THE PRICE OF HARD COVERAGE CONSTRAINTS IN ONLINE CONFORMAL PREDICTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Conformal prediction provides distribution-free uncertainty quantification, but in sequential deployment, standard coverage guarantees can still allow substantial fluctuations in the number of miscoverages over time. We study an online setting in which prediction sets are produced one round at a time and the cumulative number of miscoverages must stay within a prescribed budget throughout deployment. Such a hard coverage guarantee can be enforced by enlarging prediction sets, but how much additional prediction-set size is fundamentally necessary remains unclear. We characterize this unavoidable efficiency cost over T prediction rounds. Remarkably, the cost persists even with i.i.d. data, a known score distribution, and perfect knowledge of the population-optimal prediction set. Under suitable regularity conditions relating coverage and prediction-set size, we determine the minimum cumulative excess prediction-set size required by any online method satisfying the hard coverage constraint. This minimum cost is $\Theta(\log T)$ for expected cost at a fixed time known in advance, $\Theta(T^{1/3})$ for the largest expected cost over all fixed times up to T, and $\Theta(\sqrt{T})$ for the expected worst realized cost when the evaluation time is chosen in hindsight. We further show that these rates can be attained using past scores, without knowing the score distribution. Thus, even perfect statistical knowledge does not make sequential reliability free.

1 INTRODUCTION

Conformal prediction provides prediction sets with rigorous coverage guarantees while making minimal assumptions on the underlying predictive model (Shafer & Vovk, 2008; Angelopoulos & Bates, 2021). Its online variants have been motivated by sequential prediction problems where uncertainty estimates are repeatedly used over time, including epidemic forecasting and electricity-demand prediction (Gibbs & Candès, 2021; Angelopoulos et al., 2023; Areces et al., 2025). In such settings, reliability during deployment can matter beyond guarantees that hold only marginally or over a long horizon. Similar concerns arise more broadly in sequential decision making, where conservative online-learning methods require cumulative performance to remain above a prescribed baseline throughout operation (Wu et al., 2016). Consider an online deployment over rounds $t = 1, 2, \ldots$ with target miscoverage level $\alpha \in (0, 1)$. At round t, the method selects a prediction-set rule C_t using past observations, then observes the feature X_t and outputs $C_t(X_t)$ before observing the outcome Y_t. Even if each prediction has miscoverage probability exactly α, random fluctuations can cause the miscoverage count to temporarily exceed its nominal level αt. A natural response is to enlarge prediction sets when the miscoverage quota becomes tight, reducing the chance of another miscoverage at the cost of less informative predictions.

Document retrieval makes this tradeoff concrete. For each incoming query, a fixed retriever returns a shortlist intended to contain a target document in the collection (Intrator et al., 2024). A miscoverage occurs when that document is omitted. Larger shortlists reduce omissions but require more downstream processing. Returning the entire collection guarantees coverage, but eliminates the efficiency benefit of retrieval. Suppose the target document is revealed after each query and the cumulative fraction of omissions must remain at or below a prescribed limit throughout deployment. This *hard coverage constraint* must hold after every query, so good performance later cannot undo an earlier violation. Even a retriever with the correct miscoverage probability can violate it through random fluctuations. Preventing such violations can therefore require larger shortlists. The minimum cumulative excess prediction-set size quantifies how much retrieval efficiency must be sacrificed to

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.

enforce the constraint. This motivates our central question: *how much additional prediction-set size is fundamentally necessary to enforce hard coverage constraints throughout online deployment?*

A growing literature studies how online conformal methods adapt prediction sets using past miscoverages, scores, or distributional changes (Gibbs & Candès, 2021; 2024; Bhatnagar et al., 2023; Angelopoulos et al., 2023; Podkopaev et al., 2024; Angelopoulos et al., 2025; Areces et al., 2025; Wu et al., 2025). Recent work has also made prediction-set efficiency an explicit objective (Yang et al., 2024; Srinivas, 2026; Vaze, 2026). These methods provide sophisticated ways to learn, adapt, or control online coverage, but they do not characterize the minimum efficiency cost imposed solely by a constraint on cumulative miscoverage. To isolate this intrinsic cost, we first analyze an oracle stationary setting with i.i.d. data, a fixed nonconformity score, and complete knowledge of the data distribution. Let $M_t = \sum_{u=1}^{t} \mathbf{1}\{Y_u \notin C_u(X_u)\}$ denote the cumulative number of miscoverages by round t, and require $M_t \le \lfloor \alpha t \rfloor$ for every t. This hard constraint represents the zero-tolerance endpoint of a broader family of constraints permitting additional miscoverages. In the oracle analysis, the population coverage–efficiency frontier and its optimal prediction sets are known exactly, so estimation and distribution tracking are absent. Nevertheless, enforcing the realized coverage constraint still incurs a nonzero efficiency cost. For a horizon T, let $j(\alpha)$ denote the minimum expected prediction-set size at miscoverage α within a fixed family, allowing randomization. With μ measuring set size, define the constraint price as $\mathcal{R}_T = \mathbb{E}[\sum_{t=1}^{T} \mu(C_t(X_t))] - Tj(\alpha)$, the cumulative excess prediction-set size relative to the unconstrained population oracle.

Characterizing this minimum price raises three main challenges. First, an oracle with the correct miscoverage probability on every round can still exceed the miscoverage quota by chance. We must quantify the additional prediction-set size required to convert population-level coverage into pathwise control. Second, the cost depends critically on when efficiency is evaluated. A known evaluation horizon allows planning the remaining miscoverage quota. Controlling expected efficiency at every prefix requires preserving slack for future miscoverages, changing the optimal rate even when the final horizon is known. Our construction also works without knowing the horizon. Third, lowering the instantaneous miscoverage probability generally requires larger sets, with the enlargement governed by the local geometry of the population coverage–efficiency frontier. Thus, miscoverage control must be related to prediction-set size. Random fluctuations in realized cost create an additional difficulty when performance is evaluated retrospectively.

We connect control of cumulative miscoverage with the population coverage–efficiency frontier. An exact decomposition separates the cost associated with miscoverage slack, deviations from the frontier's supporting line, and size above the frontier. Conditioning independent Bernoulli(α) indicators on feasible paths yields an order-optimal controller at a fixed evaluation horizon. Barrier-based policies preserve slack while limiting conservatism without knowing the horizon. Martingale arguments characterize fluctuations at the worst realized prefix. Under the stated frontier conditions, these ideas yield matching upper and lower bounds for three temporal regimes. Randomized empirical ranks attain the corresponding optimal orders without knowing the score distribution under the stated regularity conditions. This separates the statistical cost of threshold estimation from the intrinsic cost of enforcing coverage. Our main contributions are summarized as follows.

- We formulate the *price of hard cumulative coverage* as the minimum cumulative excess prediction-set size, relative to the unconstrained population-optimal predictor, required solely to keep realized miscoverages within a running budget.
- We sharply characterize this price through matching upper and lower bounds, revealing three distinct temporal regimes: $\Theta(\log T)$ when performance is evaluated at a fixed time known in advance, $\Theta(T^{1/3})$ when efficiency is controlled at every fixed prefix in expectation, and $\Theta(\sqrt{T})$ when performance is evaluated at the worst realized prefix in hindsight. We establish these results using feasible-path conditioning, barrier control, and martingale arguments, and show that empirical rank-based procedures attain these orders under the same geometric conditions without knowing the score distribution.
- Experiments support the theoretical analysis of evaluation time and threshold estimation. Numerical prefix-optimal benchmarks and guarded baselines assess finite-horizon performance and horizon-specific tuning. Held-out regression experiments estimate learned-score curvature and empirical frontier gaps, with uncertainty estimates. Paired online experiments compare prediction-set size and compliance with the hard coverage constraint.

2 Related Work

Online conformal prediction and adaptation. A broad line of work studies conformal prediction in sequential and nonstationary environments, including adaptive quantile updates, strongly adaptive guarantees, control-based calibration, betting methods, decaying-step procedures, online optimization, and online-learning interpretations (Gibbs & Candès, 2021; Zaffran et al., 2022; Gibbs & Candès, 2024; Bhatnagar et al., 2023; Angelopoulos et al., 2023; Podkopaev et al., 2024; Angelopoulos et al., 2024a; 2025; Areces et al., 2025; Liu et al., 2026; Wu et al., 2025; Feldman et al., 2023). These methods develop mechanisms for adapting prediction sets to past miscoverages, scores, or distributional changes. Our oracle analysis deliberately removes this source of difficulty and instead isolates the efficiency cost induced solely by constraining the realized coverage path.

Coverage and prediction-set efficiency. Prediction-set size is a standard measure of conformal efficiency (Lei & Wasserman, 2014; Sadinle et al., 2019; Dhillon et al., 2024), and several recent works study coverage and efficiency jointly in online settings through stochastic control, competitive analysis, dynamic regret, quantile tracking, and limited-feedback selection (Yang et al., 2024; Srinivas, 2026; Vaze, 2026; Gollapudi et al., 2026). This literature establishes that coverage alone is insufficient and that the informativeness of the prediction set must also be accounted for. Our focus is different: we characterize the minimum efficiency loss that *any* policy must incur solely to enforce a hard miscoverage quota, even when the population-optimal prediction rule is already known.

Risk control, local validity, and time-uniform guarantees. Conformal prediction has been extended beyond marginal miscoverage to general risk control, nonexchangeable settings, multivalid and strongly adaptive guarantees, and validity that holds uniformly over time or under adaptive stopping (Bates et al., 2021; Angelopoulos et al., 2024b; Barber et al., 2023; Bastani et al., 2022; Bhatnagar et al., 2023; Scharfstein & Kuchibhotla, 2026; Hultberg et al., 2026; Dubey & Huo, 2026). Related statistical tools include time-uniform confidence sequences and sequential quantile estimation (Howard et al., 2021; Howard & Ramdas, 2022), while strongly adaptive and anytime notions are also central in online learning (Daniely et al., 2015; Harvey et al., 2023). These works strengthen the form or temporal scope of validity. In contrast, we impose a pathwise miscoverage quota and study the minimum prediction-set-size price required to enforce it.

Online learning under running constraints. Our work is related more broadly to safe and conservative sequential learning under performance or resource constraints (Mannor et al., 2009; Wu et al., 2016; Kazerouni et al., 2017; Garcelon et al., 2020; Du et al., 2021; Bernasconi et al., 2022; McMahan & Zhu, 2024; Sridharan & Yoo, 2025). This literature studies sample-path constraints, conservative exploration, anytime feasibility, and learning under unknown constraints. We share this fundamental-limits perspective, but our constraint acts on realized miscoverage and its cost is measured by excess prediction-set size. The frontier couples these quantities even without a learning problem. Conservative bandits protect baseline reward, while anytime-constrained MDPs and replenishment models enforce running resource budgets (Chen et al., 2026). Classical queueing control gives an inverse-square excess-cost–buffer relation (Berry & Gallager, 2002). Our finite-horizon comparison separates the population size comparator, the information available when a map is chosen, and the three temporal criteria (Appendix G). The disjoint-window lower bound covers arbitrary history-dependent policies, and empirical ranks implement the matching constructions.

3 Notation and Problem Setup

Prediction-set family and online protocol. Let $\mathcal{X}$ and $\mathcal{Y}$ denote the feature and outcome spaces, and let $\mathcal{C}$ be the collection of prediction-set rules available to the calibrator. Each $C \in \mathcal{C}$ is a measurable map $C : \mathcal{X} \to 2^{\mathcal{Y}}$, so that $C(x)$ is the prediction set issued at feature x. A common special case is a fixed nonconformity score s, with threshold sets $C_\tau(x) = \{y : s(x, y) \leq \tau\}$ (Angelopoulos & Bates, 2021). Let μ measure set size and define $\ell_C(x) := \mu(C(x))$. We assume $B := \mu(\mathcal{Y}) < \infty$ and $\ell_C(x) \in [0, B]$. The family $\mathcal{C}$ contains the full-set and empty-set maps.

Assumption 1 (Online efficiency protocol)**.** *The pairs (X_t, Y_t) are i.i.d. from a distribution D on $\mathcal{X} \times \mathcal{Y}$. The family $\mathcal{C}$ is fixed independently of these observations, possibly using independent training data. At round t, the calibrator selects $C_t \in \mathcal{C}$ from past observations and fresh independent randomization, then observes X_t, issues $C_t(X_t)$, and finally observes Y_t.*

We condition on any independent training used to construct $\mathcal{C}$. Covariate-dependent thresholding is included by treating the entire rule $x \mapsto C(x)$ as a map selected before the fresh covariate is observed.

Hard cumulative coverage. Fix a target miscoverage level $\alpha \in (0,1)$. Let $E_t = \mathbf{1}\{Y_t \notin C_t(X_t)\}$ be the miscoverage indicator, $M_t = \sum_{u \le t} E_u$ the cumulative number of miscoverages, and $W_t = \sum_{u \le t} \ell_{C_u}(X_u)$ the cumulative prediction-set size.

Definition 1 (Hard cumulative coverage)**.** A calibrator is feasible if $M_t \le b_t := \lfloor \alpha t \rfloor$ at every round t in its specified horizon, for every data stream and every realization of its internal randomization.

This pathwise requirement is independent of the i.i.d. assumption used to evaluate efficiency. It controls cumulative miscoverage at every prefix, but does not imply $M_t/t \to \alpha$, conditional coverage, or rolling-window coverage. At $\alpha = 0.1$, the first nine outputs must be full on every stream. Finite B is therefore essential for finite cost.

Population coverage–efficiency frontier. For a fixed rule $C \in \mathcal{C}$, let $m_D(C) := \Pr_{(X,Y)\sim D}\{Y \notin C(X)\}$ denote its population miscoverage probability and $V_D(C) := \mathbb{E}_{(X,Y)\sim D}[\mu(C(X))]$ its expected prediction-set size. To allow randomization, let $\Delta(\mathcal{C})$ denote the set of probability distributions over $\mathcal{C}$. For $\pi \in \Delta(\mathcal{C})$, define $m_D(\pi) := \mathbb{E}_{C\sim\pi}[m_D(C)]$ and $V_D(\pi) := \mathbb{E}_{C\sim\pi}[V_D(C)]$.

The population coverage–efficiency frontier is $j(q) = \inf_{\pi \in \Delta(\mathcal{C}):\, m_D(\pi) = q} V_D(\pi)$, $q \in [0,1]$. Thus, $j(q)$ is the minimum expected size within $\mathcal{C}$ at miscoverage q, allowing randomization. Taking $\mathcal{C}$ to be a fixed score's thresholds gives its score-restricted frontier.

Constraint price. Let $\mathcal{A}$ denote an online calibrator following the protocol above. Over a horizon of T rounds, define $\mathcal{R}_T(\mathcal{A}, D) := \mathbb{E}_D W_T - T j(\alpha)$ and $\mathcal{R}_T^\star(D) := \inf_{\mathcal{A}\ \text{feasible}} \mathcal{R}_T(\mathcal{A}, D)$. Here, $\mathcal{R}_T(\mathcal{A}, D)$ is the expected cumulative excess prediction-set size of calibrator $\mathcal{A}$ relative to the unconstrained population benchmark $Tj(\alpha)$, while $\mathcal{R}_T^\star(D)$ is the smallest such excess achievable by any feasible online calibrator. Hence $\mathcal{R}_T^\star(D)$ is the fundamental price of the hard coverage constraint. Policies in the infimum may depend on D and T, but not on unseen outcomes. Allowing randomized rules makes j convex and nonincreasing. See Appendix A.

Coverage balance. Define the miscoverage slack $G_t = \alpha t - M_t$. Since the balance increment $g_t = \alpha - E_t$ (the pinball-loss subgradient at finite thresholds) satisfies $\sum_{t \le T} g_t = G_T$, gradient equilibrium $G_T/T \to 0$ is equivalent to asymptotic coverage balance $M_T/T \to \alpha$ (Angelopoulos et al., 2025). It does not characterize the prediction-set-size cost of enforcing the pathwise constraint.

4 Algorithms and Theoretical Guarantees

We abbreviate $\mathcal{R}_T(\mathcal{A}, D)$ as $\mathcal{R}_T$. Counting controls terminal cost, while Barrier controls expected-prefix cost. Empirical ranks implement both without knowing the score distribution.

4.1 Excess set-size decomposition

Let V_t be the independent random seed at round t and let $\mathcal{H}_t = \sigma(X_u, Y_u, V_u : 1 \le u \le t)$ denote the observed history. The conditional miscoverage probability is $p_t := \mathbb{P}(E_t = 1 \mid \mathcal{H}_{t-1}) = \mathbb{E}[m_D(C_t) \mid \mathcal{H}_{t-1}]$. This averages over the current seed and fresh observation. Since the conditional law of C_t is a mixture allowed in the definition of j, $\epsilon_t := \mathbb{E}[V_D(C_t) \mid \mathcal{H}_{t-1}] - j(p_t) \ge 0$ measures how much its expected prediction-set size exceeds the frontier value $j(p_t)$. Even if each selected rule attains the frontier at its own miscoverage, randomizing between different miscoverage levels can give $\epsilon_t > 0$.

Choose $\lambda \ge 0$ such that $j(q) \ge j(\alpha) - \lambda(q - \alpha)$ for all $q \in [0,1]$. Such a λ exists because j is convex and nonincreasing. If j is differentiable at α, then $\lambda = -j'(\alpha)$. Define $\psi(q) := j(q) - j(\alpha) + \lambda(q - \alpha)$ for $q \in [0,1]$. Thus $\psi(q) \ge 0$ is the frontier's gap above this supporting line, and $\psi(\alpha) = 0$.

Theorem 1 (Excess set-size decomposition)**.** *Under Assumption 1, for every policy $\mathcal{A}$ and integer $T \ge 1$, $\mathcal{R}_T(\mathcal{A}, D) = \lambda \mathbb{E} G_T + \mathbb{E}\sum_{t=1}^T \psi(p_t) + \mathbb{E}\sum_{t=1}^T \epsilon_t$.*

Remark 1. The first term measures the cost of expected miscoverage slack, $\mathbb{E}G_T = \alpha T - \mathbb{E}M_T$. The second measures the frontier's gap above the supporting line at p_t. The third measures size

Shared protocol. Choose Counting for $t = 1, \ldots, T$ or Barrier for $t = 1, 2, \ldots$. Start with $M_0 = G_0 = 0$ and no past scores. Write $b_t = \lfloor \alpha t \rfloor$, $S_t = s(X_t, Y_t)$, and $C_\tau(x) = \{y : s(x, y) \le \tau\}$. Each round, pass the controller's q_t to Empirical ranks and update $M_t = M_{t-1} + E_t$ and $G_t = G_{t-1} + \alpha - E_t$.

