An Exponential Deterministic-Randomized Separation in Oracle-Based Online Learning

Anonymous authors
Paper under double-blind review

Abstract

Randomization is known to improve oracle efficiency in online learning, but whether it can provide a provable advantage over deterministic algorithms has remained open. We resolve this question for transductive online learning of thresholds on an unknown total order with a consistency-type empirical risk minimization (ERM) oracle. For natural extremal ERM rules, including minimal-prefix and maximal-prefix selection, every deterministic learner incurs linear total cost: on some instance, the number of mistakes M and oracle calls Q satisfy $M + Q \geq T - \varepsilon$. Consequently, any deterministic learner achieving $O(\log T)$ mistakes requires $\Theta(T)$ oracle calls. In contrast, the randomized learner of Attias et al. (2025) achieves $O(\log T)$ expected mistakes and $O(\log T)$ expected oracle calls under the same fixed rule. We prove a matching $\Omega(\log T)$ expected-query lower bound, establishing the optimal randomized query complexity. We further show that the separation depends on the oracle's selection rule: a legal feasible-median rule admits deterministic $O(\log T)$ complexity, while extremal and global-median rules force linear cost. Thus, randomization yields an exponential separation in oracle complexity, while oracle selection emerges as a fundamental complexity axis.

1 Introduction

In oracle-based online learning, statistical difficulty and access to the concept class are distinct resources. Prediction mistakes measure the former; oracle calls measure the latter. A call may return a full consistent concept or only a realizability bit. Even a class of small Littlestone dimension can require many calls, so a mistake bound alone does not describe the cost of learning. This paper isolates a further distinction: for the same class and the same interface, the oracle's choice among legal answers can determine the value of randomization. Query cost is an abstract access cost, not a running-time guarantee.

Attias, Hanneke and Ramaswami (2025) (henceforth AHR25) formalize this distinction for online and transductive learning through a consistency-type ERM oracle, which given a labeled finite set S returns some concept of the class consistent with S (as a full label vector) or reports that S is not realizable, and a weak consistency oracle, which returns only the realizability bit. For thresholds on an unknown total order (their Theorem 4.5), a deterministic learner can achieve $O(\log T)$ mistakes with $O(T)$ ERM calls, and a randomized learner can achieve $O(\log T)$ mistakes with $O(\log T)$ *expected* ERM calls. The missing step is a deterministic lower bound: their results do not show that an $O(\log T)$-mistake guarantee requires linearly many calls on this family. Their concluding question is: "Are randomized algorithms provably more powerful than deterministic ones for online learning with oracles?"

The separation. Fix the minimal-prefix ERM rule. For every deterministic learner and every T some instance satisfies $\max\{M, Q\} \geq (T - \varepsilon)/2$. Under the same rule, one randomized learner achieves $O(\log T)$ expected calls and $O(\log T)$ mistakes, and this expected-query order is optimal for an $O(\log T)$-mistake guarantee (Theorems 3.1 and 3.2). The deterministic bound is pathwise; the randomized upper and lower bounds are worst-instance expectations over the learner's private randomness.

Why fixing the rule matters. On a realizable sample, the *minimal-prefix rule* returns the smallest consistent prefix of the unknown order. Once the order is fixed, it is the memoryless function $S \mapsto$

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.

$\mathcal{O}_{\min}(S, \preceq)$. The rule is announced independently of the learner; no online choice of oracle answers is needed. Theorem 3.1 gives, for each deterministic learner, an order and target with total cost at least $T - \varepsilon$, where $\varepsilon \in \{0, 1\}$ records whether the empty prefix counts as a concept. The bound is sharp and also covers the maximal-prefix rule and any sample-dependent choice between the two. Thus a deterministic $O(\log T)$-mistake guarantee requires $T - \varepsilon - O(\log T)$ calls in the worst case, whereas AHR25 obtain $O(\log T)$ expected calls with $O(\log T)$ mistakes for every legal oracle. Measured in the Littlestone dimension $d = \Theta(\log T)$ of the class this is a gap of $O(d + 1)$ versus $2^{\Omega(d)}$. Although one returned label vector can expose many relative-order relations, our lower bound charges a call for at most one point via the *one-pin common-value* property of the extremal rules (Section 4).

Selection as a complexity parameter. Holding the class fixed exposes the role of selection. Theorem 3.4 gives a legal memoryless "feasible-median" rule, selecting the median feasible cutoff, under which a deterministic learner achieves $\lceil \log_2 T \rceil$ calls and $\lceil \log_2 T \rceil$ mistakes (for $T \geq 2$). A "global-median" rule, projecting a fixed median onto the feasible interval, again forces linear total cost. Randomization overcomes *unfavourable* selection (extremal or adversarial); favourable selection already permits deterministic learning with logarithmic mistakes and calls.

Contributions and their roles. Throughout, T is the number of instances, M the number of mistakes, Q the number of oracle calls, and $\varepsilon = 0$ under convention E (the empty prefix on the instances is a concept) and $\varepsilon = 1$ under convention N (it is not); see Section 2.

- **A sharp deterministic obstruction (Theorem 3.1).** Under a pre-declared extremal rule, every deterministic learner has $M + Q \geq T - \varepsilon$ on some instance, and $\inf_{\mathscr{A}} \sup(M + Q) = T - \varepsilon$ exactly; the rule may be minimal-prefix, maximal-prefix, or any sample-dependent choice between the two.
- **Optimal randomized query order (Theorem 3.2).** On an explicit finite distribution under the minimal-prefix rule, every randomized learner satisfies $\mathbb{E}[M] \geq ((T + 1 - \varepsilon)\, 128^{-\mathbb{E}[Q]} - 1)/2$. With AHR25, the randomized ERM query complexity of $O(\log T)$ mistakes is $\Theta(\log T)$.
- **The boundary of the separation (Theorem 3.4).** The linear bound does not extend to every legal pre-declared rule: the feasible-median rule admits deterministic $\lceil \log_2 T \rceil$ calls and mistakes for $T \geq 2$, whereas the global-median rule forces $M + Q \geq T - \lfloor (\varepsilon + T)/2 \rfloor$. The exact $T - \varepsilon$ lower bound holds for the extremal rules of Theorem 3.1 and in the worst-case legal-oracle model; it is not uniform over all legal pre-declared selection rules.

Three further results refine the picture: the bound $M + Q \geq T - \varepsilon$ transfers directly to the operational form of the lower bounds in AHR25, in which the oracle's answers are chosen after the learner and then frozen into a memoryless oracle (Corollary 4.1); deterministic lower and upper bounds at fixed query budgets are exact in the small- and large-budget regimes, with finite counterexamples to the naive exact formula in between (Theorem D.1); and with only a weak consistency oracle both deterministic and randomized learners need $\Theta(T)$ calls for $O(\log T)$ mistakes, so the advantage of randomization is tied to the oracle returning an evaluable concept (Theorem E.1).

Two independent proof mechanisms. The deterministic branch (Section 4) stabilizes each full oracle answer by at most one endpoint pin, charges lost free points to mistakes or calls, and fixes the resulting transcript to one instance; within this anchored-threshold construction the rule-specific step is the one-pin common-value lemma, and the charging and fixation steps then apply to any rule satisfying this property (Proposition A.8). A cached-transcript counterpart handles adversarial oracles (Appendix B). The independent randomized branch (Section 5) uses a stronger revelation oracle, posterior invariance and an $O(1)$ bound on the expected decrease of a logarithmic potential per call. Neither uses the general lower-bound theorem of AHR25 (their Theorem 4.4); Remark 4.3 explains why. Complete proofs are in Appendices A–E.

Why the distinction matters. Oracle-efficient learning delegates access to a class to a subroutine. Our comparison shows that its selection semantics can change access complexity even with the class and feedback interface held fixed. This identifies a specification issue for oracle-based algorithm design. Connections to contextual and hybrid learning concern this modelling principle; transferring our bounds to those models would require a separate reduction.

Table 1: Query complexity of achieving $O(\log T)$ mistakes for thresholds on an unknown order, by oracle interface and selection rule; randomized budgets are worst-instance expectations over mistakes and calls. A Θ entry gives matching orders, not exact minimax values; an O entry is an upper bound only. In the worst-case row, upper bounds hold for every legal memoryless oracle, while the deterministic bound picks, per learner, an instance and a sample-only oracle fixed before the run (Corollary 4.1); the randomized bound (Theorem 3.2), proved for the minimal-prefix rule and hence for the worst case over legal memoryless oracles, fixes the instance before the learner's randomness, so it is seed-independent. Theorem 3.1 also covers the maximal-prefix rule.

Oracle interface and selection rule	Deterministic	Randomized (expected)
ERM, minimal-prefix rule (pre-declared)	$\Theta(T)$: Thm. 3.1; AHR25 Thm. 4.5(3)	$\Theta(\log T)$: Thm. 3.2; AHR25 Thm. 4.5(4)
ERM, worst-case legal memoryless oracle	$\Theta(T)$: Cor. 4.1; AHR25 Thm. 4.5(3)	$\Theta(\log T)$: Thm. 3.2; AHR25 Thm. 4.5(4)
ERM, feasible-median rule	$O(\log T)$: Thm. 3.4	$O(\log T)$: AHR25 Thm. 4.5(4)
ERM, global-median rule	$\Theta(T)$: Thm. 3.4; AHR25 Thm. 4.5(3)	$O(\log T)$: AHR25 Thm. 4.5(4)
Weak consistency (returns a realizability bit)	$\Theta(T)$: Thm. E.1; AHR25 Thm. 4.3	$\Theta(T)$: AHR25 Thm. 4.5(2), 4.3; Lemma E.5

Position within learning theory. Known-class mistake complexity supplies the benchmark (Littlestone, 1988; Ben-David et al., 1997; Hanneke et al., 2023); oracle-efficient learning asks how access to the class can be implemented (Assos et al., 2023; Kozachinskiy & Steifer, 2024). AHR25 make calls an explicit resource. We supply the fixed-rule obstruction and matching randomized lower bound for their threshold example, then isolate selection and feedback as distinct sources of complexity. Appendix F compares the precise assumptions with contextual learning, hybrid learning, regression, and order identification.

2 From mistake complexity to oracle access

Instances $x_1, \dots, x_T$ are distinct and announced in advance; the adversary fixes a total order $\preceq$ on a finite domain X and a target concept $c_z(x) = \mathbf{1}[x \preceq z]$ with $z \in X$; the threshold class is $C_\preceq = \{c_z : z \in X\}$ and the family is $\mathcal{F} = \{C_\preceq\}$. Labels arrive one per round; the learner predicts $\hat{y}_t$ before seeing $y_t = c_z(x_t)$. Restricted to the instances, every concept is the indicator of a $\preceq$-prefix.

The known-order benchmark and the two conventions. By default we use convention E: the domain is $X = \{s\} \cup \{x_1, \dots, x_T\}$, where the sentinel s precedes all instances in every order considered, so the empty prefix on the instances is a concept; there are $T + 1$ concepts and the Littlestone dimension is $\lfloor \log_2(T + 1) \rfloor$. Convention N, in which the domain is exactly the instance set $X = \{x_1, \dots, x_T\}$ (there are T concepts; AHR25 write $\lfloor \log_2 T \rfloor$), differs from E by the bookkeeping constant ε: the lower bounds of Theorems 3.1 and 3.2 and Corollary 4.1 are stated with $\varepsilon \in \{0, 1\}$ and hold under both conventions, while Theorems 3.4, D.1 and E.1 state the two conventions explicitly. All lower bounds are proved on these finite domains, and all upper bounds hold on them. The numbers of concepts differ by one between the two conventions, while their Littlestone dimensions differ by at most one; we write $d = \lfloor \log_2(T + 1 - \varepsilon) \rfloor$ for the Littlestone dimension, so that $T - \varepsilon \geq 2^d - 1$.

Oracles. A *consistency-type ERM oracle* takes any finite $S \subseteq X \times \{0, 1\}$ and returns some $c \in C_\preceq$ consistent with S, observable as a full label vector, or $\perp$ if none exists (AHR25, Def. 2.1, second variant). A *weak consistency oracle* returns only whether such a c exists (AHR25, Def. 2.2). Every call costs one, regardless of $|S|$; calls may be made at any time. The learner knows $\mathcal{F}$, X, T and the sequence, but not $\preceq$ or z. A *selection rule* is a function $\mathcal{R}(S, \preceq)$ that is a legal consistency-type ERM oracle for every order; the *minimal-prefix rule* $\mathcal{O}_{\min}(S, \preceq)$ returns the smallest consistent prefix and the *maximal-prefix rule* $\mathcal{O}_{\max}$ the largest. Degenerate queries: $S = \emptyset$ is realizable; S assigning both labels to one point is not; under E the sentinel s has label 1 in every concept.

Quantifiers: what is fixed before learning. The lower bounds in this paper come in three quantifier forms, which we fix here once and refer to throughout.

> **Fixed rule** (Theorems 3.1, 3.4): the selection rule $\mathcal{R}$ is announced first. For every deterministic learner there exist an order and a target: $\forall \mathcal{A}\ \exists(\preceq, z)$, with oracle $S \mapsto \mathcal{R}(S, \preceq)$.
>
> **Adversarial then frozen** (Corollary 4.1): for every deterministic learner there exist an order, a target and a memoryless legal oracle $O^\star$, a function of the sample alone, on which the learner is replayed: $\forall \mathcal{A}\ \exists(\preceq, z, O^\star)$. The oracle is produced after the learner. (AHR25 note that for deterministic learners an oblivious adversary is as powerful as an adaptive one; the freezing step is where this is made precise.)
>
> **Randomized learners** (Theorem 3.2): the instance (order, target, sequence, selection rule) is fixed before the learner's private randomness (oblivious adversary), and costs are worst-instance expectations over that randomness.

The fixed-rule form is the literal reading of "an ERM oracle for C", in which the selection rule precedes the learner. The randomized learner of AHR25 works for every legal oracle, so the fixed-rule form is the one that yields a separation with the selection rule fixed first.

3 Separation, optimality, and selection

Theorem 3.1 (Deterministic lower bound; pre-declared extremal rule)**.** *Let $T \geq 1$ and fix either convention. Let $\mathcal{R}$ be $\mathcal{O}_{\min}$, or $\mathcal{O}_{\max}$ (the largest consistent prefix), or any rule $\mathcal{R}_\eta(S, \preceq)$ that applies $\mathcal{O}_{\min}$ or $\mathcal{O}_{\max}$ according to a function $\eta(S)$ of the sample alone; $\mathcal{R}$ is announced before the learner and the instance.*

(a) For every deterministic learner $\mathcal{A}$ and every instance sequence there exist $\preceq^\star$ and $z^\star$ such that, with oracle $S \mapsto \mathcal{R}(S, \preceq^\star)$, $M(\mathcal{A}) + Q(\mathcal{A}) \geq T - \varepsilon$.

(b) Consequently $\inf_{\mathcal{A}} \sup_{\preceq, z} (M + Q) = T - \varepsilon$ (the upper bound needs no calls: under E any fixed prediction rule, under $\bar{N}$ always predicting 1).

Consequently an $O(\log T)$-mistake guarantee costs $T - \varepsilon - O(\log T)$ calls under one of these rules. The randomized learner of AHR25 (Thm. 4.5(4)) achieves $O(\log T)$ expected calls with $O(\log T)$ mistakes against every legal oracle, hence under the same fixed rule. The statement concerns only the instance singled out in AHR25 (thresholds on an unknown order, consistency-type ERM oracle with full label vectors, transductive protocol, the finite domains of Section 2); it does not claim a uniform improvement for arbitrary classes or for the non-transductive setting.

Theorem 3.2 (Randomized lower bound)**.** *For every $T \geq 1$ and either convention there exist a finite distribution over (total order, target), a fixed instance sequence, and the memoryless minimal-prefix oracle $\mathcal{O}_{\min}(\cdot, \preceq)$, such that every randomized learner satisfies*

$$\mathbb{E}[M] \;\geq\; \frac{(T + 1 - \varepsilon)\, 128^{-\mathbb{E}[Q]} - 1}{2},$$

the expectations being over the distribution and the learner's randomness. In particular, if $\mathbb{E}[Q] \leq q$ (or $Q \leq q$ almost surely) then $\mathbb{E}[M] \geq \max\{0, ((T + 1 - \varepsilon)128^{-q} - 1)/2\}$, and for every randomized learner satisfying a uniform per-instance expected-query bound q there exists a fixed instance, possibly depending on the learner but not on its random seed, on which this mistake bound holds.

Theorem 3.2 shows that a worst-instance $O(\log T)$ expected-mistake guarantee under the minimal-prefix rule requires $\Omega(\log T)$ worst-instance expected calls: on an explicit distribution over orders and targets, under the same rule, every randomized learner satisfies $\mathbb{E}[M] \geq ((T+1-\varepsilon)\, 128^{-\mathbb{E}[Q]} - 1)/2$. (Table 1 on page 3 lists all five interface/rule combinations.)

Corollary 3.3. *If a randomized learner has worst-instance expected mistakes m_T and expected calls q_T, then $q_T \geq (\ln(T + 1 - \varepsilon) - \ln(2m_T + 1))/(7 \ln 2)$; for every fixed constant $\kappa \geq 0$ independent of T, if $m_T = O((\ln T)^\kappa)$ then $q_T \geq (\ln T - \kappa \ln \ln T - O(1))/(7 \ln 2) = \Omega(\ln T)$. With AHR25 (Thm. 4.5(4)) the randomized ERM query complexity of $O(\log T)$ mistakes is $\Theta(\log T)$, and with Theorem 3.1, Corollary 4.1 and the deterministic $O(T)$-call algorithm of AHR25 (Thm. 4.5(3)) the deterministic one is $\Theta(T)$ under the extremal rules of Theorem 3.1 and against worst-case legal oracles.*

Theorem 3.4 (Selection-rule dependence)**.** *The statement of Theorem 3.1 does not extend to every pre-declared memoryless rule. Let $\mathcal{R}_{\text{feas}}$ return the prefix whose cutoff is the midpoint $\lfloor(\ell + u)/2\rfloor$ of the interval $[\ell, u]$ of feasible cutoffs (and $\perp$ if the interval is empty). $\mathcal{R}_{\text{feas}}$ is a legal memoryless consistency-type ERM oracle, and for $T \geq 2$ there is a deterministic learner with $Q \leq \lceil \log_2 T \rceil$ and $M \leq \lceil \log_2 T \rceil$ on every instance; for $T = 1$ no query is needed and the learner makes at most one mistake under E and none under N (under E, $T = 1$ has two legal targets, so zero mistakes cannot be guaranteed by any learner). For the rule $\mathcal{R}_{\text{global}}$ that returns $\perp$ on unrealizable samples and otherwise projects a fixed global median cutoff $m = \lfloor(k_0 + T)/2\rfloor$ onto the feasible cutoff interval $[\ell, u]$, i.e. returns cutoff $\min\{u, \max\{\ell, m\}\}$ ($k_0 = \varepsilon$), one query and $\lceil T/2 \rceil$ mistakes suffice, while every deterministic learner has $M + Q \geq T - m$ on some instance; so no deterministic learner achieves $O(\log T)$ mistakes and $O(\log T)$ calls under $\mathcal{R}_{\text{global}}$. These are upper and lower bounds, not the exact minimax values of these rules.*

How the results fit together. Theorems 3.1 and 3.2 have independent proofs. AHR25's upper bounds complete their complexity comparison in Corollary 3.3. Theorem 3.4 explains its dependence on selection. The proof routes are:

Result	Supporting route
Deterministic	Prefix criterion (Lemma A.1); common answer (Lemma A.2); invariant and charging (Lemmas A.3–A.4); fixed-instance replay (Lemma A.5); extremal-rule extension (Lemma A.6).
Randomized	Oracle simulations (Lemmas C.1–C.2); posterior invariance (Lemma C.3); refinement and expected per-call potential bounds (Lemmas C.4–C.5).
Rule dependence	Feasible-median construction (Proposition A.7); global-median lower-bound reduction to Theorem 3.1.

Corollary 4.1 follows directly from Theorem 3.1; Appendix B supplies an alternative proof. Section 7 then separates the fixed-budget refinement from the weak-interface comparison. Table 1 (page 3) summarizes the resulting complexity landscape.

4 Deterministic lower bound: stabilize, charge, and fix

We sketch the proof of Theorem 3.1 for $\mathcal{O}_{\min}$; the complete argument, with the reduction for $\mathcal{O}_{\max}$ and $\mathcal{R}_\eta$, is Appendix A.

State space: retain compatible orders and targets. Fix an anchor σ that is constrained to be the order-minimum and hence always has label 1: under E, $\sigma = s$; under N, $\sigma = x_1$. Let $V = X \setminus \{\sigma\}$ and $n = |V| = T - \varepsilon$. The adversary maintains an ordered list L, an unordered set F and an ordered list R partitioning V; the *retained orders* are $\sigma \cdot L \cdot \pi(F) \cdot R$ for all permutations π of F, and the *retained targets* cut anywhere through F: every point of L is committed to label 1, every point of R to label 0, and the target may include any initial segment of the free permutation. Initially $F = V$. The only state changes are *endpoint pins*: a left pin moves $p \in F$ to the end of L, a right pin moves it to the front of R. Pins shrink the retained set (Appendix A, (A.5)). A worked example with three free points is given in Appendix A, §5.

Stabilize the full answer (Lemmas A.1–A.2). For every state and every query S, a refinement by zero or one endpoint pin can be chosen so that the answer $\mathcal{O}_{\min}(S, \preceq)$ is *common* to all orders retained after the refinement: either all answers are $\perp$, or all are one and the same full label vector. Write A for the positive and B for the negative sample points after rejecting contradictory samples and $(\sigma, 0)$, whose answer is $\perp$, and discarding $(\sigma, 1)$; if the remaining positive set is empty the answer is the minimum concept $\{\sigma\}$, so assume $A \neq \emptyset$. By the prefix criterion (Lemma A.1), S is realizable in an order iff no negative precedes a positive, and then the minimal prefix ends at $\max_{\preceq} A$. If some $b \in B$ precedes some $a \in A$ in every retained order, answer $\perp$ without a pin. If A and B both meet F, left-pin a free negative b: every retained order now has b before every free positive, so the answer is $\perp$. Otherwise the sample is realizable in every retained order and only the selected prefix must be stabilized: if A meets R, let p be the last point of $A \cap R$ in the fixed list R; then $p = \max_{\preceq} A$ in every retained order and its prefix $\{\sigma\} \cup L \cup F \cup R_{\leq p}$ is fixed (no pin); if A meets F but not R, then $B \subseteq R$,

and right-pinning a free positive p makes $p = \max A$ with prefix $\{\sigma\} \cup L \cup F_{\text{old}}$ (one pin); if $A \subseteq L$, the prefix through the last positive of L is fixed (no pin).

