RANDOMIZATION IS NECESSARY FOR LEARNING FROM SYMMETRIC EQUIVALENCE QUERIES

Anonymous authors
Paper under double-blind review

ABSTRACT

We resolve the dimension-only derandomization question for general symmetric equivalence queries posed in *Learning from Equivalence Queries, Revisited* at COLT 2026. Prior work gave a randomized proper learner with expected query complexity $O(d)$ for every finite class of Littlestone dimension d. We show that deterministic proper learners cannot match this dimension-only guarantee, even with a fixed target. Under target reselection, symmetric two-point generators can force deterministic complexity linear in the number of concepts even at Littlestone dimension one, while randomized learning uses fewer than three expected queries. For independently resampled distributions over total orders, the extremal deterministic complexity on the singleton class is $\Theta(\sqrt{n})$. A uniform mixture of cyclic rotations has an exact value asymptotic to $\sqrt{\pi n/2}$. A row lower-envelope inequality for the linear-order polytope yields the matching universal $O(\sqrt{n})$ upper bound. For one fixed, known total order, a lexicographic projection of the standard optimal algorithm gives a version-consistent proper learner using at most $\mathrm{Ldim}(\mathcal{H}) + 1$ queries. Complete classes attain this bound. With a fixed target, arbitrary symmetric generators have a tight worst-case $\Theta(\log n)$ hierarchy on singleton classes. Under the same fixed-target protocol, every finite class admits a generator-dependent $2\log_2|\mathcal{H}| + O(1)$ deterministic upper bound. Randomized learning remains $O(\mathrm{Ldim})$ throughout. Dimension alone therefore cannot characterize deterministic EQ learning: the target protocol and feedback structure determine the cost of determinism.

1 INTRODUCTION

Program synthesis (Solar-Lezama et al., 2006) and language-model self-debugging (Chen et al., 2024) motivate learning from corrective feedback. In Angluin's equivalence-query (EQ) model (Angluin, 1988), the learner proposes a hypothesis, which is either accepted or met with a counterexample. How counterexamples are chosen determines what the learner gains; worst-case choices give a harshly pessimistic model. In their COLT 2026 paper *Learning from Equivalence Queries, Revisited*, Braverman et al. (2026) introduced *symmetric* counterexample generators, whose disagreement-point distribution is invariant under exchanging query and target. The proof of their full-information Theorem 4 yields an explicit expected query bound of $2d + 1$ for a proper *randomized* learner on every finite class of Littlestone dimension d. Can dimension also control deterministic query costs? Their Section 5 asks:

> "In particular, is there always a deterministic learning rule for learning against arbitrary symmetric adversaries? What about the restricted but natural class of order-induced adversaries?"

We resolve the general symmetric case of this COLT 2026 derandomization question. Dimension-only guarantees are impossible for deterministic proper learners even with a fixed target (Theorem 4.3). Under target reselection, the negative answer extends to independently resampled order mixtures (Theorem 3.1).

The essential distinction. A *fixed single order* returns the first disagreement in the same known total order on every round. An *order mixture* draws a total order independently from a known

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.

distribution on each round. Both are order-induced, yet their deterministic behavior differs sharply. A fixed known order admits the optimal dimension-only bound $d + 1$. Under target reselection, an independently resampled order mixture can force $\Theta(\sqrt{n})$ queries when $d = 1$. Drawing one hidden order and keeping it throughout defines a third protocol, outside the scope of the order-mixture theorem. For general order mixtures with a fixed target, the dimension-only question remains open (O-F).

1.1 A PHASE DIAGRAM AND OUR CONTRIBUTIONS

Let $C_n = \{h_i : i \in \mathbb{Z}_n\}$ with $h_i(x) = \mathbf{1}\{x = i\}$ be the n-concept singleton class. It has $\mathrm{Ldim}(C_n) = 1$ for $n \geq 2$. Table 1 records the deterministic regimes. All constructive upper bounds assume that the generator, and any underlying order distribution, is known to the learner.

Table 1: Deterministic phase diagram. For every symmetric cell, the randomized side is $O(\mathrm{Ldim})$ by Theorem 4 of Braverman et al. (2026); in the dimension-one examples it is at most 3. The $\Theta(\sqrt{n})$ upper bound for order mixtures allows class-internal but version-inconsistent probes.

Target protocol	Fixed known single order	Independently resampled order mixture	General symmetric generator
Reselected from the current version space	Worst-case envelope $d + 1$, exactly; version-consistent proper learner	On C_n, the supremum is $\Theta(\sqrt{n})$; the cyclic mixture is $\sim \sqrt{\pi n/2}$	$\Theta(n)$ is attainable already on C_n; a threshold kernel gives $(n + 3)/2$
Fixed once	At most $d + 1$ by the stronger reselection theorem	At most $\lceil \log_2 n \rceil + 3$ on C_n; the cyclic mixture is uniformly bounded; tightness open (O-F)	The supremum on C_n is $(1+o(1)) \log_2 n$; every finite class has a $2 \log_2 \lvert\mathcal{H}\rvert + O(1)$ upper bound

Our three main contributions give this negative answer and establish deterministic bounds across the protocols in Table 1.

Randomization is necessary under target reselection (Section 3). A monotone symmetric two-point kernel forces a linear deterministic cost even at Littlestone dimension one. The cyclic order mixture gives an exact $\Theta(\sqrt{n})$ separation. More generally, the supremum over *all* independently resampled order distributions is $\Theta(\sqrt{n})$. The upper bound combines a row lower-envelope inequality for the linear-order polytope with a blocking-point learner. Its repeated probe is proper but version-inconsistent.

A tight logarithmic theory for fixed targets (Section 4). On singleton classes, every strategy is represented by its all-negative main spine. This yields an exact survival-area characterization and a constant-factor characterization by a half-decay domination number. Symmetry gives a deterministic likelihood-halving strategy using $\lceil \log_2 n \rceil + 3$ queries. For $0/1$ matrices, the problem reduces to classical tournament domination up to one query, so the bound is asymptotically tight. A separate potential argument gives $\lceil 2 \log_2 |\mathcal{H}| \rceil + 2$ for general finite classes and known conditionally symmetric generators.

A single known order admits optimal proper learning (Section 5). The classical improper $d + 1$ conversion follows from the Littlestone optimal algorithm (Angluin, 1988; Littlestone, 1988). We use a lexicographic projection lemma for proper learning. Projecting an arbitrary prediction function onto the current version space converts every first-disagreement counterexample into a counterexample to the unprojected function. The learner is version-consistent and works even under target reselection. It extends to finite multiclass labels and is optimal on complete classes. We also give a linear gap family showing that reselected first-disagreement EQ is strictly harder than transductive online prediction on the same order.

Finally, we examine the version-consistent-only order-mixture protocol (Section 6). Its exact value is a minimum over deletion permutations. Thresholding the low-probability arcs gives an exact feedback-vertex-set identity. Current directed-FVS approximation bounds then yield

$$\Omega(\sqrt{n}) \;\leq\; \sup_{\mu} D^{\mathrm{res}}_{\mathrm{cons}}(C_n, \mu) \;\leq\; O\Big(\sqrt{n \log n \log\log n}\Big)\,.$$

Thus the extremal rate on singleton classes under target reselection is resolved up to constant factors for unrestricted proper order-mixture queries. The consistent-only branch remains open up to polylogarithmic factors.

The main text presents the principal results and selected proof arguments. The appendices provide detailed proofs, exact constants, kernel calculations, and further characterizations.

Reading guide. Table 2 lists results, proof locations, and selected direct uses.

2 Model and quantifier conventions

2.1 Equivalence queries and version spaces

Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ be finite. A proper equivalence query is a hypothesis $h_t \in \mathcal{H}$. It is *version-consistent* if $h_t \in V_t$, where V_t is the set of hypotheses consistent with all previously returned labeled counterexamples. Throughout the paper, "proper" by itself means class-internal and does not imply version-consistency.

If the current target is c_t and $h_t = c_t$, the query is accepted. Otherwise a counterexample generator returns a point $x \in \mathcal{X}$ satisfying $h_t(x) \neq c_t(x)$, together with the target label $c_t(x)$. The version space is updated by $V_{t+1} = \{g \in V_t : g(x) = c_t(x)\}$. The query count Q includes the final accepted query. A nonterminating run has $Q = +\infty$.

Remark 2.1 (Label-only version-space convention). Only the returned labeled point enters the version space. The fact that a first-disagreement witness agrees with the query before the returned point is *not* recorded. The learner may remember the whole transcript. The convention only specifies which constraints later reselected targets must respect: returned labels bind, but implicitly revealed prefix labels do not. This label-only convention matches Appendix C of Braverman et al. (2026). Under the alternative full-transcript convention some results change: the separation family of Theorem 5.6 is specific to the label-only convention, while the $d + 1$ optimum of Theorem 5.2 is robust under both (Appendix A).

A generator is *symmetric* if, conditional on every public history, the distribution of the disagreement point is invariant under swapping query and target: $G(\cdot \mid h, c, \text{history}) = G(\cdot \mid c, h, \text{history})$. We use this history-expanded form. Definition 1 of Braverman et al. (2026) conditions on the preceding query sequence. All constructive deterministic upper bounds in this paper assume that the complete generator is known. This is particularly important for the fixed-target upper bound for general classes.

2.2 Reselected and fixed targets

We use two target protocols. In the *reselection protocol*, on every round the environment chooses a target $c_t \in V_t$. The choice is made after the public history is known and before the learner's private randomization for that round. For a deterministic learner, the next query is already determined by the public history. In the *fixed-target protocol*, one $c \in \mathcal{H}$ is chosen once and retained for the entire run.

Unless explicitly stated otherwise, order and kernel randomness is independent across rounds. For a deterministic learner, expectations are only over those protocol or generator draws and over any explicitly randomized legal environment. The target policy is otherwise optimized adversarially.

Write $D^{\mathrm{res}}(\mathcal{H}, G)$ for the infimum, over deterministic proper learners, of the worst-case expected query count under target reselection. Queries in this definition may be inconsistent. The corresponding fixed-target value is $D^{\mathrm{fix}}(\mathcal{H}, G)$. A subscript "cons" restricts every query to the current version space.

2.3 Two-point matrices and order mixtures

For the singleton class C_n, a query h_j against target h_i has only two possible counterexample points, i and j. Let

$$q_{ji} = \Pr[\text{the positive target point } i \text{ is returned} \mid h_j, h_i], \qquad i \neq j.$$

Symmetry is equivalent to $q_{ji} + q_{ij} = 1$. For notational convenience set $q_{ii} = 1$.

If a total order π is sampled from a distribution μ and the first disagreement is returned, then $q_{ji} = \Pr_{\pi\sim\mu}[i \prec_\pi j]$. Thus $Q = (q_{ji})$ belongs to the linear-order polytope LOP_n, the convex hull of pairwise-precedence matrices of total orders. In an order mixture a fresh $\pi \sim \mu$ is sampled on every round.

Our deterministic lower bounds allow the larger action space $h_t \in \mathcal{H}$, so they also apply to the narrower protocol of Appendix C of Braverman et al. (2026), which restricts the learner to $h_t \in V_t$. Some matching upper bounds use version-inconsistent probes. Section 6 treats the consistent-only restriction. The full protocol comparison, including early acceptance, is in Appendix A.

2.4 Notation

Let $\mathrm{GV}(\mathcal{H}, \pi) := D^{\mathrm{res}}(\mathcal{H}, G_\pi)$, where G_π is the fixed-order first-disagreement generator. For a history-invariant two-point matrix Q, let $\mathrm{Esc}(Q) := D^{\mathrm{res}}(C_n, G_Q)$, where G_Q is the induced history-invariant two-point generator. The notation $E_{\mathrm{cons}}(V; Q)$ denotes the consistent-only state value with V as the current version space, and $D^{\mathrm{res}}_{\mathrm{cons}}(C_n, \mu) = E_{\mathrm{cons}}([n]; Q_\mu)$. A matrix in place of a generator always abbreviates the induced generator.

3 Randomization is necessary

For the cyclic generator, identify the domain with $\mathbb{Z}_n$ and write $d(j, i) = (i - j) \bmod n \in \{1, \ldots, n-1\}$. The generator chooses a uniformly random cyclic starting point and returns the first disagreement encountered. Equivalently, $q_{ji} = d(j, i)/n$. This is a symmetric order mixture: it is the uniform distribution over the n cyclic rotations of one circular order.

Theorem 3.1 (Randomization is necessary)**.** *Assume that the generator is known and that the target may be reselected from the current version space on every round.*

1. *For the cyclic generator on C_n, every deterministic proper learner, even one allowed to query outside the version space, has worst-case expected query count at least*

$$D_n = 1 + \sum_{k=0}^{n-2} \prod_{r=1}^{k} \left(1 - \frac{r}{n}\right) \sim \sqrt{\frac{\pi n}{2}},$$

   and a deterministic version-consistent learner attains this value. Thus $D^{\mathrm{res}}(C_n, G_{\mathrm{cyc}}) = D_n$.

2. *Over all independently resampled order distributions,*

$$M_n := \sup_\mu D^{\mathrm{res}}(C_n, \mu) = \Theta(\sqrt{n}).$$

   The supremum is attained. The upper bound permits proper class-internal queries that need not remain version-consistent.

3. *There are known symmetric two-point generators for which $D^{\mathrm{res}}(C_n, G) = (n+3)/2$ for all $n \geq 3$. Consequently symmetric generators can force linear deterministic complexity while $\mathrm{Ldim}(C_n) = 1$.*

4. *For the cyclic generator, and more generally for every monotone translation-invariant symmetric kernel considered in Appendix C, the randomized proper strategy that queries a uniformly random h_J after the environment selects the current-round target c_t has $\mathbb{E}Q = (3n-1)/(n+1) < 3$ against every legal adaptive reselection environment. If every query must be version-consistent, Theorem 4 of Braverman et al. (2026) gives $\mathbb{E}Q \leq 2\,\mathrm{Ldim}(C_n) + 1 = 3$.*

The first two parts answer the COLT 2026 order-mixture question negatively under target reselection. The third shows that symmetry alone allows an even larger separation. The linear example is not claimed to be an order mixture. The four parts are proved by Lemma 3.2, Proposition 3.4, Theorem C.2, and Appendix B, respectively.

3.1 THE EXACT CYCLIC RECURRENCE

Set $S_{n,0} = 1$ and $S_{n,k} = \prod_{r=1}^{k}(1 - r/n) = (n-1)!/\big((n-1-k)!\, n^k\big)$.

Lemma 3.2 (Dead zones and successors)**.** *In the known cyclic generator with target reselection, and with query count including the final accepted query,*

$$D^{\mathrm{res}}(C_n, G_{\mathrm{cyc}}) = 1 + \sum_{k=0}^{n-2} S_{n,k}.$$

The lower bound holds against every deterministic proper learner, including learners that query eliminated concepts.

Proof idea. For the upper bound, a dead-zone learner queries $h_0, h_{-1}, h_{-2}, \ldots$. After $t-1$ negative answers, the dead points form a contiguous arc. Every legal target has forward distance at least t, so the next negative response has conditional probability at most $1 - t/n$. For the lower bound, a successor-dodging environment always selects the closest surviving target ahead of the predicted query. Gap counting shows that the positive-counterexample probability at round r is at most r/n against *every* deterministic learner. The two conditional products coincide term by term. The full proof and the second-order asymptotics $D_n = \sqrt{\pi n/2} + 2/3 + \frac{1}{12}\sqrt{\pi/(2n)} + O(1/n)$ are in Appendices B and C. This is a Stackelberg value: the environment is tailored to the learner, not an environment-independent saddle point.

