INFORMATION-TEMPERED PRIORS FOR GAMMA THOMPSON SAMPLING

Anonymous authors
Paper under double-blind review

ABSTRACT

Nearly equal pessimistic observations can make the optimal-arm posterior falsely precise in Gamma Thompson sampling (TS) with unknown means and shapes. Skipping the arm then freezes its posterior. Information-tempered matching (ITM) priors with fixed $\eta > 0$ retain full support and first-order probability matching with tail $k^{-1-\eta}$. Global posterior-integral comparisons, high-shape saddlepoint control, and a small-shape carrier bound establish summable pessimism odds for these priors. For every fixed finite-arm Gamma instance with a unique optimal mean, uncapped ITM-TS is uniformly efficient and attains the Burnetas–Katehakis regret constant. Regular priors globally comparable to ITM share this guarantee. For weights $w(k) \sim Ck^{-\tau}$, $C > 0$, as $k \to \infty$, reciprocal optimism near these histories is locally integrable exactly when $\tau > 1$ for every fixed sample size $n \geq 2$. With two initial pulls per arm, ordinary TS incurs $\Omega(\sqrt{T})$ lockout regret under Jeffreys' prior and a distinct $\Omega(\log T)$ coefficient constraint under the matching prior. Full-bandit experiments with 1,024 trials per instance support lower mean regret and CVaR for ITM ($\eta = .5$) than matching TS on the five-arm instance; both two-arm instances have paired intervals containing zero.

1 INTRODUCTION

Thompson sampling (TS) selects the arm with the largest sampled mean (Thompson, 1933) and can attain the information limit in one-parameter exponential families (Lai & Robbins, 1985; Korda et al., 2013). For Gamma rewards with unknown mean μ and shape k, nearly equal pessimistic observations can concentrate the optimal-arm posterior below a competitor as $k \to \infty$. Skipping the arm freezes this falsely precise posterior and can cause horizon-scale lockout. We study which prior tails permit recovery.

For unknown-variance Gaussian rewards, Honda & Takemura (2014) prove optimality under $\pi(\mu, \sigma^2) \propto (\sigma^2)^{-1-\alpha}$ for $\alpha < 0$ and large-regret lower bounds for $\alpha \geq 0$. Precision $\lambda = 1/\sigma^2$ has prior tail $\lambda^{-\tau}$, $\tau = 1 - \alpha$, so the optimal range is $\tau > 1$. Lee et al. (2023) prove uncapped Pareto TS optimal for selected matching priors. Both lower-bound constructions use a known competitor. Truncation yields guarantees for selected priors in Pareto, Gaussian, and uniform models (Lee et al., 2023; 2024); preprints study profiling, bootstrap, and Gaussian VarTS (Wang, 2026; Eckles & Kaptein, 2014; Saha & Kveton, 2023).

For fixed-shape Gamma rewards, ExpTS modifies the sampling distribution (Jin et al., 2022), and ϵ-TS combines posterior draws with empirical means (Jin et al., 2023). A workshop study analyzes UCB with a known shape lower bound (Geist, 2015). With unknown shape, fixed-sample asymptotics of the nonconjugate shape posterior describe local recovery, while the sharp-regret proof must sum pessimism odds over all sample sizes. Posterior-integral and sampling-density estimates with sample-size-uniform constants control that sum on both sides of the empirical-shape boundary.

The relevant shape coordinate is

$$A(k) = k\psi(k) - k - \log\Gamma(k), \qquad A'(k) = k\psi_1(k) - 1 > 0. \tag{1}$$

Here Γ is the Gamma function, $\psi = (\log\Gamma)'$ the digamma function, and $\psi_1 = \psi'$ the trigamma function. The second-order unconditional matching prior for μ has shape weight $A'(k)$ (DiCiccio & Young, 2010), with tail k^{-1} at the transition below. Section 2 gives its parameter transformation and matching calculation.

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.

We introduce the information-tempered matching (ITM) family

$$\pi_\eta(\mu, k) \;\propto\; \frac{A'(k)}{\mu}\big[1 + \exp\{2A(k)\}\big]^{-\eta}, \qquad \eta \geq 0. \tag{2}$$

In the coordinate $z = A(k)$, the matching shape measure is $A'(k)\,dk = dz$; ITM tempers it smoothly by $[1 + e^{2z}]^{-\eta}$. This changes the tail from k^{-1} to $k^{-1-\eta}$, retaining full support and the small-shape asymptotic weight. First-order matching holds throughout the family; $\eta = 0$ also has the cited second-order property. Matching calibrates inference for μ; the tail change permits sequential recovery.

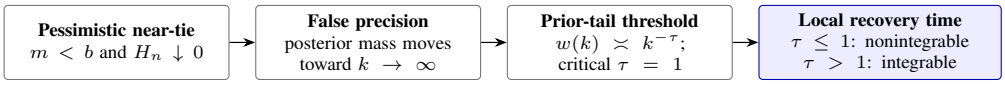

Figure 1: The false-precision mechanism and its prior-tail transition. ITM sets $\tau = 1 + \eta > 1$, making the optimal-arm pessimism odds summable while retaining the standard TS decision rule.

Contributions.

1. *Uniform sequential recovery.* Global posterior-integral comparisons, high-shape log-Beta density bounds, and a small-shape carrier bound establish summable pessimism odds across sample sizes and both shape boundaries.
2. *A class of sharp-regret priors.* For every fixed $\eta > 0$, uncapped ITM-TS is uniformly efficient and attains the instancewise information constant. Regular weights globally comparable to ITM share this guarantee. Section G.6 gives sufficient conditions at both shape boundaries under Assumption 1, including the simpler tempering $A'(k)(1 + k)^{-\eta}$.
3. *Boundary mechanism and lockout.* For each fixed $n \geq 2$, reciprocal optimism is locally integrable exactly when $\tau > 1$. With two initial pulls per arm, coupling to an unknown competitor yields $\Omega(\sqrt{T})$ and $\Omega(\log T)$ ordinary-TS lockout contributions for Jeffreys and matching priors, respectively.

The exact posterior formula determines recovery probabilities, and the scalar information projection gives the regret coefficient. Experiments measure lockout and sensitivity to η.

2 Model and information-tempered prior

For $\mu, k > 0$, let $P_{\mu,k}$ have density

$$f(x \mid \mu, k) = \frac{(k/\mu)^k}{\Gamma(k)} x^{k-1} \exp\left(-\frac{kx}{\mu}\right), \qquad x > 0. \tag{3}$$

Thus μ is the mean and k the shape, with $\mathbb{E}X = \mu$ and $\mathrm{Var}(X) = \mu^2/k$. We use $\mathrm{Gamma}(\alpha, \lambda)$ in shape–rate form and $\mathrm{InvGamma}(\alpha, \beta)$ in shape–scale form. Fix $2 \leq K < \infty$ and $[K] = \{1, \ldots, K\}$. Arm $a \in [K]$ has $\theta_a = (\mu_a, k_a)$; the unique optimum is $\theta^\star = (\mu^\star, k^\star) := \theta_{a^\star}$. With $\Delta_a = \mu^\star - \mu_a$, pseudo-regret is

$$R_T = \sum_{a \neq a^\star} \Delta_a \mathbb{E}N_a(T), \qquad N_a(T) = \sum_{t=1}^{T} \mathbf{1}\{A_t = a\}. \tag{4}$$

Every arm is pulled twice, after which TS independently samples $(\widetilde{\mu}_a, \widetilde{k}_a)$ from each posterior and chooses an arm maximizing $\widetilde{\mu}_a$.

For $P \ll Q$, write $\mathrm{KL}(P, Q) = \mathbb{E}_P[\log(dP/dQ)]$ and $D((\mu, k)\|(m, q)) := \mathrm{KL}(P_{\mu,k}, P_{m,q})$. For a suboptimal parameter $\theta = (\mu, k)$ and $b > \mu$, define

$$K_{\inf}(\theta, b) = \inf_{m \geq b,\, q > 0} \mathrm{KL}(P_{\mu,k}, P_{m,q}). \tag{5}$$

Here a policy is *uniformly efficient* if, on every fixed instance in this model class, $R_T = o(T^\alpha)$ for every $\alpha > 0$. Under this premise, the multiparameter information lower bound implies, for every suboptimal arm,

$$\liminf_{T \to \infty} \frac{\mathbb{E}N_a(T)}{\log T} \geq \frac{1}{K_{\inf}((\mu_a, k_a), \mu^\star)}, \tag{6}$$

and consequently

$$\liminf_{T\to\infty} \frac{R_T}{\log T} \geq \sum_{a\neq a^\star} \frac{\Delta_a}{K_{\inf}((\mu_a, k_a), \mu^\star)} \tag{7}$$

(Burnetas & Katehakis, 1996; Lattimore & Szepesvári, 2020). The right-hand side is the instance-level information lower bound for uniformly efficient policies.

2.1 Fisher geometry and prior tails

Taking the negative expected Hessian of the log density gives

$$I(\mu, k) = \begin{pmatrix} k/\mu^2 & 0 \\ 0 & \psi_1(k) - 1/k \end{pmatrix}, \qquad \det I(\mu, k) = \frac{A'(k)}{\mu^2}. \tag{8}$$

The cross term vanishes, and $A'(k) > 0$ because $\psi_1(k) = \sum_{j\geq 0}(k+j)^{-2} > \int_0^\infty (k+x)^{-2}dx = 1/k$. For interest parameter μ, orthogonality reduces the first-order matching equation to $\partial_\mu\{\pi(\mu, k)\mu/\sqrt{k}\} = 0$. Hence every prior $\pi(\mu, k) \propto w(k)/\mu$, including every ITM prior, is first-order matching. Example 4 of DiCiccio & Young (2010) uses ψ for the Gamma shape and λ for the rate. Under $k = \psi$ and $\mu = \psi/\lambda$, its unique second-order unconditional matching prior has shape factor $k\psi_1(k) - 1 = A'(k)$; this second-order property is specific to $w = A'$. Sequential recovery depends on the integrability of posterior pessimism odds, established for ITM with $\eta > 0$ in theorem 10. Standard special-function expansions, derived in section A, yield

$$\begin{aligned} A(k) &= \log k - 1 - k + O(k^2), & A'(k) &= k^{-1} - 1 + O(k), & k &\downarrow 0, \\ A(k) &= \tfrac{1}{2}\log\frac{k}{2\pi e} - \frac{1}{6k} + O(k^{-3}), & A'(k) &= \frac{1}{2k} + \frac{1}{6k^2} + O(k^{-4}), & k &\to \infty. \end{aligned} \tag{9}$$

Consequently A maps $(0, \infty)$ monotonically onto $\mathbb{R}$ and

$$w_\eta(k) := A'(k)[1 + e^{2A(k)}]^{-\eta} \sim \begin{cases} k^{-1}, & k \downarrow 0, \\ \frac{1}{2}(2\pi e)^\eta k^{-1-\eta}, & k \to \infty. \end{cases} \tag{10}$$

Jeffreys' rule has $w_J(k) = \sqrt{A'(k)} \asymp k^{-1/2}$ at infinity; matching has $w_0(k) = A'(k) \asymp k^{-1}$; ITM has exponent $1 + \eta$. The factor in (2) smoothly downweights the high-shape boundary while retaining full shape support.

2.2 Prior regularity

Assumption 1 (Prior regularity)**.** *w is positive and continuous on $(0, \infty)$, is bounded above and below on every compact subinterval, satisfies $w(k) = C_\infty k^{-\tau}\{1 + o(1)\}$ at infinity, and obeys $w(k) \leq C_0 k^{-\rho}$ near zero for some $\rho < 2$. The posterior after two observations is proper.*

Jeffreys, matching, and ITM satisfy Assumption 1, which supplies the posterior concentration used in the coupling and overshoot bounds.

3 Exact posterior and information projection

Let $x_{1:n} > 0$, $S = \sum_i x_i$, $L = \sum_i \log x_i$, $\bar{x} = S/n$, and

$$H_n = \log \bar{x} - \frac{L}{n} \geq 0. \tag{11}$$

Equality holds only when the observations coincide.

Theorem 2 (Posterior reduction and propriety)**.** *Under $\pi(\mu, k) \propto w(k)/\mu$,*

$$\mu \mid k, x_{1:n} \sim \mathrm{InvGamma}(nk, kS), \tag{12}$$

$$p(k \mid x_{1:n}) \propto w(k)\frac{\Gamma(nk)}{\Gamma(k)^n} n^{-nk} e^{-nkH_n}. \tag{13}$$

For Jeffreys, matching, and every ITM prior, this posterior is proper almost surely for $n \geq 2$.

Proof. The joint kernel in (μ, k) is

$$w(k)\frac{k^{nk}}{\Gamma(k)^n}\Big(\prod_i x_i^{k-1}\Big)\mu^{-nk-1}e^{-kS/\mu}.$$

Substituting $v = kS/\mu$ gives the μ integral $(kS)^{-nk}\Gamma(nk)$; together with $\prod_i x_i = S^n n^{-n} e^{-nH_n}$, this yields both displays. At zero, (36) gives, for this fixed n, $\Gamma(nk)\{\Gamma(k)^n n^{nk}\}^{-1} = k^{n-1}/n\{1 + O_n(k)\}$; multiplying by $w_\eta(k) \sim k^{-1}$ gives k^{n-2}, and Jeffreys gives $k^{n-3/2}$. Both are integrable for $n \geq 2$. At infinity,

$$\frac{\Gamma(nk)}{\Gamma(k)^n}n^{-nk} = (2\pi)^{(1-n)/2}n^{-1/2}k^{(n-1)/2}\{1 + O(k^{-1})\}.$$

Since a continuous sample has $H_n > 0$ almost surely, e^{-nkH_n} dominates each polynomial tail. The bounds at zero and infinity establish propriety. □

Posterior computation is therefore one-dimensional: sample k from (13), then draw μ exactly from (12).

3.1 A SCALAR INFORMATION PROJECTION

Set

$$h(k) = \log k - \psi(k), \qquad \mathcal{I}(r) = r - 1 - \log r. \tag{14}$$

Because $h'(k) = -A'(k)/k < 0$, h decreases from $+\infty$ to 0.

Theorem 3 (Exact information projection)**.** *For $0 < \mu < b$ and $k > 0$, there is a unique $q_b \in (0, k)$ satisfying*

$$h(q_b) = h(k) + \mathcal{I}(\mu/b). \tag{15}$$

The optimizer in (5) *is (b, q_b) and*

$$K_{\inf}((\mu, k), b) = A(k) - A(q_b). \tag{16}$$

Proof. Direct expectation of the log-density ratio gives

$$\begin{aligned} D((\mu, k)\|(m, q)) = {} & \log\Gamma(q) - \log\Gamma(k) + (k - q)\psi(k) \\ & + q\log\frac{km}{q\mu} - k + \frac{q\mu}{m}. \end{aligned} \tag{17}$$

For fixed q, $\partial_m D = q(m - \mu)/m^2 > 0$ on $m \geq b$, so $m = b$. At $m = b$,

$$\partial_q D = \psi(q) - \psi(k) + \log\frac{kb}{q\mu} - 1 + \frac{\mu}{b}, \quad \partial_{qq} D = \psi_1(q) - q^{-1} = A'(q)/q > 0.$$

The first derivative vanishes exactly at (15); strict convexity makes the root the unique minimizer. Since $\mathcal{I}(\mu/b) > 0$ and h is decreasing, $q_b < k$. Finally the root equation gives $\log(kb/(q_b\mu)) = \psi(k) - \psi(q_b) + 1 - \mu/b$. The termwise substitution displayed in section B leaves $k\psi(k) - k - \log\Gamma(k) - [q_b\psi(q_b) - q_b - \log\Gamma(q_b)]$, which is (16). □

The least-informative alternative always moves the nuisance shape downward. Equation (15) replaces constrained two-dimensional optimization with a bracketed monotone root search.

4 THE PRIOR-TAIL LOCKOUT TRANSITION

Fix a frozen posterior after $n \geq 2$ observations with empirical mean $m = \bar{x} < b$. Write

$$p_b(x_{1:n}) = \mathbb{P}(\widetilde{\mu} \geq b \mid x_{1:n}). \tag{18}$$

Independent draws recover above b after a geometric number of attempts with success probability p_b. Integrability of p_b^{-1} near rare histories determines whether the mean recovery time is finite.

4.1 SAMPLING GEOMETRY NEAR EQUAL OBSERVATIONS

For $X_i \overset{\text{iid}}{\sim} P_{\mu_0,k_0}$, the Gamma sum–proportion factorization gives

$$S = \sum_i X_i \sim \text{Gamma}(nk_0, k_0/\mu_0), \quad Z = (X_1/S, \ldots, X_n/S) \sim \text{Dir}(k_0, \ldots, k_0), \tag{19}$$

with $S \perp Z$, where Dir denotes the Dirichlet distribution and $\perp$ denotes independence. Here $m = S/n$, while

$$H_n = -\log\left[n\left(\prod_{i=1}^n Z_i\right)^{1/n}\right] \tag{20}$$

depends only on Z.

Lemma 4 (Dirichlet small ball). *Let $d = (n-1)/2$. As $u \downarrow 0$,*

$$\mathbb{P}(H_n \le u) = C_{n,k_0} u^d\{1 + O(u^{1/2})\}, \qquad C_{n,k_0} \in (0, \infty). \tag{21}$$

Proof. Write $z_i = n^{-1} + v_i$ with $\sum_i v_i = 0$. Taylor's formula, uniformly on a small tangent-space ball, gives $H_n(z) = \frac{n}{2}\sum_i v_i^2 + O(\|v\|^3)$. The Dirichlet density is positive and continuously differentiable at the uniform point. The sublevel set is therefore a $(n-1)$-dimensional ellipsoid of radius $u^{1/2}$, with a relative radial error $O(u^{1/2})$. Its volume is $Vu^{(n-1)/2}\{1+O(u^{1/2})\}$; multiplying by the density at the center proves the claim. Section C supplies the two-sided coordinate bounds underlying this expansion. □

4.2 POSTERIOR OPTIMISM

Suppose

$$w(k) = C_\infty k^{-\tau}\{1 + o(1)\}, \qquad k \to \infty, \tag{22}$$

and the posterior is proper. Put $\beta = d + 1 - \tau$.

Theorem 5 (Frozen-posterior phase transition). *Fix $0 < r_- < r_+ < 1$ and restrict to $m/b \in [r_-, r_+]$. Uniformly over this interval, as $H_n \downarrow 0$,*

$$p_b(x_{1:n}) \asymp \begin{cases} H_n^\beta, & \beta > 0, \\ 1/\log(1/H_n), & \beta = 0, \\ 1, & \beta < 0. \end{cases} \tag{23}$$

Moreover,

$$\mathbb{E}\big[p_b(X_{1:n})^{-1} \mid m/b \in [r_-, r_+]\big] < \infty \quad \Longleftrightarrow \quad \tau > 1. \tag{24}$$

Proof. Write $r = m/b$. From (13), the posterior normalizer and optimism numerator are

$$Z(H) = \int_0^\infty w(k) R_n(k) e^{-nkH}\, dk,$$
$$N(H, r) = \int_0^\infty w(k) R_n(k) e^{-nkH} P(nk, nkr)\, dk,$$

where $R_n(k) = \Gamma(nk)\{\Gamma(k)^n n^{nk}\}^{-1}$ and $P(a, x) := \Gamma(a)^{-1}\int_0^x t^{a-1}e^{-t}dt$ is the regularized lower incomplete Gamma function; $\mathbb{P}$ denotes probability. The conditional identity follows from (12): if $Y \sim \text{Gamma}(nk, 1)$, then $\mu = kS/Y$, so $\mathbb{P}(\mu \ge b \mid k, x) = \mathbb{P}(Y \le nkr) = P(nk, nkr)$.

For fixed n, Stirling's two-sided bound gives $R_n(k) \asymp k^d$ for all sufficiently large k. Hence the tail of Z is comparable to $\int_K^\infty k^{d-\tau}e^{-nkH}dk$. The explicit substitution in section C.2 gives order $H^{-\beta}$ when $\beta > 0$, $\log(1/H)$ when $\beta = 0$, and one when $\beta < 0$; the integral on $(0, K]$ converges to a positive constant.

For $r \le r_+ < 1$, the Gamma Chernoff inequality gives $P(nk, nkr) \le \exp[-nk\mathcal{I}(r)] \le \exp[-nk\mathcal{I}(r_+)]$. This dominates the numerator tail even at $H = 0$. On a fixed interval $k \in [k_1, k_2] \subset (0, \infty)$, continuity and $r \in [r_-, r_+]$ give a common positive lower bound. Dominated

convergence therefore yields $0 < c \le N(H, r) \le C < \infty$ uniformly. Dividing by $Z(H)$ proves (23).

By (19), conditioning on $m/b \in [r_-, r_+]$ does not change the law of H_n. When $\beta > 0$, lemma 4 reduces local integrability to $\int_0^\varepsilon u^{d-1-\beta} du = \int_0^\varepsilon u^{\tau-2} du$, finite exactly for $\tau > 1$. If $\beta \le 0$, (23) makes the reciprocal bounded (or logarithmic), and $\beta \le 0$ already implies $\tau > 1$. The bounds in section C.2 also show that p_b^{-1} is uniformly bounded on $H_n \ge \varepsilon$. This proves both directions. □

The cancellation $H_n^{d-1} H_n^{-(d+1-\tau)} = H_n^{\tau-2}$ explains why the threshold $\tau = 1$ is unchanged by any fixed initialization size $n \ge 2$.

If G_p is geometric on $\{1, 2, \ldots\}$ and $W_T = \min(G_p, T)$, then $\mathbb{E}(W_T \mid p) = [1 - (1-p)^T]/p \asymp \min(T, p^{-1})$. Integrating the two-sided bounds above gives

$$\mathbb{E} W_T \asymp \begin{cases} T^{(1-\tau)/(d+1-\tau)}, & \tau < 1, \\ \log T, & \tau = 1, \\ 1, & \tau > 1. \end{cases} \tag{25}$$

For $n = 2$, Jeffreys gives $\sqrt{T}$, matching gives $\log T$, and every fixed $\eta > 0$ gives a bounded local mean. The effect appears in expectations and upper tails, but not necessarily in the median.

5 Consequences for ordinary Thompson sampling

The frozen calculation yields an ordinary-TS lower bound whenever a competing arm can keep drawing above a threshold while the optimal-arm posterior remains frozen.

Theorem 6 (Ordinary-TS lockout)**.** *Consider two Gamma arms with $\mu_1 > \mu_2$, two initial pulls per arm, and a common prior satisfying Assumption 1 with tail exponent τ. Then constants $c > 0$ and $T_0 < \infty$ exist such that, for $T \ge T_0$,*

$$R_T \ge c\Delta_2 \begin{cases} T^{(1-\tau)/(3/2-\tau)}, & \tau < 1, \\ \log T, & \tau = 1. \end{cases} \tag{26}$$

Thus Jeffreys TS has an $\Omega(\sqrt{T})$ lockout contribution and is not uniformly efficient; matching TS has a distinct $\Omega(\log T)$ lockout mechanism.

Proof sketch. Choose $0 < m_- < m_+ < a < a' < \mu_2 < \mu_1$. Denote the mean and log-dispersion of the two initial optimal-arm rewards by $\bar{X}_{1,2}$ and $H_{1,2}$. The event $\{\bar{X}_{1,2} \in [m_-, m_+],\ H_{1,2} \le \varepsilon\}$ has positive probability, and its two restrictions are independent by (19). For arm 2, a positive-drift argument gives positive probability that every running empirical mean exceeds a' while $(\bar{Y}_s, H_s) \to (\mu_2, h(k_2))$. On every such reward path, Assumption 1 and the exact posterior imply $q_s := \mathbb{P}(\widetilde{\mu}_{2,s} \le a \mid Y_{1:s}) \le C(\omega) e^{-cs}$ eventually. Therefore $\prod_{s \ge 2}(1 - q_s) > 0$: with positive joint probability, every arm-2 posterior draw exceeds a.

Until the frozen optimal posterior first draws above a, TS must then choose arm 2. Conditional on the two optimal observations, this delay is geometric with parameter p_a. Multiplying (25) by the positive coupling probability proves (26). Section D proves the random-walk, posterior-tail, infinite-product, and integration steps separately. □

Under uniform efficiency, matching TS must satisfy both the information and lockout lower bounds. The same pulls serve both, so the joint bound is the larger coefficient, not their sum: the $\Omega(\log T)$ lockout order is compatible with asymptotic optimality, and coefficient optimality requires comparing the two constants.