Maintain the invariant and charge progress (Lemmas A.3–A.4). When the learner predicts $\hat{y}$ on an instance x: if $x \in \{\sigma\} \cup L$ reveal 1, if $x \in R$ reveal 0, and if $x \in F$ reveal $1 - \hat{y}$ and pin x to the corresponding side. The transcript invariant is that every recorded answer equals $\mathcal{O}_{\min}(S, \preceq)$ on every currently retained order, that all revealed labels agree with every retained target, and that no free point has been predicted. The potential $\Phi = |F|$ starts at n, decreases by at most one per call and decreases at a prediction only through a mistake; when all T predictions are made, $\Phi = 0$, so $M + Q \geq n = T - \varepsilon$.

Replay on a fixed instance (Lemma A.5); extend by duality (Lemma A.6). If $n = 0$ the bound is immediate. Otherwise simulate the learner until all predictions are completed or the n-th call is completed, whichever occurs first. Order the residual F by arrival index and take $\preceq^\star = \sigma \cdot L \cdot \pi_0(F) \cdot R$ and $z^\star$ the last point of L (or σ if $L = \emptyset$); this is a retained pair. Run the learner on $(\preceq^\star, z^\star)$ with the declared oracle $S \mapsto \mathcal{O}_{\min}(S, \preceq^\star)$. By the invariant every recorded answer equals $\mathcal{O}_{\min}(S, \preceq^\star)$ and every recorded label equals $c_{z^\star}$, so by determinism the interaction replays exactly, and the bound follows. For $\mathcal{O}_{\max}$ an anchored reversal-complement duality (Lemma A.6) maps $\mathcal{O}_{\max}$ to $\mathcal{O}_{\min}$ preserving mistakes and calls, and $\mathcal{R}_\eta$ applies one of the two procedures at each query. More generally the argument applies to any fixed legal rule with the *one-pin common-value property*: on every anchored state and every sample, some refinement by zero or one endpoint pin makes the rule's full answer constant on all retained orders. This is a sufficient condition for the linear bound, not a characterization (Section 6).

4.1 A direct consequence: adversarial then frozen oracles

Theorem 3.1 implies the operational form of the lower bounds in AHR25, in which the oracle is chosen after the learner and then frozen. Given a deterministic learner $\mathcal{A}$, take $\preceq^\star$ and $z^\star$ from Theorem 3.1(a) with $\mathcal{R} = \mathcal{O}_{\min}$ and set $O^\star = \mathcal{O}_{\min}(\cdot, \preceq^\star)$: this oracle is legal on every sample and depends on the sample only, and the sharpness statement is Theorem 3.1(b). Hence:

Corollary 4.1 (Deterministic lower bound; adversarial then frozen oracle)**.** *Let $T \geq 1$ and fix either convention. For every deterministic learner $\mathcal{A}$ there exist a total order $\preceq^\star$ on X, a point $z^\star \in X$, and a function $O^\star : 2^{X \times \{0,1\}} \to C_{\preceq^\star} \cup \{\perp\}$ which is a legal consistency-type ERM oracle on every sample and depends on the sample only, such that on this fixed instance*

$$M(\mathcal{A}) + Q(\mathcal{A}) \;\geq\; T - \varepsilon.$$

The constant cannot be improved uniformly over $T \geq 1$.

Corollary 4.2. *If a deterministic learner guarantees at most $m(T)$ mistakes on all instances and legal memoryless ERM oracles, its worst-case number of calls is at least $T - \varepsilon - m(T)$. In particular an $O(\log T)$-mistake guarantee requires $T - \varepsilon - O(\log T)$ calls, whereas the randomized learner of AHR25 (Thm. 4.5(4)) achieves $O(\log T)$ expected calls with $O(\log T)$ mistakes. In terms of $d = \lfloor \log_2(T + 1 - \varepsilon) \rfloor$: the deterministic worst-case total cost is at least $T - \varepsilon \geq 2^d - 1$ while the randomized expected total cost is $O(d + 1)$.*

Appendix B gives an alternative, self-contained proof of Corollary 4.1 by a cached-transcript construction, in which the adversary answers adaptively and then freezes its answers. Its delicate part is the freezing step: because the adaptive rule may legally return two different vectors to the same sample at different states (Appendix B exhibits such a sample), the adversary records the first answer to each sample in a cache, and the frozen oracle $O^\star$ returns cached answers on recorded samples; on an unrecorded sample it returns the minimal consistent prefix of the final order if one exists, and $\perp$ otherwise. The fixed-rule proof of Theorem 3.1 needs no cache. Corollary 4.1 has the quantifier form $\forall \mathcal{A}\ \exists(\preceq, z, O)$ of the box in Section 2: the oracle is produced after the learner and then frozen; it does not by itself yield a statement in which the selection rule is fixed before all learners, which is Theorem 3.1.

Remark 4.3 (Why Theorem 4.4 of AHR25 is not invoked). AHR25 prove a general lower bound (their Theorem 4.4) for families of all classes of a given Littlestone dimension d_{LD} by an equivalence-class construction in which "each query provides information about at most one point". Neither Theorem 3.1 nor Corollary 4.1 invokes it: its hypothesis $T \geq 2^{d_{\text{LD}}+1}$ fails for the threshold family;

and its construction uses thresholds on equivalence classes rather than on distinct totally ordered points (verbatim statement and comparison in Appendix B).

5 Randomized lower bound: a logarithmic potential

The upper bound being matched. In AHR25 (Appendix H.1), a new partition of an uncertainty region containing more than five points is built from a uniformly sampled pair queried with opposite labels, kept only when a realizable answer gives each label at least a third of the region, and reused until a mistake certifies one side and shrinks the region by a constant factor; each partition costs $O(1)$ expected calls for every legal oracle.

We sketch the proof of Theorem 3.2; details are in Appendix C. Logarithms are natural and $n = T - \varepsilon$ is the number of *hard* points, assumed $n \geq 1$ (the case $n = 0$, which occurs under N when $T = 1$, is trivial and does not use the construction); under E the sentinel $a = s$ is the minimum of the order, and under N the remaining instance a (which arrives last) is the minimum of the order and always has label 1.

Construct an oblivious hard distribution. Identify the hard points with $x_1, \dots, x_n$ in arrival order and build the balanced binary interval tree on $[n]$ (an interval of size $m \geq 2$ splits into its first $\lfloor m/2 \rfloor$ and remaining $\lceil m/2 \rceil$ indices). Draw a hidden center J uniformly from $[n]$ (recall $n \geq 1$) and independent fair labels $Y_1, \dots, Y_n$. The *priority order* π_J lists, along the root-to-J path, each sibling interval not containing J (in increasing index order) and finally J; thus every sibling met earlier has outer priority to everything in the current interval around J. The total order is $a \prec$ (points with $Y_i = 1$ in π_J order) $\prec$ (points with $Y_i = 0$ in reverse π_J order), and the target ends at the last $Y_i = 1$ point (or at a). Then $c_z(x_i) = Y_i$. The oracle is the memoryless minimal-prefix rule; all random choices precede the interaction. A worked example with eight hard points, including one auxiliary-oracle call, is given in Appendix C, §3.

Reduce to a stronger oracle (Lemmas C.1–C.2). A minimal-prefix call is simulated by one prefix-max call: for the positive hard-point set A of the query it returns the prefix through $\max_{\preceq} A$ if $A \neq \emptyset$ and the singleton anchor prefix if $A = \emptyset$, and checking the queried negative points against this full vector reconstructs the actual answer (Lemma C.1). We replace it by an even stronger *revelation oracle* that maintains an interval $I \ni J$ (initially $[n]$) and, on a query with positive set A, repeatedly reveals which child of I contains J together with all labels in the other child, stopping as soon as $A \setminus I$ contains a target-negative point, or $A \cap I = \emptyset$, or I is a singleton (which it then reveals, after which the active interval is empty); all of this costs *one* call. At every stopping point the requested prefix is determined by the revealed information (Lemma C.2), so any learner for the real oracle is simulated with the same number of calls and the same predictions, and any lower bound proved against the revelation oracle holds for the actual interface.

Preserve the posterior and account for predictions (Lemmas C.3–C.4). Conditional on the entire auxiliary history, while the active interval I is nonempty, J is uniform in I, the labels of unarrived points in I are independent fair bits, and they are independent of J (Lemma C.3); this holds for randomized learners after conditioning on the seed. Call a prediction *unrevealed* if its point is still in I when it is made; such a prediction faces an independent fair bit and costs $1/2$ in expectation, so $\mathbb{E}[M] \geq \frac{1}{2}\mathbb{E}[H_{\text{final}}]$ where H counts unrevealed predictions made so far. With r the number of not-yet-predicted points in I, the potential is $\Phi = \ln(H + r + 1)$. Prediction steps leave Φ unchanged (an unrevealed prediction moves one unit from r to H), and one binary refinement of I decreases Φ by at most $\ln 2$, because already-predicted points inside I have been paid for in H (Lemma C.4).

Bound the expected potential decrease by $O(1)$ (Lemma C.5). Intuitively, the hidden center determines a nested sequence of uncertainty intervals; one call may reveal a large sibling block, but it cannot repeatedly remove many levels unless it hits fresh random labels or spends information already paid for by earlier predictions. Conditional on any history before a call, the expected decrease of Φ during that call is at most $7 \ln 2$, whatever the query set. Call a state *early* if fewer than $\lfloor m/2 \rfloor$ of the m points of I have arrived. In an early state every point of the right child is unarrived (arrivals follow the index order, so the arrived points of I form an initial segment of I, contained in the left child), so its labels are fresh fair bits. If the query avoids the right child, the call stops when J lies

there, which has probability at least $1/2$; if it meets the right child, then with probability at least $\frac{1}{3} \cdot \frac{1}{2}$ the center is in the left child and some queried point of the right child is target-negative, which also stops the call. Hence each early refinement continues with probability at most $5/6$, the expected number of early refinements is at most 6, and from the first late state on the remaining decrease is at most $\ln 2$ since $r \le H + 1$ there. Summing over calls, $\ln(n+1) - \mathbb{E}[\ln(H_{\text{final}} + 1)] \le 7 \ln 2 \cdot \mathbb{E}[Q]$ for an adaptive random number of calls, and Jensen's inequality gives $\mathbb{E}[H_{\text{final}}] \ge (n+1)128^{-\mathbb{E}[Q]} - 1$, hence Theorem 3.2. The fixed-instance clause follows, for each such learner, because the distribution has finite support. (Appendix C also shows, in Propositions C.7–C.9, why the uniform-random-order distribution cannot yield such a bound.)

6 Selection rules locate the boundary of the separation

Theorem 3.4 combines a direct halving algorithm for the feasible-median rule with a direct upper bound and a lower-bound reduction to Theorem 3.1 for the global-median rule (Appendix A, §8).

The feasible-median rule admits deterministic logarithmic learning. Index concepts by the number k of positive instances; for every sample the feasible cutoffs form an integer interval $[\ell, u]$ (or are empty), and $\mathscr{R}_{\text{feas}}$ returns the prefix of cutoff $\lfloor (\ell + u)/2 \rfloor$. The learner (Proposition A.7) maintains $H_1 \prec U \prec H_0$, where H_1 is known to be positive, H_0 known to be negative and U the uncertain interval, initially all instances. While $|U| \ge 2$ it queries $(H_1 \times \{1\}) \cup (H_0 \times \{0\})$, a sample that need not include the labels already observed inside U; the returned concept h splits U into $U_1 = \{h = 1\}$ and $U_0 = \{h = 0\}$ with $1 \le |U_0|, |U_1| \le \lceil |U|/2 \rceil$, because the feasible interval is exactly $[|H_1|, |H_1| + |U|]$ (under N with $H_1 = \emptyset$ it is $[1, |U|]$). It predicts with h until an error. An error with $h(x) = 0$, $y = 1$ certifies every point of U_1 positive and replaces U by U_0; an error with $h(x) = 1$, $y = 0$ certifies U_0 negative and replaces U by U_1. Each error halves U (rounded up), so for $T \ge 2$ both M and Q are at most $\lceil \log_2 T \rceil$; for $T = 1$ no query is needed, with at most one mistake under E and none under N. At $T = 8$ it has $M + Q \le 6 < T - \varepsilon$ on every instance, refuting the universal extension of Theorem 3.1.

The global-median rule does not. Let $m = \lfloor (k_0 + T)/2 \rfloor$ with $k_0 = \varepsilon$ and let $\mathscr{R}_{\text{global}}$ return the feasible cutoff closest to m (formally $\min\{u, \max\{\ell, m\}\}$, and $\perp$ on unrealizable samples). One empty query returns a prefix with m positive instances; predicting with it until the first error and arbitrarily thereafter on the uncertain side gives $Q = 1$ and $M \le \lceil T/2 \rceil$, so this rule also violates the full bound $T - \varepsilon$. But it admits no deterministic $O(\log T)$-call, $O(\log T)$-mistake learner: fix an ordered block H of m instances, revealed to the learner and labeled positive, before a block U of $T - m$ instances in unknown order. On such instances every call to $\mathscr{R}_{\text{global}}$ can be simulated by at most one anchored $\mathscr{O}_{\min}$ call on U (for realizable samples with a positive in U and no negative in H, the rule returns the minimum), so Theorem 3.1 on U gives $M + Q \ge T - m \ge \lfloor T/2 \rfloor$ on some instance.

Interpretation. Three legal, memoryless, pre-declared rules for the same class and the same interface have qualitatively different deterministic complexities: the extremal rules have exact worst-case total cost $T - \varepsilon$; the global-median rule has linear total cost but not the full constant; the feasible-median rule admits $O(\log T)$ calls and mistakes. The randomized learner of AHR25 is oblivious to the rule. In this precise sense randomization substitutes for a favourable selection rule, and the selection rule is a complexity axis of oracle-based online learning, alongside the interface (Section 7) and the class. We do not define a numerical measure of "helpfulness" of a rule; the one-pin common-value property of Section 4 is a sufficient condition for the linear bound.

7 Refinements along the budget and interface axes

Fixed budgets: refine the deterministic obstruction. Against adversarial oracles as in Corollary 4.1, let $M_E^*(T, Q)$ (resp. M_N^*) be the minimax number of mistakes of deterministic learners making at most Q ERM calls, Q a nonnegative integer. Theorem D.1 (stated and proved in Appendix D) gives the exact values in the small- and large-budget regimes, a max sandwich in general, and two finite counterexamples showing that the lower bound is not exact in the middle regime. Under E the lower bounds combine Corollary 4.1 with the known-order lower bound $\lfloor \log_2(T + 1 - \varepsilon) \rfloor$ (Lemma D.10); under N the stronger $T - Q$ budget bound for $Q \ge 1$ additionally uses the first-action recurrence in Theorem D.1(i) and is not a pathwise consequence of Corollary 4.1; the upper bounds

come from a protection algorithm and from sorting followed by majority prediction; the counterexamples are established analytically, by a five-point game analysis under E and a first-action recurrence reducing the N case to it.

Weak consistency: isolate the value of a returned concept. The exponential advantage of randomization at the ERM interface depends on a successful call returning an *evaluable* concept. When the oracle returns only the realizability bit, the advantage disappears at the level of query order. Theorem E.1 (stated and proved in Appendix E) gives a deterministic learner with at most $66T$ weak-consistency calls and $\lfloor \log_2(T+1) \rfloor$ mistakes under E (at most $67(T-1)$ calls and $\lfloor \log_2 T \rfloor$ mistakes under N), and shows, from the lower bound of AHR25 (Thm. 4.3, applied with effective horizon $T - \varepsilon \geq 100$) together with a truncation lemma for expected budgets (Lemma E.5: worst-instance expected mistakes m and calls q satisfy $m + 20q \geq (T - \varepsilon)/20$), that at this interface $\Theta(T)$ calls are necessary and sufficient for $O(\log T)$ mistakes for both deterministic and randomized learners. At the level of asymptotic query order the randomized half follows from AHR25, with Lemma E.5 supplying the expected-budget formulation; the deterministic upper bound replaces their $O(T \log T)$ by $O(T)$, a linear order that AHR25 reach only with randomization (Thm. 4.5(2), by sampling-based rank estimates) and that Theorem E.1 attains deterministically, and its content is the explicit constants and the exact integer mistake bound.

8 Conclusion and open problems

We answer the question of AHR25 for the instance they highlighted: thresholds on an unknown order, in the transductive protocol and with the consistency-type ERM interface. Under a pre-declared extremal rule, and also against adversarial memoryless oracles, the deterministic query complexity of $O(\log T)$ mistakes is at least $T - \varepsilon - O(\log T)$, hence $\Theta(T)$, while the randomized complexity, with worst-instance expected mistakes and calls, is $\Theta(\log T)$ under the minimal-prefix rule and in the worst case over legal memoryless oracles. The selection-rule comparison identifies the boundary of this gap: for $T \geq 2$, a legal feasible-median rule already admits deterministic $O(\log T)$ calls and mistakes (Theorem 3.4). We do not resolve the corresponding question for general families, for the non-transductive protocol, or for arbitrary pre-declared selection rules, nor a high-probability analogue of the randomized (expected-cost) characterization.

Beyond thresholds, the relevant mechanism is the freedom to select among equally valid oracle answers. Hypotheses can agree on a submitted sample while disagreeing elsewhere; selecting among them changes the returned predictions on those points and may affect later queries. An optimization oracle has analogous freedom when several predictors minimize the submitted objective. This motivates examining selection in oracle-based contextual and hybrid learning (Syrgkanis et al., 2016; Okoroafor et al., 2026), while keeping their different objectives and feedback explicit. A natural comparison holds the class and interface fixed and varies only the legal selection rule. Our threshold results isolate this issue; extending the complexity separation requires a separate construction or reduction.

Limitations and open problems.

- **Scope.** The separation is proved for one family (thresholds on an unknown order), one protocol (transductive) and one interface (consistency-type ERM with full label vectors); it is not claimed for arbitrary Littlestone classes or for the non-transductive setting, which is the general form of the question of AHR25.
- **Rule dependence.** The linear deterministic bound holds for the extremal rules and against worst-case legal oracles, not for every pre-declared rule (Theorem 3.4). Characterizing the rules that permit deterministic logarithmic-query, logarithmic-mistake learning is the central open problem raised here; endpoint versus non-endpoint selection does not decide it, since neither median rule always returns an endpoint, yet their complexities differ.
- **Fixed budgets.** The tradeoff in the intermediate-budget regime $\lfloor T/2 \rfloor < Q < 2(T-1)$ is not fully characterized: Theorem D.1(iii) rules out $\max(T - Q, \lfloor \log_2(T+1) \rfloor)$ as a general identity, while some individual values in this regime are known.
- **Constants.** The constant 128 in Theorem 3.2 is not optimized.

AI use statement

In this work, we used generative AI tools to help develop theoretical models or conceptual frameworks, provide critical ingredients for proving mathematical claims, assist in writing proofs, design and provide feedback on the research methodology, implement methods and supplementary verification procedures, assist with translation, and interpret theoretical results and their scope. We have not used generative AI tools to formulate mathematical claims or to propose or refine hypotheses. Generating synthetic datasets, cleaning or reformatting datasets, and qualitative or thematic data analysis are not applicable to this work, which is a theoretical study without empirical datasets.

Additionally, we used generative AI tools to create or modify scientific figures, create or edit software code, draft parts of the paper, summarize or analyse existing literature, discover research topics or identify gaps, brainstorm, source or search for information, edit the paper to improve readability, identify relevant literature, format references, suggest the structure of the paper, and propose titles or keywords. The supplementary verification code comprises `one_pin_verification.py`, `exact_threshold_game.py`, `theorem_C_verification.py`, `theorem_D_verify.py`, and `theorem_Aprime_checks.py`.

We have reviewed all AI-assisted work. Three human reviewers manually checked all mathematical derivations and proofs, as well as any AI-assisted code, and confirmed their correctness. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

Ethics statement

This work is purely theoretical. It involves no human subjects, no data collection and no deployment; the authors are not aware of ethical concerns raised by its content.

Reproducibility statement

All new mathematical results have complete proofs in the appendices: Theorems 3.1 and 3.4 in Appendix A, Corollary 4.1 in Appendix B, Theorem 3.2 in Appendix C, Theorem D.1 in Appendix D and Theorem E.1 in Appendix E. Section 2 specifies the model and conventions; additional hypotheses are stated in the individual theorems, and the quantifier form of each lower bound is fixed in Section 2. The five supplementary verification scripts provided with the submission are exact-arithmetic finite checks (integers and rationals, no floating point); the accompanying file `VERIFICATION_SCOPE.md` lists, for each script, the objects enumerated, the parameter range and the assertions, and lists the steps that no script checks. In brief: `one_pin_verification.py` checks, on every anchored state with at most four points besides the anchor and every sample, that the answer rule of Lemma B.2 pins at most one point and returns an answer common to all retained orders while preserving the invariants of Lemma B.3, and replays the cached-transcript oracle of Appendix B for $T \le 4$; `exact_threshold_game.py` computes $M_E^*(5,3) = 3$ of Theorem D.1(iii) by exhaustive game-tree search (the value $M_N^*(6,4) = 3$ is proved analytically in Appendix D); `theorem_C_verification.py` checks the zero-query values of Proposition C.7 for $T \le 12$ and the equality of the auxiliary-oracle simulation with the actual prefix answer (Lemmas C.1–C.2) for $n \le 6$ hard points, and does not test the potential bound of Lemma C.5; `theorem_D_verify.py` runs the learner of Theorem E.1 on every order and target for $T \le 7$ under both conventions, checking the call bounds $66T$ and $67(T-1)$, the mistake bounds and the phase invariants, and checks the $32n$ comparison bound of the selection subroutine up to $n = 4096$; `theorem_Aprime_checks.py` computes the exact minimax value of $M + Q$ under the minimal-prefix rule for $T \le 4$ under both conventions and checks the answer rule of Lemma A.2 on every anchored state with at most five points besides the anchor.

References

Noga Alon, Shay Moran, and Shlomo Moran. Sorting from counterexamples, 2026. arXiv:2608.21579.

Dana Angluin. Queries and concept learning. *Machine Learning*, 2(4):319–342, 1988.

Angelos Assos, Idan Attias, Yuval Dagan, Constantinos Daskalakis, and Maxwell K. Fishelson. Online learning and solving infinite games with an ERM oracle. In *Proceedings of the 36th Conference on Learning Theory (COLT)*, volume 195 of *PMLR*, pp. 274–324, 2023.

Idan Attias, Steve Hanneke, and Arvind Ramaswami. Tradeoffs between mistakes and ERM oracle calls in online and transductive online learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 38, pp. 75639–75680, 2025. arXiv:2506.00135.

Idan Attias, Steve Hanneke, and Arvind Ramaswami. Regret–oracle complexity tradeoffs in agnostic online learning, 2026. arXiv:2605.07155.