3.2 THE ORDER-MIXTURE UPPER BOUND

For a row j of $Q \in \mathrm{LOP}_n$, let $\alpha_{j,1} \le \cdots \le \alpha_{j,n-1}$ be its off-diagonal entries in nondecreasing order, and let $L_{j,r} = \sum_{\ell=1}^{r} \alpha_{j,\ell}$.

Lemma 3.3 (Row lower envelope)**.** *For every $Q \in \mathrm{LOP}_n$ and $1 \le r \le n-1$,*

$$\sum_{j=1}^{n} L_{j,r} \ge \frac{r(r+1)}{2}.$$

Consequently, for every r there is a row j such that $\alpha_{j,r} \ge (r+1)/(2n)$.

Proof idea. The r-smallest-entry sum of a row is a pointwise minimum of linear functions and is therefore concave. By Jensen's inequality, it suffices to evaluate this sum at a total-order vertex. The row in position p contributes $\max\{0, r-(n-p)\}$, and summing gives $r(r+1)/2$. See Appendix B for the proof and novelty discussion, and Remark D.2 for nonredundancy.

Proposition 3.4 (Blocking-point learner)**.** *In the known independently resampled order-mixture protocol with target reselection, for $n \ge 2$ (for $n = 1$ the value is 1), a deterministic proper learner, allowed to query outside the current version space, satisfies*

$$D^{\mathrm{res}}(C_n, Q) \le T(Q) := \min_{j,r}\left(r + \frac{1}{\alpha_{j,r}}\right) \le \min_{1 \le r \le n-1}\left(r + \frac{2n}{r+1}\right) < 2\sqrt{2n}.$$

Proof idea. Query the $r-1$ weakest predecessors of a blocking row j. Every surviving target then has positive-counterexample probability at least $\alpha_{j,r}$ on every round. Repeatedly probing h_j is proper, though possibly version-inconsistent, and succeeds within an expected $1/\alpha_{j,r}$ rounds. See Appendix B. A bootstrap of the same inequality sharpens the asymptotic constant: with $M_n = \sup_\mu D^{\mathrm{res}}(C_n, \mu)$, $\sqrt{\pi/2} \le \liminf_n M_n/\sqrt{n} \le \limsup_n M_n/\sqrt{n} \le C_0 = 2.216035867\ldots$, where C_0 is characterized by the root of $y - 1 = 2\log y$. Appendix D gives details, the small extremal values and the order-atom bound.

4 FIXED TARGETS: A TIGHT LOGARITHMIC THEORY

We now fix one target for the entire run. For singleton classes and a history-invariant two-point matrix Q, all nonterminating feedback before a positive counterexample is the negative label at the current query point. Without loss of optimality, normalize the learner to query the identified target

immediately after a positive counterexample. A deterministic learner is then completely described by its infinite all-negative *main spine* $\sigma = (j_1, j_2, \ldots)$. Set

$$a_{ji} = 1 - q_{ji} \quad (j \neq i), \qquad a_{ii} = 0, \qquad R_i^\sigma(t) = \prod_{s=1}^{t} a_{j_s i}, \qquad R_i^\sigma(0) = 1.$$

Proposition 4.1 (Main-spine and half-decay characterizations)**.** *In the known fixed-target protocol on C_n, for a history-invariant two-point matrix Q and deterministic proper queries, possibly outside the current version space,*

$$V_i(\sigma) = 1 + \sum_{t \geq 1} \mathbf{1}\{j_t \neq i\} R_i^\sigma(t-1)$$

is the expected query count against target h_i, and $D^{\mathrm{fix}}(Q) = \inf_\sigma \max_i V_i(\sigma)$. With $A_i(\sigma) = \sum_{t \geq 0} R_i^\sigma(t)$ one has $A_i \leq V_i \leq A_i + 1$. For $0 < \eta < 1$, define $\kappa_\eta(Q) = \min\{k : \exists j_1, \ldots, j_k \ \max_i \prod_{s=1}^k (1 - q_{j_s i}) \leq \eta\}$. Then $\eta \kappa_\eta(Q) \leq D^{\mathrm{fix}}(Q) \leq \kappa_\eta(Q)/(1-\eta) + 1$. In particular, $\frac{1}{2}\kappa_{1/2}(Q) \leq D^{\mathrm{fix}}(Q) \leq 2\kappa_{1/2}(Q) + 1$. The half-decay number has an exact integer-program representation (Appendix F). The integer parameter is essential. Its naive fractional relaxation can have constant value for symmetric matrices whose deterministic complexity is logarithmic.

Theorem 4.2 (Universal logarithmic bound for symmetric two-point generators)**.** *Let Q be a known symmetric two-point matrix on C_n, $q_{ji} + q_{ij} = 1$. In the fixed-target protocol, a deterministic proper learner satisfies $D^{\mathrm{fix}}(Q) \leq \lceil \log_2 n \rceil + 3$.*

Proof idea. Along the all-negative spine, maintain unnormalized likelihood weights r_i. By symmetry, the r-proportional average of the post-negative-response total weight is at most half the current total. Thus some deterministic query halves it. Summing the resulting survival tails $\min\{1, n2^{-t}\}$ and including the final query gives the bound (Appendix F).

Theorem 4.3 (Tournament reduction and tightness)**.** *When $q_{ji} \in \{0, 1\}$, symmetry makes Q a tournament: orient $j \to i$ exactly when $q_{ji} = 1$, and let $\delta(T)$ be the least size of a set reaching every outside vertex along one arc. For a known tournament generator on C_n, with one fixed target and deterministic proper queries, $\delta(T) \leq D^{\mathrm{fix}}(T) \leq \delta(T) + 1$. Moreover $\sup_{Q \text{ symmetric}} D^{\mathrm{fix}}(Q) = (1 + o(1)) \log_2 n$. The bounds $\delta(T) \leq \lfloor \log_2(n+1) \rfloor$ and the existence of tournaments with $\delta(T) > \log_2 n - 3 \log_2 \log_2 n$ are classical (Erdős, 1963; Graham and Spencer, 1971). The contribution here is the reduction embedding tournament domination into a fixed-target symmetric equivalence-query generator. These logarithmic witnesses are nontransitive and are not order mixtures.*

Theorem 4.4 (Known symmetric generators on general finite classes)**.** *Let $\mathcal{H}$ be finite and let one target be fixed throughout. Suppose the complete generator is known and may depend on the public history, but at every history its conditional disagreement-point distribution is symmetric under swapping query and target. Then a deterministic proper learner satisfies $D^{\mathrm{fix}}(\mathcal{H}, G) \leq \lceil 2 \log_2 |\mathcal{H}| \rceil + 2$. The statement holds for discrete counterexample spaces and for continuous spaces under standard measurability assumptions: a standard Borel space with measurable concept evaluations and generator kernel. This case uses Radon–Nikodym densities. No uniform learner for an unknown generator is asserted. Indeed, none exists (Proposition H.1).*

Proof idea. Give each c in the version space V_y of a depth-t history y its history likelihood $\ell_y(c)$, and let $\Phi_t = \sum_{|y|=t} |V_y| \sum_c \ell_y(c)$, so $\Phi_0 = |\mathcal{H}|^2$. Average queries h proportionally to $\ell_y(h)$ and pair (h, c) with (c, h). By conditional symmetry, the paired terms share one counterexample kernel, while their label sections are disjoint in V_y. Hence some deterministic proper query halves Φ. Summing the tails $\Pr_c[N \geq t] \leq \min\{1, |\mathcal{H}|^2 2^{-t}\}$ gives the bound (Appendix F).

Proposition 4.5 (Protocol dichotomy for monotone translation kernels)**.** *Consider known monotone translation-invariant symmetric two-point kernels on C_n. (1) For the cyclic kernel and one fixed target, $D^{\mathrm{fix}} \leq 4$ for every $n \geq 2$. (2) The entire monotone family has a universal constant deterministic upper bound under a fixed target. (3) Among the two protocols "fixed target" and "target reselection", there exists a sequence of kernels with unbounded deterministic complexity if and only if the protocol is target reselection.*

The explicit universal constant $9/2$, the exact small cyclic values $(2, 7/3, 5/2)$, and an audited asymptotic interval $[65/24,\ 3.5595\ldots]$ for the fixed-target cyclic kernel are recorded in Appendix C. The third statement is existential: some reselection kernels, such as the constant $\phi \equiv 1/2$ kernel, still have bounded deterministic value.

5 A SINGLE ORDER ADMITS OPTIMAL PROPER LEARNING

Let $\pi = (x_1, x_2, \ldots)$ be a fixed known enumeration of the domain with order type at most ω. Every two distinct hypotheses in the finite class then have a π-first disagreement point. The generator always returns the π-first disagreement. The target may be reselected on every round.

For $q : \mathcal{X} \to \mathcal{Y}$ and $f \in V$, define the binary mismatch vector

$$M_q(f) = \big(\mathbf{1}\{f(x_i) \neq q(x_i)\}\big)_{i \geq 1}.$$

Let $P_\pi(q, V)$ be any lexicographic minimizer of $M_q(f)$ over $f \in V$.

Lemma 5.1 (Lexicographic projection)**.** *Let $\mathcal{Y}$ be finite, $V \neq \varnothing$, and $h \in \operatorname{lex-argmin}_{f \in V} M_q(f)$. For every $g \in V \setminus \{h\}$, if x is the π-first point at which g and h differ, then*

$$g(x) \neq q(x).$$

For binary labels the minimizer is unique and, more strongly, $h(x) = q(x)$.

Proof. Before x, the two mismatch vectors coincide. If $g(x) = q(x)$ while $h(x) \neq q(x)$, then g has the lexicographically smaller mismatch vector, a contradiction. In the binary case, if $h(x) \neq q(x)$ then the only other label is $q(x)$, which would give the same contradiction. Binary mismatch vectors also determine $f = q \oplus M_q(f)$ uniquely. □

The binary strengthening fails for three labels: on one point, $q = 0$ and $V = \{1, 2\}$ give equal mismatch vectors. The multiclass conclusion needed below is exactly that the target label differs from $q(x)$.

For a finite-label class, define the multiclass Littlestone dimension using binary branching by two distinct labels. For a nonempty version space V, let $s_V(x)$ be a label maximizing $\operatorname{Ldim}(V_{x,y})$ over the labels y with $V_{x,y} \neq \varnothing$ (equivalently, with the convention $\operatorname{Ldim}(\varnothing) = -1$), with fixed tie-breaking.

Theorem 5.2 (Optimal proper learning for a fixed known order)**.** *Let $\mathcal{H} \subseteq \mathcal{Y}^{\mathcal{X}}$ be finite, let $\mathcal{Y}$ be finite, and let $d = \operatorname{Ldim}(\mathcal{H})$. In the known fixed-single-order protocol, even when the target may be reselected from the current version space, there is a deterministic version-consistent proper learner with $\operatorname{GV}(\mathcal{H}, \pi) \leq d + 1$. It queries $P_\pi(s_V, V)$ at version space V.*

For the complete class $\mathcal{Y}^d$ on d coordinates, with $|\mathcal{Y}| \geq 2$ and for every fixed order, every deterministic proper learner, including one allowed to query outside the current version space, can be forced to use at least $d + 1$ queries. Hence $\operatorname{GV}(\mathcal{Y}^d, \pi) = d + 1$. Thus $d + 1$ is the exact optimal uniform function of multiclass Littlestone dimension.

Proof idea. By Lemma 5.1, every erroneous response to the projected SOA query lands in a non-SOA label slice, whose Littlestone dimension is strictly smaller. Hence at most d errors occur. The matching lower bound is a subcube environment that fixes one free coordinate per progress-making error. Full proofs are in Appendix E. The upper bound combines the classical SOA/Littlestone conversion with the projection lemma. We do not attribute the classical improper $d + 1$ conversion to the present work.

Proposition 5.3 (Properization under arbitrary-counterexample robustness)**.** *Fix a known total order π. Suppose a deterministic, possibly improper learner A has the following guarantee: for every fixed target c and every Littlestone-type counterexample oracle that may return* any *disagreement point and its target label, A uses at most T queries, including its final accepted query. Then replacing every proposed query q_t by $P_\pi(q_t, V_t) \in V_t$ gives a deterministic version-consistent proper learner using at most T queries in the fixed-order protocol, even when the target is reselected on every round.*

Remark 5.4 (Why "arbitrary counterexample" cannot be weakened). A guarantee only against first-disagreement oracles is insufficient. For $\mathcal{H} = \{00, 01, 11\}$, a two-query strategy succeeds against every fixed target and the first-disagreement oracle. Improper queries are allowed, but this strategy is proper. Under target reselection, its lexicographically projected simulation can be forced to use three queries (Appendix E). Theorem 5.2 is proved directly without the compiler, and SOA satisfies the stronger premise.

5.1 AN EXACT INSTANCE PARAMETER AND A TRANSDUCTIVE COMPARISON

For a fixed order, let $\mathcal{F}_\pi(h, V)$ be the set of erroneous first-disagreement updates obtainable by querying h against a target in V. Call h progressive at V if each such update strictly shrinks V. Every $h \in V$ is progressive. Define recursively $\mathrm{Fdim}_\pi(V) = 0$ for $|V| = 1$ and, for $|V| \geq 2$, $\mathrm{Fdim}_\pi(V) = 1 + \min_h \max_{V' \in \mathcal{F}_\pi(h,V)} \mathrm{Fdim}_\pi(V')$, the minimum ranging over $h \in \mathcal{H}$ progressive at V.

Proposition 5.5 (Frontier dimension and calibrations)**.** *For every finite binary class and known fixed order, under target reselection and proper queries,*

$$\mathrm{GV}(\mathcal{H}, \pi) = \mathrm{Fdim}_\pi(\mathcal{H}) + 1, \qquad \mathrm{Fdim}_\pi(\mathcal{H}) \leq \min\{|\mathcal{H}| - 1, \mathrm{Ldim}(\mathcal{H})\}.$$

Moreover, singleton classes have value 2 *(one concept:* 1*) for every order, the N-member threshold class has value $\lfloor \log_2 N \rfloor + 1$, and for every $d \geq 1$ a label-cube class has $\mathrm{Ldim} = d$ but $\mathrm{GV} = 2$ under a suitable order. Thus $d + 1$ is a worst-case envelope, not an instance-wise identity.*

Let $M_{\mathrm{td}}(\mathcal{H}, \pi)$ be the optimal worst-case number of mistakes when the learner knows the ordered instance sequence in advance, predicts one label at a time, and observes the correct label after every prediction. For deterministic learners, the fixed-target and adaptive-consistent formulations of this transductive game have the same value: the final consistent label sequence can be replayed by one concept.