Algorithm 1 Counting	**Algorithm 2** Barrier	**Algorithm 3** Empirical ranks
Parameters: miscoverage level $\alpha \in (0, 1)$, horizon $T \in \mathbb{N}_{\ge 1}$. **Round input:** $t \le T$, cumulative miscoverage count M_{t-1}. **Output:** requested miscoverage probability $q_t \in [0, 1]$.	**Parameter:** miscoverage level $\alpha \in (0, 1)$. No horizon is required. **Round input:** $t \ge 1$, miscoverage slack G_{t-1}. **Output:** requested miscoverage probability $q_t \in [0, 1]$.	**Parameter:** fixed nonconformity score s. **Round input:** t, q_t, sorted scores $S_{(1)} \le \cdots \le S_{(t-1)}$. **Output:** prediction set $C_t(X_t)$, miscoverage indicator E_t, updated scores.
1. Compute counts once. Set all $H_t(m) = 0$, then $H_T(b_T) = 1$. For $t = T, \ldots, 1$ and $m = 0, \ldots, b_{t-1}$, set $H_{t-1}(m) = H_t(m) + H_t(m+1)$. **2. Read counts.** Set $w_0 = H_t(M_{t-1})$, $w_1 = H_t(M_{t-1} + 1)$. **3. Return** $q_t = w_1/(w_0 + w_1)$.	**1. Set the barrier.** Let $L_t = 4 + \lceil t^{1/3} \rceil$. Set $h_t(z) = z(L_t - z)$ for $0 < z < L_t$, and 0 otherwise. **2. Compute weights.** $w_0 = (1 - \alpha) h_t(G_{t-1} + \alpha)$, $w_1 = \alpha h_t(G_{t-1} + \alpha - 1)$. **3. Return** $q_t = w_1/(w_0 + w_1)$.	**1.** Set $r_t = \lfloor t(1 - q_t) \rfloor$ and $\theta_t = t(1 - q_t) - r_t$. **2.** Draw an independent $V_t \sim \mathrm{Unif}[0, 1)$ and set $J_t = r_t + \mathbf{1}\{V_t < \theta_t\}$. **3.** Set C_t to the empty-set map if $J_t = 0$, the full-set map if $J_t = t$, and otherwise C_{τ_t} with threshold $\tau_t = S_{(J_t)}$. **4.** Observe X_t and issue $C_t(X_t)$. Observe Y_t, store $S_t = s(X_t, Y_t)$, and return $E_t = \mathbf{1}\{Y_t \notin C_t(X_t)\}$.

above the frontier. All three are nonnegative for a feasible policy, so small slack alone does not ensure a small constraint price. The full proof is in Appendix A.

Proof sketch. Substitute $j(p_t) = j(\alpha) - \lambda(p_t - \alpha) + \psi(p_t)$ into conditional expected size $j(p_t) + \epsilon_t$ (Lemma 1). Sum over rounds and use $\mathbb{E} \sum_{t \le T} p_t = \mathbb{E} M_T$ to obtain $\lambda \mathbb{E} G_T$.

Example. For $Y = Q(U)$ with $U \sim \mathrm{Unif}[0, 1]$ and $Q(u) = (3u + u^2)/4$, intervals $[0, \tau]$ attain the length frontier among all measurable sets: $j(q) = 1 - 5q/4 + q^2/4$ (Appendix E). Therefore $\lambda = 5/4 - \alpha/2$ and $\psi(q) = (q - \alpha)^2/4$: the second term in Theorem 1 measures squared deviations of conditional miscoverage from α.

4.2 Calibration algorithms

Fix a nonconformity score s with $C_\tau \in \mathcal{C}$ for every threshold, and let $F(u) = \mathbb{P}_D\{s(X, Y) \le u\}$. Algorithms 1 and 2 request a miscoverage probability $q_t \in [0, 1]$ using past miscoverage indicators. For known continuous F and $0 < q_t < 1$, use the threshold $F^{-1}(1 - q_t)$, where $F^{-1}(u) = \inf\{z : F(z) \ge u\}$. Requests $q_t = 0, 1$ use the full-set and empty-set maps, respectively. When F is unknown, Algorithm 3 selects thresholds from past scores (Section 4.3). Smaller q_t favors larger sets.

Fixed horizon. Let $\mathcal{A}_T$ be the set of binary miscoverage paths satisfying every quota and ending with $M_T = b_T$. After t rounds and m miscoverages, $H_t(m)$ counts the feasible ways to complete the remaining rounds. At round t, Counting sets q_t to the fraction of feasible completions with $E_t = 1$. If only one indicator allows completion, a full-set or empty-set map forces it. Under continuous i.i.d. scores, each path in $\mathcal{A}_T$ is equally likely. This is $P_\alpha(\cdot \mid \mathcal{A}_T)$, where P_α is the law of T independent Bernoulli(α) indicators. Write $\mathrm{kl}(q \| \alpha) = q \log(q/\alpha) + (1 - q) \log((1 - q)/(1 - \alpha))$ for Bernoulli relative entropy, with $0 \log 0 = 0$ and natural logarithms. Define $\mathcal{K}_T := \mathbb{E} \sum_{t \le T} \mathrm{kl}(q_t \| \alpha)$. For Counting, the chain rule gives $\mathcal{K}_T = \log[1/P_\alpha(\mathcal{A}_T)] = O(\log T)$. Since $\mathrm{kl}(q \| \alpha) \ge 2(q - \alpha)^2$, $\mathbb{E} \sum_{t \le T} (q_t - \alpha)^2 \le \mathcal{K}_T/2$. The count table uses $O(T + \alpha T^2)$ arithmetic operations and stored counts (Appendix B.1).

Unknown horizon. Algorithm 2 weights the two possible next miscoverage slack values by h_t. Near zero it favors larger sets to prevent a quota violation. Near L_t it favors smaller sets to avoid

accumulating unnecessary slack. A zero weight triggers a full-set or empty-set map. This keeps $0 < G_t < L_t$ on every stream, so λL_t bounds the slack contribution to $\mathcal{R}_t$. The potential comparison in Appendix B.2 gives $\mathbb{E}\sum_{t\le T}(q_t-\alpha)^2 = O(1+\log L_T + \sum_{t\le T} L_t^{-2})$. Choosing $L_t \asymp t^{1/3}$ balances the slack bound L_T and $\sum_{t\le T} L_t^{-2}$, both of order $T^{1/3}$. Counting reaches $M_T = b_T$ at its specified final round. Barrier instead keeps slack for future fluctuations and can continue without restarting.

Write $R_t = W_t - tj(\alpha)$ for realized cumulative excess prediction-set size. Then $\mathcal{R}_t = \mathbb{E}_D R_t$ is the constraint price defined in Section 3. Expected-prefix cost first averages R_t at each fixed time, then takes the largest value. Retrospective cost first takes the largest R_t within each run, then averages across runs. Adaptive stopping uses only observations available when stopping: for example, stop when M_t first reaches a prescribed level, or at T.

4.3 Empirical-Rank Implementation and Upper Bounds

Assumption 2 (Quadratic bound on excess set size)**.** *There exists $0 \le C < \infty$ such that every threshold map C_τ, including the full-set and empty-set maps, satisfies $0 \le V_D(C_\tau) - j(\alpha) + \lambda(m_D(C_\tau)-\alpha) \le C(m_D(C_\tau)-\alpha)^2$.*

Target efficiency and local smoothness. The assumption combines efficiency at the target (Lei et al., 2013; Sadinle et al., 2019) with a quadratic Taylor bound (Boyd & Vandenberghe, 2004, Section 9.1.2). When expected prediction-set size depends only on miscoverage, write $h(q) = V_D(C_\tau)$ for $q = m_D(C_\tau)$. At $q = \alpha$, the assumption requires $h(\alpha) = j(\alpha)$. Even for a threshold family $\mathcal{C}$, j permits mixtures and can lie below h. Smoothness or positive estimated curvature alone does not establish this equality. Conversely, this equality and the bound $|h'(q) - h'(q')| \le L|q-q'|$ for $q, q' \in [\alpha-\delta, \alpha+\delta] \subset (0,1)$, with $L \ge 0$ and $\delta > 0$, give $h'(\alpha) = -\lambda$ and a local quadratic Taylor bound. Bounded size extends it globally with $C = \max\{L/2, (B+\lambda)/\delta^2\}$ (Appendix E). The example in Section 4.1 has $C = 1/4$. A uniform additive gap $\Delta \ge 0$ in the assumption adds $T\Delta$ to the bound in Theorem 2. A gap measured only at α is insufficient.

Covariate-dependent example. Let X be uniform on $\{0,1\}$, with densities $f_0(y) = 3/2 - y$ and $f_1(y) = 3/2 - y + \sin(2\pi y)/(4\pi)$ on $[0,1]$. The score $s(x,y) = 3/2 - f_x(y)$ gives different interval lengths across covariates and attains j among all measurable randomized maps. Lemma 7 proves $4/27 \le j'' \le 12$, verifying quadratic upper and lower growth. Estimated scores require a separate efficiency check. A uniform outcome has $j(q) = 1 - q$ and all three prices below one (Appendix E.4).

Algorithm 3 estimates thresholds from past scores and stores $O(t)$ scores. Barrier itself stores only G_t and t. Define the insertion rank of S_t among the first t scores by $I_t = 1 + \sum_{u<t} \mathbf{1}\{S_u < S_t\}$. Write $\mathcal{G}_{t-1} = \sigma(I_u, V_u : u < t)$, where $V_t \sim \mathrm{Unif}[0,1)$ are independent rounding seeds, independent of the scores. The request q_t uses this history, while the threshold also uses the stored scores. Under continuous i.i.d. scores, the fresh score has t equally likely ranks. Selecting $J_t = j$ covers j of them. Randomizing between adjacent ranks gives $\mathbb{E}[J_t \mid \mathcal{G}_{t-1}] = t(1-q_t)$, avoiding systematic rounding toward larger sets.

Theorem 2 (Empirical-rank implementation)**.** *Under Assumptions 1–2 with continuous fixed-score cdf F, let $q_t \in [0,1]$ be $\mathcal{G}_{t-1}$-measurable for each $t \ge 1$ and implement it by Algorithm 3. Then $\mathbb{P}(E_t = 1 \mid \mathcal{G}_{t-1}) = \mathbb{E}[m_D(C_t) \mid \mathcal{G}_{t-1}] = q_t$ and $\mathrm{Var}(m_D(C_t) \mid \mathcal{G}_{t-1}) \le 1/(4t)$. For every integer $T \ge 1$, with $\mathsf{h}_T = \sum_{t=1}^T 1/t$, $\mathcal{R}_T(\mathcal{A}, D) \le \lambda \mathbb{E} G_T + C\{\mathbb{E}\sum_{t=1}^T (q_t-\alpha)^2 + \mathsf{h}_T/4\}$.*

Remark 2. The upper bound includes at most $C\mathsf{h}_T/4 = O(\log T)$ for threshold estimation. Here q_t uses rank history, whereas p_t uses the full observed history and averages over the current random seed. They need not agree with each other or with the selected map's miscoverage $m_D(C_t)$. For example, $q_2 = 1/2$ selects threshold S_1 without rounding, so $p_2 = m_D(C_2) = 1 - F(S_1)$. Averaging over S_1 gives the requested $1/2$. Counting and Barrier remain feasible with tied scores. The exact rank law and variance bound require continuity. The full proof is in Appendix B.

Proof sketch. Lemma 2 uses uniform insertion ranks (Shafer & Vovk, 2008) and order-statistic moments to give the requested miscoverage law and conditional variance bound. The conditional mean squared deviation of $m_D(C_t)$ from α is $(q_t-\alpha)^2 + \mathrm{Var}(m_D(C_t) \mid \mathcal{G}_{t-1})$. Apply Assumption 2 and sum: the linear terms yield $\lambda \mathbb{E} G_T$, and the conditional variances contribute at most $C\sum_{t\le T} 1/(4t)$.

Table 1: Optimal prices (Theorem 4 and, for adaptive stopping, Appendix D). All constructions use Algorithm 3. A stopping time $\tau \in \{1, \dots, T\}$ satisfies $\{\tau \le t\} \in \mathcal{H}_t$ for every t. The deterministic comparator $tj(\alpha)$ allows oracle set-size fluctuations in the retrospective cost.

Criterion minimized over feasible policies	Construction	Optimal price
(i) Terminal: $\mathcal{R}_T$	Counting (Algorithm 1)	$\Theta(\log T)$
(ii) Expected prefixes: $\max_{t\le T} \mathcal{R}_t$	Barrier (Algorithm 2)	$\Theta(T^{1/3})$
(ii′) Adaptive stopping: $\sup_\tau \mathbb{E}_D R_\tau$	Barrier (Algorithm 2)	$\Theta(T^{1/3})$
(iii) Retrospective: $\mathbb{E}\max_{t\le T} R_t$	Barrier (Algorithm 2)	$\Theta(\sqrt{T})$

Controller upper bounds. Combining Theorem 2 with Lemmas 3 and 4 gives $\mathcal{R}_T = O(\log T)$ for Counting and $\max_{t\le T} \mathcal{R}_t = O(T^{1/3})$ for one infinite run of Barrier. Applying the martingale maximal inequality to (3) gives $\mathbb{E}\max_{t\le T} R_t = O(\sqrt{T})$ for the same run (Appendix B.3). These upper bounds require neither quadratic lower growth nor knowledge of D, F, j, λ, C.

4.4 Lower bounds

Theorem 3 (Lower bounds)**.** *Suppose Assumption 1 holds and $\psi(q) \ge c_-(q-\alpha)^2$ for all $q \in [0,1]$, with $c_- > 0$. There exist $c_0 \ge 0$ and $c_1, c_2, c_3 > 0$, depending only on α, B, λ, c_-, such that every integer $T \ge 1$ and every policy $\mathcal{A}$ feasible through T satisfy (i) $\mathcal{R}_T \ge c_1 \log T - c_0$, (ii) $\max_{1\le t\le T} \mathcal{R}_t \ge c_2 T^{1/3} - 1$, and (iii) $\mathbb{E}_D \max_{1\le t\le T} R_t \ge c_3\sqrt{T}$. For (iii), $\psi(1) > 0$ suffices in place of quadratic lower growth, with c_3 depending only on $\alpha, B, \lambda, \psi(1)$.*

Remark 3. The three rates are unavoidable even with knowledge of D, T, without score continuity or Assumption 2. Together with the upper bounds, they establish the optimal orders in Table 1. For the continuous example, bound (ii) does not certify finite-horizon near-optimality at the horizons considered here. The full proof and constants are in Appendix C.

Proof sketch. Compare with independent Bernoulli(α) miscoverage indicators. Lemma 5 shows that all quotas hold with probability $O(T^{-1/2})$, whereas a feasible policy assigns this event probability one. Changing the miscoverage law therefore costs logarithmic relative entropy, which the quadratic lower bound on ψ converts into excess prediction-set size. For (ii), generate both miscoverage sequences using the same uniform random variables. Lemma 6 shows that fluctuations in each window require starting slack or reduced miscoverage probabilities. Summing over disjoint windows gives the $T^{1/3}$ bound for policies using the full history. For (iii), $\psi(1) > 0$ forces square-root fluctuations in (3), unless expected terminal cost already has that order.

4.5 Optimal rates

In each infimum, policies satisfy Definition 1 through T and may know D, T.

Theorem 4 (Optimal rates)**.** *Fix D, $\alpha \in (0,1)$, and a score with continuous cdf F. Suppose Assumptions 1–2 hold and $\psi(q) \ge c_-(q-\alpha)^2$ for every $q \in [0,1]$, with $c_- > 0$. As $T \to \infty$, the optimal prices are (i) $\inf_{\mathcal{A}\text{ feasible}} \mathcal{R}_T(\mathcal{A}, D) = \Theta(\log T)$, (ii) $\inf_{\mathcal{A}\text{ feasible}} \max_{1\le t\le T} \mathcal{R}_t(\mathcal{A}, D) = \Theta(T^{1/3})$, and (iii) $\inf_{\mathcal{A}\text{ feasible}} \mathbb{E}_D \max_{1\le t\le T} R_t = \Theta(\sqrt{T})$. The constants depend only on $\alpha, B, \lambda, C, c_-$. Algorithm 1 attains (i). One infinite run of Algorithm 2 attains (ii)–(iii). With Algorithm 3, both satisfy every quota on every stream, and Algorithm 1 has $M_T = b_T$.*

Remark 4. At a bounded stopping time, the martingale term in the realized-cost decomposition has zero expectation. The slack and accumulated nonnegative cost then give the same $T^{1/3}$ bound (Proposition 1 and Theorem 3). The full proof is in Appendix D. Appendix H extends the expected-prefix result to a fixed initial miscoverage allowance under the same assumptions.

Proof sketch. The upper bounds in Section 4.3 and the uniform lower bounds in Theorem 3 prove the three optimal orders. Table 1 summarizes the criteria and attaining policies.

5 Experiments

We examine evaluation time, threshold estimation, and coverage control in Theorems 1–4. The continuous model in Section 4.1 satisfies the quadratic conditions. We use California Housing to

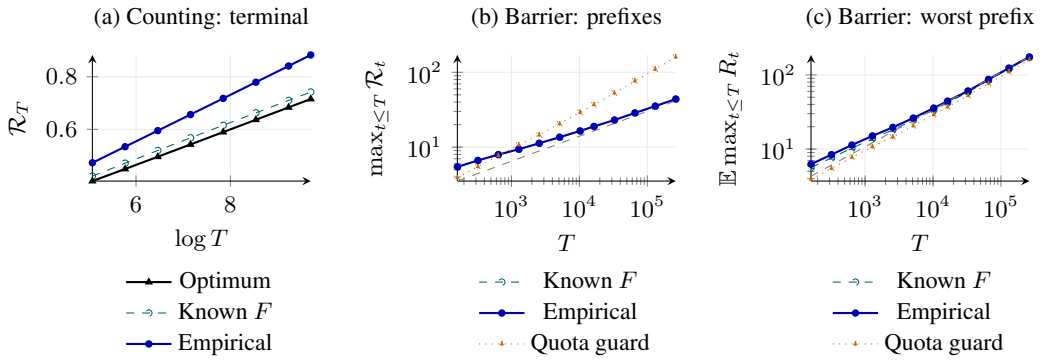


Figure 1: Three criteria for excess prediction-set size in the continuous model ($\alpha = 0.1$). (a) Counting's terminal cost versus the numerical optimum under all prefix quotas with known F, T. (b) Largest expected cost over prefixes. (c) Average of each run's largest realized cost. Known F uses population quantiles. Empirical uses Algorithm 3. Panels (b)–(c) compare default Barrier with the quota guard. Panel (a) and Barrier in (b) use numerical expectations. Other curves use 512 streams with pointwise ± 1.96 standard-error bars. Gray guides show $T^{1/3}$ in (b) and $\sqrt{T}$ in (c), with arbitrary vertical scales.

evaluate learned scores. Unless stated otherwise, *Counting* and *Barrier* mean Algorithms 1 and 2 with default settings. *Empirical* thresholds use Algorithm 3, while *known* F uses population quantiles. This comparison keeps the controller fixed. Main figures use $\alpha = 0.1$. Other targets and variants are in Appendix F.

5.1 Evaluation time changes the cost

In Figure 1a, Counting is rerun for each announced horizon. It stays close to the numerical terminal optimum, but its low terminal cost conceals much higher costs at earlier prefixes (Figure 3a). Terminal efficiency need not persist throughout a run (Corollary 1).

Panels (b)–(c) use one horizon-free Barrier controller, with $L_t = 4 + \lceil t^{1/3} \rceil$. Panel (b) takes the largest expected cost over all prefixes, computed from the controller's state probabilities. Panel (c) instead averages the largest realized excess size $R_t = W_t - tj(\alpha)$ on each of 512 independent score streams. The quota guard proposes a population threshold with miscoverage α and switches to the full set whenever another miscoverage could exceed b_t. At $T = 262{,}144$, the quota guard has higher expected-prefix cost than Barrier but lower estimated retrospective cost. Normalized curves still drift at these horizons (Appendix F.2).

