# THE PRICE OF ADAPTIVITY IN BANDIT MULTICLASS CLASSIFICATION: MINIMAL SEPARATION AND SHARP TOWER ASYMPTOTICS

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

We resolve the existence component of a foundational open problem posed at NeurIPS 2024 concerning the price of adversarial adaptivity in bandit multiclass classification. The optimal expected mistake bound against an *adaptive* adversary is at most a multiplicative $O(k)$ above the oblivious bound, where $k = |Y|$. We prove that genuine *concept classes* $H \subseteq Y^X$ can incur a strict adaptive–oblivious gap. First, a four-function class on two points with $k = 3$ has adaptive value $35/29$ and oblivious value $7/6$. No class of at most three functions separates, so four is optimal. Second, for the multi-shield towers $H_r = \{0,1\}^{[r]} \cup \{\bar{2}, \ldots, \overline{r+1}\}$, we prove a closed-form recursion for the adaptive value of every reachable state. We obtain $\mathrm{A}(H_r) = \frac{5r}{8} + O(\sqrt{r})$ and $\mathrm{B}(H_r) = \frac{r}{2} + \log_2 r - \log_2 \log_2 r + O(1)$, hence $\lim_{r\to\infty} \mathrm{A}(H_r)/\mathrm{B}(H_r) = \frac{5}{4}$ exactly. Third, our capping theory bounds natural amplification constructions. For atoms whose adaptivity ratios are uniformly bounded by $C_0$, unions with disjoint label blocks obey $\mathrm{A} \le (C_0 + 2)\mathrm{B}$ at any finite nesting depth. An exact Hamming-ball formula shows that the additive localization surcharge already reaches $\Theta(\sqrt{n})$ for two overlapping layers. Every exact value has a certificate: primal–dual certificates at each game state for adaptive play, and explicit priors and sequences for oblivious play. The included standard-library script re-checks them all. We also establish necessary conditions for linear $\Omega(k)$ separation: $q, \mathrm{B}, \mathrm{F} = \Omega(k)$, and the concept class must have at least exponential size.

## 1 INTRODUCTION

An open problem posed at NeurIPS 2024 asks how much adversarial adaptivity can increase the cost of learning genuine concept classes under bandit feedback (Filmus et al., 2024). We resolve its existence component by proving a strict gap. This question matters for systems that observe the outcome of a selected action: contextual-bandit recommendation collects feedback for the displayed item (Li et al., 2010), while reasoning systems can check generated answers or programs (DeepSeek-AI, 2025).

Online multiclass classification in the mistake-bound model (Littlestone, 1988; Daniely et al., 2015) becomes more demanding under *bandit feedback*. In each round, the learner sees an instance $x_t$, commits to a randomized prediction, and learns only *whether* it was correct (Auer et al., 2002; Kakade et al., 2008). In systems that learn from acceptance or rejection, which guarantees survive when the environment can respond to the learner? In *Bandit-Feedback Online Multiclass Classification: Variants and Tradeoffs*, published at NeurIPS 2024, Filmus et al. (2024) compared the costs of bandit feedback, adversarial adaptivity, and deterministic prediction. Their Theorem 1.2 shows that, for every class, $\mathrm{opt}^{\mathrm{adap}}_{\mathrm{bandit}} = O(\mathrm{opt}^{\mathrm{obl}}_{\mathrm{bandit}} \cdot k \log k)$ with $k = |Y|$. In fact, $O(\mathrm{opt}^{\mathrm{obl}}_{\mathrm{bandit}} \cdot k)$ follows from their Theorem D.1 and $\mathrm{F} \le \mathrm{B}$ (Open Problem 6.1). They also construct *pattern classes*, downward-closed sets of finite interaction patterns, with an $\Omega(k)$ multiplicative gap. This is not a concept class: its hypotheses constrain finitely many interactions, rather than defining functions fixed in advance on the whole domain. Their Section 4.1 asks how much of the linear price can be realized by genuine concept classes $H \subseteq Y^X$, following the distinction in Remark 1.3. We settle the existence component by exhibiting a strict gap.



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.

The obstacle is global consistency. An adaptive adversary does not commit to a target, but every feedback path must remain consistent with *some* hypothesis that is a total function on the domain. Every local trap therefore needs a class witness. An oblivious adversary can commit to that same witness in advance. The challenge is to show that adaptation buys something prior commitment cannot reproduce. Natural concept-class transplants of the pattern-class construction, including class shields on private labels and label-reuse covers, face the caps in Section 5.1.

**Contributions.** We settle the *existence* component of this NeurIPS 2024 open problem. The contributions concern minimal separation, sharp tower asymptotics, and limits to amplification.

- **Minimal separation and finite witnesses (Section 3; Appendices B and F).** A four-function class on two points separates: $\mathrm{A} = \frac{35}{29} > \mathrm{B} = \frac{7}{6}$, with ratio $\frac{30}{29}$. No class of at most three functions separates (Proposition 3.2), so four is optimal. The five-function class $D_2 = \{0,1\}^{\{x_1,x_2\}} \cup \{\bar{2}\}$ with $k = 3$ satisfies $\mathrm{B}(D_2) = \frac{7}{5} < \mathrm{A}(D_2) = \frac{16}{11}$. Its ratio is $80/77$, with matching learner and adversary certificates. We also give a single-shield family, exactly solved for $m \le 8$: $D_3$ has ratio $75/71$, and $D_4$ has ratio $5562/5255$. The ratio is globally maximized at $m = 4$. A 13-function class $C_{\mathrm{a}}$ has certified ratio in $[21296/20137,\ 4615908/4248907] \approx [1.0576, 1.0864]$.
- **A sharp asymptotic family (Section 4).** For the balanced towers $H_r = \{0,1\}^{[r]} \cup \{\bar{2}, \ldots, \overline{r+1}\}$, with $k = r + 2$, we prove an exact closed-form recursion for the adaptive value of every reachable state (Theorem 4.1). We obtain $\frac{5r}{8} \le \mathrm{A}(H_r) \le \frac{5r}{8} + \sqrt{6r-3}$ and $\mathrm{B}(H_r) = \frac{r}{2} + \log_2 r - \log_2 \log_2 r + O(1)$, so $\lim_{r\to\infty} \mathrm{A}(H_r)/\mathrm{B}(H_r) = \frac{5}{4}$ exactly. The constant $\frac{1}{8}$ is the discriminant-critical coefficient of a quadratic potential. We also pin down the coefficient 1 of $\log_2 r$ and the second-order term $-\log_2 \log_2 r$ in B. For $r = 2, 3$, the oblivious values are exact: $\mathrm{B}(H_2) = \frac{11}{6}$ and $\mathrm{B}(H_3) = \frac{18}{7}$. The corresponding exact ratios are $81/77$ and $112196/103225$.
- **Capping theory (Section 5.1).** A *uniform first-hit* principle shows that any union $H = \bigcup_i H_i$ of nonempty classes with pairwise disjoint label alphabets satisfies $\mathrm{A}(H) \le q(H) - 1 + \max_i \mathrm{A}(H_i) \le 2\mathrm{B}(H) + \max_i \mathrm{A}(H_i)$, independently of finite nesting depth (Theorem 5.2). With overlapping alphabets, the same principle localizes the game into an overlap cluster of at most $\Delta$ layers (Theorem 5.3), giving $\mathrm{A}/\mathrm{B} \le 2 + \Gamma_\Delta$. An exact analysis of Hamming-ball classes yields $\mathrm{A}(E(n,d)) = \frac{n}{2} - \mathbb{E}[(\mathrm{Bin}(n, \frac{1}{2}) - d)_+]$. The additive localization surcharge is already $\Theta(\sqrt{n})$ for two overlapping layers with $q = 2$ (Theorem 5.7). A localization-rank framework bounds the cost of label reuse from above, although it need not capture the right order (Remark E.1). An exact recursion also shows that the single-shield family is asymptotically degenerate: $\mathrm{A}(D_m)/\mathrm{B}(D_m) \to 1$ (Theorem F.2).
- **Necessary conditions (Section 5.2).** Any *linear* separation $\mathrm{A} \ge ck\mathrm{B}$ forces $q \ge \frac{ck}{6}$, $\mathrm{B} \ge \frac{ck}{12} - \frac{1}{2}$, and $\mathrm{F} \ge \frac{c}{6}\mathrm{B}$. Thus $q, \mathrm{B}, \mathrm{F} = \Omega(k)$ and $|H| \ge 2^{\Omega(k)}$. For covers with bounded-ratio atoms, it also forces localization rank $L \ge (ck - C_0)/2$ and, under positive halving, overlap degree $\Delta \ge 2^{\Omega(k)}$.

**What this does and does not resolve.** Concept classes *can* separate. Four functions already give a fixed constant factor, and an explicit family has ratio tending to exactly $\frac{5}{4}$. A *superconstant* gap, including a match for the $\Omega(k)$ pattern-class gap, remains open. Our capping theorems have constants at least 3 whenever some atom is nontrivial, so they do not constrain ratios in $(\frac{5}{4}, 3)$. They instead rule out a broad class of natural amplification constructions. Together with the necessary conditions, they identify where a larger separation would have to arise (Open Problem 6.1).

**Mechanism in one paragraph.** Every separation here uses *shield switching*. A cheap shield hypothesis, a constant function, sits beside an expensive binary core. After seeing the learner's public prediction distribution, the adaptive adversary can choose in each round between a shield threat and a core witness. The oblivious adversary must commit to one hypothesis in advance. Conversely, one principle underlies all our upper bounds: positive feedback *localizes* the hypothesis among label blocks. With disjoint blocks, the learner can wait for the first informative hit at expected total cost at most $q - 1 \le 2\mathrm{B}$. The remaining open regime lies at the boundary between these forces: label reuse must defeat localization while keeping the oblivious value low.

**Techniques.** Three devices carry the analysis. *(1) Exact state compression* places the game in a two-parameter family closed under all feedback transitions. Group averaging reduces each round to a one-parameter minimax between two affine threat lines, yielding (2). *(2) Potential/supersolution arguments* use explicit potentials for upper bounds and exchange-argument adversaries for lower bounds. They locate the critical coefficient $\frac{1}{8}$ exactly. *(3) All-negative-path accounting* fixes a prior for oblivious lower bounds and follows the all-negative feedback path. This turns the learner's behaviour into an accounting of first-prediction times. Machine verification is part of the methodology (Section 6.3).

Exact Bellman-type analyses of randomized mistake values are classical; see Filmus et al. (2023) on the randomized Littlestone dimension. We claim no novelty for the recursion method itself. Our contribution is the closed-form state compression for the tower family.

**Reading the results and proofs.** The exposition proceeds from minimal separation to tower asymptotics, then to composition bounds and necessary conditions. The index below locates their statements and proofs, together with the necessary conditions and finite certificates. Each row is a reading route; the rows do not form a single chain of proof dependencies.

| Route | Statements | Proof locations |
|---|---|---|
| Minimal separation | Theorem 3.1; Proposition 3.2 | Appendix B.1 |
| Adaptive bounds | Theorems 4.1 and 4.2 | Appendix C, through "Diagnostic domination"; Sections D.1–D.2 |
| Oblivious bounds | Theorem 4.3 | Sections D.3–D.4 |
| Ratio | Theorem 4.4 | Ratio argument in Section D.4 |
| Localization and its limits | Lemma 5.1; Theorems 5.2–5.5, 5.7 | Section 5.1; Appendix E (rank: Section E.3) |
| Necessary conditions | Theorem 5.8 | Section E.5 |

The protocol and basic tools are in Section 2. Related work is collected in Appendix A.1; further exact families and comparison tools are in Appendix F. The certificate index is in Appendix G.

## 2 Protocol and basic tools

Let $X$ be a nonempty domain and $Y = [k]$ a finite label set, $k \geq 2$. A *concept class* is a nonempty $H \subseteq Y^X$: every hypothesis labels every point of $X$ before the game starts.

**Definition 2.1** (Bandit protocol)**.** *In each round $t$: (i) the adversary reveals $x_t \in X$; (ii) the learner publishes a prediction distribution $\pi_t \in \Delta(Y)$; (iii) the adversary chooses a label distribution $\tau_t \in \Delta(Y)$; (iv) the learner draws $\hat{y}_t \sim \pi_t$ and reveals it; (v) the adversary draws $y_t \sim \tau_t$ and the learner observes only $\mathbf{1}[\hat{y}_t = y_t]$. The loss is the number of mistakes $\sum_t \mathbf{1}[\hat{y}_t \neq y_t]$.*

An *oblivious* adversary fixes an arbitrary finite or infinite sequence $(x_t, y_t)_t$ before the game, realizable by some $h \in H$: $h(x_t) = y_t$ for all $t$. We write $\mathrm{B}(H) = \inf_{\text{learner}} \sup_{\text{realizable sequence}} \mathbb{E}[\#\text{mistakes}]$. An *adaptive* adversary may choose $x_t$ as a function of the entire history and $\tau_t$ as a function of the history *and* the published $\pi_t$. Neither choice may depend on the sample $\hat{y}_t$. The adversary must satisfy the consistency requirement of Filmus et al. (2024): for every realizable feedback history $F$ and every $\pi$, every label $y \in \operatorname{supp} \tau$ must admit a realizable positive-feedback extension of $F$. We write $\mathrm{A}(H)$ for the corresponding optimal expected mistake number and $\mathrm{F}(H)$ for the full-information (label-revealed) optimal value.

We use two standard consequences proved in Appendix A. The first is the support restriction

$$\operatorname{supp} \tau_t \subseteq Y_V(x_t) := \{y : \exists h \in V,\ h(x_t) = y\}, \tag{1}$$

where $V$ is the current version space and $V_{x,y} = \{h \in V : h(x) = y\}$ denotes its fibers. The "for every $\pi$" quantifier in the consistency definition is essential: even labels with $\pi_y = 0$ may not be smuggled into $\operatorname{supp} \tau$. Second, we use monotonicity of $\mathrm{A}, \mathrm{B}$ under class inclusion and domain restriction. Oblivious adversaries are special adaptive adversaries, so $\mathrm{B}(H) \leq \mathrm{A}(H)$ always. The adaptive adversary does *not* fix a hidden target hypothesis. It must only keep every feedback history

realizable. In the class $D_2$ of Theorem B.1, predicting 0 when the hidden label is 1, and then 1 when the hidden label is 2, at the same point yields two legal negative feedbacks. Both are jointly explained by the shield $\bar{2}$, although the realized hidden labels $(1, 2)$ match no single hypothesis at that point. All results are proved for exactly this feedback-consistency model.

**Remark 2.2** (Model quantifier discipline). *All results are proved for the official consistency definition above. Theorems whose proofs use the "for every $\pi$" quantifier through* (1)*, notably Theorems 5.2, 5.3 and 5.5, do not automatically transfer to the weaker almost-sure-consistency interpretation. They survive there with uniform mixing using* positive summable *exploration rates $\varepsilon_t$ with total budget $\sum_t \varepsilon_t = \varepsilon$. This replaces $C_0 + 2$ by $C_0 + 2 + O(\varepsilon)$. A fixed exploration rate would incur unbounded cost over an infinite horizon. We identify the affected theorems rather than claim a global exemption.*

For $x \in X$ let $Y_H(x) = \{h(x) : h \in H\}$ and $q(H) = \max_x |Y_H(x)|$, the *local alphabet size*.

**Lemma 2.3** (Elementary bounds; Appendix A). *Let $H$ be a concept class, $X \neq \emptyset$.* (a) (Repetition.) *If some $x$ has $|Y_H(x)| = q \geq 2$ then $\mathrm{B}(H) \geq \frac{q-1}{2}$; hence $q(H) \leq 2\mathrm{B}(H) + 1$, and this is tight.* (b) (Finite classes.) *If $|H| = n$ then $\mathrm{A}(H) \leq \frac{n-1}{2}$, with equality when some point separates all $n$ hypotheses.* (c) (Binary entropy.) *If $q(H) \leq 2$ and $H$ is finite then $\mathrm{A}(H) \leq \frac{1}{2}\log_2 |H|$.* (d) (Binary cube.) $\mathrm{A}(\{0,1\}^{[n]}) = \mathrm{B} = \frac{n}{2}$. (e) (Zero law.) $\mathrm{B}(H) = 0$ *iff* $q(H) = 1$ *iff* $\mathrm{A}(H) = 0$.

The comparison tools are indexed in Appendix F.4: part (i) treats the full-information comparison, part (v) the expert bridge, and part (vi) the atom equalities.

## 3 A MINIMAL CONCEPT-CLASS SEPARATION

We begin with the witness theorem and the cardinality proposition. Their full proofs are in Appendix B.1. The five-function witness is Theorem B.1; its computations are in Appendix B.3.