Theorem 5.6 (EQ reselection can be strictly harder than transduction)**.** *For every nonempty finite class over a finite label set and every fixed order,*

$$M_{\mathrm{td}}(\mathcal{H}, \pi) + 1 \leq \mathrm{GV}(\mathcal{H}, \pi) = \mathrm{Fdim}_\pi(\mathcal{H}) + 1 \leq \mathrm{Ldim}(\mathcal{H}) + 1.$$

The left inequality can be strict. For $\mathcal{H}_1 = \{000, 001, 010, 101\}$ in the displayed coordinate order, $(\mathrm{GV}, M_{\mathrm{td}}) = (3, 1)$. For the k-block product $\mathcal{H}_k = \mathcal{H}_1^k$ in serial block order,

$$\big(M_{\mathrm{td}}(\mathcal{H}_k), \mathrm{Ldim}(\mathcal{H}_k), \mathrm{GV}(\mathcal{H}_k)\big) = (k, 2k, 2k + 1).$$

Consequently, for $n \geq 1$, if $\Delta_n = \max_{\mathcal{H},\pi}\big(\mathrm{GV}(\mathcal{H}, \pi) - M_{\mathrm{td}}(\mathcal{H}, \pi)\big)$ over binary classes on n ordered points, then $\lfloor n/3 \rfloor + 1 \leq \Delta_n \leq n$.

For a finite multiclass label set, the frontier dimension in Theorem 5.6 is understood through action pairs (x, y) with $y \neq h(x)$ realizable by some target agreeing with h before x. The binary induction proving $\mathrm{GV} = \mathrm{Fdim}_\pi + 1$ extends verbatim under the precise formulation of Remark E.1 in Appendix E.

The proof embeds an optimal transductive adversary into the first-disagreement environment (Appendix E). The comparison fixes the same order on both sides. It differs from comparisons that optimize separately over worst instance orders. The separation examples are stated under the label-only convention of Remark 2.1. Under the full-transcript convention, every such gap collapses, $\mathrm{GV}_{\mathrm{full}} = M_{\mathrm{td}} + 1$ for all finite classes (Proposition A.1). The two-sided $d+1$ theorem is convention-robust.

6 THE CONSISTENT-ONLY PROTOCOL

Return to independently resampled order mixtures and target reselection, but now require every query to lie in the current version space. Along an all-negative path, each query deletes itself, so the path is a permutation of the concepts. For a permutation $\sigma = (j_1, \ldots, j_m)$ of a current set V, define $\beta_t^\sigma = \min_{s>t} q_{j_t j_s}$ for $1 \leq t \leq m - 1$.

Theorem 6.1 (Permutation and feedback-vertex-set characterizations)**.** *For a known independently resampled order mixture on a singleton class, with target reselection and version-consistent proper queries,*

$$E_{\mathrm{cons}}(V) = 1 + \min_{\sigma \in \mathfrak{S}(V)} \sum_{t=1}^{m-1} \prod_{u<t} (1 - \beta_u^\sigma).$$

In particular, the unrestricted proper value is at most the consistent-only value. For $0 < x \leq 1/2$, form the directed graph $\Gamma_x(V)$ with an arc $j \to i$ whenever $q_{ji} < x$. Then

$$\min_{\sigma \in \mathfrak{S}(V)} \#\{t : \beta_t^\sigma < x\} = \tau(\Gamma_x(V)),$$

where τ is the minimum directed feedback-vertex-set size. Consequently, there is an explicit version-consistent learner whose expected number of erroneous queries is at most $\tau(\Gamma_x(V)) + 1/x$. Including the final accepted query gives $E_{\mathrm{cons}}(V) \le 1 + \tau(\Gamma_x(V)) + 1/x$.

The graph identity is an ordering form of the unanimous target-set/FVS equivalence of Feige and Kogan (2021); the linear-order-polytope cycle inequalities are classical.

Theorem 6.2 (Consistent-only order mixtures are $\widetilde{\Theta}(\sqrt{n})$)**.** *For known independently resampled order mixtures on C_n, with target reselection and version-consistent proper queries,*

$$\Omega(\sqrt{n}) \le \sup_{\mu} D^{\mathrm{res}}_{\mathrm{cons}}(C_n, \mu) \le O\Big(\sqrt{n \log n \log\log n}\Big).$$

The upper-bound learner can be constructed in polynomial time from an explicitly given rational matrix Q, using a standard directed-feedback-vertex-set approximation algorithm (Even et al., 1998).

Proposition 6.3 (A pointwise value for inconsistent probes)**.** *There is an explicit nine-concept order mixture for which, under target reselection, $E_{\mathrm{cons}} = 8108/2187$ but $\mathrm{Esc} = 100/27 = 8100/2187$. Thus allowing a proper but version-inconsistent probe has the strict pointwise value $8/2187$.*

Finally, a counting consequence of Lemma 3.3 bounds single-threshold certificates. If a threshold x certifies $\beta_t < x$ on each of the first T steps of every trajectory, then $T + 1 < 2nx$, so the certificate's survival area is at most $\min\{T, 1/x\} \le \sqrt{T/x} < \sqrt{2n}$ (Proposition G.2 in Appendix G). Such certificates cannot beat $\sqrt{n}$. This does not rule out nonuniform multiscale hazards with intermittent high-success rounds. All proofs for this section are in Appendix G.

7 Related work

The COLT 2026 paper of Braverman et al. (2026) proves the randomized $O(\mathrm{Ldim})$ upper bound for symmetric adversaries and asks whether randomization is necessary. Our answers for dimension-only guarantees separate target reselection from fixed targets, fixed known orders from independently resampled orders, unrestricted proper from version-consistent queries, and known from unknown generators. Angluin and Dohrn (2017) study random counterexamples on the *teacher's* side, with deterministic $O(\log|\mathcal{H}|)$ rules in that model given by Bhatia (2021) and Littlestone-dimension extensions by Chase et al. (2023). Our separations concern learner-side randomization under the same symmetric oracle. Among recent related works, Alon et al. (2026) determine the optimal complexity of sorting from possibly untruthful counterexamples, with no target reselection, no symmetric generators, and no deterministic–randomized separation. Dewasurendra (2026) separates randomized from deterministic risk in fixed-design list transduction, a different model from equivalence queries. Li (2026) proves an exponential deterministic–randomized gap in an ERM-oracle model that, to our knowledge, admits no reduction to ours in either direction. The improper $\mathrm{Ldim}+1$ conversion is classical (Angluin, 1988; Littlestone, 1988). The transductive model goes back to Ben-David et al. (1997), with block products as in Chase et al. (2025). Tournament domination bounds are classical (Erdős, 1963; Graham and Spencer, 1971), as are directed-cycle inequalities for the linear-order polytope (Grötschel et al., 1985). The consistent-only upper bound uses Even et al. (1998). Test-ordering survival areas (Berend et al., 2014) have exogenous rather than endogenous hazards. Appendix I gives the complete related-work discussion. Its distinctions of scope, novelty, and chronology are integral to the paper.

8 Conclusion

We resolve the dimension-only derandomization question posed by Braverman et al. (2026) at COLT 2026 for general symmetric equivalence queries. Our negative answer holds even with fixed targets and extends to independently resampled order mixtures under target reselection. A fixed known order restores the optimal proper bound $d + 1$, while fixed targets have a tight worst-case $\Theta(\log n)$ cost on singleton classes. Appendix H states five open problems (O-N, O-M′, O-A′, O-I, O-F): the extremal order-mixture constant, the consistent-only polylogarithmic gap, the fixed-target cyclic limit, hidden fixed orders, and dimension-only fixed-target guarantees for general order mixtures. Knowing the generator is indispensable: even with a fixed target, an unknown generator forces n queries on C_n (Proposition H.1).

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, and assist in writing proofs. We also used these tools to design or provide feedback on research methodology, implement methods, 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 data sets, cleaning and reformatting data sets, and qualitative and thematic data analysis are not applicable to this work. The numerical values used for verification and plotting were computed from mathematical formulas and algorithms.

Additionally, we used generative AI tools to create or modify scientific figures, create or edit software code, and draft parts of the paper. We used them to summarize or analyse existing literature, discover research topics or identify gaps, brainstorm, and source or search for information. We also used them to improve the paper's readability, identify relevant literature, format references, suggest the paper's structure, and propose a title or keywords.

We have reviewed all AI-assisted work. Three individuals manually checked all mathematical derivations and AI-assisted code and found no errors. 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 is a theoretical study of query complexity in a mathematical model of interactive learning. It involves no data sets, no human subjects, and no deployment; we do not foresee direct ethical concerns beyond the general dual-use character of learning theory.

REPRODUCIBILITY STATEMENT

All results are mathematical theorems. Complete proofs of every statement in the main text are given in Appendices A–G, with the model and all quantifier conventions fixed in Section 2. Every closed-form constant reported in the paper (exact game values, the 20-probe certificate, and the nine-concept example) has additionally been verified with exact rational arithmetic; the integer certificates are stated in the appendix in a directly checkable form.

REFERENCES

Noga Alon, Shay Moran, and Shlomo Moran. Sorting from counterexamples. arXiv preprint arXiv:2608.21579, August 21, 2026. URL https://arxiv.org/abs/2608.21579. DOI: 10.48550/arXiv.2608.21579.

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

D. Angluin and T. Dohrn. The power of random counterexamples. In *Proceedings of the 28th International Conference on Algorithmic Learning Theory (ALT)*, PMLR 76:452–465, 2017.

S. Ben-David, E. Kushilevitz, and Y. Mansour. Online learning versus offline learning. *Machine Learning*, 29(1):45–63, 1997. Preliminary version in *EuroCOLT 1995*, LNCS 904:38–52, Springer, 1995.

D. Berend, R. Brafman, S. Cohen, S. E. Shimony, and S. Zucker. Optimal ordering of independent tests with precedence constraints. *Discrete Applied Mathematics*, 162:115–127, 2014.

J. S. Bhatia. Simple and fast algorithms for interactive machine learning with random counterexamples. *Journal of Machine Learning Research*, 22(15):1–30, 2021.

Mark Braverman, Roi Livni, Yishay Mansour, Shay Moran, and Kobbi Nissim. Learning from equivalence queries, revisited. In *Proceedings of Thirty Ninth Conference on Learning Theory*, PMLR 336:808–836, June 29, 2026. URL https://proceedings.mlr.press/v336/braverman26a.html. arXiv:2604.04535. Preprint DOI: 10.48550/arXiv.2604.04535.

T. Bui, T. Lidbetter, and K. Y. Lin. Optimal pure strategies for a discrete search game. *European Journal of Operational Research*, 313(2):767–775, 2024.

H. Chase, J. Freitag, and L. Reyzin. Applications of Littlestone dimension to query learning and to compression. arXiv:2310.04812, 2023. Journal version: *Information and Computation*, 2026, doi:10.1016/j.ic.2026.105424.

Zachary Chase, Steve Hanneke, Shay Moran, and Jonathan Shafer. Optimal mistake bounds for transductive online learning. arXiv preprint arXiv:2512.12567, December 14, 2025. URL `https://arxiv.org/abs/2512.12567`. Preprint DOI: 10.48550/arXiv.2512.12567. Conference version: NeurIPS 2025. DOI: 10.52202/085713-5090.

X. Chen, M. Lin, N. Schärli, and D. Zhou. Teaching large language models to self-debug. In *The Twelfth International Conference on Learning Representations (ICLR)*, 2024. `https://openreview.net/forum?id=KuPixIqPiq`.

P. Dewasurendra. Select, do not vote: optimal list PAC learning and exact transduction. Preprint, Zenodo. First version August 11, 2026 (records 21885813 and 21885814); version 2, DOI 10.5281/zenodo.22061106, August 22, 2026.

P. Erdős. On a problem in graph theory. *The Mathematical Gazette*, 47(361):220–223, 1963.

G. Even, J. Naor, B. Schieber, and M. Sudan. Approximating minimum feedback sets and multicuts in directed graphs. *Algorithmica*, 20(2):151–174, 1998.

U. Feige and S. Kogan. Target set selection for conservative populations. *Discrete Applied Mathematics*, 305:119–132, 2021.

R. L. Graham and J. H. Spencer. A constructive solution to a tournament problem. *Canadian Mathematical Bulletin*, 14(1):45–48, 1971. For $p > k^2 2^{2k-2}$ the Paley tournament T_p has property P_k; hence its domination number is $\Omega(\log p)$.

M. Grötschel, M. Jünger, and G. Reinelt. Facets of the linear ordering polytope. *Mathematical Programming*, 33(1):43–60, 1985. See also: On the acyclic subgraph polytope, *Mathematical Programming*, 33(1):28–42, 1985.

X. Li. An exponential deterministic–randomized gap in ERM-oracle complexity for thresholds on an unknown order. arXiv:2609.10196, September 9, 2026.

N. Littlestone. Learning quickly when irrelevant attributes abound: a new linear-threshold algorithm. *Machine Learning*, 2(4):285–318, 1988. Section 4, Theorem 5, gives the mistake-bound to equivalence-query conversion; its queries may be improper.

A. Solar-Lezama, L. Tancau, R. Bodik, S. A. Seshia, and V. A. Saraswat. Combinatorial sketching for finite programs. In *Proceedings of the 12th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, pp. 404–415. ACM, 2006.

A Model details

A.1 Relation to the protocol of Braverman et al.

Appendix C of Braverman et al. (2026) permits target reselection from the current version space and restricts the learner to $h_t \in V_t$. The parent paper also permits the environment to accept an incorrect hypothesis early. Here every incorrect query must return a counterexample. This difference does not affect either set of results: a worst-case environment may always refuse early acceptance, and the parent paper's upper-bound proof does not use it. Our deterministic lower bounds allow the learner the larger action space $h_t \in \mathcal{H}$ and hence automatically apply to that narrower protocol. Some of our matching upper bounds, most notably the blocking-point upper bound for general order mixtures, use version-inconsistent probes. Section 6 treats the consistent-only restriction separately.

Table 2: Reading guide. The upper panel follows the main-text order and lists proof locations. The lower panel records selected direct uses; it is not an exhaustive dependency list.

Reading branch	Statements	Proof locations
Reselection	Theorem 3.1; Lemmas 3.2, 3.3; Proposition 3.4	Appendices B, C.1–.2, D
Fixed targets	Propositions 4.1, 4.5; Theorems 4.2–4.4	Appendices F.1–.3, .5–.6; C.2–.4
Fixed known order	Lemma 5.1; Theorem 5.2; Proposition 5.3; Remark 5.4; Proposition 5.5; Theorem 5.6	Section 5; Appendices E.1–.4, A.2
Consistent queries	Theorems 6.1, 6.2; Proposition 6.3	Appendix G.1–.3
Input result or definition	**Used in the proof of**	**Use locations**
Lemma 3.2	Theorem C.1	Appendix C.1
Lemma 3.3	Proposition 3.4; Theorem D.3	Appendices B, D
Proposition 4.1	Theorem C.4	Appendix C.3
Lemma 5.1	Theorem 5.2; Proposition 5.3	Appendix E.1
SOA definition before Theorem 5.2	Theorem 5.2	Appendix E.1
Theorem 6.1	Theorem 6.2	Appendix G.1, .3

A.2 THE TWO VERSION-SPACE CONVENTIONS

Under the full-transcript convention of Remark 2.1, the prefix agreements of a first-disagreement witness are also retained as permanent constraints. Some of our results then change. The transductive separation family of Theorem 5.6 is specific to the label-only convention. For $\mathcal{H}_1$ the full-transcript value drops from 3 to 2, and the $n = 3$ separations disappear. The complete-class optimum $d + 1$ of Theorem 5.2, however, holds under both conventions.

