \mathrm{law}(P)$ and let $\mathcal{L}_p$ be a regular conditional law of X given $P = p$. The limit is full-local-allocation-ancillary if and only if*

$$X \mid P = p \sim N\{0, \sigma^2 \mathcal{Q}_r(p)\} \quad \textit{for } \nu\textit{-almost every } p. \tag{127}$$

The statement includes singular $Q_r(0)$ in the usual degenerate-normal sense.

Proof. Under local parameter γ, the P marginal is tilted relative to ν by

$$\mathsf{w}_\gamma(p) = e^{-\gamma^\top \mathcal{Q}_r(p)\gamma/(2\sigma^2)} \int e^{\gamma^\top x/\sigma^2} \mathcal{L}_p(dx). \tag{128}$$

Conditional normality makes the integral the inverse of the prefactor, so $\mathsf{w}_\gamma = 1$. Conversely, full local allocation ancillarity gives $\mathsf{w}_\gamma(p) = 1$ for every γ, initially on a γ-dependent full-measure set. Intersect the sets for a countable dense collection of γ's. The bounded pathwise bracket in (118) gives a common local exponential moment for X; hence the conditional moment-generating function is finite and continuous on a neighborhood of zero for almost every p. The identity extends from the dense collection to that neighborhood. Uniqueness of moment-generating functions proves (127). □

Define, for $p = (p_1, p_2)$,

$$I_r(p) = \begin{cases} \dfrac{p_1 p_2}{\kappa(r)(p_1 + p_2)}, & p_1 + p_2 > 0, \\ 0, & p_1 = p_2 = 0. \end{cases} \tag{129}$$

Theorem 4 (Sharp adaptive allocation in the allocation-ancillary class)**.** *Suppose that a joint limit is full-local-allocation-ancillary and that $P_1 + P_2 \le C$ almost surely. Its local asymptotic minimax misidentification probability is*

$$R_{\mathrm{reg}}(c; \nu) = \mathbb{E}_\nu \Phi\left(-\frac{c\sqrt{I_r(P)}}{\sigma} \right). \tag{130}$$

More precisely, every recommendation sequence obeys

$$\lim_{M \to \infty} \liminf_{T \to \infty} \mathcal{R}_{T,M}(\pi, \psi; c) \ge R_{\mathrm{reg}}(c; \nu) \tag{131}$$

along the corresponding subsequence, and the observable gated randomized OLS rule constructed in the proof attains equality. Consequently, over all full-local-allocation-ancillary predictable policies and pathwise-capped stopping rules,

$$\inf_{\text{allocation-ancillary}} R_{\mathrm{reg}}(c; \nu) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma\sqrt{\kappa(r)}} \right). \tag{132}$$

For $c > 0$, equality in (132) *holds if and only if*

$$P_1 = P_2 = C/2 \quad \textit{almost surely}. \tag{133}$$

Thus deterministic full-budget balance is the unique first-order information optimum in this class. With revealed amplitudes the same conclusions hold after $\kappa(r)$ is replaced by one.

Proof. Fix $p_i > 0$. Conditional on $P = p$, Lemma 5 gives two independent Gaussian shift blocks

$$X_i \sim N\{Q_r(p_i)(u_i, v_i)^\top, \sigma^2 Q_r(p_i)\}. \tag{134}$$

Project the intercept score off the amplitude score:

$$X_i^\perp = X_{i,1} - \frac{A_r(p_i)}{B_r(p_i)} X_{i,2}, \qquad Z_i = \frac{X_i^\perp}{D_r(p_i)}. \tag{135}$$

By (124),

$$Z_i \mid P = p \sim N\left(u_i, \frac{\sigma^2 \kappa(r)}{p_i} \right), \tag{136}$$

independently of the amplitude-score coordinate. Hence

$$Z_2 - Z_1 \mid P = p \sim N\left(u_2 - u_1, \frac{\sigma^2}{I_r(p)} \right). \tag{137}$$

When an arm has zero limiting information, interpret the right side as a completely uninformative contrast. Recommending according to the sign of $Z_2 - Z_1$, with an independent fair tie decision in the zero-information case, has conditional worst error $\Phi\{-c\sqrt{I_r(p)}/\sigma\}$. Full local allocation ancillarity makes the mixing law ν parameter-independent, so integration proves the upper bound in (130).

For completeness, we give the composite lower bound rather than appealing to invariance. On the boundary write

$$\gamma = \mathsf{s} d + G\xi, \qquad d = \frac{c}{2}(-1, 0, 1, 0)^\top, \qquad G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{138}$$

where $\mathsf{s} \in \{-1, 1\}$ is the correct-label sign and ξ contains the common intercept and the two amplitudes. Give s equal prior mass and let $\xi \sim N(0, a^2 I_3)$ independently. First suppose $p_1, p_2 > 0$ and, conditional on $P = p$, whiten X by $\widetilde{X} = \mathcal{Q}_r(p)^{-1/2} X$. After integrating ξ,

$$\widetilde{X} \mid (P = p, \mathsf{s}) \sim N\{\mathsf{s}\mathcal{Q}_r(p)^{1/2} d, \Sigma_a(p)\}, \quad \Sigma_a(p) = \sigma^2 I + a^2 \mathcal{Q}_r(p)^{1/2} G G^\top \mathcal{Q}_r(p)^{1/2}. \tag{139}$$

The equal-prior conditional Bayes error is

$$\Phi\{-\sqrt{J_a(p)}\}, \quad J_a(p) = d^\top \mathcal{Q}_r(p)^{1/2} \Sigma_a(p)^{-1} \mathcal{Q}_r(p)^{1/2} d. \tag{140}$$

As $a \to \infty$, projection first off the two amplitude scores and then off the common intercept gives

$$J_a(p) \longrightarrow \frac{c^2}{\sigma^2} \frac{D_r(p_1) D_r(p_2)}{D_r(p_1) + D_r(p_2)} = \frac{c^2 I_r(p)}{\sigma^2}. \tag{141}$$

If a block is absent, work on the range of $\mathcal{Q}_r(p)$, or equivalently let that p_i decrease to zero in the displayed Gaussian experiment. The limit in (141) is then zero. Dominated convergence gives the Bayes lower bound (130) in the limit experiment.

To transfer this bound to the finite experiments, first fix a, truncate the Gaussian nuisance prior to a large cube, and then quantize it on a finite grid. Truncation changes Bayes risk by at most the discarded prior mass. On a compact cube, Gaussian total variation is uniformly continuous in the mean, so grid Bayes risks converge to the truncated-prior risk. For each finite grid the exact Gaussian likelihood and the bounded-bracket exponential-martingale bound give joint convergence and uniform integrability of its likelihood-ratio vector. Its Bayes error therefore converges. Taking grid mesh to zero, cube size to infinity, and finally $a \to \infty$ proves (131). This also shows why the order $T \to \infty$ before $M \to \infty$ is sufficient.

We next lift the attaining decision from the limit experiment to one observable finite-experiment sequence. Whenever arm i's two-column prefix design is nonsingular, let $\widehat{\mu}_{i,T}$ be its full-prefix OLS intercept and put

$$\widehat{\Delta}_T = \sqrt{T}(\widehat{\mu}_{2,T} - \widehat{\mu}_{1,T}). \tag{142}$$

For every integer $m \geq 1$, choose a continuous gate $w_m : [0, C]^2 \to [0, 1]$ such that

$$w_m(p) = 0 \quad \text{if } \min_i p_i \leq (2m)^{-1}, \qquad w_m(p) = 1 \quad \text{if } \min_i p_i \geq m^{-1}. \tag{143}$$

For example, apply a continuous piecewise-linear cutoff to $m \min_i p_i$. Given the observations and an independent terminal uniform randomizer, recommend arm 1 with probability

$$\mathfrak{q}_{T,m} = \frac{1 - w_m(P_T)}{2} + w_m(P_T) \Phi\{m \widehat{\Delta}_T\}. \tag{144}$$

If either prefix design is singular, set $\mathfrak{q}_{T,m} = 1/2$. This rule is fully observable and does not use any unknown parameter.

For fixed m, the gate removes a neighborhood of the singular set. On its support the discrete OLS contrast is a uniformly continuous function of the finite score and normalized prefix Gram matrices, and those matrices converge uniformly to $Q_r(P_i)$. Thus (144) is, up to a uniformly vanishing approximation, a bounded continuous randomization of (P_T, X_T). Joint likelihood-ratio convergence and the uniform exponential-moment bound therefore make its risk converge uniformly on every compact local parameter cube.

In the limit experiment, for each fixed p with $p_1 p_2 > 0$, first $w_m(p) \to 1$ and then $\Phi\{m(Z_2 - Z_1)\} \to \mathbf{1}\{Z_2 - Z_1 > 0\}$ almost surely as $m \to \infty$. At a boundary gap this gives conditional error $\Phi\{-c\sqrt{I_r(p)}/\sigma\}$. If $p_1 p_2 = 0$, the gate gives error $1/2$, which is the same formula because $I_r(p) = 0$. Larger gap magnitudes only reduce the limiting error. Dominated convergence over P proves attainment of (130) as $m \to \infty$.

Finally choose the smoothing index diagonally. For integer j, fixed-rule compact-uniform convergence supplies a horizon T_j after which the rule with $m = j$ is within $1/j$ of its limit uniformly on the parameter cube of radius j; take the T_j increasing. Set $m_T = \max\{j : T \geq T_j\}$. Then $m_T \uparrow \infty$, and for every fixed compact local cube the single observable rule (144) with $m = m_T$ has risk converging to (130). Taking the cube radius to infinity after $T \to \infty$ proves the asserted

finite-experiment attainment. The gate is essential when $P_i = 0$ has positive probability: an ungated OLS sign can retain selection bias along a sublinear, outcome-dependent prefix.

It remains to optimize. The arithmetic–geometric mean inequality gives

$$I_r(P) = \frac{1}{\kappa(r)} \frac{P_1 P_2}{P_1 + P_2} \le \frac{P_1 + P_2}{4\kappa(r)} \le \frac{C}{4\kappa(r)}. \tag{145}$$

Since $x \mapsto \Phi(-c\sqrt{x}/\sigma)$ is strictly decreasing for $c > 0$, (130) and (145) prove the lower bound in (132). Deterministically sampling each arm through $CT/2+o(T)$ points and using full-prefix OLS attains it. Equality in both inequalities in (145) requires $P_1 + P_2 = C$ and $P_1 = P_2$ almost surely, proving uniqueness. With revealed amplitudes, corrected prefix means have intercept information p_i instead of $p_i/\kappa(r)$, which proves the last assertion. □

Remark 1 (What "unique" does and does not mean)**.** The theorem determines the *limiting information* uniquely: full budget and balanced terminal local times. It does not say that the chronological order of the pulls must be deterministic. Every predictable rule with the same deterministic terminal prefixes has the same first-order value, and an outcome-dependent rule may deviate from balance by $o(T)$.

D.3 A FIXED-PAIR ENVELOPE FOR ARBITRARY ENDOGENOUS POLICIES

We now drop full local allocation ancillarity and no longer assume that P_T converges. This subsection derives a two-sided envelope from a single fixed least-favorable pair. Appendix E closes the global minimax value by mixing over the full nuisance family; the present calculation is kept because it shows why one fixed pair cannot prove that result by itself.

For $\varsigma \in \mathbb{R}$ and $q \in [0, 1]$, define

$$h_{r,\varsigma}(q) = q - \frac{2\varsigma}{1-r} q^{1-r} + \frac{\varsigma^2}{1-2r} q^{1-2r}, \qquad h_{r,\varsigma}(0) = 0, \tag{146}$$

$$F_{r,\varsigma}(q) = h_{r,\varsigma}(q) + h_{r,\varsigma}(1-q), \tag{147}$$

$$J_r = \inf_{\varsigma \in \mathbb{R}} \max_{0 \le q \le 1} F_{r,\varsigma}(q). \tag{148}$$

Equivalently, $h_{r,\varsigma}(q) = \int_0^q (1 - \varsigma x^{-r})^2 dx$.

Lemma 6 (Robust fixed-pair information constant)**.** *The infimum in* (148) *is attained and*

$$\frac{1}{\kappa(r)} < J_r \le 1. \tag{149}$$

Moreover, among all fixed symmetric endpoint separations $g = (g_1, g_2)$, $g_i = (g_{i,1}, g_{i,2})$, satisfying $g_{2,1} - g_{1,1} = 2c$,

$$\inf_{g} \sup_{p_1, p_2 \ge 0:\ p_1 + p_2 \le C} \sum_{i=1}^{2} g_i^\top Q_r(p_i) g_i = c^2 C J_r. \tag{150}$$

One minimizer has

$$g_1 = (-c, c\varsigma_r^\star C^r)^\top, \qquad g_2 = (c, -c\varsigma_r^\star C^r)^\top, \tag{151}$$

where $\varsigma_r^\star$ minimizes (148)*.*

Proof. For fixed coefficients, $g_i^\top Q_r(p_i) g_i$ is the integral of a square and is nondecreasing in p_i. The supremum in (150) therefore occurs on $p_1 + p_2 = C$. The objective is convex in the four coefficients and is invariant under

$$(g_{1,1}, g_{1,2}, g_{2,1}, g_{2,2}) \mapsto (-g_{2,1}, -g_{2,2}, -g_{1,1}, -g_{1,2}). \tag{152}$$

This map preserves $g_{2,1} - g_{1,1} = 2c$. Averaging any candidate with its image under (152) cannot increase the convex objective and gives $g_{1,1} = -c$, $g_{2,1} = c$, $g_{1,2} = -g_{2,2}$. Put $g_{1,2} = c\varsigma C^r$, write $p_1 = Cq, p_2 = C(1-q)$, and substitute $x = Cy$. The objective becomes $c^2 C F_{r,\varsigma}(q)$. This proves (150) and (151). Continuity in (ς, q) and coercivity in ς (already visible at $q = 1/2$) prove attainment.

Taking $\varsigma = 0$ gives $F_{r,0} \equiv 1$, hence $J_r \le 1$. At $q = 1/2$ the unique minimizer over ς is

$$\varsigma_0 = \frac{1-2r}{1-r} 2^{-r}, \qquad F_{r,\varsigma_0}(1/2) = \frac{1}{\kappa(r)}. \tag{153}$$

Thus $J_r \ge 1/\kappa(r)$. Equality would force $\varsigma = \varsigma_0$. But direct substitution gives

$$F_{r,\varsigma_0}(0) - \frac{1}{\kappa(r)} = \frac{1-2r}{(1-r)^2}(1-2^{-r})^2 > 0. \tag{154}$$

Therefore $J_r > 1/\kappa(r)$. □

Reproducible numerical width. The values of J_r reported in Table 2 do not enter any proof. They come from two independently implemented one-dimensional calculations. First, $h'_{r,\varsigma}(q) = (1 - \varsigma q^{-r})^2$, so an interior stationary point of $F_{r,\varsigma}$ away from $1/2$ satisfies

$$\varsigma = \frac{2}{q^{-r} + (1-q)^{-r}}. \tag{155}$$

At an active interior saddle, stationarity in ς also gives

$$\varsigma = \frac{1-2r}{1-r} \frac{q^{1-r} + (1-q)^{1-r}}{q^{1-2r} + (1-q)^{1-2r}}. \tag{156}$$

We solve these equations by 80-digit bisection and evaluate F. Independently, we minimize over ς the maximum over the exhaustive candidates $q = 0, 1/2$ and the roots of (155). The two procedures agree within 2×10^{-16} on the grid below.

Table 2: Numerical fixed-pair envelope constant. The last column is the multiplicative gap between the proved information constants $1/\kappa(r)$ and J_r; it is not an achieved adaptive price.

r	$1/\kappa(r)$	J_r	$\kappa(r)J_r$
.05	.002770	.003984	1.438
.10	.012346	.017704	1.434
.20	.062500	.088304	1.413
.25	.111111	.154701	1.392
.30	.183673	.250079	1.362
.40	.444444	.555556	1.250
.45	.669421	.771103	1.152
.49	.923106	.957257	1.037

The next lemma is the testing device that makes the lower bounds insensitive to the dependence between the score and its bracket created by endogenous allocation and stopping.

Lemma 7 (Completing a random Gaussian score bracket)**.** *Consider a two-point experiment with densities $L_\pm$ relative to a central law under which*

$$L_\pm = \exp\{\pm\mathcal{M}/2 - V/8\}, \qquad \mathbb{E}e^{t\mathcal{M} - t^2 V/2} = 1 \quad (t \in \mathbb{R}), \tag{157}$$

and suppose $0 \le V \le \overline{V}$ almost surely. Then its equal-prior optimal testing error satisfies

$$e^\star \ge \Phi(-\sqrt{\overline{V}}/2). \tag{158}$$

Proof. Enlarge the experiment as follows. Under the central law, conditional on the original observation, draw $W \sim N(0, \overline{V} - V)$ independently. Give the enlarged endpoint experiments densities

$$L_\pm^{\mathrm{aug}} = L_\pm \exp\{\pm W/2 - (\overline{V} - V)/8\} = \exp\{\pm(\mathcal{M} + W)/2 - \overline{V}/8\}. \tag{159}$$

The conditional exponential normalizer makes both densities integrate to one. Dropping W maps the augmented experiment to the original one, so the augmented optimal error cannot be larger.

For every $t \in \mathbb{R}$,

$$\mathbb{E}e^{t(\mathcal{M}+W)} = \mathbb{E}\left[e^{t\mathcal{M}+t^2(\overline{V}-V)/2}\right] = e^{t^2\overline{V}/2}\mathbb{E}e^{t\mathcal{M}-t^2V/2} = e^{t^2\overline{V}/2}. \tag{160}$$

Thus $\mathcal{M} + W \sim N(0, \overline{V})$. It is sufficient for the augmented pair by (159); the Gaussian likelihood-ratio test has error $\Phi(-\sqrt{\overline{V}}/2)$. Data processing proves (158). □

To avoid presupposing that an asymptotic value exists, define the lower and upper global minimax envelopes

$$\underline{\mathfrak{R}}^{\star}_{\mathrm{end}}(c, C) = \inf_{(\pi_T,\psi_T)} \lim_{M\to\infty} \liminf_{T\to\infty} \mathcal{R}_{T,M}(\pi_T, \psi_T; c),$$
$$\overline{\mathfrak{R}}^{\star}_{\mathrm{end}}(c, C) = \inf_{(\pi_T,\psi_T)} \lim_{M\to\infty} \limsup_{T\to\infty} \mathcal{R}_{T,M}(\pi_T, \psi_T; c). \tag{161}$$

Both infima range over policy sequences satisfying (118).

Theorem 5 (Universal fixed-pair lower envelope)**.** *For every sequence of parameter-free predictable allocation rules, pathwise-capped stopping times* (118)*, and recommendations,*

$$\lim_{M\to\infty} \liminf_{T\to\infty} \mathcal{R}_{T,M}(\pi, \psi; c) \geq \Phi\left(-\frac{c\sqrt{CJ_r}}{2\sigma}\right). \tag{162}$$

No allocation-convergence or regularity condition is used. Deterministic full-budget balance with OLS supplies the other endpoint, so that

$$\Phi\left(-\frac{c\sqrt{CJ_r}}{2\sigma}\right) \leq \underline{\mathfrak{R}}^{\star}_{\mathrm{end}}(c, C) \leq \overline{\mathfrak{R}}^{\star}_{\mathrm{end}}(c, C) \leq \Phi\left(-\frac{c\sqrt{C}}{2\sigma\sqrt{\kappa(r)}}\right). \tag{163}$$

If amplitudes are revealed, both global envelopes are exactly

$$\underline{\mathfrak{R}}^{\star}_{\mathrm{or}}(c, C) = \overline{\mathfrak{R}}^{\star}_{\mathrm{or}}(c, C) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma}\right). \tag{164}$$

Proof. Choose a minimizer $\varsigma^{\star}_r$ in Lemma 6 and the fixed separation g in (151). Consider the two local endpoints $\gamma^+ = g/2$ and $\gamma^- = -g/2$. They have gaps $+c$ and $-c$, respectively, and belong to the two cubes in (120) once M is large enough.

At an observation of arm i at local index s, their mean separation is

$$\delta_{i,s,T} = T^{-1/2}\{g_{i,1} + g_{i,2}(s/T)^{-r}\}. \tag{165}$$

Under the central midpoint law, put

$$\mathcal{M}_T = \sum_{t\leq\tau_T} \frac{\delta_{A_t,s_t,T}}{\sigma} Z_t, \qquad V_T = \frac{1}{\sigma^2}\sum_{t\leq\tau_T} \delta^2_{A_t,s_t,T}, \tag{166}$$

where Z_t are the standardized Gaussian innovations and s_t is the chosen arm's local index. Predictability and Gaussianity give, exactly,

$$L_{T,\pm} = \exp\{\pm\mathcal{M}_T/2 - V_T/8\}, \qquad \mathbb{E}_0 e^{t\mathcal{M}_T - t^2V_T/2} = 1 \quad (t \in \mathbb{R}). \tag{167}$$

The prefix Riemann sums for $1, x^{-r}, x^{-2r}$ converge uniformly on $[0, C]$ because $2r < 1$. Lemma 6 and the pathwise cap imply a deterministic sequence $\epsilon_T \downarrow 0$ such that

$$V_T \leq \frac{c^2C}{\sigma^2}(J_r + \epsilon_T) \quad \text{on every sample path.} \tag{168}$$

Lemma 7 therefore bounds the optimal equal-prior error of this fixed pair by

$$e^{\star}_T \geq \Phi\left[-\frac{c\sqrt{C(J_r + \epsilon_T)}}{2\sigma}\right]. \tag{169}$$

The maximum composite error is at least the average error at the two endpoints, which in turn is at least $e_T^\star$. Taking limits proves (162).

For deterministic counts $N_1 = N_2 = CT/2 + o(T)$, the two OLS intercept errors are independent, unbiased for every amplitude, and have variances $\sigma^2\kappa(r)/N_i\{1 + o(1)\}$. Their contrast therefore has $\sqrt{T}$-scaled variance $4\sigma^2\kappa(r)/C + o(1)$, uniformly over the nuisance cube. The sign rule proves the last inequality in (163).

If amplitudes are revealed, use the fixed endpoints $g_1 = (-c, 0)$ and $g_2 = (c, 0)$. Then $V_T = c^2\tau_T/(\sigma^2 T) \le c^2 C/\sigma^2$ on every path, so the same completion lemma gives the lower bound $\Phi(-c\sqrt{C}/(2\sigma))$. Balanced corrected means attain it, proving the oracle equality. □

Corollary 2 (Barrier for single-pair bracket completion)**.** *Among fixed two-point separations with opposite gaps of magnitude c, the smallest uniform pathwise bracket cap is $c^2 C J_r/\sigma^2$. Consequently, Lemma 7 applied to any one fixed pair cannot yield the balanced unknown-amplitude lower bound, because $J_r > 1/\kappa(r)$.*

Proof. Lemma 6 is the exact minimization over all fixed pairs with gap separation $2c$ and gives the displayed cap. Lemma 7 converts that cap to $\Phi\{-c\sqrt{C J_r}/(2\sigma)\}$. Equality with the balanced value would require $J_r = 1/\kappa(r)$, which (149) rules out. □

The corollary concerns the single-fixed-pair proof method; it does not show that an endogenous policy improves the composite minimax value. Appendix E supplies the stronger least-favorable prior that is missing here: a diffuse proper Gaussian mixture whose adaptive likelihood bracket converges uniformly to the balanced Schur information.

D.4 Why the terminal-allocation law does not determine risk

For completeness, we record a precise nonidentifiability result. It already holds in the revealed-amplitude submodel, so no count-only formula can repair it in the harder unknown-amplitude model.

Proposition 3 (Same terminal-count law, different optimal pair risks)**.** *There are two predictable fixed-horizon policies with the same central limiting law of $P_1 = N_1(T)/T$ and different optimal local equal-prior two-point testing errors.*

Proof. Reveal both amplitudes, set the equal-arm reference intercept to $\mu_0 = 0$, and fix $0 < \underline{p} < \overline{p} < 1$ and $t_0 > 0$. Put $n_{\underline{p}} = \lfloor \underline{p} T \rfloor$ and $n_{\overline{p}} = \lfloor \overline{p} T \rfloor$. Pull arm 1 through $n_{\underline{p}}$ observations and form the observable corrected standardized sum

$$U_T = \frac{1}{\sigma\sqrt{n_{\underline{p}}}} \sum_{s=1}^{n_{\underline{p}}} \{Y_{1,s} - \beta_1 s^{-r} - \mu_0\}. \tag{170}$$

Continue arm 1 through $n_{\overline{p}}$ observations exactly when $|U_T| \le t_0$; otherwise stop it at $n_{\underline{p}}$. Fill every remaining round with arm 2. Under the central law $U_T \sim N(0, 1)$ exactly. With B denoting the corresponding Brownian limit,

$$P_1 \Rightarrow P_1^\star = \begin{cases} \underline{p}, & |\mathsf{B}_{\underline{p}}| > t_0\sqrt{\underline{p}}, \\ \overline{p}, & |\mathsf{B}_{\underline{p}}| \le t_0\sqrt{\underline{p}}, \end{cases} \qquad \mathbb{P}(P_1^\star = \underline{p}) = w = 2\Phi(-t_0). \tag{171}$$

Compare this with an independent time-zero coin that chooses $n_{\underline{p}}$ arm-1 pulls with probability $w = 2\Phi(-t_0)$ and $n_{\overline{p}}$ otherwise, filling with arm 2. The two finite central terminal-count laws coincide exactly, and their normalized versions converge to (171). The Gaussian functional limit gives the stopped local experiment below.