**Theorem 3.1** (Four-function minimal witness). *Write elements of $\{0,2,3\}^{\{x_1,x_2\}}$ as pairs. The four-function class $H_{\mathrm{mid}} = \{\underbrace{(0,0)}_{z}, \underbrace{(2,2)}_{e}, \underbrace{(2,3)}_{f}, \underbrace{(3,2)}_{g}\}$ satisfies $\mathrm{A}(H_{\mathrm{mid}}) = \frac{35}{29}$ and $\mathrm{B}(H_{\mathrm{mid}}) = \frac{7}{6}$. Thus $H_{\mathrm{mid}}$ is a separating concept class with four functions on two points, with ratio $\frac{30}{29}$.*

*Proof sketch; full proof in Appendix B.1.* The adaptive Bellman certificate is two-sided: learner distribution $(\frac{8}{29}, \frac{14}{29}, \frac{7}{29})$ and adversary distribution $(\frac{9}{29}, \frac{17}{58}, \frac{23}{58})$ both evaluate to exactly $\frac{35}{29}$. The oblivious lower bound is a Bayes computation for the prior $(\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{3})$ and sequence $x_1, x_1, x_2$, where all three pure first-round actions cost exactly $\frac{7}{6}$. For the upper bound, an explicit learner randomizes its first prediction $(\frac{1}{6}, \frac{2}{3}, \frac{1}{6})$ on $(0,2,3)$ at the first point played, then uses three simple version-space subroutines: Pair uses one fair coin, while Follow$(z)$ and Follow$(m)$ are deterministic (Appendix B.1). We analyze the fixed true concept $h^\star$ and the three first-round branches. Obliviousness is used exactly once, when all three random branches are weighted against the *same* fixed concept. The concept names in the table follow the "Oblivious upper bound" paragraph of Appendix B.1. The conditional expected loss has these upper bounds, uniform over all instance sequences:

| $h^\star$ | $\hat{y}_1 = 0$ $(\frac{1}{6})$ | $\hat{y}_1 = 2$ $(\frac{2}{3})$ | $\hat{y}_1 = 3$ $(\frac{1}{6})$ |
|---|---|---|---|
| $z$ | 0 | $\frac{3}{2}$ | 1 |
| $m$ | 1 | $\frac{3}{2}$ | 0 |
| $e, q$ | $\frac{5}{2}$ | $\frac{1}{2}$ | $\frac{5}{2}$ |

Each row averages to exactly $\frac{7}{6}$. Exhaustive exact simulation of all instance sequences up to length 10 confirms the worst case. An adaptive adversary may shift the realized concept across the learner's random branches, so this does not contradict $\mathrm{A}(H_{\mathrm{mid}}) = \frac{35}{29}$. □

**Proposition 3.2** (Cardinality minimality). *No concept class with $|H| \leq 3$ separates: $\mathrm{A}(H) = \mathrm{B}(H)$ whenever $|H| \leq 3$. Hence the four functions of Theorem 3.1 are optimal in cardinality. Appendix B.1 proves this in two cases: $q \leq 2$ and the feedback bit reveals the label, or $q = 3 = |H|$ and Lemma 2.3(a,b) gives $1 \leq \mathrm{B} \leq \mathrm{A} \leq 1$.*

# 4 MULTI-SHIELD TOWERS AND THE SHARP $5/4$ THEOREM

We now turn from the minimal witness of Section 3 to the tower family and its asymptotic analysis.

## 4.1 STATES AND EXACT RECURSION

For $n \geq 1$, $s \geq 0$ let $D_n^{(s)} = \{0,1\}^{[n]} \cup \{\bar{2}, \ldots, \overline{s+1}\} \subseteq \{0, \ldots, s+1\}^{[n]}$, the $n$-point binary cube plus $s$ constant *shields* with pairwise distinct colors ($k = s+2$). The balanced diagonal $H_r := D_r^{(r)}$ is our asymptotic family. Playing $D_n^{(s)}$ generates *mixed states* $S_{n,s}^{d,\mathbf{b}}$ with $n$ free binary points and $d$ diagnostic points where all surviving core hypotheses agree on a fixed bit $b_i$. Diagnostic points arise from negative feedback on a binary prediction; $s$ shields also survive. Reachability requires $d \geq 1$ whenever $n = 0$.

**Theorem 4.1** (Diagnostic invariance and exact recursion)**.** *The adaptive value of $S_{n,s}^{d,\mathbf{b}}$ depends only on $(n, s)$; call it $a_{n,s}$. Then $a_{n,0} = \frac{n}{2}$ and $a_{0,s} = \frac{s}{2}$, the latter on reachable states with $d \geq 1$. For $n, s \geq 1$, with $t = \frac{n-1}{2}$, $U = 1 + a_{n-1,s}$, $V = 1 + a_{n,s-1}$,*

$$a_{n,s} = \min\left\{U,\ V,\ \frac{V\big((s+1)U - (s-1)t\big)}{s(U-t) + 2V}\right\}. \tag{2}$$

*In particular* $\mathrm{A}(D_n^{(s)}) = a_{n,s}$.

*Proof elements; full proof in Appendix C.* There are four informative free-point branches. A binary positive eliminates all shields, leaving value $t$. A binary negative leaves the half-cube plus *all* shields, a state $S_{n-1,s}^{d+1,\cdot}$ of cost $U$. This holds *even when the true hypothesis is a shield*, since negative feedback on a binary prediction does not reveal the true label. A shield positive leaves a singleton of value 0, and a shield negative removes that shield at cost $V$. Group averaging shows that symmetrizing over the binary swap and shield permutations costs the learner nothing. This reduces each free-point round to a one-parameter minimax between two affine threat lines. Their crossing yields the third term of (2), and the endpoints yield $U$ and $V$. Diagnostic points are dominated using $U \geq \frac{n+1}{2}$ and $U \geq \frac{s+1}{2}$. The adversary can always play a free point. Pure best-response labels, each with a nonempty positive fiber, complete its argument. □

Figure 1 illustrates these threats for $H_2$.

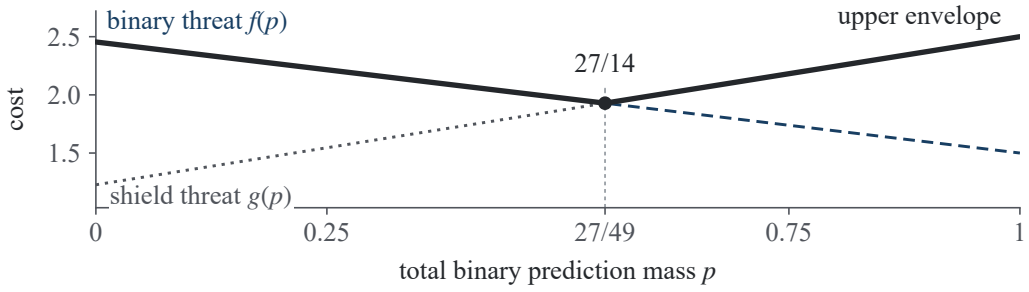


Figure 1: Binary and shield threats for $H_2 = D_2^{(2)}$ under predictions uniform within each group. The binary group has total mass $p$. Both costs include adaptive continuation. The upper envelope is minimized at $(27/49, 27/14)$. The general recursion also compares the endpoints (Theorem 4.1; Appendix C).

These are selected proof connections: Theorem 4.1 is used in the adaptive analysis of Theorem 4.2. The adaptive and oblivious bounds of Theorems 4.2 and 4.3 are then combined in Theorem 4.4. The proofs of Theorem 4.2 are in Sections D.1 and D.2. The separate proof of Theorem 4.3 is in Sections D.3 and D.4; the latter also contains the proof of Theorem 4.4.

### 4.2 ADAPTIVE ASYMPTOTICS

**Theorem 4.2** (Adaptive side). *For all* $r \geq 1$, $\frac{5r}{8} \leq \mathrm{A}(H_r) \leq \frac{5r}{8} + \sqrt{6r-3}$, *hence* $\lim_{r\to\infty} \mathrm{A}(H_r)/r = \frac{5}{8}$.

*Proof elements; Appendix D. Lower bound:* a mixing-phase adversary plays $\tau$ uniform on all currently feasible labels at fresh coordinates for $\lfloor r/2 \rfloor$ rounds. An exchange argument reduces the learner to action strings $C^c B^b$, whose excess over the binary baseline $r/2$ admits a concave-quadratic lower bound in $c$ with endpoints at least $r/8$ except for $r \in \{1, 2, 3, 5\}$, which the exact recursion values handle (Appendix D). *Upper bound:* define the two-piece potential as $\Phi(n, s) = \frac{s}{2} + \frac{n^2}{8s}$ for $s \geq 1$, $n \leq 2s$, and $\frac{n}{2}$ otherwise. Shifting it by $\sqrt{3(n+s-1)}$ gives a supersolution of (2), by monotonicity of (2) in $(U, V)$ and an exact quadratic certificate. On the diagonal, $\Phi(r, r) = \frac{5r}{8}$. The coefficient $\frac{1}{8}$ is critical: for $c < \frac{1}{8}$ the form $\frac{1}{2} + cx^2 - \frac{x}{2}$ crosses zero and no such potential exists, consistent with the slow empirical convergence of $\mathrm{A}(H_r)/r$ (0.882 at $r = 5$, 0.864 at $r = 6$), though it does not by itself determine the second-order growth. □

Tightness of the proven second-order bound $O(\sqrt{r})$ remains unknown. Over the computed intervals $400 \to 800$, $800 \to 1600$, and $1600 \to 3200$, the excess $E(r) = \mathrm{A}(H_r) - \frac{5r}{8}$ has doubling ratios approximately 1.284, 1.271, and 1.262, respectively. These finite-range observations do not determine its asymptotic order.

### 4.3 OBLIVIOUS ASYMPTOTICS

**Theorem 4.3** (Oblivious side). *For all* $r \geq 16$,

$$\frac{r}{2} + \log_2 r - \log_2 \log_2 r - \frac{3}{2} \;\leq\; \mathrm{B}(H_r) \;\leq\; \frac{r}{2} + \log_2 r - \log_2 \log_2 r + \frac{3}{2},$$

*so* $\mathrm{B}(H_r) = \frac{r}{2} + \log_2 r - \log_2 \log_2 r + O(1)$ *and* $\lim_r \frac{\mathrm{B}(H_r) - r/2}{\log_2 r} = 1$. *For all* $r \geq 2$, $\frac{r}{2} \leq \mathrm{B}(H_r) \leq \frac{r}{2} + \lceil \log_2(r+1) \rceil - \frac{1}{2}$.

*Proof elements; Appendix D. Upper bound:* a threshold learner tracks per-coordinate fair bit guesses, flipping each after a miss there. It probes the $r$ shields in uniformly random order once its first $J = \lceil \log_2(r/q) \rceil$ binary predictions, including repeated coordinates, have all erred, where $q = \log_2 r - \log_2 \log_2 r$. The rule covers arbitrary input sequences, including repeats (Appendix D). *Lower bound:* use a prior putting mass $1/\log_2 r$ on a uniform binary hypothesis and the rest on a uniform shield, with the fixed cyclic sequence $x_1, \ldots, x_r, x_1, \ldots$. Accounting along the all-negative path shows that every learner pays $\frac{r}{2} - \frac{1}{2} + \min_x \{pr2^{-x} + (1-p)x\}$. Continuous minimization gives the stated form; integrality only weakens the bound. □

### 4.4 THE RATIO AND FINITE CERTIFICATES

**Theorem 4.4** (Sharp ratio). $\lim_{r\to\infty} \frac{\mathrm{A}(H_r)}{\mathrm{B}(H_r)} = \frac{5/8}{1/2} = \frac{5}{4}$. *Explicitly, for all* $r \geq 2$,

$$\frac{5r/8}{r/2 + \lceil \log_2(r+1) \rceil - \frac{1}{2}} \;\leq\; \frac{\mathrm{A}(H_r)}{\mathrm{B}(H_r)} \;\leq\; \frac{5}{4} + 2\sqrt{\tfrac{6}{r}},$$

*the upper bound using* $\mathrm{B}(H_r) \geq r/2$. *Exact certified values for small* $r$ *appear in Table 2.*

Finite towers can exceed this limiting ratio. The recursion and threshold strategy certify $\mathrm{A}(H_{352})/\mathrm{B}(H_{352}) > 1.2617$; the integer interval certificate is given in Appendix D.4.

Figure 2 summarizes the finite certificates.

Two independent proofs establish the family's separation: the original capacity/potential argument and a cold-start reconstruction with a different adversary and potential. The adaptive values in Table 2 carry strong-duality certificates at every game state. For the oblivious values, we combine fixed-prior Bayes lower bounds and the threshold-learner bound, with exact values for $r = 2, 3$ (Section 6.3).

Tables 1 and 2, with the finite certificates, are collected in Appendix G. For the single-shield specialization, see Appendix F.1.

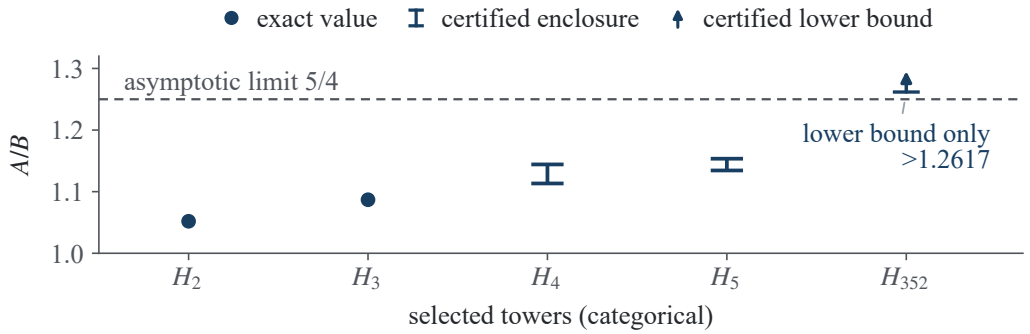


Figure 2: Certified ratios for selected balanced towers (categorical axis). Filled circles are exact values; capped segments are outward-rounded enclosures (Table 2). The arrow starts at the certified lower bound for $H_{352}$, with no upper endpoint shown (Appendix D.4). The dashed line is the asymptotic limit $5/4$ (Theorem 4.4).

## 5 Localization and limits to amplification

### 5.1 Capping bounds for composition

Throughout this subsection $X \neq \emptyset$. See Remark 2.2 for the model quantifier.

**Lemma 5.1** (Uniform first-hit)**.** *From any nonempty version space $V$, let the learner predict uniformly on the feasible set $Y_V(x_t)$ each round. Rounds with $|Y_V(x_t)| = 1$ cost nothing, since the learner predicts the unique feasible label. By* (1)*, whatever $\tau_t$ the adversary picks, the per-round hit probability is exactly $1/|Y_V(x_t)| \geq 1/q(H)$. Thus expected mistakes before the first informative positive feedback total at most $q(H) - 1$. Indeed (Appendix E), $\tau_t$ is fixed before $\hat{y}_t$ is sampled, so the hit probability is exactly $1/|Y_V(x_t)|$. A per-history hazard bound completes the proof.*

**Theorem 5.2** (Block-pure unions; depth-free capping)**.** *Let $H = \bigcup_{i=1}^{r} H_i$ with each $\emptyset \neq H_i \subseteq Y_i^X$ and the label blocks $Y_i \subseteq Y$ pairwise disjoint. Then*

$$\mathrm{A}(H) \;\leq\; q(H) - 1 + \max_i \mathrm{A}(H_i) \;\leq\; 2\mathrm{B}(H) + \max_i \mathrm{A}(H_i).$$

*Consequently, if every $H_i$ satisfies $\mathrm{A}(H_i) \leq C_0 \mathrm{B}(H_i)$ then $\mathrm{A}(H) \leq (C_0 + 2)\mathrm{B}(H)$. If every $H_i$ has $q(H_i) \leq 2$ or is a full product class, then $\mathrm{A}(H_i) = \mathrm{B}(H_i)$ by Appendix F.4(vi), giving $\mathrm{A}(H) \leq 3\mathrm{B}(H)$. These bounds are independent of nesting depth. A* finite-height *nested hierarchy of such unions is the flat union of its leaves, provided each node is the union of its children, whose blocks partition the parent's block.*

*Proof idea; details in Appendix E.* Run Lemma 5.1 until the first informative positive feedback $(x, y)$, at expected cost at most $q(H) - 1 \leq 2\mathrm{B}(H)$ (Lemma 2.3(a)). Since the blocks are disjoint, $y$ lies in exactly one $Y_i$, so $V_{x,y} \subseteq H_i$. Thus one hit localizes the layer. A restart-plus-monotonicity argument bounds the continuation by $\mathrm{A}(H_i)$, and $\mathrm{B}(H_i) \leq \mathrm{B}(H)$ turns the additive form into the $(C_0 + 2)$ ratio form. The zero law covers atoms with $\mathrm{B}(H_i) = 0$. □

Theorem 5.2 rules out every "class-shield" generalization of the towers whose shields are low-complexity classes on private label blocks: nesting never pushes the ratio beyond $C_0 + 2$. What matters is whether positive feedback localizes the block, rather than negative feedback eliminating a layer.

When layers share labels, positive feedback only localizes the hypothesis into a *cluster* of layers. For a cover $H = \bigcup_i H_i$ (all covers in this paper are finite families), let $J(x, y) = \{i : \exists h \in H_i, h(x) = y\}$, $\Delta = \max_{x,y} |J(x, y)|$ and $K_J = \bigcup_{i \in J} H_i$.

**Theorem 5.3** (One localization, overlap clusters)**.** *For any cover,* $\mathrm{A}(H) \leq q(H) - 1 + \max_{(x,y):\, y \in Y_H(x)} \mathrm{A}(K_{J(x,y)})$*; in particular* $\mathrm{A}(H) \leq q(H) - 1 + \max_{1 \leq |J| \leq \Delta} \mathrm{A}(K_J)$*, since every realized incidence set has $|J(x, y)| \leq \Delta$.*

**Theorem 5.4** (Reuse profile). *If* $0 < \mathrm{B}(H) < \infty$ *then* $\frac{\mathrm{A}(H)}{\mathrm{B}(H)} \le 2 + \Gamma_\Delta$, *where* $\Gamma_\Delta = \max\{\mathrm{A}(K_J)/\mathrm{B}(K_J) : 1 \le |J| \le \Delta,\ \mathrm{B}(K_J) > 0\}$ *(empty max* $= 0$*). Equivalently, if* $\mathrm{A} > (C+2)\mathrm{B}$ *then some cluster of at most* $\Delta$ *layers already has ratio exceeding* $C$.