5.2 The cost of estimating thresholds

The known-F and empirical curves have the same miscoverage-count law and expected slack (Theorem 2). In this quadratic model, their expected terminal costs differ by one quarter of the accumulated order-statistic variance, at most $\mathrm{h}_T/16$. At $T = 16{,}380$, estimating thresholds adds 0.141 for Counting and 0.144 for Barrier, below the bound 0.643. These numerical expectations show a small estimation cost. The near overlap of the curves does not reflect Monte Carlo uncertainty. The expected-size bound does not cover differences between retrospective maxima (Appendix F.1).

5.3 Finite-horizon control and baselines

Optimal asymptotic rates need not give the smallest finite-horizon cost. At $T = 16{,}380$, the numerical expected-prefix optimum lies in $[5.6547, 5.6598]$, whereas the default Barrier costs 18.926, about 3.34 times the feasible upper bound in excess cost (Figure 5a). Theorem 1 explains most of this cost as unused quota: its slack contribution is 18.043. The near-optimal policy builds slack early and uses it near the final round. A separately validated width variant costs 11.484. Its known-F counterpart still reserves slack for continued operation (Figure 6). The remaining gap cannot be attributed to threshold estimation alone.

Guarded adaptive conformal inference (ACI) issues the full set when a miscoverage would violate b_t. Its internal update uses the original proposal's miscoverage indicator. Figure 5b compares horizon-selected and decaying steps with the two Barrier widths. At the largest horizon, both ACI schedules have lower cost than the default Barrier, while the width variant has lower cost than decaying ACI. The finite-horizon ranking therefore depends on the settings. Horizon-selected steps use independent validation for each horizon. The decaying multiplier and Barrier width are each selected once and frozen across test horizons (Appendix F.4).

5.4 Learned scores and online performance

California Housing uses five frozen splits, each with separate training, validation, and test observations, and two predictors: Ridge and Extra Trees. Normalized responses lie in $[0, 1]$, giving $B = 1$ for interval length. The size curves in Figure 2a bend upward near the target, but mixing thresholds can reduce size at the same miscoverage. At $\alpha = 0.1$, empirical gaps above the convex envelope range from 0 to 0.00293 for Ridge and 0.000033 to 0.00183 for Extra Trees (Table 4). These diagnostics establish neither target efficiency nor global quadratic lower growth, and cannot replace the uniform additive gap Δ that adds $T\Delta$ to the upper bound (Appendices E and F.5).

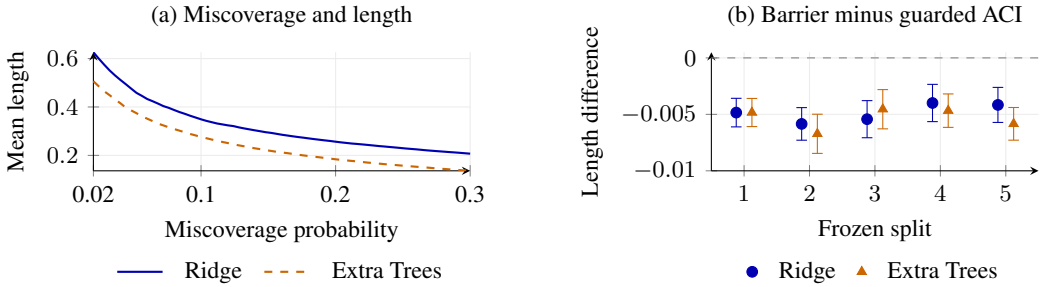


Figure 2: California Housing with frozen predictors. (a) Empirical curves averaged over five test splits. (b) Default Barrier (Algorithms 2 and 3) minus original guarded ACI at $\alpha = 0.1$. Negative values mean shorter intervals for Barrier. Each split uses 32 paired test permutations and every test row once per run. Bars are ± 1.96 ordering SEs, excluding training and population uncertainty.

Controller settings are selected using 16 validation permutations and frozen before the 32 paired test permutations. Each run starts with no past scores and uses every test observation once. At $\alpha = 0.1$, all splits select the default Barrier width. Figure 2b compares it with original guarded ACI. Barrier reduces mean interval length by 1.33% for Ridge and 1.83% for Extra Trees. With the same empirical-rank implementation for ACI, these reductions are 1.68% and 2.21%. Table 5 reports absolute lengths, paired differences, miscoverage, and full/empty frequencies. Every tested guarded prefix satisfies $M_t \leq b_t$, and empty outputs remain rare. Uncertainty intervals quantify ordering and controller randomness on fixed test sets, excluding training and population uncertainty. Appendix F.5 isolates the guard's added length using paired ACI runs with identical proposals.

6 Discussion

We characterize the efficiency cost of enforcing hard cumulative coverage in online conformal prediction. Under suitable geometric conditions, the optimal excess prediction-set size scales as $\Theta(\log T)$, $\Theta(T^{1/3})$, or $\Theta(\sqrt{T})$, depending on when performance is evaluated. These rates remain achievable using empirical ranks without knowing the score distribution.

Limitations and future work. Our efficiency guarantees rely on a stationary frontier and uniform insertion ranks, assumptions that may fail under distribution shift or when data are reused for fitting and calibration. The application examples motivate the tradeoff, but their feedback availability and frontier conditions need validation. Extending the theory to drifting distributions, alternative comparators, calibration error, and more realistic deployment constraints is an important direction for future work.

AI USE STATEMENT

The core research ideas and initial proof sketches were developed by the authors. Generative AI tools assisted in developing these sketches into detailed proofs, refining mathematical arguments, designing and implementing experiments, checking numerical results, and revising the manuscript. The authors are responsible for all final claims, proofs, code, and experimental results.

ETHICS STATEMENT

The experiments use public benchmark data and do not evaluate suitability for a particular deployment.

REPRODUCIBILITY STATEMENT

The appendices give complete proofs of the main results and the experimental protocols. The anonymous code package includes experiment scripts, fixed score pools, and implementation checks. Running `python run.py` regenerates the figures, tables, and validation outputs. Reported stream uncertainty is conditional on the fitted scores and evaluation pools.

REFERENCES

Anastasios Angelopoulos, Emmanuel Candes, and Ryan J. Tibshirani. Conformal PID control for time series prediction. In *Advances in Neural Information Processing Systems*, volume 36, pp. 23047–23074, 2023. doi: 10.52202/075280-1000. URL https://proceedings.neurips.cc/paper_files/paper/2023/file/47f2fad8c1111d07f83c91be7870f8db-Paper-Conference.pdf.

Anastasios N. Angelopoulos and Stephen Bates. A gentle introduction to conformal prediction and distribution-free uncertainty quantification, 2021. URL https://arxiv.org/abs/2107.07511. Preprint.

Anastasios N. Angelopoulos, Rina Foygel Barber, and Stephen Bates. Online conformal prediction with decaying step sizes. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pp. 1616–1630, 2024a. URL https://proceedings.mlr.press/v235/angelopoulos24a.html.

Anastasios N. Angelopoulos, Stephen Bates, Adam Fisch, Lihua Lei, and Tal Schuster. Conformal risk control. In *The Twelfth International Conference on Learning Representations*, 2024b. URL https://openreview.net/forum?id=33XGfHLtZg.

Anastasios N. Angelopoulos, Michael I. Jordan, and Ryan J. Tibshirani. Gradient equilibrium in online learning: Theory and applications. *Journal of Machine Learning Research*, 26(305):1–68, 2025. URL https://jmlr.org/papers/v26/25-0356.html.

Felipe Areces, Christopher Mohri, Tatsunori Hashimoto, and John Duchi. Online conformal prediction via online optimization. In *International Conference on Machine Learning*, volume 267 of *PMLR*, pp. 1604–1649, 2025. URL https://proceedings.mlr.press/v267/areces25a.html.

Rina Foygel Barber, Emmanuel J. Candès, Aaditya Ramdas, and Ryan J. Tibshirani. Conformal prediction beyond exchangeability. *The Annals of Statistics*, 51(2):816–845, 2023. doi: 10.1214/23-AOS2276. URL https://arxiv.org/abs/2202.13415.

Osbert Bastani, Varun Gupta, Christopher Jung, Georgy Noarov, Ramya Ramalingam, and Aaron Roth. Practical adversarial multivalid conformal prediction. In *Advances in Neural Information Processing Systems*, volume 35, pp. 29362–29373, 2022. doi: 10.52202/068431-2129. URL https://proceedings.neurips.cc/paper_files/paper/2022/file/bcdaaa1aec3ae2aa39542acefdec4e4b-Paper-Conference.pdf.

Stephen Bates, Anastasios Angelopoulos, Lihua Lei, Jitendra Malik, and Michael I. Jordan. Distribution-free, risk-controlling prediction sets. *Journal of the ACM*, 68(6):1–34, 2021. doi: 10.1145/3478535. URL https://arxiv.org/abs/2101.02703.

Martino Bernasconi, Federico Cacciamani, Matteo Castiglioni, Alberto Marchesi, Nicola Gatti, and Francesco Trovò. Safe learning in tree-form sequential decision making: Handling hard and soft constraints. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 1854–1873, 2022. URL https://proceedings.mlr.press/v162/bernasconi22a.html.

Randall A. Berry and Robert G. Gallager. Communication over fading channels with delay constraints. *IEEE Transactions on Information Theory*, 48(5):1135–1149, 2002. URL https://users.ece.northwestern.edu/~rberry/Pubs/delay.pdf.

Aadyot Bhatnagar, Huan Wang, Caiming Xiong, and Yu Bai. Improved online conformal prediction via strongly adaptive online learning. In *International Conference on Machine Learning*, volume 202 of *PMLR*, pp. 2337–2363, 2023. URL https://proceedings.mlr.press/v202/bhatnagar23a.html.

Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. URL https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf.

Yilun Chen and Wenjia Wang. Beyond non-degeneracy: Revisiting certainty equivalent heuristic for online linear programming. *arXiv preprint arXiv:2501.01716*, 2025. URL https://arxiv.org/abs/2501.01716.

Yuze Chen, Yuan Zhou, Baichuan Mo, Jie Ying, Yufei Ruan, and Zhou Ye. Online linear programming with replenishment. *arXiv preprint arXiv:2601.14629*, 2026. URL https://arxiv.org/abs/2601.14629.

Amit Daniely, Alon Gonen, and Shai Shalev-Shwartz. Strongly adaptive online learning. In *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pp. 1405–1411, 2015. URL https://proceedings.mlr.press/v37/daniely15.html.

Guneet S. Dhillon, George Deligiannidis, and Tom Rainforth. On the expected size of conformal prediction sets. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 1549–1557, 2024. URL https://proceedings.mlr.press/v238/dhillon24a.html.

Yihan Du, Siwei Wang, and Longbo Huang. A one-size-fits-all solution to conservative bandit problems. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pp. 7254–7261, 2021. doi: 10.1609/aaai.v35i8.16891. URL https://ojs.aaai.org/index.php/AAAI/article/view/16891.

Prasanjit Dubey and Xiaoming Huo. Anytime-valid federated conformal RAG for LLM swarms, 2026. URL https://arxiv.org/abs/2605.29139. Preprint.

Shai Feldman, Liran Ringel, Stephen Bates, and Yaniv Romano. Achieving risk control in online learning settings. *Transactions on Machine Learning Research*, 2023. ISSN 2835-8856. URL https://openreview.net/forum?id=5Y04GWuoJu.

Evrard Garcelon, Mohammad Ghavamzadeh, Alessandro Lazaric, and Matteo Pirotta. Conservative exploration in reinforcement learning. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pp. 1431–1441, 2020. URL https://proceedings.mlr.press/v108/garcelon20a.html.

Isaac Gibbs and Emmanuel J. Candès. Adaptive conformal inference under distribution shift. In *Advances in Neural Information Processing Systems*, volume 34, 2021.

Isaac Gibbs and Emmanuel J. Candès. Conformal inference for online prediction with arbitrary distribution shifts. *Journal of Machine Learning Research*, 25(162):1–36, 2024. URL https://jmlr.org/papers/v25/22-1218.html.

Sreenivas Gollapudi, Kostas Kollias, Kamesh Munagala, and Ali Sinop. Efficient online conformal selection with limited feedback, 2026. URL https://arxiv.org/abs/2605.14953v4. Preprint, version 4.

Hengquan Guo, Xin Liu, Honghao Wei, and Lei Ying. Online convex optimization with hard constraints: Towards the best of two worlds and beyond. In *Advances in Neural Information Processing Systems*, volume 35, 2022. URL https://papers.nips.cc/paper_files/paper/2022/hash/ec360cb73d322e80a877b7ec7e13c79a-Abstract-Conference.html.

Nicholas J. A. Harvey, Christopher Liaw, Edwin Perkins, and Sikander Randhawa. Optimal anytime regret with two experts. *Mathematical Statistics and Learning*, 6(1/2):87–142, 2023. doi: 10.4171/MSL/38. URL https://ems.press/journals/msl/articles/11474546.

Steven R. Howard and Aaditya Ramdas. Sequential estimation of quantiles with applications to A/B testing and best-arm identification. *Bernoulli*, 28(3):1704–1728, 2022. doi: 10.3150/21-BEJ1388. URL https://www.stevehoward.org/media/BEJ1388.pdf.

Steven R. Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021. doi: 10.1214/20-AOS1991. URL https://arxiv.org/abs/1810.08240.

Bror Hultberg, Dave Zachariah, and Antônio H. Ribeiro. Anytime-valid conformal risk control, 2026. URL https://arxiv.org/abs/2602.04364. Preprint.

Yotam Intrator, Regev Cohen, Ori Kelner, Roman Goldenberg, Ehud Rivlin, and Daniel Freedman. Streamlining conformal information retrieval via score refinement. In *Proceedings of the Seventh Fact Extraction and VERification Workshop (FEVER)*, pp. 186–191. Association for Computational Linguistics, 2024. doi: 10.18653/v1/2024.fever-1.22. URL https://aclanthology.org/2024.fever-1.22/.

Abbas Kazerouni, Mohammad Ghavamzadeh, Yasin Abbasi-Yadkori, and Benjamin Van Roy. Conservative contextual linear bandits. In *Advances in Neural Information Processing Systems*, volume 30, pp. 3910–3919, 2017. URL https://papers.nips.cc/paper_files/paper/2017/hash/bdc4626aa1d1df8e14d80d345b2a442d-Abstract.html.

Jing Lei and Larry Wasserman. Distribution-free prediction bands for non-parametric regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(1):71–96, 2014. doi: 10.1111/rssb.12021. URL https://www.stat.cmu.edu/~jinglei/LeiW14.pdf.

Jing Lei, James Robins, and Larry Wasserman. Distribution-free prediction sets. *Journal of the American Statistical Association*, 108(501):278–287, 2013. doi: 10.1080/01621459.2012.751873. URL https://www.stat.cmu.edu/~ryantibs/statml/lectures/Lei-Robins-Wasserman.pdf.

Tuo Liu, Edgar Dobriban, and Francesco Orabona. Online conformal prediction via universal portfolio algorithms, 2026. URL https://arxiv.org/abs/2602.03168. Preprint.

Shie Mannor, John N. Tsitsiklis, and Jia Yuan Yu. Online learning with sample path constraints. *Journal of Machine Learning Research*, 10:569–590, 2009. URL https://www.jmlr.org/papers/v10/mannor09a.html.

Jeremy McMahan and Xiaojin Zhu. Anytime-constrained reinforcement learning. In *Proceedings of the 27th International Conference on Artificial Intelligence and Statistics*, volume 238 of *Proceedings of Machine Learning Research*, pp. 4321–4329, 2024. URL https://proceedings.mlr.press/v238/mcmahan24a.html.

Aleksandr Podkopaev, Dong Xu, and Kuang-chih Lee. Adaptive conformal inference by betting. In *International Conference on Machine Learning*, volume 235 of *PMLR*, pp. 40886–40907, 2024. URL https://proceedings.mlr.press/v235/podkopaev24a.html.

Mauricio Sadinle, Jing Lei, and Larry Wasserman. Least ambiguous set-valued classifiers with bounded error levels. *Journal of the American Statistical Association*, 114(525):223–234, 2019. doi: 10.1080/01621459.2017.1395341. URL https://arxiv.org/abs/1609.00451.

Kayla E. Scharfstein and Arun Kumar Kuchibhotla. Time-uniform conformal and PAC prediction, 2026. URL https://arxiv.org/abs/2602.06297. Preprint.

Glenn Shafer and Vladimir Vovk. A tutorial on conformal prediction. *Journal of Machine Learning Research*, 9:371–421, 2008. URL https://jmlr.org/papers/v9/shafer08a.html.

Abhishek Sinha and Rahul Vaze. Optimal algorithms for online convex optimization with adversarial constraints. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://arxiv.org/abs/2310.18955.

Karthik Sridharan and Seung Won Wilson Yoo. Online learning with unknown constraints. In *Proceedings of the 42nd International Conference on Machine Learning*, volume 267 of *Proceedings of Machine Learning Research*, pp. 56790–56819, 2025. URL https://proceedings.mlr.press/v267/sridharan25a.html.

Vaidehi Srinivas. Online conformal prediction with efficiency guarantees. In *Proceedings of the 2026 Annual ACM–SIAM Symposium on Discrete Algorithms*, pp. 6696–6726, 2026. URL https://doi.org/10.1137/1.9781611978971.240.

Rahul Vaze. Simultaneous coverage and efficiency guarantee in online conformal prediction, 2026. URL https://arxiv.org/abs/2607.26577. Preprint.

Junxi Wu, Dongjian Hu, Yajie Bao, Shu-Tao Xia, and Changliang Zou. Error-quantified conformal inference for time series. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://proceedings.iclr.cc/paper_files/paper/2025/hash/cab5ae2704d3e01f06a92512a5376b87-Abstract-Conference.html.

Yifan Wu, Roshan Shariff, Tor Lattimore, and Csaba Szepesvári. Conservative bandits. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 1254–1262, 2016. URL https://proceedings.mlr.press/v48/wu16.html.

Zitong Yang, Emmanuel J. Candès, and Lihua Lei. Bellman conformal inference: Calibrating prediction intervals for time series, 2024. URL https://arxiv.org/abs/2402.05203. Preprint.

Xinlei Yi, Xiuxian Li, Tao Yang, Lihua Xie, Tianyou Chai, and Karl Johansson. Regret and cumulative constraint violation analysis for online convex optimization with long term constraints. In *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pp. 11998–12008, 2021. URL https://proceedings.mlr.press/v139/yi21b.html.

Margaux Zaffran, Olivier Feron, Yannig Goude, Julie Josse, and Aymeric Dieuleveut. Adaptive conformal predictions for time series. In *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pp. 25834–25866, 2022. URL https://proceedings.mlr.press/v162/zaffran22a.html.

A Proof of Theorem 1

The full-set and empty-set maps attain pairs of miscoverage and expected size $(0, B)$ and $(1, 0)$, and their mixtures attain every $q \in [0, 1]$. Thus $0 \le j \le B$ and $j(1) = 0$. Combining nearly optimal mixtures interpolates miscoverage and expected size. Letting their approximation errors vanish proves convexity of j. For $q < q' \le 1$, combining a mixture of miscoverage q with the empty-set map with weight $(q' - q)/(1 - q)$ raises its miscoverage to q' without increasing size. Thus j is nonincreasing and has a finite supporting slope $-\lambda \le 0$ at every interior target α.