5.1 Sequential upper bounds

We separate suboptimal-arm overshoots from optimal-arm pessimism; the next lemma charges each counted draw to the arm's next sample index.

Lemma 7 (Selected-arm posterior overshoots). *Fix $\theta = (\mu, k)$ with $\mu < b$ and a prior satisfying Assumption 1. Let $K_0 = K_{\inf}(\theta, b)$. For any $\epsilon > 0$, put $L_T = \lceil (1+\epsilon)\log(T)/K_0 \rceil$. If $\widetilde{\mu}_{a,t}$ is the fresh posterior draw of arm a at round t, then*

$$(O') \qquad \mathbb{E}_\theta \sum_{t=1}^{T} \mathbf{1}\{A_t = a,\ N_a(t-1) \geq L_T,\ \widetilde{\mu}_{a,t} \geq b\} = o(1). \tag{27}$$

Proof sketch. Choose a compact neighborhood U of θ on which $K_{\inf}(\vartheta, b) \geq (1-\delta)K_0$, with $(1+\epsilon)(1-\delta) > 1$. Exact likelihood ratios and three shape ranges give, uniformly on $\widehat{\theta}_n \in U$, $\Pi(\widetilde{\mu} \geq b \mid X_{1:n}) \leq Cn^\nu e^{-n(1-\delta)K_0}$, where Π denotes posterior probability and $\widehat{\theta}_n$ is the Gamma MLE determined by $(\bar{X}_n, H_n)$. The joint moment generating function of $(X, \log X)$ is finite around zero, so $\mathbb{P}_\theta(\widehat{\theta}_n \notin U) \leq Ce^{-cn}$. The typical contribution is at most $T \sup_{n \geq L_T} Cn^\nu e^{-n(1-\delta)K_0} = o(1)$. On the atypical event, each summand in (27) increments N_a, so each n is charged at most once and the expectation is bounded by $\sum_{n \geq L_T} Ce^{-cn} = o(1)$. Section E proves all three posterior-integral bounds. □

For the optimal arm, let $p_n(b) = \mathbb{P}(\widetilde{\mu} \geq b \mid X_{1:n})$. The second quantity in the standard TS decomposition is

$$\text{(U)} \qquad \sum_{n=2}^{\infty} \mathbb{E}_{\theta^\star} \frac{1 - p_n(b)}{p_n(b)} < \infty, \qquad b < \mu^\star. \tag{28}$$

For the density bound, define

$$h_n(k) = \psi(nk) - \psi(k) - \log n, \qquad \kappa_n(u) = h_n^{-1}(u), \tag{29}$$

and write the exact sufficient-statistic density as $f_{H_n|k}(u) = a_n(u)R_n(k)e^{-nku}$, where a_n is independent of k and R_n is the factor in theorem 5.

Proof roadmap. The sharp-regret reduction in proposition 9 combines lemma 7 with (28). Theorem 10 verifies (28) for ITM using the next bound and the small-shape carrier bound. Under shape κ, $n\kappa H_n$ has a Gamma limit for fixed n as $\kappa \to \infty$; for growing n, local-limit control of the standardized center density is uniform over $\kappa \geq 1$.

Theorem 8 (Uniform boundary saddlepoint control). *There is a universal constant $C < \infty$ such that, for all $n \geq 2$ and all $u > 0$ with $\kappa = \kappa_n(u) \geq 1$,*

$$(BSP) \qquad C^{-1} \leq \frac{a_n(u)R_n(\kappa)e^{-n\kappa u}}{\sqrt{n/\{-h_n'(\kappa)\}}} \leq C. \tag{30}$$

Here universal means independent of n, u, and κ.

The ratio in (30) is the density of standardized H_n at its center. Gauss' multiplication formula represents nH_n as a sum of $n-1$ independent log-Beta variables, giving uniform control in both regimes (section G.3). The one-sided carrier bound in lemma 16 covers $0 < \kappa \leq 1$; the global carrier estimates combine the two bounds.

Proposition 9 (Sharp-regret reduction). *Consider the finite-arm TS policy of section 2, with a unique optimal arm and a common prior satisfying Assumption 1. If* (28) *holds for every threshold $b < \mu^\star$ sufficiently close to $\mu^\star$, then every suboptimal arm satisfies*

$$\limsup_{T \to \infty} \frac{\mathbb{E}N_a(T)}{\log T} \leq \frac{1}{K_{\inf}((\mu_a, k_a), \mu^\star)}. \tag{31}$$

If the policy is uniformly efficient on the stated model class, (6) *makes this an equality.*

Proof sketch. Fix a suboptimal arm a and $b \in (\mu_a, \mu^\star)$ for which (28) holds. High posterior draws are bounded by the deterministic first L_T pulls and lemma 7. For low draws, condition on an interval with n optimal-arm observations and compare a low-draw suboptimal selection with the event that the optimal draw exceeds b while all competitors remain below b. Conditional independence gives the ratio $(1-p_n)/p_n$, even if a competitor preempts another optimistic optimal draw. Summing over n uses (28); taking $T \to \infty$, $\epsilon \downarrow 0$, and then $b \uparrow \mu^\star$ through admissible thresholds yields the information constant. The full stopped-interval calculation is in section F. □

Theorem 10 (Sharp regret of uncapped ITM)**.** *For every fixed $\eta > 0$ and every finite number of arms with a unique optimal mean, uncapped ITM-TS is uniformly efficient and, for each suboptimal arm,*

$$\lim_{T\to\infty} \frac{\mathbb{E}N_a(T)}{\log T} = \frac{1}{K_{\inf}((\mu_a, k_a), \mu^\star)}. \tag{32}$$

Uniform efficiency is instancewise as defined in section 2; finite-time constants may depend on the instance and on η.

Proof sketch. The proof turns (28) into a uniform integral bound and isolates the noncompact shape boundary. Condition on the optimal-arm sufficient statistics $(m, u) = (\bar{X}_n, H_n)$. The ratio $(1 - p_n(b))/p_n(b)$ is a quotient of two one-dimensional shape integrals. The global comparison in lemma 14 bounds both integrals at their exact finite-sample saddlepoints. Multiplication by the joint density of (m, H_n) cancels the empirical saddlepoint. BSP controls the remaining center density when $\kappa_n(u) \geq 1$, and lemma 16 covers the full range $0 < \kappa_n(u) \leq 1$. The exact correction $\log R_n(k) = nB(k)+s(nk)$ is retained through the cancellation, where $B(k) = k\log k-k-\log\Gamma(k)$ and $s(a) = \log\Gamma(a) - a\log a + a$.

For $m \leq b$, let $r = m/b$, $\widehat{k} = h^{-1}(u)$, $\kappa = \kappa_n(u)$, and $q = h^{-1}\{u + \mathcal{I}(r)\}$. The population exponent left after the cancellation is the difference between the sampling rate and the optimistic information projection. The scalar root in (15) gives

$$D((m, \widehat{k})\|(\mu^\star, k^\star)) - K_{\inf}((m, \widehat{k}), b) \geq k^\star\mathcal{I}(b/\mu^\star) =: \delta_b > 0, \qquad m \leq b. \tag{33}$$

The exact decomposition also retains the nonnegative equal-mean shape divergence. After $u = h_n(\kappa)$, the remaining boundary integral has measure $d\kappa/\kappa$ and factor $\{(1 + q)/(1 + \kappa)\}^\eta$. Its proof first splits $r = m/b$ near zero from r bounded away from zero. The latter part is then partitioned into small, compact, and high κ; the high-κ part is split once more according to whether q is bounded. The ITM factor controls bounded q through $\kappa^{-1-\eta}$, while the equal-mean shape divergence controls $q \downarrow 0$ and $q \to \infty$. For $m > b$, the downward information projection and the sampling rate have a strictly positive joint gap. These two mean regions yield constants $C, \nu, c > 0$, depending at most on the fixed $(\eta, \theta^\star, b)$ and not on n, such that

$$\mathbb{E}_{\theta^\star} \frac{1 - p_n(b)}{p_n(b)} \leq Cn^\nu e^{-cn}, \tag{34}$$

so (28) follows. Combining (28), lemma 7, and proposition 9 gives $\mathbb{E}N_a(T) = O(\log T)$ for each suboptimal arm. This bound implies uniform efficiency, so the Burnetas–Katehakis lower bound turns the upper coefficient into equality. Section G gives the complete shape-integral and boundary calculation. □

For fixed $\eta > 0$, $\widetilde{w}_\eta(k) = A'(k)(1 + k)^{-\eta}$ gives a first-order matching prior $\widetilde{w}_\eta(k)/\mu$ with ITM's sharp coefficient. By global comparison, this holds for positive continuous weights with finite positive endpoint limits of $\widetilde{w}/w_\eta$ (section G.6). The high-shape recovery integral contains $\int_n^\infty \kappa^{-1-\eta}d\kappa = n^{-\eta}/\eta$, and the pessimism budget in (28) diverges as $\eta \downarrow 0$.

Remark 11 (Compact support)**.** On a compact rectangle Θ, the information constant becomes $K_{\inf}^\Theta$. It equals $K_{\inf}$ exactly when the unique unrestricted optimizer $(\mu^\star, q_{\mu^\star})$ from theorem 3 belongs to Θ; section H proves this identity.

Remark 12 (More than two arms)**.** For fixed finite K, condition (28) concerns only the optimal-arm posterior. Conditional independence and the finite probability ratio in the stopped-interval argument make additional arms change finite sums while preserving the coefficient in (32).

6 Numerical study and discussion

Frozen-posterior phase transition. For 5,000 independent pairs from $P_{1.2,2}$ with $b = 1$, we compute p_b by adaptive one-dimensional quadrature from (12)–(13), and evaluate the conditional truncated waiting mean exactly. For M trials, CVaR_α averages the largest $\lceil(1 - \alpha)M\rceil$ values; a history is severe at horizon T if $p_b < T^{-1}$.

The predicted separation appears in the mean and upper tail while the medians remain nearly identical. At $T = 10^6$, ITM with $\eta = .5$ has mean/$\text{CVaR}_{.99}$ $1.48/24.67$, versus $33.01/3113.37$ for matching and $785.33/76642.70$ for Jeffreys.

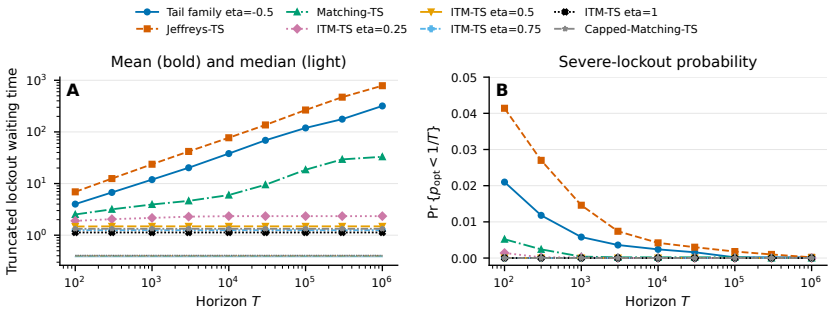

Figure 2: Frozen-posterior transition over 5,000 histories, evaluated by one-dimensional quadrature.

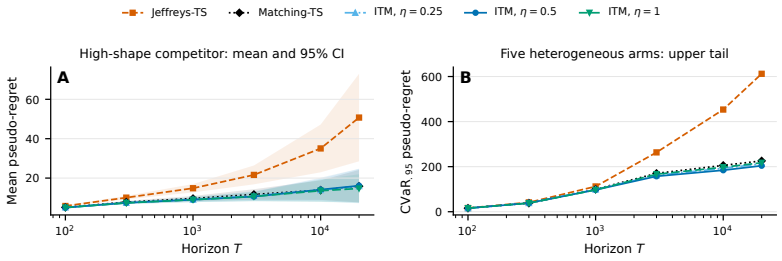

Figure 3: Ordinary TS over 1,024 common-random-number trials: high-shape mean regret with 95% intervals (left) and five-arm $\mathrm{CVaR}_{.95}$ (right). Section I gives all settings.

Full-bandit benchmark. Adaptive quadrature and Gamma-envelope rejection sampling support 1,024 trials through $T = 20{,}000$ on high-shape, small-gap, and five-arm instances. Policies share rewards and use independent posterior streams. Severe lockout means at least $.25T$ consecutive rounds without an optimal-arm pull. Section I.2 checks rare optimism by independent integration, validates the rejection target, and compares complete runs with full quadrature (64 trials per instance, $T = 2{,}000$).

Table 3 reports all mean-regret and CVaR values. For ITM at $\eta = .5$ minus matching, pointwise paired 95% intervals for mean regret are $[-2.39, -.16]$ on the five-arm instance, $[-2.77, 4.05]$ on the high-shape instance, and $[-.82, .34]$ on the small-gap instance. The five-arm instance supports improvements in mean regret and CVaR; both two-arm instances have intervals containing zero for both metrics. Rankings for $\eta = .25, .5, 1$ vary by instance and metric; all share (32).

Additional comparisons. We adapt the per-arm ϵ-TS rule (Jin et al., 2023) to the two-parameter posterior: each arm independently uses a posterior sample with probability $1/K$ and its empirical mean otherwise. The matching and ITM adaptations use less CPU time but have higher mean regret and CVaR than ordinary ITM ($\eta = .5$) on all three Gamma instances. Mean amortized CPU time is 4.77–7.70 microseconds per decision for ordinary ITM and 2.14–3.28 for adapted ϵ-ITM. In six moment-matched lognormal/Gamma-mixture settings, ITM–matching intervals include zero for both metrics; and we provide details in section J.

7 Conclusion

The false-precision failure of Gamma TS is governed by the high-shape prior tail. Tempering the matching weight by $[1 + e^{2A(k)}]^{-\eta}$ with any fixed $\eta > 0$ makes the optimal-arm pessimism odds summable, so ordinary, uncapped Gamma TS attains the information constant with unknown means and shapes. Any regular weight globally comparable to ITM inherits that constant.

AI USE STATEMENT

During manuscript preparation, the authors used GPT-6 exclusively for language polishing, grammar correction, and enhancing textual readability. No generative-AI tools contributed to research ideation, experimental design, data analysis, or the formulation of scientific claims. All suggested textual revisions were carefully reviewed, modified, and validated by the authors, who assume full responsibility for the content of this submission.

REPRODUCIBILITY STATEMENT

The main paper states the model assumptions and theorem conditions. The appendix gives complete derivations of the exact results, the small-ball and lockout proofs, the ordinary-TS coupling, the posterior-overshoot lemma, the boundary local limit theorem, and the compact-support projection identity. The supplement contains the source code, fixed random seeds, experiment configurations, numerical tests, and machine-readable summaries. Frozen-posterior probabilities use one-dimensional quadrature with separate tail integration for rare optimism events. Full-bandit draws combine adaptive quadrature with Gamma-envelope rejection sampling, with numerical validation reported in the appendix.

REFERENCES

Apostolos N. Burnetas and Michael N. Katehakis. Optimal adaptive policies for sequential allocation problems. *Advances in Applied Mathematics*, 17(2):122–142, 1996. doi: 10.1006/aama.1996.0007.

Thomas J. DiCiccio and G. Alastair Young. Objective bayes and conditional inference in exponential families. *Biometrika*, 97(2):497–504, 2010. doi: 10.1093/biomet/asq002.

Dean Eckles and Maurits Kaptein. Thompson sampling with the online bootstrap. *arXiv preprint arXiv:1410.4009*, 2014. doi: 10.48550/arXiv.1410.4009. URL https://arxiv.org/abs/1410.4009.

Matthieu Geist. A multiplicative UCB strategy for gamma rewards. In *The 12th European Workshop on Reinforcement Learning (EWRL 2015)*, 2015. URL https://ewrl.wordpress.com/wp-content/uploads/2015/02/ewrl12_2015_submission_4.pdf.

Junya Honda and Akimichi Takemura. Optimality of thompson sampling for gaussian bandits depends on priors. In *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics*, volume 33 of *Proceedings of Machine Learning Research*, pp. 375–383. PMLR, 2014. URL https://proceedings.mlr.press/v33/honda14.html.

Tianyuan Jin, Pan Xu, Xiaokui Xiao, and Anima Anandkumar. Finite-time regret of thompson sampling algorithms for exponential family multi-armed bandits. In *Advances in Neural Information Processing Systems 35*, pp. 38475–38487, 2022. doi: 10.52202/068431-2788.

Tianyuan Jin, Xianglin Yang, Xiaokui Xiao, and Pan Xu. Thompson sampling with less exploration is fast and optimal. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 15239–15261. PMLR, 2023. URL https://proceedings.mlr.press/v202/jin23b.html.

Nathaniel Korda, Emilie Kaufmann, and Rémi Munos. Thompson sampling for 1-dimensional exponential family bandits. In *Advances in Neural Information Processing Systems 26*, pp. 1448–1456, 2013. URL https://proceedings.neurips.cc/paper/2013/hash/aba3b6fd5d186d28e06ff97135cade7f-Abstract.html.

Tze Leung Lai and Herbert Robbins. Asymptotically efficient adaptive allocation rules. *Advances in Applied Mathematics*, 6(1):4–22, 1985. doi: 10.1016/0196-8858(85)90002-8.

Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020. doi: 10.1017/9781108571401.

Jongyeong Lee, Junya Honda, Chao-Kai Chiang, and Masashi Sugiyama. Optimality of thompson sampling with noninformative priors for pareto bandits. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202 of *Proceedings of Machine Learning Research*, pp. 18810–18851. PMLR, 2023. URL https://proceedings.mlr.press/v202/lee23d.html.

Jongyeong Lee, Chao-Kai Chiang, and Masashi Sugiyama. The choice of noninformative priors for thompson sampling in multiparameter bandit models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(12):13383–13390, 2024. doi: 10.1609/aaai.v38i12.29240. URL https://ojs.aaai.org/index.php/AAAI/article/view/29240.

Aadirupa Saha and Branislav Kveton. Only pay for what is uncertain: Variance-adaptive thompson sampling. *arXiv preprint arXiv:2303.09033*, 2023. doi: 10.48550/arXiv.2303.09033. URL https://arxiv.org/abs/2303.09033.

William R. Thompson. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, 25(3/4):285–294, 1933. doi: 10.1093/biomet/25.3-4.285.

Kaizheng Wang. MINTS: Minimalist thompson sampling. *arXiv preprint arXiv:2606.01655*, 2026. doi: 10.48550/arXiv.2606.01655. URL https://arxiv.org/abs/2606.01655.

A Special-function and posterior details

Throughout the appendix, the variable tending to a boundary is stated next to each asymptotic expansion. The notation $O_n(\cdot)$ means that n is fixed and that the hidden constant may depend on n. Constants in bounds that are uniform in n are written independently of n; their dependence on fixed model parameters, the threshold b, and the fixed tempering exponent η is indicated where it matters. We write γ_{E} for the Euler–Mascheroni constant and $\zeta(s)$ for the Riemann zeta function. For nonnegative quantities, $f \lesssim g$ means $f \le Cg$, $f \gtrsim g$ means $f \ge cg$, and $f \asymp g$ means both inequalities hold, with constants having the dependence stated in the surrounding text.

A.1 The information coordinate at both boundaries

As $k \downarrow 0$, the convergent local expansions are

$$\begin{aligned}\psi(k) &= -k^{-1} - \gamma_{\mathrm{E}} + \zeta(2)k + O(k^2),\\ \log\Gamma(k) &= -\log k - \gamma_{\mathrm{E}}k + \tfrac{1}{2}\zeta(2)k^2 + O(k^3),\\ \psi_1(k) &= k^{-2} + \zeta(2) + O(k).\end{aligned}$$

Substituting the first two lines into $A(k) = k\psi(k) - k - \log\Gamma(k)$ gives

$$A(k) = \log k - 1 - k + \tfrac{1}{2}\zeta(2)k^2 + O(k^3).$$

Substituting the third line into $A'(k) = k\psi_1(k) - 1$ gives $A'(k) = k^{-1} - 1 + \zeta(2)k + O(k^2)$.

As $k \to \infty$, the remainders below have the displayed orders. More precisely, for each fixed $K_0 > 0$, their big-O constants can be chosen uniformly on $k \ge K_0$ by enlarging the asymptotic constants over the compact interval between K_0 and the start of the asymptotic regime:

$$\begin{aligned}\psi(k) &= \log k - \frac{1}{2k} - \frac{1}{12k^2} + O(k^{-4}),\\ \log\Gamma(k) &= (k - \tfrac{1}{2})\log k - k + \tfrac{1}{2}\log(2\pi) + \frac{1}{12k} + O(k^{-3}),\\ \psi_1(k) &= k^{-1} + \frac{1}{2k^2} + \frac{1}{6k^3} + O(k^{-5}).\end{aligned}$$

Termwise substitution yields

$$A(k) = \tfrac{1}{2}\log\frac{k}{2\pi e} - \frac{1}{6k} + O(k^{-3}), \qquad A'(k) = \frac{1}{2k} + \frac{1}{6k^2} + O(k^{-4}).$$

In particular $e^{2A(k)} = k(2\pi e)^{-1}\{1 + O(k^{-1})\}$, which proves the constant in (10).

A.2 MARGINALIZATION AND EXACT BOUNDARY POWERS

Starting from (3), the joint posterior kernel is

$$w(k)\frac{k^{nk}}{\Gamma(k)^n}\left(\prod_{i=1}^{n} x_i^{k-1}\right)\mu^{-nk-1}e^{-kS/\mu}.$$

Under $v = kS/\mu$, $\mu = kS/v$ and $|d\mu| = kSv^{-2}dv$, so

$$\begin{aligned}\int_0^\infty \mu^{-nk-1}e^{-kS/\mu}d\mu &= \int_\infty^0 (kS/v)^{-nk-1}e^{-v}(-kSv^{-2})dv \\ &= (kS)^{-nk}\int_0^\infty v^{nk-1}e^{-v}dv \\ &= (kS)^{-nk}\Gamma(nk).\end{aligned}$$

Also

$$\prod_i x_i^{k-1} = \exp\{(k-1)\textstyle\sum_i \log x_i\} = \bar{x}^{n(k-1)}e^{-n(k-1)H_n}.$$

All factors depending on the data but not k cancel from normalization, leaving (13).

Fix an integer $n \geq 2$ and let $k \downarrow 0$. Using

$$\log\Gamma(k) = -\log k - \gamma_{\mathrm{E}}k + \frac{\zeta(2)}{2}k^2 - \frac{\zeta(3)}{3}k^3 + O(k^4)$$

at k and at nk gives

$$\log\Gamma(nk) - n\log\Gamma(k) = -\log n + (n-1)\log k + \frac{\zeta(2)}{2}n(n-1)k^2 + O_n(k^3).$$

The coefficient of k is zero: the term $-\gamma_{\mathrm{E}}nk$ from $\log\Gamma(nk)$ cancels the term $-n\gamma_{\mathrm{E}}k$ from $n\log\Gamma(k)$. Exponentiating yields

$$\frac{\Gamma(nk)}{\Gamma(k)^n} = \frac{1}{n}k^{n-1}\left\{1 + \frac{\pi^2}{12}n(n-1)k^2 + O_n(k^3)\right\}. \tag{35}$$

Separately,

$$n^{-nk} = e^{-nk\log n} = 1 - n(\log n)k + \frac{n^2(\log n)^2}{2}k^2 + O_n(k^3).$$

Multiplication of these two expansions gives

$$\begin{aligned}\left\{\frac{\Gamma(nk)}{\Gamma(k)^n}\right\}n^{-nk} &= \frac{1}{n}k^{n-1}\Big[1 - n(\log n)k \\ &\qquad + \left\{\frac{n^2(\log n)^2}{2} + \frac{\pi^2}{12}n(n-1)\right\}k^2 + O_n(k^3)\Big] \\ &= \frac{1}{n}k^{n-1}\{1 + O_n(k)\}.\end{aligned} \tag{36}$$

Here n^{-nk} multiplies the Gamma ratio. Since $w_\eta(k) = k^{-1}\{1 + O(k)\}$ and $w_J(k) = k^{-1/2}\{1 + O(k)\}$ as $k \downarrow 0$, the corresponding marginal shape kernels are respectively

$$\frac{1}{n}k^{n-2}\{1 + O_n(k)\} \quad \text{and} \quad \frac{1}{n}k^{n-3/2}\{1 + O_n(k)\}.$$

Both powers exceed -1 for every $n \geq 2$, which proves local integrability at the lower boundary. For the same fixed n, Stirling's formula as $k \to \infty$ gives

$$\begin{aligned}\left\{\frac{\Gamma(nk)}{\Gamma(k)^n}\right\}n^{-nk} &\sim \frac{\sqrt{2\pi}(nk)^{nk-1/2}e^{-nk}}{(2\pi)^{n/2}k^{nk-n/2}e^{-nk}}n^{-nk} \\ &= (2\pi)^{(1-n)/2}n^{-1/2}k^{(n-1)/2}.\end{aligned}$$

Together with $H_n > 0$ almost surely, so that e^{-nkH_n} controls the polynomial upper-boundary factor, these expressions justify every power used in theorem 2.