To show that the complete-class lower bound also holds under the full-transcript convention, we refine the environment in the proof of Theorem 5.2. It responds at the *earlier* of the earliest free coordinate and the earliest violated fixed coordinate. When the earliest violated fixed coordinate precedes every free coordinate, the returned point's prefix reveals only already-fixed coordinates, so the subcube does not shrink. Otherwise the earliest free coordinate comes first, and the environment differs there. The prefix again contains no free coordinate, so exactly one free coordinate is fixed per progress-making error. The count of d forced errors is unchanged.

More generally, the full-transcript convention collapses every gap of the kind exhibited in Theorem 5.6, not only the $\mathcal{H}_1$ examples:

Proposition A.1 (Full transcript collapses the transductive gap)**.** *For every nonempty finite class over a finite label set and every enumeration π of the domain in order type at most ω, the reselection value under the full-transcript convention satisfies*

$$\mathrm{GV}_{\mathrm{full}}(\mathcal{H}, \pi) = M_{\mathrm{td}}(\mathcal{H}, \pi) + 1.$$

Proof. A one-concept class has $\mathrm{GV}_{\mathrm{full}} = 1 = M_{\mathrm{td}} + 1$ directly. When the enumerated domain is infinite, let N be the largest of the finitely many pairwise π-first disagreement positions. The prefix $x_1, \ldots, x_N$ separates every two concepts. Every first-disagreement response to a proper query lies inside it. Once its labels are known, the consistency set is a singleton. Thus neither game changes when the enumeration is truncated to this prefix, and we may assume the domain finite.

The lower bound $M_{\mathrm{td}} + 1 \le \mathrm{GV}_{\mathrm{full}}$ follows from the adversary embedding in the proof of Theorem 5.6, whose responses are first-disagreement responses and remain legal when prefix agreements are recorded. For the upper bound, note first that under the full-transcript convention every response determines the target's labels on an initial segment of π, so the version space is always the set of concepts through a known labeled prefix. Normalize an optimal transductive learner to predict only labels consistent with the current constraints. An inconsistent prediction is a guaranteed mistake, so this loses nothing. From the current known prefix, run this learner along the hypothetical path on

which all of its future predictions are correct. Each prediction keeps the consistency set nonempty, so some $h \in V$ realizes the completed labeling. The learner queries this hypothesis. If h is accepted, the run ends. Otherwise the first disagreement x arrives together with the information that the current target agrees with h before x. This advances the internal transductive game to x and charges it exactly one mistake there. Every erroneous query is thus injectively matched with a mistake of one continuing adaptive-consistent transductive game, giving at most M_{td} erroneous queries plus the final accepted one. □

B Deferred proofs for Section 3

This appendix proves Lemma 3.2, Lemma 3.3, and Proposition 3.4, and records the randomized three-query computation of Theorem 3.1(4).

Proof of Lemma 3.2. For the upper bound, query

$$h_0, h_{-1}, h_{-2}, \dots$$

until a positive counterexample identifies a target, and then query that target. If only one concept remains, query it. Suppose the first $t-1$ exploratory queries all returned their negative query points. Those $t-1$ points form a contiguous dead zone. Every currently legal target other than the probe itself has forward cyclic distance at least t from the tth probe. A target equal to the probe is accepted outright. Thus the conditional probability of another negative response is at most $1-t/n$. The target choices need not be independent. If N is the number of erroneous queries before the final accepted query, then

$$\Pr[N \geq t] \leq S_{n,t-1},$$

and the tail-sum identity gives the asserted upper bound.

For the lower bound, fix an arbitrary deterministic learner. At a public history, let $e = n - |V|$ be the number of eliminated concepts and let h_j be the learner's next query. If $j \in V$ and $|V| \geq 2$, the closest surviving successor of j other than j has distance at most $e+1$: otherwise the first $e+1$ successors would all be dead. If $j \notin V$, then j itself is dead, so a surviving successor has distance at most e. The environment chooses such a closest surviving target, except when only one target remains.

After $r-1$ negative responses, at most $r-1$ concepts have been eliminated. The selected target therefore has positive-counterexample probability at most r/n, and hence

$$\Pr[\text{the first } r \text{ responses are negative} \mid \text{the first } r-1 \text{ are negative}] \geq 1 - \frac{r}{n}.$$

Thus $\Pr[N \geq t] \geq S_{n,t-1}$ for $1 \leq t \leq n-1$. The target choice is based only on the public history and on the deterministic next query; it does not observe the fresh random cyclic starting point. The tail-sum identity, interpreted by monotone convergence if necessary, proves the lower bound. □

This is a Stackelberg value: the lower-bound environment is tailored to the deterministic learner. It is not an environment-independent saddle point.

Proof of Lemma 3.3. For a fixed row, the sum of its r smallest entries is the pointwise minimum of all r-subset row sums. It is therefore concave in that row. By Jensen's inequality, it suffices to evaluate the sum over rows at a total-order vertex.

If a row occupies position p in the order, its r-smallest-entry sum is

$$\max\{0, r-(n-p)\}.$$

Summing over $p = 1, \dots, n$ gives $1 + \cdots + r = r(r+1)/2$. Finally $L_{j,r} \leq r\alpha_{j,r}$, so averaging over rows gives the stated blocking row. □

To the best of our knowledge, this row lower-envelope formulation has not previously been used in equivalence-query learning. The inequality itself is a folklore consequence of concavity and evaluation at vertices. Our claim concerns only the formulation and its learning application, not a new facet. Remark D.2 records which endpoints are derivable from the standard relaxations.

Proof of Proposition 3.4. Fix j, r. Query the $r-1$ concepts corresponding to the weakest off-diagonal entries in row j. A negative response deletes the queried concept. Acceptance or a positive response only helps. After these probes, repeatedly query h_j. On every repetition the favorable event is either acceptance of h_j itself, when h_j is the current target, or a positive counterexample. This event has probability at least $\alpha_{j,r}$, conditionally on every history. The expected number of repeated probes is at most $1/\alpha_{j,r}$, and one final query completes learning. The total is at most $r + 1/\alpha_{j,r}$.

After the first negative response to h_j, its repetitions may be version-inconsistent, but they remain proper. Lemma 3.3 provides a row with $\alpha_{j,r} \geq (r+1)/(2n)$. Optimizing r proves the result. □

The sharper asymptotic constant, compactness of the extremal problem, and the matching cyclic lower bound are proved in Appendix D.

B.1 The randomized three-query strategy

The randomized $O(1)$ upper bound itself follows from Theorem 4 of Braverman et al. (2026), whose proof yields $2d + 1 = 3$ here. The new result in this section is the deterministic lower bound. We include the explicit strategy below only to give a self-contained separation with an exact constant.

For a translation-invariant symmetric kernel $q_{ji} = \phi(d(j,i))$, symmetry gives

$$\sum_{d=1}^{n-1} \phi(d) = \frac{n-1}{2}.$$

After the environment selects the current-round target $c_t = h_i$, sample $J \sim \mathrm{Unif}(\mathbb{Z}_n)$ and query h_J. Conditional on every public history,

$$p_{\mathrm{correct}} = \frac{1}{n}, \qquad p_+ = p_- = \frac{n-1}{2n}.$$

After a positive counterexample, query the revealed target. If E denotes the worst conditional expected number of remaining queries,

$$E \leq 1 + p_- E + p_+,$$

and therefore

$$E \leq \frac{1+p_+}{1-p_-} = \frac{3n-1}{n+1} < 3.$$

The probabilities are independent of the current version space and of the adaptive target policy.

C Exact kernel calculations

C.1 The cyclic asymptotic expansion

Theorem C.1 (Exact cyclic value and expansion)**.** *For every $n \geq 2$, in the known cyclic generator with target reselection and deterministic proper queries,*

$$D_n = 1 + \sum_{k=0}^{n-2} \frac{(n-1)!}{(n-1-k)!\, n^k}.$$

Moreover,

$$D_n = \sqrt{\frac{\pi n}{2}} + \frac{2}{3} + \frac{1}{12}\sqrt{\frac{\pi}{2n}} + O\left(\frac{1}{n}\right).$$

Proof. The exact formula is Lemma 3.2. For the expansion, put $\varepsilon = n^{-1/2}$, $x = k/\sqrt{n}$, and

$$S_{n,k} = \prod_{r=1}^{k} \left(1 - \frac{r}{n}\right).$$

Taylor expansion of the logarithm gives, uniformly for $k \le n^{3/5}$,

$$\log S_{n,k} = -\frac{x^2}{2} - \varepsilon A(x) - \varepsilon^2 B(x) + O\big(\varepsilon^3 P(x)\big)\,,$$

where

$$A(x) = \frac{x}{2} + \frac{x^3}{6}, \qquad B(x) = \frac{x^2}{4} + \frac{x^4}{12},$$

and P is a fixed polynomial. Exponentiating,

$$S_{n,k} = e^{-x^2/2}\left[1 - \varepsilon A(x) + \varepsilon^2 C(x) + O\big(\varepsilon^3 P_1(x)\big)\right],$$

with

$$C(x) = \frac{A(x)^2}{2} - B(x) = -\frac{x^2}{8} + \frac{x^6}{72}.$$

The needed Gaussian integrals are

$$\int_0^\infty e^{-x^2/2}\,dx = \sqrt{\frac{\pi}{2}},$$

$$\int_0^\infty A(x)e^{-x^2/2}\,dx = \frac{5}{6}, \qquad \int_0^\infty C(x)e^{-x^2/2}\,dx = \frac{1}{12}\sqrt{\frac{\pi}{2}}.$$

Euler–Maclaurin contributes the endpoint term $1/2$. Hence

$$\sum_{k=0}^{n-2} S_{n,k} = \sqrt{\frac{\pi n}{2}} - \frac{1}{3} + \frac{1}{12}\sqrt{\frac{\pi}{2n}} + O\left(\frac{1}{n}\right).$$

Adding the final-query term 1 gives the theorem.

To justify the expansion, on $k \le n^{3/5}$ the summed remainder is $O(n^{-1})$ after multiplication by a Gaussian envelope. For the tail, use

$$S_{n,k} \le \exp\left(-\frac{k(k+1)}{2n}\right),$$

which makes the omitted contribution at most $n\exp(-cn^{1/5})$. □

For example, the exact formula gives $D_2 = 2$, $D_3 = 8/3$, and $D_4 = 25/8$.[1]

[1] Exact-arithmetic dynamic programs and direct hand checks agree on the reported small values; these computations are sanity checks, not part of the proof.

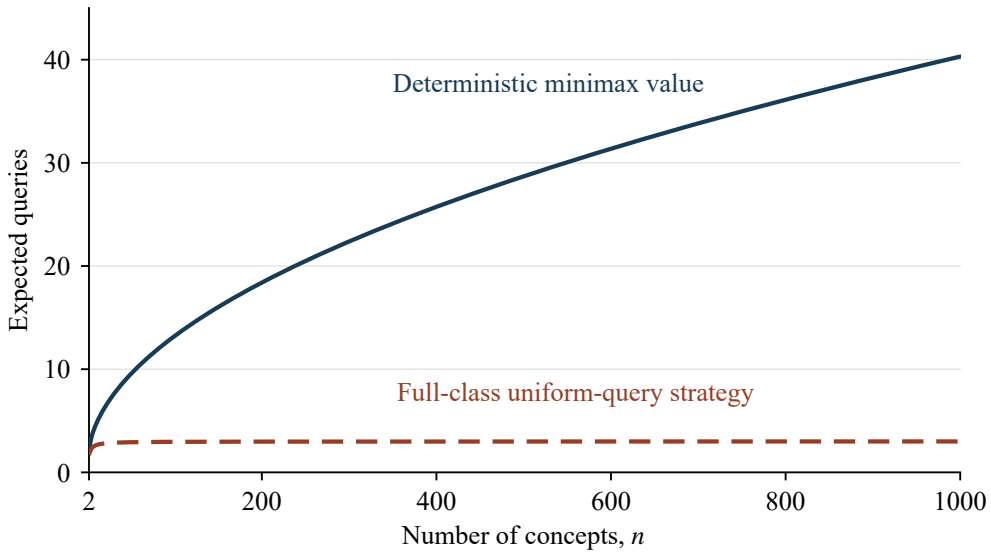


Figure 1: Finite-size costs for the known cyclic generator under target reselection. The curves directly evaluate the deterministic minimax value and the expected cost of the full-class uniform-query strategy in Theorem 3.1(1) and (4). Query counts include the final accepted query. The deterministic worst case is taken separately for each learner. See Lemma 3.2, Theorem C.1, and Appendix B.1.

C.2 General monotone symmetric kernels

Let

$$\phi_n : \{1, \dots, n-1\} \to [0, 1]$$

be nondecreasing and satisfy

$$\phi_n(d) + \phi_n(n-d) = 1.$$

The generator returns the positive target point with probability $\phi_n(d(j, i))$.

Theorem C.2 (Monotone kernels, power laws, and a linear example)**.** *Assume the generator is known, the target may be reselected, and the learner is deterministic and proper, possibly version-inconsistent.*

1. *The exact value is*

$$D_n(\phi) = 1 + \sum_{t=1}^{n-1} \prod_{r=1}^{t-1} \big(1 - \phi_n(r)\big).$$

2. *For every $\rho > 0$, let*

$$\phi_{\rho,n}(d) = \frac{d^\rho}{d^\rho + (n-d)^\rho}.$$

Then

$$D_n(\phi_\rho) \sim K_\rho n^{\rho/(\rho+1)}, \qquad K_\rho = (\rho+1)^{1/(\rho+1)} \Gamma\left(1 + \frac{1}{\rho+1}\right).$$

3. *For the threshold kernel*

$$\phi_n(d) = \begin{cases} 0, & d < n/2, \\ 1/2, & d = n/2, \\ 1, & d > n/2, \end{cases}$$

one has, for $n \geq 3$,

$$D_n(\phi) = \frac{n+3}{2}.$$

For $n = 2$ the value is 2.

4. *Every kernel above is symmetric, and the randomized uniform-query strategy has expected query count* $(3n-1)/(n+1) < 3$ *against every legal adaptive reselection environment.*

Proof. The dead-zone learner reaches its tth exploratory query only after $t-1$ distinct negative points have formed a contiguous dead zone. Every legal target other than the probe itself has distance at least t, and monotonicity gives positive-counterexample probability at least $\phi_n(t)$. A target equal to the probe is accepted outright. This gives the product upper bound.

For the converse, the successor environment selects the closest surviving target. After at most $t-1$ eliminations, the gap lemma gives distance at most t, so monotonicity bounds the positive probability by $\phi_n(t)$. Multiplication of the conditional negative probabilities gives the matching lower bound.