Lemma 1 (One-round identity)**.** *Under Assumption 1, at every round t,*

$$\mathbb{E}[\ell_{C_t}(X_t) \mid \mathcal{H}_{t-1}] = j(\alpha) - \lambda(p_t - \alpha) + \psi(p_t) + \epsilon_t.$$

Proof. Conditional on $\mathcal{H}_{t-1}$, the selected map is independent of the fresh pair and its law is an admissible mixture in the definition of j. Averaging over the current seed gives p_t as the mixture's miscoverage probability. Its conditional expected size is $j(p_t)+\epsilon_t$. Expand $j(p_t)$ using the definition of ψ. □

Proof of Theorem 1. Taking expectations in Lemma 1 and summing gives

$$\mathcal{R}_T = -\lambda \mathbb{E} \sum_{t \le T} (p_t - \alpha) + \mathbb{E} \sum_{t \le T} \psi(p_t) + \mathbb{E} \sum_{t \le T} \epsilon_t.$$

The tower property gives $\mathbb{E} \sum_{t \le T} p_t = \mathbb{E} M_T$, so the first term is $\lambda \mathbb{E} G_T$. Bounded sizes and probabilities ensure integrability for finite T. □

A.1 Shared probability tools

The following facts are used in the upper bounds, lower bounds, and empirical-rank analysis. They do not require either controller.

Relative entropy and squared deviations. For probability laws P, Q, write $\mathrm{KL}(P\|Q) = \mathbb{E}_P \log(dP/dQ)$ when P is absolutely continuous with respect to Q, and $+\infty$ otherwise. For $q, q' \in [0,1]$, its Bernoulli case is $\mathrm{kl}(q\|q') = \mathrm{KL}(\mathrm{Bernoulli}(q)\|\,\mathrm{Bernoulli}(q'))$, agreeing with Section 4.2 when $q' = \alpha$. Both denote relative entropy; lowercase kl takes probabilities, while uppercase KL takes probability laws. In particular, $\mathrm{kl}(0\|0) = \mathrm{kl}(1\|1) = 0$. For $q \in [0,1]$ and $\alpha \in (0,1)$,

$$\mathrm{kl}(q\|\alpha) \le \frac{(q-\alpha)^2}{\alpha(1-\alpha)}, \tag{1}$$

$$\mathrm{kl}(q\|\alpha) \ge 2(q-\alpha)^2. \tag{2}$$

The first follows from $\log x \le x - 1$. The second follows by integrating $\partial_q^2 \,\mathrm{kl}(q\|\alpha) = 1/[q(1-q)] \ge 4$ twice from α. Both extend to the endpoints by continuity.

Realized excess size. Recall $R_t = W_t - tj(\alpha)$. With p_t, ϵ_t from Theorem 1, define

$$D_t = \psi(p_t) + \epsilon_t \ge 0, \qquad \mathsf{Z}_t = \sum_{u \le t} \xi_u,$$

$$\xi_t = \ell_{C_t}(X_t) - \mathbb{E}[\ell_{C_t}(X_t) \mid \mathcal{H}_{t-1}] + \lambda(E_t - p_t).$$

Then $\mathbb{E}[\xi_t \mid \mathcal{H}_{t-1}] = 0$ and $|\xi_t| \le B + \lambda$. Thus Z_t is a martingale with respect to $(\mathcal{H}_t)$. Expanding D_t and summing yields the pathwise identity

$$R_t = \lambda G_t + \sum_{u \le t} D_u + \mathsf{Z}_t. \tag{3}$$

This identity holds for every policy under Assumption 1, whether or not it is feasible. Taking expectations recovers Theorem 1.

B Proof of Theorem 2 and upper bounds

We first prove the empirical-rank law and size bound in Theorem 2, then establish the controller upper bounds stated in Section 4.3. The proof uses rank symmetry, order-statistic moments, and the stated assumptions.

Lemma 2 (Rank law and moments)**.** *For continuous i.i.d. scores and $\mathcal{G}_{t-1}$-measurable requests, Algorithm 3 satisfies $\mathbb{P}(E_t = 1 \mid \mathcal{G}_{t-1}) = \mathbb{E}[m_D(C_t) \mid \mathcal{G}_{t-1}] = q_t$ and $\mathrm{Var}(m_D(C_t) \mid \mathcal{G}_{t-1}) \le 1/(4t)$.*

Proof. **Conditional miscoverage law.** For any n, the permutation ordering n i.i.d. continuous scores is uniform over the $n!$ permutations and independent of their ordered values. Its insertion ranks form a bijection with $\prod_{i=1}^{n} \{1, \dots, i\}$. Thus the I_t are independent uniforms on $\{1, \dots, t\}$, and the sorted first $t-1$ scores are independent of $\mathcal{G}_{t-1}$, including its independent rounding coins.

Write $r_t = \lfloor t(1-q_t) \rfloor$ and $\theta_t = t(1-q_t) - r_t$. Algorithm 3 selects $J_t = r_t + \mathbf{1}\{V_t < \theta_t\}$. For an interior rank, the fresh score exceeds the J_t-th past score exactly when $I_t > J_t$. The same relation holds for the literal empty-set and full-set maps at $J_t = 0, t$, so $E_t = \mathbf{1}\{I_t > J_t\}$. Past miscoverage

indicators are therefore rank-history measurable. The new rank is independent of $\mathcal{G}_{t-1}$ and V_t, giving miscoverage probability $1 - J_t/t$ conditional on $\mathcal{G}_{t-1}$ and J_t. Since $\mathbb{E}[J_t \mid \mathcal{G}_{t-1}] = t(1-q_t)$, averaging over V_t gives

$$\mathbb{P}(E_t = 1 \mid \mathcal{G}_{t-1}) = q_t. \tag{4}$$

Conditional moments. The probability integral transform gives independent uniforms $U_i = F(S_i)$. Continuity suffices: F need not be strictly increasing, as the transform preserves ordering almost surely. Let $U_{(j)}$ be the j-th order statistic of the first $t-1$ uniforms, with $U_{(0)} = 0$ and $U_{(t)} = 1$. Then $m_D(C_t) = 1 - U_{(J_t)}$, including the literal endpoints. For $1 \le j \le t-1$, the density of $U_{(j)}$ is

$$\frac{(t-1)!}{(j-1)!(t-1-j)!}\, u^{j-1}(1-u)^{t-1-j}, \qquad 0 < u < 1.$$

Integrating gives

$$\mathbb{E}U_{(j)} = \frac{j}{t}, \qquad \mathbb{E}U_{(j)}^2 = \frac{j(j+1)}{t(t+1)}, \qquad \mathrm{Var}(U_{(j)}) = \frac{j(t-j)}{t^2(t+1)}. \tag{5}$$

The formulas also hold at the endpoints, including $t = 1$.

Condition on $\mathcal{G}_{t-1}$ and abbreviate $q = q_t$, $\theta = \theta_t$, $a = t(1-q)$. The sorted values retain their unconditional law and are independent of the new coin. Therefore

$$\mathbb{E}[J_t \mid \mathcal{G}_{t-1}] = a, \quad \mathrm{Var}(J_t \mid \mathcal{G}_{t-1}) = \theta(1-\theta), \quad \mathbb{E}[m_D(C_t) \mid \mathcal{G}_{t-1}] = q. \tag{6}$$

The law of total variance gives

$$\begin{aligned}\mathrm{Var}(m_D(C_t) \mid \mathcal{G}_{t-1}) &= \frac{\mathbb{E}[J_t(t-J_t) \mid \mathcal{G}_{t-1}]}{t^2(t+1)} + \frac{\theta(1-\theta)}{t^2} \\ &= \frac{q(1-q)}{t+1} + \frac{\theta(1-\theta)}{t(t+1)} \le \frac{1}{4t}.\end{aligned} \tag{7}$$

The second equality uses $\mathbb{E}[J_t(t-J_t) \mid \mathcal{G}_{t-1}] = a(t-a) - \theta(1-\theta)$. The last inequality uses $q(1-q), \theta(1-\theta) \le 1/4$. The rounding variance vanishes when $t(1-q)$ is an integer. The mean in (6) removes the cross term in $(m_D(C_t) - \alpha)^2$. Thus

$$\mathbb{E}\sum_{t=1}^{T}(m_D(C_t) - \alpha)^2 \le \mathbb{E}\sum_{t=1}^{T}(q_t - \alpha)^2 + \frac{\mathsf{h}_T}{4}. \tag{8}$$

□

Proof of Theorem 2. Lemma 2 gives the miscoverage law and moments. Conditional on the full past and the new coin, the selected map is fixed and the fresh pair has law D. Assumption 2 gives

$$\mathbb{E}[\ell_{C_t}(X_t) \mid \mathcal{H}_{t-1}, V_t] \le j(\alpha) - \lambda(m_D(C_t) - \alpha) + C(m_D(C_t) - \alpha)^2.$$

This includes the literal endpoints $J_t = 0, t$ and the first round. The tower property gives $\mathbb{E}\sum_{t \le T} m_D(C_t) = \mathbb{E}M_T = \alpha T - \mathbb{E}G_T$. Sum and apply (8) to prove the bound. □

From controller requests to prediction-set size. Theorem 2 and (2) give

$$\mathcal{R}_T \le \lambda \mathbb{E}G_T + C\left\{\frac{\mathcal{K}_T}{2} + \frac{\mathsf{h}_T}{4}\right\}, \qquad \mathcal{K}_T = \mathbb{E}\sum_{t=1}^{T}\mathrm{kl}(q_t \| \alpha), \quad \mathsf{h}_T = \sum_{t=1}^{T}\frac{1}{t}. \tag{9}$$

Counting and Barrier request q_t using only time and past miscoverage indicators (and the announced T for Counting). Since past indicators are rank-history measurable under Algorithm 3, both requests meet the measurability condition of Theorem 2. Its conditional law and the tower property give

$$\mathbb{P}(E_t = 1 \mid E_1, \ldots, E_{t-1}) = \mathbb{E}[q_t \mid E_1, \ldots, E_{t-1}] = q_t.$$

Thus the binary-path calculations below apply to the implemented controllers. The additional term $C\mathsf{h}_T/4 = O(\log T)$ accounts for threshold estimation.

B.1 FIXED-HORIZON UPPER BOUND

Let $\mathcal{A}_T$ be the binary paths satisfying every quota and ending with $b_T = \lfloor \alpha T \rfloor$ miscoverages, and let P_α be the law of T independent Bernoulli(α) variables. The path with entries $b_t - b_{t-1}$ shows that $\mathcal{A}_T$ is nonempty.

Lemma 3. *Algorithm 1 with literal endpoint actions satisfies every prefix quota and $M_T = b_T$ on every stream. If its miscoverage indicators have the requested conditional probabilities given their binary history, their law is $Q = P_\alpha(\cdot \mid \mathcal{A}_T)$ and*

$$\mathcal{K}_T := \mathbb{E}_Q \sum_{t=1}^{T} \mathrm{kl}(q_t \| \alpha) = \log \frac{1}{P_\alpha(\mathcal{A}_T)} \le \log T + \log(T+1) + \frac{1}{\alpha(1-\alpha)T}. \tag{10}$$

Moreover, $\mathbb{E}_Q \sum_{t=1}^{T} (q_t - \alpha)^2 \le \mathcal{K}_T/2$.

Proof. The recursion for $H_t(m)$ counts valid continuations. From a state with positive count, an outcome with zero count has probability zero and is prevented by the literal full-set or empty-set map. Otherwise either outcome is valid. Induction proves feasibility and the terminal total on every stream.

Under the requested conditional probabilities, the continuation-count ratios telescope to $1/H_0(0)$: each numerator is the next state's count, and the final count is one. Thus the miscoverage law is uniform on $\mathcal{A}_T$, which is also the conditional law Q: all paths with b_T ones have the same P_α probability. The relative-entropy chain rule gives $\mathcal{K}_T = \mathrm{KL}(Q \| P_\alpha)$.

To bound $\mathcal{K}_T$, condition on $N_T = b_T$, where N_T is the total under P_α. Every cyclic orbit of such words has a quota-feasible rotation: subtract b_T/T from each entry and rotate after a largest partial sum. The new centered partial sums are nonpositive differences from this maximum, including after wrapping since the centered total is zero. Thus the integer partial sums are at most $\lfloor \alpha t \rfloor$. Each orbit has at most T words, hence $P_\alpha(\mathcal{A}_T \mid N_T = b_T) \ge 1/T$. Set $u = b_T/T$. The count b_T is a mode of Binomial(T, u), including $b_T = 0$, and therefore has probability at least $1/(T+1)$. Comparing the same count under parameters u and α gives

$$P_\alpha(\mathcal{A}_T) \ge \frac{\exp\{-T\,\mathrm{kl}(u \| \alpha)\}}{T(T+1)}.$$

Since $|u - \alpha| < 1/T$, the upper relative-entropy bound (1) proves (10). The lower relative-entropy bound (2) bounds squared request deviations. □

The binary-history law established above verifies the hypothesis of Lemma 3 for empirical ranks. Applying (9) then yields

$$\mathcal{R}_T \le \lambda(\alpha T - b_T) + C\left(\frac{\mathcal{K}_T}{2} + \frac{\mathsf{h}_T}{4}\right) = O(\log T).$$

The continuation table uses $O(T + \alpha T^2)$ arithmetic operations and storage entries. This does not bound bit complexity for its integer counts. The implementation uses rolling log counts with `np.logaddexp`, stores double-precision request arrays, and uses finite-support masks to set forced requests to exactly zero or one. At the largest Counting horizon $T = 16{,}380$, for $\alpha = (0.05, 0.1, 0.2)$, median construction times over three fresh processes per target were $(0.114, 0.170, 0.289)$ seconds. The request arrays occupied $(51.2, 102.4, 204.8)$ MiB. The largest process peak resident set size across repeats was $(98.6, 158.7, 277.1)$ MiB, including Python, NumPy, and temporary arrays. These measurements used an Apple M5 Pro with 48 GiB RAM, macOS 26.4.1 (arm64), Python 3.12.14, and NumPy 2.3.5. Timing excludes interpreter startup and file I/O. An independent integer recurrence checked 200 small cases through $T = 40$, with maximum request error 1.34×10^{-15} and all forced endpoints exact. The anonymous code package includes the benchmark script. The default run repeats the benchmark on the current machine and saves source hashes, individual measurements, and validation results.

B.2 EXPECTED-PREFIX UPPER BOUND

To apply (9), we need to control both the miscoverage slack G_T and the cumulative squared request deviations $\mathbb{E}\sum_{t\le T}(q_t-\alpha)^2$. Keeping G_t inside $(0, L_t)$ controls the first term. We bound the second through the cumulative relative entropy $\mathcal{K}_T$, as for Counting. We first analyze an auxiliary controller with a sine potential: its addition formula makes the relative-entropy bound telescope across rounds. We then transfer the bound to Algorithm 2, which uses the quadratic function h_t and does not evaluate sine.

Write $\beta = 1-\alpha$. Replace h_t in Algorithm 2 by the sine potential

$$L_t = 4 + \lceil t^{1/3} \rceil, \qquad \omega_t = \pi/L_t, \qquad \kappa_t = \log \frac{\alpha \sin(\beta\omega_t)}{\beta \sin(\alpha\omega_t)}, \tag{11}$$

$$\phi_t(z) = e^{\kappa_t z} \sin(\omega_t z) \mathbf{1}\{0 < z < L_t\}. \tag{12}$$

The sine factor is positive inside $(0, L_t)$ and zero at its boundaries. The exponential factor is chosen to make (15) hold even when $\alpha \neq \beta$. For $z = G_{t-1}$, the next slack is $z+\alpha$ if $E_t = 0$ and $z-\beta$ if $E_t = 1$. Weighting these two destinations gives the request

$$Z_t(z) = \beta\phi_t(z+\alpha) + \alpha\phi_t(z-\beta), \qquad q_t = \frac{\alpha\phi_t(z-\beta)}{Z_t(z)}. \tag{13}$$

Feasibility. At $t = 1$, $q_1 = 0$ and the full-set map gives $G_1 = \alpha \in (0, L_1)$. Since $\phi_1(G_0) = 0$, the logarithmic calculation starts at the next round. Suppose $0 < z = G_{t-1} < L_{t-1} \le L_t$. The next states $z+\alpha$ and $z-\beta$ are separated by one, and at least one lies in $(0, L_t)$: if $z-\beta \le 0$, then $0 < z+\alpha \le 1 < L_t$. If $z+\alpha \ge L_t$, then $0 < L_t - 1 \le z-\beta < L_t$. Thus $Z_t(z) > 0$. A destination outside the interval has zero weight and is excluded by the corresponding full-set or empty-set map. Induction gives

$$0 < G_t < L_t \quad \text{on every stream and every coin realization,} \tag{14}$$

which implies $M_t \le \lfloor \alpha t \rfloor$.