As proved in Appendix E, the number of layers alone cannot produce a large ratio. Any large ratio already occurs within one *realized* cluster $K_{J(x,y)}$, which contains the version space left by a single positive feedback. Arbitrary small sub-unions need not be reachable by one hit.

**Theorem 5.5** (Localization rank). *Call* $\lambda : \{$*reachable version spaces*$\} \to \{0, \dots, L\}$ *a localization rank for the cover if* $\lambda(V) = 0$ *implies* $V \subseteq H_i$ *for some* $i$*; negative feedback never increases* $\lambda$*; and whenever* $\lambda(V) > 0$ *and* $|Y_V(x)| \ge 2$*, every positive feedback strictly decreases it. Then* $\mathrm{A}(H) \le L\,(q(H) - 1) + \max_i \mathrm{A}(H_i)$. *The positive-halving condition is* $\nu(V_{x,y}) \le \lceil \nu(V)/2 \rceil$ *whenever* $V$ *is reachable with* $\nu(V) > 1$, $|Y_V(x)| \ge 2$ *and* $y \in Y_V(x)$. *Here* $\nu(W)$ *is the least number of layers whose union contains* $W$. *This condition allows* $L \le 1 + \lceil \log_2 \Delta \rceil$, *so that* $\mathrm{A}/\mathrm{B} \le C_0 + 2(1 + \lceil \log_2 \Delta \rceil)$ *whenever all atoms satisfy* $\mathrm{A}(H_i) \le C_0 \mathrm{B}(H_i)$ *and* $\mathrm{B}(H) > 0$. *Finally, call a* finite-height *rooted tree of subclasses an* overlap-closed hierarchy *for* $H$ *if its root is* $H$*, every leaf* $L$ *satisfies* $\mathrm{A}(L) \le C_0 \mathrm{B}(L)$*, and every internal node* $N$ *carries a cover of* $N$ *by its children such that every incidence cluster of that cover,* $K_{J(x,y)}$ *computed within* $N$*, is contained in a single child. Every root fiber is contained in a leaf, so* one *positive feedback localizes into a leaf and* $\mathrm{A}(H) \le q(H) - 1 + \max_{L \text{ leaf}} \mathrm{A}(L) \le (C_0 + 2)\mathrm{B}(H)$. *Depth is free, but finite height is essential (Appendix E).*

**Remark 5.6** (These caps are silent below 3). *If some atom is nontrivial (*$\mathrm{B}(H_i) > 0$*), the best constants satisfy* $C_0 \ge 1$ *and* $\Gamma_\Delta \ge 1$. *Thus each displayed multiplicative cap,* $(C_0{+}2)\mathrm{B}$, $(2{+}\Gamma_\Delta)\mathrm{B}$ *and the positive-halving coefficient, is* $\ge 3\mathrm{B}$. *The raw additive forms sit below this floor, as does the rank form* $C_0 + 2L$ *with* $L = 0$. *Then* $H$ *lies inside a single atom, giving that atom's own ratio. If every atom of a* block-pure *cover has* $\mathrm{B}(H_i) = 0$*, each atom is a single function, every point separates all* $r$ *atoms, and* $\mathrm{A} = \mathrm{B} = \frac{r-1}{2}$ *by Lemma 2.3(a,b). This fails for general covers:* $H_{\mathrm{mid}}$ *is covered by four singletons yet separates. The bounds therefore say* nothing *about whether ratios in* $(\frac{5}{4}, 3)$ *are attainable; they delimit* growth.

Could the additive term in Theorem 5.3 be replaced by $\max_i \mathrm{A}(H_i)$ plus a function of $(q, \Delta)$ only? No:

**Theorem 5.7** (Hamming-ball classes exactly; the $\Theta(\sqrt{m})$ surcharge). *Let* $E(n, d) = \{h \in \{0,1\}^{[n]} : \sum_x h(x) \le d\}$*, the Hamming ball of radius* $d$*. This is a sparse binary class, not one-dimensional thresholds. Then*

$$\mathrm{A}\big(E(n,d)\big) = \frac{n}{2} - \mathbb{E}\big[(Z_n - d)_+\big] = \mathbb{E}\big[\min\{Z_n, d\}\big], \qquad Z_n \sim \mathrm{Bin}(n, \tfrac{1}{2}).$$

*Consequently, for the two-layer cover of the full cube* $\{0,1\}^{[2m]} = E^- \cup E^+$ *by the "at most* $m$ *ones" and "at least* $m$ *ones" layers (*$q = 2$, $\Delta = 2$*, laminar incidence sets),*

$$\mathrm{A}\big(\{0,1\}^{[2m]}\big) - \max\{\mathrm{A}(E^-), \mathrm{A}(E^+)\} = W_m := \frac{m}{2} \cdot \frac{\binom{2m}{m}}{4^m} = \frac{\sqrt{m}}{2\sqrt{\pi}}\,\big(1 + o(1)\big) \ge \frac{\sqrt{m}}{4} \longrightarrow \infty.$$

*Hence no bound of the form* $\mathrm{A}(H) \le f(q, \Delta) + \max_i \mathrm{A}(H_i)$ *holds for general covers. The statistical model-selection surcharge already appears at* $k = 2$*, where bandit feedback equals full information. It is therefore unrelated to label hiding.*

A $k = 3$ variant with a shield shows the same surcharge in laminar covers with $\Delta = 2$ (Section E.4). Localization *rank* bounds the additive cost (Theorem 5.5) but can overshoot its order: the two-layer Hamming cover has minimal feasible rank exactly $2m - 1$ (Remark E.1) against a true surcharge of $\Theta(\sqrt{m})$. When the rank bound is tight remains open.

Further audited bounds are $\mathrm{A} \le 6q\,\mathrm{F}$ and $\mathrm{A} \le 6\mathrm{B}\min\{k, 2\mathrm{B} + 1\}$; the serial-game operator bound $G \le 2\rho$ with $\sup_{q=3} G/\rho = 15/14$; constant-shield towers $\mathrm{A}(D_m^{(s)}) \le \frac{m+s}{2}$ with $\mathrm{A}/\mathrm{B} \le \frac{2s+3}{s+3} < 2$; and private constant shields $\mathrm{A}(\Sigma_s K) \le \mathrm{A}(K) + s$. Appendix F.4 gives these bounds and defines the objects. Items (i)–(ii) are essential elsewhere in this paper.

### 5.2 Necessary conditions for linear separation

**Theorem 5.8** (Necessary conditions). *Suppose* $\mathrm{A}(H) \ge c\,k\,\mathrm{B}(H)$ *with* $c > 0$ *constant and* $0 < \mathrm{B}(H) < \infty$. *Then:* (i) $q \ge \frac{ck}{6}$, $\mathrm{B} \ge \frac{ck}{12} - \frac{1}{2}$, $\mathrm{F} \ge \frac{c}{6}\mathrm{B}$ *(so* $q, \mathrm{B}, \mathrm{F} = \Omega(k)$*);* (ii) $|H| \ge 2^{\Omega(k)}$

*for finite $H$;* (iii) *every layer cover of $H$ with a localization rank $L$ and* $\max_i \mathrm{A}(H_i) \le C_0\mathrm{B}(H)$, $C_0 = o(k)$, *has* $L \ge \frac{ck-C_0}{2}$, *and in positive-halving covers* $\log_2 \Delta \ge \frac{ck-C_0}{2} - 2$, *i.e.* $\Delta \ge 2^{\Omega(k)}$. *For a general ratio* $\rho = \mathrm{A}/\mathrm{B}$: $q \ge \frac{\rho}{6}$, $\mathrm{B} \ge \frac{\rho}{12} - \frac{1}{2}$, $\log_k |H| = \Omega(\rho\mathrm{B}/k)$; *for* $\omega(1) = \rho = o(k)$ *only these $\rho$-scaled forms follow.*

The proof (Appendix E.5) combines Filmus et al. (2024, Theorems D.1 and C.7) with Lemma 2.3 and Theorem 5.5.

# 6 Discussion and verification

## 6.1 Conclusion

We settle the existence component of the concept-class adaptivity problem posed at NeurIPS 2024 by Filmus et al. (2024): adversarial adaptivity has a strict cost under bandit feedback. Four functions on two points suffice to separate the adaptive and oblivious values; three cannot. An explicit tower family attains the exact asymptotic ratio $\frac{5}{4}$. First-hit localization caps disjoint-block compositions of uniformly bounded-ratio atoms at a constant factor. Such compositions cannot amplify bounded atom ratios into a superconstant separation.

The results also clarify which source of difficulty matters. The additive union cost of model selection must be distinguished from the gap $\mathrm{A} - \mathrm{B}$. Union cost is unbounded already at $k = 2$ (Theorem 5.7), where $\mathrm{A} = \mathrm{B}$, so it cannot create a gap by itself. Every separation here instead exploits *commitment*: the oblivious adversary fixes a hypothesis before the learner randomizes, and shield switching exploits that restriction. Exact-solver-in-the-loop search was essential for finding differences as small as $\frac{16}{11}$ versus $\frac{7}{5}$. The next central challenge is to determine whether concept classes admit a superconstant adaptivity ratio (Open Problem 6.1).

Recommendation and output verification motivate interest in learning from limited feedback. The formal protocol and its scope are specified in Section 2; the application examples provide context for the question studied here.

## 6.2 The remaining questions

**Open Problem 6.1.** *Determine* $\sup_H \mathrm{A}(H)/\mathrm{B}(H)$ *as a function of $k$ over classes with* $0 < \mathrm{B}(H) < \infty$. *The balanced tower family has exact asymptotic ratio* $\frac{5}{4}$ *(Theorem 4.4), while the finite instance* $H_{352}$ *has ratio greater than* $1.2617$ *(Appendix D.4). The upper bound is* $6k$, *since* $\mathrm{A} \le 6q\,\mathrm{F}$, $\mathrm{F} \le \mathrm{B}$ *and* $q \le k$ *by Appendix F.4(i). This follows immediately from Filmus et al. (2024), who state* $O(k \log k)$; *we record it for calibration, not as new. Specifically: (a) can any concept-class family achieve a superconstant ratio? (b) can concept classes match the pattern-class* $\Omega(k)$ *bound of Filmus et al. (2024)? By Section 5.1, any superconstant family must defeat one-hit localization persistently, in the* existential *sense made precise in Appendix E: a class with* $\mathrm{A} = \rho\mathrm{B}$ admits *a legal chain of* $\Omega(\rho)$ *informative positive feedbacks with per-step value drop at most* $2\mathrm{B}(H) + \varepsilon$. *This regime has neither constructions nor impossibility tools so far. Further open items are the exact second-order term of* $\mathrm{A}(H_r)$; $\mathrm{B}(H_r)$ *to within* $o(1)$; *a closed form for* $\mathrm{B}(D_m)$; *the exact* $\mathrm{B}(C_{\mathrm{a}})$.

## 6.3 Exact computational verification

Machine computation certifies the finitely many exact rational values. The asymptotic theorems (Theorems 4.2–4.4) are *proved*; machine checks cover only their finite instances. *Adaptive* values come from backward induction over version spaces with exact-fraction matrix-game solving. The submitted script recomputes each value and produces a primal–dual certificate $(M, p, \tau, v)$ at every game state, verified by an independent checker. *Oblivious* values obey no such per-state recursion. Each exact $D_n^{(s)}$ value is the meet of a fixed-prior Bayes lower bound and the universal threshold-learner bound of Appendix D. Priors, sequences and the claim-by-claim manifest appear in Appendix G. The value $\mathrm{B}(H_{\mathrm{mid}})$ is closed by its bespoke learner, which the script simulates exhaustively to length $10$. Research-process verification used three independently written solvers, cold-start re-derivations and a $46{,}719$-state out-of-sample validation of (2). This is reported as process, not as part of the submission. Appendix G also notes an erratum in the proof of Filmus et al. (2024, Lemma C.11).

### REPRODUCIBILITY STATEMENT

All results are theorems. Appendices A–G give full proofs of every main-text claim, with all assumptions stated in Section 2. Remark 2.2 specifies the exact adversary-model quantifier. The two auxiliary results, the sharper tower B lower bound and serial-game extremality, are proved in Appendix F.4. Explicit finite certificates for all exact oblivious values appear in Appendix G and Appendix F.3. The supplementary exact-arithmetic script reproduces every exact rational value using only the Python standard library. It re-derives all adaptive values with independently checked primal–dual certificates at every game state. It also checks the recursion (2) against direct solves on all 25 states $n, s \leq 5$, verifies every oblivious certificate, including the $C_{\mathrm{a}}$ mixture and the Table 2 lower bounds, and runs the exhaustive learner simulation of Theorem 3.1. Section 6.3 and Appendix G, which includes a claim-by-claim manifest, describe the verification methodology. A second standard-library script, `verify_tower_intervals.py`, certifies the finite-tower ratio bound and reproduces the reported excess doubling ratios using integer intervals.

### ETHICS STATEMENT

This theoretical study of online learning involves no human subjects, personal data, or deployment artifacts. We foresee no specific ethical concerns beyond those common to foundational machine learning theory.

### AI USE STATEMENT

In this work, we used generative AI tools to help develop theoretical models or conceptual frameworks, assist in writing proofs, design or provide feedback on research methodology and computational verification, implement methods, assist with translation, and interpret results. We did not use generative AI tools to formulate mathematical claims, provide critical ingredients for proving mathematical claims, or propose or refine hypotheses. Generating synthetic datasets, cleaning or reformatting datasets, and qualitative or thematic data analysis are not applicable to this theoretical study.

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

We have reviewed all AI-assisted work. Three human reviewers manually checked all mathematical derivations and code for correctness and found no errors. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

## REFERENCES

Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The nonstochastic multi-armed bandit problem. *SIAM Journal on Computing*, 32(1):48–77, 2002.

Shai Ben-David, Dávid Pál, and Shai Shalev-Shwartz. Agnostic online learning. In *Proceedings of the 22nd Conference on Learning Theory (COLT)*, 2009.

Guy Blanc and Gregory Valiant. Adaptive and oblivious statistical adversaries are equivalent, 2024. arXiv:2410.13548.

Amit Daniely, Sivan Sabato, Shai Ben-David, and Shai Shalev-Shwartz. Multiclass learnability and the ERM principle. *Journal of Machine Learning Research*, 16:2377–2404, 2015.

DeepSeek-AI. DeepSeek-R1: Incentivizing reasoning capability in LLMs via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025. URL `https://arxiv.org/abs/2501.12948`.

Liad Erez and Tomer Koren. From contextual combinatorial semi-bandits to bandit list classification: Improved sample complexity with sparse rewards. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025. arXiv:2502.09257. Earlier versions titled "Bandit Multiclass List Classification".

Yuval Filmus, Steve Hanneke, Idan Mehalel, and Shay Moran. Optimal prediction using expert advice and randomized Littlestone dimension. In *Proceedings of the 36th Conference on Learning Theory (COLT)*, volume 195 of *Proceedings of Machine Learning Research*, pp. 773–836, 2023.

Yuval Filmus, Steve Hanneke, Idan Mehalel, and Shay Moran. Bandit-feedback online multiclass classification: Variants and tradeoffs. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024. Full version: arXiv:2402.07453. We cite theorem numbers from the NeurIPS version; its Theorem D.1, Theorem C.7 and Lemma C.11 appear in the arXiv version as Theorem 5.1, Theorem 4.7 and Lemma 4.11 respectively.

Valentio Iverson, Tosca Lechner, and Sivan Sabato. On the importance of randomization in discriminative feature feedback. In *Proceedings of the 39th Conference on Learning Theory (COLT)*, volume 336 of *Proceedings of Machine Learning Research*, pp. 3693–3715, 2026.

Sham M. Kakade, Shai Shalev-Shwartz, and Ambuj Tewari. Efficient bandit algorithms for online multiclass prediction. In *Proceedings of the 25th International Conference on Machine Learning (ICML)*, 2008.

Lihong Li, Wei Chu, John Langford, and Robert E. Schapire. A contextual-bandit approach to personalized news article recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, pp. 661–670, 2010. doi: 10.1145/1772690.1772758. URL `https://arxiv.org/abs/1003.0146`.

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

## A Protocol foundations and related work

**Support restriction** (1)**.** If $y \in \operatorname{supp} \tau_t$, consistency requires the positive extension of the current history by $(x_t, y)$ to be realizable. Its witness $h \in H$ is consistent with all past feedback, so $h \in V$ and $h(x_t) = y$, i.e. $y \in Y_V(x_t)$. The "for every $\pi$" quantifier ensures this even for labels with $\pi_y = 0$. Conversely, if the learner predicts $\hat{y} \neq y$, the same witness survives the negative feedback, so reachable histories remain realizable.