B PROJECTION AND INFORMATION-GAP ALGEBRA

Under $P_{\mu,k}$, $\mathbb{E}X = \mu$ and $\mathbb{E}\log X = \psi(k) + \log(\mu/k)$. The pointwise log-density ratio is

$$\log \frac{f(x \mid \mu, k)}{f(x \mid m, q)} = k \log \frac{k}{\mu} - q \log \frac{q}{m} + \log \Gamma(q) - \log \Gamma(k) + (k-q)\log x - \frac{kx}{\mu} + \frac{qx}{m}.$$

Taking its expectation and substituting the two moments gives (17). The derivative steps in theorem 3 then prove the constrained minimizer.

The final cancellation in theorem 3 is termwise. Let $q = q_b$, put $r = \mu/b$, and rewrite the root equation as

$$\log \frac{kb}{q\mu} = \psi(k) - \psi(q) + 1 - r.$$

Substitution into (17) at $m = b$ yields

$$\begin{aligned} D((\mu,k)\|(b,q)) &= \log\Gamma(q) - \log\Gamma(k) + (k-q)\psi(k) \\ &\quad + q\{\psi(k) - \psi(q) + 1 - r\} - k + qr \\ &= \log\Gamma(q) - \log\Gamma(k) + k\psi(k) - q\psi(q) + q - k \\ &= \{k\psi(k) - k - \log\Gamma(k)\} - \{q\psi(q) - q - \log\Gamma(q)\} \\ &= A(k) - A(q). \end{aligned}$$

The two terms $-qr$ and $+qr$ cancel explicitly, as do the two copies of $q\psi(k)$ generated by $(k-q)\psi(k)$ and the root substitution.

To derive (33), let $m \le b$ and let q satisfy $h(q) = h(k) + \mathcal{I}(m/b)$, and abbreviate $c_b = b/\mu^\star$, $r = m/b$. Apply (16) to the empirical source (m, k):

$$K_{\inf}((m,k), b) = A(k) - A(q).$$

Insert $(\mu, k) = (m, k)$ and $(m, q) = (\mu^\star, k^\star)$ into (17), subtract $A(k) - A(q)$, and use $h(k) = h(q) - \mathcal{I}(r)$. After cancellation of the k terms,

$$\begin{aligned} &D((m,k)\|(\mu^\star,k^\star)) - [A(k) - A(q)] \\ &\quad = A(q) - A(k^\star) + k^\star\{h(q) - h(k^\star)\} + k^\star\{-\log c_b - (1-c_b)r\}. \end{aligned}$$

The first two terms form the equal-mean shape divergence. Hence the exact decomposition is

$$D((m,k)\|(\mu^\star,k^\star)) - K_{\inf}((m,k),b) = J_{k^\star}(q) + k^\star\{-\log c_b - (1-c_b)r\}, \tag{37}$$

where

$$J_{k^\star}(q) := D((m,q)\|(m,k^\star)) = A(q) - A(k^\star) + k^\star\{h(q) - h(k^\star)\}.$$

This divergence does not depend on the common mean. Its derivative is

$$J'_{k^\star}(q) = A'(q)\left(1 - \frac{k^\star}{q}\right),$$

so it decreases on $(0, k^\star)$, increases on $(k^\star, \infty)$, and vanishes only at $q = k^\star$. The expansions in section A also give

$$J_{k^\star}(q) \asymp q^{-1} \quad (q \downarrow 0), \qquad J_{k^\star}(q) = \tfrac{1}{2}\log q + O(1) \quad (q \to \infty). \tag{38}$$

Since $r \le 1$ and $0 < c_b < 1$,

$$-\log c_b - (1-c_b)r \ge -\log c_b - (1-c_b) = c_b - 1 - \log c_b = \mathcal{I}(c_b),$$

which proves (33).

C Full proof of the frozen-posterior theorem

C.1 Coordinate proof of the small-ball law

Use coordinates $v_1, \dots, v_{n-1}$ with $z_i = n^{-1} + v_i$ for $i < n$ and $z_n = n^{-1} - \sum_{i<n} v_i$. On a sufficiently small Euclidean ball, all z_i lie in a compact subset of $(0, 1)$. Taylor's theorem with a uniform third-derivative bound gives

$$\begin{aligned} H_n(z) &= -\log n - \frac{1}{n}\sum_{i=1}^{n} \log(n^{-1} + v_i) \\ &= -\frac{1}{n}\sum_{i} \left[n v_i - \frac{n^2 v_i^2}{2} + O(|v_i|^3) \right] \\ &= \frac{n}{2}\sum_{i} v_i^2 + O(\|v\|^3), \end{aligned}$$

because $\sum_i v_i = 0$. The quadratic form $Q(v) = \frac{n}{2}\sum_i v_i^2$ is positive definite on the tangent space, so $c_n\|v\|^2 \le Q(v) \le C_n\|v\|^2$. The Taylor remainder therefore satisfies $|H_n(v) - Q(v)| \le C_n Q(v)^{3/2}$ near the origin. Consequently there are $C, u_0 > 0$ such that, for $0 < u < u_0$,

$$\{Q(v) \le u(1 - C\sqrt{u})\} \subseteq \{H_n \le u\} \subseteq \{Q(v) \le u(1 + C\sqrt{u})\}.$$

Thus the relative radial error is $O(\sqrt{u})$.

The Dirichlet density in these coordinates is

$$g(z) = \frac{\Gamma(nk_0)}{\Gamma(k_0)^n} \prod_i z_i^{k_0 - 1}.$$

It is continuously differentiable and $g(z) = g(z_0)\{1 + O(\|v\|)\}$ near $z_0 = (1/n, \dots, 1/n)$. If V_Q is the volume of $\{Q(v) \le 1\}$ in these coordinates, integration over the sandwiched ellipsoids gives

$$\mathbb{P}(H_n \le u) = g(z_0) V_Q u^{(n-1)/2}\{1 + O(u^{1/2})\}.$$

The constant is finite and strictly positive.

C.2 Uniform numerator and denominator bounds

Let $R_n(k) = \Gamma(nk)\{\Gamma(k)^n n^{nk}\}^{-1}$ and $r = m/b$. Here n, r_-, and r_+ are fixed, while the limits below are in H. By regular variation, for every $\epsilon \in (0, 1)$ there is K such that

$$(1 - \epsilon)C_\infty k^{-\tau} \le w(k) \le (1 + \epsilon)C_\infty k^{-\tau}, \qquad k \ge K.$$

The fixed-n Stirling expansion gives $R_n(k) \asymp_n k^d$, $d = (n-1)/2$, after increasing K. Therefore constants $c_1, c_2 > 0$, independent of H and $r \in [r_-, r_+]$, exist for which

$$c_1 \int_K^\infty k^{\beta-1} e^{-nkH} dk \le \int_K^\infty w(k) R_n(k) e^{-nkH} dk \le c_2 \int_K^\infty k^{\beta-1} e^{-nkH} dk.$$

Write

$$I_{K,\beta}(H) = \int_K^\infty k^{\beta-1} e^{-nkH}\, dk.$$

For $a, x > 0$, $\Gamma(a, x) := \int_x^\infty v^{a-1} e^{-v} dv$ denotes the upper incomplete Gamma function; $\Gamma(a, 0) = \Gamma(a)$. If $\beta > 0$, the substitution $v = nkH$, namely $k = v/(nH)$ and $dk = dv/(nH)$, gives

$$\begin{aligned} I_{K,\beta}(H) &= (nH)^{-\beta} \int_{nKH}^\infty v^{\beta-1} e^{-v}\, dv \\ &= (nH)^{-\beta}\Gamma(\beta, nKH) \sim n^{-\beta}\Gamma(\beta) H^{-\beta}. \end{aligned}$$

Thus the exponent is $-\beta = -(d + 1 - \tau)$, including the Jacobian contribution. If $\beta = 0$,

$$I_{K,0}(H) = \int_{nKH}^\infty v^{-1} e^{-v}\, dv = \log(1/H) - \log(nK) - \gamma_{\mathrm{E}} + o(1).$$

If $\beta < 0$, then $I_{K,\beta}(0) = \int_K^\infty k^{\beta-1}dk = K^\beta/(-\beta) < \infty$, and monotone convergence gives $I_{K,\beta}(H) \uparrow I_{K,\beta}(0)$ as $H \downarrow 0$. The contribution from $(0, K)$ tends to a finite strictly positive number by posterior propriety and dominated convergence.

For the numerator, if $Y \sim \mathrm{Gamma}(nk, 1)$ and $0 < r < 1$, then for every $\lambda > 0$,

$$\mathbb{P}(Y \le nkr) \le e^{\lambda nkr}\mathbb{E}e^{-\lambda Y} = \exp\{nk[\lambda r - \log(1+\lambda)]\}.$$

Optimizing at $\lambda = r^{-1} - 1$ gives $P(nk, nkr) \le e^{-nk\mathcal{I}(r)}$. Since $\mathcal{I}$ decreases on $(0, 1)$,

$$P(nk, nkr) \le e^{-nk\mathcal{I}(r_+)}$$

uniformly in the chosen interval. At $H = 0$, the tail of the numerator is therefore dominated by $Ck^{d-\tau}e^{-nk\mathcal{I}(r_+)}$, which is integrable. On any fixed $[k_1, k_2] \subset (0, \infty)$, continuity on the compact set $[k_1, k_2] \times [r_-, r_+]$ gives

$$\inf_{k\in[k_1,k_2],\, r\in[r_-,r_+]} P(nk, nkr) > 0.$$

Near zero, (36) and the prior assumption give $w(k)R_n(k) \le Ck^{n-1-\rho}$; this is integrable because $n \ge 2$ and $\rho < 2$. These lower-boundary, compact, and upper-tail envelopes cover the whole k-axis. Uniform dominated convergence now gives $N(H, r) \to N(0, r)$ uniformly in r, with $0 < c \le N(0, r) \le C < \infty$. Dividing by the previously established order of $Z(H)$ proves (23) with two-sided constants that are uniform over $r \in [r_-, r_+]$.

The reciprocal criterion also requires control away from $H = 0$. Fix $\epsilon > 0$ and choose $K > 0$. The function $P(nk, nkr)$ extends continuously to $k = 0$ with value one, so

$$c_K := \inf_{0<k\le K,\, r\in[r_-,r_+]} P(nk, nkr) > 0.$$

Thus the numerator is at least c_K times the normalizer restricted to $(0, K]$, and for large H that restricted mass dominates the normalizer. Positivity and continuity of wR_n on $[K/2, 3K/4]$ give

$$\int_{K/2}^{3K/4} w(k)R_n(k)e^{-nkH}dk \ge ce^{-3nKH/4}.$$

Choose $L > 3K/2$. Boundedness of wR_n on $[K, L]$ gives an upper bound Ce^{-nKH} there. On $[L, \infty)$, a polynomial envelope $w(k)R_n(k) \le C(1+k)^M$ and the split $e^{-nkH} = e^{-nkH/2}e^{-nkH/2}$ give, for $H \ge 1$,

$$\int_L^\infty w(k)R_n(k)e^{-nkH}dk \le Ce^{-nLH/2}\int_L^\infty (1+k)^M e^{-nk/2}dk.$$

Both upper bounds are $o(e^{-3nKH/4})$. Hence the posterior mass of $(0, K]$ tends to one as $H \to \infty$, uniformly in r, and $p_b(H, r) \ge c_K/2$ for all sufficiently large H. On the compact set $[\epsilon, H_0] \times [r_-, r_+]$, both posterior integrals are continuous and strictly positive, so their ratio has a positive minimum. It follows that p_b^{-1} is bounded on $H \ge \epsilon$; only the small-ball integral below determines (24).

C.3 Reciprocal and truncated-waiting integrals

A dyadic-shell argument applies directly to the distribution expansion (21). Take $\epsilon > 0$ small and define $D_j = (2^{-j-1}\epsilon, 2^{-j}\epsilon]$. The remainder in (21) implies, after reducing ϵ if needed,

$$c(2^{-j}\epsilon)^d \le \mathbb{P}(H_n \in D_j) \le C(2^{-j}\epsilon)^d, \qquad j \ge 0. \tag{39}$$

Independence of m and H_n in (19) leaves these bounds unchanged after conditioning on $r = m/b \in [r_-, r_+]$. When $\beta > 0$, the uniform estimate $p_b^{-1} \asymp H_n^{-\beta}$ on D_j gives

$$\mathbb{E}[p_b^{-1}; H_n \le \epsilon \mid r \in [r_-, r_+]] \asymp \sum_{j=0}^\infty (2^{-j}\epsilon)^{d-\beta}.$$

This geometric series is finite exactly when $d - \beta > 0$, equivalently $\tau > 1$. The commonly used integral notation $\int_0^\epsilon u^{d-1-\beta}du$ represents the same dyadic comparison. For $\beta = 0$, the shell contribution is bounded by $C(1+j)2^{-jd}$ and is summable; for $\beta < 0$, p_b^{-1} is bounded.

For the truncated waiting time, let $F_T(p) = [1-(1-p)^T]/p$. For $0 < p \le 1$,

$$c\min(T, p^{-1}) \le F_T(p) \le \min(T, p^{-1}) \tag{40}$$

with a universal $c > 0$. Indeed, $1-(1-p)^T \le \min\{1, Tp\}$ gives the upper bound, whereas $1-(1-p)^T \ge 1-e^{-pT}$ is at least $(1-e^{-1})pT$ when $pT \le 1$ and at least $1-e^{-1}$ when $pT > 1$. When $\tau < 1$, $\beta > d$ and the crossover $p_b \asymp T^{-1}$ occurs at $u_T = T^{-1/\beta}$. Applying (39) on the two sides of this crossover, or equivalently summing the corresponding geometric series, gives

$$\begin{aligned}\mathbb{E}W_T &\asymp T\int_0^{u_T} u^{d-1}du + \int_{u_T}^{\epsilon} u^{d-1-\beta}du + O(1)\\ &\asymp T^{1-d/\beta} = T^{(1-\tau)/(d+1-\tau)}.\end{aligned}$$

Here the $O(1)$ term is uniform in T and comes from $H_n > \epsilon$. At $\tau = 1$, $\beta = d$ and the second integral is $\int_{T^{-1/d}}^{\epsilon} u^{-1}du \asymp \log T$. For $\tau > 1$, (24) and dominated convergence bound the expectation uniformly in T.

D EMBEDDING LOCKOUT IN ORDINARY TS

This section proves each step summarized in theorem 6.

Lemma 13 (A persistent competitor path)**.** *Fix a Gamma arm P_{μ_2,k_2}, thresholds $0 < a < a' < \mu_2$, and a prior satisfying Assumption 1. There is a positive-probability joint event on which, for every sample count $s \ge 2$,*

$$\bar{Y}_s > a' \quad \textit{and} \quad \widetilde{\mu}_{2,s} > a. \tag{41}$$

Here a fresh posterior draw is made at each s.

Proof. Step 1: the reward path. Let $Z_j = Y_j - a'$, so $\mathbb{E}Z_j = \mu_2 - a' > 0$. The Gamma moment-generating function is finite around zero. Hence, for sufficiently small $\lambda > 0$, $\varrho_\lambda := \mathbb{E}e^{-\lambda Z_1} < 1$ (differentiate at $\lambda = 0$). If a finite initial block creates reserve $R > 0$, then for the subsequent partial sums

$$\mathbb{P}\left(\exists n \ge 1 : \sum_{j=1}^{n} Z_j \le -R\right) \le \sum_{n\ge 1}\mathbb{P}\Big(e^{-\lambda\sum_{j=1}^{n} Z_j} \ge e^{\lambda R}\Big) \le e^{-\lambda R}\sum_{n\ge 1}\varrho_\lambda^n.$$

Choose R large enough that the last quantity is below one. Fix a block length $s_0 \ge 2$ and require each of the first s_0 increments Z_j to be positive and their sum to exceed R. This finite-block event has positive probability because a Gamma law has positive density on every open interval in $(0, \infty)$. Every partial mean in the block then exceeds a', and the displayed maximal inequality shows that, conditional on the block, the later partial sums stay above the line sa' with positive probability. The strong laws for Y_j and $\log Y_j$ have conditional probability one as well, so intersecting them with the survival event preserves positive probability and yields

$$\bar{Y}_s > a' \ \forall s \ge 2, \qquad (\bar{Y}_s, H_s) \to (\mu_2, h(k_2)). \tag{42}$$

Step 2: an exponential posterior left tail on each such path. Fix a path satisfying (42). Eventually $\bar{Y}_s/a \ge 1+\delta$ and H_s lies in a compact interval around $h(k_2) > 0$. Conditional on k, the event $\mu \le a$ is $G \ge sk\bar{Y}_s/a$ for $G \sim \mathrm{Gamma}(sk, 1)$. The upper-tail Chernoff bound therefore gives, for $k \ge \kappa > 0$,

$$\mathbb{P}(\mu \le a \mid k, Y_{1:s}) \le \exp[-sk\mathcal{I}(\bar{Y}_s/a)] \le e^{-s\kappa\mathcal{I}(1+\delta)}. \tag{43}$$

To control posterior mass on $k < \kappa$, define $B(k) = k\log k - k - \log\Gamma(k)$, $\phi_H(k) = B(k) - kH$, and $\chi(a) = \log\Gamma(a) - a\log a + a$. The exact identity

$$\log R_s(k) - skH = s\phi_H(k) + \chi(sk)$$

implies, on every fixed compact subinterval of $(0, \infty)$,

$$\frac{1}{s}\log\left[\frac{\Gamma(sk)}{\Gamma(k)^s}s^{-sk}e^{-skH_s}\right] = \phi_{H_s}(k) + O\left(\frac{\log s}{s}\right) \tag{44}$$

uniformly in k. Moreover $\phi'_H(k) = h(k) - H$ and $\phi''_H(k) = -A'(k)/k < 0$, so $\phi_{h(k_2)}$ has the unique maximizer k_2. Also $\phi_H(k) \to -\infty$ as $k \downarrow 0$ and as $k \to \infty$ for H bounded away from zero.

Choose $0 < \kappa < k_2/2 < K$ and a compact interval $\mathcal{H}$ around $h(k_2)$ so that, for every $H \in \mathcal{H}$, the maximizer $h^{-1}(H)$ lies in $(2\kappa, K/2)$ and

$$\sup_{k \notin [\kappa, K]} \phi_H(k) \le \phi_H\{h^{-1}(H)\} - 4c_0.$$

Such constants exist because ϕ_H is strictly concave and tends to $-\infty$ at both boundaries, uniformly for $H \in \mathcal{H}$. On the $s^{-1/2}$ neighborhood of $h^{-1}(H_s)$, Taylor's theorem, $\chi(sk) = -\frac{1}{2}\log(sk) + O(1)$, and the positive lower bound of the prior give

$$\int_0^\infty w(k) R_s(k) e^{-skH_s} dk \ge c s^{-1} e^{s\phi_{H_s}\{h^{-1}(H_s)\}}. \tag{45}$$

The three boundary pieces have explicit upper bounds. For $0 < k < 1/s$, $\Gamma(sk) \le C/(sk)$, $\Gamma(k)^{-1} \le Ck$, and $w(k) \le Ck^{-\rho}$ yield

$$\int_0^{1/s} w(k) R_s(k) e^{-skH_s} dk \le \frac{C^s}{s} \int_0^{1/s} k^{s-1-\rho} dk \le C^s s^{-s+\rho-2}.$$

Its logarithm is $-s\log s + O(s)$, so after division by (45) it is smaller than e^{-cs} for all large s. On $[1/s, \kappa]$, the bounds $|\chi(sk)| \le C(1 + \log s)$ and $w(k) \le Ck^{-\rho}$, together with the uniform rate gap, give $Cs^C e^{s\phi_{H_s}\{h^{-1}(H_s)\} - 3c_0 s}$. For $k \ge K$, put $H_- = \inf \mathcal{H} > 0$. The estimates $B(k) \le C_B + \frac{1}{2}\log(1+k)$, $\chi(sk) \le C$, and $w(k) \le C(1+k)^M$ give

$$w(k) e^{s\phi_{H_s}(k) + \chi(sk)} \le C(1+k)^M \exp\left[s\{C_B + \tfrac{1}{2}\log(1+k) - H_- k\}\right].$$

Here $M < \infty$ is any exponent in a polynomial envelope for w; such an envelope follows from $w(k) = C_\infty k^{-\tau}\{1 + o(1)\}$ in Assumption 1. The saddle values $\phi_H\{h^{-1}(H)\}$ are bounded below for $H \in \mathcal{H}$. Increase K until the expression in braces is at most $\phi_H\{h^{-1}(H)\} - 4c_0 - H_- k/2$ uniformly over this compact set of H. Integrating the remaining factor $(1+k)^M e^{-sH_- k/2}$ gives $Cs^C e^{s\phi_{H_s}\{h^{-1}(H_s)\} - 3c_0 s}$. The first two pieces alone bound the posterior mass on $k < \kappa$; including the third gives the stronger outside-compact estimate. Dividing all three boundary bounds by (45) gives

$$\Pi_s(k \notin [\kappa, K] \mid Y_{1:s}) \le C(\omega) s^C e^{-c_0 s}. \tag{46}$$

Equations (43)–(46) imply

$$q_s := \mathbb{P}(\widetilde{\mu}_{2,s} \le a \mid Y_{1:s}) \le C'(\omega) e^{-c's} \tag{47}$$

after reducing $c' > 0$ to absorb the polynomial. The finitely many earlier q_s are strictly below one because the posterior has positive density on every compact subset of $\{\mu > a, k > 0\}$.

Step 3: posterior sampling. Conditional on the infinite reward path, the construction uses one fresh posterior randomization at each successive sample count s. Along the lockout path this is exactly one draw per round because selecting arm 2 increments its sample count. These randomizations are conditionally independent. Since $\sum_s q_s < \infty$ and $q_s < 1$,

$$\mathbb{P}(\widetilde{\mu}_{2,s} > a \ \forall s \ge 2 \mid Y_{1:\infty}) = \prod_{s=2}^{\infty} (1 - q_s) > 0.$$

Indeed, for all sufficiently large s, $q_s \le 1/2$ and $\log(1 - q_s) \ge -2q_s$, so the tail of the product is positive; its finitely many initial factors are positive as well. The conditional product is a strictly positive measurable function on the positive-probability reward-path event. Its integral is therefore positive, proving the joint event (41). □

Full proof of theorem 6. Let arm 1 be optimal and choose $0 < m_- < m_+ < a < a' < \mu_2 < \mu_1$. Write

$$\bar{X}_{1,2} = \frac{X_{1,1} + X_{1,2}}{2}, \qquad H_{1,2} = \log \bar{X}_{1,2} - \frac{1}{2}\{\log X_{1,1} + \log X_{1,2}\}.$$

Let $E_1(u) = \{\bar{X}_{1,2} \in [m_-, m_+],\ H_{1,2} \le u\}$. It has positive probability for every $u > 0$; by (19), its mean and $H_{1,2}$ parts are independent. Let E_2 be the positive-probability event from lemma 13. The two reward sequences and the pre-generated posterior randomizations for the two arms are independent, so $E_1(u)$ and E_2 are independent.