Sine-potential bound. We first bound the normalizing factor $Z_t(z)$ in (13). This will bound each round's relative entropy by a change in log potential plus a term of order L_t^{-2}. Let $\bar{\phi}_t(z) = e^{\kappa_t z}\sin(\omega_t z)$ without truncation, and set $\chi(x) = \log(\sin x/x)$, with $\chi(0) = 0$. For $0 \le \omega \le \pi/4$, write $\kappa(\omega) = \chi(\beta\omega) - \chi(\alpha\omega)$, so $\kappa_t = \kappa(\omega_t)$. With $\kappa = \kappa(\omega)$,

$$\beta e^{\alpha\kappa}\sin(\alpha\omega) = \alpha e^{-\beta\kappa}\sin(\beta\omega).$$

This identity is the reason for the choice of κ_t: expanding the two shifted sines cancels their cosine terms and gives

$$\beta\bar{\phi}_t(z+\alpha) + \alpha\bar{\phi}_t(z-\beta) = \rho_t\bar{\phi}_t(z), \tag{15}$$
$$\log\rho_t = \chi(\omega_t) - \alpha\chi(\alpha\omega_t) - \beta\chi(\beta\omega_t).$$

Thus the weighted average of the untruncated potential at the two next states is a scalar multiple ρ_t of its current value. For $0 \le x \le \pi/4$,

$$-x \le \chi'(x) \le 0, \quad -x^2/2 \le \chi(x) \le 0, \quad |\kappa(\omega)| \le \omega^2/2, \quad |\kappa'(\omega)| \le \omega. \tag{16}$$

Indeed, $-\chi'(x) = \int_0^x u\sin u\,du/(x\sin x) \le x^2/(3\sin x) \le \pi x/6 \le x$. Integration proves the bound on χ. Its difference formula and $\alpha^2+\beta^2 \le 1$ prove the bounds on κ. All angles in use satisfy these estimates because $L_t \ge 5$. Monotonicity and nonpositivity of χ imply

$$0 < \rho_t \le 1, \qquad -\log\rho_t \le \omega_t^2/2. \tag{17}$$

For $0 < z < L_t$, a forbidden destination lies in $(-1, 0]$ or $[L_t, L_t+1)$, where $\bar{\phi}_t$ is nonpositive. Truncating to zero therefore gives

$$Z_t(z) \ge \rho_t\phi_t(z) > 0. \tag{18}$$

We now compare the controller's miscoverage law with independent Bernoulli(α) indicators. The sine request also uses only past miscoverage indicators, so the same rank-history and tower-property argument gives $\mathbb{P}(E_t = 1 \mid E_1, \dots, E_{t-1}) = q_t$. For $t \geq 2$, every allowed move satisfies

$$\begin{aligned}\log \frac{\mathbb{P}(E_t \mid E_1, \dots, E_{t-1})}{P_\alpha(E_t)} &= \log \phi_t(G_t) - \log Z_t(G_{t-1}) \\ &\leq \log \phi_t(G_t) - \log \phi_t(G_{t-1}) - \log \rho_t.\end{aligned}$$

The factors α or β cancel, and (18) gives the inequality. Zero-weight moves cannot occur, so the logarithms are finite. The conditional mean of the log likelihood ratio is $\mathrm{kl}(q_t \| \alpha)$. If the potential were fixed, its successive log increments would telescope exactly. Since L_t grows, we must also bound the change from ϕ_t to ϕ_{t+1} at the same slack value. For $0 < z < L_t \leq L_{t+1}$, monotonicity of $\sin x / x$ on $(0, \pi)$ and (16) give

$$\begin{aligned}\log \frac{\sin(\omega_t z)}{\sin(\omega_{t+1} z)} &\leq \log \frac{L_{t+1}}{L_t}, \\ |\kappa_t - \kappa_{t+1}| z &\leq \frac{\omega_t^2 - \omega_{t+1}^2}{2} \frac{\pi}{\omega_t} \leq \pi(\omega_t - \omega_{t+1}).\end{aligned}$$

Hence

$$\log \phi_t(z) - \log \phi_{t+1}(z) \leq \log \frac{L_{t+1}}{L_t} + \pi(\omega_t - \omega_{t+1}). \tag{19}$$

For $T \geq 2$, along every allowed path,

$$\begin{aligned}\sum_{t=2}^{T} [\log \phi_t(G_t) - \log \phi_t(G_{t-1})] &= \log \phi_T(G_T) - \log \phi_2(\alpha) \\ &\quad + \sum_{t=2}^{T-1} [\log \phi_t(G_t) - \log \phi_{t+1}(G_t)] \\ &\leq \frac{\pi^2}{2L_T} - \log \phi_2(\alpha) + \log \frac{L_T}{L_2} + \pi(\omega_2 - \omega_T).\end{aligned}$$

For the terminal term, $\sin(\omega_T G_T) \leq 1$ and $|\kappa_T| G_T \leq \pi^2/(2L_T)$. The intermediate terms are bounded by (19). The first round costs $\mathrm{kl}(0 \| \alpha) = \log(1/\beta)$. Set $C_\alpha = \log[3/(\alpha\beta)] + \pi^2$. Using $\sin(\pi\alpha/6) \geq \alpha/3$ and (17) therefore gives

$$\mathcal{K}_T := \mathbb{E} \sum_{t=1}^{T} \mathrm{kl}(q_t \| \alpha) \leq C_\alpha + \log L_T + \frac{\pi^2}{2} \sum_{t=2}^{T} L_t^{-2}. \tag{20}$$

This also holds for $T = 1$. Since $\sum_{t=1}^{T} L_t^{-2} \leq 3T^{1/3}$, Equation (9) implies

$$\mathcal{R}_T \leq \lambda L_T + C \left(\frac{\mathcal{K}_T}{2} + \frac{\mathsf{h}_T}{4} \right) = O(T^{1/3}). \tag{21}$$

The bound holds at every horizon for the same run, so $\max_{t \leq T} \mathcal{R}_t = O(T^{1/3})$.

Quadratic-potential bound. We transfer the preceding bound to $h_t(z) = z(L_t - z)\mathbf{1}\{0 < z < L_t\}$, the potential used in Algorithm 2. Write q_t^h and q_t^ϕ for the requests obtained from h_t and ϕ_t, respectively, at the same miscoverage history. Their potentials need not have similar values: only the relative weights of the two next slack values determine the request. We bound the change in these relative weights and show that replacing ϕ_t by h_t adds at most $9/(8L_t^2)$ per round to the relative-entropy bound.

Lemma 4 (Quadratic barrier)**.** *The quadratic controller has $0 < G_t < L_t$ on every stream. With empirical ranks under the assumptions of Theorem 2,*

$$\mathcal{K}_T^h := \mathbb{E} \sum_{t=1}^{T} \mathrm{kl}(q_t^h \| \alpha) \leq C_\alpha + \log L_T + \left(\frac{\pi^2}{2} + \frac{9}{8} \right) \sum_{t=2}^{T} L_t^{-2}. \tag{22}$$

Consequently $\max_{t \leq T} \mathcal{R}_t = O(T^{1/3})$.

Proof. The supports of h_t and ϕ_t agree, so the feasibility proof applies. Let Q^h, Q^ϕ be the laws of their miscoverage indicators through T. Both requests are functions of past miscoverage indicators. Theorem 2 and the tower property therefore realize their requested probabilities conditional on this binary history. The two laws allow exactly the same binary paths.

For $L = L_t$, $\kappa = \kappa_t$, and $u = z/L \in (0,1)$, set $r(z) = z(L-z)/[e^{\kappa z}\sin(\pi z/L)]$. Then using $\cot x := \cos x/\sin x$ and $\frac{d}{dx}\log\sin x = \cot x$ for $0 < x < \pi$ gives

$$(\log r)'(z) = \frac{1}{L}\left\{\frac{1}{u} - \frac{1}{1-u} - \pi\cot(\pi u)\right\} - \kappa.$$

The expression in braces has absolute value at most 2. On $0 < u \le 1/2$, the function $1/u - \pi\cot(\pi u)$ increases from 0 to 2, since its derivative is $\pi^2/\sin^2(\pi u) - 1/u^2 \ge 0$, while $1 \le 1/(1-u) \le 2$. The other half follows by antisymmetry. Equation (16) and $L \ge 5$ give $|(\log r)'| \le 2/L + \pi^2/(2L^2) < 3/L$.

When both next slack values lie inside $(0, L_t)$, they are one unit apart. The derivative bound therefore controls the change in log odds at the same miscoverage history:

$$\left|\log\frac{q_t^h}{1-q_t^h} - \log\frac{q_t^\phi}{1-q_t^\phi}\right| = |\log r(G_{t-1}-\beta) - \log r(G_{t-1}+\alpha)| \le 3/L_t.$$

In log-odds coordinates, Bernoulli relative entropy is a first-order Taylor remainder of $a(x) = \log(1+e^x)$. Since $a'' \le 1/4$, $\mathrm{kl}(q_t^h \| q_t^\phi) \le 9/(8L_t^2)$. If a destination is forbidden, both requests are the same endpoint and this relative entropy is zero. Both also request zero at $t = 1$. The chain rule gives

$$\mathrm{KL}(Q^h \| Q^\phi) \le \frac{9}{8}\sum_{t=2}^{T} L_t^{-2}.$$

To compare Q^h with P_α, we also need the sine controller's bound under Q^h, rather than under Q^ϕ. The calculation preceding (20) holds for every allowed path, before taking any expectation. Hence, for every path e in this common support,

$$\log\frac{Q^\phi(e)}{P_\alpha(e)} \le C_\alpha + \log L_T + \frac{\pi^2}{2}\sum_{t=2}^{T} L_t^{-2}.$$

Averaging under Q^h and using

$$\mathrm{KL}(Q^h \| P_\alpha) = \mathrm{KL}(Q^h \| Q^\phi) + \mathbb{E}_{Q^h}\log\frac{Q^\phi(e)}{P_\alpha(e)}$$

proves (22). All ratios are defined on the common support. Finally, (2) bounds the cumulative squared request deviations by $\mathcal{K}_t^h/2$. Together with $G_t < L_t$, Equation (9) gives (21) with $\mathcal{K}_t^h$ in place of $\mathcal{K}_t$, and therefore the expected-prefix bound for Algorithm 2. □

B.3 Retrospective upper bound and adaptive stopping

Use the realized-cost identity (3) and the variables D_t, ξ_t, Z_t defined in Appendix A.1. For Barrier and the sine variant in Appendix B.2, $G_t < L_t \le L_T$ and $\sum_{u\le t} D_u \le \sum_{u\le T} D_u$. Martingale differences are uncorrelated, so $\mathbb{E}\mathsf{Z}_T^2 = \sum_{t\le T}\mathbb{E}\xi_t^2 \le T(B+\lambda)^2$. The L^2 martingale maximal inequality therefore gives $\mathbb{E}\max_{t\le T}|\mathsf{Z}_t| \le 2(\mathbb{E}\mathsf{Z}_T^2)^{1/2}$. Taking expectations in (3) at T gives $\mathbb{E}\sum_{u\le T} D_u = \mathcal{R}_T - \lambda\mathbb{E}G_T$. Hence

$$\begin{aligned}\mathbb{E}\max_{1\le t\le T} R_t &\le \lambda L_T + \mathbb{E}\sum_{u\le T} D_u + \mathbb{E}\max_{t\le T}|\mathsf{Z}_t| \\ &\le \lambda L_T + \mathcal{R}_T - \lambda\mathbb{E}G_T + 2(B+\lambda)\sqrt{T} = O(\sqrt{T}),\end{aligned}$$

by the expected-prefix bound, proving the retrospective upper bound for the same run. With the deterministic benchmark $tj(\alpha)$, this criterion includes oracle size fluctuations. The oracle interval in Section 4.1 has constant size, so the square-root lower bound also holds without these fluctuations.

Proposition 1. *Under the assumptions of Theorem 2, Barrier and the sine variant each satisfy, for every stopping time $1 \le \tau \le T$ of $(\mathcal{H}_t)$,*

$$\mathbb{E}[W_\tau - \tau j(\alpha)] \le \lambda L_T + \mathcal{R}_T - \lambda \mathbb{E} G_T = O(T^{1/3}). \tag{23}$$

Proof. Since $\tau \le T$ and the increments are bounded, optional stopping gives $\mathbb{E}\mathsf{Z}_\tau = 0$. In (3), use $G_\tau < L_T$ and $\sum_{u \le \tau} D_u \le \sum_{u \le T} D_u$, then take expectations. □

C Proof of Theorem 3

Throughout this section, write $v = \alpha(1-\alpha)$ and $x_+ = \max\{x, 0\}$ for the positive part of a real number x. The proof uses the common identities in Appendix A.1 and the two lemmas below. It applies to arbitrary feasible policies, including policies that use the full history and know D, T.

Lemma 5 (Bernoulli survival bound)**.** *Let B_t be independent Bernoulli(α) variables with joint law P_α, let $\mathsf{S}_t = \sum_{u \le t}(B_u - \alpha)$, and set $K_\alpha = 1 + 5/v$. For every integer $T \ge 1$,*

$$P_\alpha(\mathsf{S}_t \le 0 \text{ for all } t \le T) \le \min\{1, K_\alpha/\sqrt{T}\}. \tag{24}$$

Proof. For $a \ge 1$, stop at $\tau = \inf\{t \ge 1 : \mathsf{S}_t > 0 \text{ or } \mathsf{S}_t \le -a\}$. The stopped process lies in $[-a-1, 1]$, so $v\mathbb{E}(\tau \wedge n) = \mathbb{E}\mathsf{S}^2_{\tau \wedge n} \le (a+1)^2$. Stop $\mathsf{S}_t^2 - vt$ at bounded $\tau \wedge n$ for this identity. The stopped mean is also zero. Monotone convergence gives $\mathbb{E}\tau < \infty$. Bounded convergence of the stopped positions gives $\mathbb{E}\mathsf{S}_\tau = 0$ and $\mathbb{E}\mathsf{S}_\tau^2 = v\mathbb{E}\tau$. If p_- is the lower-exit probability, then $0 \le 1 - (a+1)p_-$, whence $p_- \le 1/a$ and

$$v\mathbb{E}\tau \le 1 + p_-(a+1)^2 \le 5a.$$

A path staying nonpositive through T either exits through the lower boundary or has $\tau > T$. Its probability is at most $1/a + 5a/(vT)$. Set $a = \sqrt{T}$ to obtain (24). □

Lemma 6 (Squared deviations on a time window)**.** *For a feasible policy through T, let $\bar{q}_t = \mathbb{P}(E_t = 1 \mid E_1, \dots, E_{t-1})$ and $a = v/4$. For integers $s \ge 0$, $n \ge 1$ with $s + n \le T$,*

$$\mathbb{E} \sum_{t=s+1}^{s+n} (\bar{q}_t - \alpha)^2 \ge \frac{(a\sqrt{n} - \mathbb{E}G_s)_+^2}{n}. \tag{25}$$

Proof. For $N \sim \mathrm{Binomial}(n, \alpha)$, $n \ge 1$, we use

$$\mathbb{E}(N - n\alpha)_+ \ge a\sqrt{n}. \tag{26}$$

Indeed, $X = N - n\alpha$ satisfies $\mathbb{E}X^2 = nv$ and $\mathbb{E}X^4 = 3n^2v^2 + nv(1-6v) \le n^2 v(1+3v)$. Hölder's inequality and $\mathbb{E}X = 0$ give

$$\mathbb{E}X_+ = \frac{1}{2}\mathbb{E}|X| \ge \frac{(\mathbb{E}X^2)^{3/2}}{2(\mathbb{E}X^4)^{1/2}} \ge \frac{v\sqrt{n}}{2\sqrt{1+3v}} \ge a\sqrt{n}.$$

Realize the policy's miscoverage law using independent $U_t \sim \mathrm{Unif}[0,1]$: $E_t = \mathbf{1}\{U_t \le \bar{q}_t(E_1, \dots, E_{t-1})\}$, and couple it with $B_t = \mathbf{1}\{U_t \le \alpha\}$. This preserves the policy's miscoverage law without assuming independent miscoverage indicators. For a window $s+1, \dots, s+n$ with $s + n \le T$, let

$$N = \sum_{t=s+1}^{s+n} B_t, \qquad D_{s,n} = \sum_{t=s+1}^{s+n} \mathbf{1}\{\bar{q}_t < U_t \le \alpha\}.$$

The coupling and feasibility give $N - (M_{s+n} - M_s) \le D_{s,n}$ and $M_{s+n} - M_s \le \alpha n + G_s$. Since $D_{s,n} + G_s \ge 0$,

$$a\sqrt{n} \le \mathbb{E}(N - \alpha n)_+ \le \mathbb{E}G_s + \mathbb{E}D_{s,n} = \mathbb{E}G_s + \mathbb{E} \sum_{t=s+1}^{s+n} (\alpha - \bar{q}_t)_+.$$

Cauchy–Schwarz gives

$$\mathbb{E}\sum_{t=s+1}^{s+n}(\alpha-\bar{q}_t)_+ \le \left\{ n\mathbb{E}\sum_{t=s+1}^{s+n}(\bar{q}_t-\alpha)^2 \right\}^{1/2}.$$

Rearrange and square when $a\sqrt{n} > \mathbb{E}G_s$. Otherwise the lower bound in (25) is zero. □

Proof of Theorem 3. Let Q be the miscoverage-path law of any feasible policy and put $\bar{q}_t = Q(E_t = 1 \mid E_1, \dots, E_{t-1})$. Conditioning on past miscoverages averages over the remaining history and current randomization, leaving an admissible mixture independent of the fresh pair. Thus, for every $s \le T$,

$$\mathcal{R}_s \ge \lambda\mathbb{E}G_s + \mathbb{E}_Q\sum_{t=1}^{s}\psi(\bar{q}_t). \tag{27}$$

This includes policies using the full history and knowing D, T. Feasibility gives $G_s \ge 0$. Since $j(1) = 0$ and $j(\alpha) \ge 0$, $\psi(1) > 0$ also implies $\lambda > 0$.

(i) Terminal lower bound. Let $\mathcal{B}_T$ be the binary paths satisfying every prefix quota. Because counts are integer valued, this is precisely the event in (24). Feasibility gives $Q(\mathcal{B}_T) = 1$. The reference law P_α assigns positive mass to every binary path, so

$$\begin{aligned}\mathbb{E}_Q\sum_{t=1}^{T}\mathrm{kl}(\bar{q}_t\|\alpha) &= \mathrm{KL}(Q\|P_\alpha)\\ &= \mathrm{KL}\big(Q\|P_\alpha(\cdot \mid \mathcal{B}_T)\big) + \log\frac{1}{P_\alpha(\mathcal{B}_T)} \ge \log\frac{1}{P_\alpha(\mathcal{B}_T)}.\end{aligned} \tag{28}$$

Conditioning P_α on $\mathcal{B}_T$ divides feasible path probabilities by $P_\alpha(\mathcal{B}_T)$. Relative entropy is nonnegative. Apply the quadratic lower bound on ψ, (1), and $G_T \ge \alpha T - b_T$ in (27). Together with Lemma 5, these give

$$\mathcal{R}_T \ge \lambda(\alpha T - b_T) + c_- v\left(\tfrac{1}{2}\log T - \log K_\alpha\right)_+, \tag{29}$$

which proves (i).

(ii) Expected-prefix lower bound. Set $a = v/4$, $R = \max_{t\le T}\mathcal{R}_t \ge 0$ and choose

$$n = \left\lceil \left(1 + \frac{2R}{a\lambda}\right)^2 \right\rceil, \qquad d = c_- a^2/4.$$

Equation (27) gives $\mathbb{E}G_s \le R/\lambda \le a\sqrt{n}/2$. The same holds at $s = 0$ because $G_0 = 0$. Lemma 6 gives at least $a^2/4$ in squared deviations per complete window. Summing disjoint windows gives

$$R \ge \mathcal{R}_T \ge d\lfloor T/n \rfloor, \qquad T < n(1 + R/d).$$

Expected squared deviations add across windows. Independence is unnecessary. These inequalities also hold when $n > T$. With $A = 2\max\{1, 2/(a\lambda)\}^2$, we have $n \le A(1+R)^2$, so $T \le A(1 + 1/d)(1+R)^3$. Thus (ii) holds with $c_2 = [A(1+1/d)]^{-1/3} > 0$. This establishes the exponent, without certifying finite-horizon near-optimality.

(iii) Retrospective-maximum lower bound. For this part assume only $\delta := \psi(1) > 0$. Use D_t, ξ_t, Z_t from (3). Since $j(1) = 0$ and costs are nonnegative, $\delta = \lambda(1-\alpha) - j(\alpha) \le \lambda$. Expanding ξ_t gives

$$\xi_t = \ell_{C_t}(X_t) + \lambda E_t - \{j(\alpha) + \lambda\alpha + D_t\}.$$

On a miscoverage, $\xi_t \ge \delta - D_t$. As D_t is predictable,

$$\mathbb{E}[\xi_t^2 \mid \mathcal{H}_{t-1}] \ge \frac{\delta^2}{4}p_t\mathbf{1}\{D_t \le \delta/2\}.$$

Using $p_t \le 1$, $\sum_t \mathbb{E}p_t = \alpha T - \mathbb{E}G_T$, and $\mathcal{R}_T = \lambda\mathbb{E}G_T + \mathbb{E}\sum_t D_t$, we obtain

$$\begin{aligned}\mathbb{E}\mathsf{Z}_T^2 &\ge \frac{\delta^2}{4}\left\{\alpha T - \mathbb{E}G_T - \frac{2}{\delta}\mathbb{E}\sum_{t\le T}D_t\right\}\\ &\ge \frac{\delta^2}{4}\left\{\alpha T - \frac{2}{\delta}\mathcal{R}_T\right\}.\end{aligned}$$

Martingale differences are orthogonal, so their second moments add. Rounds with $D_t > \delta/2$ contribute at most $(2/\delta)\mathbb{E}\sum_t D_t$ to the discarded miscoverage probabilities. The last inequality uses $\delta \le \lambda$ and $G_T \ge 0$. If $\mathcal{R}_T \ge \alpha\delta T/4$, the expected maximum is at least $(\alpha\delta/4)\sqrt{T}$. Otherwise, $\mathbb{E}\mathsf{Z}_T^2 \ge \alpha\delta^2 T/8$.