Shai Ben-David, Eyal Kushilevitz, and Yishay Mansour. Online learning versus offline learning. *Machine Learning*, 29(1):45–63, 1997.

Constantinos Daskalakis and Noah Golowich. Is efficient PAC learning possible with an oracle that responds "yes" or "no"? In *Proceedings of the 37th Conference on Learning Theory (COLT)*, volume 247 of *PMLR*, pp. 1263–1307, 2024.

Steve Hanneke, Shay Moran, and Jonathan Shafer. A trichotomy for transductive online learning. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, pp. 19502–19519, 2023. arXiv:2311.06428.

Alexander Kozachinskiy and Tomasz Steifer. Simple online learning with consistent oracle. In *Proceedings of the 37th Conference on Learning Theory (COLT)*, volume 247 of *PMLR*, pp. 3241–3256, 2024.

Nick Littlestone. Learning quickly when irrelevant attributes abound: A new linear-threshold algorithm. *Machine Learning*, 2(4):285–318, 1988.

Princewill Okoroafor, Robert Kleinberg, and Michael P. Kim. Oracle-efficient hybrid online learning with constrained adversaries. In *International Conference on Learning Representations (ICLR)*, 2026.

Vinod Raman, Shenghao Xie, and Samson Zhou. Transductive and learning-augmented online regression. arXiv:2510.03917, 2025.

Vasilis Syrgkanis, Akshay Krishnamurthy, and Robert E. Schapire. Efficient algorithms for adversarial contextual learning. In *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, volume 48 of *PMLR*, pp. 2159–2168, 2016.

A Proof of Theorem 3.1 and Theorem 3.4

Proof organization. Parts 1–5 establish the minimal-prefix lower bound: the prefix criterion supports the common-value lemma, which feeds the transcript invariant, charging, and fixed-instance replay. Part 6 extends the argument to the other extremal rules; Part 7 establishes sharpness. Part 8 contrasts the median rules, with an independent feasible-median construction; Part 9 states the common-value template.

1. Statement and conventions

All lemmas below are proved here. We use the consistency-ERM interface of Section 2, including full-vector returns and unrestricted finite query samples. No weak-consistency interface is used.

For a finite ordered domain, write:

- $\mathcal{O}_{\min}$: return the smallest consistent prefix, or $\perp$ if none exists;
- $\mathcal{O}_{\max}$: return the largest consistent prefix, or $\perp$ if none exists.

Under convention **E**, the domain is

$$X = \{s, x_1, \dots, x_T\}, \qquad s \prec x \quad (x \neq s).$$

Thus every concept labels s by 1, and the smallest prefix is $\{s\}$. The $\mathcal{O}_{\min}$ case of Theorem 3.1 is treated first below.

Under convention **N**, the domain is $X = \{x_1, \dots, x_T\}$, and prefixes must be nonempty. In particular, when a query has no positive labels, $\mathcal{O}_{\min}$ returns the singleton containing the order-minimum if that singleton is consistent, and otherwise returns $\perp$. This is the appropriate nonempty-prefix version of the same rule.

Statement of Theorem 3.1 (fixed extremal ERM rules)

Fix either $\mathcal{R} = \mathcal{O}_{\min}$ or $\mathcal{R} = \mathcal{O}_{\max}$, publicly and before choosing the learner or instance. For every deterministic learner $\mathcal{A}$, and every prescribed sequence of distinct instances $x_1, \dots, x_T$, there exist a total order $\preceq^\star$ and an endpoint $z^\star \in X$ such that

$$M(\mathcal{A}, \mathcal{R}(\cdot, \preceq^\star), x_{1:T}, c_{z^\star}) + Q(\mathcal{A}, \mathcal{R}(\cdot, \preceq^\star), x_{1:T}, c_{z^\star}) \geq T - \varepsilon,$$

where $\varepsilon = 0$ under E and $\varepsilon = 1$ under N.

The oracle is the originally declared function $\mathcal{R}$; no return-value table is selected after constructing the instance.

The statement also holds for every predeclared sample-dependent extremal selector

$$\mathcal{R}_\eta(S, \preceq) = \begin{cases} \mathcal{O}_{\min}(S, \preceq), & \eta(S) = 0, \\ \mathcal{O}_{\max}(S, \preceq), & \eta(S) = 1, \end{cases} \tag{A.1}$$

where η depends on S, not on the unknown order or interaction history.

We first prove the theorem for $\mathcal{O}_{\min}$.

2. Anchored states

Use a distinguished anchor σ, which is constrained to be the order-minimum and therefore always has target label 1.

- Under E, take $\sigma = s$. It is not an online instance.
- Under N, take $\sigma = x_1$. It is an online instance whose label is fixed to 1.

Let

$$V = X \setminus \{\sigma\}, \qquad n = |V| = T - \varepsilon.$$

The construction maintains ordered lists L, R and an unordered set F, partitioning V. Its retained orders are

$$\Omega(L, F, R) = \{\sigma \cdot L \cdot \pi(F) \cdot R : \pi(F) \text{ is any permutation of } F\}. \tag{A.2}$$

The retained order–target pairs are

$$\Gamma(L, F, R) = \left\{\left(\sigma \cdot L \cdot \pi(F) \cdot R,\ \mathbf{1}_{\{\sigma\} \cup L \cup \{\pi_1, \dots, \pi_k\}}\right) : 0 \leq k \leq |F|\right\}. \tag{A.3}$$

Thus:

- every point of L is committed to target label 1;
- every point of R is committed to target label 0;
- the target may cut anywhere through the freely permutable block F.

Commitment does not mean that the label has already been revealed. In particular, query-pinned points can be assigned their eventual target labels immediately. Hence no separate blocks of pinned points are needed: pinned points are absorbed directly into L or R.

Initially,

$$L = R = \emptyset, \qquad F = V.$$

Endpoint refinements

There are two permitted one-point refinements:

$$\begin{aligned} &\text{left pin of } p \in F: \quad (L, F, R) \mapsto (L \cdot p, F \setminus \{p\}, R), \\ &\text{right pin of } p \in F: \quad (L, F, R) \mapsto (L, F \setminus \{p\}, p \cdot R). \end{aligned} \tag{A.4}$$

Both preserve nonemptiness and satisfy

$$\Gamma_{\text{new}} \subseteq \Gamma_{\text{old}}, \qquad \Omega_{\text{new}} \subseteq \Omega_{\text{old}}. \tag{A.5}$$

Indeed, a left pin restricts the old free permutation to $p \cdot \pi(F \setminus \{p\})$ and restricts the target to include p. A right pin restricts it to $\pi(F \setminus \{p\}) \cdot p$ and restricts the target to exclude p. These were already allowed choices in (A.3).

No relative order among the remaining free points is imposed.

3. Prefix realizability and the common-value query lemma

Lemma A.1 — Prefix criterion

Suppose σ is the minimum. Reject immediately if a sample S is contradictory or contains $(\sigma, 0)$. Otherwise define

$$A = \{x \neq \sigma : (x, 1) \in S\}, \qquad B = \{x : (x, 0) \in S\}.$$

Then S is realizable if and only if there are no $b \in B, a \in A$ with $b \prec a$. When realizable,

$$\mathcal{O}_{\min}(S, \preceq) = \begin{cases} \mathbf{1}_{\{\sigma\}}, & A = \emptyset, \\ \mathbf{1}_{\{x : x \preceq \max_{\preceq} A\}}, & A \neq \emptyset. \end{cases} \tag{A.6}$$

Proof. A prefix containing a must contain every $b \prec a$, proving necessity. Conversely, if $A \neq \emptyset$, the prefix through $\max A$ contains all positives and no negatives. If $A = \emptyset$, the singleton anchor is consistent. These are the smallest possible consistent prefixes. □

For a current state, write

$$b \lhd a \quad \iff \quad b \prec a \text{ in every order in } \Omega(L, F, R). \tag{A.7}$$

This relation is determined by the block order and the fixed internal orders of L, R. Two distinct free points are incomparable under $\lhd$.

Lemma A.2 — One-pin common-value query lemma

For every state and every query S, the construction can make at most one endpoint pin so that

$$\mathcal{O}_{\min}(S, \preceq)$$

has exactly the same full-vector value for every order retained after the pin.

Proof.

Contradictory samples and samples containing $(\sigma, 0)$ have the common answer $\perp$, with no refinement. Otherwise use A, B from Lemma A.1.

Case I: an inversion is already forced If some $b \in B, a \in A$ satisfy $b \lhd a$, answer $\perp$ without changing the state.

Lemma A.1 makes this the required answer for every retained order.

Case II: both label sides contain free points Suppose Case I does not apply and

$$A \cap F \neq \emptyset, \qquad B \cap F \neq \emptyset.$$

Choose $b \in B \cap F$, using the smallest arrival index to make the construction explicit, and left-pin b.

Because the sample is not contradictory, any $a \in A \cap F$ is distinct from b. Every new retained order has $b \prec a$. Therefore its prescribed oracle answer is $\perp$.

Exactly one free point was pinned.

Remaining cases Now suppose neither Case I nor Case II applies.

Every positive–negative pair has at least one endpoint outside F, so its relative order is already fixed. Absence of a forced inversion therefore implies that the sample is realizable in every current retained order. It remains to stabilize the *selected smallest prefix*.

There are three exhaustive possibilities.

Case III-R: $A \cap R \neq \emptyset$**.**

Let p be the last point of $A \cap R$ in the fixed list R. In every retained order,

$$p = \max_{\leq} A.$$

Its down-set is the fixed set

$$D = \{\sigma\} \cup L \cup F \cup R_{\leq p}, \tag{A.8}$$

where $R_{\leq p}$ is the initial segment of R ending at p.

Answer $\mathbf{1}_D$, without a pin.

This case includes samples having positives in both F and R.

Case III-F: $A \cap R = \emptyset$ **and** $A \cap F \neq \emptyset$**.**

Case II's exclusion gives $B \cap F = \emptyset$. Case I's exclusion gives $B \cap L = \emptyset$, since every left point precedes every free positive. Thus

$$B \subseteq R. \tag{A.9}$$

Choose $p \in A \cap F$ by smallest arrival index and right-pin it:

$$R_{\text{new}} = p \cdot R_{\text{old}}.$$

The point p is after every remaining free point but **before every old right point**. Since there are no positives in the old right block,

$$p = \max_{\leq} A \quad \text{for every new retained order.}$$

All negatives are in the old right block and hence follow p. The required answer is therefore

$$\mathbf{1}_D, \qquad D = \{\sigma\} \cup L \cup F_{\text{old}}. \tag{A.10}$$

This is one common full vector, obtained with one pin.

Case III-L: $A \subseteq L$**.**

If $A = \emptyset$, answer $\mathbf{1}_{\{\sigma\}}$. Otherwise let p be the last point of A in L, and answer

$$\mathbf{1}_{\{\sigma\} \cup L_{\leq p}}. \tag{A.11}$$

No pin is needed.

These cases exhaust all queries and establish the assertion. □

4. Labels, invariants, and charging

Lemma A.3 — Strong transcript invariant

The construction can maintain all of the following:

1. $\Gamma(L, F, R)$ is nonempty.
2. Every retained order–target pair agrees with every label already committed by the construction.
3. Every previous oracle answer is exactly the value of the predeclared $\mathcal{O}_{\min}$ on **every** order in the current $\Omega(L, F, R)$.
4. Every point still in F has not yet been predicted.

At a prediction on a free point, the construction can force an error while preserving these properties.

Proof.

The assertions hold initially. Queries are handled by Lemma A.2. Their refinements preserve nonemptiness and nesting by (A.5). Consequently all earlier common-value assertions remain true, and the new answer has the same property.

At a prediction event:

- if $x = \sigma$ or $x \in L$, commit $y = 1$;
- if $x \in R$, commit $y = 0$;
- if $x \in F$, observe the deterministic prediction $\hat{y}$, set

$$y = 1 - \hat{y},$$

and left-pin x when $y = 1$, or right-pin x when $y = 0$.

The last operation is an endpoint refinement. It leaves a nonempty subset of previously retained order–target pairs, all having the newly committed label. It also incurs an error and removes the predicted point from F.

A label may be committed in this offline simulation at the prediction event. Lemma A.5 below turns the result into an actual oblivious instance: the replay uses an order and target fixed before the interaction, and each label is revealed only after the learner has made its corresponding prediction. □

Lemma A.4 — Additive charging

If all T predictions are completed, then

$$M + Q \geq n = T - \varepsilon. \tag{A.12}$$

Proof.

Let p be the number of points removed from F by queries. Lemma A.2 gives

$$p \leq Q.$$

Every other initially free point is removed at its prediction event and incurs an error. Since every point in V eventually appears,

$$M \geq n - p.$$

Adding the inequalities gives (A.12).

Equivalently, the initial potential is $\Phi_0 = |F| = n$; a query decreases it by at most one, and a prediction decreases it only when an error occurs. At completion, $\Phi = 0$. □

5. Oblivious fixation and memorylessness

Lemma A.5 — Finite fixation and replay

The construction yields a single fixed order and target with the claimed cost, using the originally declared memoryless oracle.

Proof.

If $n = 0$, the assertion is immediate. Otherwise simulate the deterministic learner, stopping at the first of:

1. completion of all T predictions; or
2. completion of the n-th query.

This uses at most n query events and T prediction events. In particular, a learner that would make infinitely many queries is covered by the second stopping condition.

Let the final retained state be (L, F, R). Order the remaining F by arrival index, obtaining $\pi_0(F)$, and define

$$\preceq^\star = \sigma \cdot L \cdot \pi_0(F) \cdot R, \tag{A.13}$$

and

$$z^\star = \begin{cases} \text{last point of } L, & L \neq \emptyset, \\ \sigma, & L = \emptyset. \end{cases} \tag{A.14}$$

This is a retained order–target pair: it chooses the cut immediately after L. Hence its target agrees with all committed labels.

Now run the learner from the beginning on this fixed order, target, and the fixed oracle

$$S \mapsto \mathcal{O}_{\min}(S, \preceq^\star).$$

Induct over the recorded events. Determinism implies that the next query or prediction is unchanged whenever the preceding transcript is unchanged. For every recorded query S, Lemma A.3 gives

$$\text{recorded answer} = \mathcal{O}_{\min}(S, \preceq^\star),$$

including the entire returned vector. Recorded labels also agree with $c_{z^\star}$. Thus the interaction replays exactly through the stopping event.

If all predictions were completed, Lemma A.4 applies. Otherwise the replay already contains n queries, so $Q \geq n$.

The fixed oracle is legal on **every** sample by its definition and Lemma A.1, including samples outside the recorded transcript. No answer cache or oracle-extension table is needed.

Repeated queries are automatically consistent: every later retained family is a nonempty subset of the earlier one, and a deterministic function cannot have two different values on the same sample and the same retained order.

Finally, (A.13)–(A.14) are selected before the actual replay. The actual adversary is therefore oblivious as required in Section 2. □

Lemmas A.1–A.5 prove Theorem 3.1 for $\mathcal{O}_{\min}$.

A worked example

Let $n = 3$ under E, so $\sigma = s$, $V = \{x_1, x_2, x_3\}$ and initially $F = V$. Suppose the learner first queries $S_1 = \{(x_2, 1)\}$. Here $A = \{x_2\} \subseteq F$ and $B = \emptyset$, so Case III-F applies: right-pin x_2, giving $L = \emptyset$, $F = \{x_1, x_3\}$, $R = x_2$, and answer the prefix $\{s\} \cup F_{\text{old}} = \{s, x_1, x_2, x_3\}$ (the all-ones vector), which is $\mathcal{O}_{\min}(S_1, \preceq)$ for every retained order $s \cdot \pi(x_1, x_3) \cdot x_2$. Next x_1 arrives and the learner predicts $\hat{y} = 0$: by Lemma A.3 the label is $y = 1$ and x_1 is left-pinned, so $L = x_1$, $F = \{x_3\}$, $R = x_2$, one mistake. Suppose the learner now queries $S_2 = \{(x_3, 1), (x_1, 0)\}$: the negative point $x_1 \in L$ precedes

the positive point $x_3 \in F$ in every retained order, so Case I answers $\perp$ without a pin. Next $x_2 \in R$ arrives; suppose the learner predicts 0: its committed label is 0 (Lemma A.3), so the prediction is correct and the state is unchanged. Finally $x_3 \in F$ arrives, the learner predicts 1, the label is 0 and x_3 is right-pinned: $L = x_1$, $F = \emptyset$, $R = x_3 \cdot x_2$, two mistakes. Altogether $Q = 2$, $M = 2$ and $M + Q = 4 \geq n$; in the accounting of Lemma A.4, $p = 1$ point was removed from F by a query and $n - p = 2$ mistakes were forced. Lemma A.5 fixes $\preceq^\star = s \prec x_1 \prec x_3 \prec x_2$ and $z^\star = x_1$ (the last point of L); replaying the learner against $\mathcal{O}_{\min}(\cdot, \preceq^\star)$ returns the same two answers (x_2 is $\preceq^\star$-last, so the prefix through x_2 is everything, and $x_1 \prec^\star x_3$ makes S_2 unrealizable) and the same labels.

One full answer, at most one query pin

Event	Retained state $L \mid F \mid R$	Oracle reply / target feedback	$\lvert F\rvert$	(M, Q)
Initial	$\emptyset \mid \{1,2,3\} \mid \emptyset$	—	3	$(0,0)$
Query S_1	$\emptyset \mid \{1,3\} \mid [2]$	All-ones concept; right-pin x_2	2	$(0,1)$
Predict x_1: $\hat{y} = 0$	$[1] \mid \{3\} \mid [2]$	$y_1 = 1$: mistake	1	$(1,1)$
Query S_2	$[1] \mid \{3\} \mid [2]$	$\perp$; no pin	1	$(1,2)$
Predict x_2: $\hat{y} = 0$	$[1] \mid \{3\} \mid [2]$	$y_2 = 0$: correct	1	$(1,2)$
Predict x_3: $\hat{y} = 1$	$[1] \mid \emptyset \mid [3,2]$	$y_3 = 0$: mistake	0	$(2,2)$

State entries abbreviate x_i by i; braces denote a set and brackets an ordered list. Here $S_1 = \{(x_2, 1)\}$ and $S_2 = \{(x_3, 1), (x_1, 0)\}$.

$$s \prec x_1 \prec x_3 \prec x_2, \qquad z^\star = x_1, \qquad M + Q = 4 \geq 3.$$

Table 2: Replay of the three-point worked example under convention E (Section 4). These are the offline construction's retained blocks, not extra feedback to the learner; labels committed in L and R need not yet have been revealed. Returned concepts and target feedback are distinct: S_1 returns 1 at x_2, whose target label is 0. The two queries remove only one free point in total; the other two removals occur at mistakes. The final order and target reproduce both full answers under the pre-declared minimal-prefix rule.

6. The maximal-prefix rule and other extremal selectors

Lemma A.6 — Anchored reversal-complement duality

Keep the anchor fixed. Reverse the order of V, complement every label on V, and leave the anchor label unchanged.

Denote this transformation by D. For a concept,

$$(Dh)(\sigma) = 1, \qquad (Dh)(x) = 1 - h(x) \quad (x \in V).$$

For samples, labels on V are complemented, while labels on σ are unchanged; let $D\perp = \perp$.

Then

$$D(\mathcal{O}_{\max}(S, \preceq)) = \mathcal{O}_{\min}(DS, D \preceq). \tag{A.15}$$

Proof.

An anchored prefix containing k points of V becomes an anchored prefix containing $n - k$ points in the reversed order. This is a consistency-preserving bijection between the two concept classes, reversing inclusion. Consequently, the largest consistent prefix becomes the smallest consistent prefix. Unrealizable samples remain unrealizable. □

Given a deterministic learner for $\mathcal{O}_{\max}$, simulate it using $\mathcal{O}_{\min}$ through (A.15), complementing predictions and revealed labels on V, and leaving the anchor unchanged. Queries and errors are preserved exactly. Applying the already proved anchored minimum-rule theorem proves Theorem 3.1 for the maximum rule.

This also handles N correctly: the actual minimum instance is kept as an always-positive anchor. Ordinary reversal and complementation of **all** points would incorrectly introduce an empty concept.

For completeness, duality maps a state to

$$D(L, F, R) = (\text{reverse}\, R,\ F,\ \text{reverse}\, L). \tag{A.16}$$

Since D fixes the anchor and reverses V, it maps the retained orders of a state (L, F, R) bijectively onto the retained orders of $D(L, F, R)$, and it exchanges left pins with right pins. Given a state and a sample S, apply Lemma A.2 to the dual state and the dual sample DS: after at most one endpoint pin of the dual state, which corresponds to at most one endpoint pin of (L, F, R), the value $\mathcal{O}_{\min}(DS, D \preceq)$ is common to all retained orders $D \preceq$ of the refined dual state, and by (A.15), since D is an involution, the value $\mathcal{O}_{\max}(S, \preceq) = D\,\mathcal{O}_{\min}(DS, D \preceq)$ is common to all retained orders $\preceq$ of the refined (L, F, R). Thus the maximum rule has the same one-pin common-value property. For the selector (A.1), $\eta(S)$ is fixed independently of the compatible order, so at each query one applies the minimum or the maximum one-pin procedure to the current state according to $\eta(S)$; both procedures refine only by endpoint pins, so the nesting (A.5), the invariant of Lemma A.3, the charging of Lemma A.4 and the fixation of Lemma A.5 apply verbatim with $\mathcal{R}_\eta$ in place of $\mathcal{O}_{\min}$.

More generally, the proof applies to any fixed legal rule satisfying the following sufficient condition:

> **One-pin common-value property:** On every anchored block state and every sample, some refinement by zero or one endpoint pin makes the rule's full answer constant on all retained orders.

This is a sufficient condition, not a claimed characterization of every possible rule.

7. Sharpness and relation to AHR25

The total-cost lower bound is exact:

$$\inf_{\mathcal{A}\ \text{deterministic}} \sup_{\preceq, z} (M_{\mathcal{A}} + Q_{\mathcal{A}}) = T - \varepsilon \tag{A.17}$$

for either extremal rule.

The matching upper bounds require no oracle calls:

- under E, any fixed predictions make at most T errors;
- under N, always predicting 1 makes at most $T - 1$ errors, since every nonempty prefix labels at least one instance by 1.

Thus (A.17) includes both an upper-bound algorithm and the lower bound; it is not an assertion about an unproved exact fixed-Q tradeoff.

Our common-value lemma establishes a one-point property for genuine total-order thresholds and a fixed natural ERM selection rule; it does **not** assert that the returned vector contains only one bit. Theorem 4.4 of AHR25 is not invoked here; its statement, and why its hypothesis fails for the threshold family, are discussed once in Appendix B.