**Monotonicity.** $\mathrm{B}(V) \leq \mathrm{B}(H)$ for $V \subseteq H$ because realizable sequences nest. For A, take an adversary legal for $V$. Extend it to all $H$-realizable histories $F$ not realizable by $V$: fix any instance $x_*$ and play $\tau = \delta_{h_F(x_*)}$ for a witness $h_F \in H$ of $F$. This extension is legal for $H$ and agrees with the original on all its paths, giving $\mathrm{A}(V) \leq \mathrm{A}(H)$. For domain restriction, if $G \subseteq H|_{X_0}$ for $X_0 \subseteq X$, an adversary playing only points of $X_0$ turns the $H$-game into a $G$-game. Witnesses lift from $X_0$ to $X$, so $\mathrm{A}(H) \geq \mathrm{A}(G)$ and $\mathrm{B}(H) \geq \mathrm{B}(G)$.

*Proof of Lemma 2.3.* (a) Fix $x$ with labels $S$, $|S| = q$, and, for each $y \in S$, a witness $h_y$. The constant sequences $(x, y), (x, y), \ldots$ are realizable. Fix a learner and its random bits. Along the all-negative path at $x$, let $\rho_y \in \mathbb{N} \cup \{\infty\}$ be the first time label $y$ is predicted. If some $\rho_y = \infty$, the constant sequence with label $y$ already forces unbounded loss. Otherwise, the $\rho_y$ are distinct positive integers, so $\sum_y (\rho_y - 1) \geq \binom{q}{2}$. On the sequence with label $y$, the learner errs at least $\rho_y - 1$ times. Averaging over $y$ and over the random bits shows that some $y$ forces expected loss $\geq \frac{q-1}{2}$. For tightness, on one point with $q$ labels, testing the labels in uniformly random order makes $\frac{q-1}{2}$ mistakes in expectation against every label.

(b) With $|V| = n$, at a queried point with fiber sizes $(a_y)$, let the learner predict label $y$ with probability $a_y/n$ (concept-uniform). For true label $y$, the Bellman cost with potential $\Phi(n) = \frac{n-1}{2}$

satisfies

$$2n\Big(\Phi(n) - \tfrac{a_y}{n}\Phi(a_y) - \sum_{z\neq y}\tfrac{a_z}{n}\big(1+\Phi(n-a_z)\big)\Big) = -2n + a_y(n+2) - a_y^2 + \sum_{z\neq y} a_z^2$$
$$\geq (a_y - 1)(n - a_y) \geq 0,$$

where the middle inequality uses $\sum_{z\neq y} a_z^2 \geq \sum_{z\neq y} a_z = n - a_y$. Backward induction over any finite horizon, followed by the horizon limit, gives $\mathrm{A}(V) \leq \frac{n-1}{2}$. If some point separates all $n$ hypotheses, (a) gives the matching lower bound.

(c) If a point splits $V$ into parts of sizes $N_0, N_1$ ($N = N_0 + N_1$), set $\alpha = \frac{1}{2}\log_2 \frac{N}{N_0}$, $\beta = \frac{1}{2}\log_2 \frac{N}{N_1}$. By AM–GM, $\alpha + \beta \geq 1$, so we may choose $p_1 \leq \alpha$, $p_0 \leq \beta$ with $p_0 + p_1 = 1$. Both Bellman rows are then at most $\frac{1}{2}\log_2 N$, and induction proves the claim.

(d) For the upper bound, predict $\frac{1}{2}/\frac{1}{2}$ at each fresh coordinate. Either feedback fixes the bit, so the *expected* total is at most $\frac{n}{2}$. The oblivious lower bound $\mathrm{B} \geq \frac{n}{2}$ requires a separate argument from the adaptive bound. Put the uniform prior on the cube and fix the sequence visiting the $n$ coordinates once each. Conditioned on the past, each fresh coordinate is an independent fair bit, so every learner's expected loss is at least $\frac{n}{2}$. Averaging over the prior, some fixed realizable sequence forces at least $\frac{n}{2}$, i.e. $\mathrm{B} \geq \frac{n}{2}$. The adaptive adversary that, after seeing $\pi$, puts $\tau$ on the less likely label at a fresh coordinate also proves $\mathrm{A} \geq \frac{n}{2}$ directly. Every partial assignment extends within the full cube, so consistency holds.

(e) If $q \geq 2$, then $\mathrm{B} \geq \frac{1}{2}$ by (a). If $q = 1$, all hypotheses are the same function, so predicting it is always correct by (1). □

### A.1 Related work

Beyond Filmus et al. (2024), bandit multiclass classification in the mistake-bound and regret models goes back to Auer et al. (2002) and Kakade et al. (2008). Multiclass learnability combinatorics are studied by Daniely et al. (2015), and agnostic online learning by Ben-David et al. (2009). We use two external ingredients from Filmus et al. (2024): Theorem D.1 ($\mathrm{A} \leq 6k\,\mathrm{F}$ for every pattern class and hence every concept class) and Theorem C.7 (the realizable expert bound $\mathrm{opt}^{\mathrm{adap}}_{\mathrm{bandit}}(n,k,0) \leq k\log_k n$). Everything else is self-contained. The four-function witness and tower family are closely related to the concept-class construction of Filmus et al. (2024, Appendix F). Recent follow-ups of Filmus et al. (2024) study bandit list classification (Erez & Koren, 2025). Most of their results extend to adaptive adversaries, but they do not address concept-class separation. A different line (Blanc & Valiant, 2024) shows that adaptive and oblivious *sample-corruption* adversaries are equivalent up to polynomial factors in statistical settings. Its model and mechanism are unrelated to the online bandit protocol considered here. In discriminative feature feedback, Iverson et al. (2026) also separate adaptive from oblivious adversaries. Their feedback reveals the true label and a distinguishing feature, and their separation concerns learning with exceptions. Neither result therefore implies the other.

## B Minimal and elementary separating classes

### B.1 The four-function witness: full proof

*Adaptive value.* At $x_1$, the fiber/complement continuation values are $b = (0, \frac{1}{2}, 0)$ and $d = (\frac{3}{4}, \frac{1}{2}, 1)$ for the labels $(0, 2, 3)$. The fibers $\{z\}, \{e, f\}, \{g\}$ have values $0, \frac{1}{2}, 0$. For the complements, $\{e, f, g\}$ has two fibers at either point, with values $\frac{1}{2}$ and $0$. With mass $p$ on the doubleton label, the Bellman rows are $\frac{3}{2} - p$ and $p$, balancing at $p = \frac{3}{4}$ with value $\frac{3}{4}$. The class $\{z, g\}$ has value $\frac{1}{2}$, and $\{z, e, f\}$ has three distinct labels at $x_2$, giving value 1 by Lemma 2.3(a,b). The learner distribution $(\frac{8}{29}, \frac{14}{29}, \frac{7}{29})$ and the adversary distribution $(\frac{9}{29}, \frac{17}{58}, \frac{23}{58})$ make every row, resp. every pure prediction, equal to $\frac{35}{29}$. The other point is isomorphic under $x_1 \leftrightarrow x_2$, $f \leftrightarrow g$, giving the certified value $\mathrm{A}(H_{\mathrm{mid}}) = \frac{35}{29}$.

*Oblivious lower bound.* Put the prior $(z, e, f, g) \mapsto (\frac{1}{3}, \frac{1}{6}, \frac{1}{6}, \frac{1}{3})$ and fix the sequence $x_1, x_1, x_2$. At $x_1$ the three label masses (0, 2, 3) are each $\frac{1}{3}$. A first-round prediction of 2 costs $\frac{2}{3} + \frac{1}{3}\cdot\frac{1}{2} + \frac{2}{3}\cdot\frac{1}{2} = \frac{7}{6}$.

A hit leaves $\{e, f\}$, separated at $x_2$ at cost $\frac{1}{2}$; a miss leaves $\{z, g\}$, separated at the second $x_1$ at cost $\frac{1}{2}$. A first-round prediction of 0 or 3 costs $\frac{2}{3} + \frac{2}{3} \cdot \frac{3}{4} = \frac{7}{6}$: a miss leaves a $\frac{1}{2}/\frac{1}{4}/\frac{1}{4}$ posterior with residual Bayes cost $\frac{3}{4}$. All three pure actions, and hence any randomization, cost exactly $\frac{7}{6}$, giving $\mathrm{B}(H_{\mathrm{mid}}) \geq \frac{7}{6}$ by Yao's principle.

*Oblivious upper bound.* Let $a$ be the first point played and $b$ the other. Rename the concepts so that $q$ has $(a, b)$-labels $(2, 3)$ and $m$ has $(3, 2)$. The learner draws its first prediction at $a$ from $(0, 2, 3)$ with probabilities $(\frac{1}{6}, \frac{2}{3}, \frac{1}{6})$ and then follows its version space $V$ with three subroutines. The first uses one fair coin; the other two are deterministic. $\mathsf{Pair}(u, v)$ predicts the common label where $u, v$ agree and tosses a fair coin between their labels at a distinguishing point, with expected residual cost $\leq \frac{1}{2}$. $\mathsf{Follow}(z)$ on $V = \{z, e, q\}$ always predicts 0; a mistake reveals $V = \{e, q\}$, after which it uses $\mathsf{Pair}$. $\mathsf{Follow}(m)$ on $V = \{e, q, m\}$ or $\{e, m\}$ always predicts $m$'s label; the possible transitions $\{e, q\}, \{q\}, \{e, m\}, \{m\}$ each cost at most one mistake before a $\mathsf{Pair}$ or a singleton. A case analysis covers the fixed true concept and three first-round branches. The oblivious adversary must fix one realizable concept *before* the learner randomizes. This gives the following upper bounds on the conditional expected loss, uniform over all instance sequences:

| $h^\star$ | $\hat{y}_1 = 0$ $(\frac{1}{6})$ | $\hat{y}_1 = 2$ $(\frac{2}{3})$ | $\hat{y}_1 = 3$ $(\frac{1}{6})$ |
|---|---|---|---|
| $z$ | 0 | $\frac{3}{2}$ | 1 |
| $m$ | 1 | $\frac{3}{2}$ | 0 |
| $e, q$ | $\frac{5}{2}$ | $\frac{1}{2}$ | $\frac{5}{2}$ |

and each row averages to exactly $\frac{7}{6}$. An exhaustive exact simulation of this learner over *all* instance sequences of length up to 10 and all four concepts confirms that the worst-case expectation is exactly $\frac{7}{6}$; code is in the supplementary material. Hence $\mathrm{B}(H_{\mathrm{mid}}) \leq \frac{7}{6}$.

*Remark.* The proof of the upper bound uses obliviousness exactly once: the three first-round branches are weighted against the *same* fixed realizable concept. An adaptive adversary may shift the realized concept across the learner's random branches, which is why this does not contradict $\mathrm{A}(H_{\mathrm{mid}}) = \frac{35}{29}$.

*Proof of Proposition 3.2.* If $q(H) \leq 2$, one feedback bit reveals the true label and $\mathrm{A}(H) = \mathrm{B}(H)$ (Appendix F.4(vi)). Otherwise $q(H) = 3 = |H|$, so at some point the three hypotheses carry three pairwise distinct labels. Lemma 2.3(a) gives $\mathrm{B} \geq \frac{3-1}{2} = 1$, while Lemma 2.3(b) gives $\mathrm{A} \leq \frac{|H|-1}{2} = 1$, so $1 \leq \mathrm{B} \leq \mathrm{A} \leq 1$. For $|H| \leq 2$ the same argument yields $\mathrm{A} = \mathrm{B} \in \{0, \frac{1}{2}\}$. □

### B.2 The five-function statement and proof sketch

**Theorem B.1** (Five-function witness)**.** *Let $X = \{x_1, x_2\}$, $Y = \{0, 1, 2\}$ and $D_2 = \{0, 1\}^X \cup \{\bar{2}\}$, the four binary functions and the constant function 2. Then*

$$\mathrm{A}(D_2) = \frac{16}{11}, \qquad \mathrm{B}(D_2) = \frac{7}{5}, \qquad \textit{so} \qquad \frac{\mathrm{A}(D_2)}{\mathrm{B}(D_2)} = \frac{80}{77} > 1.$$

*Proof sketch; full details in Appendix B.3. Adaptive value:* at either point, positive feedback on a binary label eliminates the shield and leaves a one-coordinate cube of value $\frac{1}{2}$. Positive feedback on 2 ends the game. Negative feedback on a binary label leaves three survivors taking three distinct labels at the other point, with value 1 by Lemma 2.3(a,b). Negative feedback on 2 leaves the full cube of value 1. For prediction weights $(p_0, p_1, p_2)$ the cost rows are therefore $R_0 = \frac{p_0}{2} + 2(1 - p_0)$, $R_1 = \frac{p_1}{2} + 2(1 - p_1)$, $R_2 = 2(1 - p_2)$. Equalizing at $p_0 = p_1 = \frac{4}{11}$, $p_2 = \frac{3}{11}$ makes all rows $\frac{16}{11}$. The *same* distribution is an optimal equalizing (dual) certificate for the adversary, so both one-sided certificates meet at $\frac{16}{11}$. Every threatened label has a nonempty positive fiber, so consistency (1) holds throughout. *Oblivious upper bound:* a threshold learner draws $J \in \{1, 2\}$ with probabilities $\frac{3}{5}, \frac{2}{5}$ and probes label 2 once after the first $J$ binary predictions all err. This guarantees $\frac{7}{5}$ against binary targets $(1 + \mathbb{E}[2^{-J}])$ and the shield $(\mathbb{E}[J])$. *Oblivious lower bound:* the uniform prior on the five hypotheses with the fixed sequence $x_1, x_2, x_1, x_2$ has Bayes-optimal loss exactly $\frac{7}{5}$, by a three-case first-round analysis. □

**Remark B.2** (Mechanism: shield switching)**.** *After the learner publishes $\pi$, the adaptive adversary answers a binary-heavy $\pi$ with the shield threat and a shield-heavy $\pi$ with a binary witness. Every*

*threatened label remains realizable. The oblivious adversary must commit in advance, and the learner's randomized threshold exploits this commitment, resolving the existence component of the concept-class adaptivity problem in Filmus et al. (2024, Section 4.1): global consistency does* not *preclude an adaptive-versus-oblivious gap.*

### B.3 The five-function witness: full computations

#### B.3.1 $\mathrm{A}(D_2) = 16/11$

The points are symmetric. Consider a round at $x_1$ with the full class alive. The three feedback branches have the following continuation values.

- Positive feedback on a binary label (0, say): the shield dies and $V' = \{h_{00}, h_{01}\}$, a one-coordinate binary cube of value $\frac{1}{2}$ (Lemma 2.3(d)).
- Positive feedback on 2: the game ends (value 0).
- Negative feedback on binary label 0: $V' = \{h_{10}, h_{11}, \bar{2}\}$. At $x_2$ these three take the three distinct labels $0, 1, 2$, so by Lemma 2.3(a,b) the value is exactly $\frac{3-1}{2} = 1$. Label 1 is symmetric.
- Negative feedback on 2: $V' = \{0,1\}^X$, value 1 (Lemma 2.3(d)).

Thus for prediction weights $(p_0, p_1, p_2)$ the cost rows are

$$R_0 = \tfrac{p_0}{2} + 2(1-p_0), \qquad R_1 = \tfrac{p_1}{2} + 2(1-p_1), \qquad R_2 = 2(1-p_2).$$

Setting $p_0 = p_1$ and $R_0 = R_2$ gives $p_0 = \frac{4}{11}$, $p_2 = \frac{3}{11}$, and all three rows equal $\frac{16}{11}$, so the learner guarantees $\frac{16}{11}$. The adversary's certificate is the same *equalizing* distribution $\tau = (\frac{4}{11}, \frac{4}{11}, \frac{3}{11})$. Against it, every pure prediction earns exactly $\frac{16}{11}$, so no learner distribution can do better. Per-$\pi$ best responses earn at least as much by picking the pure label maximizing the corresponding row. A pure binary prediction 0 earns $\frac{4}{11} \cdot \frac{1}{2} + \frac{7}{11} \cdot 2 = \frac{16}{11}$: a miss leaves $V' = \{h_{10}, h_{11}, \bar{2}\}$ regardless of the true label, with value 1. A pure prediction 2 earns $\frac{8}{11} \cdot 2 = \frac{16}{11}$. Every threatened label carries a nonempty positive fiber, so consistency (1) holds, and the continuation values are attained recursively. Hence $\mathrm{A}(D_2) = \frac{16}{11}$ exactly.

#### B.3.2 $\mathrm{B}(D_2) \le 7/5$: the threshold learner

Initialize an independent fair coin guess $b_x \in \{0,1\}$ for each point. Whenever a binary prediction at $x$ errs, flip $b_x$. Draw $J \in \{1,2\}$ with $\Pr[J=1] = \frac{3}{5}$, $\Pr[J=2] = \frac{2}{5}$, and make at most $J$ binary predictions, using $b_{x_t}$ at the current point. If any is correct, the target cannot be $\bar{2}$, so continue binary tracking forever. If the first $J$ are all wrong, predict 2 once. A correct prediction identifies the target as $\bar{2}$; otherwise, resume binary tracking.

*Binary target $h$.* Each point contributes at most one binary mistake, each with probability $\frac{1}{2}$, so the binary mistakes total at most $\frac{m}{2} = 1$ in expectation. The probe fires only if the first $J$ binary predictions all err. If those rounds involve a repeated point, the guess has flipped to the correct bit by the second visit, so the probe never fires. Otherwise, they involve $J$ independent fair bits. Hence $\Pr[\text{probe}] \le 2^{-J}$ and $\mathbb{E}[M \mid h \text{ binary}] \le 1 + \mathbb{E}[2^{-J}] = 1 + \frac{3}{5} \cdot \frac{1}{2} + \frac{2}{5} \cdot \frac{1}{4} = \frac{7}{5}$.