Put $K = B + \lambda$. The bounded martingale differences satisfy $|\xi_t| \le K$ and $\mathbb{E}\mathsf{Z}_{t-1}^2 \le K^2(t-1)$. Expanding the fourth power conditionally and using $4K^3|x| \le 2K^2x^2 + 2K^4$ gives

$$\begin{aligned}\mathbb{E}\mathsf{Z}_t^4 &\le \mathbb{E}\mathsf{Z}_{t-1}^4 + 6K^2\mathbb{E}\mathsf{Z}_{t-1}^2 + 4K^3\mathbb{E}|\mathsf{Z}_{t-1}| + K^4 \\ &\le \mathbb{E}\mathsf{Z}_{t-1}^4 + \{8(t-1)+3\}K^4.\end{aligned}$$

Summing yields $\mathbb{E}\mathsf{Z}_T^4 \le 4K^4T^2$. Hence

$$\mathbb{E}(\mathsf{Z}_T)_+ = \frac{1}{2}\mathbb{E}|\mathsf{Z}_T| \ge \frac{(\mathbb{E}\mathsf{Z}_T^2)^{3/2}}{2(\mathbb{E}\mathsf{Z}_T^4)^{1/2}} \ge \frac{\alpha^{3/2}\delta^3}{64\sqrt{2}K^2}\sqrt{T}.$$

The decomposition gives $R_T \ge \mathsf{Z}_T$. The first prediction is full under the all-stream quota, and mixing the full-set and empty-set maps gives $j(\alpha) \le B(1-\alpha)$. Thus $R_1 \ge 0$ and $\max_{1\le t\le T} R_t \ge (\mathsf{Z}_T)_+$ pathwise. We conclude that

$$\mathbb{E}\max_{1\le t\le T} R_t \ge k\sqrt{T}, \qquad k = \min\left\{\frac{\alpha\delta}{4}, \frac{\alpha^{3/2}\delta^3}{64\sqrt{2}(B+\lambda)^2}\right\} > 0. \tag{30}$$

Under the theorem's quadratic assumption, $\delta \ge c_-(1-\alpha)^2 > 0$, completing (iii). □

C.1 Terminal and expected-prefix tradeoff

Corollary 1 (Terminal and expected-prefix tradeoff)**.** *Under the assumptions of Theorem 3, there is a constant $c_4 > 0$, depending only on α, λ, c_-, such that every policy feasible through T satisfies*

$$(1+\mathcal{R}_T)\left(1+\max_{1\le t\le T}\mathcal{R}_t\right)^2 \ge c_4 T.$$

Thus $\mathcal{R}_T = O(\log T)$ requires $\max_{t\le T}\mathcal{R}_t = \Omega(\sqrt{T/\log T})$, while $\max_{t\le T}\mathcal{R}_t = O(T^{1/3})$ requires $\mathcal{R}_T = \Omega(T^{1/3})$.

Proof. Write $P = \max_{1\le t\le T}\mathcal{R}_t \ge 0$ and use the constants $a = \alpha(1-\alpha)/4$, $d = c_-a^2/4$, and $A = 2\max\{1, 2/(a\lambda)\}^2$ from part (ii). With $n = \lceil(1+2P/(a\lambda))^2\rceil$, the same window argument gives $\mathcal{R}_T \ge d\lfloor T/n\rfloor$ and $n \le A(1+P)^2$. Consequently

$$T < n(1+\mathcal{R}_T/d) \le A(1+P)^2(1+\mathcal{R}_T/d) \le A(1+1/d)(1+P)^2(1+\mathcal{R}_T).$$

This proves the corollary with $c_4 = [A(1+1/d)]^{-1} = c_2^3$. □

All constants are independent of the policy and T, so the two asymptotic consequences also hold for policies chosen separately at each horizon. This lower bound does not assert attainability of the entire terminal–prefix tradeoff curve.

D Proof of Theorem 4

The theorem combines the constructive upper bounds in Appendix B with the lower bounds in Appendix C.

Proof. The assumptions of Theorem 4 include those of both Theorem 2 and Theorem 3. Counting implemented by empirical ranks is feasible on every stream, ends with $M_T = b_T$, and has $\mathcal{R}_T = O(\log T)$ (Appendix B.1). One infinite run of Barrier implemented by empirical ranks is feasible on every stream and satisfies $\max_{t\le T}\mathcal{R}_t = O(T^{1/3})$ and $\mathbb{E}\max_{t\le T} R_t = O(\sqrt{T})$ (Appendices B.2 and B.3). These give the upper bounds in the three infima. Theorem 3 supplies matching lower bounds uniformly over every policy feasible through T, even if it knows D, T. Taking the respective infima proves all three Θ rates, with constants depending only on $\alpha, B, \lambda, C, c_-$. The implemented controllers use only α, time, past scores and miscoverage indicators, and fresh random seeds (plus T for Counting). □

Adaptive-stopping consequence. Let $\mathcal{T}_T$ consist of integer-valued stopping times of $(\mathcal{H}_t)$ with $1 \le \tau \le T$. Under the assumptions of Theorem 4, the optimal price over feasible policies, which may know D, T, is

$$\inf_{\mathcal{A} \text{ feasible through } T} \sup_{\tau \in \mathcal{T}_T} \mathbb{E}_D[W_\tau - \tau j(\alpha)] = \Theta(T^{1/3}). \tag{31}$$

The lower bound follows from Theorem 3(ii), since deterministic times belong to $\mathcal{T}_T$. Proposition 1 gives the uniform upper bound for Algorithm 2.

E Details of the geometric assumptions

E.1 Target efficiency and smoothness

Target efficiency and local smoothness imply Assumption 2. A Taylor bound controls excess size near α, and bounded size extends this bound to all miscoverage levels.

Suppose threshold size is parametrized by miscoverage as $h : [0,1] \to [0,B]$, including the full-set and empty-set maps at $q = 0$ and $q = 1$. Since thresholds belong to $\mathcal{C}$, $h(q) \ge j(q)$. If $h(\alpha) = j(\alpha)$ and h' is L-Lipschitz on $[\alpha - \delta, \alpha + \delta] \subset (0,1)$ for some $\delta > 0$, the supporting line for j gives

$$h(\alpha + u) - h(\alpha) \ge j(\alpha + u) - j(\alpha) \ge -\lambda u.$$

Divide by $u > 0$ and $u < 0$ and take limits. Differentiability of h at α gives $h'(\alpha) = -\lambda$. For $|q - \alpha| \le \delta$, the remainder is therefore $h(q) - j(\alpha) + \lambda(q - \alpha) = \int_\alpha^q [h'(u) - h'(\alpha)]\, du$. The Lipschitz bound on the integrand yields

$$0 \le h(q) - j(\alpha) + \lambda(q - \alpha) \le \frac{L}{2}(q - \alpha)^2 \qquad (|q - \alpha| \le \delta).$$

Outside this interval the excess above the supporting line is at most $B + \lambda$, because $h(q) \le B$, $j(\alpha) \ge 0$, and $|q - \alpha| \le 1$. There $(q - \alpha)^2 \ge \delta^2$, so Assumption 2 holds with $C = \max\{L/2, (B + \lambda)/\delta^2\}$. Conversely, the assumption at $q = \alpha$ requires $h(\alpha) = j(\alpha)$. Smoothness alone does not imply this equality.

If the upper bound in Assumption 2 is relaxed to $C(q - \alpha)^2 + \Delta$ for a constant $\Delta \ge 0$ and every threshold, the proof of Theorem 2 adds $T\Delta$ to the expected-size bound. A fixed positive gap can dominate the sublinear control cost. Without target efficiency, quota feasibility, the rank miscoverage law and variance bound, and Theorem 1 remain valid under their respective assumptions. The matching upper rates need not hold. Taking $\mathcal{C}$ to be a fixed score's threshold family restricts both the comparator and admissible policies. Its frontier gives the lowest expected size attainable by mixing thresholds, which can lie below the threshold-size curve. If its value at α is attained by mixing two thresholds with miscoverages $q_- < \alpha < q_+$, the frontier is linear on $[q_-, q_+]$. Then ψ vanishes on this interval, violating strict quadratic lower growth. Restricting the comparison family alone therefore does not preserve the three matching rates. The empirical gaps in Table 4 concern this restricted comparison. They neither establish efficiency at the target nor estimate a uniform Δ.

E.2 The continuous example

We verify the frontier in Section 4.1 against all measurable sets and their randomized mixtures.

The model has quantile function Q, cdf F, and density f given by

$$Q(u) = \frac{3u + u^2}{4}, \qquad F(y) = \frac{\sqrt{9 + 16y} - 3}{2}, \qquad f(y) = \frac{4}{\sqrt{9 + 16y}}, \qquad 0 \le y \le 1.$$

For a measurable set A of length v, the sets $A \setminus [0, v]$ and $[0, v] \setminus A$ have equal length. Since f decreases, the former has no greater probability than the latter. Thus $\int_A f \le F(v)$: a set with miscoverage q has length at least $Q(1 - q)$, attained by $[0, Q(1 - q)]$. For a randomized set A, apply this bound to each realization before averaging. Convexity of Q then gives $\mathbb{E}\mu(A) \ge \mathbb{E}Q(1 - m_D(A)) \ge Q(1 - \mathbb{E}m_D(A))$. Thus thresholds attain j as defined in Section 3, giving

$$j(q) = 1 - \frac{5}{4}q + \frac{1}{4}q^2, \qquad \lambda = \frac{5}{4} - \frac{\alpha}{2}, \qquad \psi(q) = \frac{1}{4}(q - \alpha)^2.$$

This verifies both quadratic conditions with $C = c_- = 1/4$, including the full-set and empty-set maps.

E.3 THE COVARIATE-DEPENDENT EXAMPLE

The quadratic conditions also hold when optimal sets depend on the covariate. The optimal prediction-set rule uses the same density cutoff for both covariate values, giving different interval lengths.

Let X be uniform on $\{0,1\}$ and let $Y \mid X = x$ have density

$$f_0(y) = \tfrac{3}{2} - y, \qquad f_1(y) = \tfrac{3}{2} - y + \frac{\sin(2\pi y)}{4\pi}, \qquad 0 \le y \le 1. \tag{32}$$

Both integrate to one and satisfy $1/2 \le -f_x' \le 3/2$. Use $s(x,y) = 3/2 - f_x(y)$ and length cost, with $\mathcal{C}$ containing all measurable prediction-set rules. The frontier j also allows randomization over these rules.

Lemma 7 (Conditional-density thresholds)**.** *For* (32)*, the score thresholds attain the randomized population frontier at every miscoverage. For $k \in [1/2, 3/2]$, let $r_x(k)$ be the unique solution of $f_x(r_x) = k$, let F_x be the conditional cdf, and set*

$$q(k) = 1 - \tfrac{1}{2}\{F_0(r_0(k)) + F_1(r_1(k))\}, \qquad A(k) = \tfrac{1}{2}\sum_{x=0}^{1} \frac{1}{-f_x'(r_x(k))}.$$

As k increases from $1/2$ to $3/2$, $q(k)$ increases from 0 to 1. The frontier and its derivatives are

$$j(q(k)) = \tfrac{1}{2}\{r_0(k) + r_1(k)\}, \qquad j'(q(k)) = -\frac{1}{k}, \qquad j''(q(k)) = \frac{1}{k^3 A(k)} \in [4/27, 12]. \tag{33}$$

Consequently Assumption 2 holds with $C = 6$, and $\psi(q) \ge (2/27)(q-\alpha)^2$ for every $q \in [0,1]$ and every $\alpha \in (0,1)$. The score cdf is continuous.

Proof. Write $C^{(k)} := C_{3/2-k}$ for the score-threshold map, so $C^{(k)}(x) = \{y : f_x(y) \ge k\} = [0, r_x(k)]$. For any $C \in \mathcal{C}$, the integrand below is nonnegative at every covariate and outcome, so

$$\mathbb{E}_X \int_0^1 \big(\mathbf{1}\{y \in C^{(k)}(X)\} - \mathbf{1}\{y \in C(X)\}\big)(f_X(y) - k)\, dy \ge 0.$$

Average this inequality over $C \sim \pi$, with $\pi \in \Delta(\mathcal{C})$. If $m_D(\pi) = m_D(C^{(k)})$, the density terms cancel, leaving $V_D(\pi) \ge V_D(C^{(k)})$. Hence $C^{(k)}$ attains the frontier even when randomization is allowed. The full-set and empty-set maps attain the endpoints. Singleton boundaries have zero length and probability.

Implicit differentiation gives $r_x'(k) = 1/f_x'(r_x(k))$, and $f_x(r_x(k)) = k$ makes $\frac{d}{dk}F_x(r_x(k)) = k/f_x'(r_x(k))$. Hence $q'(k) = kA(k) > 0$ and $\frac{d}{dk}V_D(C^{(k)}) = -A(k)$. Dividing gives $j' = -1/k$, and differentiating again gives $j'' = 1/(k^3 A(k))$. Since $2/3 \le A(k) \le 2$ and $1/2 \le k \le 3/2$, the stated bounds follow. At $k = 1/2$ both intervals are full, and at $k = 3/2$ both have zero length and probability. Thus $q(k)$ covers the full miscoverage range. Write $k(q)$ for its inverse. Taylor's formula around α, with $\lambda = 1/k(\alpha)$, then bounds the remainder between $(2/27)(q-\alpha)^2$ and $6(q-\alpha)^2$, including the endpoints by continuity. These are the quadratic conditions in Theorem 4. Finally, each $s(x,\cdot)$ is strictly increasing with derivative in $[1/2, 3/2]$ and the conditional response density is continuous. The mixture score distribution is therefore continuous. □

E.4 THE FLAT-FRONTIER EXAMPLE

A linear frontier allows full-set and empty-set maps to use each new quota unit immediately, keeping all three prices bounded.

Let $Y \sim \mathrm{Unif}[0,1]$, take length cost, and allow all measurable sets and randomization over them. Coverage equals expected length, so $j(q) = 1 - q$, $\lambda = 1$, and $\psi \equiv 0$. Issue the empty set when $b_t - b_{t-1} = 1$ and the full set otherwise. The miscoverage indicators are deterministic on every stream, giving

$$M_t = b_t, \qquad W_t = t - b_t, \qquad R_t = \alpha t - \lfloor \alpha t \rfloor \in [0,1).$$

All three criteria are bounded by one. Assumption 2 holds with $C = 0$, but positive quadratic lower growth fails.

F EXPERIMENTAL DETAILS

In the anonymous code package, `python run.py` regenerates all six figures and six experimental tables using the included California Housing score pools. Adding `--retrain` also refits the Housing models. The package README gives the required dependencies and setup commands.

F.1 CONTINUOUS MODEL AND COST ESTIMATION

This model has explicit formulas for prediction-set size and miscoverage, so we can separate coverage control from threshold estimation.

Draw $U \sim \mathrm{Unif}[0,1]$, set $Y = (3U + U^2)/4$, and use score $s(y) = y$, intervals $[0, \tau]$, and interval length as set size. The empty and full sets have size zero and one, respectively. The density is $f(y) = 4/\sqrt{9 + 16y}$, and

$$j(q) = 1 - \tfrac{5}{4}q + \tfrac{1}{4}q^2, \qquad \lambda = \tfrac{5}{4} - \tfrac{\alpha}{2}, \qquad \psi(q) = \tfrac{1}{4}(q - \alpha)^2.$$

Appendix E verifies Assumption 2. We use targets $\alpha \in \{0.05, 0.1, 0.2\}$. Counting is recomputed at horizons $160, 320, 640, 1280, 2560, 5120, 10240, 16380$. These multiples of 20 remove integer-rounding slack. Barrier and the sine variant use $L_t = 4 + \lceil t^{1/3} \rceil$. ***Barrier*** is Algorithm 2, with quadratic potential $h_t(z) = z(L_t - z)$. The ***sine variant*** replaces only this potential by (12) for a secondary comparison. Both use Algorithm 3 unless the curve is labeled known F. The ***width variant*** uses the quadratic potential with the separately validated schedule in Appendix F.4. Learners receive only past scores and miscoverage indicators. The known law is used for the oracle comparators, evaluation, and the validation criteria below.

Numerical expectations. We track the probability of each controller state from one round to the next and compute expected sizes using uniform order-statistic moments. For empirical ranks, let $v_t := \mathrm{Var}(m_D(C_t) \mid \mathcal{G}_{t-1})$ be the variance of the selected map's miscoverage. Conditional on $\mathcal{G}_{t-1}$, this miscoverage has mean q_t for adjacent randomized ranks, and

$$v_t = \frac{q_t(1 - q_t)}{t + 1} + \frac{\theta_t(1 - \theta_t)}{t(t + 1)}, \qquad \theta_t = t(1 - q_t) - \lfloor t(1 - q_t) \rfloor.$$

For a deterministic order index $J_t \in \{0, \ldots, t\}$, with indices $0, t$ denoting the empty-set and full-set maps, $q_t = 1 - J_t/t$ and $v_t = q_t(1 - q_t)/(t + 1)$. For population thresholds, use $v_t = 0$, since $m_D(C_t) = q_t$. Thus

$$\mathcal{R}_T = \lambda \mathbb{E} G_T + \tfrac{1}{4} \sum_{t \le T} \mathbb{E}\{(q_t - \alpha)^2 + v_t\}.$$

The prefix objective maximizes this expectation over every integer prefix. Figure 3a fixes Counting's announced horizon at 16,380. Its term for size above the frontier satisfies $\mathbb{E}[\epsilon_t \mid \mathcal{G}_{t-1}] = \theta_t(1 - \theta_t)/[2t(t + 1)]$. Subtracting it from $\frac{1}{4}\{(q_t - \alpha)^2 + v_t\}$ gives the deviation term in Theorem 1.

Isolating threshold estimation. For Counting and Barrier, empirical and known-F thresholds have the same law for M_t, expected slack, and squared request deviations. Their difference in expected terminal cost is exactly $\frac{1}{4} \sum_{t \le T} \mathbb{E} v_t$, between zero and $\mathsf{h}_T/16$ by Theorem 2. At $T = 16{,}380$, this gives 0.141045 for Counting and 0.144278 for Barrier, against the upper bound 0.642566. This comparison concerns expected terminal costs, not retrospective maxima. At $T = 262{,}144$, Barrier's estimation cost is 0.206571, or 0.47% of its total expected cost.