For the pair $\mu_1 = \mp c/\sqrt{T}, \mu_2 = 0$, write $\bar{c} = c/\sigma$. The midpoint-law pair error of either limiting experiment is

$$e(\bar{c}) = \frac{1}{2}\mathbb{E}\exp\{-\bar{c}^2 P_1^\star/2 - \bar{c}|\mathsf{B}_{P_1^\star}|\} = \frac{1}{2} - \frac{\bar{c}}{2}\mathbb{E}|\mathsf{B}_{P_1^\star}| + O(\bar{c}^2). \tag{172}$$

Let $\ell = \overline{p} - \underline{p}$ and $\delta_\ell(x) = \mathbb{E}|x + \sqrt{\ell}Z| - |x|$, $Z \sim N(0,1)$. For $x > 0$,

$$\delta'_\ell(x) = -2\Phi(-x/\sqrt{\ell}) < 0. \tag{173}$$

The stopped rule continues on the central set containing the smallest $|\mathsf{B}_{\underline{p}}|$ values of probability $1 - w$. Strict rearrangement gives

$$\begin{aligned}\mathbb{E}|\mathsf{B}_{P_1^\star}|_{\mathrm{stop}} &= \mathbb{E}|\mathsf{B}_{\underline{p}}| + \mathbb{E}\{\delta_\ell(\mathsf{B}_{\underline{p}})\mathbf{1}(|\mathsf{B}_{\underline{p}}| \le t_0\sqrt{\underline{p}})\} \\ &> \mathbb{E}|\mathsf{B}_{\underline{p}}| + (1-w)\mathbb{E}\delta_\ell(\mathsf{B}_{\underline{p}}) = \mathbb{E}|\mathsf{B}_{P_1^\star}|_{\mathrm{coin}}.\end{aligned} \tag{174}$$

By (172), the stopped policy has strictly smaller optimal pair error for every sufficiently small $\bar{c} > 0$, although the marginal count laws are identical. □

Proposition 3 rules out an exact unrestricted risk formula whose only state variable is $\mathrm{law}(P)$: evaluating or optimizing such policies in general requires their joint score–information law (P, X), as in (126). This policy-wise nonidentifiability is consistent with Theorem 2, which identifies the global minimax value without claiming a count-only formula or uniqueness of the unrestricted optimizer.

E Proof of the Adaptive Two-Arm Value

This appendix proves Theorem 2. It first closes the unrestricted two-arm unknown-amplitude sandwich (163) for the below-boundary local experiment (Theorem 6) and then extends the result to every $r > 0$ over a full fixed compact amplitude interval (Theorem 7, which restates Theorem 2). The allocation is parameter-free and predictable, the stopping time satisfies the pathwise cap $\tau_T \le CT$, and no convergence or ancillarity of the terminal allocation is assumed. The key step is a pathwise identity for the likelihood ratio of two proper nuisance mixtures, whose bracket can be capped uniformly over all policies and sample paths.

E.1 A pathwise adaptive Gaussian-mixture identity

Order the local coordinates as $\gamma = (u_1, v_1, u_2, v_2)^\top$, and set

$$d = \frac{c}{2}(-1, 0, 1, 0)^\top, \qquad G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{175}$$

Thus $\gamma = \mathsf{s}d + G\xi$, $\mathsf{s} \in \{-1, 1\}$, has intercept gap $\mathsf{s}c$, while $\xi \in \mathbb{R}^3$ collects the common local intercept and the two local amplitudes. Fix $a > 0$ and give ξ the proper prior $N(0, a^2 I_3)$.

After subtracting the fixed central mean, the observation revealed at global time t has the form

$$Y_t^\circ = \ell_t^\top \gamma + \varepsilon_t. \tag{176}$$

If arm 1 is selected for its nth local pull, then

$$\ell_t = T^{-1/2}(1, (n/T)^{-r}, 0, 0)^\top,$$

and analogously for arm 2. Write $h_t = \ell_t^\top d$ and $\mathsf{z}_t = G^\top \ell_t$.

Lemma 8 (Exact adaptive mixture bracket)**.** *Let $\mathbb{P}_{+,a,T}$ and $\mathbb{P}_{-,a,T}$ be the two sequential laws obtained by first drawing $\xi \sim N(0, a^2 I_3)$, then running an arbitrary parameter-free predictable policy under $\gamma = \pm d + G\xi$. There exist a probability law $\mathbb{P}_{0,a,T}$ on the same stopped experiment, a predictable real process $h_t^\perp$, and a predictable process $\sigma_t^2 > 0$ such that*

$$\frac{d\mathbb{P}_{\pm,a,T}}{d\mathbb{P}_{0,a,T}} = \exp\left\{\pm\frac{\mathcal{M}_T}{2} - \frac{V_T}{8}\right\}, \quad \mathcal{M}_T = 2\sum_{t\le\tau_T}\frac{h_t^\perp}{\sigma_t}Z_t, \quad V_T = 4\sum_{t\le\tau_T}\frac{(h_t^\perp)^2}{\sigma_t^2}, \tag{177}$$

where Z_t are conditionally standard normal under $\mathbb{P}_{0,a,T}$. Consequently,

$$\mathbb{E}_{0,a,T}\exp\{u\mathcal{M}_T - u^2 V_T/2\} = 1 \qquad (u \in \mathbb{R}). \tag{178}$$

Moreover, on every realized adaptive design path, if H is the matrix with rows $\ell_t^\top$, $\mathsf{Z} = HG$, and $h = Hd$, then

$$\frac{V_T}{4} = h^\top(\sigma^2 I + a^2 \mathsf{Z}\mathsf{Z}^\top)^{-1} h, \tag{179}$$

$$= \frac{1}{\sigma^2}\left[d^\top Q_T d - d^\top Q_T G\{G^\top Q_T G + \sigma^2 a^{-2} I_3\}^{-1} G^\top Q_T d\right], \qquad Q_T = H^\top H. \tag{180}$$

Proof. Condition on the policy randomizer and a realized history just before time t. The action and hence (h_t, z_t) are then known. At a fixed full history, every action-kernel and stopping-kernel factor is independent of ξ and the sign because the policy is parameter-free; those factors therefore cancel from Bayes' formula. Under sign s, Gaussian conjugacy gives posterior nuisance covariance and mean

$$\Sigma_{t-1}^{\xi} = \left(a^{-2} I_3 + \sigma^{-2} \sum_{j<t} \mathsf{z}_j \mathsf{z}_j^\top\right)^{-1}, \qquad m_{\mathsf{s},t-1} = \Sigma_{t-1}^{\xi} \sigma^{-2} \sum_{j<t} \mathsf{z}_j (Y_j^\circ - \mathsf{s} h_j).$$

Thus the two one-step predictive distributions have the common variance

$$\sigma_t^2 = \sigma^2 + \mathsf{z}_t^\top \Sigma_{t-1}^{\xi} \mathsf{z}_t$$

and respective means $m_{0,t} \pm h_t^\perp$, where

$$m_{0,t} = \mathsf{z}_t^\top \Sigma_{t-1}^{\xi} \sigma^{-2} \sum_{j<t} \mathsf{z}_j Y_j^\circ, \qquad h_t^\perp = h_t - \mathsf{z}_t^\top \Sigma_{t-1}^{\xi} \sigma^{-2} \sum_{j<t} \mathsf{z}_j h_j.$$

Define $\mathbb{P}_{0,a,T}$ as the actual zero-gap proper nuisance mixture: draw $\xi \sim N(0, a^2 I_3)$, set $\gamma = G\xi$, and use the same policy and stopping rule. Its conditional observation law is $N(m_{0,t}, \sigma_t^2)$. Gaussian density ratios then give

$$\frac{d\mathbb{P}_{\pm,a,T}}{d\mathbb{P}_{0,a,T}} = \prod_{t \le \tau_T} \exp\left\{\pm \frac{h_t^\perp (Y_t^\circ - m_{0,t})}{\sigma_t^2} - \frac{(h_t^\perp)^2}{2\sigma_t^2}\right\}.$$

This is (177) with $Z_t = (Y_t^\circ - m_{0,t})/\sigma_t$. Conditional Gaussianity and the bounded number of increments give (178).

It remains to identify the bracket without conditioning on the adaptively selected row sequence. The following is a deterministic matrix identity. For any fixed rows $\mathsf{z}_1, \dots, \mathsf{z}_m$, scalars $h_1, \dots, h_m$, and

$$\mathsf{K}_t = \sigma^2 I_t + a^2 \mathsf{Z}_t \mathsf{Z}_t^\top,$$

the Schur complement of the leading block K_{t-1} in K_t is

$$\sigma_t^2 = \sigma^2 + \mathsf{z}_t^\top \left(a^{-2} I + \sigma^{-2} \mathsf{Z}_{t-1}^\top \mathsf{Z}_{t-1}\right)^{-1} \mathsf{z}_t.$$

The corresponding residual of h_t after the same block elimination is

$$h_t^\perp = h_t - \mathsf{z}_t^\top \left(a^{-2} I + \sigma^{-2} \mathsf{Z}_{t-1}^\top \mathsf{Z}_{t-1}\right)^{-1} \sigma^{-2} \mathsf{Z}_{t-1}^\top h_{1:t-1}.$$

The block-inverse formula therefore gives, pathwise,

$$h_{1:t}^\top \mathsf{K}_t^{-1} h_{1:t} = h_{1:t-1}^\top \mathsf{K}_{t-1}^{-1} h_{1:t-1} + \frac{(h_t^\perp)^2}{\sigma_t^2}.$$

Induction from $t = 1$ proves (179) for every row sequence, including a row sequence selected from the preceding responses. No assertion about the conditional law given that selected sequence is used. Applying the Woodbury identity and $Q_T = H^\top H$ gives (180). □

E.2 Uniform diffuse-prior information bound

For $\mathcal{S}_C = \{(p_1, p_2) : p_i \geq 0,\ p_1 + p_2 \leq C\}$, write

$$\mathcal{Q}_r(p) = \mathrm{diag}\{Q_r(p_1), Q_r(p_2)\}.$$

For $\lambda > 0$, define

$$J_\lambda(p) = \frac{1}{\sigma^2}\left[d^\top \mathcal{Q}_r(p) d - d^\top \mathcal{Q}_r(p) G\{G^\top \mathcal{Q}_r(p) G + \lambda I_3\}^{-1} G^\top \mathcal{Q}_r(p) d\right]. \tag{181}$$

Lemma 9 (Uniform ridge-to-Schur limit)**.** *As $\lambda \downarrow 0$,*

$$J_\lambda(p) \downarrow J_0(p) := \frac{c^2}{\sigma^2}\frac{D_r(p_1) D_r(p_2)}{D_r(p_1) + D_r(p_2)} = \frac{c^2}{\sigma^2}\frac{p_1 p_2}{\kappa(r)(p_1 + p_2)} \tag{182}$$

uniformly over $\mathcal{S}_C$, with the ratio defined as zero when $p_1 p_2 = 0$. Consequently,

$$\sup_{p \in \mathcal{S}_C} J_0(p) = \frac{c^2 C}{4\sigma^2 \kappa(r)}. \tag{183}$$

Proof. The bracketed expression in (181) is the minimum over $\zeta \in \mathbb{R}^3$ of

$$(d - G\zeta)^\top \mathcal{Q}_r(p)(d - G\zeta) + \lambda \|\zeta\|^2.$$

It therefore decreases pointwise as $\lambda \downarrow 0$ to the same minimum without the ridge penalty. Projecting first off each arm's slope leaves intercept informations $D_r(p_1)$ and $D_r(p_2)$. Projecting the vector $(-c/2, c/2)$ off their common intercept gives

$$c^2 \frac{D_r(p_1) D_r(p_2)}{D_r(p_1) + D_r(p_2)}.$$

This formula remains zero and continuous when either information is zero. Each J_λ is continuous on the compact set $\mathcal{S}_C$, and the continuous pointwise limit J_0 is approached monotonically. Dini's theorem gives uniform convergence along every sequence $\lambda \downarrow 0$, hence the stated uniform limit. Finally,

$$\frac{p_1 p_2}{p_1 + p_2} \leq \frac{p_1 + p_2}{4} \leq \frac{C}{4},$$

with equality at $p_1 = p_2 = C/2$. □

For integer $0 \leq n \leq CT$, recall the discrete prefix Gram matrix

$$Q_{r,T}(n/T) = \frac{1}{T}\sum_{s=1}^{n} g_r(s/T) g_r(s/T)^\top. \tag{184}$$

Because $1, x^{-r}, x^{-2r}$ are integrable and monotone on every $(0, C]$ when $2r < 1$, their partial Riemann sums converge uniformly:

$$\sup_{0 \leq n \leq CT} \|Q_{r,T}(n/T) - Q_r(n/T)\| \longrightarrow 0. \tag{185}$$

For fixed a, the ridge expression in (180) is continuous because its ridge matrix is strictly positive definite, so it converges uniformly over all terminal prefix-count pairs to $J_{\sigma^2/a^2}(p)$.

Lemma 10 (Uniform pathwise bracket cap)**.** *For every $\epsilon > 0$ there are a finite prior scale a and a horizon T_0 such that, for every $T \geq T_0$, every parameter-free predictable policy, every stopping time $\tau_T \leq CT$, and every sample path,*

$$V_T \leq \frac{c^2 C}{\sigma^2 \kappa(r)} + \epsilon. \tag{186}$$

Proof. First use Lemma 9 to choose a so large that

$$\sup_{p \in \mathcal{S}_C} J_{\sigma^2/a^2}(p) \leq \frac{c^2 C}{4\sigma^2 \kappa(r)} + \frac{\epsilon}{8}.$$

Then use (185) to make the finite ridge expression in (180) at most the right side plus $\epsilon/8$, simultaneously for all terminal count pairs. Therefore it is at most $c^2 C/\{4\sigma^2 \kappa(r)\} + \epsilon/4$; multiplying by four proves (186). □

E.3 Exact unrestricted minimax value

Theorem 6 (Unrestricted endogenous unknown-amplitude minimax value)**.** *For every $0 < r < 1/2$, $c > 0$, and $C > 0$, over all parameter-free predictable allocations, all stopping times satisfying $\tau_T \le CT$ almost surely, and all terminal recommendations,*

$$\underline{\mathfrak{R}}^{\star}_{\mathrm{end}}(c, C) = \overline{\mathfrak{R}}^{\star}_{\mathrm{end}}(c, C) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma\sqrt{\kappa(r)}}\right). \tag{187}$$

Proof. Fix an arbitrary policy sequence and $\epsilon > 0$. Lemma 10 selects a finite prior scale a_ϵ and a horizon T_ϵ, uniformly over that policy (indeed, uniformly over all policies), such that the two proper Gaussian nuisance mixtures at scale a_ϵ obey (186). Lemma 8 and the Gaussian-bracket completion Lemma 7 therefore give

$$\liminf_{T\to\infty} e^{\star}_{a_\epsilon,T} \ge \Phi\left\{-\frac{1}{2}\sqrt{\frac{c^2 C}{\sigma^2\kappa(r)} + \epsilon}\right\}, \tag{188}$$

where $e^{\star}_{a_\epsilon,T}$ is their equal-prior Bayes testing error.

It remains to place the proper prior inside the finite local cube. For $M > c/2$, let

$$\mathcal{K}_M = \{\xi : |\xi_0| \le M - c/2,\ |\xi_1| \le M,\ |\xi_2| \le M\}, \tag{189}$$

and condition $N(0, a_\epsilon^2 I_3)$ on $\mathcal{K}_M$. Both signs $\gamma = \pm d + G\xi$ are then supported on the respective boundary subsets of $\Gamma_{\pm,M}(c)$. Markov-kernel contraction gives

$$\mathrm{TV}(\mathbb{P}^{\mathcal{K}_M}_{\pm,a_\epsilon,T}, \mathbb{P}_{\pm,a_\epsilon,T}) \le \mathbb{P}\{\xi \notin \mathcal{K}_M\} =: \varpi_{a_\epsilon,M}. \tag{190}$$

Since equal-prior Bayes error is $\{1 - \mathrm{TV}(P_+, P_-)\}/2$, its value under the truncated mixtures is at least $e^{\star}_{a_\epsilon,T} - \varpi_{a_\epsilon,M}$. Hence, for every recommendation,

$$\mathcal{R}_{T,M}(\pi, \psi; c) \ge e^{\star}_{a_\epsilon,T} - \varpi_{a_\epsilon,M}. \tag{191}$$

Taking $\liminf_T$, then $M \to \infty$, removes the truncation error because $\varpi_{a_\epsilon,M} \to 0$. Finally let $\epsilon \downarrow 0$. The bracket bound is uniform over policies, so the resulting lower bound survives the outer infimum in (161).

Conversely, deterministic full-budget balance $N_1 = N_2 = CT/2 + o(T)$, followed by the full-prefix OLS-intercept sign rule, has worst local error converging to the right side of (187), uniformly over every fixed nuisance cube. This is the attaining rule already used for the upper endpoint of (163), and it supplies the matching upper bound for both global envelopes. □

E.4 Full fixed-amplitude class and all decay exponents

Fix a compact interval $\mathcal{B} \subset (0, \infty)$ with $\beta_0 \in \mathrm{int}(\mathcal{B})$. For $r > 0$, let $\mathcal{R}^{\pi_T,\mathcal{B}}_{T,M}(c; r)$ be, as in (14), the infimum over randomized terminal recommendations of the maximum wrong-label probability over

$$\begin{aligned}\Theta^{\mathcal{B}}_{+,T,M}(c) &= \{\mu_i = \mu_0 + u_i/\sqrt{T} :\ |u_i| \le M,\ u_2 - u_1 \ge c,\ \beta_i \in \mathcal{B}\},\\ \Theta^{\mathcal{B}}_{-,T,M}(c) &= \{\mu_i = \mu_0 + u_i/\sqrt{T} :\ |u_i| \le M,\ u_1 - u_2 \ge c,\ \beta_i \in \mathcal{B}\}.\end{aligned} \tag{192}$$

Define $\underline{\mathfrak{R}}^{\star,\mathcal{B}}_{\mathrm{end,un}}(c, C; r)$ and $\overline{\mathfrak{R}}^{\star,\mathcal{B}}_{\mathrm{end,un}}(c, C; r)$ by taking the infimum over parameter-free predictable policy sequences with $\tau_T \le CT$ almost surely of, respectively, $\lim_{M\to\infty} \liminf_T \mathcal{R}^{\pi_T,\mathcal{B}}_{T,M}(c; r)$ and the corresponding $\limsup_T$. Define the oracle envelopes $\underline{\mathfrak{R}}^{\star,\mathcal{B}}_{\mathrm{end,or}}$ and $\overline{\mathfrak{R}}^{\star,\mathcal{B}}_{\mathrm{end,or}}$ analogously, allowing the allocation, stopping, and recommendation kernels to depend on the revealed amplitude vector.

Lemma 11 (Equal-amplitude endpoint bound)**.** *For every $r > 0$, every T, and every admissible policy, the optimal equal-prior testing error between the endpoints*

$$(u_1, u_2, \beta_1, \beta_2) = (-c/2, c/2, \beta_0, \beta_0) \quad \textit{and} \quad (c/2, -c/2, \beta_0, \beta_0)$$

is at least $\Phi\{-c\sqrt{C}/(2\sigma)\}$.

Proof. Relative to their central midpoint, let

$$h_t = \frac{c}{2\sqrt{T}}\{\mathbf{1}(A_t = 2) - \mathbf{1}(A_t = 1)\}.$$

The parameter-free action and stopping kernels cancel at each stopped history. Therefore the endpoint densities have the exact form

$$L_{T,\pm} = \exp\{\pm\mathcal{M}_T/2 - V_T/8\}, \qquad \mathcal{M}_T = \frac{2}{\sigma^2}\sum_{t\le\tau_T} h_t\varepsilon_t, \qquad V_T = \frac{4}{\sigma^2}\sum_{t\le\tau_T} h_t^2 = \frac{c^2\tau_T}{\sigma^2 T}. \tag{193}$$

Predictable Gaussian increments give $\mathbb{E}_0 \exp\{u\mathcal{M}_T - u^2V_T/2\} = 1$, while the pathwise cap gives $V_T \le c^2C/\sigma^2$. Lemma 7 proves the claim. □

Theorem 7 (Unrestricted two-arm value for every decay exponent)**.** *For every $r > 0$, $c > 0$, and $C > 0$,*

$$\underline{\mathfrak{R}}^{\star,\mathcal{B}}_{\mathrm{end,un}}(c, C; r) = \overline{\mathfrak{R}}^{\star,\mathcal{B}}_{\mathrm{end,un}}(c, C; r) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma\sqrt{\kappa(r)}}\right). \tag{194}$$

The value does not depend on the particular $\mathcal{B}$ satisfying the interior condition. Moreover,

$$\underline{\mathfrak{R}}^{\star,\mathcal{B}}_{\mathrm{end,or}}(c, C; r) = \overline{\mathfrak{R}}^{\star,\mathcal{B}}_{\mathrm{end,or}}(c, C; r) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma}\right). \tag{195}$$

Proof. For $0 < r < 1/2$, the local paths $\beta_{i,T} = \beta_0 + v_iT^{r-1/2}$, $|v_i| \le M$, used in Theorem 6 lie in $\mathcal{B}^2$ for every fixed M and all sufficiently large T. They are a submodel of (192), so that theorem supplies the lower bound.

For $r \ge 1/2$ and $M \ge c/2$, the two endpoints in Lemma 11 belong to the two composite classes. Their average error lower-bounds the maximum composite error, so the lemma gives the lower bound in (194), where $\kappa(r) = 1$.

For the common upper bound, pull each arm $n_T = \lfloor CT/2 \rfloor$ times and recommend by the sign of the two OLS intercepts. Uniformly over $\mathcal{B}^2$, their difference has the exact law

$$\widehat{\mu}_2 - \widehat{\mu}_1 \sim N\left(\mu_2 - \mu_1, \frac{2\sigma^2}{D_{n_T}}\right).$$

Its worst error for $M \ge c/2$ occurs at gap $c/\sqrt{T}$, and $D_{n_T}/n_T \to 1/\kappa(r)$. This gives the matching upper bound. Corrected means give the same argument with $\kappa(r) = 1$ in the revealed experiment. For the oracle lower bound, use the same equal-amplitude symmetric endpoints for every $r > 0$: both hypotheses reveal the identical vector (β_0, β_0), so even amplitude-dependent allocation and stopping kernels cancel from their likelihood ratio. Lemma 11 therefore gives the lower bound matching the corrected-mean rule. □

Remark 2 (What is and is not closed)**.** Theorem 7 closes the two-arm minimax value for every decay exponent over a full fixed amplitude interval. It does not imply that every optimal unrestricted policy is allocation-ancillary, that deterministic balance is the unique unrestricted optimum, or that the risk of a given endogenous policy is determined by the marginal law of its terminal counts. The same-count-law counterexample of Appendix D remains valid. The result also does not cover expected-budget-only stopping or general simultaneous-near-tie endogenous $K > 2$ allocation; Appendix G treats the separated one-local-challenger class.

F Deterministic-Prefix K-Arm Identification

This appendix treats K arms with deterministic prefixes $\boldsymbol{n} = (n_1, \ldots, n_K)$. It first states and proves the exact Gaussian reduction summarized in Section 6 (Theorem 8) and then derives explicit constants for two boundary classes (Theorem 9), including the deterministic one-challenger budget that enters the $K/2$ comparison of Theorem 3. Throughout, $x_s = s^{-r}$, and a_n, b_n, D_n are as in equations (3) and (5). Fix an interior point $\theta_0 = (\mu_0, \beta_0)$ and $\rho > 0$ such that the closed ball $B(\theta_0, \rho)$ is contained in the interior of Θ_S.

F.1 THE GENERAL GAUSSIAN REDUCTION

Let $\mathcal{H} = \bigcup_{b=1}^{K} \mathcal{H}_b \subset \mathbb{R}^K$ be compact with nonempty disjoint components satisfying

$$h_j - h_b \geq 1 \quad (h \in \mathcal{H}_b,\ j \neq b). \tag{196}$$

For a vector $\boldsymbol{\upsilon}$ with positive entries, define the Gaussian ranking value

$$\mathcal{V}_{\mathcal{H}}(t; \boldsymbol{\upsilon}) := \inf_{\delta} \max_{b} \sup_{h \in \mathcal{H}_b} \mathbb{P}_{N(th, \mathrm{diag}(\boldsymbol{\upsilon}))}\{\delta \neq b\}, \tag{197}$$

where δ ranges over randomized kernels.