*Shield target $\bar{2}$.* All binary predictions err and the probe is then correct, so $\mathbb{E}[M \mid \bar{2}] \le \mathbb{E}[J] = \frac{3}{5} + 2 \cdot \frac{2}{5} = \frac{7}{5}$.

#### B.3.3 $\mathrm{B}(D_2) \ge 7/5$: a Bayes lower bound

Put the uniform prior on the five hypotheses and fix the sequence $x_1, x_2, x_1, x_2$. We reveal it to the learner in advance; this can only help it. Condition on the learner's first-round action.

*Predict 2 first:* with probability $\frac{4}{5}$ the target is binary; the round is lost and both bits remain unknown, so the risk is at least $\frac{4}{5}(1+1) = \frac{8}{5}$.

*Predict 0 first* (label 1 is symmetric): with probability $\frac{2}{5}$ the prediction is correct, leaving one unknown bit and contributing $\frac{2}{5} \cdot \frac{1}{2} = \frac{1}{5}$. With probability $\frac{3}{5}$ it is wrong, incurring immediate cost $\frac{3}{5}$,

and the posterior is uniform on $\{\bar{2}, h_{10}, h_{11}\}$. Their labels at the next point $x_2$ are the three distinct values $2, 0, 1$, so any prediction errs with probability $\frac{2}{3}$. An error leaves two equiprobable hypotheses that the remaining rounds separate at expected cost $\frac{1}{2}$. That posterior is worth $\frac{2}{3} + \frac{2}{3} \cdot \frac{1}{2} = 1$. The total is $\frac{1}{5} + \frac{3}{5} + \frac{3}{5} \cdot 1 = \frac{7}{5}$.

A randomized first round is a convex combination of these pure risks, so the Bayes risk is at least $\frac{7}{5}$. By Yao's principle, $\mathrm{B}(D_2) \geq \frac{7}{5}$. An exact dynamic program over all reachable posterior supports confirms the value $\frac{7}{5}$; code is in the supplementary material.

## C Exact recursion: full proof

Throughout, $S^{d,\mathbf{b}}_{n,s}$ is as in Section 4. We prove Theorem 4.1 by joint induction on $n + s$, simultaneously for all $d$ and $\mathbf{b}$. Negative feedback at a free point increases $d$ but strictly decreases $n + s$, so the induction is well-founded.

**Boundaries.** If $s = 0$, diagnostic labels are known and the state is an $n$-cube, so $a_{n,0} = \frac{n}{2}$ by Lemma 2.3(d). If $n = 0$ (so $d \geq 1$), the state has one surviving core hypothesis and $s$ shields. At every diagnostic point, these $s + 1$ candidates carry pairwise distinct labels. Lemma 2.3(a,b) gives the exact value $\frac{s}{2}$, independent of $d, \mathbf{b}$ and of the shield colours.

**Transitions.** Let $n, s \geq 1$ and write $t = \frac{n-1}{2}$, $U = 1 + a_{n-1,s}$, $V = 1 + a_{n,s-1}$, well-defined by induction. Predicting a currently infeasible label causes a sure mistake without changing the state, so it is dominated and excluded. At a *free* point the four informative branches are:

| | | | |
|---|---|---|---|
| predict $z \in \{0, 1\}$ | positive: | half-cube, shields dead | value $t$; |
| predict $z \in \{0, 1\}$ | negative: | half-cube + all shields $(= S^{d+1,\cdot}_{n-1,s})$ | cost $U$; |
| predict shield $c_j$ | positive: | singleton | value $0$; |
| predict shield $c_j$ | negative: | $S^{d,\mathbf{b}}_{n,s-1}$ | cost $V$. |

The second line holds *even when the true hypothesis is a shield*: negative feedback on a binary prediction does not reveal the true label. At a *diagnostic* point with core label $b$, predicting $b$ leads on positive feedback to the pure $n$-cube (value $\frac{n}{2}$) and on negative feedback to the pure shield class (value $\frac{s-1}{2}$, cost $W := \frac{s+1}{2}$). Predicting a shield $c_j$ behaves as at a free point.

**Symmetrization.** Let $(\alpha_0, \alpha_1, \beta_1, \ldots, \beta_s)$ be an arbitrary feasible prediction distribution at a free point and write $p = \alpha_0 + \alpha_1$. The pure-label cost rows are

$$C_0 = \alpha_0 t + \alpha_1 U + (1 - p)V, \quad C_1 = \alpha_1 t + \alpha_0 U + (1 - p)V, \quad C_{c_j} = pU + \textstyle\sum_{\ell \neq j} \beta_\ell V.$$

Then $\max\{C_0, C_1\} \geq \frac{C_0 + C_1}{2} = f(p) := \frac{p}{2}(t + U) + (1 - p)V$ and $\max_j C_{c_j} \geq \frac{1}{s} \sum_j C_{c_j} = g(p) := pU + \frac{s-1}{s}(1 - p)V$. The symmetric distribution ($\alpha_0 = \alpha_1 = \frac{p}{2}$, $\beta_j = \frac{1-p}{s}$) attains both averages simultaneously. Hence the free-point value is exactly $F = \min_{p \in [0,1]} \max\{f, g\}$.

**Crossing.** By induction and Lemma 2.3, $U \geq 1 + \frac{n-1}{2} > t$ and $U \geq 1 + \frac{s}{2} > \frac{s+1}{2}$. At $p = 0$, $f(0) = V > g(0)$; at $p = 1$, $g(1) = U \geq \frac{t+U}{2} = f(1)$. The two affine lines cross at $p_* = \frac{2V}{s(U-t)+2V} \in (0, 1)$ with value $\frac{V((s+1)U-(s-1)t)}{s(U-t)+2V}$. The piecewise-affine convex maximum is minimized at an endpoint or the crossing, giving (2) for the free-point value $F$.

**Diagnostic domination.** Take the optimal $p^\circ \in \{0, p_*, 1\}$ from the free point and, at a diagnostic point, put mass $p^\circ$ on the core label $b$ and split the rest uniformly over the shields. The two cost rows are $d_{\mathrm{core}}(p) = p\frac{n}{2} + (1 - p)V$ and $d_{\mathrm{sh}}(p) = p\frac{s+1}{2} + \frac{s-1}{s}(1 - p)V$. By $U \geq \frac{n+1}{2} = n - t$ and $U \geq \frac{s+1}{2}$, these are pointwise at most $f(p)$ and $g(p)$ respectively. Thus no diagnostic point exceeds $F$. Since $n \geq 1$, the adversary can always play a free point. After seeing the published distribution, its best-response pure label earns at least $\max\{f, g\} \geq F$. Every such label has a nonempty positive fiber, so (1) holds; recursive subgames use the inductive optimal adversaries. Hence the state value equals $F$ and depends only on $(n, s)$, proving diagnostic invariance and (2) simultaneously.

The single-shield statement is in Appendix F.1. Its specialization proof below refers to Appendix F.4(ii), whose monotonicity reference is the paragraph named "Monotonicity" in Section D.2.

**Single-shield specialization (Theorem F.2).** For $s = 1$, $V = \frac{n+2}{2}$ and the crossing value reads $\frac{2UV}{U-t+2V}$. From $a_{n-1,1} \le \frac{n}{2}$ (the $\frac{n+s}{2}$ bound of Appendix F.4(ii)) we get $U \le V$ and $U+t \le 2V$, so the crossing term is the minimum, giving the $D_m$ recursion. The $\delta$ recursion follows by substituting $a_m = \frac{m}{2} + \delta_m$. Positivity is immediate, and $\delta_m \le \frac{3}{m+1}$ follows for $m \ge 3$ by induction from $\delta_3 = \frac{57}{142}$ and monotonicity of $e \mapsto \frac{2+(m+4)e}{2(m+3+e)}$: the inequality $\frac{5m+12}{2(m^2+3m+3)} \le \frac{3}{m+1}$ reduces to $m^2 + m + 6 > 0$. Finally, $\mathrm{B}(D_m) \ge \frac{m}{2}$ by a uniform binary prior on $m$ distinct points, giving the ratio bound. Diagnostic invariance for $\tilde{D}_m$ is the $s = 1$ case above.

# D SHARP ASYMPTOTICS: FULL PROOFS

## D.1 LOWER BOUND $\mathrm{A}(H_r) \ge \frac{5r}{8}$

Let $T = \lfloor r/2 \rfloor$. The adversary runs a *mixing phase*. Until positive feedback occurs, it plays a fresh binary coordinate and sets $\tau$ uniform on the $q + 2$ currently feasible labels: the two binary values and the $q$ surviving shield colours. Every such label has a nonempty positive fiber, so the adversary is legal. Positive feedback ends the phase, leaving either a singleton or a pure sub-cube. Negative feedback on a binary prediction decrements the free dimension $d \mapsto d - 1$, while negative feedback on a shield prediction decrements $q \mapsto q - 1$. After $T$ all-negative rounds the adversary switches to the pure binary strategy on the remaining $d$ free coordinates, worth $d/2$.

Fix a learner and consider the all-negative path. Subtract the binary baseline, which totals $r/2$. A binary prediction in state $(d, q)$ has excess cost $g_B(q) = \frac{q}{2(q+2)}$ and a shield prediction has excess cost $g_C(d, q) = \frac{q+1-d/2}{q+2}$, which is nonnegative throughout the phase because $q \ge r - T \ge r/2$ and $d \le r$. A direct computation shows that replacing an adjacent (binary, shield) prediction pair by (shield, binary) never increases the learner's excess. The *decrease* is

$$\mathrm{Cost}(BC) - \mathrm{Cost}(CB) = \frac{d+1}{2(q+2)^2} + \frac{R}{(q+2)^2} \;\ge\; 0,$$

where $R \ge 0$ is the excess of the common suffix. Thus an optimal learner may be assumed to play $C^c B^b$ with $c + b = T$. Summing the excesses with $q = r - c$ gives

$$F(c, b) = \frac{c(r+3-c)}{2(r+2)} + \frac{q(q+2)}{2(r+2)}\Big[1 - \Big(\tfrac{q+1}{q+2}\Big)^b\Big],$$

and since $b \le q/2$, Bernoulli's inequality yields $1 - (1 - \frac{1}{q+2})^b \ge \frac{3b}{4(q+2)}$. Hence $F \ge \frac{r}{8}$ reduces to

$$P(c) := 4c(r+3-c) + 3(r-c)(T-c) \;\ge\; r(r+2),$$

a concave quadratic in $c$ whose endpoint values $P(0) = 3rT$ and $P(T) = 4T(r+3-T)$ are both at least $r(r+2)$ for even $r \ge 4$ and odd $r \ge 7$. The remaining $r \in \{1, 2, 3, 5\}$ follow directly from the exact values given by the proven recursion (2): $\mathrm{A}(H_1) = 1$, $\mathrm{A}(H_2) = \frac{27}{14}$, $\mathrm{A}(H_3) = \frac{288504}{103225}$ and the exact rational $a_{5,5}$, reported as $\approx 4.4119$ in Tables 1 and 2, all exceed $\frac{5r}{8}$. Altogether $\mathrm{A}(H_r) \ge \frac{r}{2} + \frac{r}{8} = \frac{5r}{8}$ for all $r \ge 1$.

## D.2 UPPER BOUND $\mathrm{A}(H_r) \le \frac{5r}{8} + \sqrt{6r-3}$

Define the two-piece potential

$$\Phi(n, s) = \begin{cases} \frac{s}{2} + \frac{n^2}{8s}, & s \ge 1,\ n \le 2s, \\ \frac{n}{2}, & \text{otherwise}, \end{cases}$$

which is continuous with matching $n$-slope $\frac{1}{2}$ at $n = 2s$.

*Monotonicity.* Write $\mathcal{W}_s(U, V)$ for the crossing term of (2). In the region $U \ge \frac{n+1}{2}$, $V \ge \frac{n+2}{2}$ we obtain $\partial_U \mathcal{W}_s = \frac{2V((s+1)V-st)}{D^2} \ge 0$ and $\partial_V \mathcal{W}_s = \frac{s(U-t)\,N}{D^2} \ge 0$, where $D = s(U-t) + 2V$, $N = (s+1)U - (s-1)t$. Induction gives $a_{n,s} \ge \frac{n}{2}$: applying the recursion to the floor values gives $\frac{n}{2} + \min\{\frac{1}{2}, \frac{s}{n+s+2}\} \ge \frac{n}{2}$. The second term is the crossing value, and the endpoint $U$ caps it at $\frac{n}{2} + \frac{1}{2}$. Thus inflating both children can only increase the recursion output.

*Shifted supersolution.* Claim: for $n, s \geq 1$ and any shift $e \geq 0$ with $e^2 \geq 3s$,

$$\mathcal{W}_s\big(1 + \Phi(n-1, s) + e,\ 1 + \Phi(n, s-1) + e\big) \leq \Phi(n, s) + e.$$

Substitute $U = U_0 + e$, $V = V_0 + e$, $B = G + e$, where $G = \Phi(n, s)$, $U_0 = 1 + \Phi(n-1, s)$, $V_0 = 1 + \Phi(n, s-1)$. Using $(s+1)U - (s-1)t = s(U-t) + U + t$, we can write the claim as $Q(e) := s(p_0 + e)\alpha + (V_0 + e)(\beta - e) \leq 0$ where $p_0 = U_0 - t$, $\alpha = V_0 - G$, $\beta = U_0 + t - 2G$. Equivalently, $Q(e) = C_0 + C_1 e - e^2$ with $C_0 = sp_0\alpha + V_0\beta$, $C_1 = s\alpha + \beta - V_0$. The boundary cases evaluate exactly to

| case | $C_0$ | $C_1$ |
|---|---|---|
| $n \geq 2s+1$ | $s$ | $s - \frac{n+2}{2} \leq -\frac{3}{2}$ |
| $n = 2s$ | $s + \frac{1}{4} + \frac{1}{8s}$ | $-1 + \frac{1}{8s}$ |
| $n = 2s-1$ | $s + \frac{5}{8} + \frac{1}{16s}$ | $-\frac{5}{8} + \frac{1}{4s}$ |

Each row is valid for all $s \geq 1$. In particular, the case $s = 1$, where only $n \geq 2s-1$ occurs, is covered by these rows. In the inner region $n \leq 2s-2$, which forces $s \geq 2$, we have $\alpha = \frac{1}{2} + \frac{n^2}{8s(s-1)} \in [\frac{1}{2}, 1]$, $(s-1)\alpha - G = -\frac{1}{2}$, hence $C_1 = \beta - \frac{1}{2} \leq 0$ and $C_0 = s\alpha - \frac{h}{2} + u((s+1)\alpha + G) \leq 2s + \frac{1}{2}$ with $h = \frac{(2s-n)^2}{8s}$, $u = \frac{2(2s-n)+1}{8s} < \frac{1}{2}$. In every case $C_1 \leq 0$ and $C_0 \leq 3s$, so $Q(e) \leq 3s - e^2 \leq 0$ whenever $e^2 \geq 3s$.

*Induction.* Take $e = \sqrt{3(n+s-1)}$, so $e^2 \geq 3s$. Strong induction on $n + s$, with exact boundary rows $a_{n,0} = \Phi(n, 0)$, $a_{0,s} = \Phi(0, s)$, yields $a_{n,s} \leq \Phi(n, s) + \sqrt{3(n+s-1)}$ for all $n + s \geq 1$. On the diagonal, this is $\mathrm{A}(H_r) \leq \frac{5r}{8} + \sqrt{6r-3}$. The constant $\frac{1}{8}$ is critical: for $c < \frac{1}{8}$ the form $\frac{1}{2} + cx^2 - \frac{x}{2}$ becomes negative and no quadratic potential of this shape exists.

### D.3 Oblivious lower bound

Assume $r \geq 16$ and set $L = \log_2 r$, $p = 1/L$. Fix the infinite cyclic sequence $x_1, \ldots, x_r, x_1, \ldots$ and the following prior: with probability $p$ choose a uniform binary core hypothesis; otherwise choose a uniform shield. Fix a deterministic learner of finite risk and follow its *all-negative path*, the unique history in which every feedback is negative. Among the first $r$ rounds of this path let $b$ be the number of binary predictions and $u$ the number of *distinct* shield colours predicted, so that exactly $r-b-u$ of these rounds are repeated shield probes. On the whole (infinite) path, every shield colour is eventually probed, since otherwise the constant sequence with that shield as target would force unbounded loss. Thus the *first-probe order* of all $r$ colours is well defined. Let $\mathrm{rank}(c)$ be the position of colour $c$ in it. For each colour $c$ let $\tilde{d}_c$ be the number of binary predictions preceding the first probe of $c$, except that $\tilde{d}_c := b$ if $c$ is not probed within the first $r$ rounds. For such $c$ the true count is at least $b$, so $\tilde{d}_c$ is a valid lower bound. Put $x = \frac{1}{r}\sum_c \tilde{d}_c$.