Score-stream estimates. An independent implementation checks Counting on 512 score streams per target through 16,380, using raw scores and independent rank coins. Let τ_s be the selected threshold at round s. Empty-set and full-set maps have miscoverage one and zero, respectively. At a fixed prefix t, we use

$$\widehat{R}_t = \lambda G_t + \tfrac{1}{4} \sum_{s \le t} (m_D(C_s) - \alpha)^2, \qquad m_D(C_s) = 1 - F(\tau_s)$$

for finite thresholds, with the stated endpoint conventions. Its expectation equals that of $R_t = W_t - tj(\alpha)$ because $R_t - \widehat{R}_t = \lambda \sum_{s \le t} (E_s - m_D(C_s))$ has mean zero. Only evaluation uses

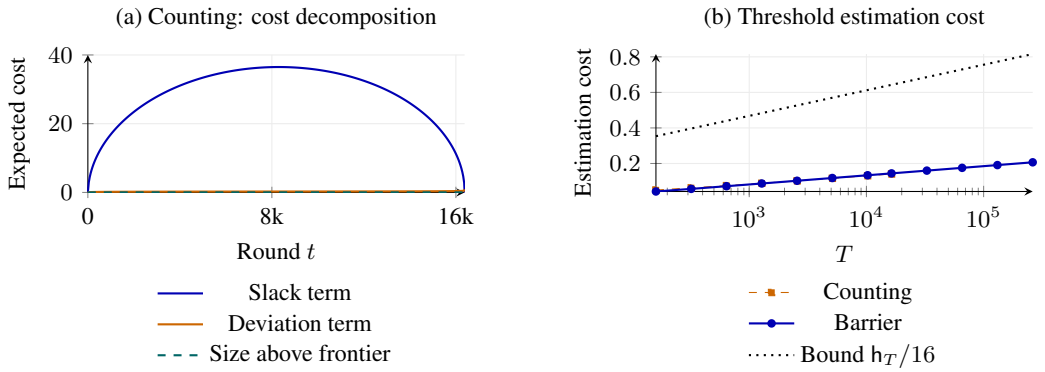


Figure 3: Direct checks of Theorems 1 and 2 in the continuous model, $\alpha = 0.1$. (a) Counting with announced horizon 16,380. Slack explains its interior peak. The terminal terms are 0, 0.8745, and 0.0086. (b) Difference in expected terminal cost when each controller replaces known-F thresholds by Algorithm 3. Counting uses each plotted horizon as its endpoint. All values are numerical expectations, not retrospective maxima. The bound is not fitted to these results.

F. This variance reduction applies at fixed prefixes. Maxima of $\widehat{R}_t$ do not estimate maxima of R_t. Pointwise Monte Carlo intervals are mean ± 1.96 standard errors (SEs) over independent streams. Methods and checkpoints on the same streams are correlated.

ACI and the quota guard. Let a_t be ACI's virtual target, γ_t its step size, and E_t^{virtual} its proposal's miscoverage indicator before the guard. Starting from $a_1 = \alpha$, ACI updates $a_{t+1} = a_t + \gamma_t(\alpha - E_t^{\text{virtual}})$ without projection (Gibbs & Candès, 2021). For $0 < a_t < 1$, its proposal uses order index $\lceil (t-1)(1-a_t) \rceil$ among past scores, with inclusive membership. An empty history or $a_t \leq 0$ gives the full-set map, and $a_t \geq 1$ gives the empty-set map. The guard enforces the same prefix quota by issuing the full set when $M_{t-1} + 1 > b_t$, and issuing the proposal otherwise. After observing the outcome, we can determine the proposal's miscoverage indicator even when the guard replaced it. The count M_t includes only miscoverages of the issued sets. The population quota guard instead proposes the known-F threshold at α.

F.2 The three temporal criteria

Expected-prefix cost takes the maximum of the expected cumulative excess size over prefixes. Retrospective cost instead takes the largest realized excess size on each stream and then averages across streams. We compute these two criteria separately for the same controllers.

Figure 1 uses the model and targets above. The longer study extends Barrier, the sine variant, and the population quota guard to 262,144 rounds. Checkpoints are the Counting horizons above, followed by 32,768, 65,536, 131,072, and 262,144. The 512 independent score streams per target are shared by Barrier and the sine variant with population and empirical thresholds, and by the population quota guard. Rank coins are independent. Figure 1 shows only the default Barrier. Figure 4 also compares the sine variant. We chose this extension after inspecting shorter runs.

Propagating controller-state probabilities computes $\max_{t \leq T} \mathcal{R}_t$. The retrospective estimate averages $\max_{1 \leq t \leq T} R_t$ over simulated score streams, using every prefix without clipping at zero. For the population quota guard, an ordinary round has prediction-set size $j(\alpha)$ and a forced full round has size one. Its realized excess cost is therefore nondecreasing, so its terminal mean estimates both prefix criteria. All five methods satisfy every tested quota. Figure 4 divides these costs by the theoretical orders. These normalized curves still vary with T, so the simulations do not establish asymptotic scaling.

F.3 Numerical oracle benchmarks

These numerical benchmarks know the distribution and horizon and may choose any population threshold. Using double precision, we compute the terminal optimum and lower and upper bounds

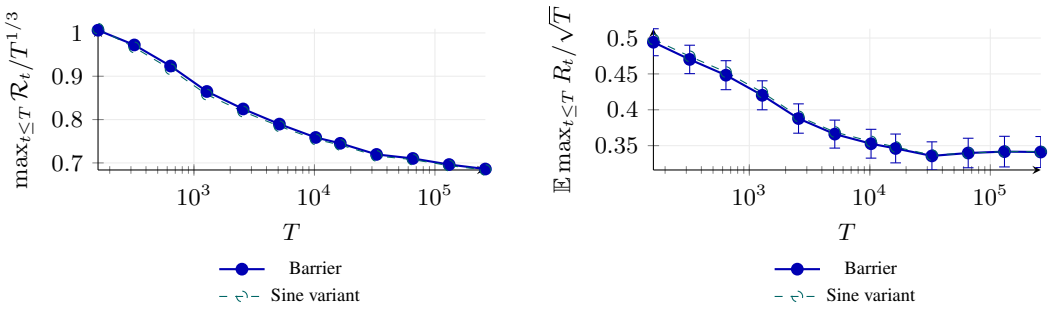


Figure 4: Default Barrier and the sine variant, both using Algorithm 3, at $\alpha = 0.1$. The left panel uses numerical expectations at every prefix. The right panel averages raw per-stream maxima. Even after cube-root normalization, the curves still drift at these horizons. Bars are pointwise approximate 95% Monte Carlo intervals.

Table 2: Both prefix criteria at all targets, $T = 262{,}144$. Expected-prefix maxima for Barrier and the sine variant are computed by state propagation. Quota-guard costs and retrospective maxima are raw trajectory estimates, mean $\pm$ Monte Carlo SE over 512 streams. Barrier and the sine variant use Algorithm 3 at the default width. The quota guard knows the distribution.

α	Method	$\max_{t\le T} \mathcal{R}_t$	$\mathbb{E} \max_{t\le T} R_t$
0.05	Barrier	42.67	130.45 ± 3.70
0.05	Sine variant	42.56	131.35 ± 3.68
0.05	Quota guard	114.74 ± 3.56	114.74 ± 3.56
0.10	Barrier	43.90	174.65 ± 5.64
0.10	Sine variant	43.76	174.94 ± 5.63
0.10	Quota guard	159.55 ± 5.51	159.55 ± 5.51
0.20	Barrier	47.89	212.77 ± 6.45
0.20	Sine variant	47.75	212.64 ± 6.49
0.20	Quota guard	201.18 ± 6.63	201.18 ± 6.63

on the expected-prefix optimum. These computations compare performance at finite horizons. They are not used in the analytic rate proofs in Appendices B–D. The reported intervals do not account for floating-point roundoff.

Terminal optimum. Let $\mathsf{V}_t(m)$ be the minimum expected remaining excess size after t rounds and m miscoverages, with $\mathsf{V}_T(m) = 0$ at feasible terminal states. The allowable requests are $\mathcal{Q}_t(m) = [0, 1]$ if $m + 1 \le \lfloor \alpha t \rfloor$, and $\mathcal{Q}_t(m) = \{0\}$ otherwise. The terminal optimum in Figure 1a is $\mathsf{V}_0(0)$ from

$$\mathsf{V}_{t-1}(m) = \min_{q \in \mathcal{Q}_t(m)} \{j(q) - j(\alpha) + (1 - q)\mathsf{V}_t(m) + q\mathsf{V}_t(m + 1)\}.$$

The recursion adds this round's excess size to the average remaining cost after the two possible outcomes. Starting from the terminal condition, it solves earlier rounds in reverse order. Infeasible branches are omitted. Write $[x]_0^1 = \min\{1, \max\{0, x\}\}$ for clipping to $[0, 1]$. When a miscoverage is permitted, the minimizer is $[5/2 - 2\{\mathsf{V}_t(m + 1) - \mathsf{V}_t(m)\}]_0^1$. Otherwise $q = 0$. We recompute it at each Counting horizon and target, without requiring the final miscoverage count to fill the quota.

Prefix optimum. At $\alpha = 0.1$, we bracket $U_T^* = \inf_{\mathcal{A} \text{ feasible}} \max_{t\le T} \mathcal{R}_t$. A weighted average of expected prefix costs cannot exceed their maximum, so minimizing such an average gives a lower bound. The expected-prefix cost of any feasible policy gives an upper bound. For nonnegative weights w_t summing to one, $L(w) := \inf_{\mathcal{A} \text{ feasible}} \sum_t w_t \mathcal{R}_t \le U_T^*$. Write $\beta_s = \sum_{t=s}^T w_t$. The same recursion computes $L(w)$ with stage cost $\beta_s\{j(q) - j(\alpha)\}$. When $\beta_s > 0$ and a miscoverage is permitted, the minimizer is $[5/2 - 2\{\mathsf{V}_s(m+1) - \mathsf{V}_s(m)\}/\beta_s]_0^1$. Take $q = 0$ when miscoverage is forbidden. Zero-weight stages have linear objectives. Convexity lets us average actions at the same round and miscoverage count without increasing cost, so the recursion covers all feasible history-dependent policies.

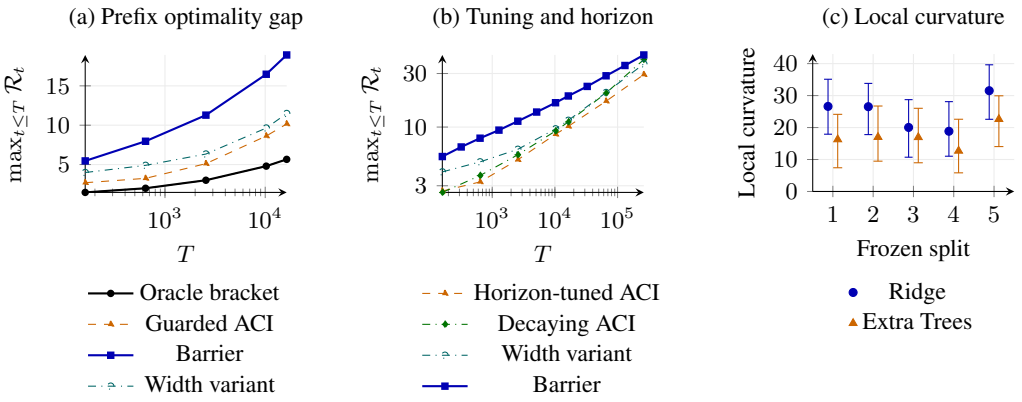


Figure 5: Finite-horizon comparisons at $\alpha = 0.1$. Barrier uses Algorithm 2's default width. Width variant uses $L_t = \max\{5, 4 + \lceil t^{1/3}/4 \rceil\}$. Both use Algorithm 3. (a) Expected-prefix costs and numerical optimum bracket. (b) All ACI curves enforce the same quota. Fixed steps are selected by horizon. The decay multiplier and variant width are selected once. Decaying ACI uses 2,048 streams (SEs in Table 3). Other costs are numerical expectations. (c) Local fits on California Housing, bandwidth 0.04, with conditional 95% paired-row bootstrap intervals (400 resamples). Positive fitted curvature does not establish population target efficiency.

Each recursion returns a feasible policy $\mathcal{A}^{(i)}$. Let $r_t^{(i)} := \mathcal{R}_t(\mathcal{A}^{(i)}, D)$ be its expected excess size at prefix t, and let z_i denote the probability of selecting this policy. A linear program minimizes $\max_t \sum_i z_i r_t^{(i)}$ over nonnegative z_i summing to one. Selecting one policy before deployment preserves every quota and gives an upper bound. The linear program also identifies time weights for the next recursion. Where L is differentiable, we refine positive time weights using $\partial L / \partial w_t = \mathcal{R}_t$. Every integer prefix is a constraint. At $T = 16{,}380$, the lower and upper values are 5.654720 and 5.659844, a 0.091% gap.

An independent linear program on a 0.05 probability grid at $T \in \{10, 20, 40, 80\}$ lies within the bracket plus $T(0.05)^2/16$. Mixing adjacent grid probabilities preserves transitions and adds at most $(0.05)^2/16$ per round. Recovered policies agree with their values from the recursion within 6×10^{-11}.

F.3.1 Cost decomposition and strictly positive slack

We decompose expected cumulative excess size into the slack and deviation terms, then measure the extra cost of requiring strictly positive slack.

Figure 6 propagates the 25 positive-weight policies in the saved mixture at $\alpha = 0.1, T = 16{,}380$ and compares them with the width variant at $L_t = \max\{5, 4 + \lceil t^{1/3}/4 \rceil\}$. Both use population thresholds. Theorem 1 gives

$$\mathcal{R}_t = \lambda \mathbb{E} G_t + \tfrac{1}{4} \mathbb{E} \sum_{s \le t} (q_s - \alpha)^2.$$

Both terms are computed at every integer prefix and averaged over the initial mixture index. Their independently computed sum agrees with cost to below 10^{-8}. This checks the numerical computation. The identity itself follows from Theorem 1.

Let $U_T^{>,*}$ be the expected-prefix optimum with $G_t > 0$ at every round. At $\alpha = 0.1$, this requires $M_t \le \lfloor (t-1)/10 \rfloor$ and a full first prediction. Deleting that round leaves an original-quota policy of length $T - 1$. Prepending a full prediction gives the converse. In the known-law i.i.d. model the discarded observation supplies only independent randomness, already permitted by the comparator. Nonnegative expected excess at each feasible prefix then gives

$$U_T^{>,*} = B - j(\alpha) + U_{T-1}^{*}. \tag{34}$$

At this target, requiring positive slack is equivalent to one initial full prediction followed by the original problem with one fewer round. For $T - 1$, truncating and renormalizing the saved time weights

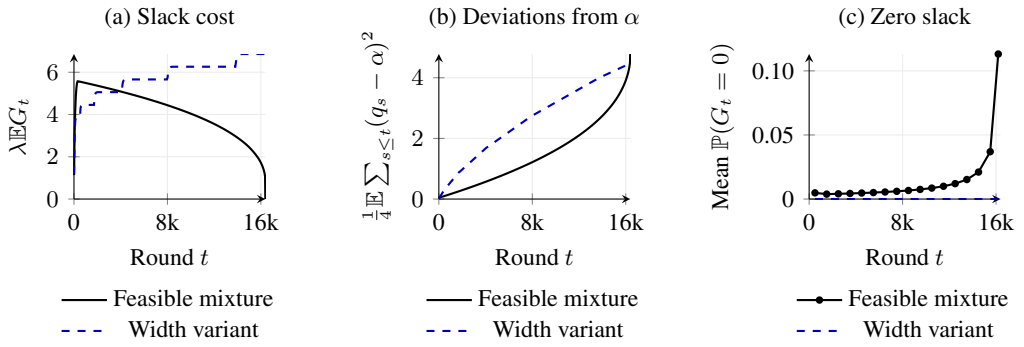


Figure 6: Structure of the recovered near-optimal prefix policy at $\alpha = 0.1, T = 16{,}380$, with population thresholds. (a)–(b) The sum is expected excess size. For readability, plotted checkpoints are quota increments. Numerical maxima use all integer prefixes. (c) Mean zero-slack probability over eligible prefixes in successive 1,000-round blocks (the final block has 380 rounds). Terminal probability is one for the mixture and zero for the width variant, which uses $L_t = \max\{5, 4 + \lceil t^{1/3}/4 \rceil\}$. These are propagated expectations.

gives a lower bound, and truncating the original mixture gives an upper bound. Consequently,

$$U_T^{>,*} \in [5.777220, 5.782344], \qquad U_T^{>,*} - U_T^* \in [0.117376, 0.127624].$$

The restriction therefore explains at most $0.127624/(11.336362 - 5.659844) = 2.25\%$ of the width variant's gap. Optimizing over all policies with strictly positive slack does not separate the remaining effects of horizon knowledge, width schedule, and potential shape.

F.4 Guarded ACI and controller selection

We compare step-size and width choices in the continuous model at $\alpha = 0.1$, with the quota guard above. Fixed ACI steps are selected at each horizon from $\{0.001, 0.005, 0.02, 0.08\}$ using 64 independent validation rank streams. The criterion is the maximum, over every prefix, of the validation-stream mean of

$$\lambda G_t + \tfrac{1}{4}\sum_{s \le t}\{(q_s - \alpha)^2 + v_s\}.$$

Uniform insertion ranks and independent rounding coins reproduce the rank-history law. Choices are frozen before independent evaluation. Maximizing after averaging distinguishes this criterion from retrospective cost. Selecting at each horizon does not give one horizon-free controller.

The *width variant*, $L_t = \max\{5, 4 + \lceil t^{1/3}/4 \rceil\}$ in Figure 5, was selected at $T = 16{,}380$ in a separate 64-stream validation study. For offsets $b \in \{0, 4\}$ and multipliers $c \in \{1/8, 1/4, 1/2, 1, 2\}$, candidates were $L_t = \max\{5, b + \lceil ct^{1/3} \rceil\}$. Both the terminal and expected-prefix validation objectives selected $(b, c) = (4, 1/4)$. We chose the grid after inspecting the default Barrier. The selected width is kept at all horizons and evaluated on 256 independent paired test streams.

To compute fixed-step ACI expectations, we track the joint distribution of the proposal miscoverage count $N_t := \sum_{u \le t} E_u^{\text{virtual}}$ and the issued-set miscoverage count M_t. The guard gives

$$N_t - M_t = \max\{N_{t-1} - M_{t-1}, N_t - b_t\}, \qquad N_0 = M_0 = 0.$$

The empirical order index depends on (t, N_{t-1}), while M_{t-1} determines whether the guard issues the full set. Order-statistic moments then give the expected issued size. States with probability below 10^{-22} are omitted in the longer runs. The recorded lost mass δ gives cost error at most $2T\delta$, below 7.5×10^{-11} in every reported calculation.

Decaying step. For the same unprojected ACI update, let c be the step multiplier in $\gamma_t = ct^{-1/3}$ (Angelopoulos et al., 2024a). At the single validation horizon 16,380, 256 independent rank streams select $c = 0.4$ from $\{0.025, 0.05, 0.1, 0.2, 0.4, 0.8, 1.6\}$ by the same expected-prefix criterion. We

Table 3: One frozen decaying-step guarded ACI schedule, $\gamma_t = 0.4t^{-1/3}$, selected on 256 validation streams at $T = 16{,}380$. Decaying ACI uses 2,048 new test streams. Parentheses give batch-bootstrap Monte Carlo SEs of the expected-prefix cost estimate. Fixed-step ACI and the width variant use numerical expectations. The fixed step is selected separately by horizon.