The local class used in the theorem is

$$\mathcal{P}_{T,b}^{K,\mathrm{un}}(c, \mathcal{H}) = \left\{ (\theta_i)_{i=1}^{K} : \mu_i = \mu_0 + \frac{c}{\sqrt{T}} h_i \text{ for some } h \in \mathcal{H}_b, \quad \theta_i \in B(\theta_0, \rho) \right\}, \tag{198}$$

where the amplitudes are otherwise unrestricted. Its minimax risk is

$$\mathcal{R}_T^{K,\mathrm{un}}(c, \mathcal{H}; \boldsymbol{n}) = \inf_{\psi} \max_{b \in \{1,\dots,K\}} \sup_{\vartheta \in \mathcal{P}_{T,b}^{K,\mathrm{un}}(c,\mathcal{H})} \mathbb{P}_{\vartheta}(\psi \neq b), \tag{199}$$

where the infimum allows randomized recommendation rules.

Theorem 8 (Exact fixed-ball K-arm reduction)**.** *For $n_i \geq 2$, $\sum_i n_i = T$, and all large T,*

$$\mathcal{R}_T^{K,\mathrm{un}}(c, \mathcal{H}; \boldsymbol{n}) = \mathcal{V}_{\mathcal{H}}\left(\frac{c}{\sigma}; \left(\frac{T}{D_{n_i}} \right)_{i=1}^{K} \right). \tag{200}$$

If $n_i/T \to p_i > 0$, the limit is

$$\mathcal{V}_{\mathcal{H}}\left(\frac{c}{\sigma\sqrt{\kappa(r)}}; \left(\frac{1}{p_i} \right)_{i=1}^{K} \right); \tag{201}$$

revealing amplitudes replaces $\kappa(r)$ by one. Optimizing over positive limiting fractions and inverting any target error in $(0, 1 - 1/K)$ gives the limiting squared-critical-signal ratio

$$T_{\mathrm{un}}^{K,\mathrm{loc}} / T_{\mathrm{or}}^{K,\mathrm{loc}} = \kappa(r). \tag{202}$$

The shrinking least-favorable amplitude sequence in the lower bound is a proof device, not a restriction of the fixed nuisance ball. For arbitrary $\mathcal{H}$, the last display concerns optimized limiting critical signals; an expansion of the actual least budget is proved only for the one-local-challenger class (Theorem 9).

Lemma 12 (Efficient intercept coordinate)**.** *Let $\boldsymbol{x}_n = (x_1, \dots, x_n)^\top$ and $\mathbf{1}_n$ be the all-ones vector. Put*

$$\boldsymbol{\ell}_n = \mathbf{1}_n - \frac{a_n}{b_n} \boldsymbol{x}_n. \tag{203}$$

Then $\boldsymbol{\ell}_n^\top \boldsymbol{x}_n = 0$, $\|\boldsymbol{\ell}_n\|^2 = D_n$, and the OLS intercept satisfies

$$\widehat{\mu}_n = \frac{\boldsymbol{\ell}_n^\top \boldsymbol{Y}_n}{D_n} \sim N\left(\mu, \frac{\sigma^2}{D_n} \right). \tag{204}$$

Its distribution is independent of β. Under the perturbation

$$\mu = \mu_0 + t, \qquad \beta = \beta_0 - t\frac{a_n}{b_n}, \tag{205}$$

the observation mean changes by exactly $t\boldsymbol{\ell}_n$. The projection along $\boldsymbol{\ell}_n$ is independent of an orthogonal complement whose law is common to every t in this submodel.

Proof. Direct calculation gives

$$\boldsymbol{\ell}_n^\top \boldsymbol{x}_n = a_n - (a_n/b_n) b_n = 0, \qquad \|\boldsymbol{\ell}_n\|^2 = n - 2a_n^2/b_n + a_n^2/b_n = D_n.$$

Inverting the two-column Gram matrix yields

$$e_1^\top \begin{pmatrix} n & a_n \\ a_n & b_n \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{1}_n^\top \\ \boldsymbol{x}_n^\top \end{pmatrix} = \frac{\boldsymbol{\ell}_n^\top}{D_n},$$

which proves (204). The mean change in (205) is $t\mathbf{1}_n - t(a_n/b_n)\boldsymbol{x}_n = t\boldsymbol{\ell}_n$. Orthogonal projections of isotropic Gaussian noise are independent. □

Lemma 13 (Uniform nuisance feasibility)**.** *For every fixed $r > 0$,*

$$\frac{D_n}{n} \longrightarrow \frac{1}{\kappa(r)}, \qquad \frac{a_n}{b_n} = o(\sqrt{n}), \qquad \max_{2\le n\le T} \frac{a_n/b_n}{\sqrt{T}} \longrightarrow 0. \tag{206}$$

Proof. When $0 < r < 1/2$, power-sum asymptotics give

$$a_n \sim \frac{n^{1-r}}{1-r}, \qquad b_n \sim \frac{n^{1-2r}}{1-2r},$$

and hence

$$\frac{D_n}{n} \to 1 - \frac{1-2r}{(1-r)^2} = \frac{r^2}{(1-r)^2} = \frac{1}{\kappa(r)}, \qquad \frac{a_n}{b_n} \asymp n^r = o(\sqrt{n}).$$

At $r = 1/2$, $a_n = 2\sqrt{n} + O(1)$ and $b_n = \log n + O(1)$. For $r > 1/2$, b_n converges while $a_n = O(n^{1-r})$, $O(\log n)$, or $O(1)$ according as $r \in (1/2, 1)$, $r = 1$, or $r > 1$. These cases yield $D_n/n \to 1$ and $a_n/b_n = o(\sqrt{n})$.

For the uniform assertion, given $\epsilon > 0$, choose N such that $a_n/b_n \le \epsilon\sqrt{n}$ for every $n \ge N$. The displayed ratio is at most ϵ for $N \le n \le T$, while its maximum over $2 \le n < N$ is a fixed constant divided by $\sqrt{T}$. □

Proof of Theorem 8. Let $\widehat{\mu}_i$ be the OLS intercept from arm i. Lemma 12 and independence across arms imply

$$Z_i := \frac{\sqrt{T}}{\sigma}(\widehat{\mu}_i - \mu_0) \sim N\left(\frac{c}{\sigma}h_i, \frac{T}{D_{n_i}}\right) \tag{207}$$

independently, for every admissible amplitude vector. Applying any rule from the Gaussian experiment (197) to Z proves the upper bound in (200).

For the reverse inequality, let $\bar{h} = \max_{h\in\mathcal{H}} \max_i |h_i| < \infty$ and restrict the amplitudes to

$$\beta_i(h) = \beta_0 - \frac{c}{\sqrt{T}} h_i \frac{a_{n_i}}{b_{n_i}}. \tag{208}$$

Uniformly over h and every deterministic allocation,

$$|\mu_i - \mu_0| \le c\bar{h}/\sqrt{T} \to 0, \qquad |\beta_i(h) - \beta_0| \le c\bar{h} \max_{2\le n\le T} \frac{a_n/b_n}{\sqrt{T}} \to 0$$

by Lemma 13. Thus the restricted submodel lies in $B(\theta_0, \rho)^K$ for all sufficiently large T. Lemma 12 decomposes each observation vector into the scalar (207) and an independent orthogonal component whose law is common to every $h \in \mathcal{H}$. Integrating any randomized full-data rule over this common ancillary produces a randomized kernel in the Gaussian experiment with exactly the same risks. The restricted submodel therefore has minimax risk equal to the right-hand side of (200). Restricting a parameter class cannot increase minimax risk, proving the matching lower bound.

If $n_i/T \to p_i > 0$, Lemma 13 gives $T/D_{n_i} \to \kappa(r)/p_i$. Gaussian laws with means in the compact set $(c/\sigma)\mathcal{H}$ converge uniformly in total variation when these diagonal covariances converge. Hence the minimax values converge to

$$\mathcal{V}_{\mathcal{H}}\left(\frac{c}{\sigma}; \left(\frac{\kappa(r)}{p_i}\right)_i\right) = \mathcal{V}_{\mathcal{H}}\left(\frac{c}{\sigma\sqrt{\kappa(r)}}; \left(\frac{1}{p_i}\right)_i\right),$$

where the equality follows by scaling the entire observation by $1/\sqrt{\kappa(r)}$. This proves (201).

When the amplitudes are revealed, the corrected sample means are independent $N(\mu_i, \sigma^2/n_i)$ and give the oracle upper bound. Holding every revealed amplitude fixed at β_0 gives the matching lower bound, because the residuals orthogonal to the sample means then have a common law. Thus the exact oracle value is

$$\mathcal{V}_{\mathcal{H}}\left(\frac{c}{\sigma}; \left(\frac{T}{n_i}\right)_i\right),$$

with the limit asserted in the theorem.

It remains to justify the allocation-optimized budget statement. Taking the infimum of the two fixed-p limiting experiments over $p_i > 0$, $\sum_i p_i = 1$, gives

$$\overline{\mathcal{V}}^{\mathrm{un}}_{\mathcal{H}}(c) = \overline{\mathcal{V}}_{\mathcal{H}}\left(\frac{c}{\sigma\sqrt{\kappa(r)}}\right), \qquad \overline{\mathcal{V}}^{\mathrm{or}}_{\mathcal{H}}(c) = \overline{\mathcal{V}}_{\mathcal{H}}\left(\frac{c}{\sigma}\right). \tag{209}$$

The generalized inverse used below is legitimate because $\overline{\mathcal{V}}_{\mathcal{H}}$ is nonincreasing. Indeed, for $0 < s < t$ and fixed p, an observation $X_t \sim N(th, \mathrm{diag}(1/p_i))$ can be degraded to the s-experiment by

$$X_s = \frac{s}{t}X_t + W, \qquad W \sim N\big(0, \{1-(s/t)^2\}\,\mathrm{diag}(1/p_i)\big),$$

with W independent. Thus every rule at signal s is simulable at signal t, so $\mathcal{V}_{\mathcal{H}}(t;(1/p_i)_i) \le \mathcal{V}_{\mathcal{H}}(s;(1/p_i)_i)$; taking the infimum over p preserves this order. At $t = 0$ all best-arm labels give the same law and the minimax error is $1 - 1/K$. Relative to that common law, the KL divergence at signal t is at most $t^2 \max_{h\in\mathcal{H}} \|h\|^2/2$, uniformly in p; Pinsker's inequality therefore gives the same limit as $t \downarrow 0$. On the other hand, (196) makes the error of balanced Gaussian ranking tend uniformly to zero as $t \to \infty$. Hence $q_{\mathcal{H}}(\alpha) \in (0,\infty)$ for $\alpha \in (0, 1-1/K)$.

Taking generalized inverses in (209) gives critical local signals $c_{\mathrm{un}} = \sigma\sqrt{\kappa(r)}q_{\mathcal{H}}(\alpha)$ and $c_{\mathrm{or}} = \sigma q_{\mathcal{H}}(\alpha)$. Since $\Delta = c/\sqrt{T}$, the associated first-order budgets are c_a^2/Δ^2, and their ratio is (202). This argument optimizes fixed limiting Gaussian experiments; it does not interchange a finite-T infimum over integer allocations with the limit. □

F.2 Explicit constants for two boundary classes

We now specialize to two boundary classes: balanced prefixes with all $K-1$ gaps simultaneously at the local scale, and the one-local-challenger class of Section 6. Put

$$E_K(t) := 1 - \int_{-\infty}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}}\Phi(z+t)^{K-1}\,dz. \tag{210}$$

For balanced prefixes $T = Kn$, let $\mathcal{P}_b^K(\Delta)$ be the set of $(\theta_i)_{i=1}^K \in B(\theta_0,\rho)^K$ with $\mu_j - \mu_b \ge \Delta$ for every $j \ne b$. Fix $\eta \in (0, \rho/4)$ and let $\mathcal{C}_{b,j}(\Delta,\eta)$ be the subset of $B(\theta_0,\rho)^K$ in which j is the only potentially local challenger:

$$\mu_j - \mu_b \ge \Delta, \qquad \mu_k - \max(\mu_b,\mu_j) \ge \eta \quad (k \notin \{b,j\}). \tag{211}$$

The ordered pair (b,j) is unknown, and the composite risk of a recommendation ψ under prefixes $\boldsymbol{n}$ is

$$\mathcal{R}_T^{K,\mathrm{pair}}(\psi;\boldsymbol{n},\Delta,\eta) := \max_{b\ne j} \sup_{\vartheta\in\mathcal{C}_{b,j}(\Delta,\eta)} \mathbb{P}_\vartheta\{\psi \ne b\}. \tag{212}$$

Theorem 9 (Explicit multiway risk and allocation-optimal boundary)**.** *Whenever the least-favorable points constructed in the proofs (Appendices F.3 and F.4) belong to the product ball, the smallest-OLS rule is exactly minimax over $\bigcup_b \mathcal{P}_b^K(\Delta)$ under balanced prefixes, with error*

$$E_K\left(\frac{\Delta\sqrt{D_n}}{\sigma}\right). \tag{213}$$

For $\Delta_T = c/\sqrt{T}$ and limiting fractions $p_i > 0$, the one-local-challenger risk tends to

$$\Phi\left[-\frac{c}{\sigma\sqrt{\kappa(r)}}\sqrt{\min_{i<j}\frac{p_ip_j}{p_i+p_j}}\right]. \tag{214}$$

Optimizing over every integer deterministic allocation gives

$$\lim_{T\to\infty} \min_{\substack{n_i\ge 2\\ \sum_i n_i = T}} \inf_{\psi} \mathcal{R}_T^{K,\mathrm{pair}}(\psi;\boldsymbol{n},\Delta_T,\eta) = \Phi\left(-\frac{c}{\sigma\sqrt{2K\kappa(r)}}\right), \tag{215}$$

and every first-order-optimal allocation has $n_i/T \to 1/K$. Revealing the amplitudes replaces $\kappa(r)$ by one.

For $\bullet \in \{\mathrm{un}, \mathrm{or}\}$, let $\mathcal{R}^{K,\mathrm{pair}}_{T,\bullet}$ denote the risk (212) *in the unknown-amplitude ($\bullet = \mathrm{un}$) or revealed-amplitude ($\bullet = \mathrm{or}$) experiment, and define*

$$\mathfrak{R}^{K,\mathrm{pair,det}}_{T,\bullet}(\Delta,\eta) := \min_{\substack{n_i \geq 2 \\ \sum_i n_i = T}} \inf_{\psi} \mathcal{R}^{K,\mathrm{pair}}_{T,\bullet}(\psi; \boldsymbol{n}, \Delta, \eta), \tag{216}$$

and the actual least budget

$$T^{K,\mathrm{pair,det}}_{\bullet}(\Delta,\alpha) := \min\{T \geq 2K : \mathfrak{R}^{K,\mathrm{pair,det}}_{T,\bullet}(\Delta,\eta) \leq \alpha\}. \tag{217}$$

Then, for every $\alpha \in (0, 1/2)$,

$$T^{K,\mathrm{pair,det}}_{\mathrm{un}}(\Delta,\alpha) \sim \frac{2K\sigma^2\kappa(r)z_\alpha^2}{\Delta^2}, \qquad T^{K,\mathrm{pair,det}}_{\mathrm{or}}(\Delta,\alpha) \sim \frac{2K\sigma^2 z_\alpha^2}{\Delta^2}, \quad \Delta \downarrow 0. \tag{218}$$

F.3 Balanced simultaneous near ties

For $T = Kn$, recall that

$$\mathcal{P}^K_b(\Delta) = \big\{(\theta_i)_{i=1}^K \in B(\theta_0,\rho)^K : \mu_j - \mu_b \geq \Delta \text{ for every } j \neq b\big\} \tag{219}$$

and let $\mathcal{R}^{\mathrm{un}}_{K,n}(\Delta)$ be its minimax misidentification risk.

Proof of the balanced claim in Theorem 9. Put $\sigma_n = \sigma/\sqrt{D_n}$. The OLS intercepts are independent $N(\mu_i, \sigma_n^2)$ under every amplitude vector. If b is best, conditioning on $Z = (\widehat{\mu}_b - \mu_b)/\sigma_n$ gives

$$\begin{aligned}
\mathbb{P}(\widehat{\mu}_b < \min_{j \neq b} \widehat{\mu}_j) &= \int \frac{e^{-z^2/2}}{\sqrt{2\pi}} \prod_{j \neq b} \Phi\left(\frac{\mu_j - \mu_b}{\sigma_n} - z\right) dz \\
&\geq \int \frac{e^{-z^2/2}}{\sqrt{2\pi}} \Phi\left(\frac{\Delta}{\sigma_n} - z\right)^{K-1} dz \\
&= 1 - E_K(\Delta/\sigma_n).
\end{aligned}$$

Equality holds when all $K-1$ gaps equal Δ. Thus the smallest-OLS rule has worst-case error $E_K(\Delta\sqrt{D_n}/\sigma)$.

For the matching lower bound, index K parameter points by their best arm b and set

$$h_i^{(b)} = \frac{1}{K} - \mathbf{1}\{i = b\}, \qquad \mu_i^{(b)} = \mu_0 + \Delta h_i^{(b)}, \qquad \beta_i^{(b)} = \beta_0 - \Delta h_i^{(b)} \frac{a_n}{b_n}. \tag{220}$$

Every nonbest intercept exceeds the best by Δ. The identity is therefore exact whenever these points belong to the product ball. Under $\Delta_T = c/\sqrt{Kn}$ they do so eventually, because $\Delta_T \to 0$ and $\Delta_T(a_n/b_n) \to 0$ by Lemma 13.

At these points the observations outside the K OLS coordinates have a common independent law. The OLS vector has covariance $\sigma_n^2 I_K$ and mean $\mu_0\mathbf{1} + \Delta h^{(b)}$. Under a uniform prior on b, the hypotheses have equal norm and their log likelihoods differ only through

$$(h^{(b)})^\top \widehat{\mu} = \frac{1}{K}\sum_{i=1}^K \widehat{\mu}_i - \widehat{\mu}_b.$$

The Bayes rule therefore chooses the smallest OLS coordinate. Its error under every symmetric point is exactly $E_K(\Delta\sqrt{D_n}/\sigma)$. The Bayes risk lower-bounds the minimax risk and matches the upper bound, proving (213). Since $D_n/n \to 1/\kappa(r)$ and $T = Kn$, the asserted local limit follows.

With revealed amplitudes, correct each observation by $\beta_i x_s$ and use its sample mean. The same upper-bound calculation has $\sigma_n = \sigma/\sqrt{n}$. For the lower bound, keep every revealed amplitude fixed at β_0 and use the same intercept vectors. This proves the oracle identity and limit. □

For reference, $E_K(0) = 1 - 1/K$, it is continuous and strictly decreasing, and it tends to zero. If $q_{K,\alpha}$ is the unique solution of $E_K(q_{K,\alpha}) = \alpha$, the exact identity also yields the balanced budget expansions

$$T_{\mathrm{un}}^{K,\mathrm{bal}}(\Delta, \alpha) \sim \frac{K\sigma^2\kappa(r)q_{K,\alpha}^2}{\Delta^2}, \qquad T_{\mathrm{or}}^{K,\mathrm{bal}}(\Delta, \alpha) \sim \frac{K\sigma^2 q_{K,\alpha}^2}{\Delta^2}. \tag{221}$$

Indeed, evaluate the exact risks at $n = (1 \pm \epsilon)\sigma^2\kappa(r)q_{K,\alpha}^2/\Delta^2$ with the appropriate integer rounding. Lemma 13 gives limiting arguments $q_{K,\alpha}\sqrt{1 \pm \epsilon}$. The least-favorable points are uniformly feasible throughout $n = O(\Delta^{-2})$, since

$$\max_{2 \le n \le C\Delta^{-2}} \Delta \frac{a_n}{b_n} \le \Delta\sqrt{\lceil C\Delta^{-2}\rceil} \max_{2 \le n \le \lceil C\Delta^{-2}\rceil} \frac{a_n/b_n}{\sqrt{\lceil C\Delta^{-2}\rceil}} \longrightarrow 0.$$

Monotonicity follows by ignoring added observations. A $(1 \pm \epsilon)$ sandwich and then $\epsilon \downarrow 0$ prove (221); the oracle proof replaces D_n by n.

F.4 One unknown local challenger

Fix $\eta \in (0, \rho/4)$, let $\mathcal{C}_{b,j}(\Delta, \eta)$ and $\mathcal{R}_T^{K,\mathrm{pair}}(\psi; \boldsymbol{n}, \Delta, \eta)$ be as in (211)–(212), and write

$$J_{ij}(\boldsymbol{n}) = \left(D_{n_i}^{-1} + D_{n_j}^{-1}\right)^{-1} \tag{222}$$

for the pooled intercept information of arms i and j.

Proof of the one-challenger claim in Theorem 9. First suppose $n_i/T \to p_i > 0$ for all i. For an upper bound, recommend the smallest OLS intercept. On $\mathcal{C}_{b,j}(\Delta_T, \eta)$, an error implies either $\widehat{\mu}_j \le \widehat{\mu}_b$ or $\widehat{\mu}_k \le \widehat{\mu}_b$ for some $k \notin \{b, j\}$. Therefore

$$\mathbb{P}(\psi \ne b) \le \Phi\left(-\frac{\Delta_T\sqrt{J_{bj}(\boldsymbol{n})}}{\sigma}\right) + \sum_{k \notin \{b,j\}} \Phi\left(-\frac{\eta}{\sigma\sqrt{D_{n_b}^{-1} + D_{n_k}^{-1}}}\right). \tag{223}$$

The sum tends uniformly to zero, and

$$\frac{J_{ij}(\boldsymbol{n})}{T} \longrightarrow \frac{1}{\kappa(r)} \frac{p_i p_j}{p_i + p_j}. \tag{224}$$

Taking the worst ordered pair gives the upper bound in (214).

For the matching lower bound, fix an unordered pair $\{i, j\}$. Write $D_i = D_{n_i}$ and similarly for a_i, b_i, and set

$$\delta_i^\mu = \frac{2\Delta_T D_j}{D_i + D_j}, \qquad \delta_j^\mu = -\frac{2\Delta_T D_i}{D_i + D_j}, \qquad \delta_\ell^\beta = -\delta_\ell^\mu \frac{a_{n_\ell}}{b_{n_\ell}} \quad (\ell \in \{i, j\}). \tag{225}$$

Construct two parameter points by setting

$$\mu_\ell^+ = \mu_0 - \delta_\ell^\mu/2, \quad \mu_\ell^- = \mu_0 + \delta_\ell^\mu/2,$$
$$\beta_\ell^+ = \beta_0 - \delta_\ell^\beta/2, \quad \beta_\ell^- = \beta_0 + \delta_\ell^\beta/2 \qquad (\ell = i, j),$$

and, for every $k \notin \{i, j\}$, use under both hypotheses

$$\mu_k = \mu_0 + \eta + \Delta_T, \qquad \beta_k = \beta_0.$$

Since $\delta_i^\mu - \delta_j^\mu = 2\Delta_T$, the plus point has arm i best with gap Δ_T, and the minus point has arm j best with the same gap. Every noncontender is at least η above the larger contender. Both points are feasible eventually, uniformly over allocations, because

$$|\delta_\ell^\beta|/2 \le \Delta_T \max_{2 \le n \le T} \frac{a_n}{b_n} = (\sqrt{T}\Delta_T) \max_{2 \le n \le T} \frac{a_n/b_n}{\sqrt{T}} \longrightarrow 0.$$

Only arms i and j differ. Their squared Mahalanobis distance is

$$d_{ij}^2 = \frac{1}{\sigma^2}((\delta_i^\mu)^2 D_i + (\delta_j^\mu)^2 D_j) = \frac{4\Delta_T^2}{\sigma^2} \frac{D_i D_j}{D_i + D_j} = \frac{4\Delta_T^2 J_{ij}(\boldsymbol{n})}{\sigma^2}.$$

The equal-prior error for two common-covariance Gaussian laws at distance d_{ij} is $\Phi(-d_{ij}/2)$. Every recommendation induces a test, so

$$\inf_{\psi} \mathcal{R}_T^{K,\mathrm{pair}}(\psi; \boldsymbol{n}, \Delta_T, \eta) \geq \max_{i<j} \Phi\left(-\frac{\Delta_T \sqrt{J_{ij}(\boldsymbol{n})}}{\sigma}\right). \tag{226}$$

Together with (223) and (224), this proves (214).