By Theorem 3.1, a deterministic learner with $O(\log T)$ worst-case errors under $\mathcal{O}_{\min}$ must make

$$T - \varepsilon - O(\log T)$$

queries on some fixed instance. AHR25 (Thm. 4.5(4)) supplies the randomized $O(\log T)$-error, $O(\log T)$-expected-query upper bound under the same fixed rule, since that algorithm works for arbitrary legal ERM returns. This establishes the intended separation without adversarial tie-breaking.

8. Arbitrary predeclared rules: a counterexample

The universal extension

"Theorem 3.1 holds for every fixed function of $(S, \preceq)$"

is false.

8.1 Median of the feasible prefixes

Index concepts by the number k of positive online-domain points:

- E: $k \in \{0, \dots, T\}$, with the sentinel always positive;
- N: $k \in \{1, \dots, T\}$.

For any sample, its feasible cutoff indices form an integer interval

$$I(S, \preceq) = \{\ell, \ell + 1, \dots, u\},$$

or are empty. Define the predeclared rule

$$\mathscr{R}_{\text{feas}}(S, \preceq) = \begin{cases} \bot, & I(S, \preceq) = \emptyset, \\ \text{prefix of cutoff } \left\lfloor \dfrac{\ell + u}{2} \right\rfloor, & \text{otherwise.} \end{cases} \tag{A.18}$$

This rule depends only on S and the order. In particular, it does not know a learner's state or free block.

Proposition A.7 — Deterministic logarithmic learning under (A.18)

Under either convention, for $T \geq 2$, there is a deterministic learner satisfying, for every order and target,

$$Q \leq \lceil \log_2 T \rceil, \qquad M \leq \lceil \log_2 T \rceil. \tag{A.19}$$

Proof.

Maintain a decomposition of the instance set into

$$H_1 \prec U \prec H_0,$$

where H_1 is an order-prefix known to have target label 1, H_0 is an order-suffix known to have target label 0, and U is the intervening interval. Initially U is the entire instance set.

For $|U| \geq 2$, query

$$S = (H_1 \times \{1\}) \cup (H_0 \times \{0\}). \tag{A.20}$$

This sample is consistent. Let h be the returned concept and set

$$U_1 = \{x \in U : h(x) = 1\}, \qquad U_0 = U \setminus U_1.$$

Write $a = |H_1|$, $m = |U|$. Under E, the feasible cutoff interval is exactly $[a, a + m]$, so $|U_1| = \lfloor m/2 \rfloor$. Under N the same holds when $a > 0$; when $a = 0$, the feasible interval is $[1, m]$, so $|U_1| = \lceil m/2 \rceil$. Therefore

$$1 \leq |U_0|, |U_1| \leq \lceil m/2 \rceil \qquad (m \geq 2). \tag{A.21}$$

Predict with h until an error occurs.

- If $h(x) = 0$ but $y = 1$, then every point in U_1 precedes x and must have target label 1. Replace

$$H_1 \leftarrow H_1 \cup U_1, \qquad U \leftarrow U_0.$$

- If $h(x) = 1$ but $y = 0$, then every point in U_0 follows x and must have target label 0. Replace

$$H_0 \leftarrow U_0 \cup H_0, \qquad U \leftarrow U_1.$$

Thus each error reduces the active interval to at most half, rounded up. If the new interval is a singleton, it consists of the just-mistaken point, whose label is now known; all target labels are then determined. Otherwise make the next query before the next prediction.

After j such errors,

$$|U| \leq \left\lceil \frac{T}{2^j} \right\rceil.$$

Consequently there are at most $\lceil \log_2 T \rceil$ errors. Queries are issued only while $|U| \geq 2$, and every query after the first is preceded by exactly one error, so before the q-th query $2 \leq |U| \leq \lceil T/2^{q-1} \rceil$, which forces $2^{q-1} < T$; hence there are also at most $\lceil \log_2 T \rceil$ queries. For $T = 1$, no queries and at most one error suffice under E; under N, predict 1 without error.

Past labels inside U need not be included in (A.20). The interface of Section 2 allows arbitrary samples; a queried hypothesis is not required to fit observations omitted from that sample. □

At $T = 8$, this gives one explicit deterministic learner with

$$M + Q \leq 6$$

for **every** order and target, under both conventions. Since $6 < 8$ under E and $6 < 7$ under N, the all-rules extension is refuted with the required quantifiers.

These are rule-specific upper bounds. No balance is assumed of an arbitrary consistency-ERM return.

8.2 The different "closest to the global median" rule

A different natural rule chooses the feasible cutoff closest to a fixed global median. This is different from (A.18).

Let

$$k_0 = \begin{cases} 0, & E, \\ 1, & N, \end{cases} \qquad m = \left\lfloor \frac{k_0 + T}{2} \right\rfloor.$$

Define $\mathscr{R}_{\text{global}}(S, \preceq) = \perp$ when $I(S, \preceq) = \emptyset$ (this covers contradictory samples and, under E, a negative sentinel); otherwise, when the feasible interval is $[\ell, u]$, let it return the prefix of cutoff

$$\text{proj}_{[\ell,u]}(m) = \min\{u, \max\{\ell, m\}\}. \tag{A.22}$$

This rule also fails the full inequality of Theorem 3.1(a), but **does not** admit simultaneous $O(\log T)$ mistakes and queries.

Upper bound sufficient to refute the bound of Theorem 3.1(a) One empty-sample query returns a prefix having m positive instances. Predict with that concept until the first error.

If an error occurs, the target uncertainty is confined to the side containing the mistaken point: the opposite side has a forced target label. Thereafter, arbitrary predictions inside the uncertain side incur at most one error per remaining point.

Hence

$$Q = 1, \qquad M \leq \max(m, T - m) = \lceil T/2 \rceil. \tag{A.23}$$

This argument is valid for any returned threshold, with its actual side sizes. Balance enters only through the declared rule's empty-query output.

For $T = 8$, (A.23) gives $M + Q \leq 5$, refuting the bound of Theorem 3.1(a) under both conventions.

Linear lower bound for the global-median rule For every deterministic learner, some instance satisfies

$$M + Q \geq T - m. \tag{A.24}$$

To prove this, let $r = T - m$. If $r = 0$ the claim is $M + Q \geq 0$ and holds trivially, so assume $r \geq 1$. If some instance in the family constructed below makes the original learner issue at least r queries (possibly infinitely many), then (A.24) already holds on that instance; otherwise the original learner makes at most $r - 1$ queries on every instance of the family, and the per-query simulation below yields a derived learner that completes the protocol. Fix an ordered block H of m instances before a block U of $T - m$ instances whose internal order is unknown. Restrict targets to label all of H positively, with an arbitrary, possibly empty, prefix target on U. The order of H can even be revealed to the learner.

An interaction with $\mathscr{R}_{\text{global}}$ can be simulated using an anchored $\mathscr{O}_{\min}$ oracle on U, with at most one minimum-oracle call per original query:

1. Contradictions and an illegal negative sentinel are rejected directly.
2. If the sample has a negative in H and a positive in U, it is unrealizable.
3. If it has a negative in H but no positive in U, any feasible cutoff lies strictly before m. Its largest feasible cutoff, or infeasibility, is computable entirely inside the known ordered block H.
4. If it has neither a negative in H nor a positive in U, cutoff m is feasible and is returned.
5. Otherwise it has a positive in U and no negative in H. When the sample is realizable, every feasible cutoff exceeds m, so (A.22) selects the **minimum** feasible cutoff. Make one minimum-oracle query on the restriction of the sample to U: return $\perp$ if it returns $\perp$; otherwise extend its labels on U by assigning 1 to all of H (and to s under E).

Simulate predictions on H using their known labels, and use the learner's predictions on U as those of the derived minimum-oracle learner. The derived learner has no more queries and no more mistakes than the original learner.

Applying the already proved E-convention minimum-rule theorem to U, with a virtual anchor, proves (A.24). Under N, the nonempty block H supplies the anchor when embedding the resulting target back into the original instance.

Since $T - m \geq \lfloor T/2 \rfloor$, a deterministic learner cannot simultaneously achieve $O(\log T)$ mistakes and $O(\log T)$ queries under (A.22).

Thus the precise conclusions are:

Predeclared rule	Deterministic conclusion
Minimum / maximum / sample-dependent extremal selectors	Exact worst-case minimax total cost $T - \varepsilon$
Closest feasible cutoff to a fixed global median	Linear total-cost lower bound; full $T - \varepsilon$ bound is false
Median of the feasible cutoffs	Deterministic $O(\log T)$ queries and mistakes

The general-rule counterexamples do not refute Corollary 4.1 or Theorem 3.1: they change the publicly declared oracle rule.

9. The one-pin common-value principle in general

The proof of Theorem 3.1 uses the extremal rules only through Lemma A.2. The following statement records the abstract form of the argument.

Proposition A.8 — one-pin common-value principle

Let $\mathcal{R}$ be a selection rule announced before the learner and the instance, that is, a function $\mathcal{R}(S, \preceq)$ that is a legal consistency-type ERM oracle for every total order $\preceq$. Suppose that for every anchored state (L, F, R) and every finite sample S there is a refinement of (L, F, R) by at most one endpoint pin such that $\mathcal{R}(S, \preceq)$ takes one common value on all orders retained by the refined state. Then, under either convention, for every deterministic learner and every instance sequence there exist a total order $\preceq^\star$ and a target $z^\star$ such that, with the oracle $S \mapsto \mathcal{R}(S, \preceq^\star)$, the learner makes $M + Q \geq T - \varepsilon$.

Proof. Lemma A.2 enters the proof of Theorem 3.1 only by supplying, at each query, a refinement by at most one endpoint pin under which the answer is common to all retained orders. Given the hypothesis, the nesting (A.5), the invariant of Lemma A.3, the charging of Lemma A.4 and the fixation of Lemma A.5 apply verbatim with $\mathcal{R}$ in place of $\mathcal{O}_{\min}$: the answer recorded at each query equals $\mathcal{R}(S, \preceq)$ for every retained order, each query removes at most one free point, every point that is still free at its prediction costs one mistake, and the order and target chosen in Lemma A.5 fix one instance on which the recorded transcript is replayed. □

Lemma A.2 shows that $\mathcal{O}_{\min}$ satisfies the hypothesis; the duality of Lemma A.6 transfers it to $\mathcal{O}_{\max}$, and Section 6 of this appendix to every sample-dependent choice $\mathcal{R}_\eta$ between the two.

B Alternative proof of Corollary 4.1

Role of this appendix. This is an alternative proof of Corollary 4.1, not a prerequisite for Theorem 3.1. The common-response lemma feeds a cached transcript invariant; charging gives the cost bound and freezing turns the transcript into a legal memoryless oracle.

Corollary 4.1: an additive deterministic lower bound with a fixed memoryless ERM oracle

Conventions and quantifiers. Let $T \geq 1$, and announce the distinct instances

$$I_T = \{x_1, \ldots, x_T\}$$

in the fixed arrival sequence $(x_1, \ldots, x_T)$. Consider either of the following conventions (Section 2):

1. **Empty-prefix convention:** $X = I_T \cup \{s\}$, where s is placed before every instance. Put $\varepsilon = 0$.
2. **Nonempty-prefix convention:** $X = I_T$. Put $\varepsilon = 1$.

For every deterministic learner $\mathscr{A}$, there exist a total order $\preceq^*$ on X, a point $z^* \in X$, and a function

$$O^* : 2^{X \times \{0,1\}} \longrightarrow C_{\preceq^*} \cup \{\bot\}$$

such that:

- O^* is a legal consistency-type ERM oracle on **every** sample S;
- $O^*(S)$ depends only on S, not on the time, query history, or target;
- the order, target, instance sequence, and oracle function are all fixed before the actual interaction; and
- on this one fixed instance,

$$\boxed{M(\mathscr{A}) + Q(\mathscr{A}) \geq T - \varepsilon.}$$

Here the learner may submit arbitrary samples, including contradictory samples, samples involving previously labeled or future instances, and samples involving the non-instance s. It may query at arbitrary times and receives complete concept label vectors, as stipulated in the model of Section 2.

Thus the additive constant is 0 under the empty-prefix convention and 1 under the nonempty-prefix convention.

The proof is self-contained below.

1. A common anchored formulation

For both conventions, introduce an anchor a and a set V of initially free instances:

Convention	a	V	$n := \lvert V \rvert$
Empty prefix	s	I_T	T
Nonempty prefix	x_1	$I_T \setminus \{x_1\}$	$T - 1$

We commit to placing a first in the total order. Every concept consequently labels a by 1.

In the nonempty-prefix convention, choosing x_1 as the least element is simply an allowed adversarial choice of order. The argument would remain valid even if this information were given to the learner for free.

A state consists of:

- an ordered list L;
- an unordered set F;

- an ordered list R;

partitioning V. Its compatible orders are

$$\mathscr{E}(L, F, R) = \{\, a \cdot L \cdot \pi(F) \cdot R : \pi(F) \text{ is a permutation of } F \,\}. \tag{B.1}$$

Points in L are committed to target label 1, and points in R to target label 0. They need not already have arrived. Points in F have not arrived and remain free.

We use a list also for its underlying set when taking unions. For a nonempty $H \subseteq X$, write $h_H = \mathbf{1}_H$.

All lemmas below apply to **both conventions**, with $n = T - \varepsilon$.

Lemma B.1 — Prefix consistency criterion

Suppose a total order τ has a as its least element. For a sample S, define

$$P(S) = \{x : (x, 1) \in S\}, \qquad N(S) = \{x : (x, 0) \in S\}.$$

If $P(S) \cap N(S) \neq \emptyset$, or if $a \in N(S)$, then S is not realizable by C_τ.

Otherwise put

$$A = P(S) \setminus \{a\}, \qquad B = N(S).$$

Then S is realizable if and only if there is no pair $b \in B, a' \in A$ with

$$b \prec_\tau a'. \tag{B.2}$$

Proof. Contradictory labels cannot be realized (Section 2). Every nonempty prefix contains the least point a, so a negative label on a is also impossible.

An inversion as in (B.2) prevents any prefix from containing a' while excluding b. Conversely, if there is no inversion and $A \neq \emptyset$, the prefix ending at $\max_\tau A$ contains every positive sample point and no negative sample point. If $A = \emptyset$, the singleton prefix $\{a\}$ is consistent. □

Notice that the oracle tests realizability by the **class**, not agreement with the target. For example, a positive singleton query remains realizable even when that point has already received target label 0.

2. The one-pin query rule

The adversary keeps an initially empty table Ans, keyed by the entire sample set S.

The partial order $\prec_K$ of a state $K = (L, F, R)$ fixes the internal orders of L, R, puts L before F before R, and leaves distinct points of F incomparable.

A useful representation is the ordered list of **atoms**

$$W_K = (\ell_1, \ldots, \ell_p, [F], r_1, \ldots, r_q), \tag{B.3}$$

omitting $[F]$ when $F = \emptyset$. The entire free set is one atom.

Lemma B.2 — Exhaustive query handling with at most one pin

For every state K and sample S, the following rule either returns an answer already cached, or constructs a new answer legal for every order in the resulting state. A fresh query removes at most one point from F.

Rule and proof.

Cached sample. If $S \in \mathrm{dom}(\mathsf{Ans})$, return its stored answer without changing the state. Preservation of this answer is proved in Lemma B.3.

Why this rule needs the cache. Let initially $F = \{u, v\}$ and let the first query be $\{(u, 1)\}$; case (iii) below may answer with the vector that is positive on both u and v. Suppose u then arrives and is labelled 1, so that the state becomes $L = [u]$, $F = \{v\}$. If the same sample $\{(u, 1)\}$ were now treated

as a fresh query, the rule may legally return the prefix $[u]$ alone. Both answers are legal, but the same sample would receive two different full vectors, so without the first-answer cache the transcript could not be frozen into a function of the sample alone (Lemma B.5). This is a property of the particular adaptive answering rule used here; the fixed-rule proof in Appendix A needs no answer cache.

For a fresh sample, first handle contradictions or a negative label on a by returning $\perp$, without changing the state. Otherwise strip the redundant positive label on a, obtaining disjoint $A, B \subseteq V$.

There are exactly three remaining cases.

(i) A forced inversion already exists

If some $b \in B, a' \in A$ satisfy

$$b \prec_K a',$$

return $\perp$, with no state change.

Every order in $\mathscr{E}(K)$ contains this inversion, so Lemma B.1 proves the answer legal.

(ii) Both labels occur in the free block

Suppose there is no forced inversion, but

$$A \cap F \neq \emptyset, \qquad B \cap F \neq \emptyset.$$

Choose one $b \in B \cap F$, breaking ties by the public arrival indices, and update

$$L' = L \cdot b, \qquad F' = F \setminus \{b\}, \qquad R' = R. \tag{B.4}$$

Return $\perp$.

Because A, B are disjoint, some $a' \in A \cap F$ is different from b. Every new compatible order places b before this a'. Hence the sample is unrealizable in every new compatible order.

Moreover,

$$\mathscr{E}(L', F', R') \subseteq \mathscr{E}(L, F, R) :$$

the new orders are exactly old-compatible orders whose free permutation begins with b. Exactly one free point is removed.

Committing the target label of b to 1, although S asks for label 0 there, is legitimate: the oracle is answering that S is **not** realizable by the chosen class. It is not required to make the target consistent with S.

(iii) No forced inversion and no mixed-label free block

Leave the state unchanged and return a prefix of the atom list (B.3):

- if $A = \emptyset$, return $h_{\{a\}}$;
- otherwise, let j be the rightmost atom containing a point of A, and return the indicator of

$$H = \{a\} \cup \{\text{all points in atoms up to and including } j\}. \tag{B.5}$$

This contains every positive sample point. It contains no negative sample point: a negative point in an earlier atom would give a forced inversion, while a negative point in the same atom as a positive point could only occur inside F, which is excluded in this case.

The set H contains either all or none of F. Consequently it is a nonempty prefix of **every** order in $\mathscr{E}(K)$.

Recording. In each of cases (i)–(iii), including the $\perp$ answers of (i) and (ii), store the new answer as $\mathsf{Ans}[S]$.

Exhaustiveness. After contradictions are removed, two distinct sample points are incomparable precisely when both lie in F. Thus, if neither (i) nor (ii) applies, every positive–negative pair is comparable and correctly ordered, or one side is empty. This is exactly case (iii). □

The claim is about preserving a completely free residual block—not about the number of bits in a returned concept. A reply may encode many comparisons concerning already committed points.

3. Prediction rule and the full invariant

The following is an **offline transcript-construction device**. Its conversion into a legal oblivious instance is proved in Lemma B.5: the replay uses an order and target fixed before the interaction, and each label is revealed only after the learner has made its corresponding prediction (Section 2).

When the deterministic learner produces prediction $\hat{y}$ on the next instance x:

- if $x = a$ or $x \in L$, set $y = 1$, with no state change;
- if $x \in R$, set $y = 0$, with no state change;
- if $x \in F$, set

$$y = 1 - \hat{y}.$$

Remove x from F, and:

$$\begin{cases} L \leftarrow L \cdot x, & y = 1, \\ R \leftarrow x \cdot R, & y = 0. \end{cases} \tag{B.6}$$

Lemma B.3 — Strong invariant, including all recorded answers

Starting from $L = R = \emptyset$, $F = V$, the construction maintains all of the following after every event.

1. **Ordered partition and revealed labels.** The lists L, R and set F partition V. No point of F has arrived. Every revealed label is 1 on $L \cup \{a\}$ and 0 on R.
2. **Universal validity of affirmative records.** For every recorded affirmative answer h_H, the same complete vector h_H is a concept in C_τ consistent with its query, for every $\tau \in \mathscr{C}(L, F, R)$.
3. **Universal validity of negative records.** For every query recorded as $\perp$, that query is unrealizable by C_τ, for every $\tau \in \mathscr{C}(L, F, R)$.
4. **An entire interval of feasible targets.** For every permutation $\pi(F)$ and every $k \in \{0, \dots, |F|\}$, the prefix with positive set

$$\{a\} \cup L \cup \{\pi_1, \dots, \pi_k\} \tag{B.7}$$

agrees with every revealed label.

In particular, a consistent order and target always exist.

Proof. Initially there are no records or revealed labels, and every prefix in (B.7) is a valid nonempty prefix.

Every state-changing operation refines the compatible-order set:

- a query pin or a free prediction labeled 1 restricts the old free permutation to begin with the extracted point;
- a free prediction labeled 0 restricts it to end with that point.

Thus previously valid affirmative and negative records remain valid in every new compatible order. New query records are valid by Lemma B.2.

For completeness, target feasibility also persists as a joint order–target statement, not merely as an order statement. If a point u is appended to L, a new free permutation π' and cut k correspond to the old free permutation

$$(u, \pi')$$

and old cut $k + 1$. If u is prepended to R, they correspond to the old permutation

$$(\pi', u)$$

and old cut k. Hence every new pair in (B.7) was an admissible old pair. An extracted free point has not previously received a label, so the new commitment does not contradict history.

For prediction events, all new targets give the just-revealed label required by (B.6). An arriving point already in L, R, or the anchor has its committed label, so no change is needed.

A repeated query changes neither state nor table. If its sample is cached, its stored answer is legal by items 2 or 3. Otherwise, the repeated sample is contradictory or negative on a, and the preliminary branch returns $\perp$ again, legally by Lemma B.1. This also proves the cache branch of Lemma B.2.

Finally, $\mathscr{C}(L, F, R)$ is nonempty because every finite set F, including the empty set, has a permutation; (B.7) always contains the anchor and therefore defines a valid concept. □

A useful strengthening. Every binary labeling of the current F remains feasible: order its desired 1-points before its desired 0-points and choose the corresponding cut in (B.7). Thus no hidden target-label restriction on the residual free points has been silently accumulated.

4. Accounting

Lemma B.4 — Query pins pay for every possibly avoided free-point mistake

Under either convention, let p be the number of points removed from F by queries, and let f be the number removed when they arrive. After all T predictions,

$$n = p + f, \qquad p \le Q, \qquad f \le M. \tag{B.8}$$

Consequently,

$$M + Q \ge n = T - \varepsilon. \tag{B.9}$$

Proof. The potential

$$\Phi = |F|$$

starts at n and ends at 0, because every point of V appears exactly once in the transductive protocol.

A query decreases Φ by at most one. Cached queries, contradictory queries, existing-inversion queries, and affirmative replies do not decrease it. Thus $p \le Q$.

A prediction decreases Φ only if its own instance is still free. Such a prediction is wrong by (B.6). Hence $f \le M$.

These are the only ways points leave F, proving (B.8) and (B.9). The anchor's possible mistake in the nonempty-prefix convention is simply an additional, uncharged mistake. □

5. Freezing into an oblivious instance

Lemma B.5 — Memoryless freezing and exact deterministic replay

For either convention, every finite constructed transcript prefix can be frozen into a fixed order, target, and globally legal memoryless ERM function reproducing that prefix.