For the power-law family, use the scale

$$a_n = n^{\rho/(\rho+1)}.$$

Uniformly for $r = O(a_n)$,

$$\phi_{\rho,n}(r) \sim \left(\frac{r}{n}\right)^{\rho}.$$

For $t = \lfloor x a_n \rfloor$,

$$\sum_{r\le t} \phi_{\rho,n}(r) \longrightarrow \frac{x^{\rho+1}}{\rho+1}, \qquad \sum_{r\le t} \phi_{\rho,n}(r)^2 = O\Big(n^{-\rho/(\rho+1)}\Big).$$

Thus the survival product converges to $\exp(-x^{\rho+1}/(\rho+1))$. For $r \le n/2$,

$$\phi_{\rho,n}(r) \ge \frac{1}{2}\left(\frac{r}{n}\right)^{\rho},$$

which gives an integrable stretched-exponential envelope. The part $r > n/2$ is exponentially small. Dominated convergence for the resulting Riemann sums yields

$$\int_0^\infty e^{-x^{\rho+1}/(\rho+1)}\,dx = (\rho+1)^{1/(\rho+1)}\Gamma\left(1+\frac{1}{\rho+1}\right).$$

The threshold formula follows by direct substitution into the exact product expression. Symmetry and the randomized bound follow from $\sum_d \phi_n(d) = (n-1)/2$. □

Remark C.3 (Monotonicity is essential). For $n = 4$, the symmetric but nonmonotone kernel

$$\phi = (0.9, 0.5, 0.1)$$

would give 2.15 if the formula of Theorem C.2 were applied formally. However, after every first query h_j, the environment can select h_{j-1}, whose positive probability is only 0.1, giving

$$\mathbb{E}Q \ge 1 + 1 + 0.9 = 2.9.$$

Thus symmetry alone does not justify the product formula.

For $n \ge 3$, the threshold kernel is not an order mixture. Along the directed Hamilton cycle of unit forward steps, all relevant q-entries are zero. Every matrix in LOP_n, however, satisfies the cycle inequality requiring their sum to be at least one. For $n = 2$ it is the uniform mixture of the two total orders.

C.3 Fixed-target cyclic constants

Theorem C.4 (Fixed-target cyclic bounds)**.** *For the known cyclic kernel and one fixed target,*

$$D_2^{\mathrm{fix}} = 2, \qquad D_3^{\mathrm{fix}} = \frac{7}{3}, \qquad D_4^{\mathrm{fix}} = \frac{5}{2},$$

and

$$D_n^{\mathrm{fix}} \le 4 \qquad (n \ge 2).$$

Furthermore,

$$\frac{65}{24} \le \liminf_{n\to\infty} D_n^{\mathrm{fix}} \le \limsup_{n\to\infty} D_n^{\mathrm{fix}} \le \frac{2278131453257646399}{639999985150744579} \approx 3.5595804783.$$

Proof. For $n = 2, 3, 4$, translation symmetry lets us fix the first query at 0. Recall $a_{ji} = 1-d(j,i)/n$ and the spine value $V_i(\sigma) = 1 + \sum_{t\geq 1} \mathbf{1}\{j_t \neq i\} R_i^{\sigma}(t-1)$ from Proposition 4.1. For $n = 2$, the spine $(0, 1)$ gives $V_0 = 1$ and $V_1 = 2$. Conversely, $V_i \geq 2$ for every target other than the first query, so $D_2^{\text{fix}} = 2$. For $n = 3$, the spine $(0, 1, 2)$ gives $V_1 = 2$ and $V_2 = 1 + 1 + a_{02} = 7/3$. Conversely, with $j_1 = 0$, if $j_2 \neq 1$ then $V_1 \geq 1 + 1 + a_{01} = 8/3$, so an optimal spine has $j_2 = 1$, and then $V_2 \geq 1 + 1 + a_{02} = 7/3$. Thus $D_3^{\text{fix}} = 7/3$. For $n = 4$, the spine $(0, 1, 2, 3)$ gives $V_1 = 2$, $V_2 = 1+1+a_{02} = 5/2$, and $V_3 = 1+1+a_{03}+a_{03}a_{13} = 19/8$. Conversely, with $j_1 = 0$, if $j_2 \neq 1$ then $V_1 \geq 1 + 1 + a_{01} = 11/4$. Thus $j_2 = 1$. If $j_3 \neq 2$ then $V_2 \geq 1 + 1 + a_{02} + a_{02}a_{12} = 23/8$. Thus $j_3 = 2$, which gives $V_2 = 5/2$ exactly. Hence $D_4^{\text{fix}} = 5/2$.

For the uniform bound, choose the following three probe positions:

n	probe positions
$3a$	$(0, a, 2a)$
$3a + 1$	$(0, a, 2a + 1)$
$3a + 2$	$(0, a + 1, 2a + 1)$.

Repeat this block of three probes periodically, in the displayed cyclic order. The probes partition the cycle into three arcs. Let A, B, C denote their lengths in scanning order. Then (A, B, C) equals (a, a, a), $(a, a + 1, a)$, or $(a + 1, a, a + 1)$ in the three residue classes. Consider a fixed target at offset $u \geq 1$ beyond the third probe, in the C-arc. Over one full period of probes, its three negative-response factors $a_{ji} = 1 - d(j, i)/n$ are

$$(f_0, f_1, f_2) = \left(\frac{C - u}{n}, \frac{A + C - u}{n}, \frac{n - u}{n} \right),$$

its per-period survival product is $P = f_0 f_1 f_2$, and its exploratory survival area under the periodic repetition is the geometric sum $(1 + f_0 + f_0 f_1)/(1 - P)$. Define the certificate

$$H_C = 3(1 - P) - (1 + f_0 + f_0 f_1),$$

so that $H_C \geq 0$ is exactly the statement that this area is at most 3. Every factor is nonincreasing in u, so the area is largest at the boundary offset $u = 1$. Substituting $u = 1$ in the three residue classes, the C-arc certificates are

$$H_C^{(3a)} = \frac{7a^3 + 17a^2 - 7a + 1}{9a^3},$$

$$H_C^{(3a+1)} = \frac{21a^3 + 91a^2 + 14a + 2}{(3a + 1)^3},$$

and

$$H_C^{(3a+2)} = \frac{21a^3 + 74a^2 + 63a + 16}{(3a + 2)^3},$$

each positive for every $a \geq 1$. Targets in the A- and B-arcs are handled in the same way, using their own ordered factor triples. These are *not* rotations of the C-arc triple when the arc lengths differ. Scanning one period from the first probe, a target at offset u in the A-arc has

$$(g_0, g_1, g_2) = \left(\frac{n - u}{n}, \frac{A - u}{n}, \frac{A + B - u}{n} \right),$$

and one in the B-arc has

$$(g_0, g_1, g_2) = \left(\frac{B + C - u}{n}, \frac{n - u}{n}, \frac{B - u}{n} \right),$$

with the certificate $3(1 - g_0 g_1 g_2) - (1 + g_0 + g_0 g_1)$ in each case. Every factor again decreases in u, so the worst offset is $u = 1$, where the six certificates, multiplied by n^3, are

n	A-arc	B-arc
$3a$	$54a^2 - 21a + 3$	$57a^2 - 21a + 3$
$3a + 1$	$60a^2 + 18a + 2$	$36a^2 + 16a + 2$
$3a + 2$	$48a^2 + 46a + 12$	$78a^2 + 66a + 16$,

all positive for every integer $a \geq 1$. A target coinciding with a probe only terminates earlier, with exploratory area at most 2. Hence every target has exploratory survival area at most 3, and one final query gives $D_n^{\text{fix}} \leq 4$.

For the asymptotic lower bound, put the uniform distribution on the fixed target and apply Yao's principle to a deterministic main spine. The first three nontrivial tail probabilities are bounded below by

$$\frac{n-1}{n}, \qquad \frac{(n-1)(n-2)}{2n^2}, \qquad C_n - \frac{1}{n},$$

where

$$C_n = \min_a \frac{3a^2 - 3an + 2n^2 - 3n + 1}{6n^2}.$$

Here, with the first spine query rotated to 0, the second at a, and $a_{ji} = 1 - d(j,i)/n$ with zero diagonal, the underlying identity is

$$\frac{1}{n}\sum_i a_{0i}a_{ai} = \frac{3a^2 - 3an + 2n^2 - 3n + 1}{6n^2},$$

valid for every a, including a repeated query $a = 0$. The third spine query deletes at most one term, of size at most $1/n$, from this average, giving the third tail bound $C_n - 1/n$. The minimum occurs at an admissible integer $a = n/2 + O(1)$, and hence

$$C_n = \frac{5n^2 - 12n + O(1)}{24n^2} \longrightarrow \frac{5}{24}.$$

Together with the final-query term, this gives

$$1 + 1 + \frac{1}{2} + \frac{5}{24} = \frac{65}{24}.$$

For the upper bound, execute the 20 probes in the order

$$(p_0, \ldots, p_{19}) = (0, 6, 9, 11, 12, 2, 3, 13, 8, 5, 1, 16, 7, 18, 10, 14, 4, 17, 15, 19),$$

using the integer probes $\lfloor p_t n/20 \rfloor$ and repeating the block of 20 probes periodically. The geometric denominator D below accounts for this repetition. For $r \in \mathbb{Z}_{20}$, define the pure-integer certificate

$$z_t = \begin{cases} 20, & p_t = r, \\ (p_t - r) \bmod 20, & p_t \neq r, \end{cases}$$

$$S_r = \sum_{t=0}^{19} \left(\prod_{u<t} z_u \right) 20^{19-t}, \qquad D = 20^{19} - 19!.$$

Within each gap, every relevant negative probability is monotone in the target offset, so the worst boundary is $u = 1$. Its value is

$$\frac{D + S_r}{D}.$$

The maximum over r occurs at $r = 0$, and after reduction this limiting certificate value is exactly

$$\frac{2278131453257646399}{639999985150744579}.$$

Within each gap every negative factor is monotone in the target offset, so it suffices to compare the right-hand limits of the survival factors immediately after each probe. A target coinciding with a probe only terminates earlier and is dominated by the neighboring gap bound. The discretized factors converge uniformly at rate $O(1/n)$ and their geometric denominators stay bounded away from zero, so the finite certificate and the limit may be interchanged, yielding the stated limsup. □

The theorem does not assert that the limit exists.

C.4 A UNIFORM FIXED-TARGET BOUND FOR ALL MONOTONE KERNELS

Theorem C.5 (Fixed-target monotone-kernel bound)**.** *For every known monotone translation-invariant symmetric kernel on C_n, with one fixed target and deterministic proper queries,*

$$D_n^{\mathrm{fix}}(\phi) \le \frac{9}{2}.$$

Consequently, within this kernel family, an unbounded sequence exists under target reselection but not under a fixed target.

Proof. Let

$$A = 0, \qquad B = \lfloor n/2 \rfloor,$$

and use the probe sequence

$$A, B, B, A, B, A, B, A, \ldots,$$

that is, the prefix A, B followed by the alternation $B, A, B, A, \ldots$ (the sequence $1, 2, (2, 1)$ in the notation of Bui et al. (2024)). Symmetry and monotonicity imply

$$\phi_n(\lceil n/2 \rceil) \ge \frac{1}{2}.$$

For every target other than the probes, at least one of A, B has success probability at least $1/2$. Splitting according to which probe is favorable and summing the two possible prefix-survival series gives an exploratory expectation at most $7/2$. A positive counterexample is followed by one final query, giving $9/2$. Targets equal to a probe terminate no later.

To calibrate the constant, consider the following auxiliary *two-probe information model*, a two-box search instance (Bui et al., 2024). On every round, the learner picks one of two probes. There are two candidate targets, one favorable to each probe. A probe ends the run against its favorable target with probability exactly $1/2$ and never against the other. After t rounds with a_t picks of the first probe and b_t of the second, the two survival probabilities are 2^{-a_t} and 2^{-b_t}. For a prefix with $a_t + b_t = t$, the pointwise minimum is

$$\min_{a+b=t} \left(2^{-a} + 2^{-b}\right) = \begin{cases} 2 \cdot 2^{-t/2}, & t \text{ even}, \\ 3 \cdot 2^{-(t+1)/2}, & t \text{ odd}. \end{cases}$$

Pairing consecutive even and odd values gives

$$\left(2 + \frac{3}{2}\right) + \left(1 + \frac{3}{4}\right) + \cdots = 3.5 \sum_{k \ge 0} 2^{-k} = 3.5 \cdot 2 = 7.$$

Since the two targets' survival areas sum to at least 7, one of them is at least $7/2$. Thus the exploratory constant $7/2$ of the strategy above is exactly optimal within this auxiliary model. This calibrates the strategy, not the kernel family. It does not assert a matching $9/2$ lower bound for monotone kernels.

The reselection side is unbounded for the cyclic, power-law, and threshold families by Theorem C.2, while the fixed-target side is uniformly bounded by the first part. □

D ORDER-MIXTURE EXTREMALITY AND CONSTANTS

Proposition D.1 (Linear-order-polytope formulation)**.** *In the known target-reselection protocol on C_n, with a fresh order drawn independently each round,*

$$M_n = \sup_{\mu} D^{\mathrm{res}}(C_n, \mu) = \max_{Q \in \mathrm{LOP}_n} \mathrm{Esc}(Q).$$

The value $\mathrm{Esc}(Q)$ is continuous on LOP_n, and a maximizing matrix has a representing distribution supported on at most $\binom{n}{2} + 1$ total orders.

Moreover, every $Q \in \mathrm{LOP}_n$ and every directed cycle $v_1 \to \cdots \to v_\ell \to v_1$ satisfy

$$\sum_{t=1}^{\ell} q_{v_t v_{t+1}} \ge 1.$$

Proof. Every order distribution produces its pairwise-precedence matrix in LOP_n, and the two-point Bellman recurrence depends only on this matrix. Conversely, every point of LOP_n is a convex combination of total-order vertices and hence has an order-distribution representation.

Continuity follows from the finite Bellman recursion, once its action values are made explicit. Let $k = |V|$ be the number of surviving concepts. A query $h \in V$ deletes itself on every negative response, so by induction on k its action value is at most k. A query $h \notin V$ has a self-loop. Let $\alpha = \min_{i \in V} q_{hi}$ be the least positive-response probability against surviving targets. The query's stationary repeated-use value solves $B = 1 + (1-\alpha)B + \alpha \cdot 1$, or $B = 1 + 1/\alpha$. The self-loop term can therefore be truncated by

$$\overline{B} = \min\{k+1,\ 1 + 1/\alpha\}$$

without changing the minimum, because a currently surviving query always offers value at most k. This gives a continuous finite-state recursion. Compactness yields a maximum, and Carathéodory's theorem gives support size at most $\binom{n}{2} + 1$.

At a deterministic total order, every directed cycle contains at least one edge whose head precedes its tail. The corresponding precedence entry is one. Averaging over total orders gives the cycle inequality. □

Remark D.2 (Nonredundancy of the row lower envelope). For $1 \le r \le n-2$, the row lower-envelope inequality of Lemma 3.3 cannot be obtained as a linear consequence solely of the complement equations and the directed-cycle inequalities. For $r = n-1$, the inequality holds with equality by the complement equations alone, so only $1 \le r \le n-2$ is at issue. Indeed, for $i \ne j$ set

$$q^{(u)}_{ji} = \frac{1}{2} + u_i - u_j.$$

These values satisfy complementarity, and on every directed cycle of length ℓ the potential differences telescope, giving

$$\sum_{t=1}^{\ell} q^{(u)}_{v_t v_{t+1}} = \frac{\ell}{2} \ge 1.$$

Sufficiently large potential differences drive the sum of the smallest row terms below the row lower-envelope bound. This construction, however, leaves the box $[0,1]$. With the box constraints added, the two endpoints do become derivable. For $r = 1$, selecting one minimum entry per row yields a functional digraph containing a directed cycle. The cycle inequality alone then gives $\sum_j L_{j,1} \ge 1$. For $r = n-2$, selecting one maximum entry per row and applying the reversed-cycle inequality together with the box bounds gives $\sum_j \max_i q_{ji} \le n-1$. Subtracting this from the total off-diagonal mass gives the required bound, since that mass is $n(n-1)/2$. Redundancy for $2 \le r \le n-3$ relative to the box-constrained relaxation is left undetermined. We therefore claim the row lower envelope only as a tool for its learning application, not as a new polyhedral fact.