We next optimize over every deterministic integer allocation. For positive fractions p define

$$m(p) = \min_{i<j} \frac{p_i p_j}{p_i + p_j}.$$

Order them as $p_{(1)} \leq p_{(2)} \leq \cdots \leq p_{(K)}$. The average of the two smallest cannot exceed the overall average, and hence

$$m(p) \leq \frac{p_{(1)} p_{(2)}}{p_{(1)} + p_{(2)}} \leq \frac{p_{(1)} + p_{(2)}}{4} \leq \frac{1}{2K}. \tag{227}$$

Equality throughout forces every $p_i = 1/K$, which also attains the bound.

This fraction argument covers arbitrary allocation sequences. Along every subsequence, pass to a further subsequence with $n_i/T \to p_i \in [0,1]$. If some $p_i = 0$, then $\min_{a<b} J_{ab}/T \leq D_{n_i}/T \leq n_i/T \to 0$. If every fraction is positive, Lemma 13 and (227) give

$$\limsup_T \min_{i<j} \frac{J_{ij}}{T} \leq \frac{1}{2K\kappa(r)}.$$

The lower bound (226) therefore shows that no deterministic allocation improves on (215). Floor/ceiling-balanced prefixes, the OLS rule, and (223) attain it. The equality case in (227) proves that every first-order optimal sequence has $n_i/T \to 1/K$.

The same convergence statements hold whenever $\Delta_T \sqrt{T} \to c \in (0, \infty)$, not only for the exact choice $c/\sqrt{T}$: in the upper and lower bounds the local term appears only as $\Delta_T \sqrt{J_{ij}}$, while the noncontender tails still vanish uniformly and the least-favorable amplitude perturbations remain $o(1)$.

With revealed amplitudes, corrected sample means replace OLS intercepts and D_{n_i} is replaced throughout by n_i. The same upper bound, two-point lower bound, and allocation optimization apply with $\kappa(r) = 1$.

Finally, consider the optimized risks and actual least budgets defined in (217). For $\alpha \in (0, 1/2)$ put

$$\mathsf{T}_{\mathrm{un}}^{K,\mathrm{pair}} = 2K\sigma^2\kappa(r) z_\alpha^2, \qquad \mathsf{T}_{\mathrm{or}}^{K,\mathrm{pair}} = 2K\sigma^2 z_\alpha^2,$$

where $z_\alpha = \Phi^{-1}(1-\alpha)$ as in the main text. For a sequence $\Delta \downarrow 0$, evaluate the optimized risks at integer budgets $T_{\bullet,\pm} = (1 \pm \epsilon)\mathsf{T}_\bullet^{K,\mathrm{pair}}/\Delta^2$ with the appropriate rounding. Then $\Delta\sqrt{T_{\bullet,\pm}} \to \sqrt{(1 \pm \epsilon)\mathsf{T}_\bullet^{K,\mathrm{pair}}}$, so the optimized limit is eventually above α at the lower budget and below it at the upper budget. The optimized risk is nonincreasing in T, since a rule can ignore an added observation. A $(1 \pm \epsilon)$ sandwich followed by $\epsilon \downarrow 0$ gives

$$T_{\mathrm{un}}^{K,\mathrm{pair,det}}(\Delta, \alpha) \sim \frac{2K\sigma^2\kappa(r) z_\alpha^2}{\Delta^2}, \qquad T_{\mathrm{or}}^{K,\mathrm{pair,det}}(\Delta, \alpha) \sim \frac{2K\sigma^2 z_\alpha^2}{\Delta^2}.$$

Their ratio tends to $\kappa(r)$, completing the proof. □

The deterministic qualifier is essential in this appendix. The least-favorable amplitude perturbations in (208) and (225) depend on the prescribed terminal prefixes. They cannot be evaluated at realized random counts in a sequential lower bound. Accordingly, these proofs establish neither nondegenerate random-limit allocation optimality nor optional-stopping optimality.

G Proof of the K-Arm Screening Theorem

This appendix proves Theorem 3, the exact unrestricted fixed-K result for the one-local-challenger class. The result permits arbitrary parameter-free predictable allocation, independent policy randomization, and stopping under a pathwise budget cap; its fixed-separation assumption on the remaining

$K-2$ arms is essential. The lower bound reuses the diffuse-prior bracket argument of Appendix E on the local pair, and the upper bound analyzes the screen–then–compare policy of Section 6.

For fixed $K \geq 2$, arm i produces the prefix

$$Y_{i,s} = \mu_i + \beta_i s^{-r} + \varepsilon_{i,s}, \qquad \varepsilon_{i,s} \overset{\text{iid}}{\sim} N(0, \sigma^2), \qquad r > 0. \tag{228}$$

Let $B(\theta_0, \rho)$ be a fixed closed ball contained in the interior of the arm parameter space, where $\theta_0 = (\mu_0, \beta_0)$. Fix $0 < \eta < \rho/4$. For an ordered pair $b \neq j$ and gap $\Delta > 0$, let

$$\mathcal{C}_{b,j}(\Delta, \eta) := \left\{ (\theta_i)_{i=1}^K \in B(\theta_0, \rho)^K : \mu_j - \mu_b \geq \Delta, \quad \mu_k - \max(\mu_b, \mu_j) \geq \eta \ (k \notin \{b, j\}) \right\}. \tag{229}$$

Thus b is the best arm, j is its only possibly local challenger, and every other arm is separated from both by a fixed margin; the ordered pair (b, j) is unknown to the policy.

At every global time, a parameter-free predictable kernel chooses an arm and reveals its next unused prefix observation. The stopping time obeys

$$\tau_T \leq CT \qquad \text{almost surely.} \tag{230}$$

For a policy and terminal recommendation ψ_T, define

$$\mathcal{R}_T^{K,\text{pair}}(\pi_T, \psi_T; \Delta, \eta) := \max_{b \neq j} \sup_{\vartheta \in \mathcal{C}_{b,j}(\Delta,\eta)} \mathbb{P}_\vartheta(\psi_T \neq b). \tag{231}$$

Recall that

$$\kappa(r) = \begin{cases} (1-r)^2/r^2, & 0 < r < 1/2, \\ 1, & r \geq 1/2. \end{cases} \tag{232}$$

G.1 Suffix intercept information

The screening pilot and the final comparison must use independent data. We first verify that discarding the pilot observations from the final OLS fit does not change the first-order intercept information.

For integers $m \geq 1$ and $q \geq 2$, put

$$a_{m,q} = \sum_{s=m+1}^{m+q} s^{-r}, \qquad b_{m,q} = \sum_{s=m+1}^{m+q} s^{-2r},$$
$$D_{m,q} = q - \frac{a_{m,q}^2}{b_{m,q}}. \tag{233}$$

Lemma 14 (Suffix Schur information)**.** *Suppose $m = m_T \to \infty$, $q = q_T \to \infty$, and, for some fixed $\omega \in (0, 1)$,*

$$m_T = T^{\omega + o(1)}, \qquad q_T/T \to q_0 \in (0, \infty). \tag{234}$$

Then, for every fixed $r > 0$,

$$\frac{D_{m_T,q_T}}{q_T} \longrightarrow \frac{1}{\kappa(r)}. \tag{235}$$

For every sufficiently large T (so that $q_T \geq 2$), OLS fitted to local indices $m_T + 1, \ldots, m_T + q_T$ satisfies exactly

$$\widehat{\mu}_{m_T,q_T} \sim N\left(\mu, \frac{\sigma^2}{D_{m_T,q_T}}\right), \tag{236}$$

independently of β.

Proof. Inverting the two-column Gram matrix with columns 1 and s^{-r} on the suffix gives intercept Schur complement (233). Gaussianity then gives (236).

It remains to calculate the limit. Because $m_T/q_T \to 0$, when $0 < r < 1/2$ the power-sum asymptotics, unchanged after deleting the initial $o(q_T)$ terms, give

$$a_{m_T,q_T} \sim \frac{q_T^{1-r}}{1-r}, \qquad b_{m_T,q_T} \sim \frac{q_T^{1-2r}}{1-2r}.$$

Consequently,

$$\frac{D_{m_T,q_T}}{q_T} \longrightarrow 1 - \frac{1-2r}{(1-r)^2} = \frac{r^2}{(1-r)^2} = \frac{1}{\kappa(r)}.$$

At $r = 1/2$,

$$a_{m_T,q_T} = 2\sqrt{q_T}\{1 + o(1)\}, \qquad b_{m_T,q_T} = \log(q_T/m_T) + O(1) = (1 - \omega + o(1)) \log T,$$

so $a^2_{m_T,q_T}/(q_T b_{m_T,q_T}) = O(1/\log T) \to 0$. If $1/2 < r < 1$, then

$$a_{m_T,q_T} = O(q_T^{1-r}), \qquad b_{m_T,q_T} \asymp m_T^{1-2r},$$

and hence

$$\frac{a^2_{m_T,q_T}}{q_T b_{m_T,q_T}} = O\{(m_T/q_T)^{2r-1}\} \to 0.$$

For $r = 1$ the same ratio is $O\{(m_T/q_T)\log^2(q_T/m_T)\} \to 0$, and for $r > 1$ it is $O(m_T/q_T) \to 0$. Thus $D_{m_T,q_T}/q_T \to 1$ throughout $r \geq 1/2$, proving (235). □

G.2 Exact unrestricted fixed-K value

Theorem 10 (Exact adaptive one-challenger value)**.** *Fix $K \geq 2$, $r > 0$, $c > 0$, $C > 0$, and $\eta \in (0, \rho/4)$. Among all parameter-free predictable policies, stopping times satisfying* (230)*, and randomized terminal recommendations,*

$$\inf_{(\pi_T,\psi_T)} \liminf_{T\to\infty} \mathcal{R}_T^{K,\mathrm{pair}}(\pi_T, \psi_T; c/\sqrt{T}, \eta)$$

$$= \inf_{(\pi_T,\psi_T)} \limsup_{T\to\infty} \mathcal{R}_T^{K,\mathrm{pair}}(\pi_T, \psi_T; c/\sqrt{T}, \eta) = \Phi\left(-\frac{c\sqrt{C}}{2\sigma\sqrt{\kappa(r)}}\right). \tag{237}$$

Both infima range over admissible policy sequences. If the amplitudes are revealed to the policy, the same statement holds with $\kappa(r)$ replaced by one. Because a policy sequence may be chosen horizon by horizon, the statement is equivalent to the horizonwise minimax formulation (23) *of Theorem 3.*

Lower bound. First take $0 < r < 1/2$. Restrict the class to the ordered pairs $(b, j) = (1, 2)$ and $(2, 1)$. Put every arm $k \geq 3$ at the fixed parameter $(\mu_0 + 2\eta, \beta_0)$. On arms 1 and 2, use the least-favorable proper Gaussian nuisance mixture from the proof of Theorem 6: the two signs have local intercept difference $\pm c/\sqrt{T}$, while their common local intercept and two local amplitudes are mixed and then truncated to a fixed cube. For every fixed truncation radius, all these points lie in $B(\theta_0, \rho)^K$ and satisfy (229) for all sufficiently large T. Indeed, their first two intercepts differ from μ_0 by $O(T^{-1/2})$, whereas every other arm is $2\eta + O(T^{-1/2})$ above them.

Observations of arms $k \geq 3$ have a law common to both signs and every value of the mixed nuisance, so their likelihood increments are one. They may affect subsequent allocation, but the adaptive conjugate-mixture identity in Lemma 8 still applies after assigning zero signal and nuisance rows to those increments. On every sample path, the informative pair uses no more than CT observations.

The diffuse-prior and horizon limits must remain in this order. Fix $\epsilon > 0$. Lemma 10 first chooses a finite proper-prior scale a_ϵ and then a horizon T_ϵ such that, simultaneously for every predictable policy and stopped history,

$$V_T \leq \frac{c^2 C}{\sigma^2 \kappa(r)} + \epsilon, \qquad T \geq T_\epsilon. \tag{238}$$

The Gaussian-bracket completion lemma (Lemma 7) therefore bounds the untruncated equal-prior mixture error from below by

$$\Phi\left[-\frac{1}{2}\sqrt{\frac{c^2 C}{\sigma^2\kappa(r)} + \epsilon}\right]. \tag{239}$$

Now condition the same $N(0, a_\epsilon^2 I)$ nuisance prior on a cube of radius M. Markov-kernel contraction changes either endpoint law in total variation by at most the Gaussian prior mass outside that cube, denoted $\varpi_{a_\epsilon,M}$. The truncated prior is supported in (229) for every fixed M and all sufficiently

large T. Consequently, the composite risk is at least (239) minus $\varpi_{a_\epsilon,M}$. Take $\liminf_{T\to\infty}$, then $M\to\infty$, and finally $\epsilon\downarrow 0$. This proves the lower bound in (237) without a residual ridge term.

For $r\geq 1/2$, use the fixed endpoints

$$(\mu_1,\mu_2)=\left(\mu_0-\frac{c}{2\sqrt{T}},\mu_0+\frac{c}{2\sqrt{T}}\right)\quad\text{or its reversal,}$$
$$\beta_1=\beta_2=\beta_0,\qquad(\mu_k,\beta_k)=(\mu_0+2\eta,\beta_0),\quad k\geq 3.$$

Only pulls of arms 1 and 2 carry likelihood signal. Relative to their zero-gap midpoint law, the exact stopped likelihood pair has bracket

$$V_T=\frac{c^2\{N_1(\tau_T)+N_2(\tau_T)\}}{\sigma^2 T}\leq\frac{c^2C}{\sigma^2}.$$

Lemma 7 gives $\Phi\{-c\sqrt{C}/(2\sigma)\}$, which is (237) because $\kappa(r)=1$. This fixed-pair argument also gives the revealed-amplitude lower bound for every $r>0$. □

Upper bound. Choose a fixed $\omega\in(0,1)$ and let $m_T=\lceil T^\omega\rceil$. For all sufficiently large T, pull every arm exactly m_T times. From these pilot prefixes compute the K full-prefix OLS intercepts $\widehat{\mu}_i^{\mathrm{pil}}$, and let $\{\widehat{b}_T,\widehat{j}_T\}$ be the unordered pair with the two smallest pilot estimates. Since each

$$\widehat{\mu}_i^{\mathrm{pil}}-\mu_i\sim N\left(0,\frac{\sigma^2}{D_{m_T}}\right),\qquad D_{m_T}\to\infty,$$

uniformly over all amplitudes, a union bound gives

$$\sup_{b\neq j}\sup_{\vartheta\in\mathcal{C}_{b,j}(c/\sqrt{T},\eta)}\mathbb{P}_\vartheta\left(\{\widehat{b}_T,\widehat{j}_T\}\neq\{b,j\}\right)\leq 2K\Phi\left(-\frac{\eta\sqrt{D_{m_T}}}{4\sigma}\right)\longrightarrow 0.\tag{240}$$

Indeed, if every pilot error is smaller than $\eta/4$, both members of the local pair have estimated loss at least $\eta/2$ below every other arm.

Let $\bar{\tau}_T=\lfloor CT\rfloor$ and

$$q_T=\left\lfloor\frac{\bar{\tau}_T-Km_T}{2}\right\rfloor.$$

For the finitely many horizons at which $\bar{\tau}_T-Km_T<4$, use any admissible rule. Henceforth $q_T\geq 2$. After screening, pull each selected arm q_T additional times and stop. The total number of pulls is at most $\bar{\tau}_T\leq CT$ on every path, while $q_T/T\to C/2$.

Fit the two final OLS intercepts using only the independent suffix observations with local indices $m_T+1,\ldots,m_T+q_T$. Conditional on the pilot and correct screening, Lemma 14 gives two independent normals with their true intercept means and common variance $\sigma^2/D_{m_T,q_T}$. Recommending the smaller suffix OLS intercept therefore has conditional error at most

$$\Phi\left(-\frac{c\sqrt{D_{m_T,q_T}}}{\sigma\sqrt{2T}}\right)\longrightarrow\Phi\left(-\frac{c\sqrt{C}}{2\sigma\sqrt{\kappa(r)}}\right),\tag{241}$$

uniformly over the class, because the local-pair gap is at least $c/\sqrt{T}$. Adding (240) proves the matching unknown-amplitude upper bound. If amplitudes are revealed, correct the suffix observations before averaging; the same argument then has $\kappa(r)=1$. □

G.3 Least-budget consequence

For an integer budget B, let

$$\mathfrak{R}_{B,\bullet}^{K,\mathrm{pair}}(\Delta,\eta):=\inf_{\pi,\psi:\,\tau\leq B\ \mathrm{a.s.}}\max_{b\neq j}\sup_{\vartheta\in\mathcal{C}_{b,j}(\Delta,\eta)}\mathbb{P}_{\vartheta,\bullet}(\psi\neq b),\qquad\bullet\in\{\mathrm{un},\mathrm{or}\},\tag{242}$$

where $\bullet=\mathrm{un}$ is the unknown-amplitude experiment and $\bullet=\mathrm{or}$ reveals the amplitude vector to the policy. For $\alpha\in(0,1/2)$, define the adaptive least pathwise budget

$$B_\bullet^{K,\mathrm{pair}}(\Delta,\alpha):=\min\{B\in\mathbb{N}:\mathfrak{R}_{B,\bullet}^{K,\mathrm{pair}}(\Delta,\eta)\leq\alpha\}.\tag{243}$$

Recall that $z_\alpha=\Phi^{-1}(1-\alpha)$.

Corollary 3 (Sharp adaptive budget)**.** *For fixed K, r, η and $\Delta \downarrow 0$,*

$$B_{\mathrm{un}}^{K,\mathrm{pair}}(\Delta,\alpha) \sim \frac{4\sigma^2\kappa(r)z_\alpha^2}{\Delta^2}, \qquad B_{\mathrm{or}}^{K,\mathrm{pair}}(\Delta,\alpha) \sim \frac{4\sigma^2 z_\alpha^2}{\Delta^2}. \tag{244}$$

Consequently, the unknown-amplitude price remains exactly $\kappa(r)$, and relative to the optimized deterministic-allocation constant $2K\sigma^2\kappa(r)z_\alpha^2/\Delta^2$ of Theorem 9, unrestricted adaptive screening improves the first-order budget by the factor $K/2$. This proves the budget statement (24) *of Theorem 3.*

Proof. We spell out the inversion so that no varying-c version of Theorem 10 is assumed. Fix $\epsilon \in (0, z_\alpha)$ and put

$$c_- = 2\sigma\sqrt{\kappa(r)}(z_\alpha - \epsilon), \qquad B_-(\Delta) = \left\lfloor \frac{c_-^2}{\Delta^2} \right\rfloor,$$
$$c_+ = 2\sigma\sqrt{\kappa(r)}(z_\alpha + \epsilon), \qquad B_+(\Delta) = \left\lceil \frac{c_+^2}{\Delta^2} \right\rceil.$$

With unit cap, Theorem 10 implies convergence of the finite-budget minimax risk at each fixed local signal c: a policy sequence can be selected horizon by horizon. Since $\Delta \le c_-/\sqrt{B_-(\Delta)}$, monotonicity in the required gap gives

$$\begin{aligned}\mathfrak{R}_{B_-(\Delta),\mathrm{un}}^{K,\mathrm{pair}}(\Delta,\eta) &\ge \mathfrak{R}_{B_-(\Delta),\mathrm{un}}^{K,\mathrm{pair}}\{c_-/\sqrt{B_-(\Delta)},\eta\} \\ &\longrightarrow \Phi\{-(z_\alpha-\epsilon)\} > \alpha.\end{aligned}$$

Similarly, $\Delta \ge c_+/\sqrt{B_+(\Delta)}$, so

$$\begin{aligned}\mathfrak{R}_{B_+(\Delta),\mathrm{un}}^{K,\mathrm{pair}}(\Delta,\eta) &\le \mathfrak{R}_{B_+(\Delta),\mathrm{un}}^{K,\mathrm{pair}}\{c_+/\sqrt{B_+(\Delta)},\eta\} \\ &\longrightarrow \Phi\{-(z_\alpha+\epsilon)\} < \alpha.\end{aligned}$$

The risk is nonincreasing in the available budget because a policy may discard extra pulls. Hence, for all sufficiently small Δ,

$$B_-(\Delta) < B_{\mathrm{un}}^{K,\mathrm{pair}}(\Delta,\alpha) \le B_+(\Delta).$$

Divide by $4\sigma^2\kappa(r)z_\alpha^2/\Delta^2$ and then send $\epsilon \downarrow 0$. The oracle proof is identical after replacing $\kappa(r)$ by one. Dividing by the deterministic constant gives $K/2$. □

G.4 Why simultaneous near ties remain open

The theorem uses the fixed margin η in an essential way: $o(T)$ pilot samples identify the only relevant pair with probability tending uniformly to one. If all K arms are separated only at order $T^{-1/2}$, no such screening is possible, and the scalar two-arm bracket argument meets a genuine multivariate obstruction, which we now make precise.

After projecting off armwise amplitude nuisances and a common intercept, a deterministic terminal fraction vector p has effective contrast information

$$\mathcal{I}_\chi(p) = \frac{1}{\sigma^2\chi}\left\{\mathrm{diag}(p) - \frac{pp^\top}{\sum_i p_i}\right\}, \qquad \chi = \kappa(r) \text{ below the boundary}, \quad \chi = 1 \text{ at or above it.} \tag{245}$$

For balance $p^{\mathrm{bal}} = (C/K)\mathbf{1}$ and the feasible allocation $p^{12} = (C/2, C/2, 0, \dots, 0)$, take the genuine centered contrasts $\mathbf{v} = e_1 - e_2$ and $\mathbf{w} = e_3 - K^{-1}\mathbf{1}$. Then, for $K \ge 3$,

$$\mathbf{v}^\top\mathcal{I}_\chi(p^{12})\mathbf{v} = \frac{C}{\sigma^2\chi} > \frac{2C}{K\sigma^2\chi} = \mathbf{v}^\top\mathcal{I}_\chi(p^{\mathrm{bal}})\mathbf{v},$$
$$\mathbf{w}^\top\mathcal{I}_\chi(p^{12})\mathbf{w} = 0 < \frac{C}{K\sigma^2\chi}\left(1 - \frac{1}{K}\right) = \mathbf{w}^\top\mathcal{I}_\chi(p^{\mathrm{bal}})\mathbf{w}.$$

Thus the two Gaussian contrast experiments are incomparable in Loewner order: balanced information does not Blackwell-dominate every pathwise allocation, and not every pathwise allocation dominates balance. There is no analogue of the two-arm proof that completes a single scalar random bracket to a universally dominating balanced Gaussian experiment. An exact unrestricted simultaneous-near-tie theorem would require a genuinely multi-hypothesis least-favorable prior or a dynamic Gaussian-experiment comparison.

Remark 3 (Scope)**.** Theorem 10 is an exact unrestricted K-arm result under arbitrary predictable allocation and pathwise-bounded stopping. Its scope is fixed K, Gaussian noise, known r and σ, parameter-free policies, a fixed separation η for all nonchallengers, and a pathwise rather than expected budget. It does not resolve simultaneous root-T near ties, random active sets, or expected-budget stopping.

H Two-Arm Identification under Stable Predictable Allocation

This appendix states and proves Theorem 11, the two-arm specialization to predictable policies whose allocation proportion stabilizes, and Proposition 4, a data-dependent policy that attains the corresponding value. The argument uses the likelihood of the full adaptive experiment; in particular, it never conditions on random terminal counts. Throughout, put $x_s = s^{-r}$ and recall

$$a_n := \sum_{s=1}^{n} x_s, \qquad b_n := \sum_{s=1}^{n} x_s^2, \qquad D_n := n - \frac{a_n^2}{b_n}. \tag{246}$$

At round $t \le T$, $A_t \in \{1,2\}$ is measurable with respect to the past $\mathcal{F}_{t-1}$, and the observed local index is

$$s_t := 1 + \sum_{u<t} \mathbf{1}\{A_u = A_t\}, \qquad Y_t = \mu_{A_t} + \beta_{A_t} x_{s_t} + \varepsilon_t, \tag{247}$$

where, conditionally on $\mathcal{F}_{t-1}$, $\varepsilon_t \sim N(0,\sigma^2)$. Internal randomization is included in $\mathcal{F}_0$ and is independent of the noises. This representation allows arbitrary predictable interleaving while preserving the prefix structure of each arm.

H.1 Statements

At each round $t \le T$, a policy chooses $A_t \in \{1,2\}$ predictably and observes the next unused point of that arm. Write $N_i(T) = \sum_{t\le T} \mathbf{1}\{A_t = i\}$. Under the equal-arm central law, call a policy sequence p-stabilizing when $N_1(T)/T \to p$ in probability. For a policy π_T and the local classes $\mathcal{P}_\pm(\Delta)$ of Section 4, let $\mathcal{R}_T(\pi_T,\Delta)$ denote the composite risk of the best recommendation based on the data generated by π_T, that is, the infimum over recommendation rules ψ of $\max\{\sup_{\mathcal{P}_+(\Delta)} \mathbb{P}_\theta^{\pi_T}(\psi = 2), \sup_{\mathcal{P}_-(\Delta)} \mathbb{P}_\theta^{\pi_T}(\psi = 1)\}$.