On E_2, if a frozen arm-1 posterior draw is at most a, the arm-2 draw is larger and TS selects arm 2. In this paragraph $p_a = \mathbb{P}(\widetilde{\mu}_1 > a \mid X_{1:2})$; it is the quantity p_b from (18) with threshold b replaced by a. Conditional on the initial arm-1 data, its posterior is unchanged until arm 1 is selected, and the successive draws from this frozen posterior are iid. Let G be the first attempt at which such a draw exceeds a. Then G is geometric with parameter p_a, and every one of the first $G-1$ attempts selects arm 2 on E_2. Thus at least $\min(G-1, T)$ suboptimal pulls occur, apart from the fixed initialization shift. The exact conditional expectation is

$$\mathbb{E}[\min(G-1, T) \mid p_a] = \sum_{j=1}^{T} (1-p_a)^j = \frac{(1-p_a)\{1-(1-p_a)^T\}}{p_a}. \tag{48}$$

On the small-p_a region relevant to divergence, the extra factor $1-p_a$ is bounded above and below by positive constants, so (48) has the same orders as (40). Hence

$$R_T \ge \Delta_2 \mathbb{P}(E_2)\mathbb{E}\left[\frac{(1-p_a)\{1-(1-p_a)^T\}}{p_a}\mathbf{1}\{\bar{X}_{1,2} \in [m_-, m_+], H_{1,2} \le u\}\right] - O(1).$$

Choose a fixed small $u_0 > 0$. Since $m/a \in [m_-/a, m_+/a] \Subset (0,1)$ on $E_1(u_0)$, applying (25) with $n = 2$, threshold a, and the expectation truncated to $H_2 \le u_0$ gives the same asymptotic orders and proves (26). □

E Proof of the selected-arm overshoot lemma

We couple the observations of the selected arm to an infinite iid sequence $X_1, X_2, \ldots \sim P_{\mu_0,k_0}$. Adaptive sampling changes only the calendar times at which this sequence is revealed. Put $\theta_0 = (\mu_0, k_0)$, assume $\mu_0 < b$, and let $K_0 = K_{\inf}(\theta_0, b) > 0$.

Step 1: a compact typical set. Let $\widehat{m}_n = \bar{X}_n$, $\widehat{H}_n = \log \bar{X}_n - n^{-1}\sum_i \log X_i$, and define $\widehat{k}_n$ by $h(\widehat{k}_n) = \widehat{H}_n$. The map $(m, H) \mapsto (m, h^{-1}(H))$ is continuous at $(\mu_0, h(k_0))$. Continuity of (15) and (16) then implies that, for every $\delta \in (0,1)$, there is a compact rectangle V around $(\mu_0, h(k_0))$ such that

$$\inf_{(m,H)\in V} \mathcal{K}(m, H; b) \ge (1-\delta)K_0, \tag{49}$$

where $\mathcal{K}(m, H; b) = K_{\inf}((m, h^{-1}(H)), b)$. Thus K_0 continues to denote only the scalar $K_{\inf}(\theta_0, b)$. Also $m/b \le r_+ < 1$ throughout V.

The joint moment generating function of the sufficient statistic is explicit:

$$\mathbb{E}_{\mu_0,k_0} e^{sX + t\log X} = \frac{(k_0/\mu_0)^{k_0}\Gamma(k_0+t)}{\Gamma(k_0)(k_0/\mu_0 - s)^{k_0+t}}, \qquad s < k_0/\mu_0, \quad t > -k_0. \tag{50}$$

It is finite on a neighborhood of $(0,0)$. Let $\lambda_0 = \mathbb{E}_{\theta_0} \log X$. By continuity of $(m, \lambda) \mapsto (m, \log m - \lambda)$ at (μ_0, λ_0), choose $\delta_m, \delta_\lambda > 0$ so that

$$|m - \mu_0| \le \delta_m, \quad |\lambda - \lambda_0| \le \delta_\lambda \quad \Longrightarrow \quad (m, \log m - \lambda) \in V.$$

Applying scalar Chernoff bounds to $\bar{X}_n - \mu_0$ and $n^{-1}\sum_i \log X_i - \lambda_0$, for both signs, and taking a union bound gives constants $C_1, c_1 > 0$ such that

$$\mathbb{P}_{\theta_0}\{(\widehat{m}_n, \widehat{H}_n) \notin V\} \le C_1 e^{-c_1 n}. \tag{51}$$

Step 2: a posterior bound on the typical set. Write

$$B(k) = k\log k - k - \log\Gamma(k), \qquad \phi_H(k) = B(k) - kH, \qquad \chi(a) = \log\Gamma(a) - a\log a + a.$$

Then $\phi'_H(k) = h(k) - H$ and $\phi''_H(k) = -A'(k)/k < 0$. Thus $\widehat{k} = h^{-1}(H)$ uniquely maximizes ϕ_H. If $r = m/b < 1$ and $q = h^{-1}\{H + \mathcal{I}(r)\}$, then q uniquely maximizes $\phi_{H+\mathcal{I}(r)}$ and

$$\phi_H(\widehat{k}) - \phi_{H+\mathcal{I}(r)}(q) = A(\widehat{k}) - A(q) = K_{\inf}((m, \widehat{k}), b). \tag{52}$$

Let $Q_n(b) = \Pi(\widetilde{\mu} \geq b \mid X_{1:n})$. Equations (13) and (12) give

$$Q_n(b) = \frac{\int_0^\infty w(k) R_n(k) e^{-nkH} P(nk, nkr)\, dk}{\int_0^\infty w(k) R_n(k) e^{-nkH}\, dk}, \tag{53}$$

where P is the regularized lower incomplete Gamma function. The Chernoff calculation in section C gives $P(nk, nkr) \leq e^{-nk\mathcal{I}(r)}$.

We bound the two integrals uniformly for $(m, H) \in V$. For $v > 0$, set

$$I_n(v) = \int_0^\infty w(k) R_n(k) e^{-nkv} dk, \qquad k_v = h^{-1}(v).$$

The values H and $H + \mathcal{I}(r)$ generated by V lie in one compact subinterval of $(0, \infty)$, and all corresponding k_v lie in a compact interval $J \Subset (0, \infty)$. Robbins' bounds give, uniformly on J,

$$R_n(k) = \sqrt{\frac{2\pi}{n}}\, k^{-1/2} \exp\{nB(k) + \rho_n(k)\}, \qquad |\rho_n(k)| \leq C/n. \tag{54}$$

Taylor expansion on $|k - k_v| \leq cn^{-1/2}$, where the prior is bounded below and $-\phi''_v$ has positive uniform bounds, yields

$$I_n(v) \geq c_2 n^{-1} e^{n\phi_v(k_v)}. \tag{55}$$

For the matching upper bound, choose $0 < k_{\mathrm{L}} < \inf J/2$ and $k_{\mathrm{U}} > 2 \sup J$. On $[k_{\mathrm{L}}, k_{\mathrm{U}}]$, (54), boundedness of the prior, and strict concavity give

$$\int_{k_{\mathrm{L}}}^{k_{\mathrm{U}}} w(k) R_n(k) e^{-nkv} dk \leq C n^\nu e^{n\phi_v(k_v)}.$$

The lower boundary is split at $1/n$. On $(0, 1/n)$, $\Gamma(nk) \leq C/(nk)$, $\Gamma(k)^{-1} \leq Ck$, and $w(k) \leq Ck^{-\rho}$ yield

$$\int_0^{1/n} w(k) R_n(k) dk \leq \frac{C^n}{n} \int_0^{1/n} k^{n-1-\rho} dk \leq C^n n^{-n+\rho-2}. \tag{56}$$

On $[1/n, k_{\mathrm{L}}]$, the exact decomposition $\log R_n(k) - nkv = n\phi_v(k) + \chi(nk)$, the bound $|\chi(nk)| \leq C(1 + \log n)$, and the fixed rate gap below the maximizer give $Cn^\nu e^{n\phi_v(k_v) - cn}$. Finally, $B(k) \leq C_B + \frac{1}{2}\log(1+k)$, $\chi(nk) \leq C$, and $v \geq v_- > 0$ give, for some finite polynomial-envelope exponent M of the prior (as justified in the coupling proof above),

$$w(k) R_n(k) e^{-nkv} \leq C(1+k)^M e^{n\{C_B + \frac{1}{2}\log(1+k) - v_- k\}}, \qquad k \geq k_{\mathrm{U}}.$$

Choose k_{U} so that $C_B + \frac{1}{2}\log(1+k) - v_- k \leq \inf_v \phi_v(k_v) - 2c - v_- k/2$ for $k \geq k_{\mathrm{U}}$, where the infimum is over the compact set of v generated by V. Splitting off $e^{-nv_- k/2}$ makes the remaining tail integral finite and gives $Cn^\nu e^{n\phi_v(k_v) - cn}$. Hence

$$I_n(v) \leq C n^\nu e^{n\phi_v(k_v)} \tag{57}$$

uniformly over all v generated by V.

The Chernoff bound in (53) now gives

$$Q_n(b) \leq \frac{I_n\{H + \mathcal{I}(r)\}}{I_n(H)}.$$

Combining (55) and (57), and then using (52), gives constants $C_3, \nu < \infty$ such that

$$Q_n(b) \mathbf{1}\{(\widehat{m}_n, \widehat{H}_n) \in V\} \leq C_3 n^\nu \exp[-n(1-\delta)K_0]. \tag{58}$$

Step 3: pull-indexed summation. Choose δ so that $(1+\epsilon)(1-\delta) > 1$, and let L_T be as in lemma 7. On the typical set, the action event is contained in the posterior-overshoot event, hence

$$\mathbb{E}\sum_{t=1}^{T}\mathbf{1}\{A_t = a, N_a(t-1) \geq L_T, \widetilde{\mu}_{a,t} \geq b, (\widehat{m}_{N_a(t-1)}, \widehat{H}_{N_a(t-1)}) \in V\}$$
$$\leq T \sup_{n \geq L_T} C_3 n^{\nu} e^{-n(1-\delta)K_0} \leq C(\log T)^{\nu} T^{1-(1+\epsilon)(1-\delta)} = o(1). \tag{59}$$

For all sufficiently large T, the function $x \mapsto x^{\nu}e^{-x(1-\delta)K_0}$ decreases on $[L_T, \infty)$; hence the supremum in (59) is attained at L_T up to the ceiling, which gives the displayed power of $\log T$ and T. When $A_t = a$, the count N_a increases. Consequently a fixed sample index n can occur in the atypical selected-arm sum at most once, and (51) gives

$$\mathbb{E}\sum_{t=1}^{T}\mathbf{1}\{A_t = a, N_a(t-1) \geq L_T, (\widehat{m}_{N_a(t-1)}, \widehat{H}_{N_a(t-1)}) \notin V\} \leq \sum_{n \geq L_T} C_1 e^{-c_1 n} = o(1). \tag{60}$$

Equations (59) and (60) prove (27); the indicator $A_t = a$ is what charges each term to a distinct sample count.

F SHARP-REGRET REDUCTION

Fix a suboptimal arm a and a threshold $b \in (\mu_a, \mu^\star)$ for which (28) holds. Let τ_n be the calendar time of the nth pull of the optimal arm, and let $\mathcal{T}_n = \{t : \tau_n < t \leq \tau_{n+1}\}$ be the corresponding stopped interval, truncated at T when necessary. Equivalently,

$$t \in \mathcal{T}_n \quad \Longleftrightarrow \quad N_{a^\star}(t-1) = n.$$

Let $\mathcal{F}_t$ denote the reward, action, and posterior-randomization history through round t. The description above depends only on $\mathcal{F}_{t-1}$. Throughout $\mathcal{T}_n$ the optimal-arm data and $p_n = p_n(b)$ are fixed. Immediately before the posterior draws at round t, let U_t be the optimal posterior mean draw and let M_t be the maximum of the suboptimal posterior mean draws. Conditional on $\mathcal{F}_{t-1}$, U_t and M_t are independent and $\mathbb{P}(U_t \geq b \mid \mathcal{F}_{t-1}) = p_n$.

Let L_t be the event that a suboptimal arm is selected and its selected draw is below b. Necessarily $M_t < b$ and $U_t < M_t$, so

$$\mathbb{P}(L_t \mid \mathcal{F}_{t-1}) \leq (1-p_n)\mathbb{P}(M_t < b \mid \mathcal{F}_{t-1}).$$

Let $S_t = \{U_t \geq b, M_t < b\}$. On S_t the optimal arm is selected, and

$$\mathbb{P}(S_t \mid \mathcal{F}_{t-1}) = p_n\mathbb{P}(M_t < b \mid \mathcal{F}_{t-1}).$$

If the common factor vanishes, the probability of L_t is also zero; otherwise

$$\mathbb{P}(L_t \mid \mathcal{F}_{t-1}) \leq \frac{1-p_n}{p_n}\mathbb{P}(S_t \mid \mathcal{F}_{t-1}).$$

Let $J_{n,t} = \mathbf{1}\{N_{a^\star}(t-1) = n\}$, which is $\mathcal{F}_{t-1}$-measurable. Multiply the last inequality by $J_{n,t}$, take expectations, and sum over t. More explicitly, the tower property gives

$$\mathbb{E}\sum_t J_{n,t}\mathbf{1}\{L_t\} = \sum_t \mathbb{E}\{J_{n,t}\mathbb{P}(L_t \mid \mathcal{F}_{t-1})\},$$

and the same identity holds with S_t. Tonelli's theorem justifies the exchange because all summands are nonnegative. An S_t event selects the optimal arm and ends $\mathcal{T}_n$, so $\sum_t J_{n,t}\mathbf{1}\{S_t\} \leq 1$ pathwise. Conditioning on the fixed optimal-arm observations therefore gives

$$\mathbb{E}[\#\text{ low-draw suboptimal selections in interval } n \mid X^\star_{1:n}] \leq \frac{1-p_n}{p_n}. \tag{61}$$

Here $X^\star_{1:n}$ denotes the first n observed rewards from the optimal arm. This ratio argument remains valid when $M \geq U \geq b$ preempts an optimistic optimal draw.

For high draws of a fixed suboptimal arm, its first $L_T = \lceil (1+\epsilon)\log T / K_{\inf}(\theta_a, b)\rceil$ pulls contribute at most L_T, and lemma 7 makes the remainder $o(1)$. Summing (61) over n therefore yields

$$\mathbb{E}N_a(T) \leq \frac{(1+\epsilon)\log T}{K_{\inf}(\theta_a, b)} + o(1) + \sum_{n=2}^{\infty}\mathbb{E}\frac{1-p_n(b)}{p_n(b)}. \tag{62}$$

Under (28), divide by $\log T$ and take $T \to \infty$, then $\epsilon \downarrow 0$ and $b \uparrow \mu^\star$ through thresholds for which (28) holds. Continuity of (16) proves proposition 9.

G UNIFORM BOUNDARY SADDLEPOINT CONTROL AND THE UNCAPPED THEOREM

G.1 EXACT CARRIER FACTORIZATION

Under $P_{\mu,k}$, the proportions $Z_i = X_i/S$ have density $\Gamma(nk)\Gamma(k)^{-n}\prod_i z_i^{k-1}$ on the simplex. On the level set $H_n(z) = u$, $\prod_i z_i = n^{-n}e^{-nu}$. The coarea formula therefore gives

$$f_{H_n|k}(u) = a_n(u)R_n(k)e^{-nku}, \qquad a_n(u) = n^n e^{nu}\int_{H_n(z)=u}\frac{d\sigma(z)}{\|\nabla_T H_n(z)\|}, \tag{63}$$

where the tangential gradient and surface measure are taken in any fixed orthonormal simplex chart. The chart Jacobian is absorbed into a_n, which is independent of k.

Differentiating the log normalizer gives

$$\frac{1}{n}\frac{d}{dk}\log R_n(k) = h_n(k), \qquad \frac{1}{n}\frac{d^2}{dk^2}\log R_n(k) = h_n'(k).$$

The trigamma multiplication formula implies

$$n\psi_1(nk) = \frac{1}{n}\sum_{j=0}^{n-1}\psi_1(k+j/n) < \psi_1(k),$$

because ψ_1 is strictly decreasing. Thus $h_n' < 0$. Moreover $h_n(k) \to \infty$ at zero and $h_n(k) \to 0$ at infinity, so (29) has a unique solution. Condition BSP is exactly a two-sided local-limit estimate for (63) at the mean $u = h_n(\kappa)$ of this one-dimensional exponential family.

G.2 GLOBAL SHAPE-INTEGRAL BOUNDS

The posterior-odds calculation uses the following deterministic Laplace bounds. They retain the high-shape prior factor throughout the comparison.

Lemma 14 (Uniform shape-integral comparison)**.** *For $v > 0$, define*

$$\ell_{n,v}(k) = \log R_n(k) - nkv, \qquad Z_{n,\eta}(v) = \int_0^\infty w_\eta(k)e^{\ell_{n,v}(k)}dk,$$

and let $\kappa = \kappa_n(v)$. For every fixed $\eta > 0$, constants c_η, C_η and n_η exist such that, for all $n \geq n_\eta$ and $v > 0$,

$$c_\eta n^{-1/2}(1+\kappa)^{-\eta}e^{\ell_{n,v}(\kappa)} \leq Z_{n,\eta}(v) \leq C_\eta n^{-1/2}(1+\kappa)^{-\eta}e^{\ell_{n,v}(\kappa)}. \tag{64}$$

If $0 < r \leq 1$, $u > 0$, and $q_n = \kappa_n\{u + \mathcal{I}(r)\}$, then

$$\int_0^\infty w_\eta(q)e^{\ell_{n,u}(q)}P(nq, nqr)dq \geq \frac{c_\eta n^{-1/2}(1+q_n)^{-\eta}e^{\ell_{n,u+\mathcal{I}(r)}(q_n)}}{1+(1-r)\sqrt{nq_n}}. \tag{65}$$

Consequently, with $\kappa = \kappa_n(u)$,

$$\frac{1-p_n(b)}{p_n(b)} \leq C_\eta\{1+(1-r)\sqrt{nq_n}\}\left(\frac{1+q_n}{1+\kappa}\right)^\eta e^{\ell_{n,u}(\kappa)-\ell_{n,u+\mathcal{I}(r)}(q_n)} \tag{66}$$

whenever $m/b = r \leq 1$.

Proof. The log-Beta representation gives

$$-nh_n'(k) = \sum_{j=1}^{n-1}\{\psi_1(k) - \psi_1(k+j/n)\}.$$

For an increment $0 < \xi \le 1$, write

$$\psi_1(k) - \psi_1(k+\xi) = \int_0^\xi 2\sum_{\ell\ge 0}(k+s+\ell)^{-3}ds.$$

If $k \le 1$, retaining $\ell = 0$ for the at least $(n-1)/2$ indices with $j/n \ge 1/2$ gives a lower bound cn/k^2. If $k \ge 1$, comparison of the sum with its integral gives a lower bound $\sum_j (j/n)/\{k(k+j/n)\} \ge cn/k^2$. For the upper bound, use $2\sum_{\ell\ge 0}(x+\ell)^{-3} \le 2x^{-3} + x^{-2}$ when $k \le 1$, and the same integral comparison followed by integration over s when $k \ge 1$. Dividing by n yields

$$\frac{c_0}{k^2} \le -h_n'(k) \le \frac{C_0}{k^2}, \qquad n \ge 2, k > 0, \tag{67}$$

with universal $0 < c_0 < C_0 < \infty$. The standard integral bounds $\log k - k^{-1} < \psi(k) < \log k - (2k)^{-1}$ also give $1/(2k) < h(k) < 1/k$. Integrating $\ell_{n,v}'(k) = n\{h_n(k) - v\}$ from its zero at κ and using (67) yields, for $y = k/\kappa$,

$$\ell_{n,v}(k) - \ell_{n,v}(\kappa) \le -c_0 n\{y - 1 - \log y\}. \tag{68}$$

For $y > 1$, this follows by integrating $h_n(s) - h_n(\kappa) \le -c_0(\kappa^{-1} - s^{-1})$; the case $y < 1$ uses the reversed limits. The upper curvature bound in (67) gives the reverse estimate

$$-C_0 n\{y - 1 - \log y\} \le \ell_{n,v}(\kappa y) - \ell_{n,v}(\kappa) \le -c_0 n\{y - 1 - \log y\}, \qquad y > 0. \tag{69}$$

The boundary expansions of A and A' and continuity on compact intervals give global constants

$$\frac{c_\eta}{k(1+k)^\eta} \le w_\eta(k) \le \frac{C_\eta}{k(1+k)^\eta}, \qquad k > 0. \tag{70}$$

Use $k = \kappa y$ in the shape integral. Equations (68)–(70) reduce its upper bound, after division by $(1+\kappa)^{-\eta} e^{\ell_{n,v}(\kappa)}$, to

$$C_\eta \int_0^\infty y^{-1}\left(\frac{1+\kappa}{1+\kappa y}\right)^\eta e^{-c_0 n\{y-1-\log y\}}dy.$$

Split the integral into $(0, 1/2)$, $[1/2, 3/2]$, and $(3/2, \infty)$. On the first interval,

$$y - 1 - \log y \ge c_s|\log y|, \qquad c_s := \frac{1}{4}, \qquad \frac{1+\kappa}{1+\kappa y} \le \frac{1}{y}.$$

The ratio inequality holds for every $\kappa > 0$ and $0 < y \le 1$, since $y(1+\kappa) \le 1 + \kappa y$. Put $c_1 = c_0 c_s$ and choose

$$n_\eta = \max\left\{2, \left\lceil \frac{2(\eta+1)}{c_1} \right\rceil\right\}.$$

For $n \ge n_\eta$, the contribution of the first interval is at most

$$C_\eta \int_0^{1/2} y^{c_1 n - \eta - 1}dy = C_\eta \frac{2^{-(c_1 n - \eta)}}{c_1 n - \eta} \le \frac{2C_\eta}{c_1 n},$$

uniformly in $\kappa > 0$. On $[1/2, 3/2]$, Taylor's formula gives $y - 1 - \log y \ge 2(y-1)^2/9$, while the prefactor $y^{-1}\{(1+\kappa)/(1+\kappa y)\}^\eta$ is at most $2^{\eta+1}$. Its integral is therefore bounded by $C_\eta 2^{\eta+1}\sqrt{9\pi/(2c_0 n)}$. For $y \ge 3/2$, the function $(y - 1 - \log y)/y$ is increasing, so

$$y - 1 - \log y \ge c_r y, \qquad c_r := \frac{1/2 - \log(3/2)}{3/2} > 0.$$

The prior ratio in this region is at most one, with $y^{-1} \le 2/3$, so the integral is bounded by $2C_\eta e^{-3c_0 c_r n/2}/(3c_0 c_r n)$. Combining the three intervals proves the upper bound in (64), with constants depending only on η. For the lower bound, restrict to $W_n = \{y : |y-1| \le (4\sqrt{n})^{-1}\}$. On W_n, (69) bounds the exponential ratio below by a positive constant, while

$$y^{-1}\left(\frac{1+\kappa}{1+\kappa y}\right)^\eta$$

is bounded below uniformly in κ. The length of W_n is $1/(2\sqrt{n})$, which supplies the matching lower bound in (64).

For (65), first combine the lower incomplete-Gamma estimate with the likelihood factor:

$$P(a, ar) \geq \frac{ce^{-a\mathcal{I}(r)}}{1 + (1-r)\sqrt{a}}, \qquad a > 0, \quad 0 < r \leq 1.$$

For $a < 1$, this follows from (96) because $r^a = e^{-a\mathcal{I}(r)}e^{a(r-1)}$ and $a^a/\Gamma(a+1)$ is bounded above and below on $(0, 1]$. With $a = nq$, the optimism integrand is therefore bounded below by

$$\frac{c\, w_\eta(q) e^{\ell_{n,u+\mathcal{I}(r)}(q)}}{1 + (1-r)\sqrt{nq}}.$$

Restrict $q = q_n y$ to $|y - 1| \leq (4\sqrt{n})^{-1}$. On this window, $y \in [3/4, 5/4]$. Equation (69) therefore bounds the exponential below by a universal positive multiple of its value at q_n. Equation (70) gives

$$w_\eta(q_n y) \geq c_\eta q_n^{-1}(1 + q_n)^{-\eta},$$

and $\sqrt{y} \leq \sqrt{5/4}$ gives

$$1 + (1-r)\sqrt{nq_n y} \leq \sqrt{5/4}\{1 + (1-r)\sqrt{nq_n}\}.$$

The Jacobian $dq = q_n\, dy$ cancels q_n^{-1}, and the window length gives $n^{-1/2}$. This proves (65). The pessimism numerator is at most $Z_{n,\eta}(u)$, so division gives (66). □

G.3 Uniform triangular-array control

The BSP constant is assembled from two disjoint regimes. For all sufficiently large n, the Fourier argument below is uniform over the entire half-line $\kappa \geq 1$. For the finitely many remaining n, the proof treats $\kappa \to \infty$ by a fixed-n Gamma-convolution limit and uses continuity on the compact part of $[1, \infty)$. Taking the largest upper bound and the smallest positive lower bound over this finite collection produces one numerical constant independent of n, u, and κ. The range $0 < \kappa < 1$ is excluded from BSP and is treated separately after the proof.