For a completed T-round construction, write the final state as $(L, \emptyset, R)$. An explicit freezing is

$$\tau^* = a \cdot L \cdot R, \qquad z^* = \mathrm{last}(a \cdot L). \tag{B.10}$$

The fixed target therefore has positive set $\{a\} \cup L$.

Let $\tau^* = (u_1, \dots, u_{n+1})$, and let $H_j = \{u_1, \dots, u_j\}$. Define

$$O^*(S) = \begin{cases} \mathsf{Ans}[S], & S \in \mathrm{dom}(\mathsf{Ans}), \\ h_{H_{j(S)}}, & S \notin \mathrm{dom}(\mathsf{Ans}) \text{ and a consistent prefix exists}, \\ \perp, & \text{otherwise}, \end{cases} \tag{B.11}$$

where

$$j(S) = \min\{j \in \{1, \dots, n+1\} : h_{H_j} \text{ is consistent with } S\}. \tag{B.12}$$

The table in (B.11) is an immutable parameter of the function, not mutable memory of calls during the actual interaction.

Proof.

Legality on every sample

By Lemma B.3, every recorded affirmative vector is a prefix concept of the final order and is consistent with its recorded sample. Every recorded $\perp$ sample is genuinely unrealizable by that same class.

For an unrecorded sample, (B.12) searches exactly all concepts of C_{τ^*}. It returns a consistent one precisely when one exists. This includes:

- $S = \emptyset$, for which the search succeeds;
- contradictory samples, for which it fails;
- samples containing a negative label on the anchor, for which it fails.

Thus (B.11) is a legal consistency ERM oracle on its whole domain. Its value depends only on S.

The non-$\perp$ returned objects are the stipulated complete label vectors. A vector recorded earlier may have had different possible endpoint names in different unfinished orders; after freezing it is exactly $c_{\max_{\tau^*} H}$. No endpoint identity was additionally promised by the interface.

Target legality

By Lemma B.3, the target in (B.10) agrees with every revealed label. It is a genuine concept, including when $L = \emptyset$: then $z^* = a$.

Replay

Run the same deterministic learner on the fixed domain, announced sequence, target c_{z^*}, and oracle O^*.

Induct on successive interaction events. The initial input is identical. If the histories agree so far, determinism implies that the learner's next operation, query set, or prediction agrees with the constructed one.

- At each query event of the constructed transcript being replayed, (B.11) returns the stored answer if the sample is recorded. Every queried but unrecorded sample is contradictory or negative on a, so (B.11) returns the same $\perp$ by Lemma B.1. This covers repetitions as well.
- At a label-revelation event, the fixed target returns exactly the constructed label.

The histories therefore remain identical. Complete returned vectors agree on every point of X, so arbitrary evaluations of returned concepts also agree.

For a finite, nonterminal transcript prefix, choose any $\tau^* \in \mathscr{C}(L, F, R)$ and any target cut from (B.7), then use the same definition (B.11). The same argument reproduces that prefix.

All adaptive choices have thus been compiled into one fixed triple. In its actual execution, each label $c_{z^*}(x_t)$ is fixed before the learner predicts, as required by the transductive protocol. □

Finiteness detail. It is enough to construct a transcript until all T predictions occur or until n oracle calls have occurred. In the latter case, freeze the current finite prefix using Lemma B.5; its replay already has $Q \geq n$. Otherwise the completed construction is finite. Thus an infinite-query behavior cannot obstruct the existential construction.

Completion of Corollary 4.1 and sharpness of the constants

Lemma B.4 gives the additive bound for the constructed transcript, and Lemma B.5 transfers it to one fixed oblivious instance.

The constants cannot be improved uniformly over $T \geq 1$:

- Under the main convention, at $T = 1$ a zero-query fixed-label predictor always has $M + Q \leq 1$. Thus no smaller universal additive constant is possible.

- Under the nonempty-prefix convention, at $T = 1$ the only target labels the sole instance by 1. Predicting 1 without querying gives $M + Q = 0$, ruling out a universal additive constant smaller than 1.

These witnesses certify only the additive constants; they are not an asserted general error–query upper tradeoff. □

6. Separation and relation to AHR25

Corollary — Deterministic low-mistake learning needs nearly T queries

Under the empty-prefix convention, suppose a deterministic learner guarantees at most $m(T)$ mistakes uniformly over the allowed fixed instances and legal memoryless ERM oracles. Then

$$Q_{\text{worst}}(\mathscr{A}, T) \geq T - m(T). \tag{B.13}$$

Under the nonempty-prefix convention, the corresponding bound is

$$Q_{\text{worst}}(\mathscr{A}, T) \geq T - 1 - m(T). \tag{B.14}$$

Proof. Apply Corollary 4.1. On the instance it provides, $M \leq m(T)$, so $Q \geq T - \varepsilon - m(T)$. □

In particular, a deterministic $O(\log T)$-mistake guarantee requires

$$Q_{\text{worst}} \geq T - \varepsilon - O(\log T). \tag{B.15}$$

By the randomized theorem of AHR25 (Thm. 4.5(4)), both expected mistakes and expected calls can instead be $O(\log T)$. Consequently:

- randomized expected total cost $M + Q$ is $O(\log T)$;
- every deterministic learner has worst-case total cost at least $T - \varepsilon$.

Equivalently, using the dimension conventions of Section 2, the gap is exponential when parameterized by $d = \Theta(\log T)$: logarithmic-in-T cost versus linear-in-T cost, or $O(d + 1)$ versus $2^{\Omega(d)}$.

This establishes a positive answer to the open question of AHR25 about the power of randomization **within the threshold-family, consistency-ERM, oblivious-instance setting of this paper**.

Why Theorem 4.4 of AHR25 is not the proof of this result

The statement of Theorem 4.4 of AHR25 is:

> "Let $\mathscr{F}$ be the family of all classes with Littlestone dimension d_{LD}. For $T \geq 2^{d_{\text{LD}}+1}$, when having access to the ERM oracle, if fewer than $T/2$ queries are made, at least $2^{d_{\text{LD}}} - 1$ mistakes will be made."

Its proof contains the sentence:

> "An essential property is that each query provides information about at most one point."

The present proof implements a related one-point charging principle, but that theorem is **not invoked as a lower-bound lemma**. Its construction uses threshold classes on equivalence classes, rather than the present threshold classes on distinct totally ordered points. Moreover, for either dimension convention of Section 2,

$$d = \lfloor \log_2(T + 1) \rfloor \quad \text{or} \quad d = \lfloor \log_2 T \rfloor,$$

the required condition $T \geq 2^{d+1}$ fails.

Here, forced-realizable queries are explicitly answered by proper full concepts, and all replies are then frozen into one memoryless oracle.

WHY THE CONSTRUCTION DOES NOT CONTRADICT THEOREM 4.5(4) OF AHR25

For a randomized learner, performing the construction separately for each random tape would generally produce

$$(\tau_\omega, z_\omega, O_\omega)$$

depending on that tape. In particular, which negative sample point is pinned may depend on the learner's random query.

This gives an instance **after seeing randomness**, not one fixed oblivious instance. The invalid quantifier exchange would be

$$\forall \omega\ \exists I_\omega \quad \Longrightarrow \quad \exists I \text{ with the same expected lower bound over } \omega.$$

Determinism permits the replay argument for one learner execution; it does not justify this exchange.

C PROOF OF THEOREM 3.2

Proof organization. Lemmas C.1–C.2 transfer the problem to the revelation oracle. Lemma C.3 gives the conditional law needed by the potential argument; Lemmas C.4–C.5 control refinement and the full adaptive call. Telescoping and Jensen's inequality yield Theorem 3.2, followed by its query consequence. Propositions C.7–C.9 examine a different distribution and are not premises of the main lower bound.

1. THEOREM 3.2

All logarithms in the proof are natural.

Let

$$\epsilon = \begin{cases} 0, & \text{under convention (E),} \\ 1, & \text{under convention (N),} \end{cases} \qquad n = T - \epsilon.$$

Under (E), there are $n = T$ prediction points and an additional sentinel $a = s$. Under (N), there are $n = T - 1$ hard prediction points and one additional prediction point a. The point a will be the minimum of the total order and will always have target label 1.

THEOREM 3.2 — EXPECTED-QUERY AND WORST-CASE-QUERY LOWER BOUNDS

For every $T \geq 1$, under either convention above, there exist:

1. a finite distribution $\mathscr{D}_T^\epsilon$ over total orders and target thresholds;
2. a fixed transductive sequence of T distinct points; and
3. an explicit deterministic, memoryless consistency-type ERM oracle

$$O(S, \preceq),$$

depending only on the sample and the total order,

such that every randomized learner satisfies

$$\boxed{\mathbb{E}[M] \;\geq\; \frac{(T + 1 - \epsilon)\, 128^{-\mathbb{E}[N_{\text{qry}}]} - 1}{2}.} \tag{C.1}$$

Here both expectations are over $\mathscr{D}_T^\epsilon$ and the learner's private randomness, and the assertion is substantive when the expected number of queries is finite.

Consequently:

- **Expected-query budget.** If $\mathbb{E}[N_{\text{qry}}] \leq Q$, then

$$\mathbb{E}[M] \geq \max\left\{0, \frac{(T + 1 - \epsilon)128^{-Q} - 1}{2}\right\}. \tag{C.2}$$

- **Worst-case-query budget.** If $N_{\mathrm{qry}} \le Q$ almost surely, the same bound holds.

In particular, for both conventions,

$$\boxed{\mathbb{E}[M] \ge \frac{1}{2}\, T \left(\frac{1}{128}\right)^{Q} - \frac{1}{2}.} \tag{C.3}$$

The expected-query hypothesis in (C.2) may be imposed merely on the average over the displayed hard distribution. In particular, the uniform, per-instance expected-query guarantee in the cost accounting of Section 2 implies it.

The construction and proof below are self-contained.

2. The hard distribution and the actual oracle

2.1 A balanced interval tree

Identify the hard points with $x_1, \dots, x_n$, in their arrival order.

Build a binary tree on the interval $[n]$. An interval of size $m \ge 2$ is split into its first $\lfloor m/2 \rfloor$ indices and its remaining $\lceil m/2 \rceil$ indices. Singletons are leaves.

For $n \ge 1$, draw

$$J \sim \mathrm{Unif}([n]), \qquad Y_1, \dots, Y_n \overset{\mathrm{iid}}{\sim} \mathrm{Bernoulli}(1/2),$$

independently.

Define a priority permutation π_J as follows:

- traverse the path from the root to leaf J;
- at every internal node, append all indices in the sibling not containing J, in increasing numerical order;
- finally append J.

Thus, every sibling encountered earlier has higher, or "outer," priority than every point in the current interval containing J.

2.2 From priorities and labels to a genuine total order

The total order is

$$a \prec (x_i : Y_i = 1, \text{ in } \pi_J \text{ order}) \prec (x_i : Y_i = 0, \text{ in reverse } \pi_J \text{ order}). \tag{C.4}$$

The target threshold ends at the last $Y_i = 1$ point in the first list, or at a if that list is empty. Therefore

$$c_z(a) = 1, \qquad c_z(x_i) = Y_i. \tag{C.5}$$

This is a total-order threshold class, not a threshold class on equivalence classes: every point occupies its own distinct position, and every point-prefix belongs to the class.

The transductive sequence is:

- under (E): $x_1, \dots, x_n$, with $a = s$ a non-instance sentinel;
- under (N): $x_1, \dots, x_n, a$.

If $n = 0$, which occurs only for $T = 1$ under (N), use the singleton order consisting of a, with target label 1.

All random choices are made before interaction. This defines $\mathscr{D}_T^{c}$, independently of the learner.

2.3 The actual, memoryless oracle

Use the **minimum consistent prefix oracle**. For completeness, the following definition is valid on any nonempty finite totally ordered domain.

Given $S \subseteq X \times \{0, 1\}$:

1. If S assigns both labels to some point, return $\perp$.
2. Let
$$P = \{x : (x, 1) \in S\}, \qquad B = \{x : (x, 0) \in S\}.$$
3. Set
$$p = \begin{cases} \max_{\preceq} P, & P \neq \emptyset, \\ \min_{\preceq} X, & P = \emptyset. \end{cases}$$
4. If some $b \in B$ satisfies $b \preceq p$, return $\perp$. Otherwise return the full vector c_p.

This rule depends only on $(S, \preceq)$, not on J, the target, the history, or the learner's private randomness.

It is a legal oracle: every threshold containing P must contain the prefix ending at p; thus an element of B in that prefix makes consistency impossible. Otherwise c_p itself is consistent. This also covers empty and one-sided samples. On our support, every concept contains a, so a query containing $(a, 0)$ is correctly rejected.

3. Removing failure answers by a stronger oracle

The next reduction is important: it handles arbitrary B, including all information carried by $\perp$ answers.

Lemma C.1 — prefix-max domination

On the constructed instances, every call to O can be simulated using one call to a stronger oracle which, on any set $A \subseteq [n]$, returns the full prefix ending at
$$\max_{\preceq} \{x_i : i \in A\},$$
or at a when $A = \emptyset$.

Proof Take A to be the hard points assigned label 1 by the sample. The stronger response is precisely the candidate minimum prefix in the definition of O. Checking the sample's 0-labeled points against that full vector determines whether O returns this prefix or $\perp$. Contradictory samples and $(a, 0)$ may be rejected immediately.

Thus one stronger call determines the original answer for every sample, regardless of its size or the arrival status of its points. □

We now strengthen this prefix-max oracle further.

The auxiliary revelation oracle

The auxiliary oracle maintains an interval I containing J, initially $I = [n]$. It reveals to the learner:

- every target label outside I;
- the priority order of all those outside points;
- the interval I itself.

Such information may include future labels. This is deliberately additional information.

For a query $A \subseteq [n]$, it proceeds as follows.

1. If $A \setminus I$ contains a point with target label 0, stop.

2. If $A \cap I = \emptyset$, stop.
3. If I is a singleton, reveal its label, set $I = \emptyset$, and stop.
4. Otherwise reveal which child of I contains J, reveal all target labels in the other child, replace I by the child containing J, and repeat.

At termination it returns the requested full prefix as well as the auxiliary information.

The entire procedure costs **one query**, however many internal revelations occur.

This is only an auxiliary, stronger information source. It is not asserted to be the actual ERM oracle. Its memory and its access to target labels do not enter the definition of $O(S, \preceq)$.

LEMMA C.2 — THE AUXILIARY ORACLE SIMULATES THE FULL PREFIX

At every stopping point above, the requested prefix is determined by the revealed information. In particular, an original learner can be simulated by a learner using the auxiliary oracle, with the same number of queries and the same prediction distribution.

Proof All points outside the current I precede all points inside I in priority order.

Suppose $A \setminus I$ contains a target-negative point. The maximum of A in the total order is then the **outermost target-negative point of** A: the one appearing earliest in π_J. It lies outside I, and its identity is known.

Its prefix contains:

- every point of I;
- every outside target-positive point; and
- exactly those outside target-negative points whose priority is no earlier than its own.

The entire vector is therefore known.

If there is no outside target-negative point and $A \cap I = \emptyset$, all points of A are revealed target-positive points. Their total-order maximum is their latest point in priority order. Its prefix contains no point of I, and its values outside I are known. The empty-A case returns the prefix $\{a\}$.

The singleton revelation makes all remaining information available. The procedure terminates because every nonterminal refinement reduces the interval size.

Combining this computation with Lemma C.1 reconstructs every original response, including $\perp$. An original learner may simply ignore the additional information. □

A WORKED EXAMPLE

Let $n = 8$. The interval tree splits $[8]$ into $[1, 4]$ and $[5, 8]$, then $[5, 8]$ into $[5, 6]$ and $[7, 8]$, and so on. Take $J = 6$: the path from the root passes through $[5, 8]$ and $[5, 6]$, so π_J lists the siblings $[1, 4]$, $[7, 8]$, $\{5\}$ and finally J: $\pi_J = (1, 2, 3, 4, 7, 8, 5, 6)$. With labels $(Y_1, \dots, Y_8) = (1, 0, 1, 0, 1, 1, 0, 1)$, (C.4) gives the total order

$$a \prec x_1 \prec x_3 \prec x_8 \prec x_5 \prec x_6 \prec x_7 \prec x_4 \prec x_2$$

(the positives $1, 3, 8, 5, 6$ in π_J order, then the negatives $7, 4, 2$ in reverse π_J order), and the target ends at x_6, so that $c_z(x_i) = Y_i$. Suppose the first action is the query $S = \{(x_3, 1), (x_7, 1), (x_2, 0)\}$. The actual oracle takes $p = \max_{\preceq}\{x_3, x_7\} = x_7$; since $x_7 \prec x_2$, the negative point x_2 lies outside that prefix and it returns the full vector c_{x_7}, which is 1 on $a, x_1, x_3, x_8, x_5, x_6, x_7$ and 0 on x_4, x_2. The auxiliary oracle, started with $I = [8]$ and $A = \{3, 7\}$, refines twice: it reveals that $J \in [5, 8]$ together with $Y_1, \dots, Y_4$, then that $J \in [5, 6]$ together with Y_7, Y_8, and stops because $x_7 \in A \setminus I$ is target-negative. As in the proof of Lemma C.2, x_7 is the outermost target-negative point of A, and its prefix consists of the anchor a, the points x_5, x_6 of I, the revealed outside positives x_1, x_3, x_8 and x_7 itself: exactly the 1-set of c_{x_7}. The single call revealed six labels, but in the potential $\Phi = \ln(H + r + 1)$ of Lemma C.4 below (here $H = 0$ and $r = |I|$, no prediction having been made) it only reduced Φ from $\ln 9$ to $\ln 3$, at most $\ln 2$ per refinement.

Offline instance: $n = 8$ **hard points, hidden center** $J = 6$

Arrival order: $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8$.
Priority permutation: $\pi_J = (1, 2, 3, 4 \mid 7, 8 \mid 5 \mid 6)$.
True threshold order:

$$\underbrace{a \prec x_1 \prec x_3 \prec x_8 \prec x_5 \prec x_6}_{\text{target prefix; cutoff } z=x_6} \prec x_7 \prec x_4 \prec x_2.$$

One auxiliary call: two refinements

$S = \{(x_3, 1), (x_7, 1), (x_2, 0)\}, \quad A = \{3, 7\}.$

$$\underbrace{\begin{array}{c} I = [1,8] \\ \Phi = \ln 9 \end{array} \xrightarrow{\text{reveal } Y_1, Y_2, Y_3, Y_4} \begin{array}{c} I = [5,8] \\ \Phi = \ln 5 \end{array} \xrightarrow{\text{reveal } Y_7, Y_8} \begin{array}{c} I = [5,6] \\ \Phi = \ln 3 \end{array}}_{\text{one auxiliary call}}$$

Stop at $[5, 6]$**:** $7 \in A \setminus I$ and $Y_7 = 0$.
No predictions yet: $H = 0$ throughout; $r = 8 \to 4 \to 2$.

Actual ERM reply: the full concept c_{x_7}

Arrival index i	1	2	3	4	5	6	7	8
Target Y_i	1	0	1	0	1	1	0	1
Returned $c_{x_7}(x_i)$	1	0	1	0	1	1	1	1

Both target and returned concept label the anchor a by 1.

Figure 1: The eight-point worked example and its first auxiliary call (Section 5). The upper strip describes the offline instance, not the learner's initial information. Priority and threshold order are different objects. Two refinements count as one call, stopping when the revealed point x_7 has target label 0 although the query requests 1. The actual memoryless ERM oracle returns only c_{x_7}; the six extra target-label revelations belong to the auxiliary simulation. The potential values describe this realization, not an expected curve or a per-call worst-case bound.

4. The posterior invariant

Lemma C.3 — conditional independence inside the active interval

At every point of the auxiliary interaction, including between internal revelations of a query, conditional on the complete auxiliary history:

1. J is uniform in the current nonempty interval I;
2. all labels of unarrived points in I are independent fair bits;
3. these unarrived labels are independent of J.

The statement remains valid for randomized learners.

Proof Initially this is the definition of the distribution.

Suppose the statement holds for I. The next child containing J has probability equal to its size divided by $|I|$. Conditional on that child, J is uniform within it.

The labels revealed in the sibling are independent of J within the retained child and independent of all unarrived labels in that child. The priority ordering of the revealed sibling is fixed by its numerical indices and does not depend on the remaining location of J.

The decision to stop or continue uses only:

- the chosen query;

- the retained child's identity;
- the revealed labels outside it; and
- whether the query has any point in the retained child.

It does not depend on the location of J within that child or on its unrevealed future labels. Hence conditioning on that decision preserves the asserted law. The final full-prefix response is determined by the revealed information, by Lemma C.2, and creates no additional conditioning.

An arriving label in I is a fair bit independent of J and the remaining unarrived labels. Revealing it therefore also preserves the invariant.

For randomized learners, include the learner's private seed in the conditioning, or apply the same induction after fixing that seed. Its independence from the hard instance supplies the base case. □

This explicitly replaces, rather than assumes, exchangeability after failure answers.

5. Charging free labels to predictions

Call a prediction **unrevealed** if its point is still in the active interval immediately before that prediction. Let H be the number of such hard-point predictions made so far; H is a cumulative count over the whole interaction and never decreases when the active interval shrinks (it must not be confused with the number k of already-predicted points inside the current interval, for which $H \geq k$).

Queries between a prediction and its label feedback can be treated as post-feedback queries in a stronger protocol: provide the label immediately after the prediction, and let a simulator ignore it until the original learner would have received it. The prediction loss is unchanged. Thus it suffices to analyze queries between prediction-feedback steps.

Lemma C.4 — mistake charge and the potential

Let the active interval have size m. Let k of its points have already been predicted, and let $r = m - k$. Then:

1. $H \geq k$;
2. $\mathbb{E}[M] \geq \frac{1}{2}\mathbb{E}[H_{\text{final}}]$;
3. the potential
$$\Phi = \ln(H + r + 1) \tag{C.6}$$
is unchanged by prediction-feedback steps;
4. one binary interval refinement decreases Φ by at most $\ln 2$.

Proof The active intervals are nested. Every already-predicted point still belonging to the current interval also belonged to the active interval when it was predicted. Hence $H \geq k$.

By Lemma C.3, an unrevealed prediction faces an independent fair target bit. Its conditional expected loss is exactly $1/2$, regardless of the prediction distribution. Summing these conditional expectations proves assertion 2.

On an unrevealed prediction, H increases by one and r decreases by one. Otherwise neither changes. This proves assertion 3.

For assertion 4, write m', k' for the retained child's size and number of already-predicted points. We have
$$m' \geq \lfloor m/2 \rfloor, \qquad k' \leq k, \qquad H \geq k.$$
Consequently,
$$H + m - k + 1 \leq 2(H + m' - k' + 1). \tag{C.7}$$
Indeed, with $d = H - k \geq 0$, the left side is $d + m + 1$, whereas the new quantity is at least $d + m' + 1$, and
$$d + m + 1 \leq 2(d + m' + 1).$$

Taking logarithms proves the assertion. □

The +1 is essential for handling empty residual sets and the last leaf without undefined logarithms.