Theorem 11 (Two-arm specialization under predictable stable allocation)**.** *For every predictable p-stabilizing policy and $\Delta_T = c/\sqrt{T}$,*

$$\liminf_{T\to\infty} \mathcal{R}_T(\pi_T,\Delta_T) \ge \Phi\left(-\frac{c\sqrt{p(1-p)}}{\sigma\sqrt{\kappa(r)}}\right). \tag{248}$$

For every $p \in (0,1)$, deterministic proportion-p sampling followed by the OLS-intercept sign attains equality. Over the class $\mathfrak{A}_{\mathrm{stab}}$ of predictable policies having a deterministic stable proportion,

$$\inf_{\boldsymbol{\pi}\in\mathfrak{A}_{\mathrm{stab}}} \liminf_{T\to\infty} \mathcal{R}_T(\pi_T,\Delta_T) = \Phi\left(-\frac{c}{2\sigma\sqrt{\kappa(r)}}\right). \tag{249}$$

The oracle statements hold with $\kappa(r)$ replaced by one. For target error $\alpha \in (0,1/2)$, the allocation-optimized local constants are

$$\mathsf{T}_{\mathrm{un}} = 4\sigma^2\kappa(r) z_\alpha^2, \qquad \mathsf{T}_{\mathrm{or}} = 4\sigma^2 z_\alpha^2, \qquad \mathsf{T}_{\mathrm{un}}/\mathsf{T}_{\mathrm{or}} = \kappa(r), \quad z_\alpha = \Phi^{-1}(1-\alpha). \tag{250}$$

Proposition 4 (Random counts with a deterministic information limit)**.** *Fix $p \in (0,1)$ and $1 \le k_T = o(T)$. Pull each arm twice, set*

$$\widehat{\Delta}_T^{\mathrm{pil}} = \frac{Y_{2,1} + Y_{2,2} - Y_{1,1} - Y_{1,2}}{2}, \qquad \mathsf{q}_T = \mathrm{clip}\{\lfloor pT\rfloor + k_T \operatorname{sign}(\widehat{\Delta}_T^{\mathrm{pil}}), 2, T-2\}, \tag{251}$$

give arm 1 exactly q_T pulls and arm 2 the remaining $T - \mathsf{q}_T$ pulls in any predictable order, and recommend the smaller full-prefix OLS intercept. The policy is p-stabilizing, each of its two possible count pairs occurs with central probability $1/2$, and

$$\mathcal{R}_T(\pi_T^{\mathrm{pilot}}, c/\sqrt{T}) \longrightarrow \Phi\left(-\frac{c\sqrt{p(1-p)}}{\sigma\sqrt{\kappa(r)}}\right). \tag{252}$$

Revealed-amplitude corrected means replace $\kappa(r)$ by one.

H.2 RANDOM-PREFIX DESIGN LEMMA

For $\varsigma \geq 0$ and $0 \leq n \leq T$, define

$$\mathcal{E}_n(\varsigma) := \sum_{s=1}^{n}(1-\varsigma x_s)^2 = n - 2\varsigma a_n + \varsigma^2 b_n, \qquad \mathcal{E}_0(\varsigma) := 0. \tag{253}$$

Lemma 15 (Random-prefix least-favorable direction)**.** *Fix $p \in (0,1]$, let $m_T/T \to p$, and set $\bar{\varsigma}_{m_T} = a_{m_T}/b_{m_T}$. Then*

$$\bar{\varsigma}_{m_T} = o(\sqrt{T}), \tag{254}$$

$$\sup_{0\leq n\leq T} \mathcal{E}_n(\bar{\varsigma}_{m_T}) = O(T), \tag{255}$$

$$\frac{\mathcal{E}_{n_T}(\bar{\varsigma}_{m_T})}{T} \longrightarrow \frac{p}{\kappa(r)} \quad \textit{whenever } n_T/T \to p. \tag{256}$$

If integer-valued random variables N_T satisfy $N_T/T \to p$ in probability under a sequence of laws, then

$$\frac{\mathcal{E}_{N_T}(\bar{\varsigma}_{m_T})}{T} \longrightarrow \frac{p}{\kappa(r)} \quad \textit{in probability}. \tag{257}$$

Proof. The standard power-sum estimates give

$$\bar{\varsigma}_{m_T} = \begin{cases} \dfrac{1-2r}{1-r}(pT)^r(1+o(1)), & 0<r<1/2, \\ \dfrac{2\sqrt{pT}}{\log(pT)}(1+o(1)), & r=1/2, \\ O(T^{1-r}), & 1/2<r<1, \\ O(\log T), & r=1, \\ O(1), & r>1. \end{cases} \tag{258}$$

This proves (254). Since a_n, b_n increase,

$$\mathcal{E}_n(\bar{\varsigma}_{m_T}) \leq T + 2\bar{\varsigma}_{m_T} a_T + \bar{\varsigma}_{m_T}^2 b_T = O(T),$$

where the final bound follows by substituting the same power-sum estimates. At $n = m_T$, $\mathcal{E}_{m_T}(\bar{\varsigma}_{m_T}) = D_{m_T}$ and $D_{m_T}/m_T \to 1/\kappa(r)$. If $n_T/m_T \to 1$, regular variation gives

$$\mathcal{E}_{n_T}(\bar{\varsigma}_{m_T}) - \mathcal{E}_{m_T}(\bar{\varsigma}_{m_T}) = o(T).$$

For completeness, when $r < 1/2$ each of $n_T - m_T$, $\bar{\varsigma}_{m_T}(a_{n_T} - a_{m_T})$, and $\bar{\varsigma}_{m_T}^2(b_{n_T} - b_{m_T})$ is $o(T)$. At $r = 1/2$, the latter two full terms are already $O(T/\log T)$; for $r > 1/2$ they are $o(T)$. This proves (256).

The deterministic sequential limit is locally uniform. Indeed, if it were not, there would be $\epsilon > 0$, $T_k \to \infty$, and deterministic integers n_k with $n_k/T_k \to p$ but

$$\left| \frac{\mathcal{E}_{n_k}(\bar{\varsigma}_{m_{T_k}})}{T_k} - \frac{p}{\kappa(r)} \right| > \epsilon.$$

Completing n_k off this subsequence by $\lfloor pT \rfloor$ would produce a deterministic sequence contradicting (256). Equivalently,

$$\lim_{\delta\downarrow 0} \limsup_{T\to\infty} \sup_{\substack{0\leq n\leq T \\ |n/T-p|\leq\delta}} \left| \frac{\mathcal{E}_n(\bar{\varsigma}_{m_T})}{T} - \frac{p}{\kappa(r)} \right| = 0.$$

For any fixed $\delta > 0$, $N_T/T \to p$ in probability makes the event $\{|N_T/T - p| > \delta\}$ negligible. Combining this with the local-uniform display and then sending $\delta \downarrow 0$ proves (257) without placing the changing laws on a common probability space. □

H.3 Adaptive lower bound and attainment

Proof of Theorem 11. We first prove the lower bound for $p \in (0,1)$. Write $p_1 = p$, $p_2 = 1-p$, choose deterministic integers $m_{i,T}/T \to p_i$, and let

$$\bar{\varsigma}_{i,T} := \frac{a_{m_{i,T}}}{b_{m_{i,T}}}.$$

For $\Delta_T = c/\sqrt{T}$, define intercept and amplitude differences

$$\begin{aligned} \delta^{\mu}_{1,T} &:= \frac{2c(1-p)}{\sqrt{T}}, & \delta^{\beta}_{1,T} &:= -\delta^{\mu}_{1,T}\bar{\varsigma}_{1,T}, \\ \delta^{\mu}_{2,T} &:= -\frac{2cp}{\sqrt{T}}, & \delta^{\beta}_{2,T} &:= -\delta^{\mu}_{2,T}\bar{\varsigma}_{2,T}. \end{aligned} \tag{259}$$

Center two parameter points at (θ_0, θ_0):

$$\begin{aligned} \mu^{+}_{i,T} &= \mu_0 - \delta^{\mu}_{i,T}/2, & \beta^{+}_{i,T} &= \beta_0 - \delta^{\beta}_{i,T}/2, \\ \mu^{-}_{i,T} &= \mu_0 + \delta^{\mu}_{i,T}/2, & \beta^{-}_{i,T} &= \beta_0 + \delta^{\beta}_{i,T}/2. \end{aligned} \tag{260}$$

Because $\delta^{\mu}_{1,T} - \delta^{\mu}_{2,T} = 2\Delta_T$, the plus and minus points have opposing asymptote gaps of exactly Δ_T. Lemma 15 gives $\delta^{\mu}_{i,T} \to 0$ and $\delta^{\beta}_{i,T} = o(1)$, so both points lie in the prescribed interior product ball for all sufficiently large T.

Let $\mathbb{P}_{+,T}$ and $\mathbb{P}_{-,T}$ be the laws of the full adaptive experiment at these points. When arm i is pulled for the sth time, the difference between its conditional means under the minus and plus laws is the predictable quantity

$$\delta_{i,s,T} = \delta^{\mu}_{i,T} + \delta^{\beta}_{i,T} x_s = \delta^{\mu}_{i,T}(1 - \bar{\varsigma}_{i,T} x_s). \tag{261}$$

We first transfer allocation stability from the central law $\mathbb{P}_0$ to the two local laws. Lemma 15 gives, uniformly over every possible realized count pair,

$$\sum_{i=1}^{2} \sum_{s=1}^{N_i(T)} \delta^2_{i,s,T} = O(1). \tag{262}$$

For $L_{+,0} = d\mathbb{P}_{+,T}/d\mathbb{P}_0$, define the predictable standardized shift

$$g_t := -\frac{\delta_{A_t,s_t,T}}{2\sigma}, \qquad Z^0_t := \frac{Y_t - \mu_0 - \beta_0 x_{s_t}}{\sigma}.$$

Because the central law is the midpoint of the plus and minus points, the adaptive Gaussian likelihood is

$$L_{+,0} = \exp\left\{ \sum_{t=1}^{T} g_t Z^0_t - \frac{1}{2} \sum_{t=1}^{T} g^2_t \right\}.$$

Equation (262) bounds $\sum_t g^2_t \le \mathsf{C}$ deterministically. Consequently,

$$\begin{aligned} \mathbb{E}_0 L^2_{+,0} &= \mathbb{E}_0 \left[\exp\left\{ 2\sum_t g_t Z^0_t - 2\sum_t g^2_t \right\} \exp\left\{ \sum_t g^2_t \right\} \right] \\ &\le e^{\mathsf{C}}. \end{aligned} \tag{263}$$

The first exponential is a mean-one predictable Gaussian exponential martingale. The same calculation holds for the minus law. Cauchy–Schwarz therefore gives contiguity of each local law with respect to the central law. The assumed stability implies, under either local law,

$$N_1(T)/T \to p, \qquad N_2(T)/T \to 1-p \quad \text{in probability.} \tag{264}$$

Under $\mathbb{P}_{+,T}$, the log likelihood ratio between the two simple hypotheses is

$$\Lambda_T := \log \frac{d\mathbb{P}_{-,T}}{d\mathbb{P}_{+,T}} = \mathcal{M}_T - \frac{1}{2} V_T, \qquad \mathcal{M}_T := \sum_{t=1}^{T} \frac{\delta_{A_t,s_t,T}}{\sigma} Z^+_t, \tag{265}$$

where

$$Z_t^+ := \frac{Y_t - \mu^+_{A_t,T} - \beta^+_{A_t,T} x_{s_t}}{\sigma}$$

are conditionally standard Gaussian under $\mathbb{P}_{+,T}$, and

$$V_T = \langle \mathcal{M} \rangle_T = \frac{1}{\sigma^2} \sum_{i=1}^{2} (\delta^\mu_{i,T})^2 \mathcal{E}_{N_i(T)}(\bar{\varsigma}_{i,T}). \tag{266}$$

By (264), the random-index conclusion of Lemma 15, and (259),

$$V_T \longrightarrow V_p^\star := \frac{4c^2 p(1-p)}{\sigma^2 \kappa(r)} \quad \text{in } \mathbb{P}_{+,T}\text{-probability.} \tag{267}$$

Moreover,

$$\max_{i,s} |\delta_{i,s,T}| \le \max_i |\delta^\mu_{i,T}|(1 + \bar{\varsigma}_{i,T}) = o(1),$$

while (262) supplies a deterministic $O(1)$ envelope for the total squared coefficients. These two facts and Gaussian tails give the conditional martingale Lindeberg condition. The martingale central limit theorem yields

$$\Lambda_T \Rightarrow N(-V_p^\star/2, V_p^\star) \quad \text{under } \mathbb{P}_{+,T}. \tag{268}$$

The equal-prior Bayes error for these two simple laws is

$$e_T^\star = \frac{1}{2} \mathbb{E}_{+,T}[\min\{1, e^{\Lambda_T}\}].$$

For $X \sim N(-V/2, V)$, Gaussian tilting gives $\mathbb{E}[\min\{1, e^X\}] = 2\Phi(-\sqrt{V}/2)$. The integrand is bounded and continuous, so (268) implies

$$e_T^\star \longrightarrow \Phi\left(-\frac{c\sqrt{p(1-p)}}{\sigma\sqrt{\kappa(r)}} \right). \tag{269}$$

Every recommendation's maximum composite error is at least its average error on this two-point subproblem, and hence at least $e_T^\star$. This proves the lower bound for $p \in (0,1)$. If $p = 0$, take $\delta^\mu_{1,T} = 2c/\sqrt{T}$ and all other δ^μ's and δ^β's zero. The midpoint likelihood energy is uniformly bounded because $N_1(T) \le T$, so the same L^2 calculation gives contiguity of both local laws with respect to $\mathbb{P}_0$. Central stability therefore transfers to $N_1(T)/T \to 0$ under the local laws. Under the plus law,

$$V_T = \frac{4c^2 N_1(T)}{\sigma^2 T} \longrightarrow 0 \quad \text{in probability,} \qquad 0 \le V_T \le \frac{4c^2}{\sigma^2}.$$

Boundedness gives $\mathbb{E}_{+,T} V_T \to 0$, and the martingale isometry then gives $\mathcal{M}_T \to 0$ in $L^2(\mathbb{P}_{+,T})$. Hence $\Lambda_T \to 0$ in probability and the Bayes error tends to $1/2$. The case $p = 1$ is symmetric, using $N_2(T)/T \to 0$.

Since $p(1-p) \le 1/4$, the optimized stable-proportion lower bound is attained only at $p = 1/2$. Deterministic balanced sampling followed by the sign of the OLS intercept difference attains it because its exact risk is

$$\Phi\left(-\frac{\Delta_T \sqrt{J(n_1, n_2)}}{\sigma} \right), \qquad J(n_1, n_2) = \left(D_{n_1}^{-1} + D_{n_2}^{-1} \right)^{-1},$$

and $J(\lfloor T/2 \rfloor, \lceil T/2 \rceil)/T \to 1/(4\kappa(r))$. More generally, deterministic counts with $n_1/T \to p$ give $J(n_1, n_2)/T \to p(1-p)/\kappa(r)$, proving the stated p-specific attainment.

For the revealed-amplitude oracle, an oracle policy is a measurable family $\{\pi_T(\beta_1, \beta_2)\}$ whose actions are predictable after the revealed pair is included in $\mathcal{F}_0$; risk is still taken over the same product parameter classes. Its p-stabilization is evaluated under the central experiment with the revealed pair $(\beta_1, \beta_2) = (\beta_0, \beta_0)$. For the lower bound, restrict to this valid equal-amplitude subproblem and repeat the construction with $\delta^\beta_{1,T} = \delta^\beta_{2,T} = 0$. Its limiting likelihood variance is $4c^2 p(1-p)/\sigma^2$, proving the oracle lower bound. Subtracting the revealed transients and comparing corrected sample means gives attainment.

Finally, the proof is unchanged for deterministic $c_T = \sqrt{T}\Delta_T \to c \in (0, \infty)$: all feasibility and likelihood-energy bounds are uniform on compact sets of c_T. A limiting error at most α therefore requires

$$c^2 \geq \frac{\sigma^2 \kappa(r) z_\alpha^2}{p(1-p)} \geq 4\sigma^2 \kappa(r) z_\alpha^2, \qquad z_\alpha = \Phi^{-1}(1-\alpha),$$

where proportion-p OLS attains the first constant and balance minimizes it at $p = 1/2$. The revealed-amplitude calculation deletes $\kappa(r)$, giving the exact local budget ratio in Theorem 11. □

H.4 A GENUINELY DATA-DEPENDENT ATTAINING POLICY

We first record a nested-prefix identity that controls the dependence created by using the pilot both for the quota and in the final OLS fit.

For $n \geq 2$, write the full-prefix OLS intercept as

$$\widehat{\mu}_n = \sum_{s=1}^{n} w_{n,s} Y_s, \qquad w_{n,s} = \frac{1 - (a_n/b_n)x_s}{D_n}. \tag{270}$$

Lemma 16 (Exact nested-prefix comparison)**.** *Let $m \geq n \geq 2$ and pad w_n with zeros to length m. Then*

$$w_n^\top \mathbf{1}_n = 1, \qquad w_n^\top \boldsymbol{x}_n = 0, \qquad \|w_n\|^2 = \frac{1}{D_n}, \tag{271}$$

and

$$\|w_m - \mathrm{pad}_m(w_n)\|^2 = \frac{1}{D_n} - \frac{1}{D_m}. \tag{272}$$

Consequently, if $m_T/T \to q > 0$, $n_T/T \to q$, and $m_T \geq n_T$, then

$$\sqrt{T}(\widehat{\mu}_{m_T} - \widehat{\mu}_{n_T}) \longrightarrow 0 \quad \textit{in } L^2, \tag{273}$$

uniformly over the intercept and amplitude parameters.

Proof. The constraints in (271) follow directly from (270); the squared norm is the OLS variance factor. Moreover,

$$w_m^\top \mathrm{pad}_m(w_n) = \frac{1}{D_m}\left\{ w_n^\top \mathbf{1}_n - \frac{a_m}{b_m} w_n^\top \boldsymbol{x}_n \right\} = \frac{1}{D_m}.$$

Expanding the squared norm gives

$$\|w_m - \mathrm{pad}_m(w_n)\|^2 = \frac{1}{D_m} + \frac{1}{D_n} - \frac{2}{D_m},$$

which is (272). Both OLS intercepts have mean μ for every (μ, β), so their difference is a linear combination of the noise alone. Its variance is σ^2 times the right side of (272). Finally, $D_k/k \to 1/\kappa(r)$ and $m_T/T, n_T/T \to q$ imply $T/D_{n_T} - T/D_{m_T} \to 0$. This proves (273); parameter-uniformity follows because the difference distribution does not depend on (μ, β). □

Proof of Proposition 4. Because $k_T/T \to 0$ and $p \in (0,1)$, clipping is inactive for all sufficiently large T. Both quotas then exceed two, so the four-pull pilot followed by deficit tracking is a valid predictable policy, and

$$\frac{N_1(T)}{T} \to p, \qquad \frac{N_2(T)}{T} \to 1 - p$$

pathwise. Under the central law the two arm curves are identical, so $\widehat{\Delta}_T^{\mathrm{pil}}$ is a centered continuous Gaussian variable. Therefore $\mathbb{P}_0(\widehat{\Delta}_T^{\mathrm{pil}} > 0) = \mathbb{P}_0(\widehat{\Delta}_T^{\mathrm{pil}} < 0) = 1/2$, proving the claimed nondegeneracy of the terminal counts.

Put $n_{1,T}^0 = \lfloor pT \rfloor$ and $n_{2,T}^0 = T - n_{1,T}^0$. On a common infinite observation sequence for each arm, let $\widehat{\mu}_{i,n}$ be its OLS intercept after n local observations. The random-count estimates are one of the two deterministic branches with counts

$$n_{1,T}^{\pm} = n_{1,T}^0 \pm k_T, \qquad n_{2,T}^{\pm} = n_{2,T}^0 \mp k_T.$$

For each arm and each branch, Lemma 16, applied with the smaller and larger of $n^{\pm}_{i,T}, n^{0}_{i,T}$, yields

$$\sqrt{T}\{\widehat{\mu}_{i,n^{\pm}_{i,T}} - \widehat{\mu}_{i,n^{0}_{i,T}}\} \longrightarrow 0 \quad \text{in } L^2.$$

The distribution of each difference is independent of all arm parameters. Irrespective of the parameter-dependent pilot sign,

$$\begin{aligned}\sqrt{T}\left|\{\widehat{\mu}_{2,N_2(T)} - \widehat{\mu}_{1,N_1(T)}\} - \{\widehat{\mu}_{2,n^{0}_{2,T}} - \widehat{\mu}_{1,n^{0}_{1,T}}\}\right| \\ \leq \sum_{i=1}^{2} \max_{\pm} \sqrt{T}\left|\widehat{\mu}_{i,n^{\pm}_{i,T}} - \widehat{\mu}_{i,n^{0}_{i,T}}\right| \longrightarrow 0 \end{aligned} \tag{274}$$

in probability, uniformly over the product parameter ball.

The centered deterministic-count difference on the second line of (274) is Gaussian with variance

$$\sigma^2\{D^{-1}_{n^{0}_{1,T}} + D^{-1}_{n^{0}_{2,T}}\}.$$

Consequently,

$$\sqrt{T}\left[\{\widehat{\mu}_{2,n^{0}_{2,T}} - \widehat{\mu}_{1,n^{0}_{1,T}}\} - (\mu_2 - \mu_1)\right] \Rightarrow N\left(0, \frac{\sigma^2\kappa(r)}{p(1-p)}\right). \tag{275}$$

Let $\widehat{\Delta}^{\mathrm{ad}}_T$ and $\widehat{\Delta}^{0}_T$ denote the random-count and baseline versions of $\widehat{\mu}_2 - \widehat{\mu}_1$, and put $R_T = \sqrt{T}(\widehat{\Delta}^{\mathrm{ad}}_T - \widehat{\Delta}^{0}_T)$. Equation (274) gives $\sup_{\vartheta} \mathbb{P}_{\vartheta}(|R_T| > \epsilon) \to 0$ for every $\epsilon > 0$. If arm 1 is best with gap $\Delta_{12} = \mu_2 - \mu_1 \geq c/\sqrt{T}$, then

$$\mathbb{P}_{\vartheta}(\widehat{\Delta}^{\mathrm{ad}}_T \leq 0) \leq \mathbb{P}_{\vartheta}\Big\{\sqrt{T}(\widehat{\Delta}^{0}_T - \Delta_{12}) \leq -c + \epsilon\Big\} + \mathbb{P}_{\vartheta}(|R_T| > \epsilon).$$

At a feasible boundary point with $\Delta_{12} = c/\sqrt{T}$, the reverse sandwich is

$$\mathbb{P}_{\vartheta}(\widehat{\Delta}^{\mathrm{ad}}_T \leq 0) \geq \mathbb{P}_{\vartheta}\Big\{\sqrt{T}(\widehat{\Delta}^{0}_T - \Delta_{12}) \leq -c - \epsilon\Big\} - \mathbb{P}_{\vartheta}(|R_T| > \epsilon).$$

The baseline normal density on the root-T scale is uniformly bounded, so the probability of the intervening 2ϵ strip is $O(\epsilon)$. The same inequalities hold after reversing the arm labels. Taking $T \to \infty$ and then $\epsilon \downarrow 0$ in (275) yields

$$\mathcal{R}_T(\pi^{\mathrm{pilot}}_T, c/\sqrt{T}) \longrightarrow \Phi\left(-\frac{c\sqrt{p(1-p)}}{\sigma\sqrt{\kappa(r)}}\right).$$

For revealed amplitudes, replace w_n by the corrected-mean weights $v_n = n^{-1}\mathbf{1}_n$. Their nested-prefix identity is $\|v_m - \mathrm{pad}_m(v_n)\|^2 = 1/n - 1/m$ for $m \geq n$. The same two-branch comparison gives (274), while the deterministic baseline variance is $\sigma^2\{(n^{0}_{1,T})^{-1} + (n^{0}_{2,T})^{-1}\}$. This proves the oracle limit with $\kappa(r)$ replaced by one. □

Why this proof stops here. If $N_1(T)/T$ has a nondegenerate random limit, the nuisance direction a_{pT}/b_{pT} cannot be chosen after that limit is realized: a parameter alternative must be fixed before the data. A least-favorable nuisance prior or a local minimax theorem with random information would be needed. At a data-dependent stopping time, both the horizon and local normalization are random, and an expected-budget constraint does not supply the pathwise quadratic-variation bound used above. Theorem 11 therefore makes no claim by itself for random limiting proportions or optional stopping. Appendix I supplies the separate ancillary pilot-quota result and explains why the marginal count law does not cover fully endogenous random information.

I Ancillary Random Information: Sharp Extension and Boundary

This appendix extends Theorem 11 from a deterministic limiting proportion to a nondegenerate random one. Two conditions are needed: the random information must be asymptotically ancillary, and the full local score must be conditionally Gaussian given the limiting proportion. We verify both

for a data-dependent $o(T)$-pilot policy that chooses a macroscopic quota. We then give an analytic counterexample showing that the marginal law of the terminal proportion alone does not determine the local experiment, and an explicit symmetric pair for which the deterministic $\kappa(r)$-rescaling fails under endogenous stopping.