Theorem D.3 (Asymptotic extremal constants)**.** *For order mixtures under target reselection and proper queries that may be version-inconsistent,*

$$\sqrt{\frac{\pi}{2}} \le \liminf_{n\to\infty} \frac{M_n}{\sqrt{n}} \le \limsup_{n\to\infty} \frac{M_n}{\sqrt{n}} \le C_0,$$

where

$$C_0 = \inf_{0<\theta<1} \frac{\sqrt{2\log(1/(1-\theta))}}{\theta} = 2.216035867\ldots.$$

If $y_* > 1$ is the root of

$$y - 1 = 2\log y,$$

then

$$y_* \approx 3.512862, \qquad \theta_* = 1 - \frac{1}{y_*} \approx 0.715332, \qquad C_0 = \frac{y_*}{\sqrt{y_* - 1}}.$$

Proof. The cyclic mixture and Theorem C.1 give the lower bound.

For the upper bound, let

$$T = T(Q) = \min_{j,r} \left(r + \frac{1}{\alpha_{j,r}} \right).$$

If $\ell < T$, then for every row j,

$$\alpha_{j,\ell} \le \frac{1}{T-\ell};$$

otherwise the pair (j, ℓ) would give a smaller value than T. Combining this with Lemma 3.3 yields, for $r < T$,

$$\frac{r(r+1)}{2} \le n \sum_{\ell=1}^{r} \frac{1}{T-\ell}.$$

Take a subsequence along which $T/\sqrt{n} \to C$. If $C = 0$ the bound is immediate. Assume therefore that $C > 0$, so $T \to \infty$ along the subsequence. The elementary blocking bound gives $T = O(\sqrt{n})$, so for each fixed $0 < \theta < 1$, $r = \lfloor \theta T \rfloor$ is at most $n-1$ for large n. Dividing the preceding inequality by n and taking Riemann-sum limits gives

$$\frac{\theta^2 C^2}{2} \le \log \frac{1}{1-\theta}.$$

The Riemann-sum error is $O_\theta(T^{-1})$. Hence

$$C \le \frac{\sqrt{2\log(1/(1-\theta))}}{\theta}$$

for every θ, and minimizing proves the upper bound. Elementary differentiation gives the displayed equation for y_*. □

Thus the cyclic leading constant is within a factor

$$\frac{C_0}{\sqrt{\pi/2}} \approx 1.768$$

of the order-mixture extremum.

Proposition D.4 (Small extremal values and atoms)**.** *For the same protocol,*

$$M_2 = 2, \qquad M_3 = \frac{8}{3}.$$

The three-point cyclic mixture is exactly extremal. If an order distribution has an atom of mass $p > 0$, then

$$\mathrm{Esc}(Q) \le 1 + \frac{1}{p}.$$

In particular, every mixture of at most two total orders has value at most 3*.*

Proof. For $n = 3$, every two-concept continuation has value 2, and the root recurrence reduces to

$$\mathrm{Esc}(Q) = 3 - \max_j \alpha_j$$

for the appropriate row minimum α_j. The $r = 1$ case of Lemma 3.3 gives $\max_j \alpha_j \ge 1/3$, and the cyclic mixture attains equality.

If one total order π_0 has mass p, query the π_0-last concept, and repeat it as a proper probe after it is eliminated. Whenever the drawn order is π_0, which happens with probability p independently each round, the first disagreement with any other target is that target's positive point. This yields a geometric contribution at most $1/p$ and one final query. A two-order mixture has an atom of mass at least $1/2$. □

Exact-arithmetic searches over several structured families found no order mixture exceeding the cyclic value for the tested $n \le 8$.[2]

[2]These searches included translation orbits, weighted rotations, and local perturbations. They are not a proof of global cyclic extremality.

E SINGLE-ORDER, MULTICLASS, AND TRANSDUCTIVE PROOFS

E.1 DEFERRED PROOFS FROM SECTION 5

Proof of Theorem 5.2. Let $h = P_\pi(s_V, V)$ and suppose the query is erroneous. The target $g \in V$ returns its label at the first point x where g and h differ. By Lemma 5.1,

$$g(x) \neq s_V(x).$$

The standard low-dimensional-branch lemma bounds the Littlestone dimension of every non-SOA branch by $\mathrm{Ldim}(V) - 1$. Indeed, if two distinct labels both retained dimension $\mathrm{Ldim}(V)$, attaching them above two depth-$\mathrm{Ldim}(V)$ trees would give a deeper tree. Hence every error strictly decreases the dimension. There are at most d errors, followed by the final accepted query.

For the complete class, the environment maintains a full subcube. If the learner's query violates an already fixed coordinate, the earliest such violation gives a legal first-disagreement response without shrinking the subcube. Otherwise the environment chooses a target that agrees with the query before the earliest free coordinate and differs there, fixing only that coordinate. This forces d errors before the final accepted query. The construction allows the target to be reselected and covers all proper class-internal queries. □

Proof of Proposition 5.3. Suppose the compiled learner receives r consecutive erroneous responses. The resulting version space V_{r+1} is nonempty. Choose any $c \in V_{r+1}$. Every returned label equals $c(x_t)$, and Lemma 5.1 shows that it differs from the unprojected query $q_t(x_t)$. Thus the entire length-r transcript is a legal arbitrary-counterexample transcript for the single fixed target c. The premise implies $r \leq T - 1$, after which the final query terminates. □

To see why the arbitrary-counterexample premise is necessary in Remark 5.4, let $\mathcal{H} = \{00, 01, 11\}$ with $x_1 \prec x_2$. Consider a learner that is allowed improper queries but makes only proper queries here. It queries 00 first, then 11 after the response $(x_1, 1)$ and 01 after the response $(x_2, 1)$. Against every fixed target and the first-disagreement oracle it uses at most two queries. Its lexicographically projected simulation can be forced under target reselection to receive errors first at x_2 and then at x_1. The round-one target 01 answers the query 00 with $(x_2, 1)$. The environment then reselects the target 11, which is consistent with the recorded label and answers the projected query 01 with $(x_1, 1)$. Only the third query is accepted. The resulting transcript is not a first-disagreement transcript for any one fixed target. Against 11, the query 00 would first disagree at x_1. Thus first-disagreement-only guarantees do not satisfy the premise of Proposition 5.3, and the compiled learner needs three queries.

E.2 THE FIXED-ORDER GAME

For binary labels and current version space V, the set of possible first-disagreement points after querying h is

$$X_\pi(h, V) = \{x_i : \exists g \in V,\ g(x_s) = h(x_s)\ (s < i),\ g(x_i) \neq h(x_i)\}.$$

If $h \in V$, every erroneous response eliminates h. If $h \notin V$, a response can leave V unchanged. Such a self-looping first action is strictly dominated in the minimax recursion. Choosing $h \in V$ always provides an action with no self-loop. Induction on $|V|$ therefore gives the frontier recursion in Proposition 5.5 and the equality $\mathrm{GV} = \mathrm{Fdim}_\pi + 1$.

For the multiclass action-pair formulation, let $E(V)$ denote the minimax number of erroneous queries from V. An action containing a self-loop cannot improve $E(V)$. Choosing $h \in V$ always gives an action with no self-loop. Induction on $|V|$ therefore gives

$$V \in \mathcal{W}_k \quad \Longleftrightarrow \quad E(V) \leq k.$$

Indeed, every no-self-loop successor $V_{x,y}$ is smaller than V, so the induction hypothesis converts the universal successor condition in the definition of $\mathcal{W}_{k+1}$ exactly into the minimax recurrence for E. Thus the binary proof is unchanged after replacing first-disagreement points by action pairs, and $\mathrm{GV} = \mathrm{Fdim}_\pi + 1$ follows for finite multiclass labels.

For an N-member threshold class, an inconsistent threshold query permits a self-loop and is not useful. The remaining interval game satisfies the balanced recurrence

$$R(N) = 1 + \min_k \max\{R(k), R(N-1-k)\},$$

for erroneous queries. Here $R(1) = 0$, and empty sides are omitted from the maximum. This yields $R(N) = \lfloor \log_2 N \rfloor$ and $\mathrm{GV} = R(N) + 1$. The same balanced-tree argument gives $\mathrm{Ldim} = \lfloor \log_2 N \rfloor$.

For every $d \geq 1$, the label-cube example may be realized by hypotheses $\{h_\star\} \cup \{h_u : u \in \{0,1\}^d\}$. On cube coordinates, h_u realizes u. On a private label point z_v, $h_u(z_v) = 1$ exactly when $u = v$. The hypothesis $h_\star$ is zero everywhere. Place all label points before the cube coordinates. Querying $h_\star$ either terminates or reveals the target's private label point, after which one final query suffices.

The label-cube construction just described realizes all 2^d binary patterns on the cube coordinates, so $\mathrm{Ldim} \geq d$. Its size is $2^d + 1$, which precludes a complete binary mistake tree of depth $d + 1$. The private label-point query gives $\mathrm{GV} = 2$.

E.3 Multiclass complete classes

For $\mathcal{Y}^d$, no multiclass Littlestone tree can repeat a coordinate along a root-to-leaf path, because all coordinates assigned on that path are fixed. Hence $\mathrm{Ldim}(\mathcal{Y}^d) \leq d$, while querying the d coordinates in a complete binary subtree gives the reverse inequality. The lower-bound environment in Theorem 5.2 forces exactly one new coordinate per progress-making error, so the multiclass value is $d + 1$.

Remark E.1 (Multiclass frontier actions). For a finite multiclass label set, the frontier dimension in Theorem 5.6 is understood through the action pairs

$$\mathcal{A}_\pi(h, V) = \{(x, y) : y \neq h(x),\ \exists g \in V,\ g \equiv h \text{ before } x,\ g(x) = y\}.$$

Starting with $\mathcal{W}_0 = \{V : |V| = 1\}$ and writing $V_{x,y} = \{g \in V : g(x) = y\}$, define

$$\mathcal{W}_{k+1} = \mathcal{W}_k \cup \{V : \exists h,\ \forall (x, y) \in \mathcal{A}_\pi(h, V) : V_{x,y} \in \mathcal{W}_k\}, \qquad \mathrm{Fdim}_\pi(V) = \min\{k : V \in \mathcal{W}_k\}.$$

An action containing a self-loop is unavailable. If $h \in V$, no action pair is a self-loop. With this action-pair formulation, the binary induction proving $\mathrm{GV} = \mathrm{Fdim}_\pi + 1$ extends verbatim.

E.4 Embedding a transductive adversary

Proof of Theorem 5.6. Let S be an optimal transductive adversary. The EQ environment maintains a private revealed prefix

$$y_{1:r}$$

and its nonempty consistency set W_r, with the invariant $W_r \subseteq V$.

If the learner's query h violates the private prefix, return the earliest violating point with its prefix label. This legal first-disagreement response repairs the query without advancing the transductive clock. It counts as an EQ error.

If h agrees with the prefix, continue the internal transductive game, using $h(x)$ as the induced prediction, until S forces the next transductive mistake. Return that point and label as the first disagreement and extend the private prefix. If the entire remaining sequence agrees with h, then

$$h \in W_n \subseteq V,$$

so the environment can legally lock target h and accept the query.

Every transductive mistake is injectively associated with an erroneous EQ query. Therefore a deterministic EQ strategy with e errors would induce a transductive learner with at most e mistakes, proving

$$M_{\mathrm{td}} + 1 \leq \mathrm{GV}\,.$$

The other inequalities follow from the frontier recursion and Theorem 5.2.

For $\mathcal{H}_1 = \{000, 001, 010, 101\}$, direct recursion gives $M_{\mathrm{td}} = 1$, $\mathrm{Ldim} = 2$, and $\mathrm{GV} = 3$. In the serial product $\mathcal{H}_k = \mathcal{H}_1^k$, the adversary always works on the earliest unfinished block. A query

violating an already finished block is answered at its earliest violating point, without advancing any other block. For a query consistent with the finished blocks, the adversary applies the one-block lower strategy inside the current block. This gives $M_{\mathrm{td}} \geq k$ and $\mathrm{GV} \geq 2k+1$. Concatenating the optimal block strategies gives the matching upper bounds, while the product tree has depth $2k$ and no greater depth. Thus

$$(M_{\mathrm{td}}, \mathrm{Ldim}, \mathrm{GV}) = (k, 2k, 2k+1).$$

With $G = \mathcal{H}_1$, use the convention $\mathcal{H}_0 = G^0$ for the one-concept class and pad every product by coordinates that are identically zero. This gives

$$\Delta_n \geq \left\lfloor \frac{n}{3} \right\rfloor + 1 \qquad (n \geq 1).$$

For the upper bound, if $M_{\mathrm{td}} \geq 1$, then

$$\mathrm{GV} - M_{\mathrm{td}} \leq (n+1) - 1 = n.$$

If $M_{\mathrm{td}} = 0$, the whole class has the same labels on the ordered sequence. Any two concepts then have no disagreement point and hence are equal, so $|\mathcal{H}| = 1$, $\mathrm{GV} = 1$, and $\mathrm{GV} - M_{\mathrm{td}} = 1 \leq n$. □

Exact dynamic programs on all small binary classes and on ternary projection instances found no violation of the stated inequalities. These checks are not used in the proof.

F Fixed-target proofs, Bellman equations, and the general potential

F.1 The integer program and proof of Proposition 4.1

The half-decay number has the exact integer-program representation

$$\kappa_{1/2}(Q) = \min \left\{ \sum_j c_j : c_j \in \mathbb{Z}_{\geq 0},\ \sum_j b_{ji} c_j \geq 1 \ \ \forall i \right\},$$

where

$$b_{ji} = \min\{1, -\log_2(1 - q_{ji})\},$$

with the value 1 when $q_{ji} = 1$. The integer parameter is needed here. Its naive fractional relaxation is insufficient: symmetric matrices can have constant fractional value but logarithmic deterministic complexity.

Proof of Proposition 4.1 and the integer-program representation. Before the positive target point is returned, every nonterminating response is the negative label at the current query point. Thus a deterministic learner follows one fixed sequence σ. Against target h_i, the probability of reaching its tth spine query is $R_i^\sigma(t-1)$. The query is erroneous exactly when $j_t \neq i$. Adding the final accepted query gives the displayed formula for $V_i(\sigma)$ and hence the infimum characterization.

Termwise,

$$\mathbf{1}\{j_t \neq i\} R_i^\sigma(t-1) \geq R_i^\sigma(t),$$

because $a_{ii} = 0$, while $\mathbf{1}\{j_t \neq i\} R_i^\sigma(t-1) \leq R_i^\sigma(t-1)$. Summing the two bounds gives $A_i(\sigma) \leq V_i(\sigma) \leq A_i(\sigma) + 1$.