6. The single-query information lemma

Lemma C.5 — bounded expected logarithmic decrease per query

Conditional on any auxiliary history immediately before a query,

$$\boxed{\mathbb{E}[\,\Phi_{\text{before}} - \Phi_{\text{after}} \mid \text{history}] \le 7 \ln 2.} \tag{C.8}$$

This holds for every query set A, including sets chosen using arbitrarily many previously observed labels.

Proof During one query, H is fixed.

Call a state **early** if

$$k < \lfloor m/2 \rfloor,$$

and **late** otherwise.

Late states

In a late state,

$$r = m - k \le k + 1 \le H + 1.$$

Therefore

$$H + r + 1 \le 2(H + 1). \tag{C.9}$$

The potential at the end of the query is at least $\ln(H + 1)$. Thus, from the first late state onward, the entire remaining potential decrease is at most $\ln 2$, regardless of how many further revelations occur.

This also covers a singleton interval.

Early states

Consider an early state in which the query needs another refinement. Let I_L, I_R be the left and right children. Here a point counts as arrived once its prediction-feedback step has occurred; since Section 5 of this appendix reduces the analysis to queries between such steps, k equals the number of arrived points of I, and since arrivals follow the numerical order, these k points form an initial segment of I.

Because arrival order is numerical order and fewer than $|I_L|$ points of I have arrived, **every point of I_R is unarrived**. Its labels are therefore independent fair bits by Lemma C.3.

There are two exhaustive cases.

- If $A \cap I_R = \emptyset$, then when $J \in I_R$, the new active interval contains no point of A, so the query stops. This event has probability

$$\frac{|I_R|}{m} \ge \frac{1}{2}.$$

- If $A \cap I_R \neq \emptyset$, consider the event that $J \in I_L$ and at least one point of $A \cap I_R$ has target label 0. The right child is then revealed, and the query stops because an outside target-negative point of A has been found. Its conditional probability is

$$\frac{|I_L|}{m}\left(1 - 2^{-|A \cap I_R|}\right) \ge \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}. \tag{C.10}$$

Thus, after each early-state refinement, the conditional probability of continuing to another early-state refinement is at most $5/6$.

Let D be the number of early-state refinements before the query stops or first reaches a late state. No independence between refinements is required:

$$\Pr(D \geq j) \leq (5/6)^{j-1}, \qquad \mathbb{E}[D] \leq 6. \tag{C.11}$$

Each refinement costs at most $\ln 2$ in potential, by Lemma C.4. The remaining late-state part costs at most another $\ln 2$. Hence

$$\mathbb{E}[\Phi_{\text{before}} - \Phi_{\text{after}}] \leq 6 \ln 2 + \ln 2.$$

This proves (C.8). □

The two mechanisms in this lemma are complementary:

- querying genuinely future points encounters an independent wrong-side point with constant probability;
- exploiting many already known labels is permitted, but their predictions have already contributed to H.

One call, two complementary accounting parts

Condition on the full auxiliary history h before the call.
The cumulative unrevealed-prediction counter H stays fixed during the call.

$m = |I|$; k counts already predicted points in I; $r = m - k$.
Early means $k < \lfloor m/2 \rfloor$; late means $k \geq \lfloor m/2 \rfloor$.

Early refinements
Before stopping or first reaching late

Unarrived right-child labels are independent fair bits (Lemma C.3).

After each required early refinement, stop with probability at least $1/6$, conditional on its current history (C.10).

Their count D satisfies $\mathbb{E}[D \mid h] \leq 6$; independence of refinements is unnecessary (C.11).

Each refinement decreases Φ by at most $\ln 2$ (Lemma C.4).

Expected early decrease
$\leq 6 \ln 2$.

Entire remainder
From the first late state onward

At that state, $H \geq k$ and

$$r \leq k + 1 \leq H + 1.$$

Hence, by (C.9),

$$H + r + 1 \leq 2(H + 1).$$

At the end, $\Phi \geq \ln(H + 1)$. Thus all remaining revelations share one total budget.

Entire remaining decrease
$\leq \ln 2$.

A call can include both parts; add their contributions.

$$\mathbb{E}[\Phi_{\text{before}} - \Phi_{\text{after}} \mid h] \leq 6 \ln 2 + \ln 2 = 7 \ln 2. \tag{C.8}$$

Figure 2: Local accounting in Lemma C.5. Conditional on the pre-query auxiliary history, early refinements before stopping or first reaching a late state contribute at most $6 \ln 2$ in expectation. From the first late state onward, the entire remaining decrease is at most $\ln 2$, irrespective of the number of later revelations or whether later states remain late. A call can include both parts. The cumulative counter H is fixed during the call; $7 \ln 2$ bounds a conditional expectation, not the pathwise amount of revelation.

7. Proof of Theorem 3.2

Initially,

$$H = 0, \qquad r = n, \qquad \Phi_{\text{initial}} = \ln(n + 1).$$

After all hard points have been predicted,

$$r = 0, \qquad \Phi_{\text{final}} = \ln(H_{\text{final}} + 1).$$

Prediction-feedback steps leave the potential unchanged. Let $\Delta_j \geq 0$ be its decrease during the j-th query, setting $\Delta_j = 0$ when that query does not occur.

Lemma C.5 gives

$$\mathbb{E}[\Delta_j] \leq 7 \ln 2 \, \Pr(N_{\text{qry}} \geq j).$$

Summing nonnegative quantities,

$$\begin{aligned} \ln(n+1) - \mathbb{E}[\ln(H_{\text{final}} + 1)] &= \mathbb{E}\left[\sum_j \Delta_j\right] \\ &\leq 7 \ln 2 \sum_j \Pr(N_{\text{qry}} \geq j) \\ &= 7 \ln 2 \, \mathbb{E}[N_{\text{qry}}]. \end{aligned} \tag{C.12}$$

This argument directly permits an adaptive, random number of queries. It does not condition the hard distribution on a low-query event.

By concavity of ln,

$$\ln(\mathbb{E}[H_{\text{final}}] + 1) \geq \mathbb{E}[\ln(H_{\text{final}} + 1)].$$

Combining this with (C.12),

$$\mathbb{E}[H_{\text{final}}] \geq (n+1) \exp(-7 \ln 2 \, \mathbb{E}[N_{\text{qry}}]) - 1. \tag{C.13}$$

Lemma C.4 now yields

$$\mathbb{E}[M] \geq \frac{(n+1)128^{-\mathbb{E}[N_{\text{qry}}]} - 1}{2}.$$

This lower bound holds even for learners receiving the auxiliary information. By Lemmas C.1 and C.2, it therefore holds for every learner using the actual memoryless oracle.

Substituting $n = T - \epsilon$ proves (C.1). Monotonicity in the query budget proves both versions of (C.2).

For $n = 0$, the instance has one known positive point and the claimed lower bound is nonpositive; it is immediate. Thus all boundary cases are included. □

Oblivious fixed-instance consequence

Suppose a learner satisfies the uniform expected-query guarantee of the cost accounting of Section 2:

$$\sup_{\leq, z} \mathbb{E}_{\mathscr{A}}[N_{\text{qry}}] \leq Q.$$

Its average query count under $\mathscr{D}_T^{\epsilon}$ is at most Q. Since the distribution has finite support, some fixed support instance satisfies

$$\mathbb{E}_{\mathscr{A}}[M] \geq \frac{(T+1-\epsilon)128^{-Q} - 1}{2}.$$

The query guarantee also holds on that instance by hypothesis.

The order, target, sequence, and oracle are already fixed before the learner's private randomness. No adaptive-adversary solidification step is needed.

8. Consequence for logarithmic query complexity

Corollary C.6

Under either convention, suppose a randomized learner has worst-case expected mistakes at most m_T and worst-case expected queries at most q_T, in the sense of the cost accounting of Section 2. Then

$$q_T \geq \frac{\ln(T+1-\epsilon) - \ln(2m_T + 1)}{7 \ln 2}. \tag{C.14}$$

In particular, for every fixed constant $\kappa \geq 0$ independent of T, if $m_T = O((\ln T)^{\kappa})$ then

$$q_T \geq \frac{\ln T - \kappa \ln \ln T - O(1)}{7 \ln 2} = \Omega(\ln T). \tag{C.15}$$

Proof Apply Theorem 3.2 to the learner and rearrange. □

Together with AHR25, Thm. 4.5(4), this proves that $O(\log T)$ expected oracle calls are optimal in order for achieving polylogarithmic expected mistakes.

Together with Corollary 4.1 and AHR25, Thm. 4.5(3), it also closes the stated comparison at $O(\log T)$ mistakes:

- randomized query complexity: $\Theta(\log T)$, in expectation;
- deterministic query complexity: $\Theta(T)$ against worst-case legal oracles and under the extremal rules of Theorem 3.1 (not under every legal pre-declared rule, Theorem 3.4).

This does **not** claim sharp constants or a complete pointwise tight characterization for every possible mistake/query budget.

Relation to the one-point property

The "one point per query" property used in the proof of Theorem 4.4 of AHR25 is not used here and generally fails for the present full-vector ERM interface: a query in our analysis may reveal a whole sibling interval, containing a linear number of labels. The replacement property is Lemma C.5: one query decreases a **paid logarithmic potential** by only a constant in expectation. Theorem 4.4 of AHR25 is not invoked here; its statement, and why its hypothesis fails for the threshold family, are discussed once in Appendix B.

9. Supplementary discussion: exact zero-query analysis for the uniform distribution

The results of Sections 9–11 are not used in the proof of Theorem 3.2; they explain why the uniform distribution cannot replace the hard distribution. The following results use the uniform distribution under convention (E), namely a uniformly random total order of the T instances together with an independent uniformly random cutoff among the $T+1$ prefixes (see Proposition C.7 for an equivalent representation); for $T \geq 2$, this differs from the distribution in Theorem 3.2.

Proposition C.7 — exact zero-query optimum

Under the uniform distribution, the minimum expected number of mistakes of any zero-query learner, including randomized learners, is

$$\boxed{F(T) = \sum_{t=1}^{T} \frac{\lfloor (t+1)^2/4 \rfloor}{t(t+1)}.} \tag{C.16}$$

In particular,

$$F(T) \geq T/4, \qquad F(T) = T/4 + \tfrac{1}{4}\ln T + O(1). \tag{C.17}$$

After observing k positive and m negative labels, the next label has posterior probability

$$\Pr(Y_{\text{next}} = 1 \mid \text{history}) = \frac{k+1}{k+m+2}. \tag{C.18}$$

Proof An equivalent representation of the uniform distribution is to draw independent

$$P, U_1, \dots, U_T \sim \text{Unif}[0,1]$$

and set $Y_i = \mathbf{1}[U_i \leq P]$. Ordering the U_i's gives a uniform permutation, and the position of P among them is uniform on $\{0, \dots, T\}$. The actual threshold point remains an element of X: it is the last instance below P, or the sentinel.

For a history with k ones and m zeros, the posterior density of P is proportional to $p^k(1-p)^m$. Taking the ratio of the elementary integrals for its first moment proves (C.18).

After r observations, the number K_r of ones is uniform on $\{0, \dots, r\}$. Therefore the expected Bayes error on round $r+1$ is

$$\frac{1}{(r+1)(r+2)} \sum_{k=0}^{r} \min(k+1, r-k+1) = \frac{\lfloor (r+2)^2/4 \rfloor}{(r+1)(r+2)}.$$

Summing proves (C.16).

A randomized prediction cannot improve on the posterior-majority decision, since its conditional loss is affine in the probability of predicting 1. Posterior majority attains every summand, proving exact optimality.

Finally, the round-t summand equals

$$\frac{1}{4} + \begin{cases} \frac{1}{4t}, & t \text{ odd}, \\ \frac{1}{4(t+1)}, & t \text{ even}. \end{cases}$$

This proves (C.17). □

Convention (N) correction. If the literal no-empty-prefix counterpart draws a uniform permutation and a uniform cut in $\{1, \dots, T\}$, its label distribution is the (E) distribution conditioned on a nonempty positive set. The total variation distance is $1/(T+1)$. Since total loss lies in $[0, T]$, its optimal zero-query value $F_N(T)$ satisfies

$$|F_N(T) - F(T)| \le \frac{T}{T+1} < 1. \tag{C.19}$$

In particular, $F_N(T) \ge T/4 - 1$.

Notice that the exact correction in (C.17) is logarithmic and positive; the exact optimum is not $T/4 + O(1)$.

10. Why the uniform distribution cannot prove Theorem 3.2

Proposition C.8 — a one-query learner for the uniform distribution

Fix $1 \le m \le T$. Predict the first m labels by posterior majority. Then query all m observed labels once and predict the returned hypothesis on all remaining points.

For the minimum-prefix rule on the uniform distribution, this algorithm has exactly

$$\boxed{\mathbb{E}[M] = F(m) + \frac{T-m}{m+2}.} \tag{C.20}$$

For **every** consistent returned threshold, it satisfies

$$\mathbb{E}[M] \le F(m) + \frac{2(T-m)}{m+2}. \tag{C.21}$$

Thus one query suffices for $O(\sqrt{T})$ expected mistakes on the uniform distribution.

Proof Use the representation in Proposition C.7. Let

$$L = \max\{U_i : i \le m,\ Y_i = 1\},$$

with $L = 0$ if there is no observed positive point.

The minimum-prefix hypothesis makes future mistakes exactly on $(L, P]$. Conditional on $P = p$,

$$\mathbb{E}[p - L \mid P = p] = \int_0^p (1-v)^m \, dv.$$

Averaging over p,

$$\mathbb{E}[P - L] = \frac{1}{m+2}. \tag{C.22}$$

The coordinate of each future point is independent of the first m points and of P, proving (C.20).

For arbitrary consistent tie-breaking, let U be the smallest observed negative coordinate, with $U = 1$ if none exists. Every consistent threshold agrees with the target outside (L, U), regardless of how its choice depends on the full order. By symmetry,

$$\mathbb{E}[U - P] = \frac{1}{m+2}.$$

Bounding mistakes by membership in (L, U) proves (C.21). □

For example, with $m = \lceil\sqrt{T}\rceil$, using $F(m) \le m/2$, (C.21) gives the explicit bound

$$\mathbb{E}[M] \le \frac{5}{2}\sqrt{T} + \frac{1}{2}. \tag{C.23}$$

This is a distribution-specific upper bound, not a worst-case upper bound over all orders. It does not contradict Corollary 4.1. It does prove that **no positive-constant $T\rho^Q$ lower bound with $Q = 1$ can hold on the uniform distribution**, even if the oracle's consistent tie-breaking is chosen adversarially.

11. A counterexample to a naive per-history contraction bound

Proposition C.9 — no history-uniform constant contraction bound for the uniform distribution

For the minimum-prefix rule on the uniform distribution (convention E), let U be the set of unarrived points whose labels are undetermined by the history h, and let U' be that set after one further query. Even after regularizing block sizes by $+1$, there is no constant C, independent of T and of h, such that $\mathbb{E}[\ln((|U|+1)/(|U'|+1)) \mid h] \le C$ for every positive-probability history h and every admissible next query. This is a uniform-distribution statement about $\ln(|U|+1)$; Lemma C.5 states a bound for $\Phi = \ln(H + r + 1)$ under the hard distribution.

Proof Choose integers $k \ge 1$, $u \ge 1$, and $T = k + u + 1$. Put $b = x_{k+1}$.

Before prediction, query the singleton positive sample $\{(b, 1)\}$. Consider the history in which:

1. the answer is the all-one vector, so b is the maximum instance in the unknown order;
2. the first k arriving labels are all 1;
3. the label of b is 0.

Under the uniform distribution this history has probability

$$\frac{1}{T} \cdot \frac{T}{T+1} \cdot \frac{1}{k+1} = \frac{1}{(T+1)(k+1)} > 0. \tag{C.24}$$

Before the next query, all u unarrived points have undetermined labels. Query the k known positive points. Let U' be the set of unarrived points above their maximum.

Removing the known maximum b, the conditional model is again a uniform random order and uniform cut, now conditioned on k specified positive labels. The expected coordinate of their maximum is $k/(k+2)$. Hence

$$\mathbb{E}[|U'| \mid \text{history}] = \frac{2u}{k+2}. \tag{C.25}$$

All points removed from the uncertainty set are certified positive; the points in U' remain label-undetermined.

By Jensen's inequality,

$$\mathbb{E}\left[\ln\frac{u+1}{|U'|+1}\,\middle|\,\text{history}\right] \ge \ln\frac{u+1}{2u/(k+2)+1}. \tag{C.26}$$

Taking $u = (k+2)^2$, the right side becomes

$$\ln\frac{(k+2)^2+1}{2k+5},$$

which diverges with k.

All parameters and probabilities here are finite and rational. For one explicit member,

$$k = 255, \quad u = 66049, \quad T = 66305,$$

the history probability is $1/16974336$, and

$$\mathbb{E}[|U'|] = 514, \qquad \frac{\mathbb{E}[|U'|+1]}{u+1} = \frac{103}{13210}.$$

Thus the expected logarithmic decrease is at least

$$\ln(13210/103) > \ln 128.$$

The parametrized family rules out every universal constant, not merely $7\ln 2$. □

The failure is not caused by taking a logarithm of zero. It persists with $+1$ regularization. The posterior on the threshold within the remaining label-undetermined set is not the fresh uniform posterior assumed by a naive potential argument.

D Deterministic mistake–query bounds at a fixed budget

Proof organization. The lower bounds combine Corollary 4.1 with the known-order benchmark; convention N additionally needs a first-action recurrence. The upper bounds use protection and sorting. Their agreement gives the exact budget regimes, while the finite games disprove the candidate identity in the intermediate regime.

Theorem D.1 (Deterministic mistake–query bounds)**.** *For a nonnegative integer Q, let $M_E^*(T,Q)$ (resp. M_N^*) be the minimax number of mistakes of deterministic learners making at most Q ERM calls, against an adversarial oracle as in Corollary 4.1.*

(i) $\max\{T-Q, \lfloor \log_2(T+1) \rfloor\} \le M_E^*(T,Q) \le \max\{T-Q, \lceil T/2 \rceil\}$*; under N,* $M_N^*(T,0) = M_N^*(T,1) = T-1$ *and for* $Q \ge 1$, $\max\{T-Q, \lfloor \log_2 T \rfloor\} \le M_N^*(T,Q) \le \min\{T-1, \max\{T-Q, \lceil T/2 \rceil\}\}$*. (The bound $T-Q$ under N is a budget statement for $Q \ge 1$, proved by a first-action recurrence $M_N^*(n,q) \ge \min\{M_E^*(n-1,q-1),\ 1+M_E^*(n-1,q),\ 1+M_N^*(n-1,q)\}$ for $n \ge 2$, $q \ge 1$ and induction; it is not a pathwise statement.)*

(ii) For $Q \le \lfloor T/2 \rfloor$, $M_E^*(T,Q) = T-Q$; for $Q \ge 2(T-1)$, $M_E^*(T,Q) = \lfloor \log_2(T+1) \rfloor$ and $M_N^*(T,Q) = \lfloor \log_2 T \rfloor$ (this budget is sufficient, not necessary).

(iii) The lower bound in (i) is not the exact value: $M_E^*(5,3) = 3$ and $M_N^*(6,4) = 3$.

Under E the lower bounds combine Corollary 4.1 with the known-order lower bound $\lfloor \log_2(T+1-\varepsilon) \rfloor$ (Lemma D.10); under N the stronger $T-Q$ budget bound for $Q \ge 1$ additionally uses the first-action recurrence in Theorem D.1(i) and is not a pathwise consequence of Corollary 4.1; the upper bounds come from a protection algorithm whose accounting is given below, and from sorting with $2(T-1)$ calls followed by majority prediction. The counterexamples are established analytically: a five-point game analysis under E, and a first-action recurrence reducing the N case to it.

1. Conventions and the counterexamples

Budgets below are nonnegative integers. A nonnegative real-valued budget B is understood as the integer parameter $Q := \lfloor B \rfloor$ before any budget case is selected or any theorem, lemma, bound, or recurrence is applied or evaluated.

Write $M_E^*(T,Q)$ and $M_N^*(T,Q)$ for the deterministic minimax mistake counts under conventions E and N, respectively. Lower bounds use the finite domains fixed in Section 2:

- E: $X = \{s\} \cup V$, where s is the least point and V consists of the T instances;
- N: $X = V$.

All non-$\perp$ oracle answers are complete label vectors. No threshold endpoint name or additional representation information is assumed.

THEOREM D.1(III) — THE CANDIDATE FORMULA IS NOT EXACT (CONVENTIONS E AND N)

In the model of Section 2,

$$M_E^*(5,3) = 3, \qquad M_N^*(6,4) = 3.$$

In particular, the candidate E formula

$$M_E^*(T,Q) = \max\{T-Q, \lfloor \log_2(T+1) \rfloor\}$$

is false. Consequently no formula that assigns the value 2 to $M_N^*(6,4)$ can be correct either.

The proof is analytic.

2. THE PROTECTION ALGORITHM

LEMMA D.2 — COMPUTABLE KNOWLEDGE AND SOUND PROTECTION

Maintain a partial order P_t on the instance set V. Its relations comprise:

1. relations learned from previous oracle answers;
2. every relation $a \prec b$ implied by an observed positive a and an observed negative b;
3. their reflexive/transitive closure.

Let A_t, B_t be the observed positive and negative sets. Define the retained version space

$$\mathcal{J}_t^E = \{I \subseteq V : I \text{ is an ideal of } P_t,\ A_t \subseteq I,\ I \cap B_t = \emptyset\},$$

and

$$\mathcal{J}_t^N = \mathcal{J}_t^E \setminus \{\emptyset\}.$$

An ideal is a downward-closed subset. Every ideal is a prefix of some linear extension of P_t: first topologically sort the ideal, then its complement. Thus these version spaces describe precisely the target vectors compatible with the information deliberately retained by the algorithm.

Define

$$D_1 = \bigcap_{I \in \mathcal{J}_t} I, \qquad D_0 = V \setminus \bigcup_{I \in \mathcal{J}_t} I, \qquad U = V \setminus (D_0 \cup D_1).$$

These sets are computable without knowing the true order. Realizability guarantees $\mathcal{J}_t \neq \emptyset$.

Under E, they have the simpler reachability descriptions

$$D_1 = \{v : \exists w \in A_t,\ v \preceq_{P_t} w\}, \qquad D_0 = \{v : \exists w \in B_t,\ w \preceq_{P_t} v\}.$$

Indeed, the smallest feasible ideal is $\downarrow A_t$; if $v \notin D_0$, then $\downarrow (A_t \cup \{v\})$ avoids B_t. Under N, enumeration of nonempty ideals also incorporates the nonemptiness inference that the reachability tests alone can miss.