Throughout, $0 < r < 1/2$, the central parameter is an interior equal-arm point (μ_0, β_0), and

$$\mu_{i,T} = \mu_0 + \frac{u_i}{\sqrt{T}}, \qquad \beta_{i,T} = \beta_0 + v_i T^{r-1/2}, \qquad \gamma_i = (u_i, v_i)^\top. \tag{276}$$

For $p \in (0, 1]$, define

$$Q_r(p) := \begin{pmatrix} p & A_r(p) \\ A_r(p) & B_r(p) \end{pmatrix}, \qquad A_r(p) := \frac{p^{1-r}}{1-r}, \qquad B_r(p) := \frac{p^{1-2r}}{1-2r}, \tag{277}$$

$$\mathcal{Q}_r(p) := \operatorname{diag}\{Q_r(p), Q_r(1-p)\}, \tag{278}$$

and note that the intercept Schur complement is

$$D_r(p) := p - \frac{A_r(p)^2}{B_r(p)} = \frac{p}{\kappa(r)}, \qquad \kappa(r) = \frac{(1-r)^2}{r^2}. \tag{279}$$

I.1 Sharp ancillary-mixture theorem

Assumption 1 (Ancillary-mixture LAN)**.** For some $\iota > 0$, the observed proportion $P_T = N_1(T)/T$ belongs to $[\iota, 1-\iota]$. There are observed $X_T \in \mathbb{R}^4$ and a probability law ν on this interval such that, centrally,

$$(P_T, X_T) \Rightarrow (P, X), \qquad P \sim \nu, \qquad X \mid P = p \sim N(0, \sigma^2 \mathcal{Q}_r(p)), \tag{280}$$

and, uniformly for $\gamma = (\gamma_1, \gamma_2)$ in compact subsets of $\mathbb{R}^4$,

$$\log \frac{d\mathbb{P}_{T,\gamma}}{d\mathbb{P}_{T,0}} = \frac{1}{\sigma^2}\gamma^\top X_T - \frac{1}{2\sigma^2}\gamma^\top \mathcal{Q}_r(P_T)\gamma + o_{\mathbb{P}_{T,0}}(1). \tag{281}$$

Equivalently, the joint central convergence includes

$$\mathbb{E}_0[f(P_T)e^{it^\top X_T}] \longrightarrow \mathbb{E}\Big[f(P)e^{-\sigma^2 t^\top \mathcal{Q}_r(P)t/2}\Big] \tag{282}$$

for every bounded continuous f and $t \in \mathbb{R}^4$.

The conditional normalizer in (281) has expectation one given P, so Le Cam's third lemma leaves $P \sim \nu$ under every bounded local parameter: the mixing variable is ancillary rather than locally tilted.

Theorem 12 (Sharp identification with ancillary random information)**.** *Under Assumption 1, the local asymptotic minimax probability of incorrect identification over*

$$\Gamma_+(c) = \{\gamma : u_2 - u_1 \geq c\}, \qquad \Gamma_-(c) = \{\gamma : u_1 - u_2 \geq c\} \tag{283}$$

is

$$R_{\mathrm{un}}(c; \nu) = \int \Phi\left(-\frac{c\sqrt{p(1-p)}}{\sigma\sqrt{\kappa(r)}}\right) \nu(dp). \tag{284}$$

More precisely, every sequence of recommendations satisfies

$$\lim_{M\to\infty} \liminf_{T\to\infty} \max_{\pm} \sup_{\gamma \in \Gamma_\pm(c) \cap [-M,M]^4} \mathbb{P}_{T,\gamma}(\text{incorrect}) \geq R_{\mathrm{un}}(c; \nu), \tag{285}$$

and the efficient intercept contrast attains equality.

If the amplitudes are revealed, the exact value is

$$R_{\mathrm{or}}(c; \nu) = \int \Phi\left(-\frac{c\sqrt{p(1-p)}}{\sigma}\right) \nu(dp). \tag{286}$$

Consequently

$$R_{\mathrm{un}}(c; \nu) = R_{\mathrm{or}}(c/\sqrt{\kappa(r)}; \nu). \tag{287}$$

For every target error in $(0, 1/2)$, the critical local signals of the two experiments differ by the factor $\sqrt{\kappa(r)}$, so the corresponding first-order budget ratio is $\kappa(r)$. Deterministic balance is uniquely optimal inside this ancillary-mixture class.

Proof. Conditional on $P = p$, the limit experiment has independent coordinates

$$X_i \sim N(Q_r(p_i)\gamma_i, \sigma^2 Q_r(p_i)), \qquad p_1 = p, \quad p_2 = 1 - p. \tag{288}$$

Define

$$X_i^{\perp} = X_{i,1} - \frac{A_r(p_i)}{B_r(p_i)} X_{i,2}, \qquad Z_i = \frac{X_i^{\perp}}{D_r(p_i)}. \tag{289}$$

Gaussian projection and (279) give

$$Z_i \mid P = p \sim N\left(u_i, \frac{\sigma^2 \kappa(r)}{p_i}\right), \tag{290}$$

independently of the corresponding nuisance-score coordinate. Hence

$$Z_2 - Z_1 \mid P = p \sim N\left(u_2 - u_1, \frac{\sigma^2 \kappa(r)}{p(1-p)}\right). \tag{291}$$

The sign rule recommends arm 1 iff $Z_2 - Z_1 > 0$. Its maximum conditional error is attained at gap magnitude c and equals the integrand in (284).

For the matching lower bound we use proper nuisance priors, avoiding any appeal to an invariant-rule compactness theorem. Write

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad d = \frac{c}{2}(-1, 0, 1, 0)^{\top}.$$

The boundary parameters can be written $\gamma = \mathsf{s}d + G\xi$, where $\mathsf{s} \in \{-1, +1\}$ is the best-arm label and $\xi \in \mathbb{R}^3$ contains the common intercept and the two slopes. Put equal prior mass on $\mathsf{s} = \pm 1$ and, independently, let $\xi \sim N(0, a^2 I_3)$.

Fix $P = p$, abbreviate $Q = \mathcal{Q}_r(p)$, and whiten the observation as $\widetilde{X} = Q^{-1/2}X$. After integrating out ξ,

$$\widetilde{X} \mid (P = p, \mathsf{s}) \sim N\Big(\mathsf{s}Q^{1/2}d, \Sigma_a(p)\Big), \qquad \Sigma_a(p) = \sigma^2 I_4 + a^2 Q^{1/2} G G^{\top} Q^{1/2}.$$

For two normals with means $\pm q$ and common covariance Σ, the equal-prior Bayes error is

$$\Phi\Big\{-\sqrt{q^{\top}\Sigma^{-1}q}\Big\}.$$

Thus the Bayes error of this proper prior is

$$e_a = \int \Phi\{-\sqrt{I_a(p)}\}\, \nu(dp), \qquad I_a(p) = d^{\top} Q^{1/2} \Sigma_a(p)^{-1} Q^{1/2} d. \tag{292}$$

The matrix $Q^{1/2}G$ has rank three. Hence, as $a \to \infty$, $\Sigma_a(p)^{-1} \to \sigma^{-2}(I - \mathrm{proj}_p)$, where proj_p is the orthogonal projector onto $\mathrm{col}(Q^{1/2}G)$. Equivalently, first projecting off the two slope scores leaves intercept informations $D_r(p) = p/\kappa(r)$ and $D_r(1-p) = (1-p)/\kappa(r)$; projecting off the common intercept then yields

$$I_a(p) \longrightarrow \frac{c^2}{\sigma^2} \frac{D_r(p) D_r(1-p)}{D_r(p) + D_r(1-p)} = \frac{c^2 p(1-p)}{\sigma^2 \kappa(r)}. \tag{293}$$

Dominated convergence in (292) therefore gives $e_a \to R_{\mathrm{un}}(c; \nu)$. Since maximum risk is at least Bayes risk for every a, this proves the lower bound in the limit experiment.

We finally transfer the bound to the original experiments. For fixed a, truncate the $N(0, a^2 I_3)$ prior to $[-L, L]^3$. The induced Bayes risk tends to e_a as $L \to \infty$, because truncation changes the joint prior–experiment law by at most the discarded prior mass. On this compact cube, quantize ξ on a finite grid. Uniformly in $p \in [\iota, 1 - \iota]$, Pinsker's inequality gives

$$\big\| N(Q(\mathsf{s}d + G\xi), \sigma^2 Q) - N(Q(\mathsf{s}d + G\xi'), \sigma^2 Q) \big\|_{\mathrm{TV}} \le \mathsf{C} \|\xi - \xi'\|,$$

so the finite-grid Bayes risks converge to the truncated-prior risk as the mesh vanishes.

For each fixed finite grid, the compact-uniform likelihood expansion in (281) and the central convergence imply joint convergence of its finite likelihood-ratio vector. Each limiting likelihood ratio has expectation one conditionally on P, hence also unconditionally; the nonnegative likelihood-ratio vectors are uniformly integrable. The finite-prior Bayes error can be written as one minus the expectation of the maximum of two finite weighted likelihood-ratio sums, so its value converges. Every rule's maximum risk over $\Gamma_+(c) \cup \Gamma_-(c)$ is at least this finite-prior Bayes risk. Taking first the grid mesh to zero, then $L \to \infty$, and finally $a \to \infty$ proves (285). This is the finite-subexperiment form of the local asymptotic minimax argument (Le Cam, 1986).

With revealed amplitudes, subtracting their contributions leaves independent intercept coordinates with informations p/σ^2 and $(1-p)/\sigma^2$. Repeating the preceding proper-prior argument with only the common-intercept nuisance gives (286) and (287). Finally $p(1-p) \le 1/4$, with equality only at $p = 1/2$; strict monotonicity of $\Phi(-x)$ proves the balance claim. □

I.2 A DATA-DEPENDENT CLASS SATISFYING THE ASSUMPTION

Theorem 13 ($o(T)$-pilot macroscopic random quota)**.** *Let $m_T = o(T)$ and fix $\iota \in (0, 1/2)$. During the first m_T global rounds, run an arbitrary parameter-free predictable pilot, and let $\mathcal{G}_T$ contain its observations and independent internal randomization. Let an integer quota q_T be $\mathcal{G}_T$-measurable and satisfy*

$$\iota T \le \mathsf{q}_T \le (1-\iota)T, \qquad \mathsf{q}_T \ge N_{1,T}^{\mathrm{pil}}, \qquad T - \mathsf{q}_T \ge N_{2,T}^{\mathrm{pil}}, \tag{294}$$

where $N_{i,T}^{\mathrm{pil}}$ is the arm-i pilot count. After the pilot, ignore all bulk outcomes for allocation and fill arm 1 to q_T and arm 2 to $T - \mathsf{q}_T$. If, centrally,

$$\mathsf{q}_T/T \Rightarrow P \sim \nu, \tag{295}$$

then Assumption 1 holds with mixing law ν. The full-prefix OLS recommendation attains (284)*.*

Proof. For an arm prefix of length $k \le m_T$, the squared mean displacement in a bounded local direction (u, v) is bounded by

$$\begin{aligned} &\sum_{s=1}^{k} \left\{ \frac{u}{\sqrt{T}} + vT^{r-1/2}s^{-r} \right\}^2 \\ &\qquad \le \mathsf{C} \left\{ \frac{k}{T} + \left(\frac{k}{T}\right)^{1-r} + \left(\frac{k}{T}\right)^{1-2r} \right\} = o(1), \end{aligned} \tag{296}$$

uniformly for $k \le m_T$ and bounded (u, v). By the Gaussian chain rule, the Kullback–Leibler divergence between the local and central pilot laws is at most the same quantity divided by $2\sigma^2$, even when the pilot allocation is predictable. Pinsker's inequality gives

$$\mathrm{TV}\{\mathrm{law}_\gamma(\mathcal{G}_T), \mathrm{law}_0(\mathcal{G}_T)\} = o(1) \tag{297}$$

uniformly on compact local sets. Thus q_T/T has the same asymptotic law ν under every bounded local parameter.

Under the central law, define

$$X_{i,T} = \frac{1}{\sqrt{T}} \sum_{s=1}^{N_i(T)} \varepsilon_{i,s} \begin{pmatrix} 1 \\ (s/T)^{-r} \end{pmatrix}. \tag{298}$$

The pilot part has covariance $o(1)$ by (296). Conditional on $\mathcal{G}_T$, the bulk noises are independent centered Gaussians. Uniformly for $p \in [\iota, 1-\iota]$,

$$\frac{1}{T} \sum_{s=1}^{\lfloor pT \rfloor} \begin{pmatrix} 1 & (s/T)^{-r} \\ (s/T)^{-r} & (s/T)^{-2r} \end{pmatrix} \longrightarrow Q_r(p). \tag{299}$$

Removing an $o(T)$ pilot prefix changes this matrix by $o(1)$, by the same power-sum estimates as (296). Conditioning on $\mathcal{G}_T$ proves (282).

For Gaussian observations and a parameter-free predictable action kernel, action probabilities cancel from the likelihood ratio. Substitution of (276) gives exactly

$$\log \frac{d\mathbb{P}_{T,\gamma}}{d\mathbb{P}_{T,0}} = \frac{1}{\sigma^2}\gamma^\top X_T - \frac{1}{2\sigma^2}\gamma^\top \mathcal{Q}_{r,T}(P_T)\gamma, \tag{300}$$

where $\mathcal{Q}_{r,T}$ is the realized discrete Gram matrix. Uniform convergence (299) proves (281). The bracket is uniformly bounded on compact local sets, while the largest incremental displacement is $O(T^{-1/2} + T^{r-1/2}) = o(1)$. Writing a likelihood ratio squared as a mean-one Gaussian exponential martingale times the exponential of its bracket gives a uniform second-moment bound. Applying the analogous bound to the inverse likelihood under the local law gives mutual contiguity.

It remains to verify OLS attainment. Put

$$a_n = \sum_{s=1}^{n} s^{-r}, \qquad b_n = \sum_{s=1}^{n} s^{-2r}, \qquad D_n = n - a_n^2/b_n. \tag{301}$$

For deterministic n, the OLS intercept weights and variance are

$$w_{n,s} = \frac{1 - (a_n/b_n)s^{-r}}{D_n}, \qquad \mathrm{Var}\{\widehat{\mu}(n)\} = \frac{\sigma^2}{D_n}. \tag{302}$$

Uniformly for $n/T \in [\iota, 1-\iota]$,

$$D_n/n \to 1/\kappa(r), \qquad a_n/b_n \le \mathsf{C}T^r. \tag{303}$$

For every $k \le m_T$,

$$\begin{aligned} T\sum_{s=1}^{k} w_{n,s}^2 &\le \frac{\mathsf{C}}{T}\{k + T^r k^{1-r} + T^{2r}k^{1-2r}\} \\ &\le \mathsf{C}'\left\{\frac{k}{T} + \left(\frac{k}{T}\right)^{1-r} + \left(\frac{k}{T}\right)^{1-2r}\right\} = o(1). \end{aligned} \tag{304}$$

Thus the pilot contribution to the $\sqrt{T}$-scaled OLS error is $o_{L^2}(1)$. Conditional on P_T, the bulk contribution is Gaussian and independent of the pilot. Therefore

$$\sqrt{T}\{(\widehat{\mu}_2 - \widehat{\mu}_1) - (\mu_2 - \mu_1)\} \Rightarrow N\left(0, \frac{\sigma^2\kappa(r)}{P(1-P)}\right) \quad \text{conditionally on } P. \tag{305}$$

The sign error at local gap c is (284); larger gaps only reduce it. Theorem 12 supplies the matching lower bound. □

For a concrete nondegenerate example, take one initial observation from each arm and choose the quota $\underline{p}T$ or $\overline{p}T$, $0 < \underline{p} < \overline{p} < 1$, according to the sign of their difference; then complete that quota. At the equal-arm central point the quotas have probability $1/2$, and the two-observation pilot is asymptotically ancillary along (276).

I.3 Boundary: the marginal law of P is not enough

The next proposition sharpens Proposition 3 by giving both limiting errors in closed form.

Proposition 5 (Same central proportion law, different local error)**.** *There are two predictable fixed-horizon policies whose terminal arm-1 proportions have the same nondegenerate central limiting law but whose optimal local two-point identification errors are different. Consequently, convergence of P_T stated without a joint score condition does not determine the local experiment.*

Proof. Reveal both amplitudes, fix $0 < \underline{p} < \overline{p} < 1$ and $t_0 > 0$, and center at a fixed, prespecified equal-arm reference point. Pull arm 1 through $\underline{p}T$. If its normalized centered partial sum $X_T(\underline{p})$ satisfies $|X_T(\underline{p})| > t_0\sqrt{\underline{p}}$, stop pulling arm 1; otherwise continue it through $\overline{p}T$. Fill every remaining round with arm 2. Along a subsequence for which $\underline{p}T, \overline{p}T$ are integers, this is predictable. Centrally,

$$P_T \Rightarrow P, \qquad P = \begin{cases} \underline{p}, & |\mathsf{B}_{\underline{p}}| > t_0\sqrt{\underline{p}}, \\ \overline{p}, & |\mathsf{B}_{\underline{p}}| \le t_0\sqrt{\underline{p}}, \end{cases} \qquad \mathbb{P}(P = \underline{p}) = w := 2\Phi(-t_0), \tag{306}$$

where B is standard Brownian motion. Compare this policy with an independent $\mathcal{F}_0$ coin choosing deterministic fraction $\underline{p}$ with probability w and $\overline{p}$ otherwise. The central limiting laws of P_T are identical.

Consider

$$\mathrm{H}_+ : \ \mu_1 = -c/\sqrt{T},\ \mu_2 = 0, \qquad \mathrm{H}_- : \ \mu_1 = +c/\sqrt{T},\ \mu_2 = 0, \tag{307}$$

and put $\bar{c} = c/\sigma$. The arm-2 data are uninformative. If Φ_2 is the standard bivariate Gaussian cdf with correlation $\sqrt{\underline{p}/\overline{p}}$, the optimal stopped likelihood test has limiting error

$$\begin{aligned} e_{\mathrm{stop}}(\bar{c}) = \Phi(-t_0 - \bar{c}\sqrt{\underline{p}}) \\ + \Phi_2(t_0 - \bar{c}\sqrt{\underline{p}}, -\bar{c}\sqrt{\overline{p}}) - \Phi_2(-t_0 - \bar{c}\sqrt{\underline{p}}, -\bar{c}\sqrt{\overline{p}}), \end{aligned} \tag{308}$$

whereas the independent mixture has error

$$e_{\mathrm{mix}}(\bar{c}) = w\Phi(-\bar{c}\sqrt{\underline{p}}) + (1-w)\Phi(-\bar{c}\sqrt{\overline{p}}). \tag{309}$$

The two functions differ analytically. Under the midpoint law, the stopped experiment has exact Bayes error

$$e(\bar{c}) = \frac{1}{2}\mathbb{E}_0 \exp\{-\bar{c}^2 P/2 - \bar{c}|\mathsf{B}_P|\} = \frac{1}{2} - \frac{\bar{c}}{2}\mathbb{E}_0|\mathsf{B}_P| + O(\bar{c}^2). \tag{310}$$

Put $\ell = \overline{p} - \underline{p}$ and

$$\delta_\ell(x) := \mathbb{E}|x + \sqrt{\ell}Z| - |x|, \qquad Z \sim N(0,1). \tag{311}$$

For $x > 0$,

$$\delta_\ell'(x) = -2\Phi(-x/\sqrt{\ell}) < 0, \tag{312}$$

so δ_ℓ strictly decreases with $|x|$. The stopped policy continues precisely on the central set containing the smallest $|\mathsf{B}_{\underline{p}}|$ values with probability $1-w$. The strict rearrangement inequality gives

$$\begin{aligned} \mathbb{E}|\mathsf{B}_P|_{\mathrm{stop}} &= \mathbb{E}|\mathsf{B}_{\underline{p}}| + \mathbb{E}[\delta_\ell(\mathsf{B}_{\underline{p}})\mathbf{1}\{|\mathsf{B}_{\underline{p}}| \le t_0\sqrt{\underline{p}}\}] \\ &> \mathbb{E}|\mathsf{B}_{\underline{p}}| + (1-w)\mathbb{E}\delta_\ell(\mathsf{B}_{\underline{p}}) = \mathbb{E}|\mathsf{B}_P|_{\mathrm{mix}}. \end{aligned} \tag{313}$$

Together with (310), this proves $e_{\mathrm{stop}}(\bar{c}) < e_{\mathrm{mix}}(\bar{c})$ for every sufficiently small $\bar{c} > 0$, although the central laws of P coincide. □

I.4 A SYMMETRIC TWO-POINT OBSTRUCTION TO UNIVERSAL κ-RESCALING

The same stopped allocation also shows that the deterministic $\kappa(r)$-rescaling need not hold for every symmetric two-point subexperiment when the limiting information is endogenous. This is a boundary result about one fixed allocation and explicit pairs, not a characterization of the full adaptive minimax value, which Theorem 2 settles.

Divide all observations by σ, so that $\sigma = 1$ and c denotes the standardized local signal throughout this subsection. For $0 < r < 1/2$, let

$$f_\varsigma(u) = 1 + \varsigma u^{-r}, \qquad \mathcal{M}_\varsigma = \int_0^P f_\varsigma(u)\, d\mathsf{B}_u, \qquad V_\varsigma = \int_0^P f_\varsigma(u)^2\, du. \tag{314}$$

The corresponding discrete symmetric endpoints have intercept shifts $\pm c/\sqrt{T}$ and amplitude shifts $\pm c\varsigma T^{r-1/2}$, so they remain in every fixed interior ball. Their exact equal-prior limiting error is

$$e_\varsigma(c) = \frac{1}{2}\mathbb{E}e^{-c^2 V_\varsigma/2 - c|\mathcal{M}_\varsigma|} = \frac{1}{2} - \frac{c}{2}\mathfrak{m}(\varsigma) + O(c^2), \qquad \mathfrak{m}(\varsigma) := \mathbb{E}|\mathcal{M}_\varsigma|. \tag{315}$$

With A_r and B_r as in (277), let

$$k_\varsigma = 1 + \varsigma A_r(\underline{p})/\underline{p}. \tag{316}$$

Conditional on $\mathsf{B}_{\underline{p}} = x$, $\mathcal{M}_\varsigma$ is normal with mean $k_\varsigma x$. Its variance is

$$\begin{aligned} \sigma_{\mathrm{s}}^2 &= \varsigma^2(B_r(\underline{p}) - A_r(\underline{p})^2/\underline{p}) && \text{on } |x| > t_0\sqrt{\underline{p}}, \\ \sigma_{\mathrm{c}}^2 &= \overline{p} - \underline{p} + 2\varsigma(A_r(\overline{p}) - A_r(\underline{p})) + \varsigma^2(B_r(\overline{p}) - A_r(\underline{p})^2/\underline{p}) && \text{on } |x| \le t_0\sqrt{\underline{p}}. \end{aligned} \tag{317}$$

Writing

$$g(m,v)=\sqrt{\frac{2v}{\pi}}e^{-m^2/(2v)}+|m|\{2\Phi(|m|/\sqrt{v})-1\}, \tag{318}$$

with $g(m,0)=|m|$, gives

$$\begin{aligned}\mathfrak{m}(\varsigma)=&\int_{|x|>t_0\sqrt{\underline{p}}} g(k_\varsigma x,\sigma_{\mathrm{s}}^2)\varphi_{\underline{p}}(x)\,dx\\&+\int_{|x|\le t_0\sqrt{\underline{p}}} g(k_\varsigma x,\sigma_{\mathrm{c}}^2)\varphi_{\underline{p}}(x)\,dx,\end{aligned} \tag{319}$$

where $\varphi_{\underline{p}}$ is the $N(0,\underline{p})$ density.

Now set $r=1/4$, $\underline{p}=1/4$, $\overline{p}=3/4$, $t_0=1$, and choose

$$\varsigma_\star=-\frac{\underline{p}}{A_r(\underline{p})}=-\frac{3\sqrt{2}}{8}. \tag{320}$$

Then $k_{\varsigma_\star}=0$, $\sigma_{\mathrm{s}}^2=1/32$, and

$$\sigma_{\mathrm{c}}^2=\frac{3}{4}-\frac{3}{2}\,3^{-1/4}+\frac{9\sqrt{3}}{32}=0.0973857611\ldots. \tag{321}$$

With $w=2\Phi(-1)$ and $q=1-w$,

$$\mathfrak{m}(\varsigma_\star)=\sqrt{\frac{2}{\pi}}\{w\sigma_{\mathrm{s}}+q\sigma_{\mathrm{c}}\}<0.21475. \tag{322}$$

For a finite lower certificate on $\mathfrak{m}(0)$, put $U=\mathsf{B}_{\underline{p}}/\sqrt{\underline{p}}\sim N(0,1)$ and $\delta_{1/2}(x)=g(x,1/2)-|x|$. By (312), this function decreases with $|x|$. Partitioning $[0,1]$ into 20 equal bins and using the right endpoint in each bin gives

$$\begin{aligned}\mathfrak{m}(0)&\ge\frac{1}{\sqrt{2\pi}}+2\sum_{j=1}^{20}\{\Phi(j/20)-\Phi((j-1)/20)\}\,\delta_{1/2}(j/40)\\&>0.6481>3(0.21475)>3\mathfrak{m}(\varsigma_\star).\end{aligned} \tag{323}$$

The conservative finite-expression bounds in (322)–(323) have a margin exceeding 3.8×10^{-3}.