Put

$$\sigma_{n,\kappa}^2 = -\frac{h_n'(\kappa)}{n}, \qquad Z_{n,\kappa} = \frac{H_n - h_n(\kappa)}{\sigma_{n,\kappa}}.$$

Since a change of variables multiplies a density by its scale, evaluating (63) at $u = h_n(\kappa)$ gives the exact identity

$$f_{Z_{n,\kappa}}(0) = \frac{a_n(u) R_n(\kappa) e^{-n\kappa u}}{\sqrt{n/\{-h_n'(\kappa)\}}}. \tag{71}$$

We prove theorem 8 through an exact independent-sum representation.

Lemma 15 (Independent log-Beta representation)**.** *Let $B_{j,n,\kappa}$, $1 \leq j \leq n-1$, be independent with $B_{j,n,\kappa} \sim \mathrm{Beta}(\kappa, j/n)$. Here $\mathrm{Beta}(a, b)$ denotes the beta distribution with positive shape parameters a, b, and $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ denotes its normalizing beta function. Then*

$$nH_n \overset{d}{=} \sum_{j=1}^{n-1} -\log B_{j,n,\kappa}. \tag{72}$$

Equivalently, the characteristic function of H_n is

$$\begin{aligned}\varphi_{n,\kappa}(t) &= n^{-it}\frac{\Gamma(n\kappa)}{\Gamma(n\kappa - it)}\left\{\frac{\Gamma(\kappa - it/n)}{\Gamma(\kappa)}\right\}^n \\ &= \prod_{j=1}^{n-1}\frac{\Gamma(\kappa - it/n)\Gamma(\kappa + j/n)}{\Gamma(\kappa)\Gamma(\kappa + j/n - it/n)}.\end{aligned} \tag{73}$$

Proof. The Dirichlet moment formula applied to $H_n = -\log n - n^{-1}\sum_i \log Z_i$ gives the first line of (73). Gauss' multiplication formula

$$\Gamma(nz) = (2\pi)^{(1-n)/2} n^{nz-1/2} \prod_{j=0}^{n-1} \Gamma(z + j/n)$$

gives the product in the second line. Each factor is $\mathbb{E}B_{j,n,\kappa}^{-it/n}$, which proves (72). □

Proof of theorem 8. Write $\alpha_j = j/n$, $Y_j = -\log B_{j,n,\kappa}$, $B_n = \sum_{j=1}^{n-1} \alpha_j = (n-1)/2$, and $V_{n,\kappa} = \sum_j \mathrm{Var}(Y_j)$. For $B \sim \mathrm{Beta}(\kappa, \alpha)$ and $Y = -\log B$, the cumulants satisfy

$$v_\alpha := \mathrm{Var}(Y) = \psi_1(\kappa) - \psi_1(\kappa + \alpha), \tag{74}$$

$$\mathrm{cum}_r(Y) = (r-1)! \sum_{\ell=0}^{\infty} \{(\kappa+\ell)^{-r} - (\kappa+\alpha+\ell)^{-r}\}, \qquad r \geq 3. \tag{75}$$

Here $\mathrm{cum}_r(Y)$ is the rth cumulant of Y, and $\psi_2 = \psi_1'$ is the order-two polygamma function. The polygamma series, $0 < \alpha \leq 1$, and $\kappa \geq 1$ give

$$\frac{\alpha}{4\kappa^2} \leq v_\alpha \leq \frac{3\alpha}{\kappa^2}, \qquad 0 \leq \mathrm{cum}_r(Y) \leq 2r!\, \alpha \kappa^{-r}. \tag{76}$$

Indeed, $v_\alpha = \int_0^\alpha -\psi_2(\kappa + s)ds$ and $x^{-2} \leq -\psi_2(x) \leq 3x^{-2}$ for $x \geq 1$; the second inequality follows by integrating the difference of the two powers in (75). Summing (76) yields

$$\frac{B_n}{4\kappa^2} \leq V_{n,\kappa} \leq \frac{3B_n}{\kappa^2}, \qquad |\lambda_r| \leq 2^{r+1} r! B_n^{1-r/2}, \tag{77}$$

where λ_r is the rth cumulant of $Z_{n,\kappa} = \{\sum_j Y_j - \mathbb{E}\sum_j Y_j\}/\sqrt{V_{n,\kappa}}$. The cumulant series is analytic for $|t| < \kappa\sqrt{V_{n,\kappa}}$, and on this disk

$$\log \mathbb{E} e^{itZ_{n,\kappa}} + \frac{t^2}{2} = \sum_{r=3}^{\infty} \frac{(it)^r \lambda_r}{r!}. \tag{78}$$

Since $\kappa\sqrt{V_{n,\kappa}} \geq \sqrt{B_n}/2$, for fixed M and n sufficiently large, (77) gives

$$\begin{aligned} \sup_{\kappa \geq 1,\, |t| \leq M} \left| \sum_{r=3}^{\infty} \frac{(it)^r \lambda_r}{r!} \right| &\leq \frac{C|t|^3}{\sqrt{B_n}} \sum_{j=0}^{\infty} \left(\frac{2|t|}{\sqrt{B_n}} \right)^j \\ &\leq C_M B_n^{-1/2}. \end{aligned}$$

Consequently,

$$\sup_{\kappa \geq 1,\, |t| \leq M} \left| \log \mathbb{E} e^{itZ_{n,\kappa}} + \frac{t^2}{2} \right| = O_M(B_n^{-1/2}). \tag{79}$$

It remains to control Fourier tails uniformly. The gamma product identity

$$\frac{|\Gamma(x+iq)|^2}{\Gamma(x)^2} = \prod_{\ell=0}^{\infty} \left(1 + \frac{q^2}{(x+\ell)^2}\right)^{-1}$$

and a one-step integral comparison give

$$\begin{aligned} |\mathbb{E} e^{iqY}|^2 &= \prod_{\ell=0}^{\infty} \frac{1 + q^2/(\kappa+\alpha+\ell)^2}{1 + q^2/(\kappa+\ell)^2} \\ &\leq \left(1 + \frac{q^2}{(\kappa+1)^2}\right)^{-\alpha}. \end{aligned} \tag{80}$$

To see the inequality, write the negative logarithm of the product as the integral over $s \in [0, \alpha]$ of $\sum_{\ell \geq 0} 2q^2/[x(x^2+q^2)]$ at $x = \kappa + s + \ell$, then bound the sum below by its integral over ℓ. Applying (80) to the standardized sum and using (77) gives

$$|\mathbb{E}e^{itZ_{n,\kappa}}| \leq \left(1 + \frac{t^2}{12B_n}\right)^{-B_n/2}. \tag{81}$$

The Fourier tail can be evaluated explicitly. Put $G_n(t) = (1 + t^2/(12B_n))^{-B_n/2}$. If $M < |t| \leq \sqrt{12B_n}$, then $\log(1+x) \geq x/2$ for $0 \leq x \leq 1$ gives $G_n(t) \leq e^{-t^2/48}$. For $|t| > \sqrt{12B_n}$, set $s = |t|/\sqrt{12B_n}$ and use $1 + s^2 \geq 2s$ to obtain, when $B_n > 2$,

$$\int_{\sqrt{12B_n}}^{\infty} G_n(t)dt \leq \frac{\sqrt{12B_n}\, 2^{-B_n/2}}{B_n/2 - 1}. \tag{82}$$

The right side tends to zero, while the first tail is bounded by the Gaussian integral $\int_M^{\infty} e^{-t^2/48}dt$. Hence

$$\lim_{M\to\infty} \limsup_{n\to\infty} \sup_{\kappa \geq 1} \int_{|t|>M} |\mathbb{E}e^{itZ_{n,\kappa}}|dt = 0. \tag{83}$$

Fourier inversion, (79), and (83) therefore yield

$$\sup_{\kappa \geq 1} \left| f_{Z_{n,\kappa}}(0) - (2\pi)^{-1/2} \right| \longrightarrow 0. \tag{84}$$

For each of the finitely many smaller n, the convolution in (72) has a positive continuous density at its interior mean, so $f_{Z_{n,\kappa}}(0)$ is positive and continuous for $\kappa \geq 1$. To identify its limit at infinity, let $W_{\kappa,\alpha} = \kappa(-\log B)$ for $B \sim \mathrm{Beta}(\kappa, \alpha)$. Its density is

$$g_{\kappa,\alpha}(y) = \frac{e^{-y}(1 - e^{-y/\kappa})^{\alpha-1}}{\kappa B(\kappa, \alpha)}, \qquad y > 0. \tag{85}$$

The relations $\kappa^{\alpha} B(\kappa, \alpha) \to \Gamma(\alpha)$ and $\kappa(1 - e^{-y/\kappa}) \to y$ show, locally uniformly on $(0, \infty)$,

$$g_{\kappa,\alpha}(y) \longrightarrow \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)}. \tag{86}$$

For $\kappa \geq 1$, the inequalities $c\min(y/\kappa, 1) \leq 1 - e^{-y/\kappa} \leq \min(y/\kappa, 1)$ and $\kappa^{\alpha} B(\kappa, \alpha) \asymp 1$ imply

$$g_{\kappa,\alpha}(y) \leq C_{\alpha}(1 + y^{\alpha-1})e^{-y}, \qquad y > 0.$$

The right side is integrable for every fixed $\alpha > 0$. Let G_{κ} be the density of $\kappa \sum_{j=1}^{n-1} Y_j$. For $n = 2$, it is the single density $g_{\kappa,1/2}$, and (86) gives the required locally uniform convergence directly. For $n \geq 3$, it is the convolution of the $n-1$ densities $g_{\kappa,\alpha_1}, \ldots, g_{\kappa,\alpha_{n-1}}$ and has $n-2$ free integration coordinates:

$$G_{\kappa}(x) = \int_{\substack{y_1,\ldots,y_{n-2}>0 \\ \sum_{j=1}^{n-2} y_j < x}} \left\{ \prod_{j=1}^{n-2} g_{\kappa,\alpha_j}(y_j) \right\} g_{\kappa,\alpha_{n-1}}\left(x - \sum_{j=1}^{n-2} y_j \right) dy_1 \cdots dy_{n-2}.$$

Fix $0 < a < b < \infty$ and let $x \in [a, b]$. Substituting $y_j = xz_j$ puts the integral on the fixed simplex $\{z_1, \ldots, z_{n-2} > 0 : \sum_{j=1}^{n-2} z_j < 1\}$, with $z_{n-1} = 1 - \sum_{j=1}^{n-2} z_j$. The Jacobian is x^{n-2}, and the product envelope is bounded, uniformly in $x \in [a, b]$ and $\kappa \geq 1$, by

$$C_{a,b,n} \prod_{j=1}^{n-1} z_j^{\alpha_j - 1}.$$

This Dirichlet density kernel is integrable because every $\alpha_j > 0$. On a compact subset away from its faces, the density convergence is uniform in x; the envelope makes the remaining integral uniformly small. Thus the convolution converges uniformly on $[a, b]$. For every fixed $n \geq 2$, this proves

$$G_{\kappa}(x) \longrightarrow f_{\mathrm{Gamma}(B_n,1)}(x) \quad \text{locally uniformly for } x > 0, \tag{87}$$

where the limiting shape is $\sum_j \alpha_j = B_n$. Direct differentiation of the Beta moments also gives $\kappa\mathbb{E}\sum_j Y_j \to B_n$ and $\kappa^2 V_{n,\kappa} \to B_n$. Since

$$f_{Z_{n,\kappa}}(0) = \sqrt{V_{n,\kappa}}\, f_{\sum_j Y_j}\left(\mathbb{E}\sum_j Y_j\right) = \kappa\sqrt{V_{n,\kappa}}\, G_\kappa\left(\kappa\mathbb{E}\sum_j Y_j\right),$$

(87) yields $f_{Z_{n,\kappa}}(0) \to \sqrt{B_n} f_{\mathrm{Gamma}(B_n,1)}(B_n) > 0$ as $\kappa \to \infty$. For each fixed n, choose K_n so that the positive limit provides common two-sided bounds on $\kappa \geq K_n$; continuity supplies such bounds on $[1, K_n]$. There are only finitely many such n, so taking their finite minimum and maximum, then combining them with (84) and (71), proves (30) with one universal constant. □

Lemma 16 (Small-shape carrier estimate)**.** *There is a universal $C < \infty$ such that, for $n \geq 2$ and $0 < \kappa \leq 1$,*

$$a_n\{h_n(\kappa)\}R_n(\kappa)e^{-n\kappa h_n(\kappa)} \leq C\sqrt{\frac{n}{-h_n'(\kappa)}}. \tag{88}$$

Proof. By (71), it suffices to bound $f_{Z_{n,\kappa}}(0)$. The log-Beta representation provides a direct Fourier bound. For $0 < \kappa \leq 1$ and $0 < \alpha \leq 1$,

$$v_\alpha = 2\int_0^\alpha \sum_{\ell\geq 0}(\kappa + s + \ell)^{-3} ds, \qquad \frac{\alpha}{4\kappa^2} \leq v_\alpha \leq \frac{3}{\kappa^2}, \tag{89}$$

$$\mathrm{cum}_r(Y) \leq \frac{r!}{2}\kappa^{2-r} v_\alpha, \qquad r \geq 3. \tag{90}$$

For the lower variance bound, if $\alpha \leq \kappa$, compare the sum in (89) with its integral; if $\alpha > \kappa$, retain the $\ell = 0$ term. The upper bound follows from $v_\alpha \leq \psi_1(\kappa) \leq 3\kappa^{-2}$. Equation (90) follows by comparing each $(\kappa + s + \ell)^{-r-1}$ with $\kappa^{2-r}(\kappa + s + \ell)^{-3}$. Hence, for $V = \sum_j v_{\alpha_j}$,

$$c\sqrt{n} \leq \kappa\sqrt{V} \leq C\sqrt{n}, \qquad \left|\mathrm{cum}_r \frac{\sum_j (Y_j - \mathbb{E}Y_j)}{\sqrt{V}}\right| \leq \frac{r!}{2}(\kappa\sqrt{V})^{2-r}. \tag{91}$$

For $|t| \leq c_0\sqrt{n}$, insert (91) in the cumulant series. Choosing c_0 so that the sum of the terms of order at least three is at most $t^2/4$ gives

$$|\mathbb{E}e^{itZ_{n,\kappa}}| \leq e^{-t^2/4}.$$

For the remaining frequencies, take $J = \{j : j/n \geq 1/2\}$ and put $s = |t|/(\kappa\sqrt{V})$. The $\ell = 0$ factor in the Gamma product gives

$$|\mathbb{E}e^{iqY_j}|^2 \leq \frac{1 + q^2/(\kappa + \alpha_j)^2}{1 + q^2/\kappa^2}, \qquad j \in J. \tag{92}$$

The set J contains at least $(n-1)/2$ indices. For all n above a fixed numerical threshold it can be partitioned into J_1, J_2 with $|J_1|, |J_2| \geq c_J n$; the finitely many smaller n are handled below. Choose a universal $\kappa_0 \in (0,1)$ small enough that the constant in the last-frequency estimate below satisfies $C\kappa_0 < 1/2$. When $\kappa \leq \kappa_0$, (92) is bounded by a constant $\varrho < 1$ for $c_0 \leq s \leq 1$, and by $2/(1+s^2)$ for $1 \leq s \leq (2\kappa)^{-1}$. For $s > (2\kappa)^{-1}$, use (92) on J_1 and (80) on J_2. This gives constants $c, C > 0$ such that

$$|\mathbb{E}e^{itZ_{n,\kappa}}| \leq \begin{cases} \varrho^{cn}, & c_0 \leq s \leq 1, \\ \{2/(1+s^2)\}^{cn}, & 1 < s \leq (2\kappa)^{-1}, \\ (C\kappa)^{cn}(1 + c\kappa^2 s^2)^{-cn}, & s > (2\kappa)^{-1}. \end{cases} \tag{93}$$

Here the three lines follow directly after substituting $q = t/\sqrt{V} = \kappa s$. On $c_0 \leq s \leq 1$, the numerator in (92) is at most $1 + C\kappa^2 s^2$, while its denominator is $1 + s^2$; choosing c_0 fixed gives $\varrho < 1$. On the second interval, $q \leq 1/2$ and $\kappa + \alpha_j \geq 1/2$, so the same ratio is at most $2/(1+s^2)$. On the last interval it is at most $C\kappa^2$ for the first subset. For the second subset, (80) and $\alpha_j \geq 1/2$ give $(1 + c\kappa^2 s^2)^{-c}$ per factor. Taking square roots and multiplying over two subsets of cardinality proportional to n gives (93). Because $dt = \kappa\sqrt{V}\, ds \leq C\sqrt{n}\, ds$, the three right-hand sides have a uniformly bounded integral for all sufficiently large n. On $c_0 \leq s \leq 1$ the integral is at most

$C\sqrt{n}\varrho^{cn}$. On $1 < s \le (2\kappa)^{-1}$, use $2/(1+s^2) \le e^{-c(s-1)}$ for $1 < s \le 2$ and its polynomial decay for $s > 2$; the result is $O(n^{-1/2})$ after multiplication by $\sqrt{n}$. On the last range, the substitution $z = \kappa s$ gives

$$C\sqrt{n}\,(C\kappa)^{cn}\kappa^{-1}\int_{1/2}^{\infty}(1+cz^2)^{-cn}dz.$$

For n large enough that $cn > 1$, the prefactor κ^{cn-1} is bounded on $(0, \kappa_0]$, and the integral is $O(n^{-1})$ by its endpoint Laplace bound. Choosing κ_0 so that $C\kappa_0 < 1/2$ makes the complete display uniformly bounded in n and $0 < \kappa \le \kappa_0$. For $\kappa_0 \le \kappa \le 1$, the proof of (81) applies with altered universal constants. For each of the finitely many remaining n, set $W_{\kappa,\alpha} = \kappa(-\log B)$ for $B \sim \mathrm{Beta}(\kappa, \alpha)$. Its density is

$$g_{\kappa,\alpha}(y) = \frac{e^{-y}(1-e^{-y/\kappa})^{\alpha-1}}{\kappa B(\kappa,\alpha)}, \qquad y > 0.$$

As $\kappa \downarrow 0$, $\kappa B(\kappa,\alpha) \to 1$ and $g_{\kappa,\alpha}(y) \to e^{-y}$ locally uniformly for $y > 0$. The behavior at zero can be controlled explicitly. Uniformly for $0 < \kappa \le \kappa_0$,

$$g_{\kappa,\alpha}(y) \le \begin{cases} C_\alpha \kappa^{1-\alpha} y^{\alpha-1}, & 0 < y \le \kappa, \\ C_\alpha e^{-y}, & y > \kappa. \end{cases}$$

The first line follows from $1 - e^{-y/\kappa} \ge c\,y/\kappa$ for $y \le \kappa$; the inequality reverses after raising to $\alpha - 1 \le 0$. Since $\kappa^{1-\alpha} \le 1$, both lines are dominated by the integrable function $C_\alpha(1 + y^{\alpha-1})e^{-y/2}$. The convolution argument used after (86) now applies with the limit density e^{-y}. Thus the density G_κ of $\kappa\sum_{j=1}^{n-1} Y_j$ converges locally uniformly on $(0,\infty)$ to the $\mathrm{Gamma}(n-1,1)$ density.

The Beta moment formulas give, for each factor,

$$\kappa \mathbb{E}Y_j \to 1, \qquad \kappa^2 \operatorname{Var}(Y_j) \to 1.$$

Consequently

$$f_{Z_{n,\kappa}}(0) = \kappa\sqrt{V}\,G_\kappa\left(\kappa\mathbb{E}\sum_j Y_j\right) \longrightarrow \sqrt{n-1}\,f_{\mathrm{Gamma}(n-1,1)}(n-1) < \infty.$$

Continuity in κ on compact subintervals of $(0,1]$ supplies a uniform bound for each of the finitely many n not covered by the preceding Fourier estimates. For $n = 2$, the same limit also follows directly from (94). Taking the maximum of these finitely many bounds and the uniform Fourier bound gives $f_{Z_{n,\kappa}}(0) \le C$, proving (88). □

This lemma is deliberately one-sided: theorem 8 supplies a two-sided standardized-center estimate only for $\kappa \ge 1$, whereas lemma 16 supplies exactly the upper carrier control needed for $0 < \kappa \le 1$. No lower small-shape BSP bound is used in theorem 10.

For $n = 2$, both fixed-n boundary limits are visible directly from $X_1/(X_1+X_2) \sim \mathrm{Beta}(\kappa,\kappa)$:

$$f_{H_2|\kappa}(u) = \frac{4^{1-\kappa}e^{-2\kappa u}}{B(\kappa,\kappa)\sqrt{1-e^{-2u}}}, \qquad u > 0, \tag{94}$$

Substitution of $u = h_2(\kappa)$ into (71) gives

$$f_{Z_{2,\kappa}}(0) \longrightarrow \begin{cases} e^{-1}, & \kappa \downarrow 0, \\ e^{-1/2}/\sqrt{\pi}, & \kappa \to \infty. \end{cases}$$

The first value agrees with the $\mathrm{Gamma}(n-1,1)$ limit in the small-shape argument, and the second with the $\mathrm{Gamma}(B_n,1)$ limit in the high-shape argument.

G.4 A UNIFORM GAMMA TAIL BOUND

Lemma 17 (Incomplete-Gamma boundary bound)**.** *There are universal $0 < c < C < \infty$ such that, for $Y \sim \mathrm{Gamma}(a, 1)$, $a \geq 1$, and $0 < r \leq 1$,*

$$c\frac{e^{-a\mathcal{I}(r)}}{1+(1-r)\sqrt{a}} \leq \mathbb{P}(Y \leq ar) \leq C\frac{e^{-a\mathcal{I}(r)}}{1+(1-r)\sqrt{a}}. \tag{95}$$

For $0 < a < 1$,

$$e^{-1}\frac{(ar)^a}{\Gamma(a+1)} \leq \mathbb{P}(Y \leq ar) \leq \frac{(ar)^a}{\Gamma(a+1)}. \tag{96}$$

Proof. Assume first that $a \geq 1$, put $x = ar$ and $\delta = (1-r)\sqrt{a}$, and let $f_a(y) = y^{a-1}e^{-y}/\Gamma(a)$. Two-sided Stirling bounds imply

$$\frac{x^a e^{-x}}{\Gamma(a+1)} = e^{-a\mathcal{I}(r)}\frac{a^a e^{-a}}{\Gamma(a+1)}, \qquad f_a(x) = \frac{1}{r}\frac{x^a e^{-x}}{\Gamma(a+1)}.$$

The factor $a^a e^{-a}/\Gamma(a+1)$ is comparable to $a^{-1/2}$ uniformly for $a \geq 1$. Hence

$$\frac{x^a e^{-x}}{\Gamma(a+1)} \asymp \frac{e^{-a\mathcal{I}(r)}}{\sqrt{a}}, \qquad f_a(x) \asymp \frac{e^{-a\mathcal{I}(r)}}{r\sqrt{a}}, \tag{97}$$

uniformly in $a \geq 1$ and $0 < r \leq 1$. For $r < 1$, the convergent series for the unregularized lower incomplete Gamma function $\gamma(a, x) := \int_0^x y^{a-1}e^{-y}dy$ is

$$\gamma(a, x) = x^a e^{-x}\sum_{j=0}^{\infty}\frac{x^j}{a(a+1)\cdots(a+j)}.$$

Keeping the $j = 0$ term and using $a(a+1)\cdots(a+j) \geq a^{j+1}$ gives

$$\frac{x^a e^{-x}}{\Gamma(a+1)} \leq \mathbb{P}(Y \leq x) \leq \frac{x^a e^{-x}}{\Gamma(a+1)(1-r)}. \tag{98}$$

Suppose first that $\delta \geq 1$. Then $x \leq a - \sqrt{a} \leq a - 1$, so f_a is increasing on $[0, x]$. Choose the terminal length

$$\ell = \frac{r}{2(1-r)}, \qquad q = \frac{\ell}{x} = \frac{1}{2a(1-r)} \leq \frac{1}{2}.$$

For $y \in [x - \ell, x]$, monotonicity of the logarithmic derivative gives

$$\begin{aligned}\log\frac{f_a(x)}{f_a(y)} &\leq \ell\left\{\frac{a-1}{x-\ell} - 1\right\} \\ &\leq \frac{1 + rq/(1-r)}{2(1-q)} \leq \frac{3}{2},\end{aligned}$$

where the last inequality uses $q/(1-r) = \{2a(1-r)^2\}^{-1} = (2\delta^2)^{-1} \leq 1/2$. Consequently, by (97),

$$\mathbb{P}(Y \leq x) \geq \int_{x-\ell}^{x} f_a(y)\,dy \geq e^{-3/2}\ell f_a(x) \gtrsim \frac{e^{-a\mathcal{I}(r)}}{\delta}.$$

The upper inequality in (98) gives the matching upper bound. Since $1 + \delta \asymp \delta$ in this regime, this proves (95) for $\delta \geq 1$.