Every previously observed point is determined. Consequently, if the current point x and another point u are both undetermined, then u has not arrived.

Query

$$S = \{(x,1), (u,0)\}.$$

- **Answer $\perp$.** This happens exactly when $u \prec x$. If $x \prec u$, the threshold c_x realizes S; the reverse ordering makes S impossible. Predict 0. If this prediction is wrong, then $x = 1$, and $u \prec x$ forces $u = 1$.

- **Answer h.** Put $H = \{v \in V : h(v) = 1\}$. Then $x \in H$, $u \notin H$, and

$$H \prec V \setminus H.$$

Predict 1. If this prediction is wrong, then $x = 0$. Every point outside H follows x, so every such point is 0, including u.

The latter inference uses the fact that H is a threshold prefix, not merely that it contains x. Both branches work for every legal consistent return value. □

Lemma D.3 — Accounting, including early termination

Run the protection algorithm with an internal query cap k: query at every undetermined arrival whenever budget remains and another undetermined point exists. Let $q \le k$ be the number of calls actually made. Then

$$\boxed{M \le T - q.}$$

If $q < k$, the valid additional conclusion is

$$\boxed{M \le q + 1.}$$

Therefore

$$\boxed{M \le \max\{T - k, k\}.} \tag{D.1}$$

Proof. Let D be the set of arrivals that were undetermined immediately before their prediction/query processing. Let b be the number of mistaken query rounds.

Each mistaken query protects a previously undetermined future point. Once protected, that point remains determined. Hence the protected points from the b mistaken queries are distinct and do not belong to D. Thus

$$|D| + b \le T.$$

At most $|D| - q$ mistakes occur on nonquery rounds, giving

$$M \le b + |D| - q \le T - q.$$

If $q < k$, budget was never exhausted. Any undetermined arrival at which no query was made must therefore have been the sole remaining undetermined point. After its label is observed, no future undetermined arrival is possible. Thus at most one member of D is a nonquery round:

$$|D| \le q + 1, \qquad M \le q + 1 \le k.$$

Combining this with the $q = k$ case proves (D.1). □

In particular, choosing

$$k = \min\{Q, \lfloor T/2 \rfloor\}$$

gives the valid general upper bound

$$M \le \max\{T - Q, \lceil T/2 \rceil\}. \tag{D.2}$$

For Theorem D.1(iii):

- E, $T = 5, Q = 3$: take $k = 2$, obtaining $M \le 3$;
- N, $T = 6, Q = 4$: take $k = 3$, obtaining $M \le 3$.

The remaining work is to prove matching lower bounds for **every** deterministic learner.

3. An analytic five-point lower bound under E

Throughout this section there are five instances, an E sentinel, and no observed labels unless explicitly stated. A poset state means that every linear extension of that poset remains available to the adversary. Providing the poset to the learner for free only strengthens the learner.

Lemma D.4 — Antichain cube

Suppose no queries remain and the available orders include all linear extensions of a poset P containing an antichain W of size r. Then the adversary can force r mistakes in the prescribed transductive sequence.

Proof. Let

$$D = \{v \notin W : v <_P w \text{ for some } w \in W\}.$$

For every $J \subseteq W$, the set $D \cup J$ is an ideal. Label D positively and $V \setminus (D \cup W)$ negatively. The labels on W are then independently selectable: every choice is a prefix in some linear extension of P.

At each arrival in W, choose the opposite of the deterministic prediction. Finally choose a linear extension in which the resulting positive ideal comes first, and take its corresponding threshold. □

Lemma D.5 — Extreme-point padding

Under E, suppose one of five points, p, can be fixed as either

- the least instance, with target label 1; or
- the greatest instance, with target label 0,

while the other four instances have arbitrary relative order and arbitrary E-threshold target. With at most one further query, at least three mistakes can be forced.

Proof. This embeds the four-point E problem, to which Corollary 4.1 gives

$$M \geq 4 - 1 = 3.$$

It remains to check that queries involving the padding point do not invalidate the embedding.

If p is least:

- a query requiring $p = 0$ and some ordinary instance to be 1 is impossible;
- if it requires $p = 0$ but no ordinary positive, return the empty-instance prefix;
- otherwise, make at most one base ERM call and return its reply, extending the reply by $p = 1$ only when it is a concept vector (a $\perp$ reply is returned unchanged).

If p is greatest:

- a query requiring $p = 1$ and some ordinary instance to be 0 is impossible;
- if it requires $p = 1$ but no ordinary zero, return the all-one concept;
- otherwise, make at most one base call and return its reply, extending the reply by $p = 0$ only when it is a concept vector (a $\perp$ reply is returned unchanged).

Conflicting labels and a negative sentinel are handled directly by $\perp$. Previously returned vectors that are valid for every order in the embedded subclass can be cached and returned without a base call.

These are legal, deterministic, memoryless simulations on the full domains. □

The same greatest-point simulation also works when the base class uses convention N; this will be used later.

Lemma D.6 — Predicting before resolving a matching costs three mistakes

Suppose the retained relations are

$$a \prec b, \qquad c \prec d,$$

with a fifth point e otherwise unrelated. If the learner predicts the first label now, the adversary can force that prediction to be wrong and then force two further mistakes, even if the full order is subsequently disclosed.

The conclusion also holds with zero or one retained comparison, since such a state can be strengthened to the displayed matching.

Proof. For any current point v and either desired label y, construct three future points H such that:

1. all four prefix labelings on H are possible while $v = y$;
2. the earliest-arriving point of H is its median.

The constructions below suffice; exchanging the two pairs handles $v = c, d$.

Current point	Desired label	Choose H	Order of the groups
a	0	$\{c, d, e\}$	$H < a < b$
a	1	$\{c, d, e\}$	$a < H < b$
b	0	$\{c, d, e\}$	$a < H < b$
b	1	$\{c, d, e\}$	$a < b < H$
e	0	$\{b, c, d\}$	$a < H < e$
e	1	$\{b, c, d\}$	$a < e < H$

Each triple has just one required comparison, $c < d$. Its earliest point can always be made the median. For the triple $\{c, d, w\}$, use respectively

$$w < c < d, \qquad c < d < w, \qquad c < w < d$$

when the earliest point is c, d, w.

Choose y opposite to the current prediction. On the triple, make the median prediction wrong. Its label leaves two possible thresholds differing on one later extreme; make that prediction wrong too. The fourth future point has a fixed label common to all four triple thresholds.

An ERM oracle reveals the class, not the target. Even full knowledge of this order cannot distinguish these remaining target thresholds before their labels arrive. □

Lemma D.7 — The second-query alternatives

Suppose the only retained comparison is $a < b$. Any query can be answered so that the remaining orders contain either:

1. a subclass covered by Lemma D.5; or
2. all extensions of two disjoint comparisons.

Proof. Remove the sentinel from the query after checking its label, and handle conflicts, empty queries, and single-sided queries directly. Such a query can receive a no-information answer, after which fixing a as the least instance gives case 1.

For a genuine mixed query, let $A \neq \emptyset$ be its positive set and $B \neq \emptyset$ its negative set.

If there are $p \in A, n \in B$ such that adding $n < p$ is consistent with $a < b$ and creates no three-element chain, answer $\perp$ and retain this relation. The resulting poset is one of:

- one comparison;
- a two-edge fork;
- two disjoint comparisons.

For a fork, put its common source globally first, or its common sink globally last. The other four points are unrestricted, giving case 1. Disjoint comparisons give case 2. If the added relation is already $a < b$, fix a as the least instance and set its target label to 1; the other four instances retain arbitrary relative order and arbitrary E-prefix targets, which satisfies the pre-existing comparison and makes the present $\perp$ universally legal, so this is case 1 as well.

Otherwise, $a \notin B$ and $b \notin A$: either membership would provide a permitted reversed pair. Moreover, if both $A \setminus \{a\}$ and $B \setminus \{b\}$ were nonempty, a pair between these sets would be disjoint from $a < b$, again permitted. Hence

$$A = \{a\} \quad \text{or} \quad B = \{b\}.$$

In the first case return the prefix $\{a\}$; in the second return $V \setminus \{b\}$. These are respectively consistent with making a least or b greatest, leaving the other four points unrestricted. □

Lemma D.8 — One query cannot eliminate a three-point antichain from the matching

From the matching state

$$a < b, \qquad c < d,$$

plus isolated e, every query has a legal answer leaving a poset with a three-point antichain.

Proof. Put

$$L = \{a, c\}, \qquad R = \{b, d\}.$$

Trivial queries can receive a no-information answer. For a mixed query with positive set A and negative set B:

- If $B \cap L \neq \emptyset$, choose $\ell \in B \cap L$ and $p \in A$. Answer $\perp$, retaining $\ell < p$. The set $R \cup \{e\}$ remains an antichain.

- Otherwise, if $A \cap R \neq \emptyset$, choose $r \in A \cap R$ and $n \in B$. Answer $\perp$, retaining $n < r$. The set $L \cup \{e\}$ remains an antichain.
- Otherwise,

$$A \subseteq L \cup \{e\}, \qquad B \subseteq R \cup \{e\}.$$

Return the prefix

$$H = \begin{cases} L, & e \in B, \\ L \cup \{e\}, & e \notin B. \end{cases}$$

This respects the matching and the query. Its three-point side is an antichain.

In each $\perp$ case, a requested negative has been placed before a requested positive, making the query genuinely unrealizable. □

Proposition D.9 — $M_E^*(5,3) \geq 3$

Proof. Fix any deterministic learner making at most three calls.

1. **Before the first call.** If it predicts, Lemma D.6 forces three mistakes. Otherwise, answer its first mixed query by $\perp$, retaining one reversed requested pair $a < b$. A trivial query can receive a no-information answer, and one comparison may then be supplied for free.
2. **After the first call.** If it predicts, Lemma D.6 again applies. Otherwise apply Lemma D.7 to its second call. An extreme-point/fork outcome invokes Lemma D.5, with at most one call left. The only remaining case is two disjoint comparisons.
3. **In the matching case.** Predicting before the last call loses by Lemma D.6. If the learner makes the last call, Lemma D.8 leaves a three-point antichain. No queries remain, so Lemma D.4 forces three mistakes.

This covers algorithms that use fewer than three calls as well.

Oblivious solidification and memorylessness. Cache every queried set and its answer; repeated queries receive their cached answer. Such repetitions consume budget without imposing additional restrictions.

Every returned nontrivial vector was retained as a prefix cut. Every nontrivial $\perp$ answer has a retained reversed pair witnessing impossibility. At the end, choose a total-order extension and a target prefix realizing the constructed labels. Define

$$O^*(S) = \begin{cases} \text{the cached answer}, & S \text{ was queried}, \\ \text{the least consistent threshold in the final order}, & S \text{ is realizable and unqueried}, \\ \perp, & \text{otherwise}. \end{cases}$$

All cached answers are legal in the final order. In the padding branch, use the fixed oracle from Corollary 4.1 through the explicit simulation of Lemma D.5.

Thus O^* is a legal function of S alone. Determinism makes the replay identical. The resulting order, target, sequence, and oracle are all fixed before that replay, as required by the oblivious-instance requirement.

All choices can be tie-broken by arrival index. Final order extensions can be obtained by lexicographic topological sorting, with the target ideal first. This is a finite constructive solidification, not a compactness argument. □

Lemma D.3 supplies the matching upper bound, proving

$$\boxed{M_E^*(5,3) = 3}.$$

4. Embedding the known-class lower bound in the transductive model

Lemma D.10 — Known-order lower bounds

For every query budget,

$$M_E^*(T, Q) \geq \lfloor \log_2(T + 1) \rfloor, \qquad M_N^*(T, Q) \geq \lfloor \log_2 T \rfloor.$$

Proof under E. Put $d = \lfloor \log_2(T + 1) \rfloor$ and choose $2^d - 1$ ordered instances. Restrict the target to their 2^d prefixes, including the empty prefix. Present these instances in breadth-first order of the perfectly balanced binary search tree; present any extra, always-negative instances afterwards.

The entire sequence is fixed in advance. At each depth, the node on the currently surviving search path splits the possible target cuts equally. Choose its label opposite to the deterministic prediction. Other nodes encountered outside the surviving interval have labels common to all surviving targets. This forces one mistake at each of the d depths.

The full order may be disclosed for free. Fix, for example, the least-consistent-threshold ERM rule. Given the order, all oracle answers are simulatable and contain no information about which target cut was chosen.

The adaptive choice of target cuts is solidified by selecting the final surviving target and replaying the deterministic interaction.

Under N, reserve the least instance as an always-positive point. Use $2^d - 1$ further instances for the same construction, where $d = \lfloor \log_2 T \rfloor$. The reserved minimum supplies the "empty" prefix on the hard instances while the target on the full domain remains nonempty. □

This is a worst-sequence statement. It does **not** assert that every known-order instance permutation has this mistake complexity.

5. The N convention also has a strict counterexample

Lemma D.11 — First-action lower recurrence under N

For $n \geq 2$ and $q \geq 1$,

$$M_N^*(n, q) \geq \min\left\{M_E^*(n - 1, q - 1),\ 1 + M_E^*(n - 1, q),\ 1 + M_N^*(n - 1, q)\right\}. \tag{D.3}$$

Proof.

If the first action is a query, it can always be answered while leaving some point p globally least and all other relative orders unrestricted:

- Mixed positive/negative query: choose p among its negatives and answer $\perp$. A globally least point cannot be negative under N.
- Negative-only query on a proper subset: choose p outside that subset and return the singleton prefix $\{p\}$.
- Negative-only query on the whole domain: answer $\perp$.
- Positive-only or empty query: return the all-one concept.
- Conflicting labels: answer $\perp$.

Disclose p as the minimum for free. Its target label is necessarily 1; the other $n - 1$ points form an E problem with at most $q - 1$ calls remaining. Cached first answers are universally legal in this subclass.

If the first action is prediction 0, give label 1, make that point least, and retain an E problem on the remaining $n - 1$ points.

If the first action is prediction 1, give label 0, make that point greatest, and retain an N problem on the remaining $n - 1$ points. The additional all-one concept introduced by the greatest point is handled by the simulation in Lemma D.5.

These three cases prove (D.3). □

Apply (D.3) with $n = 6, q = 4$. Proposition D.9 and Lemma D.10 give

$$M_E^*(5,3) = 3,$$
$$1 + M_E^*(5,4) \geq 1 + \lfloor \log_2 6 \rfloor = 3,$$
$$1 + M_N^*(5,4) \geq 1 + \lfloor \log_2 5 \rfloor = 3.$$

Therefore $M_N^*(6,4) \geq 3$. Lemma D.3, with internal cap $k = 3$, gives the reverse inequality:

$$\boxed{M_N^*(6,4) = 3}.$$

Consequently, any proposed formula assigning the value 2 to $(T, Q) = (6, 4)$ is false.

Further N conclusions

Equation (D.3), Corollary 4.1 under E, and induction on T yield the strengthened **budget** lower bound (this is a budget statement for $Q \geq 1$, obtained by the first-action recurrence and induction rather than directly from the pathwise bound of Corollary 4.1 under N: for $Q \geq 1$ the three branches of (D.3) give $M_E^*(T-1, Q-1) \geq T - Q$ and $1 + M_E^*(T-1, Q) \geq T - Q$ by Corollary 4.1 under E, and $1 + M_N^*(T-1, Q) \geq T - Q$ by the induction hypothesis, while the base case $T = 1$ has a nonpositive right-hand side)

$$M_N^*(T, Q) \geq T - Q \qquad (Q \geq 1).$$

Together with Lemma D.10,

$$M_N^*(T, Q) \geq \begin{cases} T - 1, & Q = 0, \\ \max\{T - Q, \lfloor \log_2 T \rfloor\}, & Q \geq 1. \end{cases} \tag{D.4}$$

Predicting 1 throughout makes at most $T - 1$ mistakes under N. Consequently,

$$\boxed{M_N^*(T, 0) = M_N^*(T, 1) = T - 1.} \tag{D.5}$$

This is a budget statement; it is **not** a new pathwise assertion $M + q \geq T$ when the actual query count might be zero.

6. General bounds and failure of the candidate exact formula

The results established above imply

$$\max\{T - Q, \lfloor \log_2(T+1) \rfloor\} \leq M_E^*(T, Q) \leq \max\{T - Q, \lceil T/2 \rceil\}, \tag{D.6}$$

and the N lower bound (D.4), together with

$$M_N^*(T, Q) \leq \min\{T - 1,\ \max\{T - Q, \lceil T/2 \rceil\}\}. \tag{D.7}$$

There is also exact large-budget saturation:

$$Q \geq 2(T-1) \quad \Longrightarrow \quad M_E^*(T, Q) = \lfloor \log_2(T+1) \rfloor, \quad M_N^*(T, Q) = \lfloor \log_2 T \rfloor.$$

For completeness, the explicit count follows from the sorting method of AHR25 (Thm. 4.5(3)): query a pair in one orientation, and if necessary in the reverse orientation. A successful return splits the current set into two nonempty ordered parts. There are $T - 1$ internal splits, each costing at most two calls. Once sorted, maintain the set V_t of prefix concepts consistent with the labels seen so far and predict its majority label; every mistake at least halves $|V_t|$, while the true target is never removed, so $1 \leq |V_0| 2^{-M}$ and $M \leq \lfloor \log_2 |V_0| \rfloor$, with $|V_0| = T + 1$ under E and $|V_0| = T$ under N. (The sufficient budget $2(T-1)$ is not claimed to be minimal; for instance $M_E^*(2, 1) = 1$: one pair query fixes the order of two points and halving then makes one mistake, matching the lower bound.)

Equations (D.6)–(D.7) are bounds, not an exact characterization. The strict finite counterexamples show why the candidate equality cannot be assembled from Corollary 4.1 and the known-class Halving bound alone.

Theorem 4.4 of AHR25 is not invoked here; its statement, and why its hypothesis fails for the threshold family, are discussed once in Appendix B. (For the E counterexample, $T = 5$ and $d_{\mathrm{LD}} = 2$, so its hypothesis $T \geq 2^{d_{\mathrm{LD}}+1} = 8$ fails.) Corollary 4.1 is invoked as a proved theorem; the deterministic/randomized separation is unaffected.

E The weak consistency interface

Proof organization. The deterministic upper bound combines comparisons, deterministic selection, and phase accounting. The lower bound comes from AHR25's weak-consistency theorem; Lemma E.5 separately converts its hard query cap to an expected-budget statement. This branch compares interfaces and does not enter either ERM lower-bound proof.

Theorem E.1 (Weak consistency oracle)**.** *With only the weak consistency oracle, there is a deterministic learner that, on every instance, makes at most* $66T$ *calls and* $\lfloor \log_2(T+1) \rfloor$ *mistakes under E, and at most* $67(T-1)$ *calls and* $\lfloor \log_2 T \rfloor$ *mistakes under N. Conversely, for* $T - \varepsilon \geq 100$*: any learner—deterministic or randomized—that makes at most* $(T-\varepsilon)/20$ *calls on every run (a hard cap) incurs at least* $(T-\varepsilon)/20$ *expected mistakes on some instance (AHR25, Thm. 4.3, applied with effective horizon* $T-\varepsilon$*), and any learner with worst-instance expected mistakes* m *and worst-instance expected calls* q *satisfies* $m + 20q \geq (T-\varepsilon)/20$ *(Lemma E.5 in Appendix E). Hence, at this interface,* $\Theta(T)$ *calls are necessary and sufficient for* $O(\log T)$ *mistakes for both deterministic and randomized learners.*

Remark. At the level of asymptotic query order, the randomized weak-consistency result follows from AHR25 (Thm. 4.5(2) and 4.3); the worst-instance expected-budget formulation is supplied by Lemma E.5 below. The deterministic upper bound is a modest improvement over the $O(T \log T)$ of AHR25 (Thm. 4.5(1)): the linear order, which AHR25 reach only with randomization (Thm. 4.5(2), by sampling-based rank estimates), is attained here deterministically (Theorem E.1), and its content is the explicit constants and the exact integer mistake bound; the selection procedure, the constants and the convention correction are below. Lemma E.5 (a truncation lemma for expected budgets) is ours.

Theorem E.1: deterministic linear-query learning with a weak consistency oracle

Interface and conventions. Throughout this section, the learner accesses the class only through the Boolean oracle of **Definition 2.2 of AHR25**. It never receives or evaluates a returned concept. The concept-returning interface of Section 2 is used only in the later comparison table. For a finite labeled sample S we write $\mathrm{WC}(S) \in \{\text{realizable}, \text{not realizable}\}$ for the answer of the weak consistency oracle, namely whether some $c \in C_{\preceq}$ is consistent with S.

Let the $T \geq 1$ instances be distinct and revealed in advance, as stipulated in the transductive protocol. Labels are generated by an arbitrary target

$$c_z(x) = \mathbf{1}[x \preceq z]$$

under an arbitrary unknown total order $\preceq$.

Write $\epsilon = 0$ under convention E and $\epsilon = 1$ under convention N:

- **E:** the domain contains a point preceding all T instances, so the empty prefix on the instances is available.
- **N:** the domain consists exactly of the T instances, so the empty prefix is unavailable.

Theorem E.1 — explicit bounds

There are deterministic algorithms using only weak consistency queries such that, for every admissible domain, total order, target, and transductive sequence,

Convention	Number of queries	Number of mistakes
E	$Q \leq 66T$	$M \leq \lfloor \log_2(T+1) \rfloor$
N	$Q \leq 67(T-1)$	$M \leq \lfloor \log_2 T \rfloor$.

Every oracle query made by these algorithms has exactly the form

$$\{(u,1),(v,0)\}, \qquad u \neq v.$$

Consequently, in either convention,

$$Q \leq 67T, \qquad M \leq \log_2 T + 1.$$

These are pathwise bounds: there is no randomization and no averaging over oracle answers.

LEMMA E.2 — COMPARISON INTERFACE AND ADAPTIVE CONSISTENCY

For distinct $u, v \in X$,

$$\mathrm{WC}(\{(u,1),(v,0)\}) = \text{realizable} \quad \Longleftrightarrow \quad u \prec v.$$

This holds under both E and N.

Proof. Realizability means that some $z \in X$ satisfies

$$u \preceq z \prec v,$$

which implies $u \prec v$. Conversely, if $u \prec v$, the concept c_u realizes the query. No empty-prefix assumption is needed. □

The query concerns the existence of **some** consistent concept, not consistency with the actual target. In particular, observed target labels must **not** be appended to this comparison query.

The comparison implementation is also valid against adaptively supplied comparison answers, provided the answers remain consistent with a total order. Indeed, fix any completed total order consistent with a finite comparison transcript. A deterministic comparison algorithm, run against that fixed order, follows exactly the same transcript. Thus every correctness and comparison-count guarantee proved for arbitrary fixed total orders holds for every consistent adaptive transcript.