For target error $1/2-\epsilon$, define the critical signal of the fixed ς-pair by

$$c_\varsigma(\epsilon):=\inf\{c>0:e_\varsigma(c)\le 1/2-\epsilon\}. \tag{324}$$

As $\epsilon\downarrow 0$, the critical signals of the symmetric pairs $\varsigma=0$ and $\varsigma=\varsigma_\star$ satisfy

$$c_0(\epsilon)=\frac{2\epsilon}{\mathfrak{m}(0)}+o(\epsilon),\qquad c_{\varsigma_\star}(\epsilon)=\frac{2\epsilon}{\mathfrak{m}(\varsigma_\star)}+o(\epsilon). \tag{325}$$

Consequently their squared critical-signal ratio obeys

$$\left\{\frac{c_{\varsigma_\star}(\epsilon)}{c_0(\epsilon)}\right\}^2=\left\{\frac{\mathfrak{m}(0)}{\mathfrak{m}(\varsigma_\star)}\right\}^2+o(1)>9=\kappa(1/4). \tag{326}$$

Thus κ-rescaling already fails for this explicit symmetric pair under the endogenous stopped allocation. This statement does not identify the full-composite minimax risk of that allocation and does not show that an endogenous random-proportion policy beats deterministic balanced sampling.

J Adaptive-Scope Extensions

This appendix records four secondary extensions that delimit the scope of the main results: the joint score–information limit experiment under endogenous allocation and bounded stopping (Theorem 14); fixed-K local minimax values under predictable stable allocation, below, at, and above $r=1/2$

(Theorem 15); sharp fixed-gap fixed-confidence asymptotics as $\delta \downarrow 0$ (Theorem 16); and the fixed-gap fixed-budget exponent (Proposition 6). The regimes are deliberately kept separate. Throughout, K is fixed whenever a K-arm limit is taken, and (μ_0, β_0) is an interior central point, so that every displayed local parameter class lies in one prescribed interior product neighborhood for all sufficiently large T.

For an arm prefix of length n, write

$$a_n = \sum_{s=1}^{n} s^{-r}, \qquad b_n = \sum_{s=1}^{n} s^{-2r}, \qquad D_n = n - \frac{a_n^2}{b_n}. \tag{327}$$

For $n \geq 2$, $D_n > 0$, and the full-prefix OLS intercept satisfies

$$\widehat{\mu}_i(n) - \mu_i = \frac{\sum_{s \leq n} \{1 - (a_n/b_n)s^{-r}\}\varepsilon_{i,s}}{D_n}. \tag{328}$$

If β_i is revealed, the corrected mean is

$$\widetilde{\mu}_i(n) = \overline{Y}_{i,n} - \beta_i a_n/n, \qquad \widetilde{\mu}_i(n) - \mu_i = n^{-1} \sum_{s \leq n} \varepsilon_{i,s}. \tag{329}$$

For $0 < r < 1/2$ and $p \geq 0$, put

$$g_r(x) = \begin{pmatrix} 1 \\ x^{-r} \end{pmatrix}, \qquad Q_r(p) = \int_0^p g_r(x) g_r(x)^\top dx = \begin{pmatrix} p & A_r(p) \\ A_r(p) & B_r(p) \end{pmatrix}, \tag{330}$$

where

$$A_r(p) = \frac{p^{1-r}}{1-r}, \qquad B_r(p) = \frac{p^{1-2r}}{1-2r}, \qquad D_r(p) = p - \frac{A_r(p)^2}{B_r(p)} = \frac{p}{\kappa(r)}. \tag{331}$$

Here $\kappa(r) = (1-r)^2/r^2$ below the boundary and $\kappa(r) = 1$ for $r \geq 1/2$.

We first record the decision-theoretic transfer used throughout. Its continuity assumption is the reason why the endogenous theorem is stated for compact separated ranking templates rather than for arbitrary discontinuous losses.

Lemma 17 (Compact likelihood-process minimax transfer)**.** *Let Γ be a compact metric parameter set, let $\mathcal{A}$ be finite, and let $\ell : \mathcal{A} \times \Gamma \to [0, \bar{\ell}]$ be continuous in its second argument. For each T, let a dominated experiment have central law $P_{T,0}$ and a sufficient likelihood process $\Lambda_T = (L_{T,\gamma} : \gamma \in \Gamma)$ taking values in $C(\Gamma)$. Suppose centrally*

$$\Lambda_T \Rightarrow \Lambda \quad \text{in } C(\Gamma), \qquad \sup_T \mathbb{E}_0 \|\Lambda_T\|_\infty^q < \infty \quad \text{for some } q > 1, \tag{332}$$

and suppose the limit coordinates satisfy $L_\gamma \geq 0$ and $\mathbb{E}L_\gamma = 1$. Then the minimax risks converge to the minimax risk of the dominated limit experiment with central law $\mathrm{law}(\Lambda)$ and densities L_γ.

Proof. By sufficiency, represent the finite experiment on $C(\Gamma)$ with central law $\nu_T = \mathrm{law}(\Lambda_T)$ and coordinate density $f \mapsto f(\gamma)$. For a prior Π on Γ, its Bayes value is

$$\mathfrak{B}_T(\Pi) = \int_{C(\Gamma)} \min_{\mathsf{a} \in \mathcal{A}} \int_\Gamma \ell(\mathsf{a}, \gamma) f(\gamma)\, \Pi(d\gamma)\, \nu_T(df). \tag{333}$$

The integrand is continuous in f and $\bar{\ell}$-Lipschitz in the supremum norm, uniformly over Π. Weak convergence and the moment bound in (332) therefore give $\sup_\Pi |\mathfrak{B}_T(\Pi) - \mathfrak{B}(\Pi)| \to 0$. Compactness of Γ, finiteness of $\mathcal{A}$, and continuity of the bounded loss give the elementary minimax identity

$$\inf_\delta \sup_{\gamma \in \Gamma} R_T(\delta, \gamma) = \sup_\Pi \mathfrak{B}_T(\Pi), \qquad \text{the supremum over all priors } \Pi \text{ on } \Gamma,$$

and the analogous identity in the limit experiment. One may obtain it directly by finite-dimensional minimax on finitely supported priors and finite partitions of $C(\Gamma)$ and then refine the partitions; the envelope $\bar{\ell}\|f\|_\infty$ controls the refinement. Taking suprema over priors proves the claim. □

J.1 Endogenous score–information limits under bounded stopping

Return to $0 < r < 1/2$ and use the full local path

$$\mu_{i,T} = \mu_0 + u_i/\sqrt{T}, \qquad \beta_{i,T} = \beta_0 + v_i T^{r-1/2}, \qquad \gamma_i = (u_i, v_i)^\top. \tag{334}$$

Let τ_T be a parameter-free stopping time satisfying the pathwise cap

$$\tau_T \le CT \quad \text{almost surely} \tag{335}$$

for fixed $C < \infty$. Put

$$P_{i,T} = N_i(\tau_T)/T, \qquad X_{i,T} = \frac{1}{\sqrt{T}} \sum_{s=1}^{N_i(\tau_T)} \varepsilon_{i,s} g_r(s/T),$$
$$\mathcal{Q}_r(P) = \mathrm{diag}\{Q_r(P_1), \dots, Q_r(P_K)\}. \tag{336}$$

Here $\sum_i P_{i,T} = \tau_T/T \le C$, and Q_r in (330) is used on $[0, C]$ with $Q_r(0) = 0$.

Theorem 14 (Endogenous score–information LAQ transfer)**.** *Under the central law, (P_T, X_T) is tight. Along every subsequence on which*

$$(P_T, X_T) \Rightarrow (P, X), \tag{337}$$

the local likelihood processes converge compact-uniformly to the dominated experiment with central law $\varrho = \mathrm{law}(P, X)$ and densities

$$L_\gamma(P, X) = \exp\left\{ \frac{1}{\sigma^2} \sum_i \gamma_i^\top X_i - \frac{1}{2\sigma^2} \sum_i \gamma_i^\top Q_r(P_i) \gamma_i \right\}. \tag{338}$$

In particular, $\mathbb{E}_\varrho L_\gamma = 1$, and (P_T, X_T) is sufficient for the finite local experiment. For every compact local parameter set Γ, finite action space, and bounded loss continuous on Γ, the minimax value converges along that subsequence to

$$\inf_\delta \sup_{\gamma \in \Gamma} \mathbb{E}_\varrho \{ L_\gamma(P, X) \ell(\delta(P, X), \gamma) \}. \tag{339}$$

This includes 0–1 ranking loss when Γ is a finite union of pairwise disjoint compact label components. If the joint limit is unique, the characterization holds along the whole sequence.

Proof. Stopping and allocation kernels are the same measurable functions under all parameters and cancel from the likelihood. The exact Gaussian likelihood is

$$L_{T,\gamma} = \exp\left\{ \frac{1}{\sigma^2} \gamma^\top X_T - \frac{1}{2\sigma^2} \gamma^\top \mathcal{Q}_{r,T}(P_T) \gamma \right\}, \tag{340}$$

where $\mathcal{Q}_{r,T}(P) = \mathrm{diag}(Q_{r,T}(P_1), \dots, Q_{r,T}(P_K))$ and

$$\sup_{0 \le q \le C} \|Q_{r,T}(q) - Q_r(q)\| \to 0, \qquad \sup_{T,\, q \le C} \|Q_{r,T}(q)\| < \infty. \tag{341}$$

Martingale isometry gives $\sup_T \mathbb{E}_0 \|X_T\|^2 < \infty$, while P_T lies in a compact set, proving tightness. On a subsequence satisfying (337), continuous mapping gives the candidate (338).

For $q > 1$, write

$$L_{T,\gamma}^q = \exp\left\{ \frac{q}{\sigma^2} \gamma^\top X_T - \frac{q^2}{2\sigma^2} \gamma^\top \mathcal{Q}_{r,T}(P_T) \gamma \right\}$$
$$\times \exp\left\{ \frac{q^2 - q}{2\sigma^2} \gamma^\top \mathcal{Q}_{r,T}(P_T) \gamma \right\}. \tag{342}$$

The first factor is a mean-one bounded-stopping Gaussian exponential martingale, and the second is uniformly bounded on compact γ sets. Thus the likelihoods are uniformly integrable and $\mathbb{E}_\varrho L_\gamma = 1$.

The exact log likelihood is affine-quadratic in γ and depends only on (P_T, X_T), so it converges in $C(\Gamma)$. The same exponential-martingale bound makes X_T uniformly sub-Gaussian in every fixed direction; a finite net of the unit sphere then gives a uniform exponential moment for $\|X_T\|$. Hence the moment condition of Lemma 17 holds on each compact Γ, and that lemma gives (339). Finally, (340) is a measurable function of (P_T, X_T), so the factorization criterion proves finite sufficiency. □

Let $\nu = \mathrm{law}(P)$ and let $\mathcal{L}_p$ be a regular conditional law of X given $P = p$. Under local parameter γ, the limiting allocation law is

$$\nu_\gamma(dp) = \mathsf{w}_\gamma(p)\nu(dp), \qquad \mathsf{w}_\gamma(p) = e^{-\gamma^\top \mathcal{Q}_r(p)\gamma/(2\sigma^2)} \int e^{\gamma^\top x/\sigma^2} \mathcal{L}_p(dx). \tag{343}$$

Corollary 4 (Exact ancillarity boundary)**.** *Assume $P_i \geq \iota > 0$ almost surely. The limiting allocation is ancillary under every full local intercept-and-amplitude direction γ if and only if*

$$X \mid P = p \sim N\{0, \sigma^2 \mathcal{Q}_r(p)\} \quad \textit{for } \nu\textit{-almost every } p. \tag{344}$$

Otherwise (343) *is the parameter-dependent tilt that a count-only description omits.*

Proof. Conditional normality makes the integral in (343) the inverse of its prefactor. Conversely, ancillarity gives $\mathsf{w}_\gamma(p) = 1$ for every γ. Intersect the full-measure sets for a countable dense set of directions. The exponential-norm moment inherited from the bounded likelihood bracket makes the conditional moment-generating function finite on a common neighborhood of zero. Continuity therefore extends the identity to that neighborhood, and uniqueness of moment- generating functions gives (344). □

Corollary 5 (Exact endogenous symmetric-pair error)**.** *For endpoints $\gamma^+ = g/2$ and $\gamma^- = -g/2$, the exact equal-prior Bayes error in the endogenous limit experiment is*

$$e_\varrho(g) = \frac{1}{2}\mathbb{E}_\varrho \exp\left\{-\frac{g^\top \mathcal{Q}_r(P)g}{8\sigma^2} - \frac{|g^\top X|}{2\sigma^2}\right\}. \tag{345}$$

This remains valid when P is nonancillary and when the bounded horizon is random.

Proof. Relative to ϱ, the endpoint densities are

$$L_\pm = \exp\left\{\pm\frac{g^\top X}{2\sigma^2} - \frac{g^\top \mathcal{Q}_r(P)g}{8\sigma^2}\right\}.$$

The equal-prior error is $\frac{1}{2}\mathbb{E}_\varrho \min(L_+, L_-)$, which is (345). □

The pathwise cap in (335) can be replaced by a uniform exponential-moment condition on every compact-set likelihood bracket together with tightness of the normalized information. An expected-budget condition $\mathbb{E}\tau_T \leq T$ alone supplies neither property and is not covered by the theorem.

J.2 Predictable stable allocation for fixed K

Fix $K \geq 2$ and a compact labeled ranking template $\mathcal{H} = \bigcup_{b=1}^K \mathcal{H}_b \subset \mathbb{R}^K$. Assume the nonempty compact components are pairwise disjoint and

$$h_j - h_b \geq 1 \qquad (h \in \mathcal{H}_b,\ j \neq b). \tag{346}$$

For $\boldsymbol{\upsilon} = (\upsilon_1, \ldots, \upsilon_K)$ with positive entries, define

$$\mathcal{V}_{\mathcal{H}}(t; \boldsymbol{\upsilon}) = \inf_\delta \max_b \sup_{h \in \mathcal{H}_b} \mathbb{P}_{Z \sim N(th, \mathrm{diag}(\boldsymbol{\upsilon}))}\{\delta(Z) \neq b\}. \tag{347}$$

Let a parameter-free predictable policy π satisfy, under the central equal-arm law,

$$(N_1(T)/T, \ldots, N_K(T)/T) \longrightarrow p \quad \text{in probability}, \qquad p_i > 0, \quad \sum_i p_i = 1. \tag{348}$$

Below the boundary use

$$\mu_{i,T} = \mu_0 + ch_i/\sqrt{T}, \qquad \beta_{i,T} = \beta_0 + v_i T^{r-1/2}, \tag{349}$$

and let $\mathcal{R}^{\pi,<}_{T,M}(c, \mathcal{H})$ be the minimax misranking probability over $h \in \mathcal{H}_b$ and $v \in [-M, M]^K$. At $r = 1/2$ replace the amplitude path by

$$\beta_{i,T} = \beta_0 + v_i/\sqrt{\log T} \tag{350}$$

and denote the analogous risk by $\mathcal{R}^{\pi,=}_{T,M}(c, \mathcal{H})$. Above the boundary, let $\mathcal{B}$ be any fixed compact interval contained in the interior of the amplitude parameter space and containing β_0; only the intercept is root-T local, while $\beta \in \mathcal{B}^K$. Denote this risk by $\mathcal{R}^{\pi,>}_{T,\mathcal{B}}(c, \mathcal{H})$.

Lemma 18 (Predictable score LAN below the boundary)**.** *For $0 < r < 1/2$, define*

$$X_{i,T} = \frac{1}{\sqrt{T}} \sum_{s=1}^{N_i(T)} \varepsilon_{i,s} g_r(s/T). \tag{351}$$

Under (348)*, centrally,*

$$X_T \Rightarrow N\{0, \sigma^2 \operatorname{diag}(Q_r(p_1), \dots, Q_r(p_K))\}. \tag{352}$$

Writing $\gamma_i = (u_i, v_i)^\top$, the exact likelihood ratio for $\mu_i = \mu_0 + u_i/\sqrt{T}$ and $\beta_i = \beta_0 + v_i T^{r-1/2}$ is

$$\log \frac{d\mathbb{P}_{T,\gamma}}{d\mathbb{P}_{T,0}} = \frac{1}{\sigma^2} \sum_i \gamma_i^\top X_{i,T} - \frac{1}{2\sigma^2} \sum_i \gamma_i^\top Q_{r,T}(N_i(T)/T)\gamma_i, \tag{353}$$

where

$$Q_{r,T}(q) = \frac{1}{T} \sum_{s=1}^{\lfloor qT \rfloor} g_r(s/T) g_r(s/T)^\top, \qquad \sup_{0 \le q \le 1} \|Q_{r,T}(q) - Q_r(q)\| \to 0. \tag{354}$$

Thus the convergence is compact-uniform LAN with deterministic information $\sigma^{-2} \operatorname{diag}(Q_r(p_i))$.

Proof. The parameter-free action kernels cancel from the likelihood, and expanding the Gaussian densities gives (353). The entries of $Q_{r,T}$ are cumulative Riemann sums for $1, x^{-r}, x^{-2r}$. Cut at $\iota > 0$: convergence is ordinary and uniform on $[\iota, 1]$, while both the integral and discrete mass on $[0, \iota]$ are $O(\iota^{1-2r})$. This proves (354).

In global time, X_T is a $2K$-dimensional martingale whose predictable bracket is the covariance in (352) with Q_r replaced by $Q_{r,T}(N_i(T)/T)$. It converges in probability by (348). The largest coefficient of any innovation is $O(T^{r-1/2}) = o(1)$ and the bracket is uniformly bounded. Gaussian tails give conditional Lindeberg, so the martingale CLT proves (352). The exact likelihood and uniform Gram convergence yield compact-uniform LAN. □

Theorem 15 (Adaptive fixed-K local minimax values)**.** *Under* (348)*, for $0 < r < 1/2$,*

$$\lim_{M \to \infty} \lim_{T \to \infty} \mathcal{R}^{\pi,<}_{T,M}(c, \mathcal{H}) = \mathcal{V}_{\mathcal{H}} \left(\frac{c}{\sigma \sqrt{\kappa(r)}}; \left(\frac{1}{p_i} \right)_{i=1}^K \right). \tag{355}$$

At $r = 1/2$, for every fixed M,

$$\lim_{T \to \infty} \mathcal{R}^{\pi,=}_{T,M}(c, \mathcal{H}) = \mathcal{V}_{\mathcal{H}} \left(\frac{c}{\sigma}; \left(\frac{1}{p_i} \right)_{i=1}^K \right), \tag{356}$$

and the equality remains true after $M \to \infty$. For $r > 1/2$ and each fixed compact $\mathcal{B}$ as above,

$$\lim_{T \to \infty} \mathcal{R}^{\pi,>}_{T,\mathcal{B}}(c, \mathcal{H}) = \mathcal{V}_{\mathcal{H}} \left(\frac{c}{\sigma}; \left(\frac{1}{p_i} \right)_{i=1}^K \right). \tag{357}$$

The values are attained by a Gaussian minimax kernel applied to the vector of random-prefix OLS intercepts, with an arbitrary fallback if any $N_i(T) < 2$.

Proof. First take $0 < r < 1/2$. Lemmas 18 and 17 give, on every fixed compact (h, v) set, the independent limit coordinates

$$X_i \sim N\{Q_r(p_i)(ch_i, v_i)^\top, \sigma^2 Q_r(p_i)\}. \tag{358}$$

Put

$$Z_i = \frac{X_{i,1} - A_r(p_i) X_{i,2}/B_r(p_i)}{D_r(p_i)}, \qquad \widetilde{Z}_i = \frac{X_{i,2}}{B_r(p_i)}, \qquad \varsigma_r(p_i) = \frac{A_r(p_i)}{B_r(p_i)}. \tag{359}$$

Then, independently across arms,

$$Z_i \sim N\left(ch_i, \frac{\sigma^2 \kappa(r)}{p_i} \right), \quad \widetilde{Z}_i \sim N\left(v_i + \varsigma_r(p_i) ch_i, \frac{\sigma^2}{B_r(p_i)} \right), \quad Z_i \perp \widetilde{Z}_i. \tag{360}$$

Ignoring $\widetilde{Z}$ gives the upper bound in (355).

For the matching lower bound, let R_Z be the right side of (355) and fix $\epsilon > 0$. Compact Gaussian minimax equality supplies a proper prior Π on $\mathcal{H}$ whose Bayes error is at least $R_Z - \epsilon$. Given $h \sim \Pi$, draw independent $U_i \sim N(0, a^2)$ and set

$$v_i = -\varsigma_r(p_i)ch_i + U_i. \tag{361}$$

After integrating over U, $\widetilde{Z}$ has a common label-free law, independent of Z, so the full $(Z, \widetilde{Z})$ Bayes error equals the Z Bayes error. Condition this prior on $\max_i |v_i| \le M$. Compactness of $\mathcal{H}$ implies that the discarded probability tends to zero as $M \to \infty$, and total variation contracts through the experiment. This gives a cube-supported Bayes lower bound $R_Z - \epsilon - o_M(1)$. Take $T \to \infty$, then $M \to \infty$, then $\epsilon \downarrow 0$.

The random-prefix OLS vector attains the upper bound: on $N_i(T) \ge 2$, (328) is a continuous function of the random discrete Gram matrix and $X_{i,T}$. Stabilization, uniform Gram convergence, and contiguity on every bounded local family give the limit Z_i in (359). The fallback event has probability tending to zero.

At $r = 1/2$, use the score pair

$$\left(T^{-1/2} \sum_{s \le N_i(T)} \varepsilon_{i,s}, (\log T)^{-1/2} \sum_{s \le N_i(T)} s^{-1/2} \varepsilon_{i,s} \right).$$

Its bracket tends to $\sigma^2 \operatorname{diag}(p_i, 1)$ because

$$\frac{a_{N_i(T)}}{\sqrt{T \log T}} = O_{\mathbb{P}}\Big((\log T)^{-1/2} \Big)$$

is the off-diagonal term. The maximal coefficients vanish. The slope score is asymptotically orthogonal and contains no first-order intercept signal. Compact-uniform contiguity transfers count stabilization over each fixed (h, v) cube, and Lemma 17 proves (356).

For $r > 1/2$, fixed compact amplitude shifts are uniformly contiguous because

$$\sup_{\beta \in \mathcal{B}^K} \sum_i \sum_{s \ge 1} (\beta_i - \beta_0)^2 s^{-2r} < \infty.$$

This transfers (348) uniformly over the class. For $n/T \in [\iota, 1]$, equations (328) and (329) give

$$\sqrt{T}(\widehat{\mu}_i(n) - \widetilde{\mu}_i(n)) = \frac{\sqrt{T} a_n}{b_n D_n} \left\{ \frac{a_n}{n} \sum_{s \le n} \varepsilon_{i,s} - \sum_{s \le n} s^{-r} \varepsilon_{i,s} \right\}. \tag{362}$$

Doob's inequality and the power-sum orders make the supremum of (362) $o_{\mathbb{P}}(1)$. The same calculation at $r = 1/2$ is $O_{\mathbb{P}}((\log T)^{-1/2})$, which also proves the boundary-case OLS attainment asserted above. The corrected-mean martingale CLT has variance σ^2/p_i. Hence OLS gives the Gaussian upper bound uniformly over $\mathcal{B}^K$, while fixing all amplitudes at β_0 gives the lower bound. This proves (357). □

An oracle allocation that sees the amplitudes needs a regularity condition. We call it locally p-regular if its allocation kernels ignore the revealed amplitude vector, or if its proportions converge to p uniformly over every fixed compact local amplitude set of the applicable regime; the terminal recommendation may always use the revealed values.