Now let $\delta < 1$. For the upper bound, if $1/2 \leq \delta < 1$, (98) gives $\mathbb{P}(Y \leq x) \lesssim e^{-a\mathcal{I}(r)}/\delta \lesssim e^{-a\mathcal{I}(r)}$. If $\delta < 1/2$, set $u = 1 - r$. Then $u \leq 1/2$ and

$$\mathcal{I}(1-u) = \sum_{j=2}^{\infty}\frac{u^j}{j} \leq \frac{u^2}{2(1-u)} \leq u^2,$$

so $a\mathcal{I}(r) \le \delta^2 < 1/4$ and $\mathbb{P}(Y \le x) \le 1 \le e^{1/4} e^{-a\mathcal{I}(r)}$. For the lower bound, when $1 \le a < 4$, the first inequality in (98) and (97) give $\mathbb{P}(Y \le x) \gtrsim e^{-a\mathcal{I}(r)}$. When $a \ge 4$, take $\ell = \sqrt{a}/4$. The condition $\delta < 1$ gives $r \ge 1/2$ and, for $y \in [x - \ell, x]$,

$$y \ge a - \frac{5}{4}\sqrt{a} \ge \frac{3a}{8}, \qquad \left|\frac{d}{dy} \log f_a(y)\right| = \left|\frac{a-1}{y} - 1\right| \le \frac{14}{3\sqrt{a}}.$$

Hence $f_a(y) \ge e^{-7/6} f_a(x)$ throughout this interval and

$$\mathbb{P}(Y \le x) \ge e^{-7/6} \frac{\sqrt{a}}{4} f_a(x) \gtrsim e^{-a\mathcal{I}(r)}.$$

Since $1 + \delta \asymp 1$, the central regime is complete; the case $r = 1$ follows by continuity.

Finally, when $0 < a < 1$, one has $ar \le 1$ and therefore $e^{-1} \le e^{-y} \le 1$ on $[0, ar]$. Since

$$\Gamma(a)^{-1} \int_0^{ar} y^{a-1} dy = \frac{(ar)^a}{\Gamma(a+1)},$$

the same integral gives (96) with constants independent of a and r. □

G.5 Uniform saddlepoint control implies summable pessimism odds

We prove theorem 10. Fix $\eta > 0$, an optimal-arm parameter $\theta^\star = (\mu^\star, k^\star)$, and a threshold $b \in (0, \mu^\star)$. Unless a constant is explicitly called universal or indexed only by η, it may depend on $(\eta, \theta^\star, b)$ but never on (n, m, u, r, κ, q) or T. Conditional on $(\bar{X}_n, H_n) = (m, u)$, write $p_n(b \mid m, u) = \mathbb{P}(\widetilde{\mu} \ge b \mid m, u)$ and define the posterior pessimism odds $\mathcal{O}_n(m, u) := \{1 - p_n(b \mid m, u)\}/p_n(b \mid m, u)$. Also write

$$s(a) = \log \Gamma(a) - a \log a + a, \qquad \phi_v(k) = k \log k - k - \log \Gamma(k) - kv.$$

Then

$$\ell_{n,v}(k) = n\phi_v(k) + s(nk), \qquad e^{-s(a)} \le C(1 + \sqrt{a}), \quad a > 0. \tag{99}$$

The second bound follows from $s(a) = -\log a + O(1)$ on $0 < a \le 1$ and $s(a) = \frac{1}{2}\log(2\pi/a) + O(a^{-1})$ on $a \ge 1$.

The density of $m = \bar{X}_n$ under $\theta^\star$ has the exact rate form

$$f_m(m) = \frac{1}{\mu^\star} \left(\frac{m}{\mu^\star}\right)^{-1} e^{-s(nk^\star)} e^{-nk^\star \mathcal{I}(m/\mu^\star)}. \tag{100}$$

It is independent of H_n. For $m \le b$, put $r = m/b$ and set

$$\kappa = \kappa_n(u), \quad q_n = \kappa_n\{u + \mathcal{I}(r)\}, \quad \widehat{k} = h^{-1}(u), \quad q = h^{-1}\{u + \mathcal{I}(r)\}. \tag{101}$$

The elementary inequality $\psi(x) < \log x$ gives $h_n(k) = \psi(nk) - \log n - \psi(k) < h(k)$. Therefore $h_n(q) < h(q) = u + \mathcal{I}(r) = h_n(q_n)$, and monotonicity gives $q_n < q$. Likewise $u + \mathcal{I}(r) \ge u$ gives $q_n \le \kappa$. Integrating (67) from κ to infinity gives $h_n(\kappa) \ge c_0/\kappa$. Together with $h(q) \le 1/q$, this yields $q \le C_0\kappa$.

First take $n \ge n_\eta$, where n_η is the threshold in lemma 14. Apply (66), use the maximizing property $\ell_{n,u+\mathcal{I}(r)}(q_n) \ge \ell_{n,u+\mathcal{I}(r)}(q)$, and enlarge the elementary prefactors from q_n to q. This gives

$$\mathcal{O}_n(m, u) \le C_\eta(1 + \sqrt{nq}) \left(\frac{1+q}{1+\kappa}\right)^\eta e^{\ell_{n,u}(\kappa) - \ell_{n,u+\mathcal{I}(r)}(q)}. \tag{102}$$

The carrier estimate is global: theorem 8 applies for $\kappa \ge 1$ and lemma 16 applies for $0 < \kappa \le 1$. Hence

$$a_n(u) e^{\ell_{n,u}(\kappa)} \le C\sqrt{\frac{n}{-h_n'(\kappa)}}. \tag{103}$$

Multiply (102) by $f_m(m) f_{H_n|k^\star}(u)$ and use $f_{H_n|k^\star}(u) = a_n(u) e^{\ell_{n,u}(k^\star)}$. Before applying any bound, the exponential and carrier factors are exactly

$$\begin{aligned} &a_n(u) e^{\ell_{n,u}(\kappa)} e^{-s(nk^\star)} e^{\ell_{n,u}(k^\star) - \ell_{n,u+\mathcal{I}(r)}(q)} e^{-nk^\star \mathcal{I}(m/\mu^\star)} \\ &\quad = a_n(u) e^{\ell_{n,u}(\kappa)} e^{-s(nq)} \exp\left[-n\left\{k^\star \mathcal{I}(m/\mu^\star) - \phi_u(k^\star) + \phi_{u+\mathcal{I}(r)}(q)\right\}\right]. \end{aligned}$$

The equality uses $\ell_{n,v}(k) = n\phi_v(k) + s(nk)$: the factor $e^{-s(nk^\star)}$ from the mean density cancels $e^{s(nk^\star)}$ inside $e^{\ell_{n,u}(k^\star)}$, leaving $e^{-s(nq)}$ from the denominator saddle. By (99), this last term is at most $C(1+\sqrt{nq})$. Together with the prefactor $1+\sqrt{nq}$ in (102), it is bounded by $C(1+nq)$. The population rate in braces is

$$\begin{aligned} &k^\star \mathcal{I}(m/\mu^\star) - \phi_u(k^\star) + \phi_{u+\mathcal{I}(r)}(q) \\ &\qquad = D((m,\widehat{k})\|\theta^\star) - K_{\inf}((m,\widehat{k}),b). \end{aligned} \tag{104}$$

Let $c_b = b/\mu^\star$, $\delta_b = k^\star \mathcal{I}(c_b)$, and $\lambda_b = k^\star(1-c_b)$. Equation (37) rewrites the last line exactly as

$$J_{k^\star}(q) + \delta_b + \lambda_b(1-r).$$

Finally use $m = br$, $u = h_n(\kappa)$, and (67). The Jacobians satisfy

$$\frac{dm}{\mu^\star(m/\mu^\star)} = \frac{dr}{r}, \qquad \sqrt{\frac{n}{-h_n'(\kappa)}}\,|du| = \sqrt{n\{-h_n'(\kappa)\}}d\kappa \le C\sqrt{n}\frac{d\kappa}{\kappa}.$$

We obtain

$$\mathbb{E}_{\theta^\star}[\mathcal{O}_n;\, m \le b] \le C_\eta \sqrt{n}\, e^{-n\delta_b}\mathcal{B}_{n,\eta}, \tag{105}$$

where

$$\mathcal{B}_{n,\eta} := \int_0^1 \frac{e^{-n\lambda_b(1-r)}}{r} \int_0^\infty (1+nq)\left(\frac{1+q}{1+\kappa}\right)^\eta e^{-nJ_{k^\star}(q)}\frac{d\kappa}{\kappa}\,dr, \tag{106}$$

with $q = h^{-1}\{h_n(\kappa) + \mathcal{I}(r)\}$.

Lemma 18 (Boundary integral)**.** *For every fixed $(\eta, \theta^\star, b)$ with $\eta > 0$ and $0 < b < \mu^\star$, there are $C < \infty$ and $n_0 \ge 2$, independent of (n, r, κ, q), such that $\mathcal{B}_{n,\eta} \le Cn$ for all $n \ge n_0$.*

Proof. Up to shared endpoints of zero measure, the domain is partitioned as follows: first $r \in (0, r_\circ]$ versus $r \in [r_\circ, 1]$; in the second part, $\kappa \in (0, \kappa_0]$, $[\kappa_0, 1]$, or $[1, \infty)$; and on the last range, $q \le Q$ versus $q > Q$. The constants $r_\circ, \kappa_0, Q$ are fixed from $(\theta^\star, b)$ before n varies. The relation $q \le C_0\kappa$ was proved above. Choose $q_0 < k^\star$ so that $J_{k^\star}(q) \ge c_1 h(q)$ for $0 < q \le q_0$; this is possible because both functions are asymptotic to positive multiples of q^{-1} at zero. Next choose a threshold $r_\circ \in (0,1)$ such that $h^{-1}\{\mathcal{I}(r_\circ)\} \le q_0$.

Consider first $0 < r \le r_\circ$. Since $h(q) = h_n(\kappa) + \mathcal{I}(r) \ge \mathcal{I}(r)$ and h^{-1} is decreasing,

$$q \le h^{-1}\{\mathcal{I}(r)\} \le q_0, \qquad J_{k^\star}(q) \ge c_1\{h_n(\kappa) + \mathcal{I}(r)\}.$$

The r integral is bounded explicitly by the substitution $z = \mathcal{I}(r)$: because $|\mathcal{I}'(r)| = (1-r)/r \ge (1-r_\circ)/r$,

$$\int_0^{r_\circ} r^{-1} e^{-c_1 n \mathcal{I}(r)} dr \le \frac{1}{1-r_\circ}\int_{\mathcal{I}(r_\circ)}^\infty e^{-c_1 n z} dz \le \frac{C}{n}. \tag{107}$$

For the inner integral, $q \le q_0$ and $h_n(\kappa) \ge c_0/\kappa$ give

$$\int_0^\infty (1+nq)\left(\frac{1+q}{1+\kappa}\right)^\eta e^{-c_1 n h_n(\kappa)}\frac{d\kappa}{\kappa} \le Cn\int_0^\infty e^{-cn/\kappa}\frac{d\kappa}{\kappa(1+\kappa)^\eta} \le Cn.$$

For the last inequality, split at $\kappa = n$. On $(0, n]$, the substitution $x = n/\kappa$ bounds the integral without the leading factor n by $\int_1^\infty e^{-cx}dx/x$; on $[n, \infty)$ it is at most $\int_n^\infty \kappa^{-1-\eta}d\kappa$. Thus the inner integral is $O(n)$ as displayed. Equations above bound the contribution of $r \le r_\circ$ by a constant.

Now take $r_\circ \le r \le 1$, so $r^{-1} \le r_\circ^{-1}$. Choose $\kappa_0 > 0$ so small that $C_0\kappa_0 < q_0$. On $0 < \kappa \le \kappa_0$, one has $J_{k^\star}(q) \ge c/q \ge c/\kappa$, which makes the $d\kappa/\kappa$ integral finite uniformly in r and n: its prefactor is at most $C(1+n\kappa)$, and $x = n/\kappa$ gives

$$\int_0^{\kappa_0}(1+n\kappa)e^{-cn/\kappa}\frac{d\kappa}{\kappa} = \int_{n/\kappa_0}^\infty \left(1 + \frac{n^2}{x}\right)e^{-cx}\frac{dx}{x},$$

which is uniformly bounded for $n \ge n_0$. On the compact range $\kappa_0 \le \kappa \le 1$, the global comparisons for h_n and h keep q in a fixed compact subset of $(0, \infty)$; the integral over this range is therefore $O(n)$.

It remains to treat $\kappa \geq 1$. Fix $Q > \max\{1, k^\star\}$. When $q \leq Q$,

$$(1+nq)\left(\frac{1+q}{1+\kappa}\right)^{\eta} \leq C_Q n(1+\kappa)^{-\eta},$$

whose product with $d\kappa/\kappa$ is integrable. When $q > Q$, the derivative $J'_{k^\star}(q) = A'(q)(1-k^\star/q)$ and the bound $A'(q) \geq c/q$ on $[Q,\infty)$ give

$$J_{k^\star}(q) \geq J_{k^\star}(Q) + c\log(q/Q).$$

Hence, once $cn > \eta + 2$,

$$\sup_{q>Q}(1+nq)(1+q)^{\eta}e^{-nJ_{k^\star}(q)} \leq Cn.$$

Factoring out $(1+\kappa)^{-\eta}$ again leaves the integrable measure $(1+\kappa)^{-\eta}d\kappa/\kappa$. The interval $[r_\circ, 1]$ has finite length, so all its regions contribute at most Cn, which proves the lemma. □

Lemma 19 (Finite initial pessimism odds)**.** *For every fixed $n \geq 2$ and $\eta > 0$,*

$$\mathbb{E}_{\theta^\star}\frac{1-p_n(b)}{p_n(b)} < \infty.$$

Proof. For $m > b$, conditional optimism satisfies $P(nq, nq\,m/b) \geq P(nq, nq)$. The two cases in lemma 17, evaluated at $r = 1$, give a universal constant $c_P > 0$ such that $P(a,a) \geq c_P$ for every $a > 0$. Hence $p_n(b) \geq c_P$ and the odds are bounded on this region.

Now let $r = m/b \leq 1$ and write $Z(u) = Z_{n,\eta}(u)$ for the posterior normalizer. For this fixed n, (36) and (70) give constants $c_n, C_n, q_{0,n} > 0$ such that

$$c_n q^{n-2} \leq w_\eta(q)R_n(q) \leq C_n q^{n-2}, \qquad 0 < q \leq q_{0,n}.$$

Moreover, (96) gives

$$P(nq, nqr) \geq ce^{-nq\mathcal{I}(r)}, \qquad 0 < nq \leq 1.$$

Choose $c'_n > 0$ so small that $L = c'_n/\{1+u+\mathcal{I}(r)\} \leq \min(q_{0,n}, 1/n)$. On $0 < q < L$, $e^{-nqu}P(nq,nqr) \geq ce^{-nq\{u+\mathcal{I}(r)\}} \geq ce^{-nc'_n}$. Therefore

$$\int_0^L w_\eta(q)R_n(q)e^{-nqu}P(nq,nqr)dq \geq c\int_0^L q^{n-2}dq = \frac{c}{n-1}L^{n-1},$$

and division by $Z(u)$ yields

$$p_n(b) \geq \frac{c_n\{1+u+\mathcal{I}(r)\}^{-(n-1)}}{Z(u)}. \tag{108}$$

The normalizer has explicit bounds at both ends. As $u \downarrow 0$, section C.2 with $\tau = 1+\eta$ gives

$$Z(u) \lesssim_n \begin{cases} u^{-(d-\eta)}, & 0 < \eta < d, \\ 1 + |\log u|, & \eta = d, \\ 1, & \eta > d, \end{cases} \qquad d = (n-1)/2.$$

The dyadic mass bound (39) shows directly that each expression is integrable against the law of H_n because $\eta > 0$. For large u, split the normalizer at $q_{0,n}$. The lower-boundary expansion gives

$$\int_0^{q_{0,n}} w_\eta(q)R_n(q)e^{-nqu}dq \leq C_n\int_0^\infty q^{n-2}e^{-nqu}dq = C_n\Gamma(n-1)(nu)^{-(n-1)}.$$

For $q \geq q_{0,n}$ and $u \geq 1$, write $e^{-nqu} = e^{-nq_{0,n}u/2}e^{-nqu/2}$; the second factor makes the remaining integral finite. Thus $Z(u) \lesssim_n (1+u)^{-(n-1)}$, and

$$\mathbb{E}[Z(H_n)\{1+H_n\}^{n-1}] < \infty, \qquad \mathbb{E}Z(H_n) < \infty.$$

Finally, the Gamma density of m is proportional to $m^{nk^\star-1}$ near zero, while $\mathcal{I}(m/b) = O(1 + |\log m|)$ there. It has moments of every fixed power of $\mathcal{I}(m/b)$ on $(0, b]$. Independence of m and H_n, (108), and $(x+y)^{n-1} \le C_n(x^{n-1} + y^{n-1})$ give

$$\mathbb{E}\left[\frac{1-p_n(b)}{p_n(b)}; m \le b\right] \le C_n \mathbb{E}[Z(H_n)(1+H_n)^{n-1}] + C_n \mathbb{E}Z(H_n)\, \mathbb{E}[\mathcal{I}(m/b)^{n-1}; m \le b] < \infty.$$

Together with the $m > b$ bound, this proves the lemma. □

For the region $m > b$, define the downward projection

$$K_\downarrow((m,k),b) = \inf_{0<z\le b,\, q>0} D((m,k)\|(z,q)).$$

The optimizer has $z = b$ and $h(q) = h(k) + \mathcal{I}(m/b)$, so $K_\downarrow((m,k),b) = A(k) - A(q)$. The continuous function

$$G_+(m,k) = D((m,k)\|\theta^\star) + K_\downarrow((m,k),b) \tag{109}$$

has a positive infimum on $[b,\infty) \times (0,\infty)$. The explicit Gamma KL is coercive: it diverges as $m \to \infty$, $k \downarrow 0$, or $k \to \infty$. Its compact sublevel sets contain every minimizing sequence. On such a set, $D = 0$ only at $\theta^\star$, where $K_\downarrow > 0$, while $K_\downarrow = 0$ only when $m = b$, where $D > 0$. Continuity therefore gives $\delta_+ := \inf G_+ > 0$.

For $m > b$, monotonicity and lemma 17 give $P(a, ar) \ge P(a,a) \ge c$ and $1 - P(a,ar) \le e^{-a\mathcal{I}(r)}$. Thus the posterior odds are at most $CZ_{n,\eta}\{u+\mathcal{I}(r)\}/Z_{n,\eta}(u)$. Applying (64) gives, with $\kappa = \kappa_n(u)$ and $q_n = \kappa_n\{u + \mathcal{I}(r)\}$,

$$\mathcal{O}_n(m,u) \le C\left(\frac{1+\kappa}{1+q_n}\right)^\eta e^{\ell_{n,u+\mathcal{I}(r)}(q_n) - \ell_{n,u}(\kappa)}.$$

Multiplication by the data density produces

$$\begin{aligned} &a_n(u) e^{\ell_{n,u}(k^\star)} e^{-s(nk^\star)} e^{-nk^\star \mathcal{I}(m/\mu^\star)} e^{\ell_{n,u+\mathcal{I}(r)}(q_n) - \ell_{n,u}(\kappa)} \\ &\quad = \{a_n(u) e^{\ell_{n,u}(\kappa)}\} e^{-s(nk^\star)} \exp\left\{\ell_{n,u}(k^\star) + \ell_{n,u+\mathcal{I}(r)}(q_n) - 2\ell_{n,u}(\kappa) - nk^\star \mathcal{I}(m/\mu^\star)\right\}. \end{aligned}$$

Thus (103) controls the factor in braces, and the finite-sample exponent to be bounded is exactly

$$\ell_{n,u}(k^\star) + \ell_{n,u+\mathcal{I}(r)}(q_n) - 2\ell_{n,u}(\kappa) - nk^\star \mathcal{I}(m/\mu^\star).$$

Let $\widehat{k} = h^{-1}(u)$ and $q = h^{-1}\{u + \mathcal{I}(r)\}$. Since $\widehat{k}$ and q maximize the two population log likelihoods, while κ and q_n maximize their finite-sample counterparts,

$$\begin{aligned} n\{\phi_u(\widehat{k}) - \phi_u(\kappa)\} &\le s(n\kappa) - s(n\widehat{k}), \\ \phi_{u+\mathcal{I}(r)}(q_n) &\le \phi_{u+\mathcal{I}(r)}(q). \end{aligned}$$

Substitution cancels the two copies of $s(n\kappa)$ and bounds the last display by

$$-nG_+(m,\widehat{k}) + s(nk^\star) + s(nq_n) - 2s(n\widehat{k}).$$

The factor $e^{-s(nk^\star)}$ displayed before this exponent cancels $e^{s(nk^\star)}$. Thus the finite-sample correction that remains in the full integrand is $e^{s(nq_n) - 2s(n\widehat{k})}$. To bound this finite-sample correction, integrate (67) from k to infinity and use the analogous bounds for h. This gives

$$\frac{c}{k} \le h_n(k), h(k) \le \frac{C}{k}, \qquad n \ge 2,\ k > 0. \tag{110}$$

Hence

$$\kappa \asymp \widehat{k}, \qquad q_n \asymp q, \qquad q^{-1} \asymp u + \mathcal{I}(r), \qquad \widehat{k}^{-1} \asymp u,$$

with constants independent of n, m, u. The two boundary estimates for s imply the global inequalities

$$e^{s(a)} \le C(a^{-1} + a^{-1/2}), \qquad e^{-s(a)} \le C(1 + \sqrt{a}), \qquad a > 0.$$

Therefore, for some fixed $\nu < \infty$, the complete correction satisfies

$$\mathcal{C}_n(m,u) := e^{s(nq_n)-2s(n\widehat{k})}\left(\frac{1+\kappa}{1+q_n}\right)^{\eta} \le Cn^{\nu}\{1+m+\widehat{k}+\widehat{k}^{-1}\}^{\nu}.$$

Here we used $q^{-1} \lesssim \widehat{k}^{-1} + \mathcal{I}(m/b)$, $q \le \widehat{k}$, and $\mathcal{I}(m/b) \lesssim 1+m$ on $m \ge b$.

The explicit Gamma KL gives constants $c, C > 0$ such that outside a compact subset of $[b,\infty) \times (0,\infty)$,

$$G_+(m,k) \ge c\{m + k^{-1} + \log(1+k)\} - C. \tag{111}$$

It follows that, for every $0 < \delta_0 < 1$, there are $C_{\delta_0}, \nu' < \infty$ such that

$$\mathcal{C}_n(m,u) \le C_{\delta_0} n^{\nu'} e^{n\delta_0 G_+(m,\widehat{k})}$$

for all sufficiently large n.

Finally use $m = br$, $u = h_n(\kappa)$, and (67). The measure is bounded by

$$C\sqrt{n}\,\frac{dr}{r}\,\frac{d\kappa}{\kappa}.$$

The integral of $e^{-n(1-\delta_0)G_+(br,\widehat{k})}$ against this measure is finite: (111) gives exponential decay in m and $1/\widehat{k}$, and a power $\widehat{k}^{-cn}$ at infinity; (110) transfers the same orders to κ. To retain the global gap in the exponent, factor

$$e^{-n(1-\delta_0)G_+} = e^{-n(1-\delta_0)\delta_+} e^{-n(1-\delta_0)(G_+-\delta_+)}.$$

On a sufficiently large compact set, the second factor is at most one and the measure is finite. Outside that set, enlarge it until (111) bounds $G_+ - \delta_+$ below by fixed positive multiples of m, $1/\widehat{k}$, and $\log(1+\widehat{k})$, up to constants already dominated there. The remaining one-dimensional tails are uniformly integrable for $n \ge n_0$. Splitting between these regions gives, after enlarging the polynomial power of n,

$$\mathbb{E}_{\theta^\star}[\mathcal{O}_n;\, m > b] \le C_{\delta_0} n^{\nu'} e^{-n(1-\delta_0)\delta_+}. \tag{112}$$

For each fixed preceding n, the uniform-in-(m,u) bound $p_n(b) \ge c > 0$ from lemma 17 on $m > b$ makes $\mathcal{O}_n$ uniformly bounded, so its expectation on this region is finite.