Put $K = \kappa_\eta(Q)$. The first $K-1$ terms of any spine fail to η-dominate, so for some i, $R_i^\sigma(K-1) > \eta$. Monotonicity of the survival products then gives

$$A_i(\sigma) \geq \sum_{t=0}^{K-1} R_i^\sigma(t) > K\eta,$$

which proves the lower bound. If $K = +\infty$, apply the same argument to every finite prefix. Conversely, repeat a witnessing block of length K. Each completed block multiplies every survival product by at most η, and therefore

$$A_i(\sigma) \leq K \sum_{\ell \geq 0} \eta^\ell = \frac{K}{1-\eta}.$$

The comparison with V_i gives the upper bound.

Finally, a multiset with multiplicities c_j has survival product at most $1/2$ against every i exactly when

$$\sum_j c_j\big(-\log_2(1-q_{ji})\big) \geq 1 \qquad \text{for every } i.$$

Truncating each nonnegative coefficient at 1 does not change this threshold condition, including the case $q_{ji} = 1$. Thus $b_{ji} = \min\{1, -\log_2(1-q_{ji})\}$ gives exactly the displayed integer program. □

F.2 Proof of the likelihood-halving bound

Proof of Theorem 4.2. Along the all-negative spine, maintain unnormalized likelihood weights r_i and let $W = \sum_i r_i$. If j is queried and a negative counterexample is observed, the updated total weight is

$$W'(j) = \sum_i r_i(1-q_{ji}).$$

Using symmetry and sampling j proportionally to r_j only as an averaging argument,

$$\mathbb{E}_{j\sim r/W}[W'(j)] = \frac{W^2 - \sum_i r_i^2}{2W} \leq \frac{W}{2}.$$

Hence some positive-weight deterministic query halves W. Repeating this choice gives

$$W_t \leq n2^{-t}.$$

For each fixed target, its survival probability after t negative steps is at most $\min\{1, n2^{-t}\}$. Summing these tail probabilities and including the final query gives at most $\lceil \log_2 n \rceil + 3$. □

F.3 Proof of Theorem 4.3

Proof of Theorem 4.3. Query the vertices of a minimum dominating set. If the target lies in the set, its own query terminates. Otherwise one of these queries returns the positive target point, followed by one final accepted query. This proves the upper bound.

Conversely, suppose the worst deterministic running time is M. The set of hypotheses queried in the all-negative prefix of length M must dominate every target. Otherwise an undominated target would survive all M rounds, contradicting the definition of M. Thus $\delta(T) \leq M$.

The classical average-outdegree induction gives $\delta(T) \leq \lfloor \log_2(n+1) \rfloor$. Erdős's random-tournament bound (Erdős, 1963) gives the asymptotic lower bound, while Theorem 4.2 gives the matching leading upper bound. □

These logarithmic witnesses are nontransitive and are not order mixtures. A $0/1$ precedence matrix represented by a distribution over orders would force every supported order to agree on every pair. Such a matrix would therefore be transitive.

F.4 Bellman formulation under target reselection

Proposition F.1 (Finite-state Bellman value)**.** *Let a known counterexample generator be Markovian in the sense that, given the current version space, query, and target, its distribution does not otherwise depend on history. Under target reselection and deterministic proper queries,*

$$\mathfrak{D}_G(V) = \min_{h\in\mathcal{H}} \max_{g\in V} \begin{cases} 1, & h = g, \\ 1 + \mathbb{E}_Z \mathfrak{D}_G(V \mid Z), & h \neq g, \end{cases}$$

is exactly the worst-case expected query complexity.

For a singleton class and a history-invariant two-point matrix Q, consider the environment that, after each query h_j with $|V| \geq 2$, selects a target minimizing q_{ji} over $i \in V \setminus \{j\}$. The minimization is over all of V when $j \notin V$. A one-concept version space is terminal. This environment is globally optimal, and the resulting scalar recursion is exactly $\mathrm{Esc}(Q)$. For a history-dependent generator, the full public history must be included in the Bellman state.

Proof. This is a finite stochastic shortest-path dynamic program. Querying a surviving concept ensures that every erroneous response eliminates that concept. A strategy of value at most $|V|$ therefore exists, excluding a pathological self-loop solution. Standard backward induction over version spaces then gives the least Bellman solution and the minimax value.

For singleton classes, a positive response leaves a singleton and a negative response has a continuation state independent of which target was selected. Selecting the current query itself as target accepts at cost 1. This cost is minimal, so a cost-maximizing environment with $|V| \geq 2$ never does so. Among the remaining targets the action value is monotone in q_{ji}, so choosing one with minimum positive-response probability is optimal on every round. □

Four useful calibrations are:

Q	$\mathrm{Esc}(Q)$
$q_{ji} \equiv 0$ (nonsymmetric)	n
$q_{ji} \equiv 1/2$	$3 - 2^{2-n}$
one fixed known total order	$1\ (n = 1),\ 2\ (n \geq 2)$
uniform cyclic rotations	D_n.

F.5 Proof of the general fixed-target logarithmic bound

The posterior symmetry-pairing step extends the elimination argument of Bhatia (2021, Lemmas 20, 22, and 23).

Proof of Theorem 4.4. Let $n = |\mathcal{H}|$. At a depth-t public history y, let V_y be the version space. For each $c \in V_y$, let $\ell_y(c)$ be the likelihood of that history under fixed target c, and put

$$W_y = \sum_{c \in V_y} \ell_y(c).$$

Define the global potential

$$\Phi_t = \sum_{|y|=t} |V_y| W_y.$$

Initially $\Phi_0 = n^2$.

For a candidate query h at node y, group the child histories by returned labeled counterexample. Writing $V_{x,y'} = \{g \in V_y : g(x) = y'\}$ for the label sections, the children of y under query h contribute the child potential

$$C_y(h) = \sum_{\substack{c \in V_y \\ c \neq h}} \ell_y(c) \sum_x G(x \mid h, c, y) \left|V_{x,c(x)}\right|,$$

with the inner sum interpreted as an integral in the continuous case. Target c reaches the child labeled $(x, c(x))$ with probability $G(x \mid h, c, y)$, carries its likelihood into that child, and the child's version space is the section $V_{x,c(x)}$. Average h with weights $\ell_y(h)/W_y$ and pair the ordered terms (h, c) and (c, h). Conditional symmetry $G(x \mid h, c, y) = G(x \mid c, h, y)$ gives

$$\mathbb{E}_{h \sim \ell/W} C_y(h) = \frac{1}{W_y} \sum_{\{h,c\}} \ell_y(h) \ell_y(c) \sum_x G(x \mid h, c, y) \Big(\left|V_{x,c(x)}\right| + \left|V_{x,h(x)}\right| \Big),$$

where the outer sum runs over unordered pairs of distinct concepts in V_y. Every x in the support of $G(\cdot \mid h, c, y)$ is a disagreement point of h and c, so the two sections are disjoint subsets of V_y and $|V_{x,c(x)}| + |V_{x,h(x)}| \leq |V_y|$. Hence

$$\mathbb{E}_h C_y(h) \leq \frac{|V_y|}{2W_y} \Big(W_y^2 - \sum_h \ell_y(h)^2 \Big) \leq \frac{|V_y| W_y}{2},$$

and some deterministic proper query satisfies the same inequality. If an erroneous hypothesis can be accepted early, the identity only becomes an inequality in the favorable direction. Selecting such a query independently at every node gives

$$\Phi_{t+1} \leq \frac{1}{2} \Phi_t, \qquad \Phi_t \leq n^2\, 2^{-t}.$$

For every fixed target c, the probability that at least t errors have occurred is bounded by

$$\Pr_c[N \geq t] \leq \min\{1, n^2\, 2^{-t}\}.$$

Summing the tail after $\lceil 2\log_2 n \rceil$ and including the final accepted query gives

$$\mathbb{E}_c Q \leq \lceil 2\log_2 n \rceil + 2.$$

For a continuous counterexample space, let P_c^t denote the sub-probability law, on depth-t histories, of the run against fixed target c restricted to the event of t errors, and take the dominating measure

$$\mu_t = \sum_{c \in \mathcal{H}} P_c^t.$$

Set $\ell_y(c) = \frac{dP_c^t}{d\mu_t}(y)$ and $\Phi_t = \int |V_y|\, d\mu_t(y)$. The child-potential identity and the symmetry pairing hold pointwise in the counterexample outcome and can be integrated. To choose a query achieving the averaged inequality measurably in y, fix an order on the finite class and take the first minimizer. Use an arbitrary consistent query on a μ_t-null set. The measurability assumptions of the theorem make every step well defined on a standard Borel space. □

F.6 Why the fractional domination relaxation is insufficient

Let $\kappa^*(Q)$ be the LP relaxation of the integer program for $\kappa_{1/2}$. The survival characterization gives

$$D^{\mathrm{fix}}(Q) \geq \frac{\kappa^*(Q)}{2}.$$

Poisson rounding with scale $\alpha = 4\ln(2n)$ and truncation $\theta = \ln\alpha$ gives

$$\kappa_{1/2}(Q) \leq \lceil 8\kappa^*(Q)\ln(2n) \rceil .$$

Indeed, draw independent multiplicities $X_j \sim \mathrm{Pois}(\alpha c_j^*)$ from an optimal fractional solution c^*. For a column i, using $0 \leq b_{ji} \leq 1$, $1 - e^{-\theta b} \geq b(1 - e^{-\theta})$, and the fractional feasibility $\sum_j b_{ji} c_j^* \geq 1$,

$$\Pr\Big[\sum_j b_{ji} X_j < 1\Big] \leq e^{\theta} \exp\big(-\alpha(1 - e^{-\theta})\big) = \alpha e^{1-\alpha},$$

and the union bound over the n columns is less than $1/2$ at $\alpha = 4\ln(2n)$. Also, $\Pr[\sum_j X_j > 2\alpha\kappa^*] \leq 1/2$ by Markov's inequality. Some realization therefore covers every column at cost at most $2\alpha\kappa^* = 8\kappa^*\ln(2n)$. Hence

$$D^{\mathrm{fix}}(Q) \leq 16\kappa^*(Q)\ln(2n) + 3.$$

For symmetric Q, an antisymmetric zero-sum-game argument gives $\kappa^*(Q) \leq 2$. Random tournaments nevertheless have $D^{\mathrm{fix}}(Q) = \Theta(\log n)$. Thus the logarithmic integrality gap is genuine, and a constant-factor characterization by the naive fractional domination number is false.

G Consistent-only details

G.1 Deferred proofs from Section 6

Proof of Theorem 6.1. Against a query j_t, the target that minimizes the probability of a positive counterexample is a remaining vertex attaining β_t^σ. A negative response deletes j_t, and the continuation state is independent of which minimizing target was chosen. Independence of the fresh order draws makes the survival probability the displayed product. Minimizing over deletion permutations proves the first formula.

For the graph identity, let

$$B_\sigma = \{j_t : \beta_t^\sigma < x\}.$$

Every directed cycle has a vertex that appears earliest in σ. Its outgoing cycle edge points to a later vertex, so that earliest vertex belongs to B_σ. Hence B_σ is a feedback vertex set.

Conversely, place a minimum feedback vertex set first, then list the remaining acyclic graph in reverse topological order. Every vertex outside the feedback set has no low-probability arc to a later vertex, so only feedback-set positions can have $\beta_t^\sigma < x$. This proves equality. Query the feedback-set prefix and then use the reverse topological suffix. The prefix contributes at most τ erroneous queries. The suffix has conditional positive-counterexample probability at least x on every round and hence contributes at most $1/x$ expected erroneous queries. The final accepted query adds one. □

Proof of Theorem 6.2. Every directed cycle in Γ_x has length greater than $1/x$, because the linear-order-polytope cycle inequality gives $\sum_{e\in C} q_e \geq 1$. Therefore the uniform assignment $z_v = x$ is feasible for the cycle LP and

$$\tau^*(\Gamma_x) \leq nx.$$

The approximation theorem of Even et al. (1998) gives a guarantee explicitly relative to the fractional optimum of the cycle LP. Corollary 6 and §4.4 of that paper yield

$$\tau(\Gamma_x) = O(nx \log n \log\log n)\,.$$

Combining this with Theorem 6.1 and choosing

$$x = \min\left\{\frac{1}{2}, \frac{1}{\sqrt{n \log n \log\log n}}\right\}$$

proves the upper bound.

For the lower bound, the cyclic dead-zone strategy is itself version-consistent, while Lemma 3.2 gives the same $\Omega(\sqrt{n})$ lower bound even against learners with the larger inconsistent action space. □

G.2 The nine-order product example

Let the concepts be indexed by $\mathbb{Z}_3^2$. For $s, t \in \mathbb{Z}_3$, let

$$s \prec_s s+1 \prec_s s+2$$

be the corresponding cyclic order on the first coordinate, and define the analogous order $\prec_t$ on the second coordinate. Let $\pi_{s,t}$ be the lexicographic order on $\mathbb{Z}_3^2$ using $\prec_s$ first and $\prec_t$ second. The generator samples one of the nine orders $\pi_{s,t}$ uniformly on every round.

For $j = (a, b)$ and $i = (c, d)$,

$$q_{ji} = \begin{cases} ((c-a) \bmod 3)/3, & a \neq c, \\ ((d-b) \bmod 3)/3, & a = c,\ b \neq d. \end{cases}$$

Thus every off-diagonal entry is $1/3$ or $2/3$.

Proposition G.1 (Exact product values)**.** *For this known independently resampled order mixture under target reselection,*

$$\mathrm{Esc} = \frac{100}{27}, \qquad E_{\mathrm{cons}} = \frac{8108}{2187}.$$

Proof. The graph of low arcs, those with $q_{ji} = 1/3$, is the lexicographic product $C_3[C_3]$. Its minimum feedback-vertex-set size is 5. An acyclic vertex set cannot meet all three outer blocks, because one vertex from each already forms an outer directed 3-cycle. It can contain at most two vertices from each inner directed 3-cycle. Thus every acyclic set has size at most 4, and this bound is attained by taking two vertices from each of two outer blocks.

Every vertex has four low out-neighbors, so each of the first four relevant consistent queries has hazard $1/3$. After four points are deleted, the remaining five points still contain a low-arc cycle. A $2/3$ hazard is exactly a sink in the remaining low-arc graph, and deleting a sink does not destroy an existing cycle. Consequently, of the next three relevant positions, at most two can have hazard $2/3$.