Corollary 6 (Oracle stable price)**.** *Every locally p-regular revealed-amplitude policy has local value*

$$\mathcal{V}_{\mathcal{H}}\left(\frac{c}{\sigma}; \left(\frac{1}{p_i} \right)_{i=1}^{K} \right) \tag{363}$$

in all three regimes. Define

$$\overline{\mathcal{V}}_{\mathcal{H}}(t) = \inf_{p_i > 0, \sum_i p_i = 1} \mathcal{V}_{\mathcal{H}}(t; (1/p_i)_i), \qquad q_{\mathcal{H}}(\alpha) = \inf\{t > 0 : \overline{\mathcal{V}}_{\mathcal{H}}(t) \le \alpha\}. \tag{364}$$

For $\alpha \in (0, 1 - 1/K)$, optimization over deterministic-proportion or locally p-regular oracle policies gives below the boundary

$$c_{\mathrm{un}} = \sigma \sqrt{\kappa(r)} q_{\mathcal{H}}(\alpha), \qquad c_{\mathrm{or}} = \sigma q_{\mathcal{H}}(\alpha), \qquad c_{\mathrm{un}}^2 / c_{\mathrm{or}}^2 = \kappa(r), \tag{365}$$

and ratio one at and above the boundary.

Proof. Corrected sample means obey the stable martingale CLT uniformly under the stated regularity; restriction to the common revealed vector gives the lower bound. Gaussian degradation makes $\overline{\mathcal{V}}_{\mathcal{H}}$ nonincreasing. At zero signal all labels have the same law and minimax error $1-1/K$, while balanced Gaussian ranking tends uniformly to zero as the signal diverges, so $q_{\mathcal{H}}(\alpha) \in (0,\infty)$. Theorem 15, deterministic proportion sampling, and inversion of the signal argument give (365). □

J.3 Sharp fixed-gap fixed-confidence identification

Classical fixed-confidence best-arm identification supplies the stationary benchmark. The bandit change-of-measure inequality of Kaufmann et al. (2016, Lemma 1) underlies general instance-dependent lower bounds, and Garivier & Kaufmann (2016, Theorem 1 and Theorem 14) characterize and attain the asymptotic fixed-confidence sample complexity. For two equal-variance Gaussian arms the corresponding constant is $8\sigma^2/\Delta^2$; the result below shows that it survives the transient-amplitude nuisance under a known amplitude-difference envelope. In contrast to the rest of the paper, this subsection fixes a nonzero gap and then sends $\delta \downarrow 0$. Let Θ^{env} be a two-arm parameter class with known amplitude-difference envelope

$$|\beta_2 - \beta_1| \le B_\beta < \infty. \tag{366}$$

Sampling the arms in pairs, at per-arm prefix n define

$$\widehat{\Delta}_n^{\mathrm{raw}} = \overline{Y}_{2,n} - \overline{Y}_{1,n}, \qquad \Delta_{12} = \mu_2 - \mu_1, \qquad \bar{\mathfrak{t}}_n = B_\beta a_n / n. \tag{367}$$

Then

$$\widehat{\Delta}_n^{\mathrm{raw}} = \Delta_{12} + \mathfrak{t}_n + Z_n, \qquad |\mathfrak{t}_n| \le \bar{\mathfrak{t}}_n, \qquad Z_n \sim N(0, 2\sigma^2/n). \tag{368}$$

For $\delta \in (0, e^{-1})$, put $L = \log(1/\delta)$, $\Upsilon = L^{-1/4}$, $n_1 = 2$, and

$$n_{k+1} = \lceil (1+\Upsilon) n_k \rceil, \qquad \delta_k = \frac{6\delta}{\pi^2 k^2}, \qquad w_{k,\delta} = 2\sigma \sqrt{\frac{\log(2/\delta_k)}{n_k}}. \tag{369}$$

Stop at the first epoch satisfying

$$|\widehat{\Delta}_{n_k}^{\mathrm{raw}}| > \bar{\mathfrak{t}}_{n_k} + w_{k,\delta}, \tag{370}$$

recommend according to its sign, and let the total sample size be $\tau_\delta = 2n_{k_\delta}$.

Let $\mathcal{A}_\delta^{\mathrm{un}}$ be the parameter-free procedures that terminate almost surely at every unique-best parameter and are uniformly δ-correct over Θ^{env}; define $\mathcal{A}_\delta^{\mathrm{or}}$ analogously when the amplitude vector is revealed. For $\bullet \in \{\mathrm{un}, \mathrm{or}\}$, set

$$\mathfrak{T}_\bullet(\Delta, \delta) = \inf_{\pi \in \mathcal{A}_\delta^\bullet} \sup_{\substack{\theta \in \Theta^{\mathrm{env}} \\ |\mu_2 - \mu_1| = \Delta}} \mathbb{E}_\theta \tau_\pi. \tag{371}$$

Theorem 16 (Fixed-confidence price one at every feasible fixed gap)**.** *The rule* (369)–(370) *is almost surely finite at every $\Delta_{12} \neq 0$ and uniformly δ-correct over the unique-best parameters in Θ^{env}. For every fixed $\Delta > 0$ with a nonempty gap slice,*

$$\limsup_{\delta \downarrow 0} \sup_{\substack{\theta \in \Theta^{\mathrm{env}} \\ |\mu_2 - \mu_1| = \Delta}} \frac{\mathbb{E}_\theta \tau_\delta}{\log(1/\delta)} \le \frac{8\sigma^2}{\Delta^2}. \tag{372}$$

For the matching lower bound, assume there are $\mu_\star$ and a common interior amplitude β_0 such that

$$\theta^\Delta = (\mu_\star - \Delta/2, \beta_0, \mu_\star + \Delta/2, \beta_0) \in \Theta^{\mathrm{env}}, \tag{373}$$

and, for every $\bar{\mu} \in [\mu_\star - \Delta/2, \mu_\star + \Delta/2]$, strict opposite-best alternatives $(\bar{\mu} + \epsilon, \beta_0, \bar{\mu} - \epsilon, \beta_0)$ belong to Θ^{env} for all sufficiently small $\epsilon > 0$. Then

$$\mathfrak{T}_{\mathrm{un}}(\Delta, \delta) \sim \mathfrak{T}_{\mathrm{or}}(\Delta, \delta) \sim \frac{8\sigma^2}{\Delta^2} \log(1/\delta), \qquad \delta \downarrow 0. \tag{374}$$

Thus the sharp fixed-gap fixed-confidence unknown/oracle ratio is one for every $r > 0$, under the envelope and feasibility conditions just stated.

Proof. For every epoch, the Gaussian tail bound gives

$$\mathbb{P}(|Z_{n_k}| > w_{k,\delta}) \le 2\exp\{-n_k w_{k,\delta}^2/(4\sigma^2)\} = \delta_k.$$

Because $\sum_k \delta_k = \delta$, with probability at least $1-\delta$ all epoch inequalities hold. On this event, a positive crossing implies

$$\Delta_{12} = \widehat{\Delta}_{n_k}^{\mathrm{raw}} - \mathfrak{t}_{n_k} - Z_{n_k} > \bar{\mathfrak{t}}_{n_k} + w_{k,\delta} - \bar{\mathfrak{t}}_{n_k} - w_{k,\delta} = 0,$$

and the negative case is symmetric. This proves uniform correctness. At fixed δ, $\bar{\mathfrak{t}}_{n_k}$ and $w_{k,\delta}$ tend to zero and $\widehat{\Delta}_{n_k}^{\mathrm{raw}} \to \Delta_{12}$ almost surely, proving finite stopping.

For the expected-time bound, fix $\epsilon > 0$, put

$$n_\delta = (1+\epsilon)\frac{4\sigma^2}{\Delta^2}L,$$

and let k_0 be the first epoch with $n_{k_0} \ge n_\delta$. Since $n_{k+1} \le (1+\Upsilon)n_k + 1$,

$$n_\delta \le n_{k_0} \le (1+\Upsilon)n_\delta + 1, \qquad n_{k_0}/n_\delta \to 1. \tag{375}$$

There are $O\{\Upsilon^{-1}(1+\log L)\}$ epochs before reaching order L, so $\log k_0 = o(L)$ and

$$w_{k_0,\delta} \le \frac{\Delta}{\sqrt{1+\epsilon}}\{1+o(1)\}. \tag{376}$$

The bias radius is nonincreasing. For all sufficiently small δ,

$$\frac{L + 2\log(k+1) + \log(\pi^2/3)}{L + 2\log k + \log(\pi^2/3)} \le 1 + \frac{2}{L} < 1 + \Upsilon \le \frac{n_{k+1}}{n_k},$$

so $w_{k,\delta}$ is also nonincreasing. Put

$$\Delta_\epsilon = \frac{\Delta}{2}\left(1 - \frac{1}{\sqrt{1+\epsilon}}\right) > 0.$$

For all small enough δ and every $k \ge k_0$,

$$\Delta - 2\bar{\mathfrak{t}}_{n_k} - w_{k,\delta} \ge \Delta_\epsilon. \tag{377}$$

If $\Delta_{12} = \Delta$ and the rule has not stopped at epoch k, then $|\widehat{\Delta}_{n_k}^{\mathrm{raw}}| \le \bar{\mathfrak{t}}_{n_k} + w_{k,\delta}$ and $\mathfrak{t}_{n_k} \ge -\bar{\mathfrak{t}}_{n_k}$, whence

$$Z_{n_k} \le -\{\Delta - 2\bar{\mathfrak{t}}_{n_k} - w_{k,\delta}\} \le -\Delta_\epsilon.$$

The negative gap is symmetric. Uniformly over the nuisance envelope,

$$\mathbb{P}_\theta(k_\delta > k) \le \exp\{-\Delta_\epsilon^2 n_k/(4\sigma^2)\} = e^{-c_\epsilon n_k}, \qquad k \ge k_0. \tag{378}$$

Tonelli's theorem gives

$$\begin{aligned}\mathbb{E}_\theta n_{k_\delta} &= n_1 + \sum_{k\ge 1}(n_{k+1} - n_k)\mathbb{P}_\theta(k_\delta > k) \\ &\le n_{k_0} + \sum_{k \ge k_0}(\Upsilon n_k + 1)e^{-c_\epsilon n_k} \le n_{k_0} + o(1).\end{aligned} \tag{379}$$

Multiply by two and then send $\epsilon \downarrow 0$ to prove (372).

For the lower bound, work at θ^Δ in (373). For an opposite-best alternative θ' with the same common amplitude, the stopped Gaussian likelihood chain rule gives, whenever $\mathbb{E}_\theta \tau < \infty$,

$$\mathrm{KL}(\mathbb{P}_\theta^{\mathcal{F}_\tau}, \mathbb{P}_{\theta'}^{\mathcal{F}_\tau}) = \sum_{i=1}^{2} \mathbb{E}_\theta N_i(\tau)\frac{(\mu_i - \mu_i')^2}{2\sigma^2}. \tag{380}$$

Apply the chain rule at $\tau \wedge m$ and let $m \to \infty$; finite expected stopping makes the Gaussian log-likelihood increments integrable. If the expectation is infinite, the desired lower bound is automatic. Uniform correctness and binary data processing yield

$$\sum_i \mathbb{E}_\theta N_i(\tau)\frac{(\mu_i - \mu_i')^2}{2\sigma^2} \ge \mathrm{kl}(1-\delta, \delta). \tag{381}$$

Write $\bar{N}_i = \mathbb{E}_\theta N_i(\tau)$ and $\bar{N}_\tau = \bar{N}_1 + \bar{N}_2$. Strict feasible alternatives may approach the weighted tie $\mu_1' = \mu_2' = \bar{\mu}$, and

$$\inf_{\bar{\mu}} \sum_i \bar{N}_i(\mu_i - \bar{\mu})^2 = \frac{\bar{N}_1\bar{N}_2}{\bar{N}_\tau}\Delta^2.$$

Consequently

$$\mathrm{kl}(1-\delta,\delta) \le \frac{\bar{N}_1\bar{N}_2}{\bar{N}_\tau}\frac{\Delta^2}{2\sigma^2} \le \bar{N}_\tau\frac{\Delta^2}{8\sigma^2}.$$

Since $\mathrm{kl}(1-\delta,\delta)/\log(1/\delta) \to 1$, this matches the unknown-amplitude upper bound. In the revealed experiment, corrected means give the same upper constant; the lower alternatives have identical revealed amplitude vectors and hence zero side-information KL. This proves (374). □

For a finite-sample-safe version, construct the same dense epoch grid using the branch confidence level $\delta/2$, that is, $L_{\mathrm{hyb}} = \log(2/\delta)$ and $\Upsilon_{\mathrm{hyb}} = L_{\mathrm{hyb}}^{-1/4}$. At paired prefix $n \ge 2$, let

$$\widehat{\Delta}_n^{\mathrm{OLS}} = \widehat{\mu}_2(n) - \widehat{\mu}_1(n), \qquad \widehat{\Delta}_n^{\mathrm{OLS}} \sim N\left(\Delta_{12}, \frac{2\sigma^2}{D_n}\right). \tag{382}$$

Split the error budget as

$$\delta_k^{\mathrm{raw}} = \delta_k^{\mathrm{OLS}} = \frac{3\delta}{\pi^2 k^2} \tag{383}$$

and use widths

$$w_k^{\mathrm{raw}} = 2\sigma\sqrt{\frac{\log(2/\delta_k^{\mathrm{raw}})}{n_k}}, \qquad w_k^{\mathrm{OLS}} = 2\sigma\sqrt{\frac{\log(2/\delta_k^{\mathrm{OLS}})}{D_{n_k}}}. \tag{384}$$

Stop when either

$$|\widehat{\Delta}_{n_k}^{\mathrm{raw}}| > \bar{\mathfrak{t}}_{n_k} + w_k^{\mathrm{raw}} \quad \text{or} \quad |\widehat{\Delta}_{n_k}^{\mathrm{OLS}}| > w_k^{\mathrm{OLS}} \tag{385}$$

holds, and use a crossing sign. Break a simultaneous opposite-sign crossing by any fixed rule.

Corollary 7 (Hybrid correctness and sharp constant)**.** *The hybrid rule is almost surely finite and uniformly δ-correct over Θ^{env}. Pathwise,*

$$\tau_{\mathrm{hybrid}}(\delta) \le \tau_{\mathrm{raw}}(\delta/2), \tag{386}$$

where the raw rule uses the same grid and the raw branch confidence allocation above. Hence the hybrid has the sharp upper constant $8\sigma^2/\Delta^2$. Without a known envelope (366)*, the OLS-only branch remains uniformly valid, but below $r = 1/2$ its leading total-sample constant is $8\sigma^2\kappa(r)/\Delta^2$.*

Proof. The two-sided Gaussian tail bounds for the raw and OLS contrasts are at most δ_k^{raw} and δ_k^{OLS}. Each sequence sums to $\delta/2$. On their simultaneous good event, every crossing sign equals the sign of Δ_{12}, so a union bound gives correctness without an independence assumption. The raw branch gives almost-sure finiteness and (386). Since $\log(2/\delta)/\log(1/\delta) \to 1$, Theorem 16 gives the same leading constant. Finally, $D_n/n \to 1/\kappa(r)$ below the boundary, giving the stated OLS-only constant. □

J.4 Fixed-gap exponent

The fixed-budget analogue of the previous result is that, at a fixed gap, the unknown amplitude does not change the large-deviation exponent of the misidentification probability either.

Proposition 6 (No fixed-gap exponent price)**.** *Fix $\Delta > 0$ and $0 \le \beta_i \le \bar{\beta} < \infty$. With n observations per arm, let $\mathfrak{R}_n^{\mathrm{un}}(\Delta)$ be the minimax wrong-arm probability over both gap signs on a bounded intercept interval containing the symmetric gap pair. Then*

$$\lim_{n\to\infty} -\frac{1}{n}\log \mathfrak{R}_n^{\mathrm{un}}(\Delta) = \frac{\Delta^2}{4\sigma^2}, \tag{387}$$

the same exponent as with revealed amplitudes.

Proof. Let $\mathcal{I} = [\underline{\mu}, \overline{\mu}]$ be the bounded intercept interval of Proposition 6, and assume that for some $\mu_\star \in \mathcal{I}$, both $\mu_\star - \Delta/2$ and $\mu_\star + \Delta/2$ belong to $\mathcal{I}$. Define

$$\Theta_+(\Delta) := \{(\mu_1, \beta_1, \mu_2, \beta_2) : \mu_i \in \mathcal{I},\ 0 \le \beta_i \le \bar{\beta},\ \mu_2 - \mu_1 \ge \Delta\},$$
$$\Theta_-(\Delta) := \{(\mu_1, \beta_1, \mu_2, \beta_2) : \mu_i \in \mathcal{I},\ 0 \le \beta_i \le \bar{\beta},\ \mu_1 - \mu_2 \ge \Delta\}.$$

The minimax risk $\mathfrak{R}_n^{\mathrm{un}}(\Delta)$ is the infimum over recommendations of the larger worst error over these two classes.

Choose integers $k_n \to \infty$ with $k_n/n \to 0$, and define the tail averages

$$\overline{Y}_{i,n}^{\mathrm{tail}} := \frac{1}{n-k_n} \sum_{s=k_n+1}^{n} Y_{i,s}, \qquad \overline{x}_n := \frac{1}{n-k_n} \sum_{s=k_n+1}^{n} s^{-r}.$$

Then $0 \le \overline{x}_n \le (k_n+1)^{-r} \to 0$. Under $\Theta_+(\Delta)$,

$$\mathbb{E}_\theta[\overline{Y}_{2,n}^{\mathrm{tail}} - \overline{Y}_{1,n}^{\mathrm{tail}}] = \mu_2 - \mu_1 + (\beta_2 - \beta_1)\overline{x}_n \ge \Delta - \bar{\beta}\overline{x}_n.$$

The reverse-gap case is symmetric, and the difference of the two tail averages has variance $2\sigma^2/(n - k_n)$. Its sign therefore gives

$$\mathfrak{R}_n^{\mathrm{un}}(\Delta) \le \Phi\left(-\frac{(\Delta - \bar{\beta}\overline{x}_n)\sqrt{n-k_n}}{\sqrt{2}\,\sigma}\right).$$

Using $-\log\Phi(-u) \sim u^2/2$ as $u \to \infty$ yields

$$\liminf_{n\to\infty} -\frac{1}{n}\log \mathfrak{R}_n^{\mathrm{un}}(\Delta) \ge \frac{\Delta^2}{4\sigma^2}. \tag{388}$$

For the reverse inequality, restrict to $\beta_1 = \beta_2 = 0$ and to the two parameter points

$$(\mu_1, \mu_2) = (\mu_\star - \Delta/2, \mu_\star + \Delta/2), \qquad (\mu_1, \mu_2) = (\mu_\star + \Delta/2, \mu_\star - \Delta/2).$$

They belong to $\Theta_+(\Delta)$ and $\Theta_-(\Delta)$, respectively. Their joint Gaussian laws have squared Mahalanobis distance $2n\Delta^2/\sigma^2$. Hence every recommendation has maximum error at least their equal-prior Bayes error,

$$\Phi\left(-\frac{\Delta\sqrt{n}}{\sqrt{2}\,\sigma}\right).$$

The same Gaussian-tail asymptotic gives

$$\limsup_{n\to\infty} -\frac{1}{n}\log \mathfrak{R}_n^{\mathrm{un}}(\Delta) \le \frac{\Delta^2}{4\sigma^2}.$$

Combining this with (388) proves (387). The revealed-amplitude paired problem contains the same ordinary Gaussian two-point subproblem, and corrected sample means attain the same exponent. □

K Experimental Checks

All experiments use the Gaussian model analyzed in the paper. Their purpose is to check declared finite-horizon policies and exact Gaussian calculations, not to establish finite-sample minimaxity or practical superiority for hyperparameter optimization. Complete frozen grids, seeds, batch summaries, exact-integration code, hashes, and independent verification reports are included in the supplementary artifact.

K.1 Screen–then–compare

We test the constructive policy from Theorem 3 on

$$\mu = (0, \Delta, \Delta + \eta, \ldots, \Delta + \eta), \qquad \eta = 1.5, \qquad \Delta = c/\sqrt{B},$$

with unit Gaussian noise and, for each K, a fixed amplitude profile equally spaced from .5 to 1.5. The adaptive policy uses $m_B = \lceil B^{\cdot 4} \rceil$ pilot pulls per arm (pilot exponent $\omega = .4$), selects the two smallest

pilot intercept estimates, and ranks those two arms using independent suffix data. The comparator allocates B/K full-prefix observations to every arm. Exact Gaussian integration determines the finite-budget risks of both policies, and Monte Carlo checks those calculations independently.

Across $K \in \{2, 4, 8\}$ and exponents on both sides of $r = 1/2$, the screening failure probability decreases with the budget and the squared critical-signal gain moves toward $K/2$. Convergence is slowest near $r = 1/2$, where suffix OLS has not yet reached its pointwise asymptotic efficiency. Figure 2 shows the representative mechanism; the supplementary artifact contains the full grid. These are comparisons of two declared policies on a separated-pair family, not finite-sample minimax ratios or a supremum over the nuisance.

K.2 Finite-prefix and predictable-design checks

Finite-prefix law. We evaluated the exact normalized ridge-intercept risk of equation (48) over exponents dense around $r = 1/2$ and prefixes from 64 to 1,048,576. Simulations based on independent Gaussian sufficient statistics tracked the exact calculation; for the unknown-amplitude estimator, the largest absolute standardized deviation over the 128 cells was 3.08. Figure 1 gives the direct exponent-wise comparison, while Figure 3 gives the complementary horizon-wise view: convergence to $\kappa(r)$ is highly nonuniform near the boundary. Because the finite-prefix law is available in closed form, this experiment is an implementation check rather than independent evidence for the minimax claim of Theorem 1.

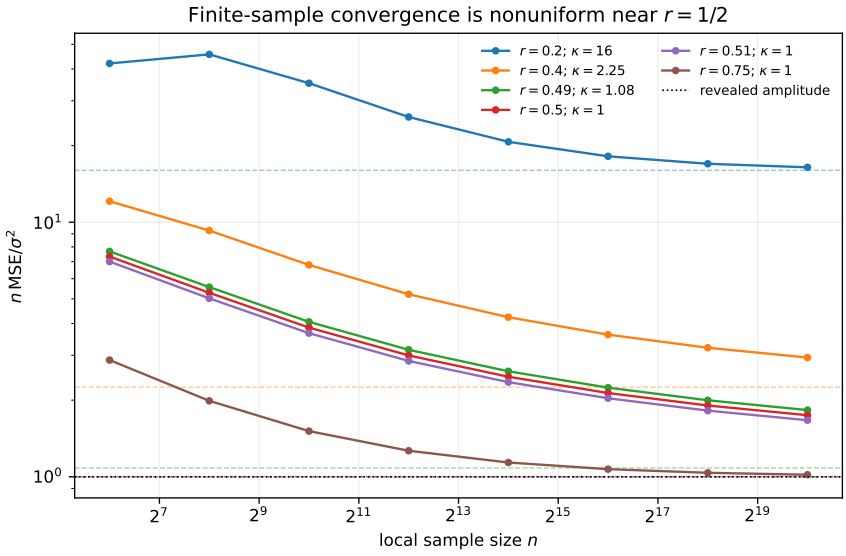

Figure 3: Exact finite-prefix normalized intercept risk. Dashed lines give the pointwise unknown-amplitude limits, and the revealed-amplitude reference is one. Convergence becomes slow near $r = 1/2$.

Predictable allocation. We also checked two data-dependent designs whose terminal proportions converge to deterministic targets: a two-arm pilot rule that perturbs the prescribed quotas by $O(\sqrt{T})$, as in Proposition 4, and a $K = 4$ pilot-ranked rule with an $O(T^{\cdot 7})$ tilt. Full-prefix OLS risks approached their deterministic Gaussian references. At the largest $K = 4$ horizon, the adaptive-to-deterministic squared critical-signal ratios ranged from .996 to 1.014 over the tested exponents. These experiments validate the chosen predictable rules only; they do not test the adaptive lower bounds or a finite-sample nuisance supremum.

K.3 Random terminal information

To illustrate why a marginal allocation law does not determine the local experiment (Proposition 5), we compare three two-arm policies with terminal arm-1 proportions $1/4$ and $3/4$ and the same central

low-quota probability $2\Phi(-1)$. The quota is selected by an independent coin, by an ancillary pilot, or by an endogenous partial-score stopping rule. At $T = 65{,}536$ and local signal $c = 1$, the first two policies have the same local risk, whereas the endogenous rule changes the joint score–information law and hence the risk: its analytic error is $.2070$, compared with $.2298$ for the coin and ancillary designs. This is a policy-specific illustration, not a global adaptive minimax comparison.

K.4 Fixed-confidence stress check

Finally, we test the fixed-gap rule of Theorem 16 under the known envelope $|\beta_2 - \beta_1| \le 2$, that is, $B_\beta = 2$. The unknown-amplitude procedure combines an OLS-intercept boundary with a raw-contrast bias-envelope boundary; the oracle subtracts the transient exactly. Across the frozen grid of confidence levels, gap signs, nuisance contrasts, and $r \in \{.25, .49, .75\}$, observed errors were compatible with the nominal guarantee. The method-to-oracle expected-sample ratio approached one near and above $r = 1/2$, while convergence remained slow at $r = .25$.

These simulations are stress checks, not proofs of uniform correctness or finite-confidence minimaxity. They assume two arms, known r and innovation variance, a fixed nonzero gap, and the stated amplitude-difference envelope.