Equations (105), (112), and lemma 18 imply

$$\mathbb{E}_{\theta^\star}\frac{1-p_n(b)}{p_n(b)} \le Cn^{\nu'} e^{-cn}, \qquad n \ge n_0, \tag{113}$$

for an instance-dependent $c > 0$, with $n_0 \ge n_\eta$. The finitely many preceding expectations are finite by lemma 19. This proves (28). Combining (28), (27), and (62) gives $\mathbb{E}N_a(T) = O(\log T)$ for every suboptimal arm. Since the number of arms is finite,

$$R_T = \sum_{a:\Delta_a>0} \Delta_a \mathbb{E}N_a(T) = O(\log T) = o(T^\alpha), \qquad \alpha > 0. \tag{114}$$

Thus uncapped ITM-TS is uniformly efficient. The upper coefficient in proposition 9 and the Burnetas–Katehakis lower bound (6) now give (32).

G.6 Prior comparison and the limit of vanishing tempering

The sharp coefficient extends to priors that are globally comparable to a fixed ITM prior. The comparison acts directly on the posterior odds, so its constants are uniform over sample sizes and observed histories.

Proposition (globally comparable priors). Fix $\eta > 0$. Let $\widetilde{w}$ satisfy Assumption 1, and suppose that constants $0 < c \le C < \infty$ satisfy

$$c\, w_\eta(k) \le \widetilde{w}(k) \le C\, w_\eta(k), \qquad k > 0. \tag{115}$$

For a history $x_{1:n}$ with $n \geq 2$ and $H_n > 0$, let $p_{n,\eta}(b)$ and $\widetilde{p}_n(b)$ denote optimism probabilities under w_η and $\widetilde{w}$, respectively. Then, for every $b > 0$,

$$\frac{c}{C}\frac{1-p_{n,\eta}(b)}{p_{n,\eta}(b)} \leq \frac{1-\widetilde{p}_n(b)}{\widetilde{p}_n(b)} \leq \frac{C}{c}\frac{1-p_{n,\eta}(b)}{p_{n,\eta}(b)}. \tag{116}$$

Ordinary TS with common prior $\widetilde{w}(k)/\mu$, two initial pulls per arm, and exact independent posterior draws is uniformly efficient on the fixed finite-arm Gamma model with a unique optimal mean. Its expected suboptimal-arm pull counts satisfy (32).

Proof. Set $m = \bar{x}$, $u = H_n$, and

$$L_n(k;u) = R_n(k)e^{-nku}, \qquad F_{n,b}(k;m) = P(nk, nkm/b).$$

For a prior factor w, define the two unnormalized posterior integrals

$$N_w = \int_0^\infty w(k)L_n(k;u)F_{n,b}(k;m)\,dk, \qquad M_w = \int_0^\infty w(k)L_n(k;u)\{1-F_{n,b}(k;m)\}\,dk.$$

The conditional inverse-Gamma identity gives $p_{n,w}(b) = N_w/(N_w + M_w)$ and hence $(1 - p_{n,w}(b))/p_{n,w}(b) = M_w/N_w$. Both integrals are positive and finite: $0 < F_{n,b}(k;m) < 1$ for $k > 0$, and their sum is the proper shape-posterior normalizer. Multiplying (115) by the nonnegative likelihood and conditional-probability factors gives

$$cN_{w_\eta} \leq N_{\widetilde{w}} \leq CN_{w_\eta}, \qquad cM_{w_\eta} \leq M_{\widetilde{w}} \leq CM_{w_\eta}.$$

Dividing these inequalities proves (116). In particular, the factor is C/c because the common posterior normalizer cancels from the odds.

Fix any instance and any $0 < b < \mu^\star$. Taking expectations and summing (116) gives

$$\sum_{n=2}^\infty \mathbb{E}_{\theta^\star}\frac{1-\widetilde{p}_n(b)}{\widetilde{p}_n(b)} \leq \frac{C}{c}\sum_{n=2}^\infty \mathbb{E}_{\theta^\star}\frac{1-p_{n,\eta}(b)}{p_{n,\eta}(b)} < \infty,$$

where finiteness is established in the proof of theorem 10. Thus $\widetilde{w}$ satisfies (28) for every required threshold. Its regularity also supplies the selected-arm overshoot bound in lemma 7. For each suboptimal arm, choose $b \in (\mu_a, \mu^\star)$ and apply (62); the result is $\mathbb{E}N_a(T) = O(\log T)$. The finite sum over arms gives $R_T = O(\log T) = o(T^\alpha)$ for every $\alpha > 0$ on every fixed instance, which proves uniform efficiency on the model class. Finally, proposition 9 gives the upper coefficient in (32), and (6) gives the matching lower coefficient. □

Corollary (two endpoint asymptotics). Let $\widetilde{w}$ be positive and continuous on $(0,\infty)$, and suppose that, for fixed $\eta > 0$ and constants $C_0, C_\infty \in (0,\infty)$,

$$\widetilde{w}(k) \sim C_0 k^{-1} \quad (k \downarrow 0), \qquad \widetilde{w}(k) \sim C_\infty k^{-1-\eta} \quad (k \to \infty). \tag{117}$$

Then (115) holds, and ordinary TS with this prior has the sharp coefficient in (32).

Proof. By (10), the continuous positive ratio $\widetilde{w}/w_\eta$ converges to C_0 at zero and to $2C_\infty/(2\pi e)^\eta$ at infinity. Choose $0 < k_- < k_+ < \infty$ so that the ratio is bounded above and below by positive constants on $(0, k_-]$ and $[k_+, \infty)$. It attains a positive minimum and a finite maximum on $[k_-, k_+]$. Taking the minimum and maximum across the three intervals proves (115). The small-shape bound in Assumption 1 holds with $\rho = 1$, and the large-shape exponent is $\tau = 1 + \eta$. Continuity gives the required compact bounds. Posterior propriety after two observations follows by comparison with the ITM normalizer, which is finite almost surely by theorem 2. Thus every part of Assumption 1 holds, and the preceding proposition applies. □

For example,

$$\widetilde{w}_\eta(k) = A'(k)(1+k)^{-\eta}$$

has endpoint constants $C_0 = 1$ and $C_\infty = 1/2$, and its ratio to ITM converges to 1 and $(2\pi e)^{-\eta}$, respectively. This prior also has full shape support and the first-order matching form $\widetilde{w}_\eta(k)/\mu$. The endpoint conditions therefore describe a class of priors with the same asymptotic regret coefficient.

Proposition (pessimism budget as $\eta \downarrow 0$). Fix $\theta^\star = (\mu^\star, k^\star)$ and $0 < b < \mu^\star$. Define

$$S_\eta(b) = \sum_{n=2}^{\infty} \mathbb{E}_{\theta^\star} \frac{1 - p_{n,\eta}(b)}{p_{n,\eta}(b)}. \tag{118}$$

For every $\eta > 0$, $S_\eta(b) < \infty$, whereas

$$\lim_{\eta \downarrow 0} S_\eta(b) = \infty. \tag{119}$$

In fact, for every fixed $n \geq 2$,

$$\lim_{\eta \downarrow 0} \mathbb{E}_{\theta^\star} \frac{1 - p_{n,\eta}(b)}{p_{n,\eta}(b)} = \infty. \tag{120}$$

Proof. Finiteness for each positive η is (28), proved above. To study the limit, fix $n \geq 2$ and a history with $H_n = u > 0$. Since $1 + e^{2A(k)} > 1$,

$$0 < w_\eta(k) \leq w_0(k) = A'(k), \qquad w_\eta(k) \longrightarrow w_0(k) \quad \text{as } \eta \downarrow 0.$$

The function $w_0(k) R_n(k) e^{-nku}$ is integrable: near zero its order is k^{n-2}, and at infinity its order is $k^{(n-3)/2} e^{-nku}$, by (36) and the fixed-n Stirling expansion. It dominates both the posterior normalizer and the optimism numerator for every $\eta > 0$, since $0 < F_{n,b}(k; m) < 1$. Dominated convergence therefore gives

$$N_{w_\eta} \longrightarrow N_{w_0} > 0, \qquad M_{w_\eta} \longrightarrow M_{w_0} > 0,$$

and hence

$$\frac{1 - p_{n,\eta}(b)}{p_{n,\eta}(b)} \longrightarrow \frac{1 - p_{n,0}(b)}{p_{n,0}(b)}.$$

Gamma samples satisfy $H_n > 0$ almost surely, so this convergence holds almost surely under $\mathbb{P}_{\theta^\star}$.

Choose fixed $0 < r_- < r_+ < 1$ and let $E_n = \{\bar{X}_n / b \in [r_-, r_+]\}$. The Gamma density of $\bar{X}_n$ is positive on $(0, \infty)$, so $\mathbb{P}_{\theta^\star}(E_n) > 0$. The matching prior $w_0 = A'$ has $\tau = 1$. Equation (24) therefore yields

$$\mathbb{E}_{\theta^\star}\big[p_{n,0}(b)^{-1} \mid E_n\big] = \infty.$$

Subtracting one and multiplying by the positive probability of E_n gives

$$\mathbb{E}_{\theta^\star} \frac{1 - p_{n,0}(b)}{p_{n,0}(b)} = \infty.$$

Apply Fatou's lemma to the nonnegative odds along any sequence $\eta_j \downarrow 0$:

$$\infty = \mathbb{E}_{\theta^\star} \liminf_{j \to \infty} \frac{1 - p_{n,\eta_j}(b)}{p_{n,\eta_j}(b)} \leq \liminf_{j \to \infty} \mathbb{E}_{\theta^\star} \frac{1 - p_{n,\eta_j}(b)}{p_{n,\eta_j}(b)}.$$

The same argument applies to any positive sequence converging to zero, which proves (120). Finally, $S_\eta(b)$ is at least its $n = 2$ summand, proving (119). □

The boundary estimate in lemma 18 displays the same parameter dependence through

$$\int_n^\infty \kappa^{-1-\eta}\, d\kappa = \frac{n^{-\eta}}{\eta}.$$

Equation (119) identifies a nonuniformity of the posterior-odds term itself, beyond the choice of this upper bound. In the regret reduction (62), $S_\eta(b)$ is an additive control term, and for each fixed $\eta > 0$ the sharp logarithmic coefficient is still $1/K_{\inf}(\theta_a, \mu^\star)$.

H COMPACT PARAMETER SPACES

Let $\Theta = [\mu_-, \mu_+] \times [k_-, k_+] \Subset (0, \infty)^2$ be a closed compact rectangle and define

$$K_{\inf}^{\Theta}(\theta, b) = \inf_{(m,q)\in\Theta:\, m\geq b} D(\theta\|(m, q)). \tag{121}$$

Because the feasible alternative set in (121) is a subset of the one in (5),

$$K_{\inf}^{\Theta}(\theta, b) \geq K_{\inf}(\theta, b).$$

The intersection of Θ with $\{m \geq b\}$ is closed and compact. If $\mu_+ < b$, this intersection is empty and $K_{\inf}^{\Theta}(\theta, b) = +\infty$. Otherwise the restricted feasible set is nonempty. For a source $\theta = (\mu, k)$ with $\mu < b$, theorem 3 identifies the unique unrestricted optimizer as (b, q_b). Therefore

$$K_{\inf}^{\Theta}(\theta, b) = K_{\inf}(\theta, b) \quad \Longleftrightarrow \quad (b, q_b) \in \Theta. \tag{122}$$

If $(b, q_b) \notin \Theta$, the restricted feasible set has positive distance from the unique unrestricted minimizer. Continuity of the Gamma KL on that set then gives $K_{\inf}^{\Theta}(\theta, b) > K_{\inf}(\theta, b)$. Thus a hard cap preserves the unrestricted information constant exactly on instances whose optimistic projection remains inside the cap; otherwise it changes the statistical model and its lower-bound constant.

I COMPUTATIONAL DETAILS

Table 1: Consequences of the high-shape prior tail. Frozen recovery orders are exact for $n = 2$; the policy column records theorems 6 and 10.

Prior	Tail exponent τ	Frozen recovery mean	Ordinary-TS consequence
Jeffreys	$1/2$	$\Theta(\sqrt{T})$	$\Omega(\sqrt{T})$
Matching	1	$\Theta(\log T)$	$\Omega(\log T)$ lockout
ITM, $\eta > 0$	$1 + \eta$	$\Theta(1)$	sharp $1/K_{\inf}$ coefficient

The code stores (n, S, L) with $L = \sum_i \log x_i$ and computes $H = \log(S/n) - L/n$. It uses a grid uniform in $y = \log k$ and normalizes $p(e^y \mid x)e^y$, including the Jacobian. Conditional simulation is exact: if $G \sim \mathrm{Gamma}(nk, 1)$, then $\mu = kS/G$ has (12). The scalar equation (15) is solved by bracketing because $h'(k) < 0$.

I.1 BENCHMARK INSTANCES

The full-bandit and sampler-sensitivity experiments use three instances:

$$\begin{aligned}
\mathcal{E}_{\text{high}} &= \{(1.2, 2), (1.0, 100)\},\\
\mathcal{E}_{\text{gap}} &= \{(1.02, 2), (1.0, 20)\},\\
\mathcal{E}_5 &= \{(1.2, 2), (1.1, 20), (1.05, 50), (.95, 5), (.9, 100)\}.
\end{aligned}$$

Each pair denotes (mean, shape). The benchmark settings and outcome measures are given in section I.4.

I.2 POSTERIOR COMPUTATION AND NUMERICAL VALIDATION

The numerical sampler uses the complete shape posterior. Its log density in $y = \log k$, including the prior and Jacobian, is

$$g_{n,w}(y; u) = \log w(e^y) + \log R_n(e^y) - ne^y u + y. \tag{123}$$

Histories with $n \leq 50$ or $H_n \leq 10^{-4}$ use adaptive quadrature in y; uncapped histories with $n > 50$ and $H_n > 10^{-4}$ use rejection sampling from a Gamma envelope. Both branches retain the selected prior. Unchanged arm histories reuse the posterior representation and draw fresh random samples each round.

Gamma envelope. Write $u = H_n > 0$ and set $a_n = (n-1)/2$ and $f_n(k) = \log R_n(k) - a_n \log k$. This function is concave. Indeed, the trigamma integral gives

$$\psi_1(k) - n\psi_1(nk) = \int_0^\infty te^{-kt} \left\{ \frac{1}{1-e^{-t}} - \frac{1/n}{1-e^{-t/n}} \right\} dt \geq \frac{1-1/n}{2k^2}.$$

For the inequality, write the bracket as $(1-1/n)/2 + F(t) - F(t/n)/n$, where $F(t) = \frac{1}{2}\coth(t/2) - 1/t$ is nonnegative and increasing: $F(0+) = 0$ and $F'(t) = t^{-2} - \{4\sinh^2(t/2)\}^{-1} > 0$. Consequently $f_n''(k) \leq 0$. For any anchor $c > 0$, write $d = f_n'(c)$ and choose it so that $\beta = nu - d > 0$. Such anchors exist because $f_n'(c) \to 0$ as $c \to \infty$. Concavity supplies the global tangent bound $f_n(k) \leq f_n(c) + d(k-c)$.

If $w(k) \leq Ck^{s-1}$ and $a_n + s > 0$, propose $K \sim \mathrm{Gamma}(a_n + s, \mathrm{rate} = \beta)$ and accept with probability

$$\exp\{\log w(K) - \log C - (s-1)\log K + f_n(K) - f_n(c) - d(K-c)\}. \tag{124}$$

The exponent is nonpositive, and multiplying the proposal density by this acceptance probability gives a constant times $w(K)R_n(K)e^{-nuK}$. Thus the accepted shape has the full target posterior; the anchor affects efficiency but does not approximate that target. The bound $A'(k) \leq 1/k$ gives $(s, C) = (0, 1)$ for matching and ITM, and $(s, C) = (1/2, 1)$ for Jeffreys. For ITM with $\eta \geq 0$, a second envelope uses

$$s = -\eta, \qquad C = (2\pi e)^\eta e^{\eta/3}.$$

For $k \leq 1$, this follows from $w_\eta(k) \leq 1/k$. For $k \geq 1$, the bounds $\psi(k) \geq \log k - (2k)^{-1} - (12k^2)^{-1}$ and $\log\Gamma(k) \leq (k - \frac{1}{2})\log k - k + \frac{1}{2}\log(2\pi) + (12k)^{-1}$ give $A(k) \geq \frac{1}{2}\log\{k/(2\pi e)\} - (6k)^{-1}$. When both Gamma shapes are positive, the implementation chooses the envelope with the smaller integral, computed from $C\Gamma(a_n + s)\beta^{-a_n - s}$. Large-shape tangent residuals use differences of Stirling remainders, avoiding subtraction of large log-Gamma terms. Capped priors use quadrature at every sample size.

Adaptive integration and rare optimism. The quadrature interval endpoints solve a prescribed log-density drop from the mode, so the mesh follows posterior concentration; the numerical upper endpoint adapts to the data. Conditional mean draws use the inverse-Gamma representation stated above. A rare optimism query contains the additional factor $P(nk, nkr)$, with $r = m/b$, whose mass can lie outside the bulk posterior window. Such queries therefore locate and scale their numerator and denominator separately. Independent reference calculations expand integration ranges, tighten tolerances, and check special functions with high-precision arithmetic. The histories include the most pessimistic observations in the frozen experiment, constructed near-ties, and concentrated posteriors. The regret theorem concerns exact posterior TS; finite quadrature meshes and floating-point arithmetic are the numerical approximations in these runs. Across 218 history–prior combinations, reference optimism probabilities span 1.71×10^{-272} to .476. With 160 grid points, the largest relative probability error is 3.28×10^{-6}; the rare-event adaptive subset has maximum relative error below 1.40×10^{-11}. These checks concern posterior queries, separately from random draws. For the rejection sampler itself, 4,480 high-precision acceptance checks and 20 independently integrated accepted densities verify the target identity; the latter reproduce optimism down to 1.71×10^{-272} with maximum relative error 5.42×10^{-12}. A separate native implementation is checked against high-precision special functions and the Python implementation. Nine seeded, 4,000-draw posterior comparisons have maximum two-sample KS distance .024.

Complete-run numerical sensitivity uses 64 trials per instance through $T = 2{,}000$, comparing the production grid/rejection sampler with 320-point quadrature at every sample size. Table 2 reports the fixed $\eta = .5$ setting; the supplement gives all five policies and a 320-point hybrid comparison that isolates the late sampling branch. The paired intervals measure finite-run differences; their widths do not establish equivalence or rare-tail convergence.

I.3 Frozen-posterior experiment

The frozen experiment uses seed 260809, $n = 2$, true $(\mu, k) = (1.2, 2)$, threshold $b = 1$, 5,000 independent data sets, 512 log-shape grid points, and horizons $T \in \{100, 300, 1000, 3000, 10^4, 3 \cdot 10^4, 10^5, 3 \cdot 10^5, 10^6\}$. It compares a heavier-tail $\tau = 1/2$ control, Jeffreys, matching, ITM with

Table 2: Numerical sensitivity for ITM with $\eta = .5$ at $T = 2{,}000$ over 64 trials. The last column gives a pointwise paired 95% interval for the mean-regret difference, production minus quadrature (5,000 resamples, seed 917264).

Instance	Production mean	Quadrature mean	Difference interval
$\mathcal{E}_{\mathrm{high}}$	7.15	6.72	$[-0.22, 1.34]$
$\mathcal{E}_{\mathrm{gap}}$	19.57	18.85	$[-0.01, 1.84]$
$\mathcal{E}_5$	30.23	29.41	$[-0.54, 2.77]$

Table 3: Pseudo-regret at $T = 20{,}000$ over 1,024 trials. Parentheses contain standard errors; $\mathrm{CVaR}_{.95}$ averages the worst 52 trials. "Severe" counts trials whose longest interval without an optimal-arm pull is at least $.25T$; Wilson intervals are included in the released CSV.

Instance	Policy	Mean (SE)	Median	$\mathrm{CVaR}_{.95}$	Severe
$\mathcal{E}_{\mathrm{high}}$	Jeffreys	50.73 (11.34)	3.80	851.23	10
	Matching	15.89 (4.22)	3.80	200.01	1
	ITM, $\eta = .25$	15.94 (4.47)	3.60	202.62	1
	ITM, $\eta = .5$	16.12 (4.40)	3.80	207.70	1
	ITM, $\eta = 1$	14.71 (3.69)	3.80	180.80	1
$\mathcal{E}_{\mathrm{gap}}$	Jeffreys	75.75 (2.89)	38.19	370.62	23
	Matching	69.33 (2.63)	37.01	345.31	2
	ITM, $\eta = .25$	69.01 (2.63)	36.61	347.37	1
	ITM, $\eta = .5$	69.13 (2.63)	36.59	345.78	1
	ITM, $\eta = 1$	68.61 (2.60)	36.79	342.25	3
$\mathcal{E}_5$	Jeffreys	65.40 (5.99)	30.32	612.15	11
	Matching	43.87 (2.41)	30.93	225.62	1
	ITM, $\eta = .25$	43.08 (2.35)	31.00	212.79	1
	ITM, $\eta = .5$	42.73 (2.28)	31.15	204.10	1
	ITM, $\eta = 1$	43.56 (2.36)	31.40	217.33	1

$\eta \in \{0.25, 0.5, 0.75, 1\}$, and capped matching. For each data set, the number plotted at horizon T is the exact conditional expected number of failures

$$\frac{(1-p_b)\{1-(1-p_b)^T\}}{p_b},$$

computed stably with `log1p` and `expm1`. The capped baseline uses $k_{\max} = 50$. Machine-readable outputs include the mean, standard error, median, 90th and 99th percentiles, empirical $\mathrm{CVaR}_{.99}$ (the mean of the largest 1% of trial outcomes), and the severe-lockout frequency $\mathbb{P}(p_b < T^{-1})$.

I.4 Ordinary-TS benchmark

The ordinary-TS benchmark uses seed 260813, 1,024 independent trials, and horizon $T = 20{,}000$. Within a trial, all policies receive common pre-generated reward streams; posterior draws use separate random streams. The released benchmark uses NumPy PCG64 reward streams and the C++ standard library Mersenne Twister for posterior draws. Posterior sampling uses a 160-point adaptive log-shape grid through 50 observations (and whenever $H_n \le 10^{-4}$ thereafter), and Gamma-envelope rejection sampling otherwise, as specified in section I.2. Results are recorded at $T \in \{100, 300, 1000, 3000, 10000, 20000\}$. The instances are defined in section I.1. The benchmark compares Jeffreys TS, matching TS, and ITM-TS with $\eta \in \{.25, .5, 1\}$. Empirical $\mathrm{CVaR}_{.95}$ averages the largest 52 trial outcomes. A severe lockout means that the longest consecutive run without an optimal-arm pull is at least $.25T$. Reported 95% confidence intervals equal the sample mean plus or minus 1.96 standard errors. Prefix estimates at 96, 192, 384, 768, and 1,024 trials are accompanied by percentile intervals from 2,000 nonparametric bootstrap resamples with seed 260821; severe-lockout rates use 95% Wilson intervals. Table 3 reports the final checkpoint; Figures 4 and 5 show the horizon trajectories and prefix stability.

Table 4: Paired differences, ITM with $\eta = .5$ minus matching, at $T = 20{,}000$. Negative values favor ITM. Intervals are pointwise 95% bootstrap intervals.