*Binary world.* Each of the $r$ points is a fresh independent fair bit, so the first binary prediction at each point errs with probability $\frac{1}{2}$ and the binary predictions contribute at least $\frac{r}{2}$ in expectation. A point that never receives a binary prediction incurs an error on every one of its cyclic recurrences, which is even worse. Every shield probe errs in this world. A first probe of colour $c$ occurring within the first $r$ rounds is preceded by $\tilde{d}_c$ binary predictions and is actually reached with probability $2^{-\tilde{d}_c}$, and each of the $r - b - u$ repeated probes is reached with probability at least $2^{-b}$. Count the colours not probed within $r$ rounds at their surrogate value $\tilde{d}_c = b$, a bookkeeping term rather than a reach probability. Then

$$\mathbb{E}_{\mathrm{bin}}[\text{shield errors}] \ \geq \sum_{c \text{ probed in } \leq r \text{ rounds}} 2^{-\tilde{d}_c} + (r-b-u)\,2^{-b} \ = \ \sum_{c=1}^{r} 2^{-\tilde{d}_c} - b\,2^{-b} \ \geq \ r2^{-x} - \tfrac{1}{2},$$

where the identity holds because exactly $r - u$ colours carry the surrogate value. The last step follows from convexity of $z \mapsto 2^{-z}$ and $b2^{-b} \leq \frac{1}{2}$. Thus $\mathbb{E}_{\mathrm{bin}}[M] \geq \frac{r}{2} + r2^{-x} - \frac{1}{2}$.

*Shield world.* If the true shield is colour $c$, the game follows the all-negative path until the first probe of $c$, and every round before that probe errs: at least once for each of the $\mathrm{rank}(c) - 1$ colours probed earlier, and at least $\tilde{d}_c$ times on binary predictions. For colours first probed after round $r$,

the surrogate $b$ is again a lower bound. Hence $M(c) \geq \mathrm{rank}(c) - 1 + \tilde{d}_c$. Since the ranks are a permutation of $\{1, \ldots, r\}$, averaging over the uniform shield gives $\mathbb{E}_{\mathrm{const}}[M] \geq \frac{r-1}{2} + x$.

Combining these bounds, every deterministic learner suffers $\mathbb{E}[M] \geq \frac{r}{2} - \frac{1}{2} + pr2^{-x} + (1-p)x \geq \frac{r}{2} - \frac{1}{2} + \min_{x \geq 0}\{pr2^{-x} + (1-p)x\}$, and continuous minimization, with integrality of the $\tilde{d}_c$ only weakening the bound, gives $(1 - \frac{1}{L})\big[L - \log_2(L-1) + \log_2 \ln 2 + \frac{1}{\ln 2}\big] \geq L - \log_2 L - 1$. Averaging over the learner's randomness and applying the probabilistic method inside the prior's (finitely supported) realizable sequences,

$$\mathrm{B}(H_r) \;\geq\; \frac{r}{2} + \log_2 r - \log_2 \log_2 r - \frac{3}{2} \qquad (r \geq 16).$$

### D.4 Oblivious upper bound and the ratio

The threshold learner of Appendix B.3 generalizes as follows. The rule covers arbitrary fixed input sequences, including repeating ones. Fix a (possibly random) threshold $J \geq 1$. Maintain an independent fair guess $b_x \in \{0, 1\}$ for each point, flipped whenever a binary prediction at $x$ errs. Predict $b_{x_t}$ until $J$ binary predictions have been made, *repeated coordinates included*. If any of them is correct, the target is binary: discard all shields and track bits forever. If all $J$ err, probe the $r$ shields once each in uniformly random order. A correct probe identifies the target. If all $r$ probes fail, every shield is eliminated, and bit tracking resumes.

*Binary target.* Bit tracking loses at most $\frac{1}{2}$ per point in expectation: the first visit is a fair guess, and after a miss the flip makes all later visits correct. This gives an expected total of at most $\frac{r}{2}$ across the whole game, before and after any probe phase. If the first $J$ binary predictions contain a repeated coordinate, the earliest repetition is already flipped to the correct bit, that prediction is correct, and the probe phase never starts. Thus it starts only if $J$ *distinct* coordinates all miss on their first visit, an event of probability at most $2^{-J}$, and then costs at most $r$ probe errors. Hence $\mathbb{E}[M] \leq \frac{r}{2} + r\,\mathbb{E}[2^{-J}]$.

*Shield target.* Every binary prediction errs, since both binary labels are wrong. The probe phase therefore starts after exactly $J$ binary errors, and the uniformly random probe order pays $\frac{r-1}{2}$ expected errors before hitting the target. Thus $\mathbb{E}[M] \leq \mathbb{E}[J] + \frac{r-1}{2}$.

Thus $\mathrm{B}(H_r) \leq \max\{\frac{r}{2} + r\mathbb{E}[2^{-J}],\ \frac{r-1}{2} + \mathbb{E}[J]\}$ for every threshold distribution. The same argument gives, for $D_n^{(s)}$,

$$\mathrm{B}\big(D_n^{(s)}\big) \;\leq\; \max\Big\{\tfrac{n}{2} + s\,\mathbb{E}[2^{-J}],\ \tfrac{s-1}{2} + \mathbb{E}[J]\Big\},$$

which is the upper-bound half of every exact oblivious value of a class $D_n^{(s)}$ reported in this paper, as detailed in Appendix G. The four-function class $H_{\mathrm{mid}}$ is instead closed by its bespoke learner in Appendix B.1. With $q = L - \log_2 L$ and the deterministic $J = \lceil \log_2(r/q) \rceil$ we get $r2^{-J} \leq q$ and $J - \frac{1}{2} \leq q + \log_2(L/q) + \frac{1}{2} \leq q + \frac{3}{2}$, whence $\mathrm{B}(H_r) \leq \frac{r}{2} + \log_2 r - \log_2 \log_2 r + \frac{3}{2}$ for $r \geq 16$. The cruder choice $J = \lceil \log_2(r+1) \rceil$ gives $\mathrm{B}(H_r) \leq \frac{r}{2} + \lceil \log_2(r+1) \rceil - \frac{1}{2}$ for all $r \geq 2$. Together with $\mathrm{B} \geq \frac{r}{2}$ and the two-sided bounds on A this proves Theorems 4.3 and 4.4.

**A finite tower above the limiting ratio.** Applying (2) on the grid $Q^{-1}\mathbb{Z}$, with $Q = 10^{12}$ and outward rounding at each state, gives

$$229.003712397003 \leq \mathrm{A}(H_{352}) \leq 229.003712397146.$$

For completeness, the rounding preserves a rigorous enclosure. Let $T_{n,s}(U, V)$ denote the minimum in (2). The boundary values lie exactly on the grid. The lower grid value at $(n, s)$ remains at least $n/2$: at the smallest admissible inputs,

$$T_{n,s}\left(\frac{n+1}{2}, \frac{n+2}{2}\right) = \frac{n}{2} + \min\left\{\frac{1}{2}, \frac{s}{n+s+2}\right\} \geq \frac{n}{2},$$

and $n/2$ itself lies on the grid. Thus downward rounding preserves this floor, so both lower child inputs stay in the monotonicity region of Section D.2. Induction on $n + s$ then shows that evaluating at the lower and upper child bounds, followed by outward rounding, encloses the exact value. All

certified operations use integers. Taking $J = 6$ in the threshold bound above yields $\mathrm{B}(H_{352}) \le \max\{176 + 352/64, 351/2 + 6\} = 363/2$. Consequently,

$$\frac{\mathrm{A}(H_{352})}{\mathrm{B}(H_{352})} \ge \frac{76334570799001}{60500000000000} > 1.2617 > \frac{5}{4}.$$

This certifies a lower bound on the finite ratio, not its exact value or the maximum over towers. It is consistent with the limit in Theorem 4.4. The supplementary script `verify_tower_intervals.py` reproduces the enclosure and the threshold bound.

# E Localization and necessary conditions

## E.1 Proofs of Theorems 5.2 and 5.3

Run the uniform first-hit strategy of Lemma 5.1. Rounds with $|Y_V(x_t)| = 1$ cost nothing and leave $V$ unchanged. In every other round the hit probability is exactly $1/|Y_V(x_t)| \ge 1/q(H)$ regardless of $\tau_t$. This follows from (1) and the fact that $\tau_t$ is chosen before $\hat{y}_t$ is sampled. A per-history hazard argument then shows that the mistakes before the first informative positive feedback are stochastically dominated by a geometric variable of mean $q(H) - 1$. Upon the first positive feedback $(x, y)$ the new version space is $V' = V_{x,y}$. In the block-pure case $y$ lies in exactly one $Y_i$, so $V' \subseteq H_i$. In general $V' \subseteq \bigcup_{i \in J(x,y)} H_i = K_{J(x,y)}$ with $1 \le |J(x, y)| \le \Delta$. The continuation is bounded by $\mathrm{A}(V') \le \mathrm{A}(H_i)$ (resp. $\mathrm{A}(K_{J(x,y)})$) via the restart-and-monotonicity argument of Appendix A: the remaining adversary, restricted to histories extending the prefix, is a legal adversary for the subclass, and an $\varepsilon$-optimal continuation strategy with fresh randomness completes the bound. Let $\varepsilon \downarrow 0$. Finally $q(H) - 1 \le 2\mathrm{B}(H)$ by Lemma 2.3(a), with the zero law covering $q = 1$ atoms, and $\mathrm{B}(H_i) \le \mathrm{B}(H)$, giving the $(C_0 + 2)$ forms. For the depth claim, in a *finite-height* nested hierarchy in which every node is the union of its children and the children's blocks partition the parent's block, every hypothesis descends along its branch to a leaf. The label blocks of any two distinct *leaves* are disjoint, since they separate at their lowest common ancestor. Thus the root class is the flat disjoint-block union of its leaves, and Theorem 5.2 applies once. Finite height is essential here as well: an infinite single-branch chain of copies of a class $K$ below one child of the root never contributes a leaf, so the leaf union can miss $K$ entirely and the leaf-based bound fails for $K \cup \{\bar{2}\}$ with $K$ a large cube.

## E.2 Proof of Theorem 5.4

If $\mathrm{B}(H) > 0$ then $q \ge 2$ by the zero law. For each cluster, either $\mathrm{B}(K_J) = 0$, in which case $\mathrm{A}(K_J) = 0$, or $\mathrm{A}(K_J) \le \Gamma_\Delta \mathrm{B}(K_J) \le \Gamma_\Delta \mathrm{B}(H)$ by monotonicity. Substitution into Theorem 5.3 gives $\mathrm{A}(H) \le 2\mathrm{B}(H) + \Gamma_\Delta \mathrm{B}(H)$.

**Informative high-value chains (the existential statement of Open Problem 6.1).** For every $V \subseteq H$ with $q(V) \ge 2$, the first-hit strategy gives $\mathrm{A}(V) \le q(V) - 1 + \sup\{\mathrm{A}(V_{x,y}) : |Y_V(x)| \ge 2,\ y \in Y_V(x)\}$. Rounds with $|Y_V(x)| = 1$ cost nothing and leave $V$ unchanged, and the first informative positive feedback leaves $W_{x,y} \subseteq V_{x,y}$ for an informative $(x, y)$. Hence, if $\mathrm{A}(H) = \rho \mathrm{B}(H)$ with $0 < \mathrm{B}(H) =: \beta < \infty$ and $0 < \varepsilon \le \beta$, choosing at each step an informative fiber within $\varepsilon$ of the supremum yields states $V_0 = H, V_1, V_2, \ldots$ with $\mathrm{A}(V_t) \ge \mathrm{A}(H) - t(2\beta + \varepsilon) \ge \mathrm{A}(H) - 3t\beta$, using $q(V_t) - 1 \le 2\mathrm{B}(V_t) \le 2\beta$. For $t \le \lfloor \rho/6 \rfloor$ this stays $\ge \mathrm{A}(H)/2 > 0$, so each $V_t$ still has an informative point and $\mathrm{B}(V_t) \le \beta$. This gives a legal chain of $\Omega(\rho)$ informative positive feedbacks with high adaptive value throughout.

## E.3 Proof of Theorem 5.5

Ignore rounds with $|Y_V(x)| = 1$, which cost nothing and leave the version space unchanged along the relevant branch. While $\lambda > 0$, each charging phase costs at most $q - 1$ in expectation before its first informative positive feedback by Lemma 5.1. Negative feedback never increases $\lambda$. Each phase-ending positive feedback strictly decreases the integer $\lambda$, so at most $L$ phases occur. Once $\lambda = 0$ the state is inside an atom and the continuation costs at most $\max_i \mathrm{A}(H_i)$. For the positive-halving corollary define $\lambda(V) = 0$ if the covering number $\nu(V) = 1$, $\lambda(V) = \lceil \log_2 \nu(V) \rceil$ if

$2 \le \nu(V) \le \Delta$, and $1 + \lceil \log_2 \Delta \rceil$ otherwise. Subclasses never increase $\nu$. The first informative positive feedback lands in a cluster with $\nu \le \Delta$, and halving with $\lceil \log_2 \lceil n/2 \rceil \rceil \le \lceil \log_2 n \rceil - 1$ gives $L \le 1 + \lceil \log_2 \Delta \rceil$. For overlap-closed hierarchies we prove the depth-free form directly. First, *every root fiber is contained in a leaf*: fix $(x, y)$ with $F = H_{x,y} \neq \emptyset$. The root's cover has a child $N_1$ containing the incidence cluster $K_{J(x,y)} \supseteq F$. Since $F \subseteq N_1$, the fiber of $N_1$ at $(x, y)$ is again exactly $F$, so the definition applies at $N_1$. Iterating down the finite tree yields a leaf $L \supseteq F$. Now run the first-hit strategy at the root. Its expected cost is bounded by $q(H) - 1 \le 2\mathrm{B}(H)$ by Lemma 2.3(a), with the zero law covering $q(H) = 1$. The first informative positive feedback $(x, y)$ leaves $V' \subseteq H_{x,y} \subseteq L$ for some leaf $L$, and monotonicity gives $\mathrm{A}(V') \le \mathrm{A}(L) \le C_0 \mathrm{B}(L) \le C_0 \mathrm{B}(H)$, whence $\mathrm{A}(H) \le (C_0 + 2)\mathrm{B}(H)$. Thus one positive feedback already reaches a leaf and the hierarchy adds no depth cost. A hierarchy notion under which depth genuinely costs would require strictly weaker closure conditions, and we leave its formulation open. Finiteness of the height is essential for the fiber-descent step: an infinite spine of nodes all labeled $H$, each with one $H$-child and one leaf child $\{h_0\}$ for a fixed $h_0 \in H$, satisfies the closure condition: every cluster is contained in the $H$-child. It has $C_0 = 0$, yet for $H = \{0, 1\}^{[4]}$ the claimed cap $q - 1 + \max_L \mathrm{A}(L) = 1$ would contradict $\mathrm{A}(H) = 2$. The descent never reaches a leaf there.

### E.4 Proof of Theorem 5.7

The Hamming-ball formula specializes the binary characterization of Filmus et al. (2023).

*Exact formula.* Since $q(E(n, d)) \le 2$, bandit feedback reveals the true label. With $r$ unqueried coordinates and residual budget $b$, the value $v(r, b)$ satisfies $v(r, 0) = 0$, $v(r, r) = \frac{r}{2}$ and, for $1 \le b < r$,

$$v(r, b) = \min_{p \in [0,1]} \max\{p + v(r - 1, b),\ 1 - p + v(r - 1, b - 1)\},$$

where consistency of both adversary branches is witnessed by padding with zeros. By induction $v(r - 1, b) - v(r - 1, b - 1) = \Pr[Z_{r-1} \ge b] \in [0, 1]$, so the two lines balance at an interior $p^*$ and $v(r, b) = \frac{1 + v(r-1,b) + v(r-1,b-1)}{2}$. The closed form $v(r, b) = \frac{r}{2} - \mathbb{E}[(Z_r - b)_+]$ then follows from $\mathbb{E}[(Z_r - b)_+] = \frac{1}{2}\mathbb{E}[(Z_{r-1} - b)_+] + \frac{1}{2}\mathbb{E}[(Z_{r-1} - b + 1)_+]$. The needed a-priori bound $v(r, b) \le v(r, b - 1) + 1$ comes from running the smaller-budget strategy and paying one extra mistake at the first budget violation, after which all remaining labels are forced. Revisiting a resolved coordinate or stopping early never helps the adversary: revisits are answered without error, and $v$ is monotone, $v(r, b) - v(r - 1, b) = \frac{1}{2} - \frac{1}{2}\Pr[Z_{r-1} \ge b] \ge 0$ by the closed form.

*Surcharge.* $\{0, 1\}^{[2m]} = E^- \cup E^+$ with $E^- = E(2m, m)$ and $E^+$ its label-swapped mirror. Each point is incident to both layers through both labels, so $q = 2$, $\Delta = 2$, and every incidence set $J(x, y)$ equals $\{-, +\}$, a trivially laminar family. The cube has value $m$ by Lemma 2.3(d), while $\mathrm{A}(E^{\pm}) = m - \mathbb{E}[(Z_{2m} - m)_+]$. The telescoped identity $(j - m)\binom{2m}{j} = m[\binom{2m-1}{j-1} - \binom{2m-1}{j}]$ evaluates $W_m = \mathbb{E}[(Z_{2m} - m)_+] = \frac{m}{2}\binom{2m}{m}4^{-m}$. The Wallis-type induction $\binom{2m}{m}4^{-m} \ge \frac{1}{2\sqrt{m}}$, using the ratio comparison $\frac{2m+1}{2m+2} > \sqrt{\frac{m}{m+1}}$, gives $W_m \ge \frac{\sqrt{m}}{4}$. Since bandit feedback with $k = 2$ equals full information, the surcharge is purely a statistical model-selection cost.