For numerical hazard sequences, moving a $2/3$ hazard ahead of an adjacent $1/3$ hazard can only decrease the survival area. The product after both positions is unchanged, but the intervening survival product acquires the smaller failure factor earlier. Hence the smallest possible hazard sequence is

$$\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{2}{3}, \frac{2}{3}, \frac{1}{3}.$$

It is realized by the permutation

$$(0, 1, 2, 6, 8, 7, 3, 4, 5),$$

where an element denotes $3 \cdot (\text{outer coordinate}) + (\text{inner coordinate})$. Its term-by-term exploratory area is

$$\begin{aligned} G_{\text{cons}} &= 1 + \frac{2}{3} + \frac{4}{9} + \frac{8}{27} + \frac{16}{81} + \frac{16}{243} + \frac{16}{729} + \frac{32}{2187} \\ &= \frac{5921}{2187}, \end{aligned}$$

and therefore

$$E_{\text{cons}} = 1 + G_{\text{cons}} = \frac{8108}{2187}.$$

For unrestricted proper queries, clear the four low out-neighbors of one fixed point and then repeatedly probe that point. The survival calculation gives

$$E = 1 + \frac{65}{27} + \frac{16}{81}\left(\frac{3}{2}\right) = \frac{100}{27}.$$

Conversely, every state whose induced low-arc graph contains a directed cycle has value at least $5/2$. To see this, a query outside the version space has a self-loop and gives value at least $5/2$ directly. A query in the version space with a low outgoing target gives at least

$$1 + \frac{1}{3} + \frac{2}{3} \cdot 2 = \frac{8}{3},$$

whereas a low-arc sink can be deleted without destroying the cycle and gives, inductively,

$$1 + \frac{2}{3} + \frac{1}{3} \cdot \frac{5}{2} = \frac{5}{2}.$$

During each of the first four progress steps, at least one of the queried point's four low out-neighbors survives, and after any four deletions a low-arc cycle remains. Applying the Bellman operation

$$T(x) = 1 + \frac{1}{3} + \frac{2}{3}x$$

backward therefore yields

$$T^4\left(\frac{5}{2}\right) = 1 + \frac{65}{27} + \frac{16}{81}\left(\frac{3}{2}\right) = \frac{100}{27}.$$

This matches the explicit strategy, and the exact difference is $8/2187$. □

G.3 Proof details for the FVS upper bound

For $x \le 1/2$, the graph Γ_x has no pair of opposite arcs, since symmetry would otherwise imply $q_{ji} + q_{ij} < 2x \le 1$. Every directed cycle has length strictly greater than $1/x$ by the cycle inequality. The cycle LP for directed FVS is

$$\min \sum_v z_v \quad \text{subject to} \quad \sum_{v \in C} z_v \ge 1 \text{ for every directed cycle } C, \qquad z_v \ge 0.$$

The uniform assignment $z_v = x$ is feasible and has cost nx. The weighted directed-FVS approximation theorem of Even et al. (1998) gives a guarantee explicitly relative to the fractional optimum of this cycle LP. By Corollary 6 and §4.4 of that paper, it returns an integral set of size

$$O(\tau^* \min\{\log \tau^* \log\log \tau^*, \log n \log\log n\}).$$

Here and in Theorem 6.2, all logarithms are truncated below at 1, so the expressions are meaningful for small parameters. An acyclic Γ_x has $\tau = 0$ and needs no approximation call. Substitution in Theorem 6.1 gives Theorem 6.2.

G.4 THE MULTISCALE OBSTRUCTION

Proposition G.2 (Limitation of a single-threshold lower-bound certificate)**.** *Let $Q \in \mathrm{LOP}_n$ and $0 < x \le 1/2$. If Γ_x has minimum outdegree at least $d \ge 1$, then $d + 1 < 2nx$. Consequently, a lower-bound argument that certifies $\beta_t < x$ on each of the first $T \ge 1$ steps of every trajectory must satisfy $T + 1 < 2nx$. Its single-threshold survival area is $(1 - (1 - x)^T)/x$. For this area to be $\Omega(T)$, one needs $Tx = O(1)$, which implies $T^2 = O(n)$. In every case the area is at most $\min\{T, 1/x\} \le \sqrt{T/x} < \sqrt{2n}$, so single-threshold survival certificates yield at most $O(\sqrt{n})$ lower bounds.*

If every vertex of Γ_x has at least $d \ge 1$ outgoing arcs, then its d smallest row entries are all below x. Therefore

$$L_{j,d} < dx$$

for every row. Lemma 3.3 gives

$$\frac{d(d+1)}{2} \le \sum_j L_{j,d} < ndx,$$

and division by $d \ge 1$ yields $d + 1 < 2nx$.

Now suppose every deletion permutation of $[n]$ satisfies $\beta_t^\sigma < x$ at each of its first $T \ge 1$ positions. Then every vertex of Γ_x has at least T outgoing arcs. To see this, suppose some vertex v had at most $T - 1$ low out-neighbors. List those out-neighbors first and v immediately after them, at some position $t \le T$. Every vertex listed after v then satisfies $q_{vj_s} \ge x$, so $\beta_t^\sigma \ge x$, a contradiction. Hence the minimum outdegree of Γ_x is at least T, and the first part applied with $d = T$ gives $T + 1 < 2nx$. For the area bound, each of the T terms of the survival sum is at most 1 and the full geometric sum is at most $1/x$, so the area is at most $\min\{T, 1/x\} \le \sqrt{T/x}$, and $T < 2nx$ gives $\sqrt{T/x} < \sqrt{2n}$. This proves Proposition G.2.

H OPEN PROBLEMS

O-N. Is the cyclic leading constant extremal among order mixtures? That is, does $M_n \sim \sqrt{\pi n/2}$ hold, and can the extremizing order distributions be classified?

O-M′. Can the factor $\sqrt{\log n \log\log n}$ be removed in the version-consistent-only order-mixture protocol? One route would be a specialized bound $\tau(\Gamma_x) = O(nx)$ near $x \asymp n^{-1/2}$. Another is a multiscale survival-area argument.

O-A′. Does the fixed-target cyclic value $D^{\mathrm{fix}}(C_n, G_{\mathrm{cyc}})$ converge as $n \to \infty$, and what is its exact limit?

O-I. Suppose one fixed order is used on every round but is hidden from the learner. On the singleton class this protocol degenerates. Draw the hidden order uniformly at random and let the environment always reselect the π-last surviving concept. A query of any other surviving concept h_j is answered $(j, 0)$ and eliminates only h_j. A query of an already-eliminated concept is also answered $(j, 0)$. This is because the π-largest survivor is last in every survivor set that contains it and hence is never eliminated. Conditionally on any all-negative history, the relative order of the surviving concepts remains exchangeable, so the π-last survivor is uniform among them. Any query, randomized or not, is accepted and terminates the run with probability at most $1/|V|$. Averaging over the hidden order, every learner needs at least $(n + 1)/2$ expected queries. Thus, for each learner, some fixed order yields expected cost at least $(n + 1)/2$. Querying uniformly in the current version space matches this bound by the induction $F_m \le 1 + (bF_{m-1} + a)/m \le (m + 1)/2$, where the current target has b surviving predecessors and a successors. The randomized minimax value of this protocol is therefore exactly $(n + 1)/2$. Against a deterministic learner, the hidden order can be chosen so that its queries always sit early, forcing n. Deterministic and randomized learners alike thus need $\Theta(n)$ queries. There is no separation of asymptotic order, only the constant-factor gap between n and $(n + 1)/2$. The projection lemma requires knowledge of the scanning order. For which general classes does a hidden fixed order admit query complexity below the trivial version-space bound, and can hiding the order separate proper from improper, or deterministic from randomized, learning?

O-F. For a fixed target and a known independently resampled order mixture on C_n, is there a deterministic proper learner with $O(1)$ expected queries uniformly over all mixtures—and, for general finite classes, a dimension-only bound under order-mixture generators? Theorem 4.2 gives $\lceil \log_2 n \rceil + 3$ for all symmetric two-point generators, and Theorem C.5 gives the uniform constant $9/2$ for monotone translation kernels. For general order mixtures, neither an $\omega(1)$ lower bound nor a uniform constant upper bound is known. The logarithmic witnesses of Theorem 4.3 are not order mixtures. This is the only cell of Table 1 whose asymptotic order on singleton classes remains undetermined.

We previously asked here whether one deterministic rule can use $O(\log |\mathcal{H}|)$ queries for a fixed target and an *unknown* online symmetric generator. In fact, the answer is negative. Theorem 4.4 is necessarily nonuniform in the generator:

Proposition H.1 (Generator knowledge is indispensable)**.** *Every deterministic proper learner A that does not depend on the generator satisfies*

$$\sup_{\pi,\, i} \mathbb{E} Q_A(C_n, G_\pi, h_i) \;\geq\; n,$$

the supremum ranging over fixed total orders π and fixed targets h_i, where G_π is the symmetric, history-independent first-disagreement generator of π. Enumerating the concepts attains n, so the optimal generator-independent fixed-target value on C_n is exactly n. This contrasts with the $\lceil \log_2 n \rceil + 3$ of Theorem 4.2 available when the generator is known.

Proof. Simulate A for $n-1$ steps, answering every query h_j with the negative counterexample $(j, 0)$. These queries name at most $n-1$ distinct concepts, so some concept h_i is never queried. Fix the target h_i and a total order π that places i last. Every simulated query h_j has $j \neq i$. The two concepts disagree exactly on $\{i, j\}$, and j precedes i in π, so G_π returns $(j, 0)$. The real run against (G_π, h_i) therefore reproduces the simulated transcript for $n-1$ queries, none of which is accepted. At least one further query is needed, so $Q_A \geq n$. For the upper bound, query the concepts in any fixed order. A negative response eliminates the queried concept. A positive response reveals the target, which is queried next. Thus at most n queries always suffice. □

I Extended related work

Symmetric equivalence queries. Braverman et al. (2026) prove the randomized $O(\mathrm{Ldim})$ upper bound and an $\Omega(\mathrm{Ldim})$ converse for suitable symmetric adversaries. Their Section 3.1.1 explains why the general proof produces a randomized learner and gives a deterministic rule in the random-counterexample special case. Section 5 of their COLT 2026 paper asks whether randomization is inherent for general symmetric and order-induced adversaries. We resolve the dimension-only question for general symmetric generators even with a fixed target (Theorem 4.3), and for independently resampled order mixtures under target reselection (Theorem 3.1). The fixed-target order-mixture case remains open (O-F). Our answers separate four issues: target reselection versus a fixed target, a fixed known order versus independently resampled orders, unrestricted proper versus version-consistent proper queries, and known versus unknown generators.

Angluin and Dohrn (2017) study equivalence queries with *distributional* random counterexamples. Their Max-Min algorithm identifies the target in at most $\log_2 n$ expected queries in that model. Their model is not equivalent to our cyclic-start generator. Moreover, the randomness there comes from the *teacher's* side, through the counterexample oracle. Our separations concern *learner-side* randomization against a fixed random generator. We compare deterministic and randomized learners under the same symmetric oracle. We do not claim the broad usefulness of random counterexamples as new. Within the teacher-random model itself, Bhatia (2021) gives simple deterministic and randomized rules with $O(\log |\mathcal{H}|)$ asymptotically optimal average learning time, and Chase et al. (2023) extend the model's guarantees through Littlestone dimension, including to infinite classes. These are the closest deterministic precursors we know of. They operate in the distribution-conditioned random-counterexample special case and do not address deterministic learning against general symmetric or order-mixture generators under target reselection, the setting of our lower bounds. A three-point check makes the model gap concrete. On singletons, a distribution-conditioned random counterexample with base weights w yields $q_{ji} = w_i/(w_i + w_j)$. Thus the odds ratios q_{ji}/q_{ij} multiply to 1 around every directed triangle. The three-concept cyclic mixture has all three forward probabilities

equal to $1/3$, with odds-ratio product $1/8$. The cyclic generator therefore lies outside that special case. For the cyclic singleton construction, our claims are its exact value, its $\sqrt{n}$ separation at Littlestone dimension one, and its lower bound matching the order-mixture upper bound up to constants under target reselection.

Closely related recent work. Alon et al. (2026) (arXiv, August 21, 2026) study equivalence queries for learning an unknown linear order. Counterexamples are incorrectly ordered pairs, and up to k of them may be untruthful. They determine the optimal complexity $\Theta(n \log n + nk)$. Their setting has no target reselection, no symmetric counterexample generators, and no randomized-versus-deterministic separation. The two sets of results are disjoint. Dewasurendra (2026) studies list PAC learning and fixed-design list transduction. It characterizes the optimal randomized risk as μ_ℓ/m and the optimal deterministic risk as $\lceil \mu_\ell \rceil/m$, and separates the two by a sharp factor $\ell + 1$ on a one-coordinate class. That randomized-versus-deterministic separation concerns transductive risk in list prediction, not equivalence-query complexity, and neither set of results contains the other.

Li (2026) proves an exponential deterministic–randomized gap in a different oracle model: transductive online prediction of thresholds on an unknown total order of T points. The learner never proposes hypotheses. It calls a memoryless consistency-type ERM oracle that returns the minimal or maximal consistent prefix. The cost is mistakes plus oracle calls. Every deterministic learner pays $M + Q \geq T - \varepsilon$ against a fixed target, while a randomized learner needs only $\Theta(\log T)$ expected calls. The hardness comes from the hidden order and the oracle's selection rule. Our lower bounds instead concern equivalence queries against known symmetric or order-mixture generators and, in the linear and $\sqrt{n}$ cases, require target reselection. With a fixed target, the same kernels are learnable in $O(1)$ queries. We are not aware of a reduction between the two models in either direction, and the existing theorems do not contain each other. Both nevertheless exhibit the theme that the oracle's tie-breaking rule, not the class alone, governs deterministic complexity.

Relative to all of these, the contributions claimed here remain exactly three: (i) deterministic lower bounds against symmetric and order-mixture generators under target reselection, with the extremal $\Theta(\sqrt{n})$ theorem on singleton classes; (ii) the fixed-known-order optimal proper $d + 1$ theorem via lexicographic projection; and (iii) the tight logarithmic fixed-target theory on singleton classes and a general logarithmic upper bound.

Classical EQ and Littlestone dimension. The improper query bound $\mathrm{Ldim} + 1$ follows from classical work of Angluin and Littlestone (Angluin, 1988; Littlestone, 1988). The fixed-order contribution here is the lexicographic projection mechanism and the resulting optimal version-consistent proper theorem. The upper bound also follows by specializing the deterministic score argument of Braverman et al. (2026, Section 3.1.1) to geometrically decreasing weights along the known order. We found no prior occurrence of this projection in the equivalence-query literature, but make no claim beyond the literature search described by the cited scope.

Transductive learning. The transductive model goes back to Ben-David et al. (1997). Block-product constructions also appear in work comparing transductive and online dimensions, including Chase et al. (2025). Our comparison fixes the same order and contrasts target-reselecting first-disagreement EQ with transductive prediction. This differs from their worst-order comparison.

Tournament domination. The domination-number upper bound and the logarithmic lower-bound tournaments are classical, due respectively to elementary average-outdegree arguments and work of Erdős (Erdős, 1963). Explicit constructions are associated with Graham and Spencer (1971). Our contribution is the reduction from fixed-target symmetric equivalence queries to tournament domination.

Linear-order polytopes and feedback vertex sets. Directed-cycle inequalities belong to the classical linear-order polytope literature, including work of Grötschel et al. (1985). To the best of our knowledge, the row lower-envelope inequality and its blocking-point learning application have not appeared in this form. We claim it only as a tool for the learning application, with the derivable endpoints recorded in Remark D.2. The consistent-only upper bound uses the directed-FVS approximation theorem of Even et al. (1998). Constant-factor results for special graph families do not apply without a proof that the threshold graphs Γ_x belong to those families.

Test ordering and survival area. Expressions of the form $\sum_t \prod_{s<t}(1 - r_s)$ occur in optimal test-ordering problems with fixed rejection probabilities (Berend et al., 2014). Our hazards are different: each β_t is endogenously determined by the remaining suffix of a permutation. We claim no independent novelty for survival-area expressions as such. Our contribution is their endogenous, suffix-determined use in the consistent-only protocol.