Instance	Mean difference	Interval	$\text{CVaR}_{.95}$ difference	Interval
$\mathcal{E}_{\text{high}}$	0.24	$[-2.77, 4.05]$	7.69	$[-51.21, 82.91]$
$\mathcal{E}_{\text{gap}}$	-0.20	$[-0.82, 0.34]$	0.47	$[-3.28, 4.51]$
$\mathcal{E}_5$	-1.15	$[-2.39, -0.16]$	-21.52	$[-44.48, -4.39]$

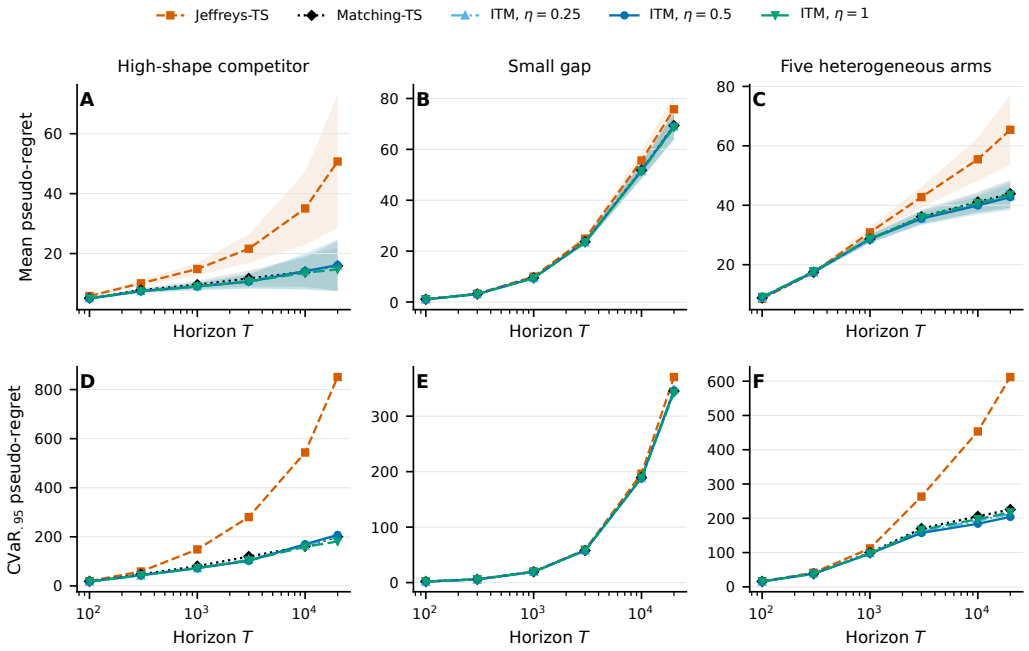

Figure 4: Full ordinary-TS benchmark. The top row gives mean pseudo-regret with 95% confidence intervals; the bottom row gives $\text{CVaR}_{.95}$. Columns correspond to $\mathcal{E}_{\text{high}}$, $\mathcal{E}_{\text{gap}}$, and $\mathcal{E}_5$.

We also compare policies by resampling complete trial identities jointly, which retains their common reward streams. Table 4 gives ITM minus matching differences and pointwise percentile intervals from 5,000 paired bootstrap resamples (seed 260906). Each resample recomputes both tail means using its largest $\lceil .05M \rceil$ outcomes. The released CSV includes all policies and the corresponding starvation-gap comparisons.

I.5 Cap sensitivity and Fourier check

The cap study uses seed 260814, 96 trials, horizon $T = 10{,}000$, and the same random-stream design. It compares matching priors truncated at $k_{\max} \in \{10, 30, 100, 300\}$ with matching TS and uncapped ITM at $\eta = .5$ on the high-shape instance $\{(1.2, 2), (1.0, 100)\}$ and the moderate-shape instance $\{(1.4, 5), (1.0, 2)\}$. In the high-shape instance, the final (mean regret, $\text{CVaR}_{.95}$) pair is $(9.93, 67.08)$ for ITM, compared with $(14.00, 81.20)$, $(12.19, 95.16)$, $(11.44, 90.40)$, and $(11.97, 105.56)$ for caps $10, 30, 100, 300$. On the moderate-shape instance, the six means lie between 21.50 and 22.57. Figure 6 shows how the cap ranking changes across the two instances.

The exact characteristic function in (73) also permits direct Fourier inversion. Over $n \in \{4, 8, 16, 32, 64\}$ and $\kappa \in \{1, 3, 10, 100, 10^4\}$, the standardized center density ranges by n as follows:

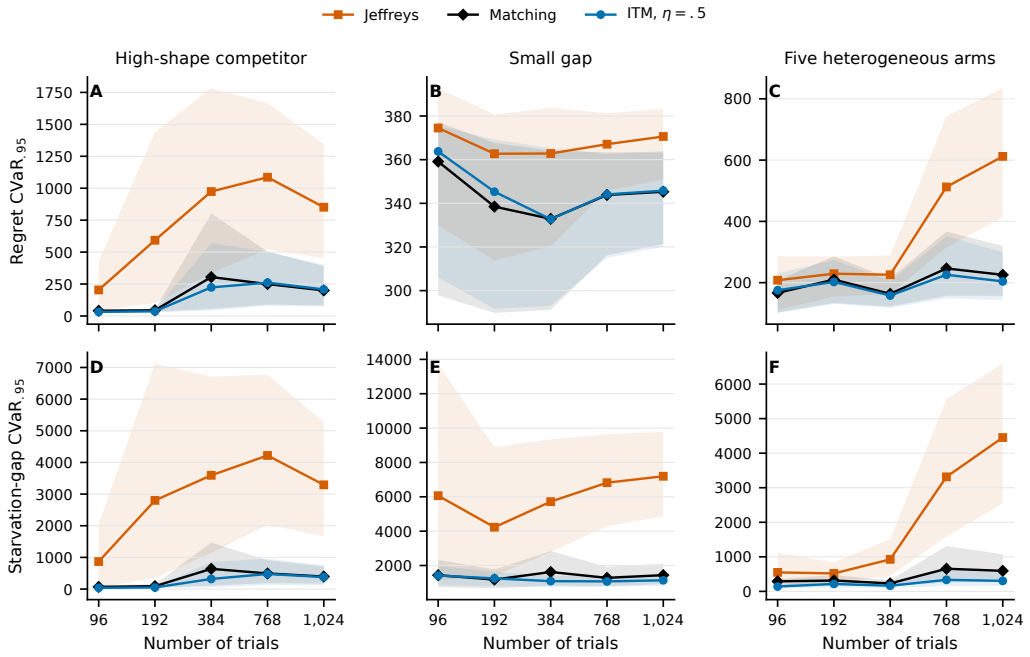

Figure 5: Prefix stability of full-bandit tail summaries. Lines use the first 96, 192, 384, 768, and 1,024 seeded trials; shaded regions are 95% bootstrap intervals. The top row reports regret $\mathrm{CVaR}_{.95}$ and the bottom row reports $\mathrm{CVaR}_{.95}$ of the longest interval without an optimal-arm pull.

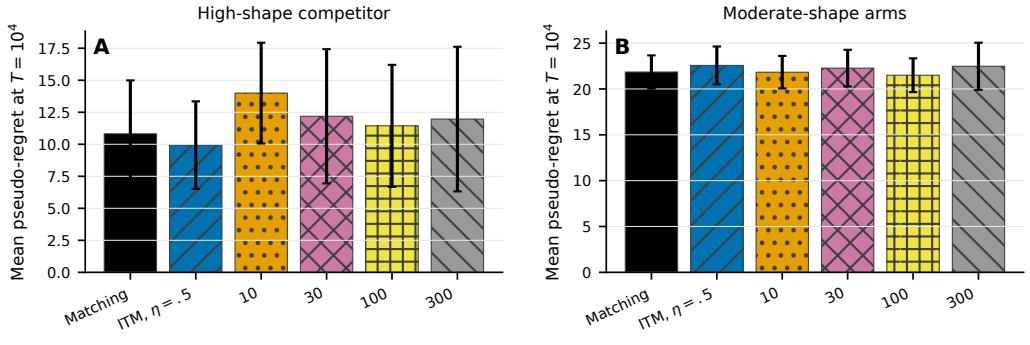

Figure 6: Sensitivity to a fixed upper shape cap. Bars give mean pseudo-regret at $T = 10^4$ and error bars give 95% confidence intervals. Uncapped ITM uses $\eta = .5$.

n	4	8	16	32	64
minimum	.374688	.388668	.394135	.396617	.397798
maximum	.378299	.389582	.394537	.396803	.397888

Across the tested shape grid, the values approach the Gaussian center density $(2\pi)^{-1/2} = .398942$ as n increases, consistent with (84). Configuration files, seeds, trial-level CSV files, summaries, and plotting scripts accompany the paper.

J ADDITIONAL POLICY AND REWARD-MODEL EXPERIMENTS

J.1 A COMMON POSTERIOR UNDER TWO DECISION RULES

The comparison crosses two prior factors, w_0 and $w_{.5}$, with ordinary TS and the per-arm ϵ-TS rule of Jin et al. (2023). We adapt that rule to the two-parameter Gamma posterior in (13). At each round and for each arm, independently draw $B_{a,t} \sim \text{Bernoulli}(1/K)$ and use the score

$$V_{a,t} = \begin{cases} \widetilde{\mu}_{a,t}, & B_{a,t} = 1, \\ \bar{X}_{a,N_a(t-1)}, & B_{a,t} = 0. \end{cases}$$

The arm with the largest score is selected; ties are broken uniformly at random. A posterior sample is computed only when $B_{a,t} = 1$. Thus the four policies separate the prior choice from the frequency of posterior sampling. Ordinary TS corresponds to $B_{a,t} = 1$ for every arm. Both rules use the same Gamma-envelope and 160-point quadrature implementation from section I. The decision-rule adaptation is evaluated here with both mean and shape unknown.

Each of the three instances in section I.1 uses 1,024 trials, two initial pulls per arm, and horizon $T = 20{,}000$. Within a trial, all policies share the complete reward array of each arm; its next reward is indexed by that arm's pull count. Posterior draws and Bernoulli decisions use separate random-number streams. The Gamma controls retain the original reward and posterior seeds, allowing trajectory-by-trajectory comparison with section I.4. The seed is 260813, and checkpoints are $100, 300, 1000, 3000, 10000, 20000$. Policy settings are fixed in `configs/v14_gamma.json` and `configs/v14_misspecification.json`.

J.2 DEPARTURES FROM THE GAMMA REWARD MODEL

For each base arm (μ, k), two alternative reward laws preserve its mean μ and variance μ^2/k. The first is lognormal:

$$\log X \sim \mathcal{N}\left(\log \mu - \tfrac{1}{2}s^2, s^2\right), \qquad s^2 = \log(1 + 1/k).$$

The second is a mixture of Gamma laws in the paper's mean–shape notation:

$$Q_{\mu,k} = \tfrac{1}{3}P_{\mu,k/2} + \tfrac{2}{3}P_{\mu,2k}.$$

Its component means both equal μ, and its variance is $\frac{1}{3}(2\mu^2/k) + \frac{2}{3}(\mu^2/(2k)) = \mu^2/k$. The lognormal variance follows from $\text{Var}(X) = \mu^2(e^{s^2} - 1)$. The three original mean configurations therefore keep their reward gaps and variances, while higher moments change. Each alternative has 1,024 trials per configuration at the same horizon. Every policy continues to fit the Gamma likelihood; these runs measure its finite-horizon behavior under model misspecification.

J.3 OUTCOME INTERVALS AND COMPUTATIONAL COST

Mean pseudo-regret, $\text{CVaR}_{.95}$, the longest interval without an optimal-arm pull, and severe-lockout frequency are computed from complete trajectories. CVaR averages the largest 52 of 1,024 outcomes. For the new paired comparisons, 5,000 bootstrap samples resample whole trial identities jointly across all four policies, with seed 260916. Each resample recomputes both the mean and the CVaR tail. The reported percentile intervals are pointwise 95%; severe-lockout frequencies use Wilson intervals. The released CSVs contain comparisons against both ordinary matching TS and ordinary ITM-TS.

The implementation records native thread CPU time for each complete trajectory. Dividing by $T - 2K$ gives amortized CPU time per decision; reported medians and 95th percentiles summarize these trial-level averages. Reward generation and Python orchestration are outside the timed region. The sampler records posterior calls, quadrature calls, rejection calls, and proposal counts. All four policies use the same compiled numerical kernel and instrumentation.

The runs used an AMD EPYC 9V74 processor, Linux x86_64, and GCC 13.3 with `-O3` optimization. Eight independent workers ran one native thread each. The source package records the compiler, kernel, seeds, and completed-trial counts.

Table 5: Gamma reward comparisons at $T = 20{,}000$, with 1,024 trials per instance. ITM uses $\eta = .5$; the adapted ϵ rules use $1/K$. "Severe" counts trials with an optimal-arm starvation interval of at least $T/4$. The matching and ITM controls use the trajectories in table 3.

Instance	Policy	Mean regret	CVaR$_{.95}$	Severe
$\mathcal{E}_{\text{high}}$	Matching	15.89	200.01	1
	ITM	16.12	207.70	1
	ϵ-Matching	27.23	373.45	2
	ϵ-ITM	23.73	321.38	2
$\mathcal{E}_{\text{gap}}$	Matching	69.33	345.31	2
	ITM	69.13	345.78	1
	ϵ-Matching	84.80	388.34	4
	ϵ-ITM	84.95	388.14	4
$\mathcal{E}_5$	Matching	43.87	225.62	1
	ITM	42.73	204.10	1
	ϵ-Matching	100.66	897.46	4
	ϵ-ITM	94.55	816.27	2

Table 6: Paired Gamma comparisons of each adapted ϵ rule with ordinary ITM-TS at $\eta = .5$. Positive differences favor ordinary ITM. Intervals are pointwise 95% bootstrap intervals. The original prior comparison is in table 4.

Instance	Contrast	Mean Δ	Interval	CVaR Δ	Interval
$\mathcal{E}_{\text{high}}$	ϵ-Matching $-$ ITM	11.11	[7.03, 16.25]	165.75	[89.88, 262.90]
	ϵ-ITM $-$ ITM	7.60	[4.56, 11.70]	113.68	[56.94, 191.32]
$\mathcal{E}_{\text{gap}}$	ϵ-Matching $-$ ITM	15.67	[13.08, 18.43]	42.56	[27.12, 62.80]
	ϵ-ITM $-$ ITM	15.81	[13.15, 18.52]	42.36	[27.16, 62.27]
$\mathcal{E}_5$	ϵ-Matching $-$ ITM	57.94	[46.69, 70.70]	693.36	[529.63, 864.51]
	ϵ-ITM $-$ ITM	51.83	[41.80, 63.00]	612.17	[476.03, 757.47]

Gamma rewards. Both adapted ϵ rules have higher mean regret and CVaR than ordinary ITM on all three instances, with paired intervals excluding zero (table 6). On $\mathcal{E}_5$, for example, adapted ϵ-Matching increases mean regret by 57.94 (95% interval $[46.69, 70.70]$) and CVaR by 693.36 ($[529.63, 864.51]$). Table 4 gives the corresponding ordinary-prior comparisons.

Reward-model departures. Across the six lognormal and Gamma-mixture settings, all ITM–matching intervals include zero for both mean regret and CVaR (tables 10 and 11). The finite-horizon effect of tempering is therefore unresolved in these settings. The adapted ϵ-Matching rule has higher mean regret and CVaR than ordinary ITM in all six settings, with the paired intervals excluding zero.

Computational tradeoff. On Gamma rewards, ordinary ITM uses 4.77–7.70 microseconds of amortized CPU time per decision; the adapted ϵ-ITM rule uses 2.14–3.28 microseconds (table 7). The latter requests posterior samples for an expected one arm per round, compared with K under ordinary TS. Its lower sampling cost accompanies larger regret and upper-tail losses in these experiments.

Table 7: Amortized native CPU time per decision for Gamma rewards, in microseconds. Mean, median, and p95 summarize trial-level CPU time divided by $T - 2K$. Grid is the percentage of posterior draws using quadrature; proposals is the mean number per accepted rejection draw.

Instance	Policy	Mean	Median	p95	Grid (%)	Proposals
$\mathcal{E}_{\mathrm{high}}$	Matching	4.51	4.33	6.20	40.08	1.88
	ITM	4.77	4.53	6.85	40.67	2.05
	ϵ-Matching	2.40	2.31	3.48	36.90	1.89
	ϵ-ITM	2.51	2.38	3.74	38.22	2.07
$\mathcal{E}_{\mathrm{gap}}$	Matching	5.37	5.17	6.69	1.50	1.92
	ITM	6.14	5.94	7.72	1.41	2.43
	ϵ-Matching	2.89	2.77	3.76	2.49	1.92
	ϵ-ITM	3.28	3.11	4.21	2.39	2.43
$\mathcal{E}_5$	Matching	6.94	6.86	9.31	55.25	1.92
	ITM	7.70	7.59	10.49	55.27	2.36
	ϵ-Matching	1.96	1.91	2.77	61.19	1.92
	ϵ-ITM	2.14	2.08	2.99	61.13	2.35

Table 8: Lognormal rewards with the same arm means and variances as the Gamma instances. Horizon, trials, prior factors, and decision rules are as in table 5.

Instance	Policy	Mean regret	$\mathrm{CVaR}_{.95}$	Severe
$\mathcal{E}_{\mathrm{high}}$	Matching	12.55	133.72	0
	ITM	12.02	124.34	0
	ϵ-Matching	27.03	358.28	1
	ϵ-ITM	20.02	230.78	0
$\mathcal{E}_{\mathrm{gap}}$	Matching	75.07	390.23	9
	ITM	74.55	389.90	8
	ϵ-Matching	93.39	396.91	19
	ϵ-ITM	92.54	396.71	18
$\mathcal{E}_5$	Matching	58.63	490.44	5
	ITM	58.04	476.89	4
	ϵ-Matching	142.46	1390.50	16
	ϵ-ITM	139.50	1385.44	19

Table 9: Gamma-mixture rewards with the same arm means and variances as the Gamma instances. Settings follow table 5.

Instance	Policy	Mean regret	$\mathrm{CVaR}_{.95}$	Severe
$\mathcal{E}_{\mathrm{high}}$	Matching	16.96	220.93	2
	ITM	16.49	213.61	2
	ϵ-Matching	22.55	287.33	2
	ϵ-ITM	19.58	238.62	1
$\mathcal{E}_{\mathrm{gap}}$	Matching	65.94	323.08	0
	ITM	65.79	325.04	0
	ϵ-Matching	82.62	383.07	1
	ϵ-ITM	82.52	382.80	0
$\mathcal{E}_5$	Matching	42.01	193.19	1
	ITM	41.37	175.90	0
	ϵ-Matching	89.50	737.50	2
	ϵ-ITM	89.39	751.10	2

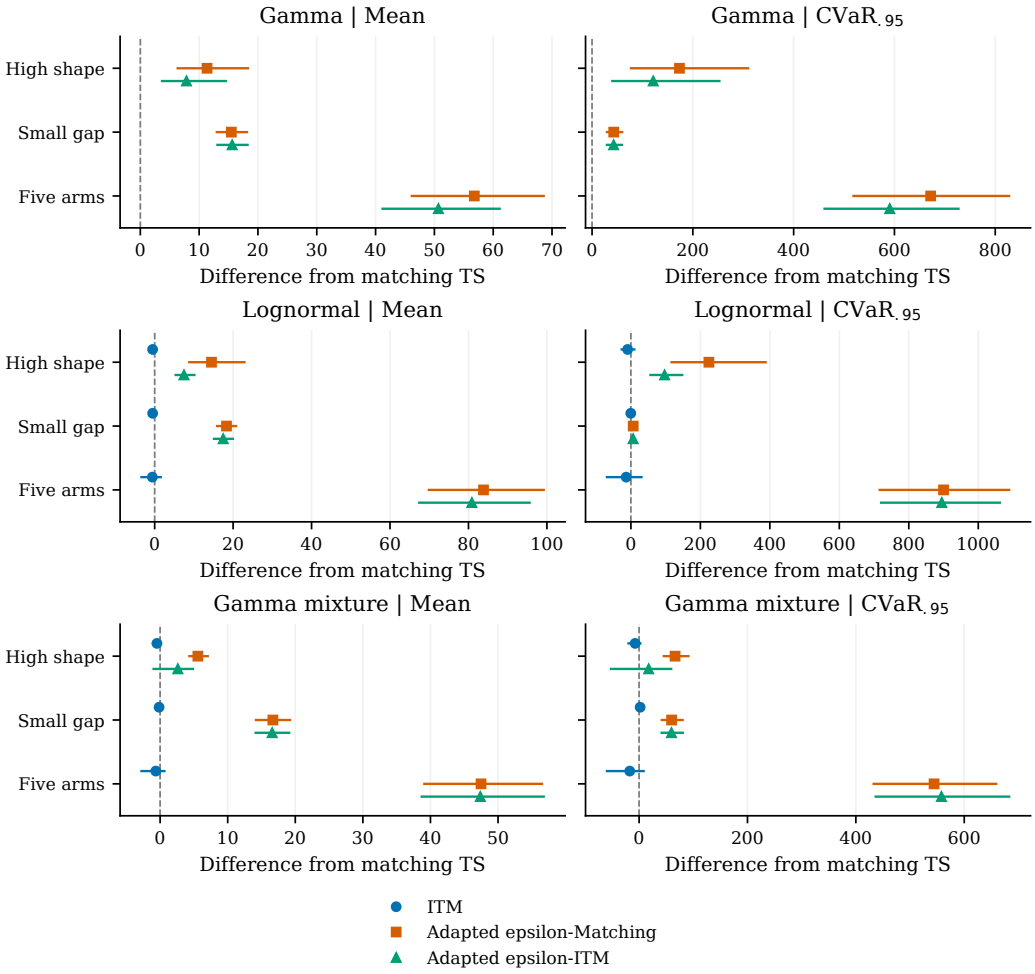

Figure 7: Paired policy comparisons with ordinary matching TS under three reward families. The Gamma ITM–matching comparison is in table 4. Points show mean and $\text{CVaR}_{.95}$ differences; lines give pointwise 95% bootstrap intervals. Negative values favor the plotted policy. Every ITM variant uses $\eta = .5$, and each setting has 1,024 trials at $T = 20{,}000$. Axis ranges vary across panels.

Table 10: Paired differences under lognormal rewards, with pointwise 95% bootstrap intervals. Each contrast subtracts the second policy from the first.

Instance	Contrast	Mean Δ	Interval	CVaR Δ	Interval
$\mathcal{E}_{\text{high}}$	ITM − Matching	-0.52	[-1.62, 0.64]	-9.38	[-29.90, 13.26]
	ϵ-Matching − ITM	15.01	[8.66, 24.32]	233.93	[114.14, 410.39]
	ϵ-ITM − ITM	8.00	[5.43, 11.25]	106.44	[59.68, 165.10]
$\mathcal{E}_{\text{gap}}$	ITM − Matching	-0.52	[-1.25, 0.14]	-0.33	[-2.68, 1.83]
	ϵ-Matching − ITM	18.83	[16.08, 21.61]	7.01	[3.68, 15.23]
	ϵ-ITM − ITM	17.99	[15.34, 20.71]	6.81	[3.46, 14.81]
$\mathcal{E}_5$	ITM − Matching	-0.59	[-3.66, 1.84]	-13.55	[-72.33, 33.81]
	ϵ-Matching − ITM	84.43	[70.63, 99.80]	913.61	[734.98, 1093.98]
	ϵ-ITM − ITM	81.46	[67.74, 96.34]	908.55	[729.57, 1071.56]

Table 11: Paired differences under Gamma-mixture rewards, with pointwise 95% bootstrap intervals. Contrast orientation follows table 10.

Instance	Contrast	Mean Δ	Interval	CVaR Δ	Interval
$\mathcal{E}_{\text{high}}$	ITM − Matching	-0.47	[-1.24, 0.17]	-7.32	[-21.63, 4.40]
	ϵ-Matching − ITM	6.05	[4.56, 7.95]	73.73	[48.56, 105.94]
	ϵ-ITM − ITM	3.09	[-0.68, 5.67]	25.01	[-49.53, 72.17]
$\mathcal{E}_{\text{gap}}$	ITM − Matching	-0.15	[-0.58, 0.25]	1.96	[-0.88, 6.17]
	ϵ-Matching − ITM	16.84	[14.12, 19.56]	58.03	[37.78, 79.50]
	ϵ-ITM − ITM	16.73	[14.09, 19.42]	57.76	[37.68, 79.64]
$\mathcal{E}_5$	ITM − Matching	-0.64	[-2.93, 0.81]	-17.29	[-61.50, 10.39]
	ϵ-Matching − ITM	48.13	[39.14, 57.93]	561.61	[442.12, 686.36]
	ϵ-ITM − ITM	48.02	[38.92, 57.99]	575.20	[444.71, 709.96]