*A $k = 3$ variant with all three labels present.* Let $H = \{0, 1\}^{[2m]} \cup \{\bar{2}\}$, covered by $H_{\pm} = E^{\pm} \cup \{\bar{2}\}$. Then $q = 3$, $\Delta = 2$, and every incidence set equals $\{-, +\}$, so the family is laminar. By monotonicity and Appendix F.4(iii), which shows that adding one constant shield on a fresh label costs at most 1, $m - W_m \le \mathrm{A}(H_{\pm}) \le m - W_m + 1$ and $m \le \mathrm{A}(H) \le m + 1$, whence

$$W_m - 1 \ \le\ \mathrm{A}(H) - \max\{\mathrm{A}(H_-), \mathrm{A}(H_+)\} \ \le\ W_m + 1,$$

the same $\Theta(\sqrt{m})$ surcharge.

**Remark E.1** (The rank bound is not tight)**.** *For the two-layer Hamming cover $\{0, 1\}^{[2m]} = E^- \cup E^+$ of Theorem 5.7, the minimal feasible localization rank is exactly $2m - 1$, while the true additive surcharge is $W_m = \Theta(\sqrt{m})$. The certificate of Theorem 5.5 can overshoot the actual cost by an unbounded factor: $L_*(q - 1)/W_m = \Theta(\sqrt{m})$ with $q = 2$.* Lower bound $L_* \ge 2m - 1$: *every reachable version space is a subcube with $a$ ones and $b$ zeros fixed. Querying $2m - 2$ fresh coordinates along a legal all-positive path with $a = b = m - 1$ keeps the version space off both atoms, since it still contains total weights $m - 1$ and $m + 1$. Each of these positive feedbacks occurs*

*at a point with two feasible labels, so any valid rank strictly decreases $2m-2$ times and is still $\geq 1$ at the end.* Upper bound: *set $\lambda(V)=0$ if $a \geq m$ or $b \geq m$, in which case $V \subseteq E^+$ resp. $E^-$. Set $\lambda(V)=2m-1-a-b$ otherwise. This is a valid rank: since $q=2$, while $\lambda>0$ every informative feedback, whether positive or negative, fixes a fresh coordinate and strictly decreases $\lambda$. Once $\lambda=0$ the state is inside an atom and $\lambda$ stays 0. Finally, $\lambda(\{0,1\}^{[2m]})=2m-1$.*

### E.5 Proof of the necessary conditions

*Proof of Theorem 5.8.* (i) By Filmus et al. (2024, Theorem D.1), $\mathrm{A} \leq 6k\,\mathrm{F}$, so $ck\mathrm{B} \leq 6k\mathrm{F}$ gives $\mathrm{F} \geq \frac{c}{6}\mathrm{B}$. Next, $\mathrm{A} \leq 6q\,\mathrm{F}$ (Appendix F.4(i)) together with $\mathrm{F} \leq \mathrm{B}$ yields $ck\mathrm{B} \leq 6q\mathrm{B}$, or $q \geq \frac{ck}{6}$. Then $\mathrm{B} \geq \frac{q-1}{2} \geq \frac{ck}{12} - \frac{1}{2}$ (Lemma 2.3(a)) and $\mathrm{F} \geq \frac{c}{6}\mathrm{B}$ give all three claims.

(ii) The halving algorithm bounds the deterministic full-information mistake number by $\log_2|H|$, so $\mathrm{F} \leq \log_2|H|$ and $|H| \geq 2^{\mathrm{F}} \geq 2^{\Omega(k)} = k^{\Omega(k/\log k)}$.

(iii) By Theorem 5.5, Lemma 2.3(a) and the cover premise, $\mathrm{A} \leq L(q(H)-1) + \max_i \mathrm{A}(H_i) \leq (2L+C_0)\mathrm{B}(H)$, whence $ck \leq 2L+C_0$ and $L \geq \frac{ck-C_0}{2}$. In positive-halving covers one may take $L \leq 1+\lceil\log_2\Delta\rceil$ (Theorem 5.5), so the same chain gives $ck \leq C_0 + 2(1+\lceil\log_2\Delta\rceil)$, hence $\lceil\log_2\Delta\rceil \geq \frac{ck-C_0}{2} - 1$ and $\log_2\Delta \geq \frac{ck-C_0}{2} - 2$.

For the $\rho$-parameterized forms, the same two steps give $q \geq \frac{\rho}{6}$ and $\mathrm{B} \geq \frac{q-1}{2} \geq \frac{\rho}{12} - \frac{1}{2}$. For the class size, split into two cases. If $|H| \geq k$, the expert bridge (Appendix F.4(v)) gives $\rho\mathrm{B} = \mathrm{A} \leq Ck\log_k(k|H|) = Ck(1+\log_k|H|) \leq 2Ck\log_k|H|$, so $\log_k|H| \geq \rho\mathrm{B}/(2Ck)$. If $2 \leq |H| \leq k$, then Lemma 2.3(b) and the elementary inequality $\frac{x-1}{2} \leq k\log_k x$ for $2 \leq x \leq k$ apply. To see the latter, $x \mapsto \frac{\ln x}{x-1}$ decreases, so the ratio of the two sides is minimized at $x=k$, where it is $\frac{2k}{k-1} > 1$. Together these give $\rho\mathrm{B} = \mathrm{A} \leq \frac{|H|-1}{2} \leq k\log_k|H|$, so $\log_k|H| \geq \rho\mathrm{B}/k$. In both cases $\log_k|H| = \Omega(\rho\mathrm{B}/k)$. □

## F Further exact families and comparison tools

### F.1 The single-shield family

**Proposition F.1** (Single-shield family; exact ratios)**.** *For $D_m = \{0,1\}^{[m]} \cup \{\bar{2}\}$ ($k=3$): $\mathrm{A}(D_3) = \frac{135}{71}$, $\mathrm{B}(D_3) = \frac{9}{5}$; $\mathrm{A}(D_4) = \frac{2472}{1051}$, $\mathrm{B}(D_4) = \frac{20}{9}$. The certified ratios $\frac{80}{77}, \frac{75}{71}, \frac{5562}{5255}$ ($m=2,3,4$) increase in this range. Nevertheless, the family is asymptotically degenerate (Theorem F.2): $\mathrm{A}(D_m)/\mathrm{B}(D_m) \to 1$.*

**Theorem F.2** ($D_m$ recursion and degeneracy)**.** *Let $a_m = \mathrm{A}(D_m) = a_{m,1}$. Then $a_1 = 1$ and $a_m = \frac{2(m+2)(1+a_{m-1})}{m+7+2a_{m-1}}$ for $m \geq 2$, reproducing $a_2 = \frac{16}{11}$, $a_3 = \frac{135}{71}$, $a_4 = \frac{2472}{1051}$ and predicting $a_5 = \frac{3523}{1254}$, independently confirmed. Writing $a_m = \frac{m}{2} + \delta_m$,*

$$\delta_m = \frac{2+(m+4)\delta_{m-1}}{2(m+3+\delta_{m-1})}, \qquad 0 < \delta_m \leq \frac{3}{m+1} \quad (m \geq 3),$$

*and since $\mathrm{B}(D_m) \geq \frac{m}{2}$, $1 \leq \frac{\mathrm{A}(D_m)}{\mathrm{B}(D_m)} \leq 1 + \frac{6}{m(m+1)}$ for $m \geq 3$, so the ratio tends to 1. Moreover, any number of diagnostic points leaves the adaptive value unchanged.*

Thus the certified ratio increase in Proposition F.1 is a small-$m$ phenomenon, and the peak is *global*: $\max_{m\geq 1} \mathrm{A}(D_m)/\mathrm{B}(D_m) = \frac{5562}{5255}$, attained only at $m=4$. Exact certificates for $\mathrm{B}(D_5), \ldots, \mathrm{B}(D_8)$ and the tail argument via $\mathrm{A}(D_9)$ and Theorem F.2 are in Appendix G. A closed form for $\mathrm{B}(D_m)$ remains open.

The specialization proof is the final paragraph of Appendix C. Appendix G contains the finite certificates and the global-peak argument.

### F.2 A larger finite witness

**Proposition F.3** (A larger constant witness)**.** *A 13-function class $C_{\mathrm{a}}$ on six points with $k=3$, described in Appendix F.3, satisfies $\mathrm{A}(C_{\mathrm{a}}) = \frac{42592}{20137}$, $\frac{1688}{867} \leq \mathrm{B}(C_{\mathrm{a}}) \leq 2$, and $\frac{\mathrm{A}}{\mathrm{B}}(C_{\mathrm{a}}) \in$*

$[\frac{21296}{20137}, \frac{4615908}{4248907}] \approx [1.0576, 1.0864]$. *The lower bound on* $\mathrm{B}$ *has an explicit finite certificate: a rational-weight mixture of* 46 *realizable labeled sequences whose exact Bayes value is* $\frac{1688}{867}$ *by Appendix F.3. The data and checker are in the supplementary material.*

### F.3 $C_{\mathrm{a}}$: DEFINITION AND ANATOMY

The domain is $\{x_0\} \cup B_0 \cup B_1$ with $|B_0| = 3$, $|B_1| = 2$, and $Y = \{0, 1, 2\}$. The class has three fibers:

$$\underbrace{\{h:\ h(x_0) = 0,\ h|_{B_0} \in \{0,1\}^3,\ h|_{B_1} \equiv 0\}}_{F_0\ (8)} \cup \underbrace{\{h\colon h(x_0) = 1,\ h|_{B_0} \equiv 1,\ h|_{B_1} \in \{0,1\}^2\}}_{F_1\ (4)} \cup \{\bar{2}\}.$$

The value $\mathrm{A}(C_{\mathrm{a}}) = \frac{42592}{20137}$ has strong-duality certificates at every state, reproducible with the submitted script. The bounds $\frac{1688}{867} \le \mathrm{B}(C_{\mathrm{a}}) \le 2$ are proved below. The two non-shield fibers *reuse* the labels $0, 1$: at every $B_0$ point, fiber 0 can output 1 while fiber 1 always does. Thus one positive feedback bit does not localize the fiber *with respect to this three-fiber cover*, and the uniform first-hit cap of Theorem 5.2 does not apply to it. The theorem does apply to the coarser two-block regrouping $C_{\mathrm{a}} = (F_0 \cup F_1) \cup \{\bar{2}\}$, whose blocks use the disjoint label alphabets $\{0, 1\}$ and $\{2\}$. Since $\mathrm{A}(F_0 \cup F_1) = \mathrm{B}(F_0 \cup F_1) \le \mathrm{B}(C_{\mathrm{a}})$ (Appendix F.4(vi) and monotonicity), it yields $\mathrm{A}(C_{\mathrm{a}}) \le q - 1 + \mathrm{B}(C_{\mathrm{a}}) \le 3\mathrm{B}(C_{\mathrm{a}})$. This is consistent with, and far weaker than, the certified values. For the three-fiber cover, $\Delta = 2$. The sub-union $F_0 \cup \{\bar{2}\}$ is, up to diagnostic points, a copy of $D_3$ (ratio $\frac{75}{71}$), and the certified ratio of $C_{\mathrm{a}}$ exceeds $\frac{75}{71}$ only modestly. However, $F_0 \cup \{\bar{2}\}$ is *not* an incidence cluster of this cover: no realized $K_{J(x,y)}$ equals this union. Positive feedback may leave a version space *contained* in it, e.g. the fiber $F_0$ at $x_0$, but never delivers it as a cluster. The comparison is therefore a numerical observation about where the separation resides, not a one-hit localization argument. This makes $C_{\mathrm{a}}$ the prototype of the label-reuse regime of Open Problem 6.1.

**Lower bound $\mathrm{B}(C_{\mathrm{a}}) \ge \frac{1688}{867}$: an explicit mixture certificate.** An oblivious adversary may commit to a *distribution* $\mu$ over realizable labeled sequences, so for every $\mu$, $\mathrm{B}(C_{\mathrm{a}}) \ge \inf_{\text{learner}} \mathbb{E}_{\sigma\sim\mu}[\#\text{mistakes}]$. The right-hand side is computable exactly by a posterior-support recursion in which the learner observes both the revealed instance $x_t$ and the feedback bit. At each round, the support first splits by the observed instance; the learner's best prediction then splits it by feedback. Predicting a label absent from the support is dominated, by superadditivity of the unnormalized Bayes risk. Our certificate is a mixture of 46 explicit labeled sequences, each of length 15 and realizable by a member of $C_{\mathrm{a}}$ by construction, with rational weights. Its exact value is $\frac{1688}{867} \approx 1.9469$. For comparison, the best *product* certificate we know, using one prior and one fixed instance sequence, gives $\mathrm{B}(C_{\mathrm{a}}) \ge \frac{19}{10}$. On the fixed sequence $b_1b_2b_3b_1b_2b_3c_1c_2$ (writing $B_0 = \{b_1, b_2, b_3\}$, $B_1 = \{c_1, c_2\}$), put mass $\frac{1}{10}$ on each of the seven $F_0$-functions whose $B_0$-pattern is not 111, mass $\frac{1}{40}$ on each $F_1$-function, and $\frac{1}{5}$ on the shield. The exact Bayes value is $\frac{19}{10}$. We do not know whether a product certificate can reach $\frac{1688}{867}$. The data (sequences and weights) and exact checker are in the supplementary material. The checker also re-verifies realizability of every sequence and both certificates.

**Upper bound $\mathrm{B}(C_{\mathrm{a}}) \le 2$: a two-budget miss-testing learner.** Write $C_{\mathrm{a}} = K \cup \{\bar{2}\}$ with $K$ the 12 binary functions. Since $q(K) = 2$, we have $\mathrm{A} = \mathrm{B} =: \mathrm{R}$ on binary subclasses by Appendix F.4(vi). The learner keeps a pseudo-version $V \subseteq K$ (initially $K$) and a miss budget $j = 2$. While $j > 0$, it predicts binary labels. *Positive* feedback rules out the shield and starts exact binary play on the realized fiber. *Negative* feedback on prediction $a$ at $x$ updates $V \leftarrow V_{x,1-a}$ and decrements $j$. For a binary target, this update is exact, since $q = 2$ forces $h(x) = 1 - a$. At $j = 0$, it probes label 2 once. A hit identifies the shield; a miss eliminates it, and exact binary play resumes on $V$. If the target is the shield, every binary prediction errs and the probe is correct, giving at most 2 mistakes. For a binary target, the guarantee is the recursion $T_0(V) = 1 + \mathrm{R}(V)$,

$$T_j(V) = \sup_x \min_{p\in\Delta(\{0,1\})} \max_{y\in Y_V(x)} \big[p_y\,\mathrm{R}(V_{x,y}) + (1 - p_y)\big(1 + T_{j-1}(V_{x,y})\big)\big],$$

whose $\sup_x$ makes it valid against adaptively chosen instances, in particular against every fixed sequence. With $\mathrm{R}(K) = \frac{7}{4}$ and the fiber values $\mathrm{R} = \big(\frac{3}{2}, 1\big), \big(1, \frac{3}{2}\big), \big(\frac{13}{8}, \frac{1}{2}\big)$ at $x_0$, a $B_0$ point and a $B_1$ point respectively, an exact evaluation gives $T_1$ values $\big(2, \frac{3}{2}\big)$, $\big(\frac{3}{2}, 2\big)$, $\big(\frac{17}{8}, 1\big)$ and local $T_2$

scores 2, 2, $\frac{29}{16}$, whence $T_2(K) = 2$ (machine-checked in the submitted script). Hence $\mathrm{B}(C_{\mathrm{a}}) \le \max\{2,\, T_2(K)\} = 2$.

### F.4 Comparison tools and further caps

For completeness we state and prove the remaining bounds.

(i) $\mathrm{A} \le 6q\,\mathrm{F}$: pointwise label-injection compresses $H$ into $[q]^X$ while preserving $\mathrm{A}, \mathrm{F}$. The $\mathrm{A} \le 6k'\mathrm{F}$ bound of Filmus et al. (2024, Theorem D.1) applies with $k' = q$. Here $\mathrm{F} \le \mathrm{B}$ combines two standard steps. *Full-information adaptive = oblivious:* given any learner $L$, let $L'$ predict, at each $(x, y)$-history, from the marginal law of $L$'s prediction given that history, with fresh randomness each round. Against a fixed labeled sequence the $(x, y)$-history is deterministic, so $L$ and $L'$ incur identical expected loss. Against $L'$ adaptivity is useless, since $L'$'s future behaviour depends only on the public $(x, y)$-history. Its realized predictions carry no information about its future, so its supremal expected loss equals the supremum over deterministic realizable sequences. Hence $\sup_{\mathrm{adap}} \mathrm{loss}(L) \ge \sup_{\mathrm{seq}} \mathrm{loss}(L) = \sup_{\mathrm{seq}} \mathrm{loss}(L') = \sup_{\mathrm{adap}} \mathrm{loss}(L')$, and optimizing over $L$ gives $\mathrm{F} = \mathrm{F}^{\mathrm{obl}}$. Restricting to the public history is essential: a learner whose future depends on its own past sampled predictions can be exploited by an adaptive adversary, though never below the optimal value. *Simulation:* a full-information learner can simulate the bandit feedback bit from the revealed label, so $\mathrm{F}^{\mathrm{obl}} \le \mathrm{B}$. With Lemma 2.3(a) this yields $\mathrm{A} \le 6\mathrm{B}\min\{k, 2\mathrm{B}+1\}$.