For a finite instance set, a consistent partial order has a total-order extension: repeatedly remove a minimal remaining element. This is the only extension fact needed here.

LEMMA E.3 — EXACT SELECTION IN AT MOST $32n$ COMPARISONS

There is a deterministic comparison algorithm

$$\mathrm{SELECT}(A,k), \qquad 1 \le k \le n = |A|,$$

which returns the element of rank k using at most $32n$ comparisons.

Therefore, an exact median of rank

$$k = \left\lceil \frac{n}{2} \right\rceil$$

can be found using at most $32n$ weak consistency queries, and partitioning the other elements relative to it costs at most $n-1$ additional queries.

Algorithm.

- If $n \le 64$, insertion-sort the elements and return the element of rank k.
- Otherwise, let $m = \lfloor n/5 \rfloor$. Form m complete groups of five; leave at most four elements outside these groups.
- Insertion-sort each complete group and collect its median.
- Recursively select the median q of those m medians, using rank $a = \lceil m/2 \rceil$.
- Compare every other original element with q.
- Return q, or recurse into the appropriate side with the appropriately adjusted rank.

All grouping and tie-independent choices use the original arrival-index order, making the procedure deterministic.

Correctness. The partition determines the exact rank of q. The desired element is therefore either q itself or the desired adjusted rank in exactly one of the two sides. Induction proves correctness.

Comparison count. Let $C(n)$ be the worst-case number of comparisons, with $C(0) = 0$. For $n \le 64$, insertion sort uses at most

$$\frac{n(n-1)}{2} \le \frac{63}{2} n \le 32n.$$

Now suppose $n \ge 65$. Sorting the complete groups costs at most $10m$ comparisons, and partitioning around q costs at most $n-1$.

There are at least

$$3(a-1)+2 = 3a-1$$

elements strictly below q. There are at least

$$3(m-a)+2 \geq 3a-1$$

elements strictly above q; the last inequality follows directly for both parities of m.

Thus the side used by the second recursive call has size s satisfying

$$s \leq n-3a.$$

Consequently,

$$m+s \leq n+m-3\left\lceil \frac{m}{2} \right\rceil \leq n-\frac{m}{2} \leq \frac{9n}{10}+\frac{2}{5},$$

where the last step uses $m \geq (n-4)/5$.

Induction now gives

$$\begin{aligned} C(n) &\leq C(m)+C(s)+10m+n-1 \\ &\leq 32(m+s)+3n-1 \\ &\leq 32\left(\frac{9n}{10}+\frac{2}{5}\right)+3n-1 \\ &= \frac{159n+59}{5} \\ &\leq 32n, \end{aligned}$$

because $n \geq 65 > 59$. This proves the stated explicit constant. □

For $n \geq 2$, the selected rank satisfies

$$\frac{3n}{10} \leq \left\lceil \frac{n}{2} \right\rceil \leq \frac{7n}{10}.$$

For $n = 2$ this is immediate; for $n \geq 3$, use $\lceil n/2 \rceil \leq (n+1)/2 \leq 7n/10$. The literal middle-rank interval contains no integer when $n = 1$; a singleton instead requires no comparisons.

Lemma E.4 — a phase contains at most one mistake

The learner maintains a partial map K of **certified labels**. Every certification will be proved correct. Let

$$U = \{x_1, \ldots, x_T\} \setminus \mathrm{dom}(K).$$

Thus every point in U is unobserved and uncertified. Certification may be conservative: U need not contain exactly the logically undetermined points, and it need not be an interval in the unknown order.

At the start of a phase, freeze

$$V = U, \qquad n = |V| \geq 1.$$

Using Lemma E.3, find the element p of rank $k = \lceil n/2 \rceil$ in V, and partition

$$L = \{u \in V : u \prec p\}, \qquad R = \{u \in V : p \prec u\}.$$

Throughout this phase:

- Predict the certified label on points already in $\mathrm{dom}(K)$.
- On an uncertified point in $L \cup \{p\}$, predict 1.
- On an uncertified point in R, predict 0.
- Certify each observed label.
- Following a mistake:
- if the prediction was 1, certify every point of $\{p\} \cup R$ as 0;
- if the prediction was 0, certify every point of $L \cup \{p\}$ as 1.

End the phase, before another prediction, whenever

$$|U| \le \rho(n), \qquad \rho(n) := \left\lfloor \frac{n-1}{2} \right\rfloor.$$

Then every certification is correct, and each phase contains at most one mistake.

Proof. Certified points cannot cause mistakes, assuming the certification invariant.

Consider a mistake on an uncertified point x. The following table lists all possibilities.

Mistake	Valid threshold implication	Active set after certification
$x \in L$, prediction 1, label 0	Every point in $\{p\} \cup R$ is 0	$U' \subseteq L \setminus \{x\}$
$x = p$, prediction 1, label 0	Every point in $\{p\} \cup R$ is 0	$U' \subseteq L$
$x \in R$, prediction 0, label 1	Every point in $L \cup \{p\}$ is 1	$U' \subseteq R \setminus \{x\}$

For example, in the first case,

$$z \prec x \prec p,$$

so p and every point above it are negative. In the third case,

$$p \prec x \preceq z,$$

so p and every point below it are positive. These implications also show that new certifications never conflict with previous correct certifications.

Since $|L| = k - 1$ and $|R| = n - k$, the corresponding size bounds are

$$k-2, \qquad k-1, \qquad n-k-1.$$

Cases involving an empty side simply cannot occur. In every possible case,

$$|U'| \le \max\{k-1, n-k-1\} = \left\lfloor \frac{n-1}{2} \right\rfloor = \rho(n).$$

Therefore, the first mistake immediately ends the phase. There can be no second mistake in that phase.

Importantly, these are containments in the **original phase sets** L, R. They remain valid even if many points have already been removed by correct predictions. No lower bound on the number of *newly* removed points is required. □

Proof of Theorem E.1

Convention E

Initially, K is empty and the active set has size T. Repeatedly execute the phases of Lemma E.4. If U becomes empty, all remaining predictions use certified labels.

Let the positive phase-start sizes be

$$n_1, n_2, \dots, n_J.$$

Then $n_1 = T$, and the phase-ending rule gives

$$n_{i+1} \le \left\lfloor \frac{n_i - 1}{2} \right\rfloor, \qquad \text{hence} \qquad n_{i+1} + 1 \le \frac{n_i + 1}{2}.$$

Since $n_J \ge 1$,

$$2 \le n_J + 1 \le \frac{T+1}{2^{J-1}},$$

which implies

$$J \le \lfloor \log_2(T+1) \rfloor.$$

Lemma E.4 therefore yields

$$M \le J \le \lfloor \log_2(T+1) \rfloor.$$

The query cost of a phase of size n_i is at most

$$32n_i + (n_i - 1) \le 33n_i.$$

Furthermore, $n_{i+1} \le n_i/2$, so

$$\sum_{i=1}^{J} n_i \le 2T.$$

Thus

$$Q \le 33 \sum_i n_i \le 66T.$$

Convention N

Here X consists exactly of the T instances. First find its minimum a by a sequential minimum scan, using exactly $T - 1$ comparisons.

Every target satisfies $a \le z$, so $c_z(a) = 1$. Certify a as positive before online prediction begins.

Run the same phase algorithm on the remaining $T - 1$ active points. The preceding proof, with initial active size $T - 1$, gives

$$M \le \lfloor \log_2 T \rfloor$$

and

$$Q \le (T - 1) + 66(T - 1) = 67(T - 1).$$

For $T = 1$, this means zero queries and zero mistakes under N. Under E, the singleton phase uses zero queries and makes at most one mistake.

Consequently, under either convention, $Q \le 67T$ and $M \le \lfloor \log_2(T + 1) \rfloor \le \log_2 T + 1$; the sharper convention-specific bounds are the ones stated above. □

Convention correction. The E algorithm itself also works unchanged under N, with the same $66T$ query bound and at most $\lfloor \log_2(T + 1) \rfloor$ mistakes. The minimum-scan variant obtains the sharper N bound $\lfloor \log_2 T \rfloor$. Thus the integer mistake correction is exactly the replacement of $T + 1$ by T, a difference of at most one, matching the two Littlestone dimensions of Section 2.

Weak consistency: matching linear lower bounds

The quoted lower bound

The statement of Theorem 4.3 of AHR25 is:

> Consider any family $\mathscr{F}$ of concept classes of the form $C \subset \{0, 1\}^X$, where the family $\mathscr{F}$ has the property that for every labeling function $f : \{x_1, x_2, \dots, x_T\} \to \{0, 1\}$, there exists some concept class $C \in \mathscr{F}$ and some concept $c \in C$ such that $c(x_t) = f(x_t)$ for all $t \in [T]$ (i.e., all 2^T possible binary labelings of the sequence x are captured by the family $\mathscr{F}$). For $T \ge 100$, any (possibly randomized) algorithm that makes at most $T/20$ queries to the weak consistency oracle will incur an expected mistake bound of at least $T/20$. In particular, this theorem holds for general classes like Littlestone classes, and also for special families of classes, including thresholds, k-intervals, and d-Hamming balls (see the next section).

Applicability under E

Use the finite domain

$$X = \{s, x_1, \dots, x_T\},$$

with s preceding all instances. For an arbitrary labeling f, order the points as

$$s \prec \{\text{positive instances, in arrival-index order}\} \prec \{\text{negative instances, in arrival-index order}\}.$$

If there is a positive instance, choose the last positive instance as z; otherwise choose $z = s$.

This realizes every one of the 2^T labelings. Therefore, for $T \ge 100$, a hard cap of $T/20$ weak consistency queries entails a worst-case expected mistake bound of at least $T/20$. Deterministic algorithms are included as a special case.

Applicability under N: the necessary one-point correction

The full T-point N family does **not** realize the all-zero labeling, so the hypothesis of Theorem 4.3 of AHR25 should not be applied to those T points without adjustment.

Instead, designate the first actual instance s as the minimum and restrict to orders with that property. Its label is always 1. On the remaining $T-1$ instances, every binary labeling is realizable by arranging positives before negatives and choosing $z = s$ when all remaining labels are zero.

Given a learner for the full N problem, simulate it on

$$(s, x_2, \dots, x_T),$$

supply the known first label 1, and count its mistakes on the remaining sequence. Queries are unchanged, and the remaining mistakes are no greater than its total mistakes.

Theorem 4.3 of AHR25 therefore applies with effective horizon

$$h = T - 1.$$

For $T \geq 101$, a hard cap of $(T-1)/20$ queries entails at least $(T-1)/20$ expected mistakes in the worst case.

The designated minimum is an actual instance of the full N problem, not an additional unavailable query point.

The oracle is fixed and memoryless

There is no adversarial choice of a Boolean answer once the order is fixed.

The lower-bound invocation is Theorem 4.3 of AHR25 in its oblivious-adversary model. No adaptive oracle construction is being substituted for an oblivious instance.

Lemma E.5 — expected-query budgets also require linear queries

Let

$$h = T - \epsilon \geq 100.$$

Suppose a possibly randomized learner has worst-instance expected mistake bound m and worst-instance expected query bound q. Then

$$\boxed{m + 20q \geq \frac{h}{20}.}$$

In particular,

$$q \geq \frac{h}{400} - \frac{m}{20}.$$

Proof. Work on the effective h-round problem above, including the N simulation when necessary.

Set

$$b = \left\lfloor \frac{h}{20} \right\rfloor.$$

Construct a truncated learner that follows the original learner, using the same private random bits, until it would make query $b+1$. At that point it makes no further queries and uses an arbitrary fixed prediction for all remaining rounds.

The truncated learner has a hard query cap b. On any fixed instance, let A be the event that truncation occurs. Coupling the two runs gives

$$M_{\text{trunc}} \leq M_{\text{original}} + h\mathbf{1}_A.$$

Also,

$$\Pr(A) \leq \frac{\mathbb{E}[Q_{\text{original}}]}{b+1},$$

because A implies $Q_{\text{original}} \geq b+1$.

Theorem 4.3 of AHR25 supplies a fixed instance on which

$$\mathbb{E}[M_{\text{trunc}}] \geq \frac{h}{20}.$$

On that same instance,

$$\frac{h}{20} \leq m + \frac{h}{b+1}q \leq m + 20q.$$

This proves the claim. □

Thus a logarithmic expected mistake guarantee requires

$$q \geq \frac{T-\epsilon}{400} - O(\log T) = \Omega(T).$$

For example, the hard-cap statement alone implies that

$$q \leq \frac{h}{800} \quad \Longrightarrow \quad m \geq \frac{h}{40}$$

under an **expected-query** budget.

If the unqualified query budget of Theorem 4.3 of AHR25 is read directly as the expected quantity Q defined in Section 2, its quoted $T/20$ implication applies directly to that budget as well. The argument above avoids needing that interpretation: the required expected-query $\Omega(T)$ lower bound follows even from the hard-cap reading alone.

Corollary E.6 — tight weak-consistency query order

For either E or N, the query complexity needed to guarantee $O(\log T)$ mistakes is

$$\boxed{\Theta(T)}$$

for both deterministic and randomized learners.

For randomized learners this remains true when query and mistake budgets are worst-instance expectations. The upper bound may be taken from Theorem 4.5(2) of AHR25, or, more strongly regarding query tails, from the deterministic algorithm of Theorem E.1.

Interface comparison and the precise separation statement

The table concerns $O(\log T)$ mistakes. Randomized query bounds use worst-instance expectation, as in the cost accounting.

Oracle interface	Deterministic query complexity	Randomized query complexity
Consistency-type ERM: returns a full concept or $\perp$	$\Theta(T)$ for worst-case legal oracles and for the extremal rules of Theorem 3.1 (not uniform over all pre-declared rules; see Theorem 3.4) — Theorem 3.1 and Corollary 4.1 give $Q \geq T - \epsilon - O(\log T)$; the sorting algorithm of AHR25 gives $O(T)$	$\Theta(\log T)$ — the $O(\log T)$ upper bound is AHR25, Thm. 4.5(4); the matching lower bound is Theorem 3.2, proved for the minimal-prefix rule and hence for the worst case over legal memoryless oracles
Weak consistency: returns only realizable / not realizable	$\Theta(T)$ — Theorem E.1 together with Theorem 4.3 of AHR25	$\Theta(T)$ — Theorem 4.5(2) of AHR25 (or Theorem E.1) together with Theorem 4.3 of AHR25 and Lemma E.5

Interpretive proposition — scope of the interface separation

In the model of Section 2 and for logarithmic mistake guarantees on unknown-order thresholds:

1. Deterministic learners require linear-order query complexity under either interface: at the weak consistency interface unconditionally, and at the ERM interface against worst-case legal oracles and against the pre-declared extremal rules of Theorem 3.1. The ERM statement is not uniform over all legal pre-declared selection rules: the feasible-median rule admits deterministic $O(\log T)$ queries and mistakes (Theorem 3.4).

2. Randomized learners can use $O(\log T)$ expected queries when a successful query returns an evaluable full concept, by AHR25, Thm. 4.5(4).
3. Randomized learners still require $\Omega(T)$ expected queries when only the realizability bit is returned.

Thus the established linear-versus-logarithmic improvement from randomization is available with the **concept-returning ERM interface**, but disappears at the level of query order with the **weak consistency interface**.

This conclusion does not require Theorem 3.2. Theorem 3.2 is needed only to upgrade the randomized ERM upper bound to a matching $\Theta(\log T)$ characterization. The separation is exponential when expressed in the parameter $d = \Theta(\log T)$.

The statement is about these two interfaces, this family, and this mistake regime. It does not assert that randomization cannot improve constants or other weak-oracle tradeoffs.

Theorem 4.4 of AHR25 is not invoked here; its statement, and why its hypothesis fails for the threshold family, are discussed once in Appendix B. The weak-oracle lower bound above comes from Theorem 4.3 of AHR25, and the deterministic ERM lower bound from Theorem 3.1 and Corollary 4.1.

F Related work

F.1 Related work

Mistake bounds and the transductive protocol. Littlestone (1988) studies mistake bounds and the Littlestone dimension; the transductive ("off-line") variant, in which the sequence of instances is announced in advance, was introduced by Ben-David et al. (1997), and Hanneke et al. (2023) show that, for a fixed known binary class, the deterministic realizable transductive minimax mistake bound is $O(1)$, $\Theta(\log T)$, or T. These mistake bounds assume a known class and do not charge oracle calls. For thresholds on a known order of T points the deterministic minimax mistake bound is $\lfloor \log_2(T + 1 - \varepsilon) \rfloor$, the benchmark here; the question is what it costs, in oracle calls, to reach it when the order is unknown and the class is accessible only through an oracle. Raman et al. (2025) characterize ℓ_1 transductive learnability of $[0, 1]$-valued classes by finite fat-shattering dimension at every positive scale, give regret bounds, and separate it from online learnability; their learner has access to the class.

Oracle-efficient online learning. Assos et al. (2023) give realizable online-learning guarantees using consistency and value oracles, bounding both mistakes and consistency-oracle calls (their Theorem 6), and Kozachinskiy & Steifer (2024) learn with a consistent oracle with mistake bounds depending only on the Littlestone dimension. AHR25 investigate explicit mistake–query tradeoffs, made the number of calls itself the resource, treated the weak consistency interface alongside the ERM interface, and proved the upper bounds recalled above; earlier oracle-efficient results are surveyed there. Daskalakis & Golowich (2024) study the weak consistency oracle in the PAC setting; Syrgkanis et al. (2016) obtain oracle-efficient adversarial contextual learning from an optimization oracle, where the interaction between an oracle's tie-breaking and the learner also appears, in a different model. Okoroafor et al. (2026) study a hybrid model with i.i.d. features, convex Lipschitz losses and full-function feedback (the adversary selects a labelling function r_t from a fixed known class before x_t is drawn, and the learner observes r_t after the round); their algorithm runs with a linear optimization oracle for the hypothesis class, makes $O(T^2)$ oracle calls over T rounds, and has regret governed by a Rademacher complexity. The resource optimized there is regret, not a lower bound on calls.

Regret–oracle tradeoffs and learning orders. In the agnostic, non-transductive setting against adaptive adversaries, Attias et al. (2026) prove query-budget regret lower bounds for an agnostic ERM oracle (their Theorem 5.1); for Littlestone dimension $d = 1$ these imply $\Omega(T/Q)$ when $1 \le Q = o(\sqrt{T})$ (the general form is discussed in Appendix F). Alon et al. (2026) identify an unknown order on n items from counterexamples, at most k of which may be untruthful, using $\Theta(n \log n + nk)$ queries; their goal is the order itself, ours is to predict threshold labels through an oracle for the class. Appendix F discusses these works, and the query models of Angluin (1988), in more detail.

F.2 Further model and theorem comparisons

Mistake bounds and the transductive protocol. Littlestone (1988) studies mistake bounds and the Littlestone dimension; the transductive ("off-line") variant, in which the sequence of instances is announced in advance, was introduced by Ben-David et al. (1997), and Hanneke et al. (2023) show that, for a fixed known binary class, the deterministic realizable transductive minimax mistake bound is $O(1)$, $\Theta(\log T)$, or T. These mistake bounds assume a known class and do not charge oracle calls. For thresholds on a known order of T points the deterministic minimax mistake bound (worst case over arrival orders) is $\lfloor \log_2(T + 1 - \varepsilon) \rfloor$, and this remains the information-theoretic benchmark here; the question studied in this paper is what it costs, in oracle calls, to reach it when the order is unknown and the class is accessible only through an oracle. Raman et al. (2025) study transductive online regression: under the ℓ_1 loss they characterize transductive online learnability of $[0, 1]$-valued classes by finiteness of the fat-shattering dimension at every positive scale, give associated regret bounds, and establish a separation from non-transductive online regression for classes with finite fat-shattering but infinite sequential fat-shattering dimension; their learner has access to the function class and no oracle is involved, so the separation is between protocols, not between deterministic and randomized oracle-based learners.

Oracle-efficient online learning. Assos et al. (2023) give realizable online-learning guarantees using consistency and value oracles, bounding both the number of mistakes and the number of consistency-oracle calls (their Theorem 6), and Kozachinskiy & Steifer (2024) learn with a consistent oracle that returns some concept agreeing with the examples seen so far, with mistake bounds depending only on the Littlestone dimension. AHR25 investigate explicit mistake–query tradeoffs, made the number of calls itself the resource, treated the weak consistency interface alongside the ERM interface, and proved the upper bounds recalled in Section 1; earlier oracle-efficient results are surveyed there. Daskalakis & Golowich (2024) show oracle-efficient realizable binary PAC learning using a realizability-bit oracle; their data are sampled i.i.d., and the oracle accepts submitted labeled samples, so their setting is distributional rather than online. Syrgkanis et al. (2016) obtain oracle-efficient adversarial contextual learning from an optimization oracle; the interaction between an oracle's tie-breaking and the learner appears there too, in a different model. Okoroafor et al. (2026) consider hybrid online learning with full-function feedback: features are i.i.d. from an unknown distribution, losses are convex and Lipschitz, the adversary selects a labelling function r_t from a fixed known class before x_t is drawn, and the learner observes r_t after the round; their algorithm is efficient given a linear optimization oracle for the learner's hypothesis class (used in a Frank–Wolfe reduction with a truncated entropy regularizer), makes $O(T^2)$ oracle calls over T rounds, and achieves regret governed by the Rademacher complexity of a composite class. The oracle there answers optimization questions about the learner's own class, the resource optimized is regret, and no lower bound on the number of calls is studied.

Regret–oracle tradeoffs and learning orders. In the agnostic, non-transductive setting against adaptive adversaries, Attias et al. (2026) prove query-budget regret lower bounds for learners with an agnostic ERM oracle (their Theorem 5.1): for every $d \geq 1$ there is a family of classes of Littlestone dimension d on which every randomized learner making at most Q oracle calls has expected regret at least $(T/2)\min\{1, d/(Q+1)\} - Q/2$; this implies $\Omega(dT/Q)$ when $Q \geq 1, d \leq Q+1$ and $Q = o(\sqrt{dT})$, and at $Q = 0$ the formula gives $T/2$. The setting, the class and the cost measure differ from ours. Alon et al. (2026) identify an unknown linear order on n items from counterexamples: the learner proposes a complete order and receives either confirmation or a wrongly ordered pair, at most k of the counterexamples may be untruthful with k not known in advance, and the query complexity is $\Theta(n \log n + nk)$. Their goal is the order itself; ours is to predict threshold labels through an ERM or realizability oracle for the class, and the cost is mistakes plus calls. In the classical query models of Angluin (1988) the oracle answers questions about the target (membership, equivalence); the oracles here answer questions about the class (consistency), and the target is revealed only through the online labels. We do not claim any equivalence between these models; they are cited to place the selection behaviour of an oracle, as a design parameter of the feedback, in a broader landscape.