T	Decaying ACI	Selected fixed-step ACI	Width variant
160	2.617 (0.024)	2.691	3.986
640	3.699 (0.023)	3.246	4.898
2,560	5.664 (0.028)	5.114	6.335
10,240	9.228 (0.039)	8.627	9.639
16,380	11.116 (0.039)	10.158	11.484
65,536	20.100 (0.040)	17.008	20.308
262,144	39.977 (0.050)	29.362	36.525

freeze it for $2{,}048$ new paired test streams through $262{,}144$, without retuning at checkpoints. The prefix cost is estimated by maximizing the test-stream mean over every integer prefix. SEs use 400 bootstrap resamples of 64 independent batch-mean curves, with 32 streams per batch. They measure simulation variability without removing the upward bias from maximization. This miscoverage-level update is an empirical comparator, not an application of a rate theorem for threshold updates.

F.5 California Housing

We examine the relation between interval length and miscoverage, then compare online interval lengths under the quota using the same frozen predictors.

The California Housing dataset has $20{,}640$ observations and eight features. Seeds $71000, \ldots, 71004$ give five frozen splits, each with $10{,}320$ training, $5{,}160$ validation, and $5{,}160$ test observations, disjoint within a split. The two predictors are standardized Ridge with penalty one and Extra Trees with 128 trees and minimum leaf size five. Standardization and fitting use training data only. The response is the recorded capped house value divided by six, with imposed domain $[0, 1]$. Predictions are clipped to this domain. Models stay fixed during diagnostics, controller selection, and online runs.

Empirical curves. For held-out response Y_i, prediction c_i, and score $S_i = |Y_i - c_i|$, threshold $\tau \geq 0$ issues $[c_i - \tau, c_i + \tau] \cap [0, 1]$. On n held-out rows we compute

$$\widehat{V}(C_\tau) = \frac{1}{n}\sum_{i=1}^{n}\{\min(\tau, c_i) + \min(\tau, 1 - c_i)\}, \qquad \widehat{m}(C_\tau) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\{S_i > \tau\}.$$

We retain every distinct observed score as a threshold, include the empty-set map, remove larger-size points at identical miscoverage, and take the lower convex envelope $\widehat{j}$. Write $\widehat{h}$ for linear interpolation of the miscoverage–length points. The gap $\widehat{h}(\alpha) - \widehat{j}(\alpha)$ measures empirical length above the frontier. The full-set map remains available to the guard even when dominated under the empirical law. This envelope compares only thresholds of the frozen score and their mixtures, not all measurable prediction maps.

Curvature is the second derivative of a least-squares quadratic fitted to $\widehat{h}$ at 81 equally spaced levels around α. For nominal bandwidth $b \in \{0.02, 0.04, 0.06\}$, the fitting bandwidth is $\min\{b, 0.8\alpha\}$, with $b = 0.04$ primary. We evaluate $\alpha \in \{0.05, 0.1, 0.2\}$ on validation and test separately. Four hundred paired-row bootstrap resamples recompute the curves and envelopes. The percentile ranges condition on the fitted model, treating held-out rows as sampling units. They omit training uncertainty and geographic dependence. A positive local fit does not establish population target efficiency or global quadratic growth, and the empirical envelope itself is polygonal.

Table 4 reports the target gaps and bootstrap ranges. At $\alpha = 0.1$, gaps range from 0 to 0.00293 for Ridge and from 0.000033 to 0.00183 for Extra Trees. The ranges describe sampling variation in the curves and envelopes, not a calibrated test of target efficiency. Estimating the envelope can shift the bootstrap distribution: for Extra Trees split 2, the point estimate lies below the percentile range.

Table 4: Target-gap diagnostics on the five frozen test splits at $\alpha = 0.1$. Entries are in units of 10^{-3} normalized length. Ranges are the 2.5th and 97.5th percentiles of 400 paired-row bootstrap resamples, conditional on each fitted predictor. They do not establish target efficiency for the population frontier.

	Ridge		Extra Trees	
Split	Gap	Percentile range	Gap	Percentile range
1	0.000	[0.000, 5.657]	1.728	[0.063, 3.996]
2	1.457	[0.000, 3.261]	0.033	[0.052, 3.916]
3	2.932	[0.030, 6.563]	1.832	[0.017, 7.021]
4	0.904	[0.094, 4.198]	0.214	[0.170, 4.804]
5	2.902	[0.035, 5.313]	0.492	[0.005, 2.328]

Small estimates and positive lower endpoints do not establish whether target efficiency holds. These pointwise gaps cannot replace the uniform Δ in Appendix E.

Online evaluation. Each run permutes all $5{,}160$ observations without replacement and starts with an empty score history. For each model, split, and target, 16 validation permutations select the ACI step from $\{0.001, 0.005, 0.02, 0.08\}$ and the quadratic width multiplier c from $\{1/8, 1/4, 1/2, 1\}$ by mean issued length. The latter uses $L_t = \max\{5, 4 + \lceil ct^{1/3} \rceil\}$. Choices are frozen for 32 paired test permutations. Test observations are excluded from fitting and selection. Zero-radius intervals are zero-length singletons, distinct from empty outputs. Score ties are included.

Figure 2b averages paired length differences within each split. Its SE is their sample standard deviation over 32 permutations divided by $\sqrt{32}$. Percentage reductions average the five split-specific relative reductions equally. Uncertainty concerns ordering and controller randomness on fixed test sets. Repeated permutations and overlapping training splits are not independent population replications. Every tested prefix for the guarded methods and Barrier satisfies the quota.

Matching the rank implementation. Alongside original guarded ACI, a variant passes its clipped request $q_t = \min\{1, \max\{0, a_t\}\}$ through Algorithm 3. Its virtual target a_t remains unprojected. The update uses the unguarded proposal's miscoverage indicator even when the guard issues the full set. Both variants use the same frozen pools and paired permutations, with independent rank coins. We test matched ACI at the original selected step and at its own validation choice from the same grid, frozen before testing. At $\alpha = 0.1$, both select $\gamma = 0.005$ in all ten model/split pairs, so these comparisons coincide. The fixed-step comparison changes both the quantile implementation and the resulting virtual targets. Requests are not held fixed.

Table 5 shows that matching ranks slightly increases ACI's mean length. Barrier's reduction relative to matched ACI averages 1.68% for Ridge and 2.21% for Extra Trees across splits. All methods have similar miscoverage, and Barrier's empty-output frequency is below 0.003%. For each length difference, the pooled conditional SE is the square root of the sum of five squared within-split SEs, divided by five. Intervals are the mean ± 1.96 SE. They concern randomization on the frozen data, not population performance. The default reproduction run generates both step-selection comparisons at all three targets and saves all validation and test runs.

Paired quota-guard ablation. We reuse the same five frozen splits, both predictors, and all 32 test permutations at $\alpha = 0.1$. Each ACI variant uses its previously selected $\gamma = 0.005$. No model, parameter, or test ordering is reselected. Because the guarded method updates its virtual state with the proposal's miscoverage indicator, removing the guard leaves the proposal sequence unchanged. We therefore record two issued streams for identical proposals, feedback, and, for matched ACI, rounding coins. Only one stream replaces proposals by the full set when another miscoverage would violate the quota. Their cumulative length difference equals the sum of full-set enlargements on guarded rounds, including zero increments when the proposal is already full.

For both variants and predictors, every unguarded run violates some prefix, despite terminal miscoverage near 0.1. Peak excess is $\max_{t \le T}(M_t - b_t)_+$: its mean is about eight miscoverages and its maximum is twelve. The guarded streams and Barrier have zero violations. For matched ACI,

Table 5: California Housing at $\alpha = 0.1$. Both ACI variants are quota-guarded. ACI + ranks uses Algorithm 3, and Barrier uses its default width. All ten model/split pairs select $\gamma = .005$ for both ACI variants, so the fixed-step and separately tuned comparisons coincide. Entries average five frozen splits with 32 paired test permutations each. Length is normalized. Miscoverage, full, and empty frequencies are percentages. Differences subtract ACI + ranks, with conditional 95% ordering intervals excluding training and population uncertainty. Every tested prefix is feasible.

Predictor	Method	Length	M_T/T	Full	Empty	Difference [95% interval]
Ridge	Original ACI	0.3644	9.84	1.68	0.0000	-0.0013 $[-0.0016, -0.0010]$
Ridge	ACI + ranks	0.3657	9.84	1.94	0.0000	—
Ridge	Barrier	0.3595	9.79	0.46	0.0013	-0.0062 $[-0.0068, -0.0055]$
Extra Trees	Original ACI	0.2910	9.84	1.69	0.0000	-0.0011 $[-0.0015, -0.0007]$
Extra Trees	ACI + ranks	0.2921	9.84	1.89	0.0000	—
Extra Trees	Barrier	0.2856	9.79	0.52	0.0024	-0.0064 $[-0.0071, -0.0058]$

Table 6: Paired quota-guard ablation on California Housing ($\alpha = 0.1$, $T = 5{,}160$). Each entry averages five frozen splits and 32 test permutations per split. ACI pairs use identical proposals and the previously selected $\gamma = 0.005$. The M_T/T column reports terminal miscoverage. Violations count runs with any infeasible prefix. Max excess is the largest $(M_t - b_t)_+$ over all tested rounds and runs.

Predictor	Method	Length	M_T/T (%)	Violations (%)	Max excess
Ridge	ACI	0.3542	10.00	100.0	12
Ridge	ACI + guard	0.3644	9.84	0.0	0
Ridge	Barrier	0.3595	9.79	0.0	0
Extra Trees	ACI	0.2795	10.00	100.0	12
Extra Trees	ACI + guard	0.2910	9.84	0.0	0
Extra Trees	Barrier	0.2856	9.79	0.0	0

the guard adds 0.01127 normalized length for Ridge and 0.01236 for Extra Trees, with paired conditional 95% intervals $[0.01071, 0.01182]$ and $[0.01177, 0.01296]$. For original ACI, the guard adds 0.01017 and 0.01150 normalized length, with paired conditional 95% intervals $[0.00962, 0.01072]$ and $[0.01087, 0.01212]$, respectively (Table 6). The paired intervals use the same within-split ordering SEs and pooling rule as above. These differences isolate the guard for the fixed implementation, not the infimum over feasible policies defining the price.

All $4{,}160$ candidate/repetition rows for the original guarded implementations reproduce exactly. Independent sorted-list replay checks complete first-permutation traces for Ridge split 1 and Extra Trees split 5, including the proposal and issued-set miscoverage indicators, every prefix, and both guarded and unguarded lengths. The anonymous code package contains the protocol and scripts. The default run saves input hashes, per-stream results, paired summaries, and verification results.

Empty outputs. Counting uses empty sets to complete its final count, and Barrier can use them to reduce large slack. Our size objective assigns them zero cost. Their low observed frequency does not establish rates under a useful-output restriction. Such a restriction must specify the allowed family or a minimum useful size. Nonempty singletons can still have zero length.

Table 7: Full paired guard ablation at $\alpha = 0.1$. The two ACI variants each use $\gamma = 0.005$, fixed by prior validation. Peak excess means $\max_t(M_t - b_t)_+$. Its mean and maximum are taken across the 160 runs per predictor. Full and empty frequencies are percentages of rounds.

Predictor	Method	Length	M_T/T (%)	Mean peak	Max peak	Full (%)	Empty (%)
Ridge	ACI	0.3542	10.00	8.16	12	0.02	0.0000
Ridge	ACI + guard	0.3644	9.84	0.00	0	1.68	0.0000
Ridge	ACI + ranks	0.3544	10.00	8.06	12	0.09	0.0016
Ridge	ACI + ranks + guard	0.3657	9.84	0.00	0	1.94	0.0000
Ridge	Barrier	0.3595	9.79	0.00	0	0.46	0.0013
Extra Trees	ACI	0.2795	10.00	8.41	12	0.02	0.0000
Extra Trees	ACI + guard	0.2910	9.84	0.00	0	1.69	0.0000
Extra Trees	ACI + ranks	0.2797	10.00	8.22	12	0.09	0.0028
Extra Trees	ACI + ranks + guard	0.2921	9.84	0.00	0	1.89	0.0000
Extra Trees	Barrier	0.2856	9.79	0.00	0	0.52	0.0024

G Comparison with related control formulations

Hard prefix constraints and cost–buffer tradeoffs appear in control and resource allocation. Table 8 compares the information available at each decision, the comparator, and when cost is evaluated.

Constrained online convex optimization also studies regret–violation tradeoffs (Guo et al., 2022; Yi et al., 2021; Sinha & Vaze, 2024). These works bound cumulative violations. The "hard" metric in Guo et al. (2022) sums positive violations, so conservative rounds cannot cancel them. This differs from satisfying the quota at every realized prefix. We measure the cost relative to the population coverage–efficiency frontier, even when the distribution is known.

Table 8: Closest control formulations and their evaluated objectives.

Work	Constraint and decision information	Comparator and time objective
Berry & Gallager (2002), Propositions 4.2, 4.5	Nonnegative queue. Current channel and buffer observed. Stationary policies under its regularity assumptions.	Long-run average power above the unconstrained power–rate frontier, with an inverse-square gap as average delay grows.
Chen & Wang (2025), Theorem 3.1	Fixed initial inventory. Current reward and demand observed. Distribution and horizon known.	Expected terminal reward gap to a hindsight linear program. Rates depend on the request distribution.
Chen et al. (2026), Theorems 5.3, 5.8	Inventory nonnegative at every prefix. Current reward, demand and replenishment observed. Unknown distribution.	Expected terminal reward gap to a prefix-feasible hindsight integer program. Logarithmic and square-root regimes on different instance classes.
McMahan & Zhu (2024), Theorem 3	Almost-sure prefix budgets. Action precedes realized cost. Policies may use cumulative cost history.	Optimal expected total reward among feasible policies over a finite horizon. State augmentation supports planning and learning.

Here the prediction-set rule is selected before the fresh observation, and the comparator is $tj(\alpha)$. Changing the time criterion under the same geometric conditions yields the three orders in Theorem 4. Lemma 6 bounds squared probability adjustments on each window for every feasible policy, including those using the full history and knowing D, T. Summing over disjoint windows gives the expected-prefix lower bound without requiring stationary or independent actions. Empirical ranks implement the upper bounds without knowing the score distribution.

H A fixed additional miscoverage allowance

A fixed allowance permits additional miscoverages while retaining the same comparator and expected-prefix cube-root order.

For an integer $k \geq 0$, let $\mathcal{A}_k(T)$ be the policies satisfying $M_t \leq b_t + k$ for every $t \leq T$, on every stream and every realization of their randomization. Retain $\mathcal{R}_t = \mathbb{E}W_t - tj(\alpha)$ and $G_t = \alpha t - M_t$.

Proposition 2 (Expected-prefix cost with a fixed allowance)**.** *Under the assumptions of Theorem 3, every $\mathcal{A} \in \mathcal{A}_k(T)$ satisfies*

$$\max_{1 \leq t \leq T} \mathcal{R}_t(\mathcal{A}, D) \geq c_2 T^{1/3} - 1 - 2\lambda k,$$

where c_2 is the coefficient in the proof of Theorem 3(ii), independent of k. Under the assumptions of Theorem 4, for each fixed integer $k \geq 0$,

$$\inf_{\mathcal{A} \in \mathcal{A}_k(T)} \max_{1 \leq t \leq T} \mathcal{R}_t(\mathcal{A}, D) = \Theta(T^{1/3}) \qquad (T \longrightarrow \infty).$$

Proof. Write $\bar{q}_t = \mathbb{P}(E_t = 1 \mid E_1, \ldots, E_{t-1})$. Equation (27) still holds, and feasibility gives $G_s + k \geq 0$. Thus, for $1 \leq s \leq T$,

$$\mathcal{R}_s + \lambda k \geq \lambda \mathbb{E}(G_s + k) + c_- \mathbb{E} \sum_{t \leq s} (\bar{q}_t - \alpha)^2.$$

In particular, $\max_{t \leq T} \mathcal{R}_t \geq -\lambda k$. Put $R = \max_{t \leq T} \mathcal{R}_t + 2\lambda k \geq \lambda k \geq 0$. Then $\mathbb{E}(G_s + k) \leq R/\lambda$ for $0 \leq s \leq T$, including $s = 0$ since $G_0 + k = k$, and $c_- \mathbb{E} \sum_{t \leq T} (\bar{q}_t - \alpha)^2 \leq R$.

For a window $s+1, \ldots, s+n$ contained in $1, \ldots, T$, feasibility gives $M_{s+n} - M_s \leq \alpha n + G_s + k$. The proof of Lemma 6, with $a = \alpha(1-\alpha)/4$, gives

$$\mathbb{E} \sum_{t=s+1}^{s+n} (\bar{q}_t - \alpha)^2 \geq \frac{(a\sqrt{n} - \mathbb{E}(G_s + k))_+^2}{n}.$$

Choose $n = \lceil (1 + 2R/(a\lambda))^2 \rceil$ and $d = c_- a^2/4$. Since $\mathbb{E}(G_s + k) \leq a\sqrt{n}/2$, each complete window contributes at least $a^2/4$ in expected squared deviations. Summing disjoint windows yields $R \geq d\lfloor T/n \rfloor$ and $T < n(1 + R/d)$. These inequalities also hold when $n > T$. With $A = 2\max\{1, 2/(a\lambda)\}^2$, we have $n \leq A(1+R)^2$, and therefore $T \leq A(1 + 1/d)(1+R)^3$. Thus $R \geq c_2 T^{1/3} - 1$, with the same $c_2 = [A(1 + 1/d)]^{-1/3}$ as before. Substituting the definition of R proves the lower bound.

The original Barrier policy belongs to $\mathcal{A}_k(T)$ and gives $\max_{t \leq T} \mathcal{R}_t = O(T^{1/3})$ under the upper-bound assumptions. For fixed k, the additive $2\lambda k + 1$ in the lower bound is eventually dominated by $c_2 T^{1/3}$, proving the asserted order. □

Relative to $tj(\alpha)$, additional miscoverages may make excess size negative. The decomposition gives only $\mathcal{R}_t \geq -\lambda k$. This result fixes k before $T \to \infty$. Its finite-horizon lower bound is uniform in k, which also gives the following consequence.

Corollary 2 (Initial allowances of order $o(T^{1/3})$)**.** *Under the assumptions of Theorem 4, let k_T be nonnegative integers with $k_T = o(T^{1/3})$. If the allowance k_T is available from the first round and is held fixed throughout the run, then*

$$\inf_{\mathcal{A} \in \mathcal{A}_{k_T}(T)} \max_{1 \leq t \leq T} \mathcal{R}_t(\mathcal{A}, D) = \Theta(T^{1/3}).$$

Proof. Substitute $k = k_T$ into the uniform lower bound of Proposition 2. Its subtractive term $1 + 2\lambda k_T = o(T^{1/3})$ cannot change the order. The original Barrier is feasible for every nonnegative allowance and supplies the upper bound, with constants independent of k_T. □

For example, $k_T = \lfloor \kappa T^\beta \rfloor$ with fixed $\kappa > 0$ and $0 \leq \beta < 1/3$ satisfies this condition. An initial allowance k_T differs from an allowance k_t that changes during a run. Larger initial allowances, time-varying allowances, and their terminal and retrospective criteria require further analysis. In particular, a fixed multiple of $\sqrt{T}$ does not guarantee that every Bernoulli path stays within the enlarged quota.