(ii) Constant-shield towers: the potential $\frac{n+s}{2}$ is a supersolution of (2). By the monotonicity established in the "Monotonicity" paragraph of Section D.2, it suffices to feed in the inflated children $U = V = 1 + \frac{n+s-1}{2}$, where the crossing term evaluates to $\frac{U(s^2+3s+2n)}{s^2+4s+2n+2} \le U - \frac{1}{2} = \frac{n+s}{2}$. The inequality reduces to $ns \ge s$, giving $\mathrm{A}(D_m^{(s)}) \le \frac{m+s}{2}$. Moreover $\mathrm{B}(D_m^{(s)}) \ge \frac{s+1}{2} + \frac{m-1}{s+2}$: give each shield prior mass $\frac{1}{s+2}$ and the uniform binary core total mass $\frac{2}{s+2}$, and fix the sequence $x_1$ repeated $s+1$ times followed by $x_2, \ldots, x_m$. At $x_1$ the $s+2$ label classes, consisting of $s$ shield colours and the two binary values, are equiprobable. Thus the permutation argument in the proof of Lemma 2.3(a), capped at $s+1$ rounds, gives expected loss at least $\frac{s+1}{2}$ on the first segment. Conditioned on a core target, of mass $\frac{2}{s+2}$, the remaining $m-1$ fresh coordinates stay independent fair bits, adding $\frac{2}{s+2} \cdot \frac{m-1}{2} = \frac{m-1}{s+2}$. Combined with $\mathrm{B} \ge \frac{m}{2}$ from the cube and $\mathrm{A} \le \frac{m+s}{2}$, maximizing over $m$ gives a peak where the two lower bounds cross, at $m = s+3$. Thus $\mathrm{A}/\mathrm{B} \le \frac{2s+3}{s+3} < 2$ over all $(m, s)$.

(iii) Private constant shields: writing $\Sigma_s K$ for $K$ together with $s$ constant functions in $s$ fresh, pairwise distinct labels, $\mathrm{A}(\Sigma_s K) \le \mathrm{A}(K) + s$ by shield-by-shield testing with restart.

(iv) Serial-game operator: for a single point queried repeatedly with feasible-label continuation profile $g = (g_1, \ldots, g_q)$, let $G(g)$ be the adaptive value of the repeated-root game. In its Bellman recursion, a hit on label $i$ continues at value $g_i$, and a miss on a prediction of $i$ removes label $i$. Let $\rho(g) = \max_{\emptyset \ne S \subseteq [q]} [\frac{|S|-1}{2} + \mathrm{avg}_{i \in S}\, g_i]$ be the oblivious scheduling value. Then $G \le 2\rho$. *Proof:* order $g_1 \ge \cdots \ge g_q \ge 0$ and let the learner test the labels deterministically in this order. Each miss removes the tested label, so if the true label is the $i$-th tested one the total cost is $i - 1 + g_i$. By the definition of $\rho$ applied to the top-$i$ subset, $\rho \ge \frac{i-1}{2} + \mathrm{avg}_{j \le i}\, g_j \ge \frac{i-1}{2} + g_i$, whence $i - 1 + g_i \le (i-1) + 2g_i \le 2\rho$ for every $i$. □ At the $q = 3$ profile $(\frac{3}{2}, \frac{3}{2}, 0)$ the ratio $G/\rho = \frac{15}{14}$ is attained: the root matrix, with the learner as rows, is $M = \left(\begin{smallmatrix} 3/2 & 5/2 & 5/2 \\ 5/2 & 3/2 & 5/2 \\ 3 & 3 & 0 \end{smallmatrix}\right)$, and $p = (\frac{3}{7}, \frac{3}{7}, \frac{1}{7})$, $\tau = (\frac{5}{14}, \frac{5}{14}, \frac{2}{7})$ equalize every row and column at $\frac{15}{7}$, while $\rho = 2$, also machine-checked in the submitted script. This profile is *extremal*: $\sup_{q=3} G/\rho = \frac{15}{14}$, uniquely at permutations of $(\frac{3}{2}, \frac{3}{2}, 0)$. *Proof.* Adding a constant $c \ge 0$ to all continuations adds $c$ to both $G$, by induction over $|S|$, and $\rho$, so it suffices to bound normalized profiles $g = (a, b, 0)$, $a \ge b \ge 0$, for which $\rho = \max\{a, \frac{1+a+b}{2}, 1 + \frac{a+b}{3}\}$. The exact two-label value is $G(u, v) = \max\{u, v, \frac{1+u+v}{2}\}$. Row $i$ of the root matrix has diagonal $g_i$ and constant off-diagonal $c_i := 1 + G(g_{-i})$. Put $d_i = c_i - g_i$ and $s = a + b$. If $a \ge c_1$, testing label 1 first gives $G \le a = \rho$. Otherwise all $d_i > 0$ and the learner distribution $p_i \propto 1/d_i$ equalizes every column at $F = \frac{\sum_i c_i/d_i - 1}{\sum_i 1/d_i} \ge G$. Four regions complete the proof. (1) $a \le 1$: here $c_i = \frac{3+s-g_i}{2}$, $\rho = 1 + \frac{s}{3}$ and $c_i - \rho = \frac{d_i}{3}$, so $\sum_i \frac{c_i - \rho}{d_i} = 1$, or $F = \rho$. (2)

$b \geq 1$, so $d := a - b < 1$. An exact simplification gives $F = \frac{(s+2-d^2)(s+3)}{2(s+4-d^2)}$, decreasing in $d^2$, while $\rho = \max\{\frac{s+1}{2}, \frac{s+3}{3}\}$. Comparing on $2 \leq s \leq 3$ and $s \geq 3$ gives $F/\rho \leq \frac{15}{14}$ with equality only at $s = 3, d = 0$, that is, $(a, b) = (\frac{3}{2}, \frac{3}{2})$. (3) $b \leq 1 \leq a$, $a - b \leq 1$: with $D = d_1d_2 + d_1d_3 + d_2d_3 > 0$, $\frac{15}{14}\rho - F = \frac{(s+3)Q}{56D}$ where $32Q = (16a - 3b - 21)^2 + 192 + (1 - b)(41b + 39) > 0$. (4) $b \leq 1$, $a \geq b + 1$, $a < c_1$: substitute $b = 1 - u$, $a = 2 - u + \frac{ut}{2}$ with $0 < u \leq 1$, $0 \leq t < 1$. Then $\frac{15}{14}\rho - F = \frac{P(u,t)}{112D}$, and all sixteen coefficients of the degree-$(3, 3)$ Bernstein expansion of $P$ on $[0, 1]^2$ are $\geq 56$, so $P > 0$. All identities were re-checked in exact symbolic arithmetic. □

(v) Expert bridge: there is an absolute constant $C$ such that every finite class with $|H| \geq 2$ satisfies $\mathrm{A}(H) \leq C\, k \log_k(k|H|)$. After the pointwise label compression of (i), map each hypothesis to its advice vector and invoke the realizable expert bound $\mathrm{opt}^{\mathrm{adap}}_{\mathrm{bandit}}(n, k, 0) \leq k \log_k n$ of Filmus et al. (2024, Theorem C.7). Positive and negative bandit feedback translate to the expert game constraint by constraint. Theorem C.7 requires $n \geq k$. After the compression of (i) the alphabet size is $q \leq |H| = n$, so the condition holds. This uses only the upper-bound Theorem C.7, which is independent of the lower-bound lemma discussed in Appendix G.

(vi) Atom equalities: if $q(H) \leq 2$ then $\mathrm{A}(H) = \mathrm{B}(H)$. With two feasible labels the feedback bit reveals the true label, and the adaptive and oblivious Bellman recursions coincide. Likewise, full product classes $[q']^{[m]}$ satisfy $\mathrm{A} = \mathrm{B} = \frac{m(q'-1)}{2}$: the potential $\sum_j (|S_j| - 1)/2$ over surviving per-coordinate label sets telescopes for both protocols.

## G Verification methodology and an erratum note

### G.1 Finite-value tables

Machine verification: (2) agrees with a direct exact game-tree solver on all 25 states $n, s \leq 5$ (Table 1), and its out-of-sample prediction for $a_{6,6}$ exactly matches an independent 46,719-state game-tree computation.

Table 1: The exact values $a_{n,s} = \mathrm{A}(D^{(s)}_n)$ for $1 \leq n, s \leq 5$, computed by the recursion (2) and confirmed entry by entry by a direct exact game-tree solver (all values are rationals, e.g. $a_{2,1} = \frac{16}{11}$, $a_{2,2} = \frac{27}{14}$, $a_{3,3} = \frac{288504}{103225}$; shown to six decimals). The diagonal (bold) is the asymptotic family $\mathrm{A}(H_r)$.

| $n\backslash s$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1 | **1.000000** | 1.500000 | 2.000000 | 2.500000 | 3.000000 |
| 2 | 1.454545 | **1.928571** | 2.411765 | 2.900000 | 3.391304 |
| 3 | 1.901408 | 2.338471 | **2.794904** | 3.262978 | 3.738580 |
| 4 | 2.352046 | 2.749635 | 3.174998 | **3.618311** | 4.073830 |
| 5 | 2.809410 | 3.168968 | 3.562276 | 3.978835 | **4.411919** |

Table 2: Exact certified values for the balanced towers $H_r$; the ratio intervals for $r = 4, 5$ are outward-rounded four-decimal enclosures of the exact rational intervals.

| $r$ | $\mathrm{A}(H_r)$ | $\mathrm{B}(H_r)$ | A/B |
|---|---|---|---|
| 2 | 27/14 | 11/6 | **81/77** |
| 3 | 288504/103225 | 18/7 | **112196/103225** |
| 4 | $\approx 3.618311$ | [253/80, 13/4] | [1.1133, 1.1442] |
| 5 | $\approx 4.411919$ | [153/40, 35/9] | [1.1344, 1.1535] |

### G.2 CERTIFICATES AND VERIFICATION DETAILS

The *adaptive* values in Tables 1 and 2 are computed by backward induction over version spaces with exact-fraction matrix-game solving.

**What the submitted script checks (manifest).** The supplementary material contains two standard-library Python verification scripts and the $C_{\mathrm{a}}$ mixture certificate data file. The original script, `verify_all_values.py`, recomputes every exact adaptive value in this paper ($D_2$, $H_{\mathrm{mid}}$ and its continuation values and root certificates, $D_m$ for $m \le 5$, $C_{\mathrm{a}}$, $E(n,d)$ instances and $W_m$), emitting at every game state a primal–dual certificate $(M, p, \tau, v)$ verified by an independent checker. It checks the recursion (2) against direct certified solves on all 25 states $n, s \le 5$, verifies each row of the oblivious-certificate table below (Bayes value and threshold mixture) and the $\mathrm{B}(H_4), \mathrm{B}(H_5)$ lower bounds, and runs the exhaustive simulation of the $H_{\mathrm{mid}}$ learner over all instance sequences of length up to 10. It checks $T_2(K) = 2$ (hence $\mathrm{B}(C_{\mathrm{a}}) \le 2$), the $C_{\mathrm{a}}$ mixture certificate value $\frac{1688}{867}$ and the realizability of its 46 sequences, and evaluates the serial-game profile, $G = \frac{15}{7}$, $\rho = 2$. Expected output: `ALL CHECKS PASSED`. The second script, `verify_tower_intervals.py`, certifies the $H_{352}$ enclosure and ratio bound in Appendix D.4. Its `--empirical` option also reproduces the excess doubling ratios reported in Section 4 using outward-rounded integer arithmetic.

**Research-process verification (not part of the submission).** During the research, three independently written solvers, using different data structures and LP methods, agreed on all values. Cold-start audits re-derived, among others, $a_{2,2} = \frac{27}{14}$ and $a_{3,3} = \frac{288504}{103225}$ from definitions alone, and re-proved the $\frac{5}{4}$ theorem through a different adversary and potential. The recursion (2) was validated out of sample by an exact rational match at $(6,6)$ against an independent 46,719-state computation. We report these as process. The submission's checkable evidence is the script and the certificates above and below.

**Explicit certificates for the exact oblivious values.** Each exact oblivious value of a class $D_n^{(s)}$ in this paper is the common value of a fixed-prior Bayes lower bound and the threshold-learner upper bound of Appendix D. The value $\mathrm{B}(H_{\mathrm{mid}}) = \frac{7}{6}$ is instead certified by the prior/sequence of Appendix B.1 for the lower bound and the bespoke learner there for the upper bound, and $\mathrm{B}(C_{\mathrm{a}})$ by Appendix F.3. The lower bounds use the recursion over posterior supports

$$R(t, V) = \min_y \big\{ w(V_y^-) + R(t+1, V_y^+) + R(t+1, V_y^-) \big\}, \qquad R(T+1, \cdot) = 0,$$

where $w$ is the normalized prior mass, $V_y^+ = V_{x_t, y}$ and $V_y^- = V \setminus V_y^+$. The Bayes risk of the fixed sequence is $R(1, H)$, and $\mathrm{B} \ge R(1, H)$ by Yao's principle, since revealing the sequence in advance only helps the learner. The table lists, for each class $D_n^{(s)}$, the prior (per binary hypothesis / per shield), the fixed input sequence, and the matching threshold mixture. In every row the Bayes value and both threshold bounds equal the stated B exactly as checked by machine in exact arithmetic using the script in the supplementary material.

| class | bin. prior | shld prior | sequence | threshold $J$ | B |
|---|---|---|---|---|---|
| $D_2$ | $1/5$ | $1/5$ | $x_1x_2x_1x_2$ | $1{:}\frac{3}{5},\ 2{:}\frac{2}{5}$ | $7/5$ |
| $H_2$ | $1/6$ | $1/6$ | $x_1x_2x_2x_1x_2$ | $1{:}\frac{2}{3},\ 2{:}\frac{1}{3}$ | $11/6$ |
| $D_3$ | $1/10$ | $1/5$ | $(x_1x_2x_3)^2$ | $1{:}\frac{1}{5},\ 2{:}\frac{4}{5}$ | $9/5$ |
| $H_3$ | $1/14$ | $1/7$ | $x_1x_2x_3x_2x_1x_2x_3$ | $1{:}\frac{3}{7},\ 2{:}\frac{4}{7}$ | $18/7$ |
| $D_4$ | $1/18$ | $1/9$ | $(x_1\cdots x_4)^2$ | $2{:}\frac{7}{9},\ 3{:}\frac{2}{9}$ | $20/9$ |
| $D_5$ | $1/36$ | $1/9$ | $(x_1\cdots x_5)^2$ | $2{:}\frac{1}{3},\ 3{:}\frac{2}{3}$ | $8/3$ |
| $D_6$ | $1/68$ | $1/17$ | $(x_1\cdots x_6)^2$ | $3{:}\frac{15}{17},\ 4{:}\frac{2}{17}$ | $53/17$ |
| $D_7$ | $1/136$ | $1/17$ | $(x_1\cdots x_7)^2$ | $3{:}\frac{7}{17},\ 4{:}\frac{10}{17}$ | $61/17$ |
| $D_8$ | $1/264$ | $1/33$ | $(x_1\cdots x_8)^2$ | $4{:}\frac{31}{33},\ 5{:}\frac{2}{33}$ | $134/33$ |

In particular $\mathrm{B}(D_m)$ is known exactly for all $m \le 8$, and the ratio peak of Proposition F.1 is *global*: $\max_{m \ge 1} \mathrm{A}(D_m)/\mathrm{B}(D_m) = \frac{5562}{5255}$, attained only at $m = 4$. *Proof.* The certified exact ratios for $m \le 8$ have their unique maximum $\frac{5562}{5255}$ at $m = 4$, and $\mathrm{A}(D_1)/\mathrm{B}(D_1) = 1$. For $m = 9$, the recursion of Theorem F.2 gives $\mathrm{A}(D_9) = \frac{52133}{11096}$, and $\mathrm{B}(D_9) \ge \frac{9}{2}$ by the cube, so $\mathrm{A}(D_9)/\mathrm{B}(D_9) \le \frac{52133}{49932} < \frac{5562}{5255}$. For $m \ge 10$, Theorem F.2 gives $\mathrm{A}(D_m)/\mathrm{B}(D_m) \le 1 + \frac{6}{m(m+1)} \le \frac{58}{55} < \frac{5562}{5255}$. □

The lower ends of the Table 2 intervals have the same kind of certificate. These are lower bounds only; the upper ends are the threshold values $\frac{13}{4}$ and $\frac{35}{9}$:

| class | bin. prior | shld prior | sequence | Bayes value |
|---|---|---|---|---|
| $H_4$ | $11/320$ | $9/80$ | $(x_1 \cdots x_4)^2$ | $253/80$ |
| $H_5$ | $1/64$ | $1/10$ | $(x_1 \cdots x_5)^2$ | $153/40$ |

**Remark G.1** (An erratum note on the proof of Filmus et al. (2024, Lemma C.11))**.** *This note concerns a constant in a proof, not the lemma's truth. The stopping-time argument for the adaptive lower bound* $\mathrm{opt}^{\mathrm{adap}}(n,k,r) \geq \frac{(k-1)r}{2}$ *(Lemma C.11 of the NeurIPS version; Lemma 4.11 of arXiv:2402.07453) does not yield its stated constant at* $(k,r) = (3,2)$*: against the uniform learner the construction gives only* $\frac{38}{27} < 2$*. The bound's* order $\Omega(kr)$ *is recoverable by a Markov-argument variant, and independent verification shows the lemma's statement itself is true at this parameter: a minimal-probability adversary achieves* $\geq 2$*. Nothing in this paper depends on the constant: we only ever use the* $\Omega(kr)$ *form, and the expert upper bound Theorem C.7 used in Appendix F.4(v) is independent of this